Long-Term Human-Robot Interaction: The Personal Exploration Rover and Museum Docents[1]

Kristen Stubbs [a,2], Debra Bernstein [b], Kevin Crowley [b], and Illah Nourbakhsh [a]
[a] *Robotics Institute, Carnegie Mellon University*
[b] *Learning Research and Development Center, University of Pittsburgh*

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 people spend more time with a robot, it is expected that how they make sense of the robot - their "cognitive model" of it - may change over time. In order to identify factors that will be critical to the future development of a quantitative cognitive model of long-term human-robot interaction, a study was conducted involving the Personal Exploration Rover (PER) museum exhibit and the museum employees responsible for it. Results of the study suggest that these critical factors include how people experience successes and failures with the robot (as opposed to how they understand its capabilities) and how people anthropomorphize the robot and talk about anthropomorphization.

Keywords. human-robot interaction, informal learning, educational robotics

1. 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 to assist humans with 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

[1]This work is partially supported by NASA/Ames Autonomy, Intel Corporation, and an NSF Graduate Student Fellowship to K. Stubbs.
[2]Correspondence to: Kristen Stubbs, Robotics Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213; Tel.: +1 412 268 8813; E-mail: kstubbs@cmu.edu.

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.

2. Research Goals

The primary goal of this research is to help establish how people's understanding of a robot – their cognitive model of the robot – changes over time. This work can then be used to help generate a quantitative model of long-term human-robot interaction. In order to identify the factors that will be most important in the development of such a model, this study focuses 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. A number of robots have been created that might eventually be used by humans for long periods of time to provide therapy or other assistive services for humans (see [5], [6], [7]); however, none of these robots have been tested with people for long periods of time. One robot that has been studied over a relatively long period of time is *Cero* ([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' study of the Personal Exploration Rover and museum docents is unique in that it focuses on the relationship between a group of people and a particular type of robot over a period of months, placing emphasis on understanding how the docents' understanding of and interacting with the robot may change over time.

Constructing a complete cognitive model of long-term human-robot interaction is beyond the scope of this study. Instead, the focus of this research is on identifying factors that will play a crucial role in the future development of such models. In order to meet this goal, 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, and PERs have been installed in six science centers around the United States.

3. The Personal Exploration Rover

The Personal Exploration Rover (PER) is the third rover designed and built as part of the Personal Rover Project [12]. 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.

Physically, the PER is reminiscent of the MER in its overall mechanical design (Fig. 1(a)). 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.

(a) The PER examines a rock.

(b) A docent talks about the PER with two young visitors.

Figure 1. The PER at (a) the National Science Center and (b) the Smithsonian National Air and Space Museum.

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 trackball 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. The reliability and robustness of the PERs combined with their use in museum exhibits around the United States provide an ideal setting for observing and analyzing long-term human-robot interaction.

There are a three different groups of individuals who have had interactions with the PERs since the PER project began. These are 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 [13] 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. These interactions include setting up the PERs at the start of the day, changing their batteries, diagnosing and repairing problems, and talking about the PERs and their exhibit to museum visitors (Fig. 1(b)). In addition, museum employees together form a group of naive initial users who will learn 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.

4. Methodology

For this study, the authors' goal was to develop a methodology that would enable them to answer the following types of questions about employees' cognitive models of the PER:

- How does the employee's conception of robot intelligence change over weeks of interaction?

- How do employees anthropomorphize the robot, if at all?
- As employees gain more experience working with the robot, how do their descriptions of its capabilities change?
- How do employees see the connection between the PER and the MER?

In order to answer these questions, the authors conducted periodic interviews with museum employees from December 2003 through June 2004. These open-ended interviews were conducted once before the PER exhibit had been installed, one to two weeks after the exhibit had been installed, approximately one and a half months after installation, and approximately three and a half months after installation. The exact questions asked to employees at each interview varied slightly, but each employee had an equal opportunity to comment on all question topics. Eighteen museum employees at four PER installations were interviewed; of these, only eleven were able to complete three 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 [14] and [15]. The coding scheme presented here 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.
   - Failures
     This theme is applied to comments about how the robot failed and the ability of employees to diagnose and solve problems.
   - Reliability
     This theme is used to describe comments about the robot's robustness and resistance to failure.
   - Criteria for intelligence
     This theme focuses on what reasons museum employees give for saying that the PER is intelligent or unintelligent.
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. Previous work on robot anthropomorphization over the short term can be found in [16] and [17].
   - 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 [18].

Distribution of Content Codes

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

Figure 2. 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. *Indicates a statistically significant change (one-way repeated-measures ANOVA).

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.
   - 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], a study of the educational impact of a course on robotic autonomy.

