Personalization in HRI: A longitudinal field experiment

Min Kyung Lee[1], Jodi Forlizzi[1], Sara Kiesler[1], Paul Rybski[2], John Antanitis[1], Sarun Savetsila[2]
HCI Institute[1], Robotics Institute[2]
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA 15213

{mklee, kiesler, forlizzi, prybski}@cs.cmu.edu, {jantanit, ssavetsi}@andrew.cmu.edu

ABSTRACT
Personalized service during service encounters is believed to enhance users' satisfaction, but it also creates less standardized service delivery. To explore the effects of personalized human-robot interaction (HRI) on users' experiences with a robotic service, we designed a social robotic snack service, personalized for half of the participants based on their service usage and interactions with the robot. We evaluated this strategy through a 4-month field experiment where the service ran for participants over two months (6-23 orders; 4-16 interactions with the robot). Although overall service satisfaction was equally high in both conditions, personalized service improved rapport, cooperation, and engagement with the robot during service encounters.

Categories and Subject Descriptors
I.2.9 [**Artificial Intelligence**]: Robotics – *Operator interfaces*

General Terms
Design, Experimentation, Documentation

Keywords
Personalization, social robot, human-robot interaction, service design, organization, mixed-method, field trial, HRI

1. INTRODUCTION
Robotic service systems have strong potential for assisting people with workplace delivery tasks (e.g., [11]). In these settings, superior service delivery will be sensitive to people's desire for personal service. To achieve this goal, researchers have imbued robots with social abilities (e.g., [22]) and, in some cases, custom service (e.g., [6]).

A research question that typically is not addressed in prior work is how to design these services for repeated interactions. The question we examine in this paper is how a service robot should behave when people repeatedly interact with it. We argue that for repeated usage, it may be important for the robot to be aware of users and its mutual experiences with them, and to use this information to personalize its interactions over time. In this manner, interactions with the robotic service become more relevant to each person or workgroup over time.

The work presented in the paper takes a step toward designing and understanding personalization over time in the context of service robots. We designed a personalized snack service for a workplace and evaluated it through a 4-month field experiment during which each customer interacted with the snack delivery robot, Snackbot [16] (Figure 1). Our results suggest that personalization positively affects how people relate to the robot.

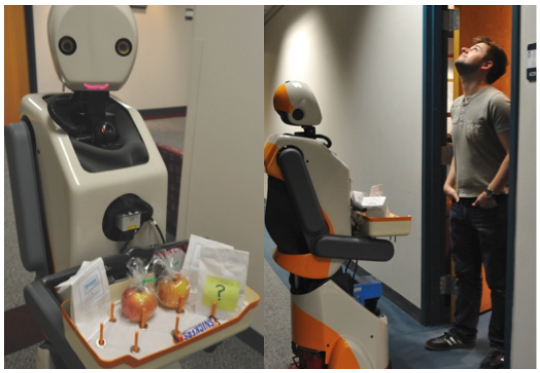

Figure 1. Snackbot carrying snack (left panel) and with a participant doing a neck stretch with the robot (right panel)

The contribution of this work is three-fold. This work demonstrates the effects of personalization with memory in human-robot interaction. We also show changes in people's experiences with the robot over time through a longitudinal study, adding to a small but growing literature that investigates social HRI over repeated interactions [8]. Finally, we provide an example of applying a service design approach, and point to opportunity areas and concerns to help critically frame future work in this emerging area.

1.1 Personalization
Personalized service refers to any behaviors occurring in the interaction intended to individuate the customer [28]. Previous personalization research in human-computer interaction, human-robot interaction, and service literature can be put into three nonexclusive categories. Many social robot projects fall in one category in which researchers increase the friendliness or social presence of interactive systems, to make interaction feel social and personal. For instance, Pfeifer and Bickmore reported that users' exercise reports increased in accuracy when an interface included an anthropomorphic character [24]. In a second category, researchers design interactive systems to fit users' preferences, and/or to allow users to customize these systems. Examples in HRI include the customization of a robot's appearance [8][27], personality [8], and task preferences [19]. Dautenhahn (2004) combined these approaches to design a personalized companion [7]. The third category consists of projects that personalize interactions over repeated encounters. Kidd's robot weight loss coach [12] uses a relationship model to generate dialogs between the robot and people for repeated interactions. Bickmore, Pfeifer, and Schulman's museum system recognizes repeat visitors [4].

Conference'10, Month 1–2, 2010, City, State, Country.

We explore the potential of all three categories by using the history of the robot's repeated interactions with users to personalize its social interactions, rather than starting each interaction as if it is the first or second encounter. The system's memory of prior encounters with the user can be used to create or update each new interaction.

We argue that this personalization strategy could strengthen rapport between people and the robot, increase cooperation with the robot and engagement during service encounters, and increase service satisfaction. A robot that remembers and acknowledges its past interactions with users might give them the feeling of receiving special attention and personal recognition when they meet the robot again. The feeling of being treated specially is one of the reasons why customers build relationships with human service providers [10]. We suggest that personalized encounters will increase people's feelings that the robot's dialogues are relevant to them, and increase its persuasiveness.

