"How was your stay?": Exploring the Use of Robots for Gathering Customer Feedback in the Hospitality Industry

Michael Jae-Yoon Chung and Maya Cakmak

***Abstract*— This paper presents four exploratory studies of the potential use of robots for gathering customer feedback in the hospitality industry. To account for the viewpoints of both hotels and guests, we administered need finding interviews at five hotels and an online survey concerning hotel guest experiences with 60 participants. We then conducted the two deployment studies based on deploying software prototypes for Savioke Relay robots we designed to collect customer feedback: (i) a hotel deployment study (three hotels over three months) to explore the feasibility of robot use for gathering customer feedback as well as issues such deployment might pose and (ii) a hotel kitchen deployment study (at Savioke headquarters over three weeks) to explore the role of different robot behaviors (mobility and social attributes) in gathering feedback and understand the customers' thought process in the context that they experience a service. We found that hotels want to collect customer feedback in real-time to disseminate positive feedback immediately and to respond to unhappy customers while they are still on-site. Guests want to inform the hotel staff about their experiences without compromising their convenience and privacy. We also found that the robot users, e.g. hotel staff, use their domain knowledge to increase the response rate to customer feedback surveys at the hotels. Finally, environmental factors, such as robot's location in the building influenced customer response rates more than altering the behaviors of the robot collecting the feedback.**

I. Introduction

Gathering customer feedback is a critical component of the hospitality industry. Hotels have long amassed guest feedback to measure customer satisfaction and loyalty as well as staff performance. The collected feedback helps hotels to monitor service quality, make necessary improvements, and ultimately stay ahead of the competition [1].

We believe service robots can be an effective medium to elicit and gather guest feedback in hotels since they can serve as neutral liaisons between hotels and guests. The robots also draw people's attention in public spaces [2]–[6], and their interactive behaviors can be precisely controlled to enforce hotel brand standards or elicit certain emotional responses [7].

Finding a novel real-world use case for service robots is not trivial. Deploying robots in workplaces may require structural and procedural changes in workplace design and use [8], consequences that are difficult to foresee and costly to discover post-deployments. In addition, it is difficult to gather quality feedback from potential users in up-front research sessions because most people lack firsthand experiences with robots in their workplace [9].

The authors are with Computer Science & Engineering, University of Washington, Seattle, WA 98195-2350 {mjyc, mcakmak}@cs.washington.edu

Fig. 1. (Left) The Relay robot running the breakfast survey behavior at the P1's hotel. (Right) The robot asking for employee feedback about the provided meal at the Savioke headquarters. The robot briefly stopped navigating to respond to the employee.

With these challenges in mind, we present four studies exploring the customer feedback collection use case for robots in the hospitality industry that feature the Savioke Relay, a commercial service robot. We discuss the overall approach in Sec. III and share our experiences and learned lessons as findings and design implications from each study in the subsequent sections (Sec. IV, Sec. V, Sec. VI, Sec. VII).

II. Related Work

Our work is part of the research being done in the field of human-robot interaction (HRI), which focuses on designing and evaluating real-world robot applications.

Over the last decade, HRI researchers have employed product design methods such as *need finding* and *participatory design* to understand users' needs and related constraints and requirements for domains of interests, such as home organization [10] and guiding visitors in tourist sites [11], airports [12], and office buildings [13]. Findings from these studies have been abstracted into themes [12], [13], frameworks [10], and visual maps of user experiences [11] and used to derive design recommendations. While such research work presents insightful recommendations, we believe it is also critical to evaluate a potential application early in the deployment process and in real-world settings (for example, via prototyping) to learn critical design and implementation lessons soon as possible.

A body of HRI work evaluates robotic systems in the wild [2]–[4], [14]–[17]. Researchers in this domain ask questions concerning feasibility and performance of the proposed

systems, e.g., "Can the robotic system successfully perform the given task?" and "How well does it work?" This research attempts to understand why the proposed systems work (or do not work) and to identify factors influencing robot usage. For example, researchers verified the feasibility of deploying fully autonomous robots as guides in public places [2], [14], [15], a walking group assistants in a care site [17], and as scouts in an office building [16]. They also investigated potential benefits, such as enhancing the motivation of older adults [17]. Compared to these studies, our work puts more emphasis on understanding the requirements and opinions of multiple stakeholders of our proposed use case before developing it.

Our work is also related to work that applies a design approach when developing a service robot application. Three prominent examples are Snackbot [18], a delivery robot in an office building; TOOMAS [19], a shopping guide robot in a home improvement store; and Sacarino [20], a bellboy robot in a hotel. The research teams for these projects employed design processes that involved multiple phases of prototyping with potential users in target environments to get continuous feedback during development. Recently, researchers proposed adapting a *lean UX* design approach from the startup field for identifying a commercial social robot application and applied the proposed method to rapidly prototyping an assistant robot in Sydney International Airport [21]. Like these work, our research also emphasizes the importance of the users and the context of a robot application. However, we focus on a much less explored domain: a *guest feedback collection* use case involving a *non-anthropomorphic* robot in *hotel* environments.

