International Journal of Social Robotics (2024) 16:1901–1918
https://doi.org/10.1007/s12369-024-01119-z

New Comer in the Bakery Store: A Long-Term Exploratory Study Toward Design of Useful Service Robot Applications

Sichao Song[1] · Jun Baba[1] · Yuki Okafuji[1] · Junya Nakanishi[2] · Yuichiro Yoshikawa[2] · Hiroshi Ishiguro[2]

Accepted: 21 February 2024 / Published online: 8 October 2024

Abstract

In this study, we report a 6-month empirical study on a service robot deployed in a bakery shop. Recently, potential applications of service robots have been increasingly explored. However, further empirical knowledge is required to determine the optimal approach to design service robots for useful applications. We also address "usefulness" from two perspectives: the effects of a robot on customers' shopping behavior and the practical benefits the robot could provide for human workers in its working environment. The results show that our robot achieved long-term effects on product recommendations for customers who visited the bakery store on a regular basis (weekly) but not for other customers. A thematic analysis of the interviews reflected the practical values that the staff expected from the robot. Based on these findings, we we outline key considerations for designing effective long-term service robot applications.

Keywords Service robot · Sales promotion · Product recommendation · Long-term · In the wild · Real-world applications

1 Introduction

In the past decade, there has been a growing discourse on social robots. Unlike industrial robots, social robots can socialize with people, including expressions that convey emotions and natural conversations [12, 22, 32]. A well-known example is SoftBank's humanoid pepper robot, which communicates via gestures and natural language. Owing to their social communication capabilities, social robots are considered beneficial for various purposes, such as education [2, 6, 27], health [34, 35], domestic applications [19, 28, 30, 43], and public services [1, 53, 54, 56]. Social robots are particularly useful in service industries. Robots are relatively novel, particularly when deployed in residential or public spaces. Therefore, the use of robots has attracted attention [49] and has improved customers' overall experience [45]. These service robots can be defined as those that "operate partially or fully autonomously to perform services useful to the well-being of humans and equipment, excluding manufacturing operations." [8, 10]. Several potential applications of service robots have been considered such as hotel services [46, 51], information [11, 36], and recommendations [33, 56]. Specifically, service robots are a broad category of robots designed to assist humans by performing tasks such as cleaning, delivery, healthcare, and customer service. However, social robots are particularly designed for social interaction with humans. They are capable of emotional communication and of performing tasks that facilitate social interactions [40]. A service robot can be considered social if it can interact with humans in a social context and provide services [15]. For example, a robot can converse with humans in a human-like manner.

Service robots are considered particularly promising for product recommendation. Due to their ability to capture interest and attract attention, robots can engage a considerable number of potential customers, thereby creating numerous

✉ Sichao Song
song_sichao@cyberagent.co.jp

Jun Baba
baba_jun@cyberagent.co.jp

Yuki Okafuji
okafuji_yuki_xd@cyberagent.co.jp

Junya Nakanishi
nakanishi@irl.sys.es.osaka-u.ac.jp

Yuichiro Yoshikawa
yoshikawa@irl.sys.es.osaka-u.ac.jp

Hiroshi Ishiguro
ishiguro@irl.sys.es.osaka-u.ac.jp

1 AI Lab, CyberAgent, Inc., Shibuya, Tokyo 150-0042, Japan

2 Graduation School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan

interaction opportunities [3, 10, 49]. Similarly, service robots can be designed to be highly engaged with human interlocutors during interactions, and attract customers to the store and its products [17]. Moreover, service robots can use persuasive strategies to recommend products and increase sales [62].

Research on service robots for product recommendations remains largely at a relatively simple stage. Laboratory-based studies can offer limited clues for real-world applications and the design of customer-robot interactions. However, more empirical knowledge is still required to uncover customer behaviors toward service robots and understand how such robots could function in the real-world context of retail environments. Several studies have explored this concept [10, 56, 59–62]. For example, [10] placed a pepper robot in front of a chocolate store, and investigated its effect on product recommendations. They showed that their robot elicited more interactions than a tablet service kiosk, caused more people to enter the store, and increased the total revenue. [60] applied two service robots to promote sales in a retail store, one of which was placed outside the store to welcome foot traffic into the store and announce product information, while the other robot was placed inside the store mainly to recommend products. While there was no rise in customer footfall, the presence of in-store robots positively impacted the business, significantly boosting sales. [62] also explored the question of whether androids can perform a role equivalent to that of salespeople in a real working environment. They deployed a "female" android as a salesperson in a department store to sell sweaters. They reported that their robot served nearly twice as many customers as a human clerker and successfully sold 43 sweaters in ten days.

However, these studies were relatively short term (lasting a few days or weeks). The novelty effect [16, 58] could inevitably contribute to these findings, and whether customers can form long-term bonds with a robot and how the robot's effects change over time remain unclear. [18] explored the reasons why people refuse or abandon the use of robots in a long-term view. They suggested that robot designers should not only create enjoyable experiences in the short term but also provide relevant functionality to retain users over the long term. In another study [30], the authors reported the long-term evaluation of a social robot in real homes. They found that users' evaluations of the robot dropped initially and subsequently increased after the robot had been used for a longer period. From these findings, we can presume that people behave differently toward a service robot during long-term deployment.

Crucially, in our current study, the long-term effects we examine are not tied to the duration of the robot's physical presence in the environment, but rather to the frequency and regularity of people's interactions with the robot. The experience of a person who meets a robot for the first time or many times is less relevant than how long the robot has already been deployed. To the best of our knowledge, few studies have addressed this issue because most studies have focused on the effects in chronological order, as they observed the same users throughout their experiments [16, 29, 30, 55, 64]. These research methods are definitely important but could be less adequate for evaluating service robot applications, as there could be different types of customers, varying in their frequency of visiting the store and the number of times they have met the robots.

In addition to customers, consideration of the input of other parties is required in this context. [47] investigated the expectations and success criteria of the management and retailers for adopting a service robot in a shopping mall. They found that stakeholders considered it important for the robot to create a warm and fun atmosphere in the mall and provide practical help to both customers and staff. A service robot employed in an existing store would probably need to work collaboratively with the human staff. During long-term deployment, the relationship between the staff and robot may vary with time. Therefore, we observed changes in staff members' attitudes and work styles during the robot's deployment, assessing whether and how they found the robot helpful. The practical benefits that robots could provide to staff have often been neglected, as researchers and service providers tend to pay attention mainly to customers. Robots designed to offer practical support to human staff can help improve job performance and contribute to creating a better overall shopping experience for customers.

1.1 Research Objectives (ROs)

In summary, the fundamental issue is that it is still not clear how the interactions between the customers and the robot and between the store staff and the robot will change because there is no previous research investigating when service robots are introduced into actual stores over an extended period of time. We believe that it will become feasible to apply these findings to the optimal design of in-store service robots that will be useful by clarifying the relationships between humans, customers and staff, and robots over the long term. To address this, we placed a service robot in a bakery store for two days each week (Friday and Saturday) for six months. We adopted an observational study [26] to explore user behavior and shape our applicable design implications. The main task of the robot was to recommend products to customers and engage them in conversations. To enable the robot to handle various practical scenarios and overcome technical limitations, such as speech recognition in the wild [37, 56, 57], we adopted the Wizard of Oz (WoZ) method [4, 52], in which an operator remotely controls and talks through the robot. The following are the research objectives (RO):

- RO1: How will a service robot's product recommendations influence customers' purchase behavior in the long-term?
- RO2: What kinds of supports can a service robot provide for both customers and staff to be useful?

Previous studies investigated the long-term use of robots in fields such as homes, education, and care ([23, 50, 63]). However, only a few studies have been conducted on service robots installed in store environments. Regarding RO1, [48] claimed that few studies had investigated long-term persuasion of robots and conducted a study to investigate persuasion strategies for social robots. However, their experiment was not conducted in a real store configuration and the 3-week period was insufficient to reveal long-term effectiveness. Our work addresses these limitations in previous work by conducting a 6-month long-term empirical study in a real store environment. Regarding RO2, [47] showed that customers expect a service robot to provide guidance and information, an entertaining experience, and maintain the company. However, service robot to welcome customers and promote products, help them with simple tasks, and offer entertaining experiences, especially to children. However, these findings were obtained from workshops and interviews and have not been confirmed in the actual field. Our study addresses this gap by examining actual stores. Based on the results of our experiment, the expectations of customers and staff advocated in this study were largely aligned. In addition, [42] claimed that more empirical research is required to build an overarching theory, specifically regarding the long-term usage of service robots on actual behaviors, well-being, potential downsides, and risks for customers and service employees.

For RO1, weekly customer questionnaires were collected. We asked customers whether they purchased the products recommended by the robot, their frequency of store visits, how many times they met our robot, and their impressions of it. For RO2, we primarily evaluated the long-term impact of the robot on customers from their shopping behavior and impression of the robot and arranged post-experiment interviews with the staff members as well as the robot operator to reveal their thoughts and experiences working with the robot. Our interviews were semi-structured, aiming to deeply understand if and how our robot influenced the store's atmosphere, offered valuable support to the staff, and impacted their job performance.

"Usefulness" can be defined as the subjective probability that using a technology improves the way a customer completes a given task [9]. In this study, we considered the usefulness of a robot in a store environment to refer to how a service robot will improve the way customers and staff complete their tasks. The customers' main task is to shop. The usefulness of robots lies in improving shopping efficiency and experience. The main staff tasks were job tasks. Robots are useful for supporting customer service, product preparation and production, and promotional event tasks. As demonstrated in [47], customers expect a service robot to provide guidance and information, provide an entertaining experience, and keep the company, while staff expect a service robot to welcome customers and promote products, help them with simple tasks, and offer entertaining experiences, particularly for children. We consider the expectation that both customers and staff will perceive the usefulness of service robots.

