Understanding Expectations of a Robot's Identity Through Multi-User Interactions

Wilma Alice Bainbridge, Shunichi Nozawa, Ryohei Ueda,
Yohei Kakiuchi, Kotaro Nagahama, Kei Okada, Masayuki Inaba

The JSK Robotics Laboratory

Department of Mechano-Informatics, The University of Tokyo
Engineering Building No. 2
7-3-1, Hongo, Bunkyo-ku, Tokyo, Japan 113-8656
+81 3-5841-7416

wilma.bainbridge@gmail.com

ABSTRACT

We conducted two experiments looking at how to read user expectations of a robot's identity within multi-user environments. Multi-user environments are unpredictable and fast-paced, which can become a challenge for roboticists to interpret. However, they also present a rich landscape of data, and we propose methodologies to retrieve user reactions to the robot through sensor data. We also emphasize the necessity of using the results from these methodologies to define a robot's identity based on user expectations. In Experiment 1, we found that sensor data taken from a handshake with the robot can be used to find differences in the views of different demographic groups towards interaction with our robot. In Experiment 2, we expand the subject pool and reaffirm the usefulness of sensor data in multi-user environments, while also using questionnaire data to create an identity for our robot.

I.2.9 [**Robotics**]: Operator interfaces – *methods to research user expectations and intuitive interaction, multi-user environments, robot identity, biofeedback*

1. INTRODUCTION

Humanoid robots are almost ready for the real world - many are safe, durable, and able to perform complex interactions with humans, objects, and environments. However, there are two important issues that have rarely been addressed within human-robot interaction (HRI) research. First, as robots make their entrance into the real world, they will need to be able to handle environments filled with multiple users. Second, robot's identities and personalities have often been designed based on roboticists' intuitions rather than based on the user's expectations. This study proposes new ideas on methodology for examining user impressions on human-robot interactions within a multi-user setting. We also propose molding a robot's identity based on user expectations of a robot.

Recent HRI research has just begun to focus on placing robots within ecologically valid multi-user environments rather than the traditional one-on-one interaction model [1, 2]. Many of these works focus on the engineering obstacles in having a robot interact smoothly with multiple people, and have not yet explored ways of measuring human reactions to the robot. It is especially tricky to extract data from multi-user environments because users have brief, unscripted actions with the robot. Often it is even impossible to gather surveys from subjects, forcing researchers to rely on other measures. However, these multi-user environments are rich untapped sources of information, because they can provide large amounts of data from several subjects in an ecologically valid setting without any lengthy training of the subjects.

While much of current HRI research looks at subjects' feelings on an interaction, few studies ask the user to define the robot's identity based on their interaction. Roboticists often pre-create a robot's identity before an interaction, assigning it a name, gender, voice, etc. However, it is important that a robot's identity be suited to match the expectations of the general public, rather than the assumptions of an engineer. Expectations of a robot's identity can vary greatly based on a user's age, gender, or culture. We propose measuring users' reactions to a robot to determine how to best create a robot to match user expectations. We also examine how expectations of robot identities differ across these demographic groups. Previous studies have shown that the collection of questionnaire data alone often ignores other signs of a human's feelings towards a robot [3], so we employ both sensor data and questionnaire data to fully assess a user's reactions to a robot within a multi-user setting.

Figure 1. An example of the interaction

The HRP-2 is shown here during Experiment 2, shaking hands with a participant while under observation by the experimenter.

These two main goals - creating natural interactions within multi-user environments, and designing methods to define robotic identities based on user expectation - are very lofty goals that will continue to be themes of HRI research in the years to come. The current study aims to serve as a pilot study for new methodological ideas at how to answer these questions, and as a starting point for discussion.

2. GENERAL METHOD

2.1 Overview

We performed two studies to examine expectations and reactions to a robot's identity within a multi-user setting. Experiment 1 took place at an alumni reunion of our laboratory, and assessed the usefulness of sensor data for understanding human reactions to our robot within this multi-user environment. For Experiment 2, we expanded the model of Experiment 1 and asked people across the University of Tokyo Hongo campus to interact with the robot. For Experiment 2, we examined sensor data as well as questionnaire data. The results of Experiment 2 inspired us to also perform a brief survey across campus to investigate how to match expectations of a robot's voice to its appearance.