The interview transcripts were coded according to the procedure used in [9] and [10]. Out of the lines that were eligible for one of the nine thematic codes, 92.6% of the lines were described unambiguously by one of the thematic codes. The high rate of lines that could be unambiguously described by codes supports the validity of this coding scheme and suggests that the scheme fit the data well.

5. Results

All together, the forty-four interviews generated 2,821 lines that were coded according to the scheme described above. The data from the eleven employees who were able to complete three interviews were used to compute matched-sample statistics. This technique of transforming qualitative data into quantitative data is adapted from [11]. The percentages of each theme that were recorded for each interview can be seen in Fig. 2.

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

5.0.1. 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. 2). 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.

5.0.2. People and the PER

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. 2). 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.

5.0.3. 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 model tended to shift away from this higher-level concept over time.

6. 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 has a large impact on a person's cognitive model.

However, as seen in Fig. 2, 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 course on robotics autonomy [9], 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. It is also possible that employees with different roles had different reactions to the robot, but there is insufficient data to support this kind of analysis.

Based on the changes that were observed in this study, some of the key factors that should be considered when constructing a cognitive model of how people understand robots include:

- A robot's actual failures and successes may be more important than its purported capabilities. In order to aid people in developing accurate cognitive models, it is best to keep robot behavior transparent. Providing this transparency into the robot's successes and failures will allow users to develop the best possible cognitive model, one based on their own experiences rather than on extensive pre-training.
- Anthropomorphism is a broad concept, frequently associated with a number of other concepts, such as reliability. While it is clear that anthropomorphization is an important part of a person's cognitive model of a robot, exactly what role anthropomorphism plays in that model remains an open question.
- 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. A quantitative model of long-term human-robot interaction will need to recognize this distinction between "interactive device as robot" and "interactive device as collaborator" as a person moves from one to the other.

To further advance this research on long-term human-robot interaction, a study on the interaction between scientists and a remotely located "robotic astrobiologist" is currently in progress [19]. 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. The next step is to formalize a quantitative model of human-robot interaction. A robot equipped with this model and an adaptive architecture may then be able to generate more fruitful interactions with the humans around it.

Acknowledgements

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] Fong, T., Nourbakhsh, I., and Dautenhahn, K. (2002). A survey of socially interactive robots: Concepts, design, and applications. Tech. report CMU-RI-TR-02-29, Robotics Institute, Carnegie Mellon University.

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

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

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

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

[6] Mitsui, T., Shibata, T., Wada, K., and Tanie, K. (2002). Psychophysiological effects by interaction with mental commit robot. *Journal of Robotics and Mechatronics*, 1(1), 20–26.

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

[8] Severinson-Eklundh, K., Huttenrauch, H., and Green, A. (2003). Social and collaborative aspects of interaction with a service robot. *Robotics and Autonomous Systems, Special Issue on Socially Interactive Robots*, 42(3-4).

[9] Nourbakhsh, I., Crowley, K., Wilkinson, K., and Hamner, E. (2003). The educational impact of the Robotic Autonomy mobile robotics course. Tech report CMU-RI-TR-03-18, Robotics Institute, Carnegie Mellon University.

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

[11] Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, 6(3), 271–315.

[12] Nourbakhsh, I., Hamner, E., Bernstein, D., Crowley, K., Porter, Hsiu, T., Dunlavey, B., E., Ayoob, E., Lotter, M., Shelley, S., Parikh, A., and Clancy, D. (2004). The Personal Exploration Rover: The ground-up design, deployment and educational evaluation of an educational robot for unmediated informal learning sites. Tech. report CMU-RI-TR-04-38, Robotics Institute, Carnegie Mellon University.

[13] Bernstein, D. (2004). Parent, docents, and robots: Examining mediation at a Mars rover exhibit. *Islands of Expertise: An Approach to Exploring the Cognitive Psychology of Childhood*, K. Crowley (Chair). Symposium conducted at the meeting of the Visitor Studies Association, Albuquerque, New Mexico.

[14] Schauble, L., Gleason, M., Lehrer, R., Bartlett, K., Petrosino, A., Allen, A., Clinton, C., Ho, E., Jones, M., Lee, Y., Phillips, J., Seigler, J., and Street J. (2002). Supporting science learning in museums. *Learning Conversations in Museums*, G. Leinhardt, K. Crowley, and K. Knutson (Eds.) Lawrence Erlbaum Associates.

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

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

[17] Mori, M. (1982). *The Buddha in the Robot*, Tuttle Publishing.

[18] Paris, S. (2002). *Perspectives on Object-Centered Learning in Museums*, Lawrence Erlbaum Associates.

[19] The Limits of Life in the Atacama (2004). http://www.frc.ri.cmu.edu/atacama.