The Personal Exploration Rover: Educational Assessment Of A Robotic Exhibit For Informal Learning Venues*

ILLAH NOURBAKHSH
The Robotics Institute, Carnegie Mellon University, Pittsburgh, P.A., USA and NASA/Ames Research Center, Moffet Field, CA, USA. E-mail: illah@ri.cmu.edu

EMILY HAMNER, ELLEN AYOOB, ERIC PORTER, BRIAN DUNLAVEY
The Robotics Institute, Carnegie Mellon University, Pittsburgh, Pa, USA

DEBRA BERNSTEIN, KEVIN CROWLEY
Research And Development Center, University Of Pittsburgh, Pittsburgh, Pa, USA. E-mail: crowleyk@pitt.edu

MARK LOTTER, SKIP SHELLY
LotterShelly, Pittsburgh, PA, USA

THOMAS HSIU
Gogoco LLC, Sunnyvale, CA, USA

DANIEL CLANCY
NASA/Ames Research Center, Moffet Field, CA, USA

Robotics brings together learning across mechanism, computation and interaction using the compelling model of real-time interaction with physically instantiated intelligent devices. The project described here is the third stage of the Personal Rover Project, which aims to produce technology, curriculum and evaluation techniques for use with after-school, out-of-school and informal learning environments mediated by robotics. Our most recent work has resulted in the Personal Exploration Rover (PER), whose goal is to create and evaluate a robot interaction that will educate members of the general public in an informal learning environment and capitalize on the current enthusiasm and excitement produced by NASA’s Mars Exploration Rovers (MERs).

We have two specific goals: teaching about the role of rovers as tools for scientific exploration and teaching about the importance of robot autonomy. To this effect we have designed an interactive, robotic museum exhibit that has been deployed at five locations across the United States, including the San Francisco Exploratorium and the Smithsonian National Air and Space Museum. Here we introduce the robot hardware and software designed for this task and the exhibits developed, then detail the educational assessment methodology and results, which encompass exhibit impact on museum visitors at two installation sites.

INTRODUCTION

CRITICAL ENABLING TECHNOLOGIES for long-term, high competence mobile robotics have made significant strides over the past few years. In conjunction with this greatly increased potential for mobile robots to interact intelligently with humans, human–robot interaction is experiencing significant growth as a field of scholarly endeavour [12,13]. Through the Personal Rover Project, we have focused specifically on the application of interactive, physically embodied robotic technology to informal learning environments [11]. This agenda has been motivated by our and others’ quantitative and anecdotal results that show that educational robotics can trigger significant learning across broad learning themes that extend well beyond STEM (science, technology, engineering and mathematics) and into associated lifelong skills of problem-solving and communication [2, 10, 14, 16, 17, 22, 23, 25, 26, 27, 29].

Educational robotics, while a fast-growing and important present-day endeavour, has concentrated primarily upon mediated, formal learning venues. Robot contests such as BotBall and US FIRST provide mediated structure for students in classroom settings and after-school programs [28,29]. Formal integration of research robots and field robot prototypes into the curriculum has also been quite successful, where time with the robot is rare and therefore valuable and carefully structured [5, 6, 21]. Intensive, challenge-
based curriculum has even been shown quantitatively to demonstrate statistically significant broad learning acquisition, as prior work in the Personal Rover Project has shown [23]. In the present project our focus was to explore the role of technology in learning in the context of short-term, unmediated interactions as can be found in the high-volume setting of science museums. The first challenge was to choose a specific application.

Motivated by the broad expected exposure of the Mars Exploration Rover (MER) missions, targeted to land in January 2004, we elected to launch a technology-based educational experience that would be widespread in the informal learning venue of a number of science centers across the country. This ambitious level of implementation demands robotic technology that can survive robustly without expert roboticists on call.

Dubbed the Personal Exploration Rover (PER), our resulting interactive science rover experience is meant for prolonged use in unmediated settings, by novice users, without demonstrating the fragility and susceptibility to failure often seen in interactive robotics devices. The PER is designed to meet its specific educational objectives within the context of the NASA MER missions. These objectives are:

- to show that rovers are tools for doing science by enabling visitors to act as mission scientists, using the PER to conduct a science operation; and
- to enable visitors to appreciate the role of autonomy on board rovers.

In the hope of evaluating these educational objectives, science centers offer a prime venue because these informal learning spaces offer both transient and long-term interaction opportunities over a sufficiently large body of visitors such that statistically meaningful conclusions regarding interaction and education can be drawn.

The PER exhibit was designed from the ground up by a team led by Carnegie Mellon University consisting of government, industry and academic partners. NASA/Ames and Intel Corp. provided funding; Intel also provided the Intel Stargate arm-based single board computer. Gogoco and Lotter-Shelly provided professional mechanical design and graphic design. Botrics provided electronics engineering services. The Learning Research and Development Center (LRDC) provided formal educational evaluation.

The Personal Exploration Rover has been designed as a robotic introduction to the technologies that enable NASA’s missions and as an immersive tool for experiencing the challenges faced by NASA mission scientists. The PER exhibit installations, aimed specifically at the informal learning environment of science museums and tech museums, present museum visitors with the challenge of searching for signs of life on discrete rocks placed in a physically instantiated Mars yard. Using a carefully designed user interface to communicate with the rover, visitors interpret panoramic imagery and orthographic, overhead imagery to identify their science target, then observe as the PER approaches the rock, scans to find the target’s exact position, maneuvers autonomously for a close approach, then conducts an ultraviolet test for organofluorescent signs of life (Fig. 1).

