Long-Term Human-Robot Interaction: The Personal Exploration Rover and Museum Docents*

Kristen N. Stubbs
Robotics Institute
Carnegie Mellon University
kstubbs@cmu.edu

Debra Bernstein
Learning Research and Development Center
University of Pittsburgh
dlb36@pitt.edu

Kevin Crowley
Learning Research and Development Center
University of Pittsburgh
crowleyk@pitt.edu

Illah Nourbakhsh
Robotics Institute
Carnegie Mellon University
illah@cs.cmu.edu

***Abstract*— As an increasing number of robots have been designed to interact with people on a regular basis, research into human-robot interaction has become more widespread. At the same time, little work has been done on the problem of long-term human-robot interaction, in which a human uses a robot for a period of weeks or months. As the person spends more and more time interacting with the robot, it is expected that the means by which they make sense of the robot and its behavior - their "cognitive model" of the robot - may change over time. In order to examine this cognitive model, a study was conducted involving the Personal Exploration Rover (PER) museum exhibit and the museum employees responsible for them. The results of the study suggest a number of relevant components of a cognitive model for human-robot interaction which will be further examined in an upcoming study on scientist-rover interaction.**

***Index Terms*— human-robot interaction, informal learning, educational robotics**

I. INTRODUCTION

The number of robots designed to interact with humans has increased in recent years, giving rise to the field of "human-robot interaction" as a domain of scientific interest [1]. Within this domain, researchers have designed robots to interact and collaborate with humans in a variety of ways. For example, the Sony *AIBO* is intended for use as a toy [2], *Robovie* was designed to help teach English to Japanese schoolchildren [3], and still other robots have been created with the goal of assisting humans in urban search and rescue [4].

Despite covering a wide range of activities, it is important to note that most of these robots do not interact with their human users for more than a few minutes or hours at a time. However, if robots are being built with the intention of interacting with people over the long-term, it is crucial to investigate how people understand, model, and interact with robots over long periods of time. This is an interesting and challenging research problem as it requires access to robots that will function properly with minimal maintenance for months on end and at the same time have a rich interaction modality with human beings.

*This work is partially supported by NASA/Ames Autonomy, Intel Corporation, and an NSF Graduate Student Fellowship to K. Stubbs

II. RESEARCH GOALS

The primary goal of this research is to help establish how a person's understanding of a robot, their cognitive model of the robot, changes over time. A "cognitive model" refers to how a person thinks about a robot in terms of such things as assessing its capabilities, predicting its behavior, and interpreting and resolving failures. In order to develop this model, we focus on the human user and how he or she makes sense of a robot after a period of regular interaction lasting weeks or months.

While numerous robots have been designed to be used by humans over long periods of time, few long-term human-robot interaction studies have been conducted at this time. *Paro*, a seal-shaped robot, has been used as a means of therapy for both the elderly and for hospitalized children [5], [6]. Ideally, *Paro* might become a therapeutic toy for a sick child for a period of weeks or months, but experiments of that length with *Paro* have not yet been conducted. Similarly, Reference [7] describes an EVA assistant robot designed to aid astronauts, another system which might someday be used for weeks at a time, but it has not yet been studied for such a long period.

To ensure the success of robots designed to interact with humans for significant periods of time, roboticists must be able to understand how a person's cognitive model of a robot works and use this model to improve the quality of robots' interactions with people. Some researchers have already begun to examine how humans and robots interact over long periods of time. One robot that has been studied over a relatively long period of time is *Cero*, developed by [8]. In this study, a motion-impaired user utilized *Cero* to help her carry out various tasks over a period of months; however, this research mainly focused on communication and mediated interaction. The authors found that the morphology of the robot has significant impact on human behavior, particularly when directionality aids humans in inferring the goals of *Cero* in terms of motion down halls, facilitating interaction in close spaces. An open question is the role of such morphology in long-term interaction, where first impression inferences will not dominate.

Reference [9] examined the educational impact of interactions with robots on high school students as part of a summer course on robotic autonomy. While these authors demonstrated significant learning along themes both within engineering disciplines and life-long learning concepts, the analysis results from evaluation of students who have been provided with daily, mediated interaction with robots through instruction and curriculum deployment. In contrast, we are interested in unmediated interactions, during which the trajectory of interaction over the long term is less structured and may result in a very different type of relationship between humans and robots. As humans work with a robot for a period of weeks or months, their cognitive model of the robot may change significantly. This includes such factors as their expectations of what the robot is capable of and what they desire the robot to do. Understanding how this cognitive model changes over time will enable researchers to design robots that are able to have successful interactions with humans over a period of weeks, months, or years.