2.2 Methods

For both studies, the robot followed the same pattern of behavior. It searched for a human face, and after finding a face for five consecutive frames, it initiated a randomly chosen action: a bow, a handshake, a wave, or inaction. During the interaction, the robot's head followed faces that it found, appearing to create eye contact with the subject, and it continued following the faces during its greeting action. Refer to Figure 1 for an example from Experiment 2 of the handshake gesture.

Data were taken from many different sources. We chose measures that could be easily taken by the robot within a limited amount of time, and that seemed relevant to interpreting human emotion. Our main focus for sensor data was during handshakes with the robot, as there is plenty of data that can be taken just from the contact of the human's hand with the robot's hand. Sensor data included the user's directional forces on the robot's arm (Newtons), tactile measurements from the robot's hand (a unitless analog measure), temperature data measured from the robot's index and middle fingers (Celsius), and the distance to the subject's face taken from its cameras and face recognition software (millimeters). Sensor and video data resulted in about 1 GB of data per minute, and thus several thousand lines of data were collected from each subject. Biographical data for the subject were hand-coded by the experimenter and included the user's gender, country of origin, age group, familiarity with the laboratory and department, and vocal reactions to the robot. Opinion data were taken through a questionnaire to the subjects for Experiment 2.

Data were then analyzed for significance across groups using statistical analysis tests including the Independent Samples T-test, Analysis of Variance (ANOVA), and Pearson's Correlation. For the purposes of these two experiments, data were separated into groups for analysis (such as handshake data versus non-handshake data) rather than separated by subject, in order to capture the wide range of information within each interaction. Groups for analysis were determined by divisions within the biographical data (age, gender, nationality, laboratory familiarity) and by binary measures taken during the interaction (such as negative reaction versus positive reaction). Sensor data were compared between groups using all pieces of data that either had a face tracking measurement or were a part of the handshake behavior.

2.3 The Robot: HRP-2

The robot we used for our experiments is the HRP-2, a bipedal humanoid robot developed by Kawada Industries through funding from the Japanese Ministry of Economy, Trade and Industry [4]. It is 154 cm tall, weighs 58 kg, and has 30 degrees of freedom. The joints in its hands and arms are flexible so that they accommodate to forces from the user. This results in a safe and natural interaction with the robot, and actions such as a handshake with the robot are comfortable for the user. For both experiments, the HRP-2 was connected only to a power source, and supported itself using auto-balancing in its legs.

The HRP-2 that we used for each experiment is modified for the laboratory (The HRP2JSK), including stereovision, a head with seven degrees of freedom, and multiple movable fingers [5]. It was controlled using code written in Euslisp [6] with a ROS architecture [7]. For this experiment, we mainly collected data from sensors already built into the robot, such as the force and tactile sensors. All joints in the HRP-2 are equipped with force sensors, and there are tactile sensors on both hands. We also modified the hand of the HRP-2 to include temperature sensors in its index and middle fingers (points that made the most contact with a human hand during a handshake). Within this paper, the data from these sensors will be hereafter referred to as *temp0* (index finger) and *temp1* (middle finger). The tactile sensors used to measure tactile forces on the hand will be referred to in this paper as *tactile0* and *tactile1*.

3. EXPERIMENT 1

3.1 Methods

The HRP-2 greeted alumni who came to visit the laboratory as a part of an alumni day. As participants walked into our laboratory, the HRP-2 stood at the door and initiated a greeting as described in the Overall Methods. Participants were all previous members of our laboratory and so had experience with robots. Current members of our laboratory also interacted with the robot. In total, 27 people interacted with the robot (24 male, 3 female) over a period of one hour. Seven were current members of the laboratory, while twenty were alumni. Interactions were very brief, and did not last longer than a few minutes per subject. Sensor data were collected several times a second. After selecting data with only handshake or face tracking measurements, subject data ranged from 118 to 1520 data pieces per subject, with an average number of 437 ($SD = 372$).