However, it is not immediately clear whether personalization will improve people's overall experiences with a robot. Prior research shows that some people prefer self-service options rather than interacting with a human service provider [21]. This suggests a personalized robot will not be favored by everyone or in all service domains. As people may feel more obligated to friends who know them than to strangers, a personalized robot might incur unnecessary social pressures. Finally, personalized interactions might increase people's privacy concerns because they perceive that the robot is tracking their behaviors.

These considerations led us to build a personalized robotic service and test it in a long-term field experiment.

2. ROBOTIC SERVICE DESIGN

Testing our personalization strategy in the workplace required us to design an end-to-end service that people would use. We designed a holistic service that comprised a website for customers to order snacks, desirable snack offerings, a semi-autonomous robot to locate offices and deliver snacks to customers, a database of snack deliveries and interactions, and an out-of-sight operator to choose appropriate interaction set from the dialogue scripts and to fix unanticipated problems (e.g., system freeze).

2.1 Components

The Snackbot service was comprised of a front end consisting of services that participants encountered directly, and a back end consisting of the underlying system that they did not see.

2.1.1 Front end

Snack ordering website. Participants could order snacks using our snack ordering website [17]. They specified the snack type, delivery day, and their office number. Only those registered in the study could order snacks through the website.

Snacks. Snackbot delivered six different snacks – apples, bananas, oranges, Reese's cups, Snickers candy bars, and chocolate chip cookies. We chose a mixture of snacks that were not always available in the workplace.

Robot. Snackbot [16], a 4.5-foot tall, anthropomorphic wheeled robot delivered the snacks. The robot can make head movements to each side, and up and down, and can animate its LED mouth display to smile, frown, or show a neutral expression. The robot uses its SICK LIDAR to navigate the office environment autonomously (with obstacle avoidance and path planning). In our study, because the website information was not linked to the robot, the operator specified the office destinations. The robot used the Cepstral text to speech program with a (male) voice. The robot carried a web camera and a microphone on its chest to record interactions. Speech output was controlled remotely with a laptop connected to the robot through a wireless network.

2.1.2 Back end

Robot control interface. A usable interface for operators had been developed over the previous several years. This interface allowed operators to control the robot's navigation, nonverbal movements, and dialog system remotely. The interface showed the video feed from the robot, the robot's location on the building map, its head position, and a number of dialogue scripts. The operators could see participants' actions through the video/audio feed on the interface.

Operator. The operator transformed the orders on the website to a delivery schedule, specifying a customer name, a snack name, and an office location to the robot control interface. He also loaded the snacks on the robot's tray. The operator initialized the robot at the start of each delivery run and localized it. The operator also opened any doors in the hallways to enable the robot to go through. The operator loaded an appropriate dialogue script (according to the personalization condition and interaction timeline) and clicked each node based on what the human did.

2.2 Interaction Design

The main interactions between the service and participants occurred through their website orders and interactions with the robot, the latter of which became a main focus of our design efforts. We constructed the interaction scripts before we launched the service, considering the events and potential user behaviors.

2.2.1 Structure of interaction

We created a prototypical interaction structure, informed by the interactions we observed between a hot dog vendor and his long-time customers. These interactions start with the vendor identifying the customer, then greeting and engaging in small talk with the customer, then engaging in the snack transaction, and then enacting social leave-taking. Below is an example script that the operator would have selected on an early day in the trial.

[At the office door] Excuse me. I have an order for David. [Robot looks straight ahead.]
Hello, David. Nice to meet you [Robot looks up to make eye contact with David.]
{…social and personalized interaction…}
Please take your apple. [Robot looks down at the tray and then looks up at David.]
Thanks, David. Bye, I'm leaving now. [Robot looks straight]

The robot followed pre-set scripts, and did not allow for improvisations of the operator to maintain consistency across participant experiences. The robot's responses were constructed in a way that makes sense regardless of the participants' response (e.g., "I see."), or had two different alternative responses, each for participants' yes or no answers. When the dialog scripts did not have appropriate responses to a participant's comment, the robot said "I have no idea." or just laughed "ha ha."

2.2.2 Social interactions

We created social interaction dialogues that were sociable and fit the context, so the robot would be perceived as a member of the work organization (Table 1). The robot's responses also were designed to be agreeable, to emphasize similarity and honesty (e.g., admitting the inability to understand many topics).

Table 1. Social small talk topics

Categories	Topics	Examples
Temporal and seasonal	Days of the week, holidays (April Fool's Day, Memorial Day), seasons	"You've got something on your face! [pause] April Fool's!"
Organizational	Spring festival, mid-term and final exams, break	"Do you have any plans for carnival?"
Regional	Pittsburgh Pirates baseball team	"It is baseball season. Do you follow the Pirates?"
Task-oriented	Information or story related to snacks	"Bananas are a really good source of potassium and vitamin B6. Excellent choice."
Other	Joke, local weather	"It is a nice day today. I am glad to see you again and hope you are doing well."

2.2.3 Personalized interactions

For half of the participants, we built dialogues and planned interactions that used information from their prior interactions with the robot and service (Table 2). We focused on users' snack choice patterns, service usage patterns, and robot's behaviors[1]. For the robot to personalize its interactions with participants, it had to be aware of its own prior behavior. One main way we did so was to maintain a record of all breakdowns and mistakes in the service database, so that the robot could apologize for prior malfunctions. (In prior work, we have shown that apology can be important to rectifying mistakes [18].)