Installations operated at five national science centers in early 2004, including the Smithsonian National Air and Space Museum (NASM) and the San Francisco Exploratorium. Operation continues at several sites, including NASM, and will spread further in future months. In the first two months of 2004, Personal Exploration Rovers effected more than 20,000 autonomous science target approaches as directed by museum visitors. More than 30 miles of rover travel were completed, with idle times approaching 0% of museum operating hours at the Exploratorium. Key enabling technology advances include the areas of power management, terrain inference, science target approach and software architecture. This paper describes the specific results of educational analysis of the PER exhibit. First, however, we present contextual information regarding robot design and interaction design, both completed de novo for the purposes of the PER exhibit.

**ROBOT DETAILS**

The mechanical chassis of the PER (Fig. 2) loosely resembles the configuration of the two MER instances currently exploring Mars. Like MER, there is a six-wheel suspension supporting a rectilinear electronics box. An ultraviolet (UV)-fluorescent light mounted on the exterior of this box enables the PER to illuminate target rocks in order to test for organofluorescence. The lid of the electronics box is shaped to be reminiscent of the ‘winged’ solar panels on the MER deck. The lid supports the PER’s camera and optical infrared (IR) range-finder. These are mounted together on a pan and tilt head that is on a short mast at the front of the rover. The camera is used for panoramic imaging and close-up target imaging. The rover uses the rangefinder to scan for obstacles in
its path during traverses and to identify the exact distance and bearing to target rocks. Based on prior results regarding diagnostic transparency, great care was taken to design the PER’s pan/tilt head so that it clearly demonstrates the robot’s direction of attention. This static design aesthetic, combined with appropriate dynamics as the head pans and tilts to search for obstacles and science targets, facilitates inferences made by museum visitors regarding the level of attention PER pays to its surroundings. For example, as the PER moves forward its head continuously scans left and right, aiming the rangefinder at the terrain the rover is about to traverse. Visitors easily recognize that PER is ‘looking for targets’ or ‘looking for obstacles’ even though they may be wholly unaware of the specific sensor mechanisms used by PER. A catadioptric, parabolic mirror assembly would have enabled 360° vision and thus obviated the need for a pan/tilt mechanism, but such inferences of capability and internal robot processing would have been unacceptably sacrificed.

Unlike the MER, the PER was designed to be relatively inexpensive so that many PERs could be built for multiple simultaneous exhibitions at an affordable price point, as with previous Personal Rover Project robots [15]. Rather than designing the PER to have a similar scale to the MER platforms, we chose to minimize the size of the PER, subject to off-the-shelf microprocessor, sensor and motor constraints, so that relatively small museum Mars yards would nevertheless yield rich interactions. Overall, the height of the PER is approximately 36 cm, the length is 33 cm and the width is 34 cm. The approximate weight, fully loaded, is 15 lbs. All time-limited parts used on the PER were designed to be easily replaced. For instance, each position-controlled joint (of which there are six in total) is powered by an unmodified, stock servomotor used by various hobby communities. These parts would prove to be the sole source of repeated repair and, due to their off-the-shelf nature, museums were able to replace servos in-house.

The PER’s main processor, the Intel Stargate board (www.xbow.com), runs the Linux operating system and communicates via 802.11b wireless Ethernet with the Java-based mission control interface running on a PC. Wireless communication combined with a battery pack that can power the rover for approximately 10 hours enables the PER to operate without a tether, making it a more realistic emulation of the MER. For greater detail concerning PER hardware and software architecture refer to [24].

EXHIBIT INTERACTION

A multidisciplinary team consisting of interaction designers, roboticists, and programmers collaborated to design and implement the intended museum interaction. Three goals were set for the exhibit; as well as supporting the project goals of teaching about robot autonomy and robots as scientific tools, the interaction should be easily completed by visitors in less than three minutes in order to facilitate throughput in view of visitor flow requirements.

Although a series of static storyboards were used to identify candidate interaction trajectories, a critical aspect of the exhibit interaction design process involved real-world sparse testing. Before
the final rover hardware was complete, a prototype four-steer robotic vehicle was fabricated for preliminary testing (Fig. 3). This prototype would serve multiple purposes simultaneously. First and foremost, this prototype used the candidate servomotors, drive gearmotors, rangefinder, USB camera device and microprocessors selected for the PER, serving as a burn-in system test for these off-the-shelf components. Second, this prototype exhibited the same kinematic motion capabilities of the final system, enabling high-fidelity testing of the interaction system even though the final PER instantiation would not be complete for several months. Several cycles of public usage of the prototype rover using candidate interfaces were completed, helpfully identifying the most critical adaptations of the interface required for smooth operation by untrained users. The interaction trajectory described below, together with the final design solutions, embody the conclusions drawn from this series of iterative test and refinement cycles [3,18].

Museum interaction

A typical museum interaction begins when the visitor presses the button on the kiosk. The rover then takes a 360° panorama, which is displayed on the kiosk screen (Fig. 4). The user selects a target rock by clicking on the panoramic image. This identifies the angle to the target. The user’s next step is to select the location of the rover and the target rock on a ‘satellite map’ in order to estimate the distance from rover to target. When users are satisfied with the mission specifications they send the mission to the rover for autonomous execution.

The rover first turns to face the target, then drives the specified distance, all the while checking for obstacles in its path by panning and tilting its head as it moves forward. Upon reaching the end of the path, the rover scans the area in front of it to locate the rock. If it finds the rock, the rover will do a series of adjustments and scans to ensure that it is well aligned with the target. It then drives to within a few centimeters of the target and turns on...
its UV light to analyze the rock for organofluorescent signs of life. This is simulated with UV fluorescing paint, which has been applied to a subset of the yard’s rocks. The rover sends an image of the rock back to the user for scientific analysis, and the user makes the final determination of whether there is evidence for life on the rock.

Design solutions
To maximize the users’ learning experiences and create a fun and educational interaction, the designers focused on the interface language, interaction cues, physical orientation, real-time feedback, and the visual interface. Through rapid prototyping of the designs and a series of informal user tests, the team was able to quickly eliminate problematic concepts and arrive at the following sampling of solutions.