In order to meet this goal of understanding long-term human-robot interaction, the authors chose to study the Personal Exploration Rover (PER), a small robot designed to operate in science centers across the United States [10]. The PER was an excellent focus for a long-term human-robot interaction study for a number of reasons. The PER was designed to operate in a museum environment under heavy usage for weeks and months at a time. In addition, PERs have been installed in six science centers around the country, making cross-site comparisons possible.

While museum visitors may interact with the robot for a span of only a few minutes, it is museum employees and docents who develop long-term relationships with the PER. These employees interact with the rovers in a variety of ways, including setting them up at the start of the day, changing their batteries, diagnosing and repairing problems, and explaining about the PERs and their exhibit to museum visitors (Fig. 1). These interactions happen every day that the exhibit is open to the public, which may span many months. The reliability and robustness of the PERs combined with their use in museum exhibits around the country provide an ideal setting for observing and analyzing long-term human-robot interaction.

III. The Personal Exploration Rover

The Personal Exploration Rover (PER) is the third rover designed and built as part of the Personal Rover Project [11]. The goal of this project is to design and build interactive robots capable of educating and inspiring children. The PER was designed as a tool to educate the public about certain aspects of NASA's Mars Exploration Rover (MER) mission. The goals of the PER are to demonstrate to the public that rovers are tools used for doing science and to illustrate the value of on-board rover autonomy.

Fig. 1. A docent at the Smithsonian National Air and Space Museum talks about the PER with two young visitors.

Fig. 2. The PER examines a rock at the National Science Center.

Physically, the PER is reminiscent of the MER in its overall mechanical design (Fig. 2). The PER is a six-wheeled robot that uses a rocker-bogie suspension system similar to that used on the MERs. The PER is equipped with a camera and range finder mounted on a pan-tilt head as well as an ultraviolet light for conducting simulated scientific testing.

The PER museum exhibit consists of a PER deployed inside a simulated Martian environment (the "Mars yard") complete with several large rocks as "science targets" and an interactive kiosk, equipped with a track ball and a single button. The premise of the exhibit is that visitors will use the robot to search for life within the Mars yard. The robot is able to test for signs of life using a simulated organofluorescence test, in which the robot shines a UV light on a rock. As the robot conducts the test, it sends a picture of the rock back to the kiosk, where visitors look for a "glow" indicating the presence of (simulated) organic material.

Fig. 3. The mission generation screen of the PER's kiosk interface.

When a visitor presses the button to begin the interaction, the PER takes a 360 degree panoramic image of the Mars yard. Using the panorama and overhead, orthographic map, the visitor then specifies the direction and distance to a target rock (Fig. 3). Once those instructions have been sent to the rover, the PER navigates toward the target rock and autonomously drives up for a close approach. It then turns on its UV light and sends a picture of the rock back to the visitor, who can see whether or not "life" is present.

Besides being designed to educate museum visitors, the creators of the PER made sure to keep the needs of museum employees in mind as well. The PERs were designed to withstand many days of continuous use in museums. As of the present time, the PERs have made over 50,000 autonomous close approaches. One of the major goals of the project was to produce a robot that would not fail despite this kind of heavy use. At the present time, the PERs' mean time to failure exceeds two weeks [10]. In addition, the PERs' diagnostic software make it easy for museum employees to solve many smaller, common problems in-house.

There are a number of different groups of individuals who have had interactions with the PERs since the PER project began. These include the creators of the PERs at Carnegie Mellon University, museum employees at the PER installation sites, and the museum visitors who use the PER exhibit. Reference [12] is a study of how visitors interact with and react to the PER exhibit, but these interactions rarely last more than several minutes. Museum employees, including administrators, explainers, and technical support people, were chosen to be the focus of this study due to their regular interactions with the PERs over a period of months. In addition, museum employees together form a group of naive initial users who will grow over time and develop cognitive models that they initially may not have had. These two characteristics make them a group well-suited for a study of long-term human-robot interaction.

IV. Methodology

In order to determine how museum employees' cognitive changed over time, the authors conducted periodic interviews with museum employees from December 2003 through June 2004. These interviews were conducted once before the PER exhibit had been installed, one to two weeks after the exhibit had been installed, roughly one month after that, and roughly two months after that. Eighteen museum employees at four PER installations were interviewed. While it was hoped that all employees would be interviewed four times, due to their busy schedules, in the end five employees were interviewed only once, two employees were interviewed twice, nine employees were interviewed three times, and two employees were interviewed four times for a total of forty-four interviews.