1.2 Contributions

In this study, we present a 6-month empirical study in which we investigate our service robot from two perspectives: those of customers and human staff. In particular, we discovered that customers' frequency of store visits significantly influenced the long-term effects of the service robot. We also heard from the staff and accordingly suggested some important points for achieving practical benefits. With the empirical knowledge obtained from both customers and staff, this study can contribute to designing useful service robot applications in public places.

The remainder of this paper is organized as follows.

2 Empirical Setting

2.1 Service Robot System

We chose a small humanoid robot, named "SOTA" (Vstone Co., Ltd) for this study. SOTA is 28 cm tall and has a childlike appearance. Both hands have two degrees of freedom, allowing the SOTA to perform simple gestures. Although SOTA is not capable of facial expressions, three LED lights on the face, showing a mouth and two eyes, can be used to assist in natural conversations with people. Additionally, SOTA can rotate its body, which allows it to face the interlocutor and perform gaze behaviors. SOTA uses an Intel(R) Edison CPU, which is a computer-on-module system for IoT devices with integrated Bluetooth and Wi-Fi communications.

As shown in Fig. 1, we constructed a WoZ system using SOTA. The system consists of two main parts: a robot controller and an operator interface. The robot controller manages the system using speech recognition, behavioral control, and video communication through an operator interface. Specifically, a behavior-control program was developed and run on a mini-PC placed inside a control box under SOTA. A set of gestures mapped using keywords was predefined in the control program. When the operator says a target keyword such as "hello," the robot would behave accordingly, for example by raising one of its hands. For

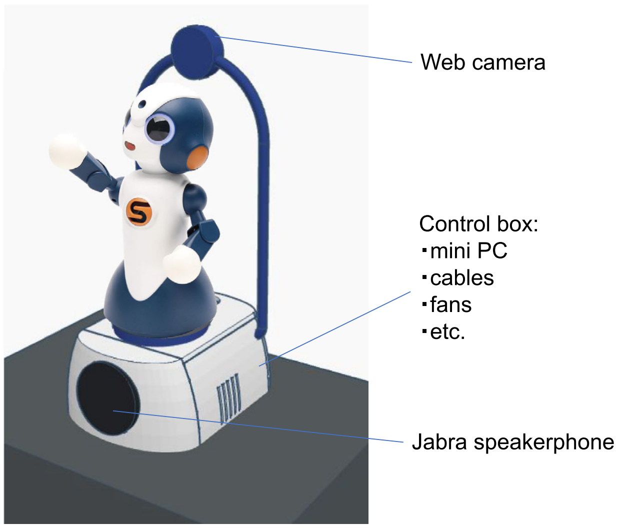


Fig. 1 3D image of the tele-operated robot system

remote operation, an ultra-wide-angle webcam was installed at the back of SOTA for a holistic view of the interactions. Additionally, voice-changing software was used to shift the pitch of the operator's voice to match SOTA's child-like appearance of the SOTA. This helped conceal the presence of the human operator from the robot's interlocutors.

The SOTA has a child-like appearance, which often attracts people's interest owing to its cuteness. There is a risk that a childlike appearance may be perceived as a lack of competence [39]. However, previous research [3, 49] has shown that the performance of SOTA is comparable to that of humans. Therefore, we consider that the potential lack of performance owing to SOTA's childlike design would not be an issue for the robot to be competent for product recommendation.

2.2 The Bakery Store

We conducted a long-term field experiment at a bakery store in Osaka, Japan. The store is relatively small, selling over 50 types of bread. It is located in a residential area near a subway station. Consequently, various types of customers visit stores, including families with children, students, company employees, and the elderly. For a normal customer, a typical routine inside the store could involve picking up a tray, looking around, checking products in a clockwise direction, and then returning to the entrance where the counter is located to complete a transaction. Figure 2 shows the layout of the bakery shop and experimental setup. We placed our service robot in a corner near the entrance of the store. This allowed the robot to monitor all the customers in the store.

For the most part, the staff perform four categories of daily tasks: bread baking and post-processing freshly baked breads for the day, preparing ingredients for the next day, managing sales events, and processing transactions. The first type of job begins early in the morning and ends at around noon. The staff were busy baking bread and putting freshly baked bread on the shelfs. The second type of job is typically performed from the afternoon until close. The staff must prepare ingredients, mostly vegetables, and classify them according to the recipes for different breads. For the third type of task, sales (called a time-sale event) often last for approximately two to three hours until closing, and the start time depends on how many breads are left to sell for the day. Finally, the staff must handle transactions at the counter located at the entrance over the course of the day.

Generally, three to four staff members work on bread baking and post-processing tasks in the early morning. From 10 a.m. to approximately 1 p.m., two workers remained in the store to perform the remaining bread-baking tasks for a few types of bread. After 1 p.m., only a single worker manages all the remaining tasks, including preparing ingredients for the next day and managing timed sales. Although the morning component has a heavier task volume, the afternoon component may be more stressful.

3 Study Design and Method

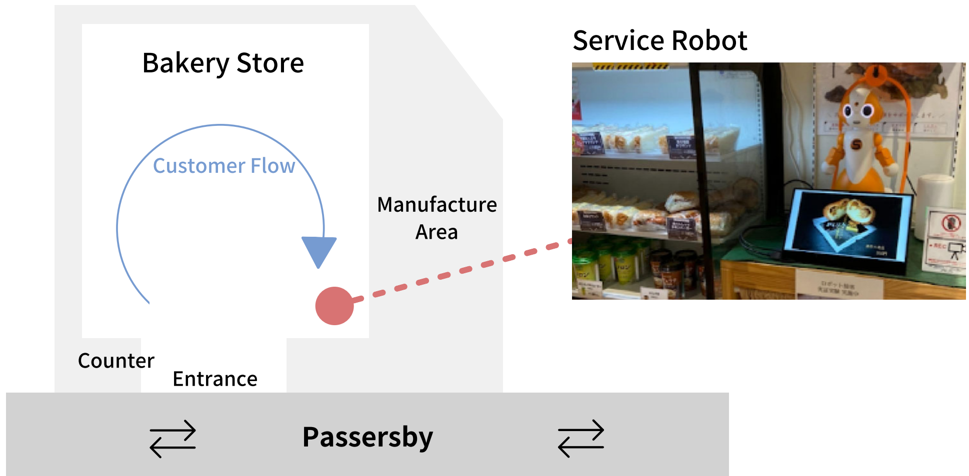
A previous work [60] was a one-week study conducted in the same bakery store over three months before this experiment. We believe that this did not influence the findings of this study.

In [60], we conducted a short-term case study to explore the potential of service robots for improving store sales. We also analyzed the behavior of both robots and customers in the store, and identified the key actions that contributed to the product recommendation effectiveness of the robots. By contrast, this study focuses on the design of useful robot services in the context of long-term service robot applications. We observed the interactions, functionalities, and behaviors that benefit both customers and store staff, summarizing the key implications for designing useful real-world service robot applications. Although both this study and the previous one were conducted in the same field, they had distinct research objectives and challenges. The novelty and contributions of this study lie in its investigation of the long-term effects of service robots by focusing on the usefulness toward different user groups, customers and staff.

3.1 Tasks

The main task of the robot was to recommend products to customers. Product recommendations is an an act of service because (1) they can provide a tailored shopping experience for each customer, making the process more efficient and enjoyable; (2) they can save customer time because they do

Fig. 2 Bakery store layout and service robot settings



not have to search through a multitude of options by themselves; and (3) they can introduce customers to products they might not have found on their own, thereby potentially increasing their overall satisfaction with the shopping experience. Therefore, product recommendations can be considered an important task for service robots in a store.

In particular, the robot recommends six types of bread out of over 50, with prices ranging from about 250 to JPY 350. As shown in Fig. 2, the display was placed in front of the robot. To facilitate the bread recommendation, we prepared photos of the products, including their price information. When the robot started talking about a particular bread, the corresponding photo was presented on the display to attract attention and induce interest. Moreover, the robot conversed with the customers. Occasionally, customers were particularly interested in the robot and attempted to chat with it or simply test its conversational capabilities. At these times, the robot responded to the customers and chat with them to offer a fun interaction experience.

In addition to the pre-designed tasks, we encouraged the robot to flexibly handle various practical scenarios and create new tasks for itself as long as such tasks did not hinder the main task of product recommendation. Practical situations occur when robots are deployed in uncontrolled environments. Situations related to customer behavior, requests from stores, or the environment can arise. This allowed us to explore which service robots should be capable of long-term deployment.

3.2 Wizard of Oz

In human-computer/robot interaction research, the WoZ technique is useful for exploring design ideas and observing user behaviors during prototyping without significant cost [13]. This helps overcome limitations such as speech recognition in noisy environments, natural language conversation, and handling complex scenarios [52]. Therefore, we adopted the WoZ technique in this study. Specifically, we required our robot to have high-quality natural language conversation capabilities and the ability to deal with various contexts and tasks that had not been specifically planned during the design phase. Therefore, the WoZ approach enabled our robot to offer satisfactory service to customers, manage practical needs raised by the local staff, and remain operational during the 6-month period of the experiment. We considered the adoption of WoZ as a necessity for achieving the objectives of this study because we would not have been able to observe the long-term effects of the robot if it could not achieve its basic performance.