Interface language. The prospective audience can potentially cover a broad range of scientific expertise, so minimal formal scientific and technical terminology is used. Instead, a simple, inquisitive, game-like tone supports the interaction.

Interaction cues. The default screen display in the kiosk is a loop that provides a visual overview of the impending mission and what the user might be expected to do. The kiosk itself has a track ball and a button, similar to an arcade game. The mission begins when the user presses the button. A linear interaction follows as the mission is progressively disclosed to the user.

Physical orientation. To help the user orient between the Mars yard and the screen display (Fig. 5), a Martian sun is painted on the wall of the Mars yard; it is visible from both the kiosk and in the panoramic view on screen. In addition, the rock positions, rock shapes, and the shape of the yard provide feedback and help users interpret the orthographic map. An animation is used to communicate the 360° nature of the panoramic image.

Real-time feedback. A Mission Builder screen display (Fig. 4) was created to reinforce the educational aspects of mission building. The display tracks users’ progress in real-time until they are ready to submit the mission to the rover. As the rover executes the mission, a rover’s-eye view camera allows the visitor to experience the mission from the rover’s perspective. The Rover Mission subwindow at bottom right remains during execution, providing data regarding rover operations, distance traveled and angles turned.

Visual interface. A consistent color palette is used to unify the screens. Static and animated elements on the screen are designed to provide focal points for the users, depending on the actions required. Consistent, clear typography provides visual hierarchy and improves readability [4].

MUSEUM INSTALLATIONS
To date the PER exhibit has been deployed at five museum locations across the USA: the Smithsonian National Air and Space Museum, the Smithsonian Udvar-Hazy Center, the San Francisco Exploratorium, the National Science Center, and the NASA/Ames Mars Center. For a two-week period the Exploratorium also shared their exhibit with the Randall Museum. The exhibits opened between December 29, 2003 and January 24, 2004 and ran for two months or more. As of January 2005 the Udvar-Hazy and NASA/Ames exhibits continue to operate and the exhibit is scheduled to open at the Japan World Expo in March 2005. Although each museum was provided with guidance regarding exhibit construction (i.e. Mars yard fabrication, kiosk fabrication), variation in both exhibit design and execution has been significant across installations, leading to the potential for comparative analyses of the effectiveness of identical robotic technology as implemented in a variety of modes. The most distinct areas of variation are in interaction format and Mars yard design, summarized below.

Interaction format
The format of the exhibit in terms of docent activity is left up to the individual museum. As a result we have observed three different styles of interaction. At NASM, interaction with the exhibit is fully mediated by a dedicated docent. At the Udvar-Hazy Center, the exhibit is used for structured teaching activities with school groups. The Exploratorium, National Science Center, and NASA/Ames allow visitors to explore the exhibit without mediation and in a freeform manner. This variation in the level of guidance, most extreme between the Smithsonian and the Exploratorium, justified joint educational analysis of these two installations, as described in the Exhibit Analysis section.

Mars yards
Each museum designed and produced its own Mars yard or yards for the exhibit, subject to yard
design constraints expressed by the PER team to ensure exhibit success. The Mars yards are specifically designed with the PER’s capabilities and the desired exhibit interaction in mind. The rocks and hills in the terrain are all traversable by the rover, demonstrating the animation of its rocker-bogie suspension system, except for four to five very large rocks, which serve as the scientific targets. The yards are surrounded by hip-height walls decorated with Martian landscapes and horizons from NASA’s Pathfinder mission. Each yard also displays a sun on one wall to help the visitors orient themselves when using the exhibit. This sun icon is apparent when viewing the physical yard, when viewing the rover-generated panorama and is iconically represented on the ‘satellite map’ overhead view of each yard. Thus across all three representations of physical space the sun serves as a landmark for orientation and familiarity. As reported earlier, early user tests identified confusion regarding rover orientation to be a principal obstacle to effective human–rover interaction; thus the need for such a clear landmark across all representations.

Museums have one to two yards ranging from 256 square feet to 72 square feet. The yards are constructed in various ways, including spray painted Styrofoam; layered paint, glue, sand, wood and plaster; small lava rocks and sand; and layered Styrofoam, polymesh and dryvit compound (Fig. 6). Of particular interest is the fabrication methodology used by the Smithsonian NASM and Udvar-Hazy sites. Local high school students, working in teams, researched the topography of the Mars Pathfinder landing site, then recreated a portion of this landing site as a school project using shaped foam and stucco. Thus the Mars yard creation process itself was transformed into a learning experience and outreach opportunity by these museums.

EXHIBIT USE PATTERNS

Quantitative statistics regarding exhibit use were collected automatically at installations by the exhibit software itself and by sampled passive observation. Both quantitative results and informal observations guided the more formal educational exhibit evaluation that followed. These statistics identify the demographics of the exhibit users and the manner in which the exhibit was used. Significantly, the statistics show that time on task is extremely close to the design target of 3 minutes and, more importantly, virtually all exhibit users were able to complete the entire mission successfully. Together these statistics indicate that the distribution of time on task is not, as is often the case in museum exhibits, exponential but rather unimodal and narrow. Users who are engaged by the PER exhibit remain engaged through mission completion, then helpfully release control to the next museum visitor in queue. Details of both user demographics and mission use statistics follow.

Audience

Exhibit use observations were conducted at the Exploratorium and the National Air and Space Museum. At both locations, the exhibit was in nearly constant use. Over roughly 4.5 hours of observation, 184 people interacted with the exhibit. This included 71 adult users (36 females and 35 males), and 113 child users (28 females and 85 males). The majority of exhibit users were in groups, and the average group size was 3.06 ($\sigma$ 1.22), with a total of 64 groups using the exhibit during this period. Group members often took turns conducting rover missions. Although more boys than girls were present at the exhibit, 61% of boys and 71% of girls attending the exhibit operated the rover.