3.2 Methods

To confirm the validity of our data analysis techniques, we compared the sensor data between when the robot was involved in a handshake behavior versus non-contact behaviors. As expected, temperature (*temp0:* $t(7150) = 30.21$, $p < 0.001$; *temp1:* $t(7150) = 26.26$, $p < 0.001$) and tactile data from tactile 1 ($t(7150) = 33.44$, $p < 0.001$) were higher, distance to the subject's face was closer during the handshake ($t(1061) = 3.21$, $p < 0.001$), and forces on the arm were higher except for in the x direction ($p < 0.001$). We then compared other group divisions on the same measurements for while they were engaged in a handshake with the robot, as this was the only time they were actively touching the robot. Group divisions were decided based on biographical information we could collect without a survey (gender and relationship to the laboratory) and vocalizations of the subjects during the interaction. All subjects in this experiment were Japanese, so we did not look at the effects of culture. Males measured higher for force against the arm in all directions of x, y, z, roll, pitch, and yaw (for all, $p < 0.001$). This likely reflects a gender difference in strength. Females had higher hand temperatures (temp*0:* $t(11127) = 16.25$, $p < 0.001$; temp*1:* $t(11127) = 3.26$, $p < 0.001$), lower tactile measurements (t*actile0:* $t(11127) = 10.14$, $p < 0.001$; *tactile1:* $t(11127) = 10.25$, $p < 0.001$), and closer faces ($t(881) = 4.80$, $p < 0.001$). Alumni of the laboratory showed a similar trend versus current lab members; higher temperatures (*temp0:* $t(11808) = 16.65$, $p < 0.001$; *temp1:* $t(11808) = 14.07$, $p < 0.001$), lower tactile measurements (*tactile0:* $t(11808) = 26.74$, $p < 0.001$; *tactile1:* $t(11808) = 4.93$, $p < 0.001$), and a closer face distance ($t(1471) = 1.246$, $p < 0.001$). We reviewed video of the experiment and coded subjects who made positive remarks (such as, "cool") versus negative remarks ("scary"), excluding current members of the laboratory. Four subjects made positive remarks, while three made negative remarks. Subjects who made negative remarks versus those who made positive remarks had similar data patterns to the other group divisions, with again higher temperatures (*temp0:* $t(3060) = 12.02$, $p < 0.001$; *temp1:* $t(3060) = 17.22$), lower tactile measurements (*tactile0:* $t(3060) = 15.86$, $p < 0.001$; *tactile1:* $t(3060) = 4.68$, $p < 0.001$), and closer face distances ($t(293) = 2.31$, $p < 0.001$). Refer to Table 1 for a comparison of the average measurements for each group.

Table 1. Sensor Data for Experiment 1.

Average temperature and tactile measurements for different groups are shown, with higher means between group divisions bolded.

Group	Temp 0	Temp 1	Tact 0	Tact 1	Face Dist
All subjects	25.8	24.8	8.8	15.7	1413.8
Handshake	**27.4**	**26.4**	**22.2**	**83.9**	1337.1
No handshake	25.5	24.4	19.9	7.2	**1417.2**
Female	**27.0**	**25.5**	9.4	34.1	781.2
Male	26.2	25.4	**16.6**	**51.9**	**1442.5**
Alumni	**26.7**	**25.7**	12.1	47.0	1447.1
Current members	26.0	25.1	**26.5**	**53.7**	**1521.0**
Negative	**26.2**	**25.5**	7.1	57.9	1387.6
Positive	25.4	24.2	**9.9**	**70.6**	**1514.4**

3.3 Discussion

The similar pattern in sensor data (higher temperature, lower tactile measurements, and closer face distance) for women, old (versus current) members of the laboratory, and people expressing negative comments reflects a likely similar reaction to the robot. This pattern especially stands out, because while a handshake (versus no interaction with the robot) will result in both higher temperature and tactile measurements, these groups show an opposite pattern of high temperature but low tactile measurements. We hypothesize this pattern could reflect violations of expectation of the robot for the subject. The violation of expectation could be negative, such as stress, or it could be positive, such as excitement. Higher finger and hand temperatures have been correlated with stress and arousal in psychology research [8]. Lower tactile measurements could indicate a reluctance to touch the robot or a difference in handshake style. The differences in face distance could indicate several possible emotions, and further research is necessary. Previous research has found that subjects who feel more negatively about a robot are more willing to invade its personal space [3]. However, this close distance to the robot could also indicate close examination of the robot based on curiosity, or even a comfort with the robot. All of these ideas are at most speculative, but the fact that there are significant differences in sensor data across demographic groups demonstrates the potential usefulness of this methodology. Further study is required to determine the relationship between these biological measures and feelings towards a robot, but these results show an interesting pattern ripe for future investigation, and that a robot's sensor data can be useful even in quick, multi-user interactions to extract differences between demographic groups.