III. Approach

This paper explores the use of robots for gathering customer feedback in hotels. We address the following research questions:

- Can we use robots to gather feedback from hotel guests?
- How should we design robotic systems to gather better customer feedback?

The answer to the first question depends on the context in which robots interact with customers. Hence, it is important to understand the physical and situational context as well as customer opinions. It is also important to understand the needs of service industry workers and their current practices for gathering customer feedback. Finally, it is critical to tackle these questions through field deployments to capture the context in which the customer experiences the service and decides whether and how to respond to a robot's solicitation for feedback. To that end, our research focused on the Savioke Relay robot (Sec. III-A), which were deployed in approximately 70 hotels in January 2018.[1]

We conducted the following four studies:

1) *Need finding interviews* with hotel management (n=5) at five hotels that already used a Relay robot for guest room delivery. We explored the current practices hotels use to gather customer feedback and the contexts in which the robot could gather feedback (Sec. IV).
2) *An online survey* with varying-frequency hotel customers (n=60). We explored guests' perceptions of the Relay room delivery robot and their willingness to respond to the robot in different situations cs responding to other feedback methods (Sec. V).
3) *Passive observations and follow-up interviews* (n=5) at three of the five hotels in 1). We deployed the prototypes of robot-based customer feedback applications for 3-4 months explored the value that the robot added, interaction patterns with the robot, and real-world challenges in gathering feedback from customers with robots (Sec. VI).
4) *Passive observations, follow-up questionnaires, and analysis of measurements* in a three-week deployment at a kitchen area of the Savioke headquarters. We explored the role of different robot behaviors (mobility and social attributes) in gathering feedback and identified design constraints (Sec. VII).

A. The Relay Robot

Savioke Relay[2] is an autonomous mobile robot that delivers small items from the front desk to guest rooms in hotels. The robot is approximately three feet tall and weighs 100 pounds, has a lockable interior bin, and displays a touchscreen mounted facing forward.

The default behavior of the robot is as follows. The robot stays in its docking station and charges its battery when not in use. Upon receiving a delivery request, the front desk clerk loads the robot's bin and sends it to the specified location. The robot travels to the guest's room by traversing hallways, doorways, and elevators while avoiding obstacles and people. When it reaches the destination room, it phones the room to notify the guest of its presence, and it opens its lid when the guest opens the room door. Once the guest retrieves the item, it interacts with the guest to confirm the pickup and then returns to its docking station. Hotel staff interacts with Relay through a web interface to send a delivery, monitor the robot's status and location, and download a delivery history.

Although the Relay robot was designed for a specific application, it can be considered a general-purpose mobile robot that uses parts (e.g., a mobile base and touchscreen) common to other autonomous indoor robots, such as those from Vecna, Cobalt Robotics, and Fellow Robots.[3]

IV. Need Finding Interviews at Hotels

In 2016, we conducted need finding interviews with five employees who held administrative staff jobs at five hotels that had been using a Relay robot [22]. The interviews took a place in the participants' hotels and were structured to learn: (i) current practices for collecting guest feedback, (ii) opinions on the idea of gathering feedback with the robot, and (iii) guest experiences with using the robot at their hotels.

[1] spectrum.ieee.org/view-from-the-valley/robotics/industrial-robots/ces-2018-delivery-robots-are-fulltime-employees-at-a-las-vegas-hotel

[2] www.savioke.com

[3] www.vecna.com, www.cobaltrobotics.com, www.fellowrobots.com

Fig. 2. (Left, Middle) Pictures of the kiosk and the Relay robot used in the online survey to illustrate Kiosk and Robot FD. (Right) A screenshot from the video that demonstrates room delivery to convey Robot RM.

We learned that the hotels actively collect guest feedback and heavily rely on traditional means to collect data, such as post-stay email surveys or collecting data from online forums such as TripAdvisor.[4] The collected data were most frequently used to prevent or recover dissatisfied customers. We also found that the day of the week, the season and guest demographics most affected the robot usage the most. Hotels were interested in using the Relay robot to solicit guest feedback in a lobby or to ask customer satisfaction questions after a room service delivery. After collecting feedback, hotels suggested that the Relay robot could alert the staff members about issues or resolve problems as they occur. We summarize our findings and design recommendations from [22] below.

Based on our findings, we propose the following design recommendations. Overall, robots should play a role in helping hotels widely disseminate positive customer feedback. Robots should take advantage of being in the context of the service to encourage customers to easily express their opinions in the moment and on the property. They should be designed to make guests feel more comfortable giving negative feedback. Robots should always respond to customer feedback. By default, interactions for gathering feedback should be short and, if possible, entertaining to accommodate the short attention spans of modern customers. Robots could leverage their status as *novelty items* to engage customers and encourage them to respond to questions. Ultimately, robots should adjust their strategies for eliciting engagements and interacting with customers based on the types of customers.

V. Online Guest Scenario Survey

Viewpoints of guests are important to hotels as well as to the robotics company providing the service. Hence, we conducted an online survey that explored potential guests' attitudes and motivations towards robot-based feedback solicitations compared to other solicitations.