We employed a female voice actor in her 20's as the operator. She was a skilled performer and communicator and was therefore qualified for scripted conversations (such as product recommendations) as well as libbed talks (chitchats). For customers, we set the concept of our robot as an autonomous conversational system. The operator talked in a performative manner and her pitch shifted higher in accordance with the child-like appearance of the robot. We conducted pre-meetings before the experiment, in which we discussed the simple characteristic design of the robot with an operator. During the experiment, the operator spoke to convey the characteristics. The local staff and owner of the bakery store were informed of the use of WoZ to achieve smooth collaboration and permission to deploy the robot. This setup would result in robot-mediated human-human interaction between the staff and the operator. However, previous work [7] has suggested that people may treat humans, autonomous robots, and tele-operated robot teammates similarly, but perceive tele-operated robots as less intelligent than human teammates. Regardless, we instructed the operator to remain in the character of the robot while talking to the staff member. We also instructed the staff to treat the robot as an autonomous system. In addition, we did not provide detailed information on our robot system or the experimental settings to the staff. Therefore, we believe that our experimental setup is sufficient for obtaining the intended data.

Specifically, the conversations of the WoZ operator's conversations were semi-scripted. Specifically, we referred to semi-scripted conversations as the structure of conversations of interaction, but the specific details and responses were not fixed or adapted to the situation. This allowed us to achieve a balance between control and flexibility because real-world scenarios can be considerably complex and difficult to script in advance. One of the main tasks of the robot was product recommendation, and we instructed the operator to follow a basic operation pipeline that (a) welcomes a customer when entering the store, (b) recommends products to the customer, and (c) thanks the customer when leaving. The operator might switch his/her target customer for a recommendation when a customer enters the store during the operation, and the current target does not respond to the robot. In cases where the customer stopped at the robot, we asked the operator to respond based on the product information sheet[1] if the customer talks about the products, introduces products if the customer remains silent but only looks at the robot, and responds flexibly if the person chitchats with the robot. When unexpected interactions occurred (including chitchat), the operator was allowed to respond flexibly as long as it kept the robot in its character (a childlike autonomous robot). We did not script the chitchats in advance because the topics could be too broad. For example, when chatting about traveling, the operator should avoid saying "I have travelled to the place" but rather than using phases such as "the place seems nice" because the robot cannot travel by itself; the operator should avoid coughing or laughing without muting the microphone because it may expose the existence of the human operator. In addition, the operator would say, "I am sorry that I do not understand" when she does not know how to respond properly. In addition, we permitted the operator to learn and perform new tasks, such as those requested by the staff, if such tasks did not hinder the robot's main product recommendation task. This allowed us to explore the expectations of staff for a useful service robot. Although it was challenging to provide a uniform dialog for customers with varying motivations for real-world robot services, we consider that this did not significantly influence the robot's primary task of product recommendation, and it provided valuable insights into the types of interactions that occur during our long-term service. These flexible conversations can contribute to our understanding of useful robot designs from the customer's perspective.

[1] The product information sheet was provided by the bakery store and contained detailed information about the products, including price, ingredients, descriptions, and examples utterances for recommendations, etc.

3.3 Evaluation

We primarily relied on two measurements in our evaluation: a customer questionnaire, and staff and operator interviews. As a quantitative approach, a customer questionnaire allowed us to track changes in the robot's long-term effects on customer behavior. By contrast, interviews helped us uncover deep insights into the relationships developed between the robot and staff as a qualitative approach.

We used the customer questionnaire to gather the following data:

- Visiting frequency: a customer's frequency of visiting the store, with options including on a weekly basis (more than once a week), on a monthly basis or less (less than three times a month), and first visit.
- Robot experience: phases regarding how many times a customer had met the robot in the store, including at the first meeting (once), in the adaption phase (2 or 3 times), and in the acquaintance phase (more than 4 times).
- Purchasing behavior: whether a customer purchased at least one of the breads that were recommended by the robot.
- Impression: a customer's impressions of the robot, including its attributes of intelligence, usefulness, ease of communication, fun, attachment, and impact [24].

Specifically, we assessed long-term customer behavior based on the answers to the question "Did you purchase at least one of the breads that were recommended by the robot?" The purchase rate (PR) is the ratio of the number of customers who bought the recommended bread to the total number of customers who answered the questionnaire. We considered PR to be a meaningful variable for measuring the long-term effects of our robot. The effect of store management on sales promotion is of utmost interest. Therefore, the product recommendation effect of the robot is a crucial factor for the store, and PR is an indicator of how the robot's product recommendation affects overall customer purchasing behavior. However, PR alone seldom provides valuable patient information. During long-term deployment, different types of customers would visit a store, varying in their frequency of visiting the store and the number of times they met the robot in the store. Visit frequency reflects customer characteristics, such as shopping habits. In contrast, the robot experience shapes a customer's impression of the robot as a mere exploration effect [65], which suggests that the greater the frequency of contact with the robot, the greater its influence. Consequently, a customer's experience with a robot affects the robot's persuasive influence.

The questionnaire items for impressions were modified based on [24]. Although the study scenarios were different, [24] investigated the user's impression of a robot during long-term use with the same objective as ours. Specifically, we translated and adapted some items from it to fit our work, for example, we modified the item "ease of use" to "ease of conversation." We intentionally designed the questionnaire to be easy to answer and kept the number of questions under ten to ensure that all questions could be answered within approximately 1–2 min. To collect unique customer data, we stationed an experimenter outside the store to ask customers who were leaving the store to cooperate, ensuring that they had not previously completed a questionnaire. In total, 378 valid responses were obtained.

We announced to all customers that this was an experiment to be conducted through a notification board. This study was conducted on an opt-out basis with participants who did not wish to participate in the experiment. No participant was excluded from the experiment. Therefore, we considered the effect of opt-out on the experimental results to be minimal.

We performed post-experiment interviews with both the local staff and the WoZ operator. Semi-structured interviews were conducted to collect the following data:

- Job environment: the environment of the store, including the atmosphere and human relations.
- Job tasks: the typical tasks and job flow for the staff and the robot.
- Relationship with customers: the relationship between the staff/robot and the customers with regards to, for example, communication, psychological distance, and trust.
- Relationship with the robot/staff: the relationship between the staff members and the robot.

Three local workers participated in the interviews. All of them worked shifts on Fridays and/or Saturdays across the 6-month long experiment, from beginning to end, and had both morning (from 7 a.m. to 1 p.m.) and afternoon shifts (from 1 p.m. until close). To avoid the risk of COVID-19 transmission, interviews were conducted online using a Zoom video communication tool. Interviews were conducted in two stages. Initially, the interviewer explained the interview's purpose and flow and asked the staff about their personal information, educational background, and tenure at the store to establish a personal context. This was followed by a brief icebreaking conversation to make the participants comfortable. This part was not used in the analysis, but rather aimed to contextualize the responses. The second part focuses on the experience of the staff with the robot. In addition to the staff members, we conducted interviews with WoZ operators.

Table 1 Demographics of the staff members (S) and the operator (P) who attended the interviews

ID	Sex	Age	Educational background
S1	Female	19	Student (Bachelor)
S2	Female	19	Student (Bachelor)
S3	Female	21	Student (Bachelor)
P	Female	29	Voice actor

We asked questions similar to those in the staff interviews but from the robot's perspective. The main purpose was to search for evidence to support the workers' statements. Table1 lists interviewees' demographic characteristics.

Each interview lasted for one hour, and a total of four hours of recorded data were obtained. The recordings were transcribed by native Japanese speakers, resulting in a transcription of over 2600 words. The transcripts were then analyzed using an inductive thematic analysis technique [14, 20]. Both the interviewers and researcher who performed the thematic analysis were co-authors of this work. The thematic analysis aimed to uncover and highlight whether and how our robot was beneficial for the different stakeholders during long-term deployment, not only for the customers but also for the staff, and to identify some important points of interest to develop useful service robot applications in the real world. The initial coding generated codes that appeared in at least two out of three staff interviews or in both staff and operator interviews. These codes were clustered into three themes based on the value of the support they represented: shopping, utilitarian, and mental support.

4 Results

4.1 Long-Term Customer Behavior

4.1.1 Product Recommendation

We explored PRs in relation to customers' frequency of visits and their experience with robots. The purchase rate over visiting frequency (PR-F) is the ratio of the number of customers who bought the recommended bread related to the three visiting frequencies: first visit (PR-F_{fi}), monthly visit (PR-F_{mo}), and weekly visit (PR-F_{we}). The results for the PR-Fs were 17.90% for PR-F_{fi}, 27.30% for PR-F_{mo}, and 27.10% for PR-F_{we}. There were no significant differences observed in the PR-Fs. In contrast, the purchase rate of the robot experience (PR-E) is the ratio of the number of customers who bought the recommended bread related to the three phases of experience: the first meeting phase (PR-E_{fi}), the adapting phase (PR-E_{ad}), and the acquaintance phase (PR-E_{ac}). The results of the PR-Es were 23.30% for PR-E_{fi}, 26.40% for PR-E_{ad}, and 38.00% for PR-E_{ac}. The Chi-square test

