Considering the Bystander's Perspective for Indirect Human-Robot Interaction

Katherine M. Tsui, Munjal Desai, and Holly A. Yanco
University of Massachusetts Lowell
Department of Computer Science
One University Aveue
Lowell MA 01854
Email: {ktsui,mdesai,holly}@cs.uml.edu

Keywords – Experiment, trust, social etiquette

I. INTRODUCTION

As robots become more pervasive in society, people will find themselves actively interacting with robots, and also rushing past them without any explicit interaction. People are able to maneuver in crowded situations by speeding up or slowing down to slip in between open pockets where people are not standing or walking. Our research focuses on this indirect bystander interaction.

Scholtz defines a bystander as a person who "does not explicitly interact with a robot but needs some model of robot behavior to understand the consequences of the robot's actions" and does not have formal training about the robot [1], [2]. We investigated the level of trust that a bystander has of a robotic system in a corridor passing scenario by asking people to watch short videos of such scenarios where the hallway is only wide enough to accommodate two entities (either human or robot). Our goal was to understand the bystander's mental model of how a robot should behave when passing a human, the bystander's expectation of the robot to adhere to social protocol, and the overall trust a bystander has of the robot to do the right thing.

II. EXPERIMENT

We manipulated two independent variables: the type of robot and how the robot behaved when passing a person in the corridor. We used three distinct robots: a Kyosho Blizzard, an iRobot ATRV-Jr, and a custom robot wheelchair (Fig. 1). The robot wheelchair had a rider either riding passively on the wheelchair (indicating robot control) or actively driving the wheelchair. We used four distinct passing behaviors: 1) the robot stopped, 2) the robot continued on its path but slowed down, 3) the robot continued on its path at the same velocity (neutral), and 4) the robot sped up on its path.

For each of the 16 conditions (4 robot types × 4 robot passing behaviors), we made a short video clip (Fig. 2). Sample videos can be found at http://www.cs.uml.edu/robots/ 2009-hallway-passing. In each of the videos, an off-camera person controlled the trajectory and speed of the robot. To ensure that the robot's behaviors were perceivable, we validated the content of the videos through coding (κ=86.67, excluding chance). Two coders watched each of the 16 videos and wrote a free response description of what occurred in the video.

Fig. 1. (Left to right) Blizzard, ATRV-Jr, power wheelchair (with person driving), and robot wheelchair (with robot driving)
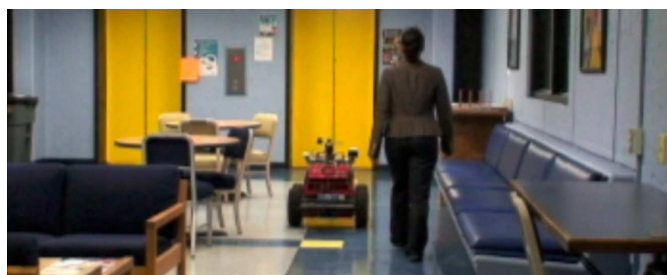

Fig. 2. A bystander passes by Jr. in a corridor

We conducted a web survey with the video clips using Amazon.com's Mechanical Turk system [3]. Participants viewed all 16 video clips. To minimize ordering and learning effects, we created 16 similar versions of the survey in which the order of the robots was rotated and the robot behavior was randomized. Fourteen people participated in each version of the survey, totaling 224 participants for this experiment. Participants were allowed up to 24 hours to complete the survey. The average length of this experiment was 1 hour 16 minutes (SD 3 hours and 28 minutes); the median time was 27 minutes. Participants were paid $1.00 through Mechanical Turk.

After each of the video clips, we collected rating information from the participant using a five-point semantic differential scale. To investigate the participant's mental model, we asked the participant to rate how appropriately the robot behaved (Q1), where 1 was "completely as I expected," 4 was "not at all as I expected," and 5 was "I had no expectations/I did not know what to expect." To investigate the participant's expectation of the robot to adhere to social protocol (Q2),

we asked the participant to rate his/her perception of the robot's manners, where 1 was "very polite," 3 was "neutral; neither polite nor impolite," and 5 was "very impolite." To investigate the participant's overall trust of the robot, we asked the participant to rate how much the participant trusted the robot to do the right thing (Q3), where 1 was "I completely trust the robot," 3 was "I neither trust nor distrust the robot," and 5 was "I completely distrust the robot."

In the post-experiment questionnaire, we asked the participants to distribute 100 points over the four robots along three independent dimensions: appropriate behavior (Q4), compliance with social etiquette (Q5), and trustworthiness (Q6). Additionally, we asked the participants to describe their reasoning for the point distribution. We asked the participants to describe each of the robots using three adjectives or phrases, and solicited for any questions or comments about the experiment.

III. RESULTS AND DISCUSSION

Of the 224 participants, there were 104 males, 119 females, and 1 person who did not report gender. Thirty-eight participants (17%) reported prior robot experience, including using robot house work devices (e.g., Roomba), industrial robots (e.g., packaging), and robotic toys (e.g., Lego). The average age of the participants was 32.2 years (SD=10.6).