Mission statistics

Based on logs automatically generated by the Exploratorium and NASA Ames kiosks between December 29th, 2003 and April 14th, 2004 we are able to report additional information about exhibit use. [All of the kiosks generate logs, but these results are based upon NASA/Ames and Exploratorium analyses only.] The exhibits were in use 75.4% of the time while they were open (331 hours idle and 1017 hours in use). Out of 26 200 missions, only 525 (2.0%) timed out before the end of the Mission Builder screen, meaning that 98% of users were able to design a mission and send it to the rover successfully. When a mission is unsuccessful, users are given the option of trying again or quitting. Only 499 (1.9%) missions timed out at this stage, showing that users were highly engaged even when their mission failed to find the target rock. The average mission length was approximately 2 minutes 20 seconds (139.7 seconds $\sigma$ 60.1 seconds). This is the length of time for a single set of instructions to be selected by the user, sent to the rover, and executed. On average each user engaged the PER in 1.6 missions ($\sigma$ 0.94), thus the overall individual time on task is approximately 4 minutes, significantly exceeding
the 1.4-minute engagement time typically seen at interactive science exhibits [9].

About half of the missions (52.7%) ended with the rover successfully locating a rock (Fig. 7). The next most common outcome was the detection of an obstacle (23.1%), meaning that the rover encountered an obstacle more than 150 centimeters from the expected target distance. The rover went ‘out of range’, i.e. detected a hill wall blocking its path, only 18.1% of the time. In 3.4% of the missions, the mission ended due to a robot error such as failed communication. The rover was unable to locate any rock or hill wall 2.7% of the time.

In summary it is clear both from time on task values, time-out rarity and mission success rates that visitors are able to make use of the PER exhibit effectively, even in the unmediated cases of the Exploratorium and NASA/Ames installations. It is further clear that for children, there is no obvious statistical gender gap in terms of engagement with the PER exhibit. Both of the above conclusions are encouraging in that the PER exhibit attracts and engages the target population. The next question, addressed in the following section, is whether this exhibit uses technology in an educationally positive manner.

EXHIBIT ANALYSIS

Traditional school-based assessments of learning are often inappropriate for use in informal learning environments [1]. As groups of visitors use and talk about exhibits, they are constructing a shared understanding of the content. Following recent theoretical and empirical work in museum learning [8, 20], our analyses focus on this naturally occurring talk as the best indicator of whether the exhibit is successful in terms of its educational goals.

The PER exhibit is interesting to a wide range of visitors, but we focus here upon impact for one of the most common user groups: families visiting the museum with children. In this paper we first analyze videotapes of families using the exhibit in order to describe the extent to which their conversations reflect the intended educational themes. Second, we analyze post-exhibit interviews with children in order to describe the extent to which they understood those same themes after using the exhibit.

Method

Research was conducted at the Exploratorium in February 2004 and at the Smithsonian National Air and Space Museum (NASM) in April 2004. At the Exploratorium families interacted with the exhibit on their own, although staff were generally available to answer visitor questions. At NASM a docent was stationed next to the control kiosk, in order to provide information about the PER (and MER missions) and to assist visitors as they engaged with the exhibit. Thus, these sites provide a contrast in how the exhibit functioned in stand-alone vs. supported environments.

We analyze the activity of 43 families recruited at the two target sites: Twenty-nine at the Exploratorium and 14 at NASM. For recruiting purposes, a ‘family’ was defined as a parent or guardian (over age 18) and at least one child between the ages of 4 and 14. The average age of children at the Exploratorium was 8.8 years (SD = 2.1; range = 4.8–12.1 years). This sample included 12 girls and 17 boys. The average age of child participants at NASM was 8.8 years (SD = 1.1; range = 6.9–10.3 years). This sample included 4 girls and 10 boys. [Due to camera failure, data from one of the NASM child interviews was not usable.]

Participants at the Exploratorium spent an average of 6 minutes, 38 seconds at the exhibit, of which 5 minutes, 1 second was spent at the kiosk, operating the rover. Exploratorium participants completed an average of 2.3 missions, of which 55% were successful. Participants at NASM spent an average of 15 minutes, 9 seconds at the exhibit, of which 4 minutes, 18 seconds were spent at the kiosk. NASM participants completed an average of 1.4 missions, of which 88% were successful.

Families were approached at the entrance to the exhibit in each museum, and invited to participate in the research study. Interested families were asked to sign a consent form. Participating families were videotaped as they used the exhibit (including while they waited in line to operate the PER). In order to record exhibit conversations, one child in each family was asked to wear a wireless microphone. Upon completion of exhibit use, one child and one parent from each family were interviewed separately.

The child interview consisted of a set of open-ended questions about the Mars mission, the MERs, and the PERs. At the beginning of each interview, children were shown pictures of Spirit and Opportunity, the Mars Exploration Rovers, and asked to identify the rovers and the goal of
their mission. Children were then asked to explain how they thought the rovers worked. For example, children were asked to predict how action is initiated for the rovers, whether the rovers needed to be ‘smart’ to accomplish their goals, whether the rovers were capable of autonomous behavior, and why NASA would decide to send robots (instead of astronauts) to explore on Mars. Questions about autonomy and whether the rovers were ‘smart’ were repeated verbatim for the MER and PER. The question about initiating action was only asked of the MER. For the PER, children were asked to describe what they did in the exhibit, and whether or not the PER had a successful mission. When children reported that the rover or the person controlling the rover was responsible for the mistake. The average length of child interviews was 6 minutes, 23 seconds at the Exploratorium and 7 minutes, 50 seconds at NASM.

The parent interview also consisted of a set of open-ended questions regarding the MER missions: Parents were asked to describe what they knew about the MER; their family’s level of interest in the MER missions; and what they thought their child learned from the PER exhibit. Due to space limitations in the current article, we report analyses of the parent interview data in [24].