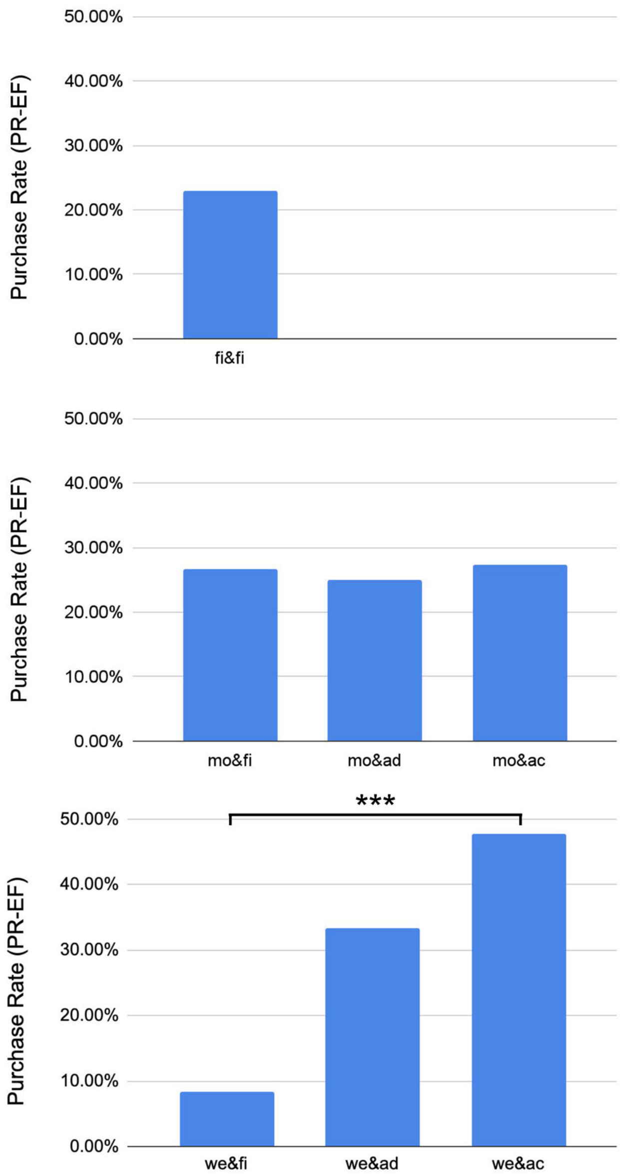


Fig. 3 Results of purchase rates with regard to robot experience and visiting frequency (PR-EFs). Regarding robot experience, fi refers to first meeting, ad refers to adaption phase, and ac refers to acquaintance phase; regarding visiting frequency, fi refers to first visit, mo refers to monthly basis or less, and we refers to weekly basis

revealed significant differences between the three conditions in PR-E ($\chi^2(2) = 6.38$, $p < .05$). Residual analysis, in comparison with the mean across all conditions, indicated that PR-E$_{ac}$ exhibited a marginally but significantly higher PR-E ($r = 2.47$, $p = .08$). A post-hoc binomial test indicated that PR-E$_{ac}$ exhibited a significantly higher PR-E than both PR-E$_{fi}$ ($p < .05$) and PR-E$_{ad}$ ($p < .05$).

We further considered the details of how PR-E varied depending on PR-F, because interaction effects may have occurred for each group of customers who had different visiting frequencies to the store; PR could vary in relation to their experience with the robot (PR-EF). As shown in Fig. 3, for customers who visited the store for the first time and had their first meeting with the robot, the purchase rate of PR-EF$_{fi\&fi}$ is 22.89%. For customers who visited the store on a monthly basis or less, the PR-EFs were 26.67% for PR-EF$_{mo\&fi}$, 25.00% for PR-EF$_{mo\&ad}$, and 27.27% for PR-EF$_{mo\&ac}$. For customers who visited the store weekly, the results are 8.33% for PR-EF$_{we\&fi}$, 33.33% for PR-EF$_{we\&ad}$, and 47.83% for PR-EF$_{we\&ac}$. In particular, for customers who visited the store weekly, the Chi-squared test showed significant differences among the three conditions in PR-EF ($\chi^2(2) = 10.95$, $p < .01$). The residual analysis, compared to the mean across all conditions, indicated that PR-EF$_{we\&fi}$ had a significantly lower PR-EF ($r = -3.15$, $p < .01$), whereas PR-EF$_{we\&ac}$ had a significantly higher PR-EF ($r = 2.90$, $p < .05$). A post-hoc binomial test indicated that PR-EF$_{we\&ac}$ had a significantly higher PR-EF than PR-EF$_{we\&fi}$ ($p < .001$).

4.1.2 Impressions of the Robot

We evaluated the customers' impressions of the robot in relation to their visiting frequency and robot experience. For visiting frequency, no significant results were observed for any attribute; for robot experience, one-way ANOVA tests revealed significant main effects on all the six attributes (Intelligence: $F = 6.32$, $p < .01$, Usefulness: $F = 3.01$, $p < .05$, Ease of communication: $F = 4.99$, $p < .01$, Fun: $F = 5.93$, $p < .01$, Attachment: $F = 3.20$, $p < .05$, Impact: $F = 7.03$, $p < .01$). In particular, the robot experience had a significant influence on customers who visited the store frequently on a weekly basis, as shown in Fig. 4, although the ratings of the robot by customers who visited the store on a monthly basis or less did not change with the number of times they had met the robot. They rated the robot lower than other customers when they had limited experience with it, particularly with regard to its intelligence and usefulness. However, when they met the robot several times and entered the acquaintance phase, their impressions increased significantly.

4.2 Staff and Operator Interviews

From the results of our thematic analysis, we identified three key forms of support our robot offered to customers and staff: Shopping support, which enhanced the customers' shopping experience, making it more seamless. Utilitarian support, aiding staff in performing their duties more efficiently and effectively. Moral or mental support, contributing to a fun store atmosphere and reducing staff stress.

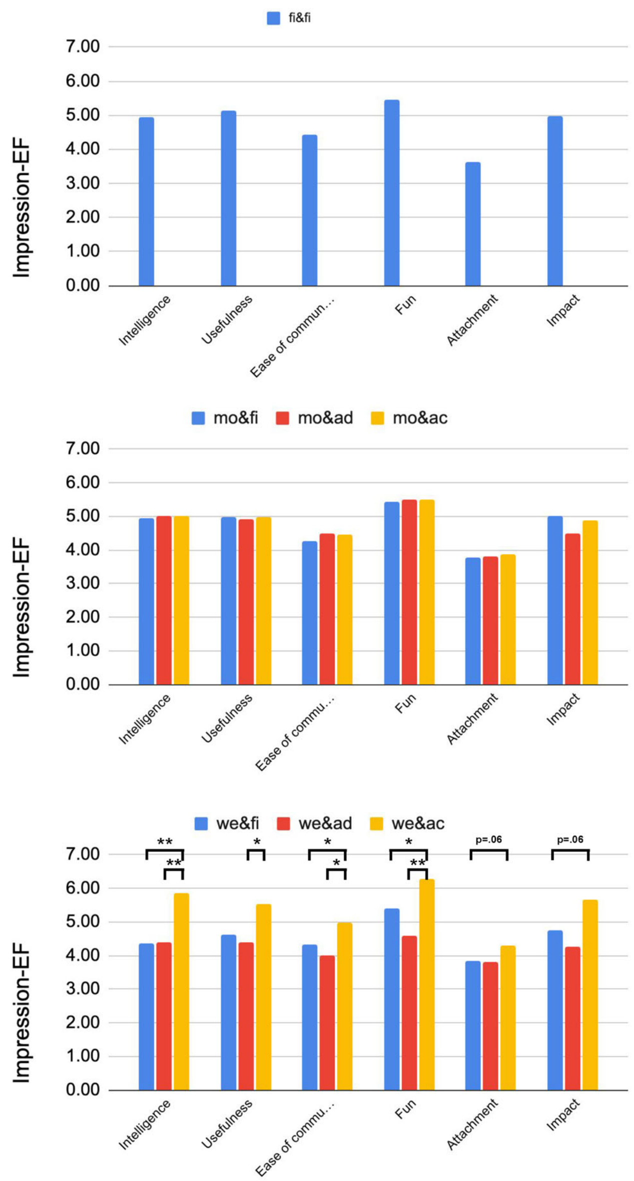


Fig. 4 Results of impressions with regard to robot experience and visiting frequency (Impression-EFs). Regarding robot experience, fi refers to first meeting, ad refers to adaption phase, and ac refers to acquaintance phase; regarding visiting frequency, fi refers to first visit, mo refers to monthly basis or less, and we refers to weekly basis

4.2.1 Shopping Support

We define shopping support as the support provided by a robot to improve the customer's shopping experience.

> "(The robot) explains details about the time-sale event, and as a result, the probability of a customer bringing a bread that was not included in the event to the counter was reduced." (S1)

The bakery store held a regular event called "time-sale," the purpose of which was to reduce the amount of bread left unsold after the store closed, as most breads would not last to the next day and had to be thrown away. Specifically, the staff checked the number of breads remaining at approximately 5 p.m. and decided when to start and what kinds of bread should be included in the time-sale list. They then wrote each day's rules on a standing signboard and placed it in front of the entrance to the store to indicate the start of a timesale event. However, this approach led to some confusion. While the event's rules were similar each day, they varied in details, particularly concerning the types of bread included and the extent of their discounts.

Customers who did not visit the store frequently tended to verify the details of the rules before entering. In contrast, customers who frequently visited stores often overtrusted their past experiences, as the rules were generally similar. This resulted in troubling situations in which they occasionally brought wrong bread to the counter and failed to receive discounts. Because of hygiene issues, customers were not allowed to place the breads back on the shelves and had to purchase them even if they failed to receive discounts and did not want to buy the breads after learning the correct price.

These troubling situations may seem avoidable, but can be unexpectedly difficult. Case in which only a single staff member managed various types of tasks during the afternoon shift Staff members are required to manage several tasks in parallel. All three workers who participated in the interviews mentioned that they could be particularly busy during the timesale and therefore had no time to explain the rules to customers. Therefore, they left customers with their own and took care of them only when they approached the counter to make a transaction. The staff reflected that customers could use support to receive information from the store in an easier way, and help them decide which bread to buy. They realized that the lack of effective communication methods between stores and customers was a common issue for bakery stores.