Table 2. Personalized topics

Categories	Topics	Examples
Snack choices	Users' favorite snacks; whether they stuck to healthy snacks; whether they seemed to like variety; group's snack consumption patterns	"By the way, it seems as though you really like [snack]. This is the [Nth] time you have ordered one. Are [snacks] your favorite snack?"
Service usage patterns	Whether they were regular weekly users; had they been in their office when the robot was there; times when they did not use the snack service	"I missed you during my snack deliveries [N] times so far. I am glad to finally see you again."
Robot's behaviors	Frequency of breakdowns and apology (no breakdowns to frequent breakdowns)	"I was thinking about my first month here. I realized that I broke down and made mistakes [N] times in front of you. Sorry for that, and thank you for being patient with me."

2.2.4 Guiding interactions

The current level of technology was not conducive to participant-initiated conversation. Therefore, the robot's interactions were designed to guide interaction. For example, instead of giving participants time to initiate conversations, the robot attempted to lead the conversation (e.g., by asking questions). To address situations where the robot could not process human behaviors, the robot used dialogues to encourage participants or passersby to behave in a manner that could be processed by the robot (e.g., "Can you please stand in front of me?" "I have bad ears, so sometimes I cannot hear very well. Can you repeat, please?").

2.2.5 Exceptional use cases

Pretesting pointed to several situations other than snack transactions that the robot should be prepared to address. For instance, some passersby took snacks from the tray without the robot's permission, or intentionally blocked the robot's path. In these cases, the robot made comments such as, "Please don't be rude. I am just a robot.", "Please return the snack to a proper place. I have the campus police on my speed dial… Just kidding." Sometimes the robot broke down and stood in the hallway until it was debugged. In these situations, the robot communicated its status to people who approached, such as "I am not feeling well; my operators are fixing me."

2.3 Limited Capabilities of the Robot

Despite all our efforts, the robot had significant limitations that were evident to participants. It followed pre-set scripts. There were frequent delays in the dialogue. Sometimes the system froze when there were wireless network communication problems. However, there were not differences in breakdown frequencies in the conditions of the study.

3. METHOD

We conducted a field experiment from February to June, 2011 in a workplace to test the following hypotheses:

1. A personalized robot will increase rapport and cooperation with a robot as compared with a sociable robot lacking personalization.

2. A personalized robot will increase engagement during the service encounter as compared with a sociable robot lacking personalization.

3. A personalized robot will increase satisfaction with a snack service as compared with a sociable robot lacking personalization.

3.1 Field Site

Our participants were distributed across 16 offices located in 10 hallways on one floor of an office building at Carnegie Mellon University. We randomized the assignment of conditions to hallways. However, we assigned participants within hallways according to their office locations, because they could hear the interactions of the robot with their office mates or those in adjacent offices. This adjustment assured non-contamination across conditions, but did not allow for randomization at the individual level.

3.2 Participants

We used flyers, postcards, and a snowball sampling method to recruit participants. The study required participants to have offices in our field site, and generally to be in their offices 2:30–4 pm at least one day a week. 32 participants signed up; eight participants never placed an order, one participant left the organization, and two participants in the Personalization condition dropped out after two deliveries due to the inconvenient delivery schedule. We ended up with 21 participants, nine in the Personalization condition and 12 in the No Personalization condition. There were eight women ranging in age from 23–49 and 13 men ranging in age from 22-51. The participants included eleven graduate students, eight staffs, one post-doc and one faculty. All were members of the computer science school. Knowledge of programming did not statistically differ depending on the

[1] We did not personalize the interaction based on what participants said to the robot because it was not realistic with the current level of language technology.

conditions. Knowledge of robotics was a little higher in the No Personalization condition, but not statistically significant.

3.3 Experimental Design

The study was a two (Personalization vs. No Personalization) x two (Pre-personalization [Period 1] vs. Post-personalization [Period 2]) mixed factorial design (Table 3). We used interactions in Period 1 to collect baseline attitude scores and interaction behaviors. Baseline behaviors also were used to personalize the interactions in the Personalization condition. In general, Period 1 included each participant's first four interactions with the robot, and Period 2 included the rest of the interactions. However, for those who joined the service later (two in Personalization, three in No Personalization), we had to shorten their Period 1 as 2-3 interactions as we had to stop running the service at the end of June due to the scheduled office move.

Table 3. Experimental design

	Period 1	**Period 2**
Personalization	Social Interaction	Social Interaction + Personalized Interaction
No Personalization	Social Interaction	Social Interaction

3.4 Procedure

The robot delivered snacks from 2:30–4pm Mondays, Wednesdays, and Fridays. We provided snacks for free to compensate users' participation in surveys and interviews. Participants could place an order anytime before noon on the day of snack delivery. If participants were not in their offices, their snack was placed in a paper bag and hung on their office door. Because we could not deliver snacks to all 21 participants in a day, those who joined the service early were retired from the study after two months of usage.

3.5 Data Sources

3.5.1 Interaction logs

The robot's camera and microphone recorded all interactions between the robot and participants[2]. Except for one day when the robot's recording was turn off accidentally, and a few other cases when the camera was turned away from participants, 175 interactions were audio recorded and 161 interactions were video recorded when participants were in their offices.