Results

In this paper, we focus on the question of how the exhibit supported its two stated educational objectives: (1) allowing visitors to explore the role of robots in mission science; and (2) enabling visitors to appreciate the nature of robot autonomy. We will describe conversational coding schemes we created and applied to the interaction and interview data. Unless otherwise specified, comments about the MER and the PER were given equal weight in coding. Reliability was assessed by comparing codes from two independent raters on 20% of data. Inter-rater reliability for each coding scheme exceeded 85%.

The role of robots in mission science.

One of the goals of the PER exhibit was to provide a tangible connection to the unfolding story of the search for signs of life on Mars. This story includes both the possibility of finding life on Mars, and the excitement of using robots to conduct exploration. We developed four coding categories to capture exhibit talk related to the role of robots in mission science.

The first coding category captured exhibit talk about the MERs and the goals of their mission. An example of this type of talk is provided below.

Now that is what they sent to Mars . . . . I heard last night they were running one of the wheels so they could make a trench. (Parent, Exploratorium)

A second category included direct comparisons made between the design and capabilities of the PER and MER.

The real ones on Mars don’t go much quicker than this. (Docent, NASM)

. . . . you noticed this one had a light in front of it to do its science? The real one actually has an arm that reaches out and checks out the rock. (Docent, NASM)

The next two categories were created to capture talk about robots as part of a collaborative team. The third category focused on communication, specifically the mediating nature of programming and telecommunications.

So you’re going to pretend that you’re gonna be one of those computer guys, okay, and you’re going to do some signals so that the rover can move around like it was on Mars. (Parent, Exploratorium)

A fourth category, Collaborating with Robots, captured talk about how robots and people can work together and exchange information.

If you look on the computer screen, it shows you what the camera on the rover is seeing. (Parent, Exploratorium)

So did you have to give it an exact directional . . . . or do you just say there’s a rock over here and it locks on the rock? (Parent, Exploratorium)

So now it’s going to ask you to make a map for it. (Parent, Exploratorium)

Conversations at the PER exhibit

Table 1 presents the percentage of conversational groups discussing each topic, broken down by museum. [As a unit of analysis, the conversational group includes anyone present at the exhibit with the child. At the Exploratorium, the conversational group generally included the child, parent(s), siblings, and any other exhibit users with whom the child interacted. At NASM, the conversational group included the child, parent(s), siblings, other exhibit users, and a docent. These data suggest that the PER exhibit supported conversations about the Mars mission and general robotics at both sites. However, conversational groups at NASM, which included a docent, were significantly more likely to talk about the Mars science.

<table>
<thead>
<tr>
<th>Themes</th>
<th>Exploratorium</th>
<th>NASM</th>
</tr>
</thead>
<tbody>
<tr>
<td>About the Mars Mission*</td>
<td>55%</td>
<td>93%</td>
</tr>
<tr>
<td>Comparisons between MER and PER*</td>
<td>24%</td>
<td>79%</td>
</tr>
<tr>
<td>Communicating with robots</td>
<td>45%</td>
<td>72%</td>
</tr>
<tr>
<td>Collaborating with robots</td>
<td>86%</td>
<td>93%</td>
</tr>
</tbody>
</table>

*Indicates a statistically significant difference between the Exploratorium and NASM groups.
mission and to make explicit comparisons between the MER and the PER."

Further analysis revealed that parents generally initiated the same amount of thematic talk at both the Exploratorium and NASM exhibits (with the exception of talk about collaboration with robots, which was initiated more often by parents at the Exploratorium), and that the additional talk observed at NASM was coming from the docents. As one might expect, this additional docent talk was often general and driven by the script docents used. In contrast, parent comments, particularly in the Exploratorium, were more often specific and targeted to child experience. For example, a mother and 4-year old boy were getting ready to initiate a new mission at the Exploratorium. After setting up the mission, the mother turned to her son to encourage him to push the ‘Go’ button and begin the mission: ‘OK, now look, you want to tell him to go?’ The child nodded, leaned over to the Mars Yard, looked straight at the PER, and shouted: ‘Go!’ Even though they had just used the interface together, the mother realized that her son did not really understand that the computer was mediating the human–robot interaction. So, she tried to address his misconception: ‘Look at that, he’s following directions (points towards yard). You communicated with him through the computer. . . . you were able to give him accurate directions, just by moving and clicking.’ Additional analysis on the differences between parent and docent talk at the PER exhibit is provided in a later section.

Child interviews

The four categories described in the previous section were also used to code children’s post-exhibit interviews. Regardless of site, children came away from the exhibit demonstrating fairly high levels of thematic knowledge. Almost all children demonstrated basic knowledge of the MER mission (Exploratorium, 93%; NASM, 100%) and of collaborating with robots (Exploratorium, 97%; NASM, 100%). Most children (Exploratorium, 72%; NASM, 69%) were also able to describe devices people can use to communicate with robots (e.g., computers and, in the case of rovers in space, satellites). Although we never directly asked children to compare MER and PER, 21% of Exploratorium children and 38% of NASM children made spontaneous comparisons between the two. None of these differences were statistically significant.

In conjunction with the conversational analysis, these findings suggest that the mission-based exhibit format was successful in encouraging visitors to engage with the idea of robots as partners in scientific exploration. Because we did not pre-test children, we cannot make strong causal claims about learning from the PER. However, we can make strong claims about the exhibit being successful in supporting specific connections to the MER missions, suggesting that the PER was a catalyst for conversations that were probably based on news accounts of the ongoing MER missions. From the perspective of the museum community, where exhibitions take years to develop and are rarely linked to current events, the PER exhibit demonstrates an innovative strategy for informal science education.