After the interviews were transcribed, a coding scheme was designed to reflect the museum employees' thoughts about the robot. The development of a coding scheme for categorizing types of utterances with respect to learning and museums can be found in [13] and [14]. Our own coding scheme is based upon both the content of the interviews as well as previous related work. The following nine themes are included in the coding scheme, grouped into three major categories. Each of these nine themes contains a number of sub-codes, but for the purposes of this paper the data have been collapsed up to the super-category level. The three major categories and nine content codes are as follows:

1) Technical talk about the PER
   - Capabilities of the robot
     This theme represents comments about what the PER can and cannot do in terms of its physical components, its behaviors, and its kiosk interface. Comments about the rover's capabilities are classified as to whether they are general and abstract ("It moves") or specific ("I talk about the different servos"). It is hypothesized that as museum employees spend more time working with the PER, their comments will become more specific.
   - Failures
     We hypothesize that how museum employees think about and handle failures is one of the most important aspects of their cognitive model of the PER. This theme is applied to comments about how the robot failed and the ability of employees to diagnose and solve problems.
   - Reliability
     On the other hand, this theme is used to describe comments about the robot's robustness and resistance to failure and whether employees used general or specific language in discussing it.
   - Criteria for intelligence
     This theme focuses on what reasons museum employees give for saying that the PER is intelligent or unintelligent: its programming, its autonomy, or its ability to follow human instructions. Whether or not employees decide that the

PER is intelligent, this theme reveals how they think about the robot and what criteria were most important to them when making this decision.

2) People and the PER
   - Robot anthropomorphization
     This theme encompasses remarks that museum employees make that the PER "wants", "feels", or "knows" something or that employees or visitors are treating the PER as if it were a living being, such as by calling to it or naming it. People's tendency to anthropomorphize robots has been examined by such groups as [15] and [16], but this study gives an opportunity to examine robot anthropomorphization over a longer period of time than this previous work.
   - Visitor description
     This theme is used to characterize comments made by museum employees about how visitors are interacting with the exhibit and how they are treated by employees, either as passive or active learners [17].

3) PER-MER connections
   - Relationship to the MER mission
     This theme is used for comments museum employees make about how the PER is related to the MERs and their mission. This includes statements about how the PER increases visitors' interest in the MER, how the PER itself is similar to or different from the MER, and how the mission that visitors conduct is similar to or different from NASA's MER mission.
   - The role of a robot
     This theme is used to represent how museum employees perceive the role of the PER and/or the MER; whether it is a tool used by humans or a machine that collaborates with humans.
   - Taking different points of view
     This theme encompasses the museum employees' seeing the world from the perspective of the PER or of a NASA mission scientist. This theme is adapted from the theme of "Identification with technology" as introduced by [9]. This theme was found in written comments by students participating in a three-week course on robotic autonomy. As museum employees have worked with the PER for much longer than three weeks, it is hypothesized that they too may start to see the world from the perspective of the robot or from the perspective of the NASA scientists who are also using rovers.

Each line of the interview transcripts was coded with one of these themes as well as a second code indicating what type of relationship or interaction was being discussed. This relationship description code was one of the following:

- Employees themselves (descriptions of employees' work history, educational background, etc.)

Fig. 4. A docent guides a visitor through the process of sending a mission to the PER.

- Employee-robot (talk about using the PER, describing the PER, etc.)
- Employee-visitor (talk about visitor's questions, interacting with visitors, etc.) (Fig. 4)
- Employee-scientist (talk about relating to NASA MER mission scientists)
- Visitor-visitor (descriptions of visitors interacting with each other)
- Visitor-robot (talk about visitors' using the PER, visitors' thoughts about the PER, etc.)
- Visitor-scientist (talk about visitors' relating to NASA scientists)

Overall, out of the lines that could have been given one of the nine thematic codes, an appropriate code was identified for 92.6% of them. Out of the lines that could have been given one of the relationship description codes, 99.8% of them were successfully coded. This high rate of lines that could be described by codes supports the validity of this coding scheme.

V. RESULTS

All together, the forty-four interviews generated 2,821 lines that were coded according to the scheme described above. As only two employees were able to give four interviews, the data from the eleven who were able to complete three interviews were used to compute matched-sample statistics. The percentages of each theme that were recorded for each interview can be seen in Fig. 5.