A. Survey Design

To help participants contextualize the decision about responding to different kinds of customer feedback solicitations, we provided a motivating scenario. We chose one negative and one positive guest experience scenario adapted from the scenarios[5] commonly used in hospitality research [23]. Each participant read one of the two scenarios. We instructed participants to assume the situation described in the scenario had just happened to them. We asked them how likely they would be to respond to each of the following feedback solicitations (1: Extremely unlikely to respond; 5: Extremely likely to respond):

- **Email:** You received an online survey after you left the hotel.
- **Kiosk:** You noticed a kiosk near the front desk which says "How is your stay?"
- **Robot Front Desk (FD):** You noticed a robot near the front desk which says "How is your stay?"
- **Robot Room (RM):** You ordered a snack from the front desk and a robot delivered the snack to your room. After handing off the snack, the robot asks *"How is your stay?"*

We included an optional open-ended question asking for an explanation of each response. For Kiosk, Robot FD, and Robot RM, we included explanatory images and a video to give participants an accurate sense of what the kiosk and robot would look like (Fig. 2). Note that the four solicitation methods were selected based on Savioke's and their customer hotels' business interests.

B. Participants

Participants were recruited through Amazon Mechanical Turk. After agreeing to participate, online participants were directed to the single-page form that contained the scenario and the four questions about the four feedback alternatives. To control for quality of responses, we did not allow a person to participate in our survey more than once and rejected people who incorrectly answered the question used to identify those who randomly selected answers. We offered $0.01 for participation and continued recruiting until we had 30 responses for each scenario. A total of 60 people (22 M, 38 F) responded in less than 2 weeks. Their age groups distributions were: 10% in 19-24, 30% in 25-34, 30% in 35-44, 30% in 45-54, 16.67% in 45-54, 10% in 55-64, 3.33% in 65-74. Respondents' answers to the question regarding the frequency of staying at a hotel in any given year ranged from 1 to 20, with a median of 2 times a year.

C. Findings

The means and standard deviations of the responses across the four solicitations were: M = 3.58 & SD = 1.43 (Email); M = 3.38 & SD = 1.58 (Robot RM); M = 3.25 & SD = 1.67 (Kiosk); and, M = 2.68 & SD = 1.69 (Robot FD). The same statistics across the two scenarios were: M = 3.22 & SD = 1.52 (Positive); and, M = 3.23 & SD = 1.72 (Negative). Fig. 3 shows the distribution of responses for each feedback solicitation method in each scenario. We conducted open coding for the open-ended question responses. We summarize our findings below.

1) Factors Influencing Guests' Willingness to Respond: The most frequently mentioned reason for responding to a solicitation method was for participants to share hotel experiences (e.g., *"I would want the hotel to know that my room*

[4]https://www.tripadvisor.com/

[5]For more details, see Appendix A. in the Supplementary Materials available at https://goo.gl/xz1xk8.

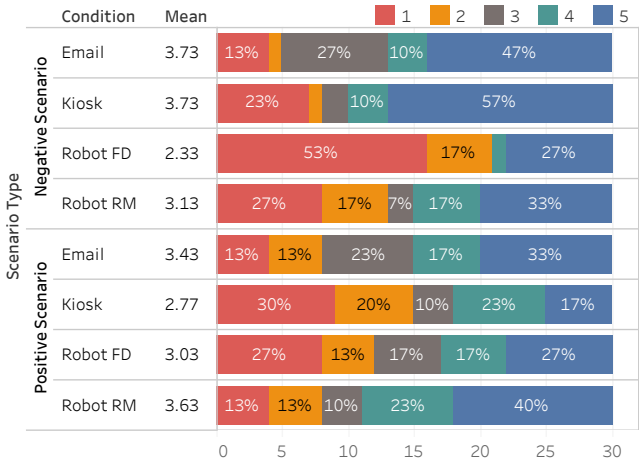

Fig. 3. The distribution of responses for different feedback solicitations in positive and negative guest experience scenarios (1: Extremely unlikely to respond; 5: Extremely likely to respond).

was unacceptable."). The second most mentioned reason was convenience. This was also the main reason why Email and Robot RM were rated highly. For example, a participant mentioned, *"If I'm already interacting with the robot, I may as well answer the question."* In fact, inconvenience was the top reason for participants' unwillingness to respond to a solicitation method.

Some participants had privacy concerns about using the Kiosk and Robot FD, e.g., *"Less likely, but I'd feel the staff might be watching over my shoulder."* They mentioned that they liked the Email and Robot RM methods because of increased privacy, e.g., *"I would respond because it is still kind of private and (the robot is) not out in the busy lobby."* Given our findings from the need finding interviews (Sec. IV), we noticed an understandable conflict between hotels and guests: hotels want more data from guests, but guests value privacy want it to be respected.