The nature of robot autonomy

The second main objective of the exhibit was to help visitors explore rover autonomy. Although museum visitors will come to the exhibit with some prior knowledge about robots, most have probably not interacted with a robot that possessed true autonomous properties [19]. Thus, the exhibit provides a unique opportunity for visitors to re-evaluate concepts of robots that have perhaps been built largely upon fictional autonomous robots (e.g., R2D2 and C3PO) or non-autonomous robots (e.g., industrial robots).

We developed three coding categories relevant to the goal of appreciating rover autonomy. The first category, Rover Design, included talk about the technology used to build rovers, rover size, and the importance of rover autonomy. For example:

See it [PER] has two motors. One is at the wheels . . . to move it forward, and the other is the other on the top, which is to turn the wheel. You see— it has two motors. The design is very simple, actually. (Parent, NASM)

The second coding category captured talk about the types of activities rovers could perform, such as taking pictures and examining rocks.

There’s some pictures that it’s taking. (Parent, Exploratorium)

Now the rover’s starting his mission, so what he’s doing is taking pictures all the way around himself to create a 360 degree panorama. (Docent, NASM)

The final category captured talk about the autonomous activities of the rovers. This category included discussions of rovers sensing things in the environment (e.g., looking for rocks), rovers avoiding obstacles, planning their own routes, and achieving goals with minimal user input.

This rover also has a great deal of autonomy, meaning he can think for himself . . . he’s going to go the distance you gave him. When he’s done following your commands, then he does the thinking by himself to find the rock. (Docent, NASM)

Conversations at the PER exhibit

As shown in Table 2, visitors at both museums were coded as addressing all three themes, although each was addressed significantly more frequently at NASM. Analysis of the source of exhibit conversation revealed that parents at both the Exploratorium and NASM discussed these topics with similar frequency. As for talk about robots and mission science, the presence of docents was responsible for the increased frequency of thematic talk about autonomy at NASM.
Child interviews

Children’s interview transcripts were first coded using two of the categories described above: rover design and rover activities. Children were able to speak knowledgeably about rover design at both the Exploratorium (52%) and NASM (77%) and to speak about rover activities at both Exploratorium (55%) and NASM (85%). Although there was a suggestion that children at NASM were more likely to demonstrate knowledge in these two categories, neither of these differences proved significant.

In constructing a measure to assess children’s ideas about the third theme, rover autonomy, we needed to account for the fact that children were often inconsistent and uncertain when deciding whether a robot would be capable of particular autonomous behaviors. To do this, we constructed an autonomy score. For each statement indicating comprehension of the autonomous operations of the rover, a child was given one positive point. [The following references were used in order to develop guidelines for coding statements as autonomous: Smothers (1997); The Mars autonomy project: www.frc.ri.cmu.edu/projects/mars; Wikipedia, online encyclopedia: http://en.wikipedia.org/wiki/Autonomous_robot; What is autonomy technology?: http://ic.arc.nasa.gov/projects/remote-agent/activities/polo/docs/mission/1-whatis-autonomy-tech.html.] For each statement indicating the opposite belief, namely that the rovers were incapable of independent action and operated via remote control, a child was given one negative point. These points were summed independently for statements about the MER and PER, thus each child was assigned two autonomy scores.

Examples of statements from different children that were coded as indicating rover autonomy, and thus receiving a positive point, included:

- I clicked it and it didn’t go far enough and it [PER] looked around and it found the rock anyway. (10 year-old boy, Exploratorium)
- It [MER] has a smarter capability to say, and it’s able to move around those. It can detect an obstruction and it will go around it instead of going straight through it. (10 year-old boy, Exploratorium)
- I think they [MERs] might need to do things for themselves because if the computer crashes they have no way of contacting it. So it must have a boot up or something that will make it go by itself and know what to do. (10 year-old boy, Exploratorium)

Examples of statements that received a negative autonomy point were:

- He [PER] can move when we tell him to do and when we don’t tell him to he doesn’t move. (8 year-old girl, Exploratorium)
- People probably have to tell it [MER] how much to, how many degrees to turn and how much more to go, and maybe control the instruments. (10 year-old boy, NASM)
- Because it [MER] doesn’t know which rock because it doesn’t have any eyes. It only has a camera that the humans are controlling so only they’ll know where it is. (9 year-old boy, NASM)

As shown in Figs. 8 and 9, between a third and a half of the children came away from the exhibit with a positive autonomy score for PER (Exploratorium, 41%; NASM, 46%) and MER (Exploratorium, 31%; NASM, 54%). As one might expect, PER and MER scores were significantly correlated ($r = 0.48, n = 42, p = 0.001$). The differences in scores between museums were not significant.

Additional analyses were conducted to look for potential relationships between children’s autonomy scores and the other categories of robot/Mars mission talk described above. This was only done...
with data from children at the Exploratorium, as there were too few children from NASM to allow a further breakdown of the data. Analysis revealed that children with positive autonomy scores (i.e., children who described the rovers as capable of some autonomous action) were more likely to make comparisons between the MER and the PER. This was true for children with high PER autonomy scores, \( \chi^2 (1, N = 34) = 5.02, p = 0.025 \), as well as high MER autonomy scores, \( \chi^2 (1, N = 33) = 7.34, p = 0.007 \). As autonomy is an important commonality between the MER and the PER, perhaps children were more likely to make comparisons between the rovers when they were aware of their autonomous attributes. Children with positive PER autonomy scores were also somewhat more knowledgeable about rover design, \( \chi^2 (1, N = 34) = 4.14, p = 0.042 \). However, it is important to note that there were no other significant relationships between children’s autonomy scores and other topical categories. It would seem that an understanding of robot autonomy is potentially available for any child who comes to use the exhibit, regardless of their prior knowledge about the Mars mission or about robots in general.