Using the data from the eleven museum employees who were interviewed three times, a one-way repeated-measures ANOVA was computed to determine whether or not there were statistically significant differences across time, accounting for individual differences between employees.

The results of this data analysis can be grouped according to the three major content categories described above, with focus on technical language about the PER,

Distribution of Content Codes

Code	Interview *1*	*2*	*3*
Capabilities	14.5%	10.9%	13.5%
Failures	17.0%	17.3%	16.7%
Reliability	1.1%	7.1%	6.1%
Intelligence	1.7%	6.4%	4.3%
Anthropomorphization	1.1%	10.1%	18.4%
Visitor description	34.1%	31.5%	35.8%
MER mission	11.1%	8.5%	4.3%
Role of robot	12.2%	4.1%	0.5%
Different POV	7.1%	4.0%	0.5%

Fig. 5. For each interview and content code, the value listed is equal to the ratio of the number of times that that content code was used out of the total number of lines coded.

interactions between the PER and people, and connections between the PER and the MER.

1) Technical Talk about the PER: Over the course of the interviews, there were many significant changes in coding frequencies relating to technical talk about the PER robot itself. Employees talked significantly more about the Reliability ($df = 2$, $F = 5.01$, $p < 0.05$) theme and discussed failures more frequently than any other topic besides museum visitors (Fig. 5). In addition, when describing failures, the use of specific technical terms increased significantly ($df = 2$, $F = 6.73$, $p < 0.01$) without a significant increase in the use of general terminology. At the same time, there were no significant changes in talk about the PER's capabilities. This suggests that as the employees became more familiar with the PER, they tended to focus on the robot's actual successes and failures rather than what it was supposed to be capable of achieving.

2) People and the PER: Not surprisingly, over time, employees talked more about their own interactions with the PER ($df = 2$, $F = 18.8$, $p < 0.0001$) and about their interactions with visitors ($df = 2$, $F = 4.62$, $p < 0.05$). Talk about anthropomorphization and instances in which museum employees anthropomorphized the PERs also increased significantly ($df = 2$, $F = 11.14$, $p < 0.01$) as did talk about why the PERs are or are not intelligent ($df = 2$, $F = 4.43$, $p < 0.05$); however, Anthropomorphization was the only content code that increased across all three interviews (Fig. 5). In addition, talk about anthropomorphization was significantly positively correlated with talk about visitors, reliability, and intelligence ($N = 44$, $p < 0.05$, $p < 0.01$, and $p < 0.01$, respectively). These results suggest that as employees spent more time with the PER, anthropomorphization was an important part of their cognitive model, one that was related to talk about several other key themes.

3) PER-MER Connections: Talk relating the PER to the MER became less frequent as the interviews progressed ($df = 2$, $F = 4.46$, $p < 0.05$). This suggests that the focus of the employees' cognitive models tended to shift away from this higher-level concept over time.

VI. Conclusion

The fact that there were many significant changes in employees' talk about the PERs between the first and second interviews suggests that regular interaction with a robot for even a couple of weeks may have a large impact on a person's cognitive model of the robot.

However, as seen in Fig. 5, the only content code that increased across all three interviews was Anthropomorphization. The fact that more content codes did not exhibit this same trend may be due to a number of factors. The PERs themselves do not exhibit a very wide range of behaviors, and so they may not have required employees to spend a significant amount of time interpreting and adapting to them. In addition, unlike the students in the Robotics Autonomy course, the employees were not challenged to solve a wide variety of problems with the PER on a regular basis. Without the need to apply their knowledge of the PER in a variety of situations, it is possible that employees' cognitive models were not tested in such a way as to cause a greater number of significant changes.

Based on the changes that were observed in this study, some of the key characteristics of a person's cognitive model of a robot include:

- A robot's actual failures and successes may be more important than its purported capabilities
- Anthropomorphism is a broad concept, frequently associated with a number of other concepts, such as reliability.
- Talk about higher-level concepts, such as the idea of robotic intelligence, declined over time but this decrease was matched by an increase in talk about anthropomorphism. This suggests that people may be thinking of the robot less as a machine and more as a collaborator.

As a next step with respect to research on long-term human-robot interaction, a study on the interaction between scientists and a remotely located "robotic astrobiologist" is currently in progress [18]. The scientists will spend approximately three weeks using a rover to search for life in the Atacama Desert of Chile. Both the experiences of the scientists in the US and of the roboticists in Chile will be observed and recorded in order to obtain a complete picture of how this robot-human system operates over time. This new study will examine some of the same themes discussed here as well as additional issues with potential applications to space exploration and other kinds of remote science operations.