2) Influence of the Scenario Type: The influence of the scenario type varied across the four solicitations. When the participants read the *positive* scenario and were presented with Email or Kiosk, they were not enthusiastic about informing the hotels (Email: $p < 0.001$, Kiosk: $p < 0.00001$; one-tailed, paired t-test); more than half said they would be unlikely to respond (i.e., responded with < 3) because they would not have the time or did not want to further engage with the hotel. In contrast, more than half of the participants who read the *negative* scenario and were presented with Email, Kiosk, or Robot RM would inform the hotel about their stay (responded with > 3).

The participants were more enthusiastic about responding to the two methods involving robots when they read the positive scenario than when they read the negative one (both Robot FD and Robot: RM $p < 0.0001$; one-tailed, paired t-test). Some participants would be likely to respond to the robot (i.e., responded with > 3) predicted that interacting with it would be pleasant and would make their experience at the hotel unique. In contrast, the participants who read the negative scenario most commonly questioned the possibility of the presented scenario, for example, by commenting *If they had enough money for a robot, they would have enough money for a cleaning staff to do a good job.*

3) Perception of the Robot-based Solicitations: Regardless of the scenario type, participants mentioned similar reasons for liking or disliking the two robot-based solicitations. When participants were willing to respond, the reasons they offered included the novelty effect (e.g., "*I will fill out of curiosity.*"), the entertainment value of the robot (e.g., "*The robot would interest me and be a fun addition to my stay.*"), and the feeling of obligation (e.g., *"Because the robot needs a response."*). When they were not willing to respond, they mentioned that they disliked the robot (e.g., *"A feedback robot? That's just weird; the world isn't ready for that."*), did not trust the robot (e.g., *"I wouldn't trust the robot."*), or envisioned a potential difficulty using the robot (e.g., *"No clue how to interact with it."*).

D. Design Implications

The process of providing feedback to the robot should be as convenient as possible for the guest, for example, by keeping survey questions short or asking a question at the end of a different interaction to allow the guest to ignore the question. Robots should respect the privacy of guests even at the cost of losing data they can provide to the hotels since the goal of both hotels and robotics companies is to satisfy customers. For example, robots should collect guest feedback in private settings and restrict using sounds or movements that could reveal guest responses in open spaces.

VI. Hotel Deployments

Three interviewees–P1, P2, and P5[6]–from the need finding interviews (Sec. IV) were interested in using a robot-based feedback solicitation at their (different) hotels. This gave us the opportunity to test our idea in the field and thereby better understand real-world challenges.

A. Breakfast Room Survey

P1 requested that we enable the Relay robot to ask a few customer satisfaction survey questions to the hotel guests in the breakfast room area. We collaborated with a robot UX designer at Savioke to prototype a breakfast room survey robot's behavior and refined it with P1 to meet hotel's requirements. The prototype behaved as follows. Upon launch, the robot navigated to a predefined location near the entrance of the breakfast room. The robot then stayed in place and displayed the question, "Good morning! How was your breakfast?" together with a five-star rating response field. When a customer response was detected, it asked the second question, "How is your stay?" and responded on receiving five stars with a happy facial expression and a dance. If no response was received within five minutes, the robot played whistle sounds and displayed the two messages, "Hello! I'm Relay, a delivery robot." and "Need anything? Dial 0 from

[6]See [22] for descriptions of P1, P2, and P5.

your room and I'll bring it to you!" in sequence to elicit attention from passing guests.

In May 2017, we installed the break room survey behavior to the Relay robot in P1's hotel (as shown in Fig. 1 (Left)) and instructed all staff members there about how to start and stop the survey behavior and handle the potential problems. While the robot's status information, such as its location and remaining battery charge, was available on the web interface, the hotel staff was not able to see the customer response history (due to logistical reasons). However, P1 insisted on deploying the breakfast room survey behavior as is to increase utilization of the robot. To gain insight into how the hotel used the provided survey behavior, we monitored their usage for four months both remotely and through two on-site visits, on Tuesday and Saturday of the 3rd week of September 2017, to observe the robot in context. We were not allowed to talk to guests for logistical reasons.

B. Low-Ratings Alert for Guest Room Delivery

P2 and P5 requested that we enable the Relay robot to ask customer satisfaction questions after each guest room delivery and alert staff members on receiving negative responses from guests. We collaborated with a Savioke robot UX designer and prototyped a feature that adds the *"How is your stay?"* star rating question after a delivery confirmation interaction and sends email alerts on receiving ratings below three stars.

We provided the low-rating alert feature to both hotels in June 2017. We configured the feature to send emails to the staff mailing list and logged the usage of the feature for four months. On the first week of September 2017, we interviewed P2 and P5 and one staff person at each hotel.

C. Findings

Despite not providing actual feedback from customers, the staff members at P1's hotel used the breakfast room survey extensively. Over the four months, they used the survey every day except for three days post-deployment (123 days). On average, the survey ran for 229 minutes, and 43 questions were answered per day. On our two visit days, we observed approximately 9 guest-robot interactions on the first day and 22 on the second day. The robots received responses to the customer satisfaction question 1707 times at the P2's hotel and 709 times at the P5's hotel during the four months deployment. Of those, 46 and 17 (2.70% and 2.40%), respectively, received less than three stars.