**How parents and docents talked about autonomy**

Research in the field of museum learning suggests that parents can serve an important bridging function between what a museum intends for children to understand about an exhibit and what children actually do understand [7]. Of course, that bridging function is also the primary job description of museum docents. But while docents and parents may find themselves in similar roles in a museum, each group brings unique skills to the task. Museum docents are often trained in the content of the exhibit, while that is rarely the case for parents. On the other hand, parents are much more familiar with their child’s interests, knowledge, and learning history than are docents. The goal of the current analysis is to determine if these differences led parents and docents to approach exhibit content in different ways.

In this section we report our analysis of how adults (parents and docents) talked to children about autonomy. Robotic autonomy statements were chosen to undergo additional coding for two reasons: (1) autonomy is a difficult concept, and parents and docents used a variety of strategies to explain it to children; and (2) an understanding of robotic autonomy is an important learning outcome for the PER exhibit. Each statement was coded in terms of two dimensions. Statements were categorized as either referencing a specific instance of rover activity (targeted), or as a general statement about the rover (general). Statements were also categorized as pointing out an autonomous feature of the rover (feature-level autonomy; e.g., the ability to locate rocks or avoid obstacles) or introducing the concept of autonomy at a higher conceptual level (high-level autonomy; e.g., definitions of autonomy). Examples of these different types of statements are provided below.

- He just looked around to see if he could find the rock that you wanted him to go to. (Targeted, feature-level autonomy statement by an Exploratorium parent)
- There it goes. Oh, it does the thinking itself. (Targeted, high-level autonomy statement by an Exploratorium parent)
- It won’t run into the wall, because it’s got sensors that will tell it . . . that will stop it before it gets to the wall. (General, feature-level autonomy statement by a NASM docent)
- So you have to tell the rover where to go, and it has to be smart enough to go find it on its own. (General, high-level autonomy statement by a NASM docent)

Eleven parents (10 from the Exploratorium, 1 from NASM) and 13 docents (1 from the Exploratorium, 12 from NASM) made statements about autonomy during exhibit interactions. The 11 parents produced a total of 14 autonomy statements, while the 13 docents produced a total of 37. Eighty-six percent of parent autonomy statements and 59% of docent autonomy statements were targeted to a specific instance of rover activity. Seventy-nine percent of parent autonomy statements and 59% of docent autonomy statements were targeted to a specific instance of rover activity. Seventy-nine percent of parent statements addressed autonomy at the feature level, as did 57% of docent statements.

It would seem that the training that the docents received allowed them to generate more high-level
autonomy statements at the exhibit. However, the
finding that parents are more likely to target their
statements about autonomy to specific instances of
rover activity is consistent with previous research
that parents provide ‘just-in-time’ explanations for
scientific phenomenon [8]. Such explanations allow
parents to provide children with information at the
moment it is needed and to shape children’s inter-
pretation of what they are doing and seeing
in the museum. Additionally, these explanations
build upon the shared experience between parent
and child, leaving open the possibility of the
families following-up on the information at a
later date.

Summary

This assessment suggests that the exhibit was an
effective forum for involving visitors in explora-
tions of the role of robots in mission science and of
robots as autonomous agents. Analysis of family
conversation suggests that visitors were expanding
on relevant themes as they used the exhibit.
Families talked about the ongoing Mars mission,
they compared the MER and PER, they discussed
communicating and collaborating with robots, and
they talked about robot design, technology, and
autonomy. Interviews with children following the
exhibit suggested that almost all children were
aware of the MER missions and that many of
them also were able to connect the exhibit experi-
ence in specific ways to the mission. This finding
suggests that the format of the exhibit, with chil-
dren conducting their own missions, was effective
both in holding visitor attention and communicat-
ing educational content.

Children did not end their experience with a
uniformly robust view of autonomy. Although
some recognized autonomous characteristics of
the rovers, most children held inconsistent theo-
ries. More than half still held views that the
rovers are primarily operated through remote
control. We did not necessarily believe that a
single exhibit experience would be a sufficient
base for children to develop fully correct theories
of autonomy. Rather, the exhibit experience is
probably best seen as a chance for families to
work out some of these issues in the context of
an authentic autonomous rover. It may be the case
that making the autonomous functions of the
rover more explicit, either by providing signage
to direct visitor’s attention to the rover’s autono-
mous capabilities, or by providing a direct explana-
tion of robotic autonomy, would help families
explore this concept more effectively.

CONCLUSIONS

The Personal Exploration Rover has served as a
rewarding demonstration of educational robotics
applied to the informal learning space. Given
current goals in relation to the NASA Mars
Exploration Rover mission, this team designed a
new educational rover from the ground up, tested
and refined a graphical interaction system,
engaged multiple high-traffic museums across the
country, shepherded installation and maintenance
of the resulting exhibit and performed quantitative
and qualitative evaluation of the exhibit’s efficacy.

In summary, this project demonstrates that robotic
technology has compelling value in the museum
setting, and that concrete educational results can
be achieved and measured in such a setting.

Our lessons learned from this experience span
the disciplines of robot morphology, interaction
design, and mission design to maximize educa-
tional efficacy. In the area of interaction design,
three most important lessons learned concern
language, mission design and visualization.

Increasingly inquisitive, game-like screens demo-
strated a friendliness that attracted and retained
users more effectively than ‘realistic’ technically
designed interfaces deployed during early user
tests. In terms of mission design, we found that a
focused narrative and linear mission progression
both focused and guided interaction to a reward-
ing conclusion. Initial user tests with a more open-
ended mission scenario failed to provide as much
of a compelling ‘voyage’ for the visitor. Finally,
in terms of visualization, robot perspective-taking
is an altogether novel activity for museum visitors,
and so the concept of robot-relative orientation
and reference frames were a significant stumbling
block in early testing. Providing visualizations that
juxtapose the robot together with obvious physical
landmarks such as uniquely shaped rocks and a
sun on the horizon were critical clues to help users
orient and formulate correct plans.