4. EXPERIMENT 2

4.1 Methods

In order to reaffirm the effectiveness of our methodology, we conducted a second, similar experiment, but on a larger-scale. Specifically, we aimed for a longer experiment time, a larger subject pool, and working with people who had never met a robot before. We placed the HRP-2 outside in the University of Tokyo Hongo campus for three hours, during the University's Homecoming event for alumni in all departments visiting the campus. People who walked by the robot were asked to briefly "meet" the robot as a part of a demonstration for the laboratory. As subjects came closer, the robot then tracked the subject's face and initiated a greeting with the subject, as described above in the Overall Methods. During Experiment 1, some people remarked that they wanted to speak with the robot, so we had the robot speak a simple greeting when it initiated its handshake and wave gestures. The robot greeted people in Japanese, using an average male speech generation voice [9], and said the Japanese equivalent of, "Nice to meet you, my name is HRP-2". In total, 70 people interacted with the robot (49 male, 21 female). Participants came from all parts of the school and were of all ages, with an average age of 25 years. Unlike Experiment 1, Experiment 2 included both Japanese and foreign subjects. Seven were from Western countries, while sixty-three were from East Asia.

After interacting with the robot, subjects were asked to fill out a one-page questionnaire. The questionnaire was divided into three parts: 1) *biological information* - the subject's gender, age group, and if they had interacted with a robot before, 2) *adjectives about the interaction* - cool, natural, scary, fast, interesting, and cute, 3) *descriptors of the robot* - human versus robotic, masculine versus feminine, childish versus adult, Japanese versus foreign, and how much the voice matched the robot. Subjects rated the adjectives on a scale of 1 (low) to 10 (high). Fifty-five people filled out the questionnaire (33 male, 13 female, 9 no response). Fifteen had interacted with a robot before, while forty had not.

4.2 Results

There were slight differences in the sensor data taken from Experiment 2 versus Experiment 1. We coded extra information including whether subjects reciprocated the robot's actions when waved or bowed to. Due to last-minute difficulties with the motors in our robot's hand, we had to switch to a different HRP-2 that did not have its face distance measurement actively working. The tactile sensors were also differently calibrated from the HRP-2 in Experiment 1, so the range for the data is much higher and narrower (239 - 242), still as an analog unitless value. The narrow range of the data made us question the accuracy of the tactile data, but they had a strong correlation with temperature as expected from Experiment 1 (*tactile0 to temp0*: $r = 0.98$, $p < 0.001$; *tactile0 to temp1*: $r = 0.97$, $p < 0.001$; *tactile1 to temp0*: $r = 0.98$, $p < 0.001$; *tactile1 to temp1*: $r = 0.97$, $p < 0.001$), so we believe the sensors were functioning well. The robot's appearance was the same as the robot used in Experiment 1.

We looked at similar groupings to Experiment 1 and found some interesting similarities and differences in the results. A summary of the results between groups can be seen in Table 2. A comparison of lab members who interacted with the robot to people freshly meeting the robot showed a similar pattern to Experiment 1. People new to the robot had higher temperatures (*temp0:* $t(50510) = 31.86$, $p < 0.001$; *temp1:* $t(50510) = 31.15$, $p < 0.001$) and put lower force ($p < 0.001$ for all directions except for y and yaw) on the robot, but unlike Experiment 1, had a higher tactile measurement than lab members (*tactile0:* $t(50510) = 28.34$, $p < 0.001$; *tactile1:* $t(50510) = 28.54$, $p < 0.001$). Gender differences were opposite of Experiment 1; females had lower temperature (*temp0:* $t(50510) = 1.61$, $p < 0.001$; *temp1:* $t(50510) = 2.23$, $p < 0.001$), higher tactile measurements (*tactile0:* $t(50510) = 4.67$, $p < 0.001$; *tactile1:* $t(50510) = 4.57$, $p < 0.001$), and higher force on the arm ($p < 0.001$ in all directions) than males. Females also more frequently reciprocated the robot's waves and bows compared to males ($\chi^2(2, 70) = 13.66$, $p < 0.001$). We also looked at a possible effect of culture on sensor data. East Asian subjects had higher temperatures (*temp0:* $t(49096) = 5.94$, $p < 0.001$; *temp1:* $t(49096) = 5.40$, $p < 0.001$) and lower tactile measurements (*tactile0:* $t(49096) = 3.12$, $p < 0.001$; *tactile1:* $t(49046) = 3.09$, $p < 0.001$), compared to Western subjects.