Robot type. We compared the point distribution for expectation (Q4) and trust (Q6) against another type of robot using paired one-tailed t-tests with α=0.05. We wanted to see if the participants trusted one type of robot more than the others. Participants reported that they trusted the wheelchair with the person driving the most ($\bar{X}$=33.87 points, SD=19.16), the robot wheelchair ($\bar{X}$=23.43, SD=10.7), Blizzard ($\bar{X}$=22.49, SD=12.47), and Jr. ($\bar{X}$=20.21, SD=10.2) the least. The participants significantly trusted the person driving the wheelchair more than all the other types of robots (p<0.01), and the robot wheelchair significantly more than Jr. (p<0.01) regardless of their comparable size and autonomous operation. Both results are consistent with our study of attitudes towards automation [4]. We found that the participants trusted Blizzard more than Jr. (p<0.01) which we hypothesize is due to the small size of the Blizzard.

Looking at how appropriately the robot behaved (Q4), we found that participants reported that Blizzard performed most appropriately ($\bar{X}$=26.42 points, SD=11.09), wheelchair with the person driving ($\bar{X}$=25.74, SD=9.56), the robot wheelchair ($\bar{X}$=24.19, SD=8.49), and Jr. ($\bar{X}$=23.65, SD=9.04) the least. Interestingly, Blizzard outranked all the other robot types. Participants rated Blizzard as better matching their expectations over the robot wheelchair and Jr. (p<0.04); the difference between the expected behavior of the wheelchair with the person driving and Blizzard was not significant. Aside from this inconsistency, the robot types follow the above trend.

Robot behaviors. We also wanted to see if people would trust one type of behavior more than the others. We conducted statistical analysis of the ratings of Q1, Q2, and Q3 against another types of the behaviors. We examined how well people rated the robot's actions as matching their expectations (Q1).

In 7.6% of responses, participants reported "no expectation/did not know what to expect," which were removed for this analysis. For all of the robot types, people reported that they found the neutral behavior to be the most expected slow, stop, and fast the least expected. We computed unpaired one-tailed t-tests with α=0.05 comparing each robot behavior against each other and found significance in all cases, where p<0.01.

We also looked at how the participants rated the robots' adherence to social etiqutte (Q2). The participants reported that they found the stop behavior to be the most polite ($\bar{X}$=1.89, SD=1.03), slow ($\bar{X}$=2.00, SD=0.89), neutral ($\bar{X}$=2.48, SD=0.88), and fast ($\bar{X}$=3.50, SD=1.09) the least; 1 was the highest politeness rating and 5 was the least. We computed paired one-tailed t-tests with α=0.05, comparing each robot behavior against each other, and found significance in all cases, where p<0.01. We hypothesized that people would find the fast behavior to be the most impolite as this abrupt behavior is infrequently employed. However, we did not expect that people would rate the stop or slow behaviors as more polite than the neutral behavior because the corridor was wide enough to accommodate two entities side by side.

From the ratings of Q3, the participants reported that they trusted the stop behavior the most ($\bar{X}$=2.20, SD=0.93), slow ($\bar{X}$=2.24, SD=0.82), neutral ($\bar{X}$=2.36, SD=0.85), and fast ($\bar{X}$=3.20, SD=1.06) the least; 1 was the highest trust rating and 5 was the least. We computed paired one-tailed t-tests with α=0.05 comparing each robot behavior against each other. People trusted both the stop and slow behavior significantly more than the neutral (both p<0.01) or fast (both p<0.01) behaviors. The participants trusted the neutral behavior more than the fast behavior (p<0.01). We believe that participants trusted the slower behaviors the most and the faster behaviors the least because a slower velocity allows for a larger reaction time in case of an unexpected behavior.

IV. CONCLUSIONS AND FUTURE WORK

We found that for our four types of robots, all should move in the manner that people do: continue at a relatively constant speed unless there is need to yield, in which case the robot should slow or stop. Our next step is to investigate corridor passing scenarios in which the area is only wide enough for one person or robot, thus causing one to yield.

REFERENCES

[1] J. Scholtz, "Theory and Evaluation of Human Robot Interactions," in *Proc. of the Hawaii Intl. Conf. on Systems Sciences*, 2003.
[2] J. Scholtz and S. Bahrami, "Human-Robot Interaction: Development of an Evaluation Methodology for the Bystander Role of Interaction," in *IEEE Intl. Conf. on Systems, Man and Cybernetics*, vol. 4, 2003.
[3] Amazon, "Amazon Mechanical Turk: Artificial Artificial Intelligence," 2009, http://www.mturk.com. Accessed Dec. 2009.
[4] M. Desai, K. Stubbs, A. Steinfeld, and H. Yanco, "Creating Trustworthy Robots: Lessons and Inspirations from Automated Systems," in *"New Frontiers in Human-Robot Interaction" Workshop at the Convention of the Society for the Study of Artificial Intelligence ad Simulation of Behaviour (AISB)*, 2009.