1) Values of Robot-Based Guest Feedback Collection: As predicted in the need finding interviews (Sec. IV), P1 reported that the robot's ability to provide unique experiences to the guests was its most valuable aspect. P1 nonetheless acknowledged the potential benefit of the data collected by the robot: *"Getting a report that shows me the overall scores, that would be great. That way, at least I could track what days people are not happy with."* Regarding the low ratings alert feature, both P5 and P2 were satisfied with the feature and mentioned they were able to capture 2-3 unhappy customers per month.

2) Hotels Used Their Domain Knowledge: We learned that the time and location of running the breakfast survey behavior were carefully selected by the hotel staff. The robot was located in front of the breakfast room area, which was located right next to the elevator. Hence, the robot was seen by people going in and out of the breakfast room as well as guests waiting for the elevator; most people noticed the robot immediately or via the whistle sounds it played. We also noticed the network effect: whenever a guest started interacting with the robot, it raised the attention of the other passing or waiting guests. When we continued our observation at P1's hotel in the afternoons, we observed almost no activities near the breakfast room.

At P2's hotel, a staff person mentioned they paid extra attention over the weekend because the hotel usually receives a higher number of complaints while the staff person from P5 reported that they paid extra attention to input from guests with a membership.

3) Privacy Issues: Due to logistical reasons, our two prototypes did not rigorously follow one of our own guidelines: the robot should respect customer privacy. For example, the robot responded with a whistle sound and a dance on receiving five stars while surveying in the breakfast area, which allowed the people around the robot to notice what rating the person had given. In addition, the low ratings alert feature did not give customers the choice of notifying the front desk when they responded with less than three stars on the "How is your stay?" question. We learned that the hotels took advantage or were unaware of the consequences. For example, P1 mentioned that they like to monitor the guest interacting with the robot to not only to identify unhappy customers but also to understand the status of the hotel in general by eavesdropping on nearby conversations around the robot. Regarding the low ratings alert, all four interviewees mentioned that they always followed up with guests to recover potentially unhappy guests. No one considered the guests who do want to be contacted by the hotels.

D. Design Implications

The user interface for the robot should be designed to protect guest privacy. For example, on receiving complaints, the robot should ask whether the guest is comfortable with its informing hotel staff about the complaints. To maintain the robot's position as a neutral liaison between hotels and guests; the interface should not reveal the guest's feedback if they do not want to inform the hotel. The interface for the robot should support using the hotel users' domain knowledge, e.g., providing an option to customize the messages used during survey, schedule survey behavior, or change survey location.

VII. Kitchen Deployment

As a final step in our exploration, we wanted to better understand: (i) the impact of the Relay robot's behavior on eliciting customer feedback, and (ii) the opinions of customers who are experiencing a service. In the hotel

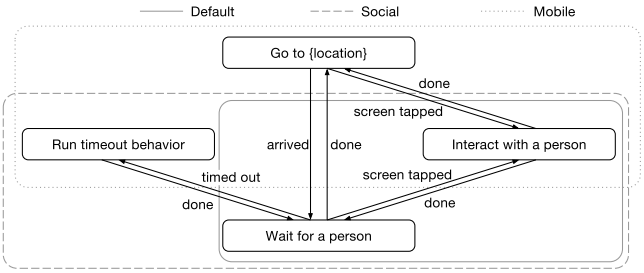

Fig. 4. Finite state machines that implement the three robot behaviors: Baseline, Social, and Mobile.

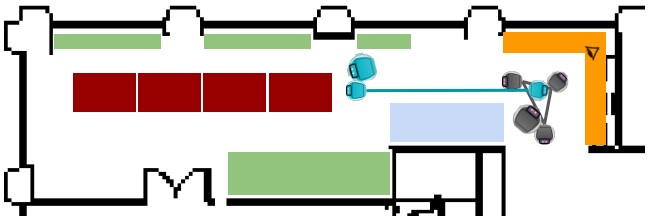

Fig. 5. The layout of the kitchen where the Relay collected feedback about meals. Large icons show the poses of the robot in Baseline and Social behaviors (gray: breakfasts, blue: lunch). Small connected icons show the waypoints in Mobile behavior.

deployment study (Sec. VI), testing out different robot behaviors at hotels or talking to guests was not an option for logistical reasons. Therefore, we deployed a survey robot at the kitchen of Savioke's headquarters.