Subjects' average responses on the questionnaire items are shown in Table 3. We examined correlations of biographical data to the survey results. Age group was closely correlated to several questions in the survey. People in older age groups tended to rate the robot higher for being feminine ($r = -0.33$, $p < 0.05$), but lower for being cool ($r = -0.45$, $p < 0.005$), interesting ($r = -0.41$, $p < 0.005$), cute ($r = -0.34$, $p < 0.05$), and having its voice match its appearance ($r = -0.39$, $p < 0.01$). We also examined correlations between questionnaire answers. Some correlations validated assumptions about how subjects would respond; subjects who found the robot masculine also said the man's voice matched the robot ($r = 0.34$, $p < 0.05$), subjects who found the robot natural said it was more human-like ($r = 0.35$, $p < 0.01$), and there was a direct positive correlation amongst the qualities of cute, cool, and interesting ($p < 0.01$ for each comparison). There was also a significant correlation between subjects who found the robot foreign and those who found it scary ($r = 0.43$, $p < 0.001$), possibly reflecting interesting cultural perceptions of a robot. We did not find any significant differences in opinion data based on gender or previous experience with robots.

Table 2. Sensor Data for Experiment 2.

Average temperature and tactile measurements for different groups are shown, with higher means between group divisions bolded.

Group	Temp 0	Temp 1	Tact 0	Tact 1
All subjects	**20.0**	**20.3**	**239.88**	**240.80**
Female	21.4	21.7	**240.02**	**240.93**
Male	**21.6**	**21.9**	239.64	240.57
General Subject	**21.9**	**22.3**	**240.05**	**240.98**
Lab member	16.2	16.7	236.03	236.92
East Asian	**21.6**	**22.0**	239.71	240.63
Western	20.8	21.2	**240.06**	**240.98**

Table 3. Questionnaire results.

Numbers indicate average scores given by subjects on a scale of 1 (low) to 10 (high). Averages are bolded together with the term they were closer to. The questionnaire was delivered in Japanese, and so there may be slight differences in nuance of the English translations used in this paper.

Word	Average	Opposite	Word	Average	Opposite
Cool	**7.57**	Uncool	**Cute**	**6.43**	Not cute
Natural	**5.47**	Unnatural	**Voice matches**	**6.82**	Doesn't match
Scary	**4.48**	**Not scary**	Human-like	**4.43**	**Machine-like**
Fast	**4.62**	**Slow**	**Masculine**	**8.20**	Feminine
Interesting	**8.30**	Boring	Childish	**3.44**	**Adult**

4.3 Follow-up Survey

The results of Experiment 2 prompted us to do a brief survey across campus to determine which computer-generated voice (from AquesTalk's library [9]) best matched the robot's appearance. This survey allowed us to choose a voice for the robot for future interaction experiments that would best match expectations of the robot. We asked 76 people (54 male, 22 female, average age 27.6 years) across campus to see a picture of the robot and then choose one of five different voices (two female, two male, and one very machine-like male robotic voice) for the robot. The survey was written in visual programming language Lazarus [10], and conducted on a multi-touch Windows 7 tablet computer. We used only a picture for this survey rather than video or interaction with the robot, to get a large number of opinions from across campus. Female subjects most frequently chose the second female voice (59.1%) while male subjects most frequently chose the robotic male voice (27.8%) but were more evenly distributed in their choices. People who chose the female voice said it was soft, approachable, and easy to interact with. People who chose the robotic voice said it matched best because it was most stereotypically robotic. These opinion differences reflect very opposite differences in expectations for the robot – for it to be comfortable for interactions with humans, or for it to be as robot-like as possible. It will be interesting to investigate in future studies what factors cause different perspectives on the role a robot's identity should fill. Despite the differences in trend between gender, overall, the most popular voice was the female voice (31.6%), and this voice will be used for future human-robot interaction studies with our HRP-2. We will also try testing how perceptions of the robot's voice change when interacting with the real robot versus selecting a voice based on solely a picture.