Staff members noticed the changes caused by the robots. During the time-sale event, the robot explained the rules for the day to customers, in addition to its main product recommendation task. This information serves as a useful resource for customers to decide on what to buy. The WoZ operator also mentioned that customers seemed to be more explicitly influenced and were less hesitant. Moreover, the operator mentioned that some customers who were confused about the rules came to the robot to ask questions. This was confirmed by the staff, and S2 raised the point that "It seemed that many customers felt more at ease and more comfortable asking the robot quick questions rather than us." The staff reported that the ratio of customers who brought the wrong bread to the counter decreased.

> "Parents often left their children chitchatting with the robot so that they could focus on shopping." (S3)

Many store visitors were parents. Depending on the situation and human characteristics, shopping with children could be stressful for some people, and focusing on checking product details might be difficult. The staff explained that bakery store customers with children could face particular hygiene issues, as the parents needed to watch the children to prevent them from arbitrarily touching their bread. Customers must complete their purchases quickly.

Children are particularly interested in robots [60]. All interviewees mentioned that our robot was particularly popular among children, as they showed excitement and chitchat with it extensively. Consequently, some parents chose to leave their children with the robot while they continued to shop. The staff noticed that customers who relied on the robot to care for their children shopped comparatively longer and checked the products more carefully. Interestingly, both staff members and the operator realized that a special relationship was established between parents, robots, and their children. Occasionally, when the robot told the children that one particular bread was brilliant, the children would turn back and ask their parents to buy it for them. After the parent put the bread on the tray, the children returned to the robot, telling them that they had bought it. The parents also asked their children about the types of bread they wanted to eat, and the children turned to the robot and asked for recommendations. As S2 mentioned that "The robot's recommendation power seemed quite strong for the children."

In conclusion, shopping support focuses primarily on customers. Customers had conversations with the robot about products and store events (e.g., timesale events) to receive information and support. The robot was useful in reducing customer shopping efforts. For instance, a robot can provide information about time-sale events, making it easier for customers to understand the information and help them choose products. Additionally, for customers with children, the robot serves as a playmate, allowing parents to concentrate on shopping. In addition, product recommendations from robots are helpful in making customers' shopping more efficient.

4.2.2 Utilitarian Support

We define utilitarian support as the type of support provided by a robot to assist staff in their job tasks.

While shopping support addresses the benefits for customers, utilitarian support focuses on the practical benefits our robot provides to staff in their job tasks.

> "(The robot) really helped me a lot during the time-sale and allowed me to focus on other tasks without worrying about the mistakes made by the customers." (S1)

The robot not only assisted customers with information and decision-making but also provided substantial support to the staff, enhancing their work efficiency. All staff who participated in the interviews emphasized and appreciated the practical support provided by the robot, sharing parts of their tasks, and simplifying their jobs. For example, the robot reduced the ratio of customers who brought the wrong bread and/or asked them about the details of the rules of the timesale event. This largely released the staff and enabled them to focus more on the tasks at hand. In particular, S1 used the word "teamwork" to reflect the relationship established between her and the robot.

The usefulness of the robot was also reflected in the reduction in bread left unsold after closing the store. Starting from the mid-afternoon and often together with time sale, the robot made special recommendations by announcing breads for which numerous products remained unsold. The staff reported that such special recommendations from robots often successfully reduced the amount of wasted bread. Before our robot was deployed in a store, the staff often had to make recommendations themselves if a type of bread had many items left. The robot performs these roles and helps reduce the difficulty of their tasks.

> "I was able to keep updated to the current situations, such as customers visiting or leaving, while working behind the workbench, thanks to the robot's speeches." (S2)

Often, the staff had to work in the production area. In such cases, they had difficulty noticing the counter areas. Consequently, they must constantly pay attention to the current scenario inside the store to recognize customers visiting or leaving for transactions. This can be cognitively demanding. The staff mentioned that the robot's speech helped them considerably with this issue. Specifically, the robot welcomed a customer when they entered the store, recommended bread according to their current standing position, and told the customer when they finished shopping and headed to the counter. By listening to the robot's words, the staff could recognize the current status of a customer and decide if they needed to immediately move to the counter. This reduces the cognitive load on workers and allows them to focus on their current tasks.

In conclusion, utilitarian support was primarily focused on staff. The staff often had conversations with the robot about products for recommendation, freshly baked bread, store events, and other simple tasks that they would like the robot to assist. The robot was useful in taking over some of the staff's job tasks and making it easier for them to make situational judgements based on the robot's actions. For instance, the robot took over the task of explaining and answering questions about time-sale events, freeing the staff to handle other tasks, such as preparing bread ingredients. In addition, because the robot greeted and conversed with customers, the staff did not have to pay close attention to whether the customers were coming in or leaving, but could judge the situation from the robot's voice. For example, when the robot said "Thank you," the staff could tell that the customer was about to leave and go to the counter to complete the transaction.

4.2.3 Mental Support

We define mental support as the type of support that the robot provides to both customers and staff to improve the atmosphere in the store.

> "The presence of the robot made me feel less lonely." (S3)

As mentioned previously, a store typically has only a single staff member for the afternoon shift. Although the number of tasks was not overly large for one person to manage, the mental stress was relatively high. All staff members who participated in the interviews addressed this point, as they sometimes felt lonely, stressed, and less motivated. S3 described a typical situation, that when she was the only person in the store, as "very quiet, although music was playing." She mentioned that such a situation, although not bad, was awkward and lacked vitality. Rather, during the experiment, the robot often talked, announcing the recommended breads even if no customer was present at that time. The staff felt that the store was full of vitality and energy, and that they had companions, particularly during the afternoon shift.

> "I felt the robot created a fun atmosphere in the store." (S2)

Not only did the robot support the staff mentally, but it also created a fun atmosphere in the store and improved the shopping experience for customers in general. The child-like appearance of the robot allowed it to easily establish a close psychological connection with its customers. Both the staff and operator mentioned that the appearance of the robot and the way it talked created a very bright and fun atmosphere in the store. Customers who chat with the robot often laugh and seem satisfied with it. Moreover, many customers who did not talk to the robot often looked at it and were curious about it.

Additionally, robots often became a topic and triggered conversations between customers and staff members. Staff mentioned that many customers proactively talked to them, particularly during counterprocessing transactions. Although mostly short conversations, the customers, particularly those who frequently visited the store, first commented on the robot such as "(The robot) is so cute" and then said something else positive about the store such as "I often visit here and your breads are delicious." These cases were rare before the deployment of the robot. Customers used the robot as an opportunity to break the walls between them and the store and sought closer relationships. As S2 noted, the robot served as a bridge to establish friendly relationships between the customers and workers.

In conclusion, mental support demonstrated its usefulness by mentally supporting both staff and customers. For the staff, particularly those who were typically alone, having the robot around reduced feelings of loneliness, and the staff felt more secure knowing that the robot was there to help if any trouble occurred. As described in the interviews, they often had chitchat with robots, particularly when there was no one else. For customers, the presence of the robot made it easier to converse with the staff, which helped them build relationships with the store. Furthermore, the robot created a fun atmosphere in the store, which was a welcome presence for both staff and customers. Many customers liked chitchat with robots for themselves and for their interests.

5 Discussion

5.1 RO1: Long-Term Effects of Our Robot

Our study addresses both the long-term effects of service robots on customers and their practical benefits to staff. Analysis of the customer questionnaires revealed that only customers who frequently visited the store seemed to be particularly influenced by robots. Customers who visited the store less frequently showed no substantially different behavior, regardless of how many times they had met the robot. As shown in Fig. 3, the PRs of the recommended breads remained approximately the same at 26% for customers who visited the store on a monthly basis or less, regardless of their robot experience, and were close to 22.89%, which was the PR for customers who visited the store for the first time. However, customers who visited the store weekly had a surprisingly low PR of only 8.33% for the first time when they met the robot. However, the PRs increased drastically afterward, reaching a significantly higher level of 47.83% in the acquaintance phase. Previous studies [16, 29, 30, 55, 64] mainly focused on the effects of robots in chronological order, as they observed the same users throughout their experiments. However, for service robot applications in public places, our results indicate that customer visiting frequency to the store should be considered an important factor in evaluating the effectiveness of the robot.

These results must be interpreted with caution, as we cannot simply conclude that our robot had no effect on customers who visited the store at a lower frequency. Rather, we treated the PR of 22.89% as a baseline representing the overall percentage of customers who bought the target bread the first time they visited the store and met the robot. We further examined how the visiting frequency to the store and the robot experience influenced customer purchasing behavior compared with the baseline. Visit frequency reflects whether a customer has relatively fixed shopping habits. Customers who visited the store more often indicated that they had developed shopping habits. Similarly, robot experience can affect the degree of influence of the robot. Frequent encounters between a customer and a robot can foster a closer relationship and greater trust, leading to a more pronounced influence from the robot.

We presume that customers who visited the store on a monthly basis or less rarely formed fixed shopping habits, and were comparatively open to the robot's recommendations the first time they met the robot. However, customers who frequently visited the store likely had their own shopping habits, meaning that they had largely decided which bread to buy in their minds before entering the store. Therefore, the results show that these customers had a significantly lower PR for the first time when they encountered the robot. They maintained their shopping lists and were less likely to change their behavior as soon as they heard the robot's recommendations. However, the robot experience shapes customer behavior differently. For customers who visited the store less frequently, the period between the two times they met the robot (current and last) could be long, for example, several weeks or even months. This long span of time hindered the establishment of close relationships between them and the robot, and as a result, diminished the effects of the robot in the long run. The results confirmed that the PRs for customers who visited the store on a monthly basis or stayed close to the baseline PR regardless of the number of times they met the robot. However, for customers who visited the store more frequently, the period between the two occasions they met the robot was signioficantly shorter, which allowed them to build long-term relationships with the robot. The effects of the robot persisted in the long run and became more powerful as the bond between customers and robots strengthened.