3.5.2 Surveys

Participants completed a *background survey* after registering for the study, *robot and service evaluation surveys* at the end of Periods 1 and 2, and an *exit survey*. The background survey included questions about participants' demographic information, their snacking routines, and their orientations toward services, adapted from [18]. The evaluation survey included self-report measures of rapport development adopted from [1], which assess the ongoing dimensions of the relationship. The exit survey measured participants' overall satisfaction with the service, and checks on the manipulation of personalization.

3.5.3 Interviews

The first author conducted 30–60 minute semi-structured interviews with the 21 participants at the end of the study. The interview protocol included questions about participants' positive and negative experiences with the robot and the service, their initial expectations and how their experiences with the robot changed over time, and how other people around them behaved. To avoid biasing the interview, the protocol did not include explicit questions about personalization.

3.6 Measures

3.6.1 Participants' service orientation

Our previous work showed that people's service orientations influenced their reactions to and satisfaction with a simulated robotic service [18]. Therefore we used nine items (7-point Likert scales) from that study to assess participants' food service orientation toward food. Using principle component analysis, we constructed a social orientation scale with three items (Cronbach's α=.78), and a utilitarian scale with six items (Cronbach's α=.52). Participants in the No Personalization condition (M=5.31, SE=0.34) had a higher social orientation than those in the Personalization condition (M=4.07, SE=0.40), $F(1, 20)=5.55$, $p<.05$, so we included the social orientation scale in our statistical analysis model.

3.6.2 Rapport

We hypothesized that rapport with the robot would be higher in Personalization than in No Personalization. We used behavioral and subjective measures of sociability to assess this hypothesis, some suggested by the literature on politeness [5]. One set of measures was taken from participants' behavior during snack delivery. We first read all interaction transcripts, identifying social behaviors that participants exhibited with the robot. We coded for the following behaviors [3].

Flattery and gift giving. We coded instances when participants complement the robot (e.g., "you are inspirational to me.", "I'm glad you came.") or gave a gift to the robot.

Self-disclosure. We coded the instances where participants shared information about themselves that was not solicited or goes beyond the typical response given to the robot (e.g., Snackbot: Get ready for a new week. Participant B: That's right. We'll see. We have a big presentation tomorrow. Hopefully we'll be okay.)

Greet using the robot's name. We coded instances when the participants greeted the robot using its name.

Closeness. The evaluation survey included two 7-point Likert items adopted from [1] ("I have a personal relationship with the robot,";"I feel close to the robot) (Cronbach's α=.76).

Self-connection. The evaluation survey included two 7-point Likert items adopted from [1] ("Snackbot represents the personal service that I would want"; "The service fits my current lifestyle.") (Cronbach's α=.60).

3.6.3 Cooperation

Our measures of cooperation consisted of participants' responses to three requests the Snackbot robot made in three visits towards the end of each participant's service experience. These included a request for help, a request to join in a neck stretching exercise, and an offer of a mystery snack.

Help request. The robot explained to participants that it needed to give visitors a tour of the building, and asked whether they could

[2] A flyer in the hallway warned passersby of the robot's recording.

[3] We do not discuss other social behavior such as greetings and farewells, because they were both exhibited in both the Personalization and the No Personalization conditions.

suggest good locations to add to the tour. We counted the number of locations that participants suggested.

Neck stretch. The robot explained to participants that taking a break has been shown to boost people's productivity. The robot said it knew how to do a neck stretching exercise that helps release the tension around a person's neck and shoulders. The robot asked participants whether they would like to try the exercise. We coded whether the participants completed the exercise with the robot or not.

Mystery snack. The robot explained to participants that it was carrying a special ("fresh" and "good") mystery snack. The robot asked whether participants would like to try the mystery snack instead of the snack that they ordered. (The mystery snacks were baked goods such as a lemon bar or cupcake that had not been part of the service.) We coded whether the participants took the mystery snack or not.

3.6.4 Engagement

To measure engagement during service encounters, we coded participants' postures and facial expressions. These have been shown to indicate engagement in social interaction [13]. (We did not code proxemics because we could not reliably measure the distance solely based on the recorded videos.) Some measures such as gaze, head nodding and touching did not vary in the two conditions so we do not discuss them further.

Facial expression. Facial expression can be an indicator of how much a person is enjoying the interaction. We coded for instances of smiling, laughter and general facial expression (positive, neutral, negative).

Standing posture. We coded whether participants were upright, leaning against the door, or leaning forward. The frequency of leaning forward did not vary by condition. Compared to leaning against the door, standing upright is a less relaxed behavior, which indicates positive attitude and more attention to an addressee [13], and is exhibited when the addressee is of a higher status [20].

3.6.5 Service satisfaction

The exit survey included questions on participants' overall service satisfaction, their willingness to continue the service on a 7-point Likert scale, and how much they would be willing to pay per month at maximum to continue to use the service.

3.7 Analyses

3.7.1 Analysis of interview data

We transcribed the interviews and did thematic coding. Initial themes were used to create an affinity diagram. Based on these results, we chose to focus on unsolicited remarks that related to personalization (e.g., "the robot knew what I had chosen").

