John Ryan

Human Robot Interaction for Reading, Literacy, and Library Patronage

Introduction: Human Robot Interaction is generally the study of how humans interact with machines that have autonomy. For example, a study on how workers at a packing facility engage with a robotic arm that bundles and wraps boxes together would fall into the Human Robot Interaction category. One important subcategory of Human Robot Interaction studies people's perception of social robots, or robots designed for the sole purpose of social interaction. It is necessary to understand and predict how people will react to social robots if they become more commonplace in society. Disregarding the wide range of reactions seen from people while interacting with social robots could be problematic, especially dealing with their perceived safety and anxiety around robots. Researchers have found promise in using social robots to alleviate stress and anxiety in children [1], and success with engaging audiences with expressive storytelling [2]. However, there is a need for more data and sustained long term studies to determine if there is value in using social robots in these settings, and determine the risks involved with deployment of social robots. The goal of this study is to gather data on people's reactions to social robots over an extended period of time, categorizing results by age, gender, and level of previous experience with robots. This data will be used to determine the usefulness of a social robot as a storytelling device, and the effects of long term exposure to social robots. It is predicted that people with previous exposure to social robots will feel more safe due to being more used to interacting with them, however less interested and engaged due to the novelty wearing off.

Methods: In order to accomplish this eight different children's books were programmed to be read by Pepper, a humanoid social robot. Pepper was sent to a local library to read books to patrons, and gather survey data after the interaction. A modified version of the GODSPEED questionnaire [3] was used to measure how humanlike, lively, likeable, and intelligent the robot seemed as well as how safe participants felt during the story that was read to them. Although the GODSPEED survey was previously developed and tested, it was decided to modify the wording on the survey to make it more easily understandable for people with lower English proficiency. Although simplifying language may sacrifice specificity of individual questions, for the purpose of this study it was decided the benefits of a more palatable survey outweighed this burden. There were 31 different survey responses on the first round of data collection, two of which were incomplete giving a cleansed sample size of n = 29. There will be more data collected at a different library to ensure meaningful conclusions can be drawn from the data, as well as to validate the modified version of the GODSPEED survey that was developed. My main contribution to this research was taking the lead on programming Pepper to read a story to children. I programmed two of the eight stories, and was the first to discover and share many different techniques used in every group member's final code. I shared with groupmates how to show images and play audio on Pepper, and how to preload assets to reduce latency when running the story reading programs. I also developed a program to turn off motor stiffness and record encoder values on Pepper to playback movements physically done instead of tediously coding each motor position to achieve the desired movement (Figure 1). This would streamline the creation of custom motions for story performances.

Data: The survey employs a five point Likert scale for participants to report how they felt while interacting with Pepper. On the 19 question survey, responses were grouped into the humanlike, lively, likeable, and safety categories. Cronbach's alpha was computed to measure internal consistency, or to ensure each question within the defined characteristic groups were measuring the same thing. With $\alpha \geq 0.7$ generally accepted, four of the five planned categories are acceptable (Table 1), while the three questions planned to be grouped into an intelligence characteristic had an $\alpha = -0.081$. The demographic information was encoded to have each response correspond to a number. For gender, male was represented with 5, and female was represented with 10. For level of experience with robots, having interacted with the same or similar robot was represented with 10, and not having experience with a similar robot was represented with 5. Ages were binned with 3-5 years represented by 0, 6-7 years represented by 10, 8-19 years represented by 20, and 20+ years represented by 30. The distance between points on the Likert scale was approximated to be the same to treat the ordinal data as an interval. The Pearson correlation coefficient r between each demographic category and measured response was computed (Table 1).

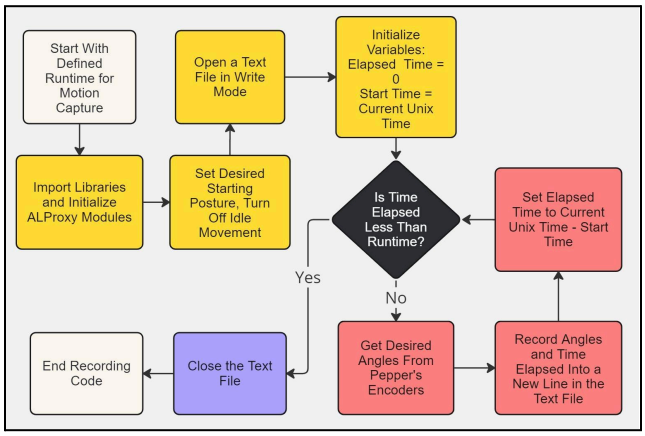

Figure 1: Flowchart of Motion Recording Algorithm

Results: A significance test for nonzero correlation was computed finding a critical value of $|r| \geq 0.355$ for $\alpha = 0.05$. Thus, a significant negative correlation of -0.381 was found between previous experience with robots, and perceived safety in the encounter. Also, a near significant negative correlation of -0.351 was found between gender (scaling male to female), and safety. This means that females, and people who had previous experience with robots were more likely to report feeling unsafe interacting with the robot. This correlation shows the opposite trend to what was predicted, and people did not report feeling more safe as they became accustomed to the robot. It is interesting to note that all correlations for experience with robots were negative, meaning survey responses were generally more negative for all categories. This could be due to the novelty effect predicted, however only one of the correlations was significantly negative. My role in the data analysis was computing the cronbach's alpha, as well as the significance test for the correlations. I programmed these functions in MATLAB for future use.

Discussion: Based on the initial analysis done on the first round of data, it will be interesting to see if there is a significant trend for people to view Pepper more negatively when they have previous experience with robots. More data, and better suited statistical analysis is necessary to draw conclusions, however future research could involve a study over time to better understand this effect. The responses of participants interacting with Pepper at regular intervals could be analyzed to see how survey responses regarding different characteristics change over time. I would predict that people would develop more concrete opinions over time, less likely to change later in this potential study. This could be used to find the optimal frequency of exposure to social robots to improve people's overall experience, and get a more accurate idea of how people would react to long term exposure and implementation of social robots into everyday life.

Characteristic	Cronbach's Alpha	Gender Correlation	Experience Correlation	Age Correlation
Human-like	0.769	-0.081	-0.162	-0.235
Lively	0.748	0.167	-0.052	-0.343
Likeable	0.878	-0.145	-0.219	0.163
Perceived Safety	0.942	-0.351	-0.381	-0.184

Table 1: Demographic Correlations for Grouped Questions

References:

[1] M. K. Crossman et al.,"The influence of a socially assistive robot on mood, anxiety, and arousal in children," Profes. Psych. Research & Pract., vol. 49, pp. 48-56, Feb. 2018.

[2] K. Westlund et al.,"Flat vs. Expressive Storytelling: Young Children's Learning and Retention of a Social Robot's Narrative," Front. in Hum. Neurosc., vol. 11, Jun. 2017.

[3] C. Bartneck et al.,"Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots," Int. J. of Soci. Robo., vol. 1, pp. 71-81, Nov. 2008.