A. Study Design

We believe the Relay robot has two key properties that offer competitive advantages when gathering feedback relative to other feedback modes: *mobility* to move towards potential respondents, and *social agency* to increase the engagement using social cues [6]. By varying these properties, we designed the following three behaviors:

- *Baseline* was designed to mirror the experience of using a kiosk, like the Happy-or-Not Smiley Terminal.[7] The robot stayed at a predefined location and asked a question about the meal using a five-star rating selection menu. We removed Relay's face, i.e., the eyes and speech bubble shown in Fig. 2, with a white background and did not use any sounds or movements. If a person answered the question with a rating, the robot responded by displaying "Thank you" for 5 seconds.
- *Social* was designed to make people perceive the Relay robot as a social agent. While its general behavior was similar to that of Baseline, the Social robot used more animated messages, including sounds, in-place movements, and LED light patterns. We kept Relay's face; all messages were displayed in the robot's speech bubble by using facial expressions. When nobody interacted with the robot for more than five minutes, the robot encouraged people to rate their meal with sounds and texts (i.e., 'Run timeout behavior' in Fig. 4).
- *Mobile* was designed to increase the chance of people noticing the robot by setting the robot in constant motion. The robot moved between two or three locations, encouraging people to leave their feedback. Once a person tapped the screen, the robot stopped and interacted with the person as in Baseline.

All experiments took place in Savioke's kitchen area, where the company provides meals to its employees at least three times a week. The kitchen area, approximately 700 square feet, was consisted of three sub-areas: a countertop area, a fridge area, and a table area (Fig. 5). Savioke had approximately 45 employees at its headquarters, who were mostly engineering and sales personnel.

To capture the situational variability, we conducted experiments during three company-provided meals: *Monday breakfast*, *Monday lunch*, and *Wednesday breakfast*. The two breakfasts were served on the countertop in the kitchen and delivered by an office administrator between 8:30 am and 9:30 am. The Monday lunch was served on the largest bar table in the kitchen and delivered by a catering company between 11:30 am and 12:00 pm.

To understand their employee-perspectives after they rated a meal, we asked them to express their agreement (1: Strongly agree; 5: Strongly disagree) on a questionnaire with the following statements:

- **Effort:** Relay has made it easy to report my opinion.
- **Ease of use:** Relay was easy to use.
- **Perception of the robot:** Relay was pleasant to interact with and be around.
- **Change in perception of the service:** I felt better about my meal experience after reporting my opinion.

We included an optional open-ended question asking employees to explain their answers. The questions were adapted from the questionnaires used in related work [24], [25].

B. Procedure

The experiments were administered for three consecutive weeks, beginning the second week of August 2017. The order of conditions was counterbalanced using a Latin square design, crossing the three robot behaviors with the three meal types. We worked with the office administrator to decide on the menu, selected to introduce variance as possible across the sessions.

Prior to the first experiment, we installed a Nest Cam Indoor camera and informed everyone in the company about the purpose and the duration of the study. We also shared our video recording policy and encouraged employees to share their privacy concerns. Each experiment session began when the food was served on the countertop or the bar table and ended when the food was taken away from the kitchen by the office administrator. The researcher started and stopped the robot behavior and video recording whenever a session was begun and ended. After each session, the researcher analyzed the recorded video and collected rating data to

[7]https://www.happy-or-not.com

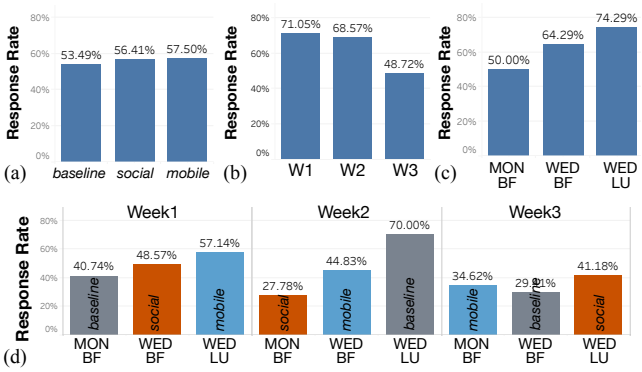

Fig. 6. Response rate results in our three week deployment at Savioke headquarters of three different robot behaviors, accumulated by: (a) robot behaviors, (b) weeks, and (c) meal types; and response rates (d) broken down by session.

identify people who gave a star rating to the robot and send them the post-session questionnaire. We also encouraged all employees, whether or not they interacted with the robot, to directly email their feedback directly to the researcher.

C. Findings

During the three-week deployment, a total of 57 employees visited the kitchen area, and a total of 38 interacted with the robot. We determined the survey response rate by analyzing the recorded video and the data collected by the robot. We calculated response rate as $|M \cap R| \,/\, |M|$, where M is a set of employees who had a meal provided by the company and R is a set of employees who used the robot to report their feedback. We also reviewed answers to the open-ended questions, email feedback sent to the researcher, and video recordings of the sessions.

1) Response Rate: We compared the response rate across the three robot behaviors, the meal types, and the week numbers (Fig. 6a–c). We saw no significant impact of robot behaviors on the response rate; however, we saw influences of the meal types and the week numbers. The response rate dropped over the three weeks, potentially due to the novelty effect (Fig. 6b). The influence of meal type on response rate was the strongest (Fig. 6c), especially on the first and second weeks (Fig. 6d); we suspect that Savioke's organizational structure, e.g. holding of company-wide meetings on Monday morning, was the root cause.