Although we were not able to follow up every customer and confirm whether they purchased the recommended bread every time they visited the store, we tracked purchases of the recommended products through customer questionnaires. Although it was infeasible to monitor the purchase behavior of each customer during every store visit in a real-world configuration, we could capture the proportion of customers in each category (based on the visiting frequency and number of robot interactions) who purchased the recommended products. For instance, in the category of frequent visitors (on a weekly basis), those who interacted with the robot more often were more likely to purchase the recommended products. Therefore, from a statistical standpoint, we conclude that these customers are more receptive to a robot's recommendations. We assumed that those who were not interested in the product and chose not to purchase it were equally represented across all categories.

The trends in customer impressions of the robot showed similar results, supporting the observed differences in their purchase behaviors. For customers who visited the store on a monthly basis or less, their ratings on the robot had similar scores over the six attributes compared with the baseline ratings and did not change with the number of times they had met the robot. However, for customers who visited the store more frequently (weekly), their impressions of the robot changed significantly based on their experience. In particular, those customers had relatively low scores of their ratings of the robot, specifically regarding the attributes of intelligence and usefulness, at the first time they met the robot. A plausible reason could be that, because they had their own shopping habits for the first time when they visited the store, they did not perceive the robot's recommendations as necessary and useful to them. Interestingly, as customers interacted with the robot more frequently, long-term relationships formed. They began to perceive the robot as both fun and useful, becoming increasingly open to its recommendations, more likely to attempt different products, and willing to alter their shopping habits.

The interview results reflected the long-term effects of our robot from different perspectives. Both the staff and the operator noticed that several customers, particularly children, formed friendships with the robot. They regularly visited the robot, greeted it, and typically chat with it. They wanted the robot to remember previous conversations. Interestingly, they often purchase bread incidentally. Either they asked the robot for a recommendation, or the robot used customized recommendation speeches such as "I recommend this bread for you this time because (reasons). This is different from what you tried last time!"" These customers mostly would buy it. Furthermore, more customers showed a less explicit, but more positive relationship with the robot. The staff pointed out that many regular customers tended to look at the robot with smiles or even greeted it using verbal and/or nonverbal cues such as speaking, waving hands, and nodding heads.

5.2 RO2: From Newcomer to Veteran

At the beginning of the experiment, our robot was only responsible for recommending products to customers. As a newcomer to the bakery store, it provided a novel experience and caused confusion among the local staff. This confusion reflected an unstable relationship, as they did not yet know much about the robot or appropriate ways to communicate with it. The lack of relevant pre-knowledge led the staff to choose to follow basic human-human social protocols, typically talking with the robot in a relatively formal way using honorific words and maintaining social distance from it.

Seeing the performance of the robot over a short period from the first day of deployment, the staff began to trust and form expectations from it and started to consider how it might offer help in their daily tasks. One worker mentioned in the interview that she did not have high expectations regarding the benefits of the robot. She agreed that the robot was cute and could create a fun atmosphere in the store, but questioned its practical value because it could not move and was not capable of physical tasks. However, after seeing the smoothness and power of the conversations with the robot, she had a different impression and was motivated to explore the practical help it could provide her based on its conversational functions.

Consequently, the staff started to "train" the robot with detailed information such as the ingredients and stories of the bread. Such training not only helped the robot improve its recommendation skills, but more importantly, offered an opportunity for the staff to build rapport and trust in the robot. The interviews indicated that communication became more frequent and less stressful, and friendly behaviors from the staff, such as talking in a casual way or touching the robot's head or hands, were observed. Staff members mentioned that they attempted to use the robot to support some tasks. Finally, after stable relationships are established, the robot begins to fulfill new work roles.

The findings revealed an important point: the role of a service robot designed by its developers may not sufficiently meet the practical needs of its users, which include not only customers but also other stakeholders, such as the staff of the facility. Robots can provide customers with more value by equipping themselves with useful utilities. However, for the staff who would stay close and work together with the robot, the practical benefits that the robot could provide to them often seem to have been neglected, as the developers and service providers tended to focus mainly on the customer side for whom to offer services. As [41] mentioned, performing everyday tasks together can be an example of routine behavior relevant to the maintenance of social relationships, and we suggest that robot designers pay attention to their robots' capabilities to work together and offer task support to the staff. A successful relationship between the robot and staff has a high potential to improve their overall job performance, including efficiency and motivation, and consequently contribute to customers' overall shopping experience. We need to reconsider the roles of service robots and provide more value to the different stakeholders relevant to the application.

5.3 Toward Design of Useful Service Robot Applications

From the above discussion, we note that customers and staff may have different expectations toward the value and usefulness of service robots. Specifically, a customer often looks for a smooth and convenient shopping experience and would be happy if the facility (e.g., a store) offers added value, such as fun and a pleasant atmosphere. Consequently, a service robot may require utility functions for shopping assistance as a minimum requirement. Such a robot may provide information, Q&As, and other utilities as an effective communication method to bridge the gap between customers and the store. In addition, the robot can engage in discourse and other entertainment activities, creating a more fun and lively atmosphere to improve the overall shopping experience, which could contribute to customer satisfaction and purchase intention.

Meanwhile, facility staff may prioritize the robot's benefits for their job tasks. Our findings suggest that staff expectations of the robot may align closely with those of another human staff member. Essentially, functionalities that share the existing tasks of staff and provide both physical and mental support are perceived as practically useful. Therefore, designers and researchers must consider this important point and build service robots that can learn and adapt to various new requirements to ultimately achieve satisfactory long-term use. In addition, entertainment activities and the fun and lively atmosphere created by the robot would undoubtedly bring positive added value to the staff as well.

Moreover, the three types of support discussed above had a common characteristic in that they relied on long-term adaptation among stakeholders. In the experiment, the robot successfully provided benefits and support to both the customers and staff. However, this did not occur immediately after the robot was deployed. The staff mentioned that, although they felt that the robot was fun and cute, they were not sure what the robot was capable of and how and whether they could work together with it at the beginning of the experiment. After a certain period, when the staff and the robot had adapted to each other, the staff then started to teach the robot new knowledge about breads and tasks other than recommendations, expecting more practical support for their tasks. Subsequently, the robot started being perceived more as a team member, rather than a "cute and strange being" as described by S3. We suggest that service robots be designed with learning capabilities to make adaptation progress among them and other stakeholders, particularly the staff, successfully.

For the customers, the novelty effect of the robot helped with succeeding in a positive interaction experience for the first time they met with the robot. Specifically, SOTA's cute appearance has a novel effect, which easily attracts the interest of customers visiting the store for the first time. However, the novelty effect worsened for customers who visited the store multiple times. Specifically, previous studies [25, 31] on CSCW and HCI have investigated the novelty effect. They reported that the initial spike in technology usage stabilized within a few weeks. Similarly, previous research [5, 41] in HRI contexts also noted a novelty effect in the initial stages of interactions, which faded after a short period of time. In this study, we did not directly evaluate the novelty of the robot. Based on our observations during the experiment and previous findings on the novelty effect, we consider that the novelty effect would wear off for customers who visited the store multiple times. Therefore, for a robot to keep being effective in a long-term operation, attention must be paid to the interaction design beyond the "novelty." Therefore, building long-term relationships with customers is important for the long-term effects of robots. A robot must be able to build social bonds with customers and provide them with useful functionalities and support to encourage their continuous use [41]. The staff noticed that some customers, particularly children, regularly visited the store, chitchat with the robot, and often purchased the breads recommended by the robot incidentally. The operator reflected that establishing such long-term relationships requires time and effort. Moreover, the customer questionnaire results suggested that our robot achieved long-term effects on product recommendations, particularly for customers who visited the store on a weekly basis. These findings indicate design policies in which service robots should take particular care to build and maintain long-term relationships with regular customers and may provide added value to them to make the most of the robots' effects in the long run. Remembering customers and topics of conversations they had in the past could be quite effective, although ethical issues need to be addressed. However, this does not imply that other types of customers should be ignored. Robots can use the novelty effect to increase their influence and improve their overall experience.

We consider the following potential indicators, in addition to purchase rate, of long-term usefulness: (1) the time a customer has had a conversation with the robot and whether or not the customer is willing to meet the robot again could be treated as indicators of long-term usefulness for customers; and (2) whether or not a staff member perceives the robot as a teammate could be used as an indicator for the staff. In other studies [24, 29, 30], long-term indicators such as social acceptance were used to trace individuals. However, in this study, we were not able to constantly ask an individual to participate in a survey in the field; therefore, we did not consider these indicators.

In addition, from a marketing perspective, social robots can offer unique value propositions to both customers and staff. For customers, social robots can provide a tailored shopping experience for each customer, save customer time, improve efficiency, and increase their overall satisfaction, while for staff, social robots can handle tasks such as information and product recommendations with consistent service quality, freeing the staff to focus on other important tasks. Compared to less sophisticated technologies, such as digital kiosks, social robots can better attract attention and offer wealthy interactive and real-time feedback. Embodied and emotional communication can significantly increase customer service and engagement. A previous study [10] compared the effectiveness of a humanoid service robot with that of a tablet service kiosk and showed that the robot was more effective in attracting passersby and converting them into buyers. In addition, welcoming customers and promoting products are listed as tasks that retailers expect robots to perform [47]. This can be considered a unique value that can only be offered by social robots.