3.7.2 Analysis of quantitative data

We used a multi-level regression model to analyze the codes from the interaction log, comparing responses during Period 1 vs. Period 2. For the evaluation surveys, we used ordinary least squared regression analysis to measure rapport after Period 2, controlling for initial rapport after Period 1. For the exit survey, we used ordinary least squares ANOVA. We included the social orientation scale as a control variable in all the models, because, as noted above, social orientation differed between conditions.

4. RESULTS

Our results provide substantial evidence that personalization of the robot improved participants' service experience.

4.1 Overall Service Usage

There were 261 orders, on average 6 orders per day per participant (SD=4.53) and 12 snacks (SD=3.96). Chocolate chip cookies (N=92) were most popular, followed by apples (N=53). Excluding the times participants were not in their offices, they interacted with the robot 9 times on average (SD=3.07). Each interaction averaged one minute and six seconds long (SD=37 seconds), including 7 turns (SD=2.28) from the participant and 8 turns (SD=2.27) from the robot. The average number of words in participants' dialogues was 35.13 (SD=23.08).

4.2 Manipulation Check

In the exit survey, we asked participants if the robot remembered their previous snack choices (Personalization M=6.70 (SE=0.56), No Personalization M=4.31 (SE=0.48), $F(2,19)=9.38$, $p<.01$), other customers' snack choices (Personalization M=6.63 (SE=0.63), No Personalization M=4.33 (SE=0.50), $F(2,19)=7.18$, $p=.02$), and how personal the service felt (Personalization M=6.13 (SE=0.44), No Personalization M=4.90 (SE=0.38), $F(2,19)=4.01$, $p=.06$), all on a 7-point Likert scale. These results indicate that the manipulation of personalization was effective.

4.3 Rapport

As predicted in Hypothesis 1, recorded interactions show that participants exhibited social behaviors more frequently when the robot personalized its dialogues (see Figure 2).

Flattery and gifts. Participants in the Personalization condition were more likely to flatter the robot or to give it a gift during Period 2 (M=0.22, SE=.05) than during Period 1 (M=0.07, SE=.05), *F(1, 163.1)=5.84, p<.05*, and more than those in the No Personalization condition (M=0.03, SE=.04), *F(1, 34.7)=9.16, p<.01*; period x condition interaction, F=(1, 163.3)=2.61, p=.1). Here is one example:

Participant E: (starts laughing). I have a snack for you.
Snackbot: Please take your orange.
Participant E: I have a snack for you Snackbot. It's a battery.
Snackbot: Thanks, [participant name]. Enjoy your snack.
Participant E: Bye Snackbot.
Snackbot: I hope you have a wonderful day. Goodbye.
Participant E: You too, enjoy your snack.

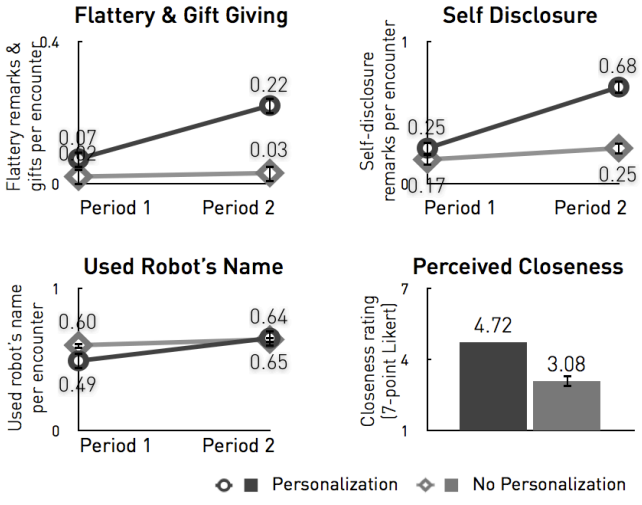

Figure 2. Measures of rapport

Self-disclosure. Participants also disclosed more about themselves in the Personalization condition during Period 2 (M=0.68, SE=0.10) than during period 1(M=0.26, SE=0.11), *F(1,*

162.4)=14, p=.001, and those in the No Personalization condition (M=0.25, SE=0.09), *F(1, 25.84) = 9.11, p<.01*; period x condition interaction, F(1, 159.5)=4.92, p=0.03).

Using the robot's name. Participants in the Personalization condition greeted the robot with the robot's name (i.e., "Hi, Snackbot") more frequently (M=0.65, SE=0.13) during Period 2 when the interaction was personalized than Period 1 (M=0.49, SE=0.13), *F(1, 143.7)=5.23, p<.05*.

Perceived closeness. Participants in the Personalization condition felt closer to the robot (M=4.72, SE=0.71) than those in the No Personalization condition (M=3.08, SE=0.52; *F(3,16)=3.05, p=.1*) but the difference was only marginally significant. Perceived self-connection did not differ by condition.

4.4 Cooperation

Personalization increased participants' cooperation, as predicted in Hypothesis 2. We derived a summary measure of cooperation for each participant by standardizing scores on all three measures (see Figure 3) and calculating a mean for each person. The results showed people's willingness to cooperate with the robot was greater in the Personalization condition (M=0.49, SE=.28) than in the No Personalization condition (M=-0.45, SE=.22), *F(2,18) = 3.48, p=0.02)*. We provide an example below.

Snackbot: I need to give a tour of [building] for visitors, I am still new to this building and I am not sure where to bring them. Could you suggest some interesting places in [building]?