2) Comparison of Robot Behaviors: Although the differences in response rates across the robot behaviors were not significant, we observed differences in how people noticed the robot, which can be considered the first step to interacting with it. People were more likely to notice the robot when it moved; 92% of employees in the kitchen noticed the robot running the Mobile behavior, but less than 58% noticed the robot running the Baseline and Social behaviors. The increased movement potentially contributed to the highest response rate (Fig. 6a) but decreased the usability for some employees; we observed five people who failed to tap the robot to stop its motion on their first try. We saw no significant differences between the Baseline and Social behaviors among our quantitative measurements.

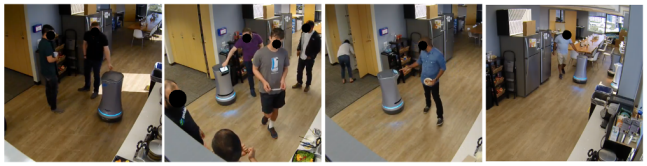

Fig. 7. People interacting with the Relay robot running the Mobile behavior.

3) User Perspectives: We received a total of 87 responses to the post-session questionnaire from 37 unique participants. The means and standard deviations for the post-session questionnaire were: M = 4.36 & SD = 0.94 (Effort); M = 4.67 & SD = 0.66 (Ease of use); M = 4.26 & SD = 0.96 (Perception of the robot); M = 3.46 & SD = 1.17 (Change in perception of the service). In answers to the open-ended questions regarding the effort, the most common positive remarks concerned the robot's ability to gather feedback in the moment, e.g., *"He was right there when the experience was still fresh in my mind."* The most common negative remarks were about the robot's inability to collect richer feedback, e.g., *"I wouldn't have been able to give concrete feedback other than to say it's good or bad."*

Regarding the ease of use, only three respondents who experienced the non-responsive touchscreen commented that the robot was difficult to use. Although we saw no negative comments regarding the perception of the robot from the open-ended questionnaire, one employee's email described discomfort concerning the mobile behavior: *"The robot in the kitchen is getting agitated (beeping more and more frequently as the morning goes on) and pushy (He started following me out of the kitchen and whether it was intentional or not, seemed to be trying to get me to take the survey.)"*.

In the answers to the "Change in perception" open-ended question, some people did not find that the robot affected their perception of the provided service, e.g., *"The food is what matters"*. They noted the lack of explanation on how their feedback would be handled, e.g., *"No confirmation on where the feedback is going, how it would be actionable."* They also appreciated the opportunity to give feedback, e.g., *"It seems good to know that I can provide input, whether it's good or not, to make things better next time."*

D. Design Implications

While using navigation movements could be useful, robots should avoid moving excessively so as not to disturb customers in the vicinity. Robots should have an option for guests to provide richer feedback. They should also clearly explain how the feedback will be handled or ask if/how the customers want followed-up contact.

VIII. Limitations

We acknowledge the following limitations in our research. First, we interviewed only five hotel employees in the need finding study, which is a small number of data points. In

the online survey study, we did not consider a wide range of solicitation methods, such as those based on mobile phones mentioned by the hotel staff [22]. Finally, the kitchen deployment study was conducted at Savioke headquarters, the company that developed the Relay robot. Our findings may be biased due to the participants' familiarity with the robot and our design implications may not apply directly to the customer feedback collection use case in hotels due to the differences between office and hotel environments.

IX. Conclusion

We presented four studies that explored the use case of collecting guest feedback with service robots in the hospitality industry. Based on our findings, we summarize our design recommendations as follows. Robots should collect customer feedback in the context of the service and keep their interactions with guests brief to make feedback collection convenient for guests and useful for hotels. To protect the privacy of guests, they should be able to opt out of sharing their feedback with hotel staff. They should also be able to provide more feedback if they so desire. The user interface for robots should support configuring the robot behaviors for the robot users to use their knowledge about their organization to manage engagement levels of the robot.

We believe that robot-based customer feedback collection could be a viable commercial application that takes advantage of both functional and social aspects of service robots. While we do not claim our findings and design recommendations generalize to all hotels, we hope this investigation will stimulate follow-up explorations of the customer data collection use case in the hospitality industry and beyond.

ACKNOWLEDGMENT

This work was supported by Savioke, Inc. and partially supported by the National Science Foundation, Awards IIS-1552427 "CAREER: End-User Programming of General-Purpose Robots." We thank Steve Cousins, Tessa Lau, Leila Takayama, Adrian Canoso, Will Vaughan, Natalie Friedman, and Jean Fong for inspirational discussions during the early phase of the project.

References

[1] J. Weed, "Checking in after checkout," *The New York Times*, May 2013. [Online]. Available: https://www.nytimes.com/2013/05/28/business/hotels-work-harder-to-collect-customer-responses.html

[2] W. Burgard, A. B. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun, "Experiences with an interactive museum tour-guide robot," *Artificial intelligence*, vol. 114, no. 1-2, pp. 3–55, 1999.

[3] R. Gockley, A. Bruce, J. Forlizzi, M. Michalowski, A. Mundell, S. Rosenthal, B. Sellner, R. Simmons, K. Snipes, A. C. Schultz, *et al.*, "Designing robots for long-term social interaction," in *Proc. of IROS*, 2005, pp. 1338–1343.