4.4 Discussion

The sensor data from this experiment present several interesting possible interpretations. First of all, the opposite trend in the data based on gender stands out. This difference could come from a number of factors – potential differences in the subject pool (the female engineer alumni of Experiment 1 versus the general female population of Experiment 2), differences in the robot (the added voice), or perhaps differences in the experiment presentation (for Experiment 1, every alumnus met the robot, while for Experiment 2, only people who actively approached the robot became subjects). However, there is still the similar pattern of a higher temperature, lower tactile, and lower force in one gender between both experiments. As women in Experiment 2 were much more likely to reciprocate a robot's gestures than men, it seems possible that the women in this experiment felt more comfortable with the robot. However, further investigation into gender differences in expectations and feelings towards robots will be necessary.

One other interesting pattern is the higher temperature and lower tactile data of East Asian subjects versus Western subjects. This could reflect a possible cultural difference in comfort with interactions with robots, or a cultural difference in hand-shaking, as it is a much more common greeting in the West.

We also found a difference in the tactile results from Experiment 1 and 2 for members of the lab versus non-members. However, we believe the difference may not be particularly interesting, as many lab members in Experiment 2 used handshakes with the robot to test its function, and were thus not "fully involved" handshakes, unlike with Experiment 1.

These interpretations, however, are again only explorative and it is impossible to make conclusions about gender or cultural perceptions of robots based on only these results. However, the results of Experiment 2 demonstrate that even within a different setting, we can still find significant differences in sensor data between demographic groups.

Based on the questionnaire and follow-up survey results, we can paint a clear picture of the identity people assign to the robot. To the average subject, the robot was viewed as a masculine, foreign, adult. However, ideas of the robot's identity changed across subject demographics (especially age), and reflect differing expectations in how a robot should appear. There is also the interesting discrepancy of subjects finding the robot masculine and saying that the robot's voice matches, but selecting a feminine voice in the follow-up survey. This perhaps points to a need for a robot to be dynamic, and able to adjust its identity to match its user's expectations, as they may differ strongly based on subject.

5. GENERAL DISCUSSION

This study accomplished two main tasks that will be important to the future design of human-robot interactions. First, we proposed a methodology for retrieving emotion data from subjects when interacting with a robot. We found that using sensor data based upon measures from a person - their hand temperature, the directional forces of their hand, their face distance - can provide a quick look into people's unconscious reactions towards a robot. While the current study only begins to get at possible differences in robot perceptions, we hope to refine this methodology to further explore potential demographic differences and how they relate to psychological phenomena in future work. We also hope to expand the data we collect to include other information typically used in psychological studies, such as skin conductance, voice data, and tactile measurements across the entire robot's body. Combining the sensor data taken during these quick interactions with questionnaire and behavioral data can create a comprehensive image of a user's feelings and expectations for a robot, and in the future robots could learn to dynamically adjust to these signals in order to meet a user's expectations.

The second main contribution of our study is demonstrating one potential method for robots to collect data quickly within unconstrained multi-user environments. Using sensor data allowed our robots to collect thousands of samples of data from individuals in only minutes of time. From reviewing the robot's camera data, it is easy to separate out individuals' data and collect basic biographical information (such as gender) without having to have an experimenter actively collect data on-site or distribute surveys. While these two experiments were focused solely on the collection and analysis of these sensor data, this methodological approach could be used during any human-robot interaction. For example, a quick handshake before and after public demonstrations of a robot could be used to collect user opinion data quickly, and to allow the robot to adjust to its users' expectations. The ability to take data quickly allows researchers to collect HRI reaction data in fast-paced, complicated, multi-user settings, as demonstrated in Experiments 1 and 2.