This kind of attention to understanding people and how they think about robots is crucial in order to develop technologies that will remain useful to people for long periods of time. Robots themselves may one day be able to use this kind of knowledge to generate more fruitful interactions with the humans around them.

Acknowledgment

The authors would like to thank the staff of the Smithsonian Air and Space Museum, the San Francisco Exploratorium, the National Science Center, and the NASA Ames Visitors' Center for all of the time and energy that they have given to this research. This study would not have been possible without their generous support.

References

[1] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots: Concepts, design, and applications," Carnegie Mellon University, Tech. Rep. CMU-RI-TR-02-29, December 2002.

[2] "Sony Entertainment Robot AIBO." [Online]. Available: http://www.aibo.com

[3] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro, "A practical experiment with interactive humanoid robots in a human society," in *Third IEEE International Conference on Humanoid Robots (Humanoids 2003)*, 2003.

[4] M. Lewis, K. Sycara, and I. Nourbakhsh, "Developing a testbed for studying human-robot interaction in urban search and rescue," in *Proceedings of the 10th International Conference on Human Computer Interaction (HCII'03), Crete, Greece*, June 2003.

[5] K. Wada, T. Shibata, T. Saito, and K. Tanie, "Effects of robot assisted activity to elderly people who stay at a health service facility for the aged," in *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, October 2003.

[6] T. Mitsui, T. Shibata, K. Wada, and K. Tanie, "Psychophysiological effects by interaction with mental commit robot," *Journal of Robotics and Mechatronics*, vol. 1, no. 1, pp. 20–26, February 2002.

[7] R. Burridge, J. Graham, K. Shillcutt, R. Hirsh, and D. Kortenkamp, "Experiments with an EVA assistant robot," in *Proceedings of the 7th International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS-03)*, 2003.

[8] K. Severinson-Eklundh, H. Huttenrauch, and A. Green, "Social and collaborative aspects of interaction with a service robot." *Robotics and Autonomous Systems, Special Issue on Socially Interactive Robots*, vol. 42, no. 3-4, 2003.

[9] I. R. Nourbakhsh, K. Crowley, K. Wilkinson, and E. Hamner, "The educational impact of the Robotic Autonomy mobile robotics course," Carnegie Mellon University, Tech. Rep. CMU-RI-TR-03-18, December 2003.

[10] I. Nourbakhsh, E. Hamner, D. Bernstein, K. Crowley, E. Ayoob, M. Lotter, S. Shelley, T. Hsiu, E. Porter, B. Dunlavey, and D. Clancy, "The Personal Exploration Rover: Educational assessment of a robotic exhibit for informal learning venues," *International Journal of Engineering Education, Special Issue on Robotics Education*, In press.

[11] E. Falcone, R. Gockley, E. Porter, and I. Nourbakhsh, "The Personal Rover Project: The comprehensive design of a domestic personal robot," *Robotics and Autonomous Systems*, vol. 42, pp. 245–258, 2003.

[12] D. Bernstein, "Parent, Docents and Robots: Examining Mediation at a Mars Rover Exhibit," August 2004, in K. Crowley (chair), Islands of Expertise: An Approach to Exploring the Cognitive Ecology of Childhood. Symposium conducted at the meeting of the Visitor Studies Association, Albuquerque, NM.

[13] L. Schauble, M. Gleason, R. Lehrer, *et al.*, "Supporting science learning in museums," in *Learning Conversations in Museums*, G. Leinhardt, K. Crowley, and K. Knutson, Eds. Lawrence Erlbaum Associates, 2002.

[14] K. Crowley and M. Jacobs, "Buildings islands of expertise in everyday family activity," in *Learning Conversations in Museums*, G. Leinhardt, K. Crowley, and K. Knutson, Eds. Lawrence Erlbaum Associates, 2002.

[15] J. Goetz, S. Kiesler, and A. Powers, "Matching robot appearance and behavior to tasks to improve human-robot cooperation," in *Proceedings of the 12th IEEE Workshop on Robot and Human Interactive Communication (RO-MAN)*, 2003.

[16] M. Mori, *The Buddha in the Robot*. Tuttle Publishing, 1982.

[17] S. G. Paris, Ed., *Perspectives on Object-Centered Learning in Museums*. Lawrence Erlbaum Associates, 2002.

[18] "The Limits of Life in the Atacama," 2004. [Online]. Available: http://www.frc.ri.cmu.edu/atacama