[4] T. Kanda, D. F. Glas, M. Shiomi, H. Ishiguro, and N. Hagita, "Who will be the customer?: A social robot that anticipates people's behavior from their trajectories," in *Proc. of UBICOMP*, 2008, pp. 380–389.

[5] C. Shi, M. Shiomi, C. Smith, T. Kanda, and H. Ishiguro, "A model of distributional handing interaction for a mobile robot." in *Proc. of RSS*, 2013, pp. 24–28.

[6] L. Moshkina, S. Trickett, and J. G. Trafton, "Social engagement in public places: a tale of one robot," in *Proc. of HRI*, 2014, pp. 382–389.

[7] A. M. Rosenthal-von der Pütten, N. C. Krämer, L. Hoffmann, S. Sobieraj, and S. C. Eimler, "An experimental study on emotional reactions towards a robot," *International Journal of Social Robotics*, vol. 5, no. 1, pp. 17–34, 2013.

[8] B. Mutlu and J. Forlizzi, "Robots in organizations: the role of workflow, social, and environmental factors in human-robot interaction," in *Proc. of HRI*, 2008, pp. 287–294.

[9] C. Pantofaru and L. Takayama, "Need finding: A tool for directing robotics research and development," in *RSS Workshop on Perspectives and Contributions to robotics from the human sciences*, 2011.

[10] C. Pantofaru, L. Takayama, T. Foote, and B. Soto, "Exploring the role of robots in home organization," in *Proc. of HRI*, 2012, pp. 327–334.

[11] D. E. Karreman, E. M. van Dijk, and V. Evers, "Using the visitor experiences for mapping the possibilities of implementing a robotic guide in outdoor sites," in *Proc. of RO-MAN*, 2012, pp. 1059–1065.

[12] M. Joosse, M. Lohse, and V. Evers, "How a guide robot should behave at an airport insights based on observing passengers," *CTIT Technical Report Series*, no. TR-CTIT-15-01, 2015.

[13] S. Azenkot, C. Feng, and M. Cakmak, "Enabling building service robots to guide blind people a participatory design approach," in *Proc. of HRI*, 2016, pp. 3–10.

[14] D. Bohus, C. W. Saw, and E. Horvitz, "Directions robot: in-the-wild experiences and lessons learned," in *Proc. of AAMAS*, 2014, pp. 637–644.

[15] D. Karreman, G. Ludden, and V. Evers, "Visiting cultural heritage with a tour guide robot: a user evaluation study in-the-wild," in *Proc. of ICSR*, 2015, pp. 317–326.

[16] M. J.-Y. Chung, A. Pronobis, M. Cakmak, D. Fox, and R. P. Rao, "Autonomous question answering with mobile robots in human-populated environments," in *Proc. of IROS*, 2016, pp. 823–830.

[17] D. Hebesberger, C. Dondrup, T. Koertner, C. Gisinger, and J. Pripfl, "Lessons learned from the deployment of a long-term autonomous robot as companion in physical therapy for older adults with dementia: A mixed methods study," in *Proc. of HRI*, 2016, pp. 27–34.

[18] M. K. Lee, J. Forlizzi, P. E. Rybski, F. Crabbe, W. Chung, J. Finkle, E. Glaser, and S. Kiesler, "The snackbot: documenting the design of a robot for long-term human-robot interaction," in *Proc. of HRI*, 2009, pp. 7–14.

[19] N. Doering, S. Poeschl, H.-M. Gross, A. Bley, C. Martin, and H.-J. Boehme, "User-centered design and evaluation of a mobile shopping robot," *International Journal of Social Robotics*, vol. 7, no. 2, pp. 203–225, 2015.

[20] R. Pinillos, S. Marcos, R. Feliz, E. Zalama, and J. Gómez-García-Bermejo, "Long-term assessment of a service robot in a hotel environment," *Robotics and Autonomous Systems*, vol. 79, pp. 40–57, 2016.

[21] M. Tonkin, J. Vitale, S. Herse, M.-A. Williams, W. Judge, and X. Wang, "Design methodology for the ux of hri: A field study of a commercial social robot at an airport," in *Proc. of HRI*. ACM, 2018, pp. 407–415.

[22] M. J.-Y. Chung and M. Cakmak, "Exploring the use of robots for gathering customer feedback in the hospitality industry," in *HRI Workshop on Social Robots in the Wild*, 2018.

[23] T. J. Levesque and G. H. McDougall, "Service problems and recovery strategies: an experiment," *Canadian Journal of Administrative Sciences*, vol. 17, no. 1, pp. 20–37, 2000.

[24] M. Dixon, K. Freeman, and N. Toman, "Stop trying to delight your customers," *Harvard Business Review*, vol. 88, no. 7/8, pp. 116–122, 2010.

[25] M. Heerink, B. Krose, V. Evers, and B. Wielinga, "Measuring acceptance of an assistive social robot: a suggested toolkit," in *Proc. of RO-MAN*, 2009, pp. 528–533.