Our study also found some potentially interesting demographic differences in expectations towards robots that would be interesting starting points for future investigation. There appears to be a gender difference in attitudes towards robots, but it is still difficult to tease apart the direction of this gender difference. One previous study found that males viewed robots as more human-like, while females viewed robots as more machine-like and unsocial [11]. However, we found some hints in Experiment 2 that females may feel more comfortable with robots. We also found potential differences in expectation for a robot's role and gender from our follow-up survey to Experiment 2. These results are still very preliminary, and further study will be necessary to examine gender differences in robot expectations,. We also found other demographic differences that are ripe for future investigation. Some HRI research has proposed cultural differences in views towards robotics between the West and the East [12], and our results from Experiment 2 also present a possible difference in comfort with a robot. The questionnaire results from Experiment 2 also uncover a possible age difference in expectations of robots; while younger people expect robots to be exciting and modern, older people may be less concerned with the "cool factor" of a robot.

Overall, these results present a glimpse into possible methods to investigate expectations and reactions to robots within natural, multi-user environments, and we propose using these results to shape a robot's identity. This is only the first step in investigating effective methods for adapting robots to match human expectations within a natural multi-user environment, and this will be an interesting field for discovery in future research.

6. ACKNOWLEDGMENTS

Wilma Bainbridge would like to thank the Fox International Fellowship, Gordon Grand Fellowship, and Yale University for their support for her research at the University of Tokyo.

7. REFERENCES

[1] Matsusaka, Y., Tojo, T., Kubota, S., Furukawa, K., Tamiya, D., Hayata, K., Nakano, Y. and Kobayashi, T. 1999. Multi-person conversation via multi-modal interface. In *Proceedings of the Sixth European Conference on Speech Communication and Technology* (Budapest, Hungary, 1999).

[2] Kondo, Y., Takemura, K., Takamatsu, J. and Ogasawara, T. 2010. Multi-person human-robot interaction system for android robot. In *Proceedings of the IEEE/SICE International Symposium on System Integration* (Sendai, Japan, December 21-22, 2010).

[3] Bainbridge, W. A., Hart, J., Kim, E. S. and Scassellati, B. 2009. The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics*. 1-2, 2009 – 2010.

[4] Hirukawa, H., Kanehiro, F., Kaneko, K., Kajita, S., Fujiwara, K., Kawai, Y., Tomita, F., Hirai, S., Tanie, K., Isozumi, T., Akechi, K., Kawasaki, T., Ota, S., Yokoyama, K., Honda, H., Fukase, Y., Maeda, J., Nakamura, Y., Tachi, S. and Inoue H. 2003. Humanoid robotics platform developed in HRP. In *Proceedings of the 2003 IEEE-RAS International Conference on Humanoid Robots* (Humanoids 2003), 2003.

[5] Okada, K., Ogura, T., Haneda, A., Kousaka, D., Nakai, H., Inaba, M. and Inoue, H. 2004. Integrated system software for HRP2 humanoid. In *Proceedings of the International Conference on Robotics and Automation* (ICRA 2004), 3207 - 3212.

[6] Matsui, T. and Inaba, M. 1990. Euslisp: an object-based implementation of Lisp. *Journal of Information Processing*. 13, 3 (1990).

[7] Quigley, M., Conley, K., Gerkey, B. P., Faust, J., Foote, T., Leibs, J., Wheeler, R. and Ng, A. Y. 2009. ROS: an open-source robot operating system. In *ICRA Workshop on Open Source Software,* 2009.

[8] Baker, L. M. and Taylor W. M. 1954. The relationship under stress between changes in skin temperature, electrical skin resistance, and pulse rate. *Journal of Experimental Psychology*. 48, 5 (1954).

[9] Aquest. 2010. AquesTalk – Text-to-speech synthesis middleware. Retrieved October 1, 2010 from: http://www.a-quest.com/download/index.html.

[10] Lazarus. 2011. Free Pascal Lazarus project. Retrieved November 1, 2010 from http://www.lazarus.freepascal.org/.

[11] Schermerhorn, P., Scheutz, M., Crowell, C. R. 2008. Robot social presence and gender: do females view robots differently than males? In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction* (Amsterdam, The Netherlands, March 12-15, 2008).

[12] Evers, V., Maldonado, H C., Brodecki, T. L., Hinds, P. J. 2008. Relational vs. group self-construal: untangling the role of national culture in HRI. In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction* (Amsterdam, The Netherlands, March 12-15, 2008).