5.3.1 Design Implications

A previous study [30] discussed the mere-exposure effect [65] on people's evaluations of a robot during its long-term use in homes. The authors found that users' enjoyment increased when they saw robots as familiar interactants, and familiarizing themselves with robots caused them to experience more meaningful social interactions. Our observations confirmed similar results and suggested that the visiting frequency could influence a customer's familiarity with a robot and further affect its effectiveness. Interestingly, it is difficult for customers who visit a store less frequently to familiarize themselves with robots and maintain social relationships with them. This leads to an important question for robot designers and researchers "How can we encourage customers to visit and interact with a robot more frequently to maintain their relationship with the robot?"

The most common reason for the discontinuance of robot use was that people could not find added value for the robot [41], and people often abandoned a robot if they thought the robot had no utility value [38]. Our findings suggest that a service robot may offer various types of support to customers, such as shopping assistance (e.g., providing information, answering questions, and childcare), making it more useful and encouraging long-term user interaction.

In addition to the utility value, humans perform various activities to maintain relationships with others, and a robot needs to do the same. Simple interactions, such as greetings and short enjoyable conversations, can be good examples for robots to maintain basic relationships with their users. Moreover, strategic behaviors such as recalling past conversation topics and interactants' personal information (e.g., names) could be effective in maintaining social bonds and increasing trust and intimacy [21]. We believe that these social interactions form the essential value of a social robot, which differentiates it from other gadgets [38].

Our results of sales based on product recommendations suggest that increasing customers' frequency of store visits could positively enhance the robot's performance on product recommendations. However, previous studies [60] found that the utilization of robots alone may hardly contribute to an increase in the number and frequency of visits. However, previous research [44] explored the use of a hybrid system consisting of both physical robots and smartphone platforms to provide richer interaction opportunities and experiences. We believe that designers can make good use of such strategies to motivate customers to visit stores more often. This can, in turn, positively contribute to improving the performance of service robots in long-term operations.

Based on our observations, we defined three types of support that a service robot can provide to both customers and staff. To improve customers' shopping efficiency and offer a better shopping experience, robots are recommended to provide information, such as sales events, which needs to be easy to understand. Additionally, designers could consider equipping robots with assistive functionalities to allow customers to concentrate better on shopping. For instance, the robot could provide entertainment services and act as a playmate for children, thereby offering parents a respite from babysitting duties. To better assist staff in their job tasks, we suggest first looking into the store environment and workflows of the staff and designing the robot's functionalities accordingly. In particular, our observations reveal that the staff often treat the robot as a teammate and want it to share parts of their job tasks. The designer could pay attention to repetitive and routine tasks to ensure that the robot could help relieve the staff and better concentrate on higher-level tasks. Furthermore, designers can make good use of a robot's social capabilities, for example, chitchats and entertaining performances, and this could help to create a fun and hospitable atmosphere in the store, which not only offers a welcome presence for customers but also reduces mental stress for the staff.

Although we used WoZ to overcome technical limitations, our observations revealed that most of the interactions discussed above did not require the robot to have high conversational capabilities. We consider that many of these interactions can be implemented by autonomous robot systems, whereas some may be technically challenging at present. We summarize our design implications and hope that robot designers and researchers will find these beneficial.

- *An interaction can be implemented by current autonomous systems.* Many tasks that our robot performed during the experiment could be implemented by autonomous conversational systems. For example, the majority of the robot's utterances of product recommendation were unidirectional, as well as greetings and farewells; enjoyable interactions with children can rely on entertainment contents such as interactive games or storytelling; some tasks for staff support, such as event announcement, can also be implemented autonomously. These simple tasks are typically paid less attention to by the robot designers but could play essential roles in the building and maintaining of social relationships between the robot and its users.
- *An interaction can potentially be implemented by current autonomous systems while facing technical challenges.* For example, the robot's conversations of product recommendation could adapt to a customer's status, such as position and attribute, to improve effectiveness, and this depends on the robot's ability of customer recognition; some tasks for staff support would also require inference of the staff's status and intent. These tasks could be technically challenging but potentially realizable by current technologies, and we suggest the designers consider making use of them for improving the task performance of their robots.
- *An interaction can hardly be implemented by autonomous systems due to the current limitations of technology.* For example, high-quality conversations of chitchat can be a difficult task for a current conversational system to achieve; learning new tasks and knowledge, such as detailed information about a product, through conversations and demonstrations is still challenging, particularly in a wild environment. However, we suggest that designers and researchers address these challenges because they could potentially bring considerable added value to future service robots.

6 Conclusion

In this study, we conduct a long-term experiment on the usefulness of service robots in baking stores. Specifically, we focused on two groups, customers and staff, and investigated their long-term interactions with robots. We observed that our robot had long-term product recommendation effects, particularly for customers who visited the bakery store on a regular basis (weekly), but not for other customers.. Moreover, for the customers, the robot provided support to their shopping and contributed to the establishment of relationships between them and the store. For the staff, the robot was described as a useful teammate and provided support both to their job tasks and mentality. Based on these findings, we propose design implications for useful service-robot applications.

6.1 Limitations and Future Work

This study has several limitations that should be addressed in the future. First, we used WoZ to overcome the technical limitations. Although our results revealed that many interactions and tasks of our robot can be implemented by autonomous systems, there are some interactions that could be potentially effective, but are currently difficult to fully realize without the help of human operators. Therefore, future research should use WoZ with caution and pay attention to the current technical limitations to better generalize the findings to autonomous service robot systems. Secondly, the experiment was conducted in a baking store. Other facilities with different services and products may engage in various customer behaviors. Therefore, future research should continue to explore various contexts to obtain more empirical observations and knowledge. Third, although we focused on our robot's long-term effects on customer behavior, for practical reasons, we could not collect detailed sales data and perform customer interviews because of the restrictions from the commercial facility (reasons: COVID-19 and disturbing other customers, etc.) We consider many of the difficulties could be trade-offs for conducting field experiments but suggest future work to collect various kinds of data to improve validity. In addition, we shared the WoZ settings with the staff and the owner of the store. Although previous research [7] suggests that people may treat autonomous robots, humans, and tele-operated robot as teammates, a bias could inevitably occur in our findings. Researchers could explore improved study designs to examine autonomous service robots, focusing on how these robots can develop relationships with and provide practical value to staff and other stakeholders.

Author Contributions SS, JB, YO, and JN conceptualized the study. All authors designed the experiments. SS, JB, and JN programmed the software. SS collected data. SS analyzed data. All authors interpreted results. SS wrote the manuscript.

Funding Open Access funding provided by Osaka University. This work was supported by JST Moonshot R&D Grant Number JP-MJMS2011.

Data Availability The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest This study received funding from the company CyberAgent, Inc. Authors SS, JB, and YO are employees of CyberAgent, Inc., and the remaining authors are employed at Osaka University. The funder was not involved in the study design, collection, analysis, interpretation of data, writing of this article, or the decision to submit it for publication. The authors declare that they have no other competing interests.

Ethical Approval This experiment was approved by the Research Ethics Committee of Osaka University (Reference number: R1-5-9).

References

1. Aaltonen I, Arvola A, Heikkilä P et al (2017) Hello pepper, may i tickle you? children's and adults' responses to an entertainment robot at a shopping mall. In: Proceedings of the companion of the 2017 ACM/IEEE international conference on human-robot interaction, pp 53–54
2. Ahmad MI, Khordi-Moodi M, Lohan KS (2020) Social robot for stem education. In: Companion of the 2020 ACM/IEEE international conference on human-robot interaction, pp 90–92
3. Baba J, Song S, Nakanishi J et al (2021) Local vs. avatar robot: performance and perceived workload of service encounters in public space. Front Robot AI 8:778753
4. Baxter P, Kennedy J, Senft E et al (2016) From characterising three years of HRI to methodology and reporting recommendations. In: 2016 11th ACM/IEEE international conference on human-robot interaction (HRI), IEEE, pp 391–398
5. Belpaeme T (2020) Advice to new human-robot interaction researchers. Human-robot interaction: evaluation methods and their standardization, pp 391-398
6. Belpaeme T, Kennedy J, Ramachandran A et al (2018) Social robots for education: a review. Sci Robot 3(21):eaat5954
7. Bennett M, Williams T, Thames D et al (2017) Differences in interaction patterns and perception for teleoperated and autonomous humanoid robots. In: 2017 IEEE/RSJ international conference on intelligent robots and systems (IROS), IEEE, pp 6589–6594
8. Bertacchini F, Bilotta E, Pantano P (2017) Shopping with a robotic companion. Comput Human Behav 77:382–395
9. Blut M, Wang C, Wünderlich NV et al (2021) Understanding anthropomorphism in service provision: a meta-analysis of physical robots, chatbots, and other AI. J Acad Market Sci 49:632–658
10. Brengman M, De Gauquier L, Willems K et al (2021) From stopping to shopping: an observational study comparing a humanoid service robot with a tablet service kiosk to attract and convert shoppers. J Bus Res 134:263–274
11. Brščić D, Ikeda T, Kanda T (2017) Do you need help? a robot providing information to people who behave atypically. IEEE Transact Robot 33(2):500–506
12. Cabibihan JJ, Javed H, Ang M et al (2013) Why robots? a survey on the roles and benefits of social robots in the therapy of children with autism. Int J Soc Robot 5(4):593–618
13. Carros F, Meurer J, Löffler D et al (2020) Exploring human-robot interaction with the elderly: results from a ten-week case study in a care home. In: Proceedings of the 2020 CHI conference on human factors in computing systems, pp 1–12
14. Clarke V, Braun V, Hayfield N (2015) Thematic analysis. Qual Psychol A Pract Guide Res Methods 222(2015):248
15. Dautenhahn K (2007) Socially intelligent robots: dimensions of human-robot interaction. Philos Transact Royal Soc B Biol Sci 362(1480):679–704
16. Davison DP, Wijnen FM, Charisi V et al (2020) Working with a social robot in school: a long-term real-world unsupervised deployment. In: 2020 15th ACM/IEEE international conference on human-robot interaction (HRI), IEEE, pp 63–72
17. De Gauquier L, Brengman M, Willems K et al (2021) In or out? a field observational study on the placement of entertaining robots in retailing. Int J Retail Distrib Manag 49(7):846–874
18. De Graaf M, Allouch SB, Van Diik J (2017) Why do they refuse to use my robot?: reasons for non-use derived from a long-term home study. In: 2017 12th ACM/IEEE international conference on human-robot interaction (HRI, IEEE), pp 224–233
19. De Graaf MM, Allouch SB, Klamer T (2015) Sharing a life with harvey: exploring the acceptance of and relationship-building with a social robot. Comput Human Behav 43:1–14