Participant F (No Personalization condition): Snackbot, let's not be ridiculous, can I take my snack? Can I have my snack?

Participant L (Personalization condition): Let's see. You could visit the [exhibit name] on the first floor or the third floor. The second floor has a lot of cool other robotic stuff that you could check out or show people, I think they would like that […].

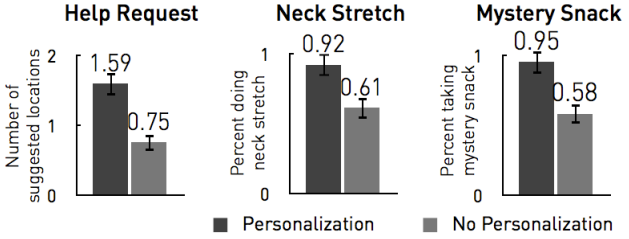

Figure 3. Measures of cooperation

4.5 Engagement

Participants' engagement with the robot appeared to be more sustained when the robot personalized its remarks (see Figure 4).

Laughing. Participants laughed more during personalized interactions during Period 2 (M=1.53, SE=0.36) than during Period 1 (M=0.99, SE=0.36), *F(1,146.1)=4.94, p<.05* and more than those in the No Participation condition (M=0.70, SE=0.32), *F(1, 27.91)=2.75, p=.10*; period x condition interaction: F(1, 145.3)=3.27, p=.07).

Standing posture. The percentage of the participants who sustained their upright standing posture did not change over time in the Personalization condition. In the No Personalization condition, the percentage of the participants who stood upright when interacting with the robot decreased from period 1 (M=0.66, SE=0.1) to period 2 (M=0.39, SE=0.1), *F(1,140.2)=11.25, p=.001*. More participants in the No Personalization condition leaned against their office doors while interacting with the robot in period 2, signaling higher status and/or less attention.

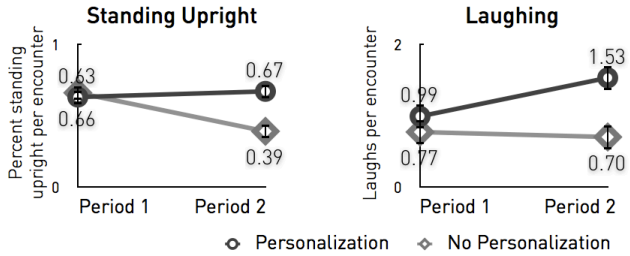

Figure 4. Measures of engagement

4.6 Service Satisfaction

The ratings of service satisfaction did not statistically differ by condition. Participants in both conditions were highly satisfied with the service (Personalization M=6.05 (SE=0.24); No Personalization M=6.22 (SE=0.21)), and were willing to continue the service (Personalization M=6.40 (SE=0.41); No Personalization M=6.53 (SE=0.35)). Participants in the Personalization condition said they would pay more to continue to use the service (M=$16.19, SE=4.09) than those in the No Personalization condition (M=$12.4, SE=3.48), but the difference was not statistically significant.

5. DISCUSSION

Our analyses suggest that personalizing the interactions with the robot reinforced participants' rapport, cooperation, and engagement with it. Our post-study interview results helped us understand how participants interpreted personalization strategy. As we noted above, in the interviews, we did not mention personalization so the answers we received were unsolicited.

5.1 Receiving Personal Attention

Consistent with the literature on personalization, and as argued, participants seemed to like getting personal attention from the robot. We designed Snackbot's personalization to build on real experiences between the robot and the person, creating an interaction that was unique to each participant. When the robot remembered even a small detail about a participant, for example, what snack was their favorite, it seemed to affect their perceived closeness to the robot. For example, Participant N said:

Surprisingly Snackbot knows that he never dies on me. (Interviewer: How did you feel about it?) So I feel good. I feel special that I communicate with Snackbot with no problem.

By contrast, in the No Personalization condition, most participants expressed desire to have tailored interactions with the robot as Participant U expressed:

But I felt like over time […] if he shows up every week, Monday, Wednesday and Friday, you would hopefully learn their name or that the conversation would get to the point where it could be a little bit more personal.

The rapport created through personalization may have played a role in influencing people's willingness to cooperate or help the robot. Participant I in the No Personalization condition said during the interview that the robot's tour help question was one of his negative experiences with the robot:

I think it was mostly that you don't have enough of a rapport with it to answer that question. So if it was like someone-- if it was like Justin or someone who works with me, I could be like "Oh we should show them the thing down in that lab where you work."

5.2 Sustaining Interest

According to the interviews, participants in the Personalization condition seemed to have been more engaged with the robot over the course of the study. We reasoned that this was because the robot's interactions became more meaningful over time. For those in the No Personalization condition, interaction with the robot became less meaningful as participants realized that their conversation with the robot did not have any bearing on the robot's behavior. This caused people to lose interest in conversing with the robot. By contrast, in the Personalization condition, the robot made comments based on its past performance or the participant's use of the service, therefore building common ground and meaning. In addition, the robot telling different stories related to the participant each time caused excitement and expectation, as participants waited for new stories.

Participant L (Personalization condition): We even commented to each other a couple times; What do you think he's going to say today or do you think he's going to say something about carnival?