When the above three problems are overcome,
one can provide an experience sufficiently compel-
ling such as to change the dynamics of user
engagement times, leading to relatively long times
on task with a sharp and unimodal distribution.

Further exhibit statistics suggest that, among
children, girls and boys are both engaged by this
robotic exhibit, to such a degree that virtually all
users succeed in the completion of an entire
scientific rover mission. Educational evaluation
suggests that the exhibit effectively serves as a
platform for family discussions about the MER
missions and robotics, and that children come
away from the exhibit with measurable knowledge
in these areas.

As robotic technology advances, such interdisci-
plinary teams of engineers, interaction designers
and education specialists will be capable of invent-
ing and executing ever more compelling exhibits
and curricula for both formal and informal learn-
ing venues. We hope that this project can serve as a
motivation for future teams not only to research,
dream and invent, but also to harden, fabricate
and install so that thousands can benefit from
these educational technology ventures.

Acknowledgements—The PER project would not have been
possible without funding and equipment provided by NASA/
Ames Autonomy and Intel Corporation. We would like to thank
all those at the museums who believed in this project and helped bring the PER to the public, especially: Doug Baldwin, Corinne Cannon, Jeff Cross, Maylene Duenas, Jim Frye, Robyn Higdon, Andrew McClellan, Nicole Minor, Victor Morales, Jennifer O’Brien, Mike Reeves, Tom Roach, Priscilla Strain, Nick Veronico, Noel Wanner, Ollie Washington, Steven Williams, and Cheryl Zimmerman. We would also like to thank the following individuals for their contributions and support: Jim Butler, Catherine Eberbach, Edward Epp, Rachel Gockley, Jean Harpley, Anuja Parikh, Kristen Stubbs, and Peter Zhang.

REFERENCES

Illah Nourbakhsh is an Associate Professor of Robotics in The Robotics Institute at Carnegie Mellon University. He is on leave for the 2004 calendar year and is at NASA/Ames Research Center serving as Robotics Group lead. He received his Ph.D. in computer science from Stanford University in 1996. He is co-founder of the Toy Robots Initiative at The Robotics Institute. His current research projects include educational and social robotics, electric wheelchair sensing devices, believable robot personality, visual navigation and robot locomotion. At the Jet Propulsion Laboratory he was a member of the New Millennium Rapid Prototyping Team for the design of autonomous spacecraft. He is a founder and chief scientist of Blue Pumpkin Software, Inc. Illah recently authored the MIT Press textbook, *Introduction to Autonomous Mobile Robots*.

Emily Hamner received her BS degree in Computer Science from Carnegie Mellon University in 2002. She is currently a Research Assistant at the Robotics Institute at Carnegie Mellon University where she manages the Mobile Robot Programming Lab. She serves as project coordinator for the PER project. Her research focuses on interface design, human–robot interaction, and the application of robots in education.

Debra Bernstein is a doctoral student in the Cognitive Psychology program at the University of Pittsburgh. She received an MA degree in Educational Psychology from Columbia University. Her past work has included research and evaluation of educational television and other media. She is currently a graduate student researcher at the University of Pittsburgh Center for Learning in Out of School Environments (UPCLOSE) at the Learning Research and Development Center. Her research focuses on technology in education and learning in informal environments.

Kevin Crowley is an Associate Professor of Cognitive Studies and Cognitive Psychology at the University of Pittsburgh, and the director of the University of Pittsburgh Center for Learning in Out of School Environments (UPCLOSE) at the Learning Research and Development Center. He received his Ph.D. in Psychology from Carnegie Mellon University. His current research focuses on family learning about science and technology in out of school settings, the development of scientific thinking, and other issues in cognitive development. His recent publications include: *Designing for Science: Implications from Everyday, Classroom and Professional Settings* (2001), co-edited with C. Schunn and T. Okada, and *Learning Conversations in Museums* (2002), co-edited with G. Leinhardt and K. Knutson.

Mark Lotter is co-founder of LotterShelly LLC. Mark develops solutions that support business objectives and delight customers. Often the solutions integrate digital and traditional technologies. Mark has years of experience improving customer interfaces to complex information spaces for clients in a variety of public sectors as well as in research. Mark has successfully used the design process to restore the viability of ‘old economy’ companies—those that need to transform their information systems, product and services to secure a position in a changing marketplace. He has also helped new technology-driven companies bring their ideas to market. His work includes design for the Robotics Institute at Carnegie Mellon University, the Abreon Group, ABN AMRO, CoManage and the Software Engineering Institute.

Skip Shelly is co-founder of LotterShelly LLC. Skip has led interdisciplinary design teams to envision and develop solutions that meet business goals and affect customers. He has used design to unify the efforts of executives, technologists, product managers and marketing in global investment banking, financial services, business intelligence, defense, network security, energy, healthcare, civil organizations, industrial services and consumer products sectors. Skip elevates the design process to a strategic asset that creates new business value. He excels at seeing possibilities that are hidden in plain sight and supporting their adoption. He is skilled at developing business, facilitating teamwork and building a client’s capability to design on their own. Skip has worked with global, regional and local clients including Kodak, ABN AMRO, Dun and Bradstreet, PNC Bank and 3rc.

Thomas Hsiu is founder and president of Gogoco LLC, a small design consultancy with specialties in mobile robotic platform design and fabrication. He has contributed to several
previous CMU sponsored robotic research projects including the Trikebot and My Personal Rover (MPR). His experiences range from the development of consumer electronics to the creation of animatronics for motion pictures, such as Anaconda, Deep Blue Sea and The Perfect Storm. He holds an MS and BS in Mechanical Engineering from Stanford University with emphasis on Product Design and Smart Product Design.

Eric Porter received his BS and MS degrees in Computer Science from Carnegie Mellon University in 2003 and 2004, respectively. While at Carnegie Mellon, he performed research in mobile robotics and computer vision. He is currently employed by Oracle Corporation.