20. Dereshev D, Kirk D, Matsumura K et al (2019) Long-term value of social robots through the eyes of expert users. In: Proceedings of the 2019 CHI conference on human factors in computing systems, pp 1–12
21. Duck S (1998) Human relationships. Sage, Newcastle upon Tyne
22. Duffy BR (2003) Anthropomorphism and the social robot. Robot Auton Syst 42(3–4):177–190
23. Fernaeus Y, Håkansson M, Jacobsson M et al (2010) How do you play with a robotic toy animal? a long-term study of pleo. In: Proceedings of the 9th international conference on interaction design and children, pp 39–48
24. Fink J, Bauwens V, Kaplan F et al (2013) Living with a vacuum cleaning robot. Int J Soc Robot 5(3):389–408
25. Gallacher S, O'Connor J, Bird J, et al (2015) Mood squeezer: lightening up the workplace through playful and lightweight interactions. In: Proceedings of the 18th ACM conference on computer supported cooperative work & social computing, pp 891–902
26. Gockley R, Bruce A, Forlizzi J et al (2005) Designing robots for long-term social interaction. In: 2005 IEEE/RSJ international conference on intelligent robots and systems, IEEE, pp 1338–1343
27. Gordon G, Breazeal C, Engel S (2015) Can children catch curiosity from a social robot? In: Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction, pp 91–98
28. de Graaf MM, Allouch SB (2015) The evaluation of different roles for domestic social robots. In: 2015 24th IEEE international symposium on robot and human interactive communication (RO-MAN), IEEE, pp 676–681
29. de Graaf MM, Allouch SB, van Dijk JA (2016a) Long-term acceptance of social robots in domestic environments: insights from a user's perspective. In: 2016 AAAI spring symposium series
30. de Graaf MM, Allouch SB, van Dijk JA (2016) Long-term evaluation of a social robot in real homes. Interact Stud 17(3):462–491
31. Hazlewood WR, Stolterman E, Connelly K (2011) Issues in evaluating ambient displays in the wild: two case studies. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp 877–886
32. Hegel F, Muhl C, Wrede B et al (2009) Understanding social robots. In: 2009 second international conferences on advances in computer-human interactions, IEEE, pp 169–174
33. Herse S, Vitale J, Tonkin M et al (2018) Do you trust me, blindly? factors influencing trust towards a robot recommender system. In: 2018 27th IEEE international symposium on robot and human interactive communication (RO-MAN), IEEE, pp 7–14
34. Hung L, Liu C, Woldum E et al (2019) The benefits of and barriers to using a social robot paro in care settings: a scoping review. BMC Geriatr 19(1):1–10
35. Iroju O, Ojerinde OA, Ikono R (2017) State of the art: a study of human-robot interaction in healthcare. Int J Inf Eng Electron Bus
36. Kaipainen K, Ahtinen A, Hiltunen A (2018) Nice surprise, more present than a machine: Experiences evoked by a social robot for guidance and edutainment at a city service point. In: Proceedings of the 22nd international academic mindtrek conference, pp 163–171
37. Kennedy J, Lemaignan S, Montassier C et al (2017) Child speech recognition in human-robot interaction: evaluations and recommendations. In: Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction, pp 82–90
38. Kertész C, Turunen M (2017) What can we learn from the long-term users of a social robot? In: International conference on social robotics, Springer, pp 657–665
39. Lacey C, Caudwell C (2019) Cuteness as a 'dark pattern'in home robots. In: 2019 14th ACM/IEEE international conference on human-robot interaction (HRI), IEEE, pp 374–381
40. Larsson S, Liinason M, Tanqueray L et al (2023) Towards a socio-legal robotics: a theoretical framework on norms and adaptive technologies. Int J Soc Robot 1–14
41. Leite I, Martinho C, Paiva A (2013) Social robots for long-term interaction: a survey. Int J Soc Robot 5(2):291–308
42. Lu VN, Wirtz J, Kunz WH et al (2020) Service robots, customers and service employees: what can we learn from the academic literature and where are the gaps? J Serv Theor Pract 30(3):361–391
43. Luria M, Hoffman G, Zuckerman O (2017) Comparing social robot, screen and voice interfaces for smart-home control. In: Proceedings of the 2017 CHI conference on human factors in computing systems, pp 580–628
44. Matsumura R, Shiomi M (2022) An animation character robot that increases sales. Appl Sci 12(3):1724
45. Nakanishi J, Baba J, Kuramoto I, et al (2020a) Smart speaker vs. social robot in a case of hotel room. In: 2020 IEEE/RSJ international conference on intelligent robots and systems (IROS), IEEE, pp 11391–11396
46. Nakanishi J, Kuramoto I, Baba J et al (2020) Continuous hospitality with social robots at a hotel. SN Appl Sci 2(3):1–13
47. Niemelä M, Heikkilä P, Lammi H (2017) A social service robot in a shopping mall: expectations of the management, retailers and consumers. In: Proceedings of the Companion of the 2017 ACM/IEEE international conference on human-robot interaction, pp 227–228
48. Okafuji Y, Baba J, Nakanishi J et al (2021) Persuasion strategies for social robot to keep humans accepting daily different recommendations. In: 2021 IEEE/RSJ international conference on intelligent robots and systems (IROS), IEEE, pp 1950–1957
49. Okafuji Y, Ozaki Y, Baba J, et al (2022) Behavioral assessment of a humanoid robot when attracting pedestrians in a mall. Int J Soc Robot 1–17
50. Ostrowski AK, Breazeal C, Park HW (2022) Mixed-method long-term robot usage: Older adults' lived experience of social robots. In: 2022 17th ACM/IEEE international conference on human-robot interaction (HRI), IEEE, pp 33–42
51. Reis J, Melão N, Salvadorinho J et al (2020) Service robots in the hospitality industry: The case of HENN-NA hotel, Japan. Technol Soc 63(101):423
52. Riek LD (2012) Wizard of oz studies in HRI: a systematic review and new reporting guidelines. J Human-Robot Interact 1(1):119–136
53. Saad E, Neerincx MA, Hindriks KV (2019) Welcoming robot behaviors for drawing attention. In: 2019 14th ACM/IEEE international conference on human-robot interaction (HRI), IEEE, pp 636–637
54. Sabelli AM, Kanda T (2016) Robovie as a mascot: a qualitative study for long-term presence of robots in a shopping mall. Int J Soc Robot 8(2):211–221
55. Scassellati B, Boccanfuso L, Huang CM et al (2018) Improving social skills in children with ASD using a long-term, in-home social robot. Sci Robot 3(21):eaat7544
56. Shiomi M, Shinozawa K, Nakagawa Y et al (2013) Recommendation effects of a social robot for advertisement-use context in a shopping mall. Int J Soc Robot 5(2):251–262
57. Shiomi M, Kanda T, Howley I et al (2015) Can a social robot stimulate science curiosity in classrooms? Int J Soc Robot 7(5):641–652
58. Smedegaard CV (2019) Reframing the role of novelty within social HRI: from noise to information. In: 2019 14th ACM/IEEE international conference on human-robot interaction (HRI), IEEE, pp 411–420
59. Song S, Baba J, Nakanishi J et al (2021) Teleoperated robot sells toothbrush in a shopping mall: a field study. In: Extended abstracts of the 2021 CHI conference on human factors in computing systems, pp 1–6
60. Song S, Jun B, Nakanishi J et al (2022) Service robots in a bakery shop: a field study. In: 2022 IEEE/RSJ international conference on intelligent robots and systems (IROS), IEEE, pp 134–140
61. Tonkin M, Vitale J, Ojha S et al (2017) Would you like to sample? robot engagement in a shopping centre. In: 2017 26th IEEE interna-

tional symposium on robot and human interactive communication (RO-MAN), IEEE, pp 42–49
62. Watanabe M, Ogawa K, Ishiguro H (2015) Can androids be salespeople in the real world? In: Proceedings of the 33rd annual ACM conference extended abstracts on human factors in computing systems, pp 781–788
63. Weiss A, Pillinger A, Tsiourti C (2021) Merely a conventional 'diffusion' problem? on the adoption process of anki vector. In: 2021 30th IEEE international conference on robot & human interactive communication (RO-MAN), IEEE, pp 712–719
64. Westlund JMK, Park HW, Williams R et al (2018) Measuring young children's long-term relationships with social robots. In: Proceedings of the 17th ACM conference on interaction design and children, pp 207–218
65. Zajonc RB (1968) Attitudinal effects of mere exposure. J Pers Soc Psychol 9(2p2):1