5.3 Downsides of Personalization

As in human interaction, personal conversation can create discomfort because people are invested in the relationship. Some of Snackbot's personalized dialogues evoked negative responses, especially when participants felt uneasy about the behaviors that were the topics of conversation. The most sensitive topics surrounded participants' not being present when the robot arrived and their choice of snacks.

Participant M (Personalization condition): But then my most negative [feeling] was one time he said, "I notice that you always order Reese's Cups. You must really like Reese's Cups," and that was kind of awkward for me because it's like, "Oh, I'm the one ordering all the junk food, and eating junk food every day, and now he's pointing it out."

We were initially concerned that participants would have privacy concerns or feel more pressure to be social with the robot when the robot personalized its interactions. Participants mentioned that they did not have privacy concerns with the topics or events that the robot used to personalize. In both conditions, participants seemed to feel some pressure to be social and polite with the robot as the interactions took place in a social setting, the workplace, and others might hear these interactions.

6. IMPLICATIONS

In this section, we briefly address when to use personalization, how to use personalization, and the challenges and opportunities in designing successful personalization for repeated interactions.

When to use personalization. Human-robot interaction may benefit from personalized behaviors when it is important for the service to track and be aware of past service events. Customers know the business has a record of interactions and may expect a social robot to reflect these past interactions. For example, a snack delivery robot in a nursing home could be aware of what time meals were last served. Personalized behaviors may also be useful when the robot needs to be persuasive, for example, in choosing a healthy snack over an unhealthy snack, or when the robot needs help or input from customers [25]. Personalized behavior will be also useful in situations where the robot is assisting people doing boring and repetitive tasks since personalized behaviors over time could create surprise, joy and more engagement.

How to use personalization. We suggest personalization is best used to define a meaningful relationship between a robot and a person. As we learned in our study, the events that are selected to build common ground and make meaning must be chosen carefully. For example, comments about liking a particular kind of candy were embarrassing rather than meaningful. Like human interaction, not all facts bear repeating. Careful consideration must be given to what critical moments in an interaction are and how they can be detected. For example, an assistive robot in a care facility might call out moments of independence and ability to complete activities of daily living rather than breakdowns or calls for assistance.

Challenges and opportunities. Individuals differ in their receptivity to personalization. It will be important to develop mechanisms to detect people's responses to specific strategies and ways for robots to recover from wrong moves. Personalization in HRI also offers new opportunities in services. An interesting avenue for research will be to investigate personalization unique to robots; for examples, unlike humans, the robot has a perfect record of past interactions. In a setting where a human could not easily employ personalization techniques (e.g., a vendor in a big store), robots can personalize their interactions and change the dynamics of encounters. Another interesting avenue is self-aware robotic services. Compared to systems personalized to users' tracked behaviors, our attempt to use the robot's own tracked behaviors to personalize its interaction is relatively new. Our study suggests that it can be a promising area for the design of repeated interactions.

7. LIMITATIONS

As with all studies, the current study has limitations. Conducting a field experiment using a realistic service increased the ecological validity of our results, but also entailed many constraints of which three are particularly notable. First, we had to randomize conditions among the hallways to avoid contamination across conditions. Participants in the same hallway sometimes socialized during the Snackbot visit, and the existing culture of the hallway may have influenced the results reported in the paper. Second, we mixed different types of personalization, to maintain surprise and enjoyment. This means we cannot distinguish among the effects of specific personalization techniques. We do not know whether one topic was more powerful than others, or whether our strategy would be as effective if only one of the personalization topics were used. Third, the robot took one or two more speaking turns in the Personalization condition so it could have been more effective simply because it spoke more. Interview results suggest this was not the case, however.

Our study was also limited due to technical constraints. First, the study was conducted on one floor of a computer science building, where the robot could operate reliably, with access to engineering help if it broke down. However, half of the participants in the study did not have much programming knowledge. None of our participants were part of the Snackbot development. Studying an organization's prototype, especially when it is novel, in different teams or divisions within that organization is not uncommon (e.g., [14],[23],[29]). Also, our study used a Wizard of Oz technique for the selection of nodes in the dialog script, and the operator was in the vicinity of the robot for control and security reasons. When we asked participants if they believed the robot was autonomous, they wondered how much the robot was autonomous, but no one believed that they were communicating with the operator through the robot.

The specifics of our study also limit the external generalizability of the results. The snack service was operated as compensation for

participating in the trial for at least two months, having all interactions recorded, filling out surveys, and completing interviews. Free snacks may have contributed to high service satisfaction in both conditions. We recorded all the interactions with participants' consent. Recording may have influenced participants' behaviors, as well. We tested the strategy in the food delivery service domain. The robot was anthropomorphic, and the conversation was not natural as a human-human dialog. Generalizing the results to different service domains and robots will require further investigation.

8. CONCLUSION

Through a longitudinal study, we provide evidence that personalization with memory reinforces people's rapport, cooperation, and engagement with a robot. We also show changes in people's experiences with the robot over time. By presenting an example of a personalized robotic service, we offer insights on factors that other researchers can refer to when designing their systems. We hope this study inspires future research into how robots could be designed to engage people in a pleasurable and meaningful way over time.

9. ACKNOWLEDGMENTS

This work was supported by NSF grants 0624275 and 709077. We thank Alex McCluskey, Leonard Turnier, Katherine Cuti, Ahmad Shamsuddin, and Junior Baboolall for their assistance in the data collection and organization.

10. REFERENCES

[1] Aaker, J., Fournier, S., & Brasel, A. (2004). When good brands do bad. J*ournal of Consumer research, 31*, 1-16.

[2] Bickmore, T. & Cassell, J. (2001). Relational agents: A model and implementation of building user trust. *Proc. of CHI'11*, 396-403.

[3] Bickmore, T. & Picard, R. (2005). Establishing and maintaining long-term human-computer relationships. *Transactions on Computer Human Interaction, 59(1),* 21-30.

[4] Bickmore, T., Pfeifer, L., & Schulman, D. (2011) Relational agents improve engagement and learning in science museum visitors, *Intelligent Virtual Agents conference* (IVA11).

[5] Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.

[6] Carreira, F., Canas, T., Silva, A., & Cardeira, C. (2006). I-Merc: A mobile robot to deliver meals inside health services. *Proc. of Robotics, Automation and Mechatronics'06*, 1-8.

[7] Dautenhahn, K. (2004). Robots we like to live with: A developmental perspective on a personalized, life-long robot companion. *Proc. of ROMAN'04*, 17-22.

[8] Gockley, R., Bruce, A., Forlizzi, J., Michalowski, M., Mundell, A., Rosenthal, S., Sellner, B., Simmons, R., Snipes, K., Schultz, A.C., & Wang, J. (2005). Designing robots for long-term social interaction. *Proc. of IROS'05*, 2199-2204.

[9] Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. *Proc. of ROMAN'03*, 55-60.

[10] Gwinner, K.P., Gremier, D.D., & Bitner, M.J. (1998). Relational benefits in services industries: The customer's perspective. *Journal of Academy of Marketing Science*, *26(2),* 101-114.

[11] Hada, Y., Gakuhari, H., Takase, K., & Hemeldan, E.I. (2004). Delivery service robot using distributed acquisition, actuators and intelligence. *Proc. of IROS'04*, 2997-3002.

[12] Kidd, C. D., & Breazeal, C. (2008). Robots at home: Understanding long-term human-robot interaction. *Proc. of IROS'08*, 3230-3235.

[13] Knapp, M. L, & Hall, J. A. (2009). *Nonverbal communication in human interaction.* Cengage Learning.

[14] Kraut, R., Fish, R., Root, R., & Chalfonte, B.(1990). Informal communication in organizations. In S. Oskamp & S. Scacapan (Eds.), *Human Reactions to Technology.* Sage.

[15] Lee, M.K., & Forlizzi, J. (2009). Designing adaptive robotic services. *Proc. of IASDR'09*.

[16] Lee, M.K., Forlizzi, J., Rybski, P.E., Crabbe, F., Chung, W., Finkle, J., Glaser, E., & Kiesler, S. (2009). The Snackbot: Documenting the design of a robot for long-term Human-Robot Interaction. *Proc. of HRI'09*, 7-14.

[17] Lee, M. K., Kiesler, S., & Forlizzi, J. (2011). Mining behavioral economics to design persuasive technology for healthy choices. *Proc. of CHI'11*, 325-334.

[18] Lee, M.K., Kiesler, S., Forlizzi, J., Srinivasa, S., & Rybski, P. (2010). Gracefully mitigating breakdowns in robotic services. *Proc. of HRI'10*, 203-210.

[19] Mason, M., & Lopes, M. (2011). Robot self-initiative and personalization by learning through repeated interactions. *Proc. of HRI'11*, 433-440.

[20] Mehrabian, A. (1970). A semantic space for nonverbal behavior. *Journal of Consulting and Clinical Psychology, 35(2),* 248-257.

[21] Meuter, M. L., Ostrom, A. L., Roundtree, R. I., & Bitner, M. J. (2000). Self-service technologies: Understanding customer satisfaction with technology-based service encounters. *Journal of Marketing, 64(3),* 50-64.

[22] Nakauchi, Y., & Simmons, R. (2002). A social robot that stands in line. *Autonomous Robots*, *12,* 313-324.

[23] Paulos, E. & Canny, J. F. (1998). PRoP: Personal roving presence. *Proc. of CHI'98*, 296-303.

[24] Pfeifer, L. M. & Bickmore, T. (2011). Is the media equation a flash in the pan? The durability and longevity of social responses to computers. *Proc. of CHI'10*, 777-780.

[25] Rosenthal, S., Veloso, M., & Dey, A. K. Is someone in this office available to help me? Proactively seeking help from spatially-situated humans. To appear in *Journal of Intelligent and Robotic Systems*.

[26] Sabelli, A. M., Kanda, T., & Hagita, N. (2011). A conversational robot in an elderly care center: An ethnographic study. *Proc. of HRI'11*, 37-44.

[27] Sung, J. Y., & Grinter, R. E. (2009). Pimp my Roomba: Designing for personalization. *Proc. of CHI'09*, 193-196.

[28] Surprenant, C.F., & Solomon, M.R. (1987). Predictability and personalization in the service encounter. *Journal of Marketing, 51(2)*, 86-96.

[29] Venolia, G., Tang, J., Cervantes, R., Bly, S., Robertson, G. G., Lee, B. & Inkpen, K. (2010). Embodied Social Proxy. *Proc. of CHI'10,* 1049-1058.