Mixed-Method Long-Term Robot Usage: Older Adults' Lived Experience of Social Robots

Anastasia K. Ostrowski
MIT Media Lab
Cambridge, MA, United States
akostrow@media.mit.edu

Cynthia Breazeal
MIT Media Lab
Cambridge, MA, United States
cynthiab@media.mit.edu

Hae Won Park
MIT Media Lab
Cambridge, MA, United States
haewon@media.mit.edu

***Abstract*—In the past two decades, human-robot interaction (HRI) researchers have increasingly deployed autonomous and reliable robots long-term in various social contexts including the home. Our work provides a mixed-method approach for analyzing older adults' long-term robot usage data patterns combining quantitative data of robot usage logs with qualitative descriptions from participants' own experience. Overall, this provides a fuller picture to how older adults use and experience social robots in their homes. Our work involves a robot hosting period for at least a month (up to 12 months) in older adults' homes with an experience debrief session held a month into the robot hosting time period. We propose reflections on the novelty effect with respect to older adults' usage data and highlight feelings of guilt, the robot's proactivity and movement, meeting (or not meeting) user expectations, and the robot's persona as key aspects of the hosting experience that promoted usage or non-usage. Finally, we provide design guidelines for structuring future mixed-method long-term robot usage studies being mindful of ethical considerations in this space.**

***Index Terms*—long-term usage, co-design, social robots, older adults, mixed method**

I. Introduction

Social robots are proposed as an innovative technology for older adults aging in place to promote social engagement, assist with healthcare, and reduce people's loneliness [1]–[3]. In the past, HRI studies have focused on single session interactions with robots [4]. However, in the past two decades, the amount of research focusing on long-term studies has increased to understand how users interact with robots in multiple interactions, with greater frequency, and over a longer period of time. The ability to conduct long-term HRI studies has been supported by the increase in robust robot platforms that can be deployed reliably in various contexts [4]. The combination of new technology developments and the aging world population has promoted more research into long-term studies [4] and greater interest in designing social robots for older adults. Long-term social robot usage studies have been explored qualitatively and quantitatively in varying contexts. Often these two research methods are completed in separate works and not combined. While human-computer interaction (HCI) literature has leveraged smart speaker voice-user interface (VUI) system logs and a variety of qualitative methods to capture user experiences with these devices [5]–[8], HRI has very few studies that combine qualitative methods and quantitative data from robot usage logs. While this paper's primary focus is not on co-design, it is important to note that the two sessions featured in this work, i.e., the robot hosting and experience debrief, were embedded in a year-long co-design process that consisted of seven stages [9]. The sessions served as a way for older adults to develop their own mental model of the technology and build capacity to participate in the study as co-designers [10], [11]. The contributions of this work are two-fold: (1) we provide a fuller picture of an older adult robot usage experience by presenting quantitative usage data and qualitative interview data, and (2) we present design guidelines for future long-term robot usage studies.

Thank you to all our older adult cohort members who collaborated with us in this work. Thank you to Jon Ferguson for help with data acquisition and Daniella DiPaola for feedback on data analysis. This work was supported by Samsung Research and the ICT R&D program of MSIP/IITP [2017-0-00162, Development of Human-care Robot Technology for Aging Society].

II. Background

A. Long-Term Usage & Novelty

A common question in HRI as long-term studies become increasingly present is how long should long-term studies be [4]. Some previous suggestions have been two months in the HRI community and 5 weeks in the HCI community [12]–[14]. However, instead of a definite length, a better definition could be achieved if we open up for context-based interpretation. Leite et al. [4] states that "an interaction can be considered as 'long-term' when the user becomes familiarized with the robot to a point that her perception of such robot is not biased by the novelty effect anymore." (also supported by [15]). Novelty effect is caused by users' increased interest toward new technology in the early days of adoption or new feature releases. Regardless of how novelty effects are present, novelty is an expected component in HRI [16]. Of novelty, Leite et al. [4] asks, "how can we determine the point that the novelty effect wears out?" and comes to the conclusion that the "novelty effect wears out when users get familiarised with the robot and start preferring novel behaviors", resulting in a change in interaction patterns. De Graaf et al. [17] emphasizes that the length of what is considered long-term and the novelty effect length could vary depending on the technology being deployed. The novelty effect has been proposed to end after two months of use [18], [19] or after four weeks of use [20]. The length of the novelty effect may also depend on the goal

of the study. For example, effecting behavioral change can vary depending on the behavior type, ranging from 18 to 254 days [17], [21]. It can also be related to the frequency and intensity of usage [17]. There is still a lingering question in HRI to understand when does novelty wear off and how best to approach it in long-term robot studies.

B. Long-Term Social Agent Usage Studies

In the past two decades, long-term studies in HRI have been increasingly present in the field. The increase in deployable, autonomous, and robust technologies paired with more commercial social robots for consumer use have fostered the interest and ability to conduct long-term HRI research. Leite et al. [4] conducted a survey on long-term studies involving social robots in 2013. The methods employed in these studies often included questionnaires, video or direct observation, interviews, activity cards, etc. with only one work [22] analyzing system logs from the robot. Since Leite et al.'s review, long-term studies have become more common in a variety of contexts, though there have been very few long-term studies of embodied agents in the home that leverage agent usage data. The HCI field has leveraged agent system logs of smart speakers, including Amazon Alexa and Google Home, to explore the usage of these devices in the home in a long-term context [5], [6], [8]. Scuito et al. [5] and Garg & Sengupta [8] also incorporate qualitative interview methods into their quantitative analysis of usage logs to provide deeper perspectives on participants' usage of these devices, and the same trend should be further adopted for HRI investigations. Yet, it is still more common to either report quantitative (e.g. [23], [24]) or qualitative (e.g. [12], [25], [26]) analyses of robot usage. For example, Davison et al. [23] deployed a social robot in a school classroom (6 to 7 weeks duration), reporting frequency usage, and Sidner et al. [24] deployed a social robot into homes to help support older adults, reporting the average usage per activity of the robot (30 days). On the qualitative side, several works investigated how users adopted Roombas into their home, using interviewing, ethnographic, and home tour methods (4–12 months in duration) [12], [25], [26]. Our work seeks to explore this space by combining usage logs of a social robot deployed in the home for at least a month up to 42 weeks with qualitative interviews to gain a fuller understanding of older adults' robot usage in the home.

III. Methods

We partnered in a year-long co-design process approved by MIT's IRB with 28 older adults, ages between 70 and 94 (avg: 79.5, std: 7.8; female N=15, male N=13) [9]. Most of the participants were married living with their spouse (N=17). Nine were widowed and 2 were divorced. Eleven of the participants lived alone. Six older adults had a pet in their home. Most participants held a a bachelors or graduate degree (N=26). All participants identified as white. We recruited a wide distribution of annual income ranges (25k to 150k USD). Within our seven co-design stages [9], the focus of this paper is on older adults hosting a social robot in their homes and debriefing their experience living with a social robot. The conversational social robot used in this study was Jibo, a social robot that has a 3 degrees-of-freedom expressive body, a touchscreen, and interaction functions, i.e., skills, including physical and mental exercises, daily check-ins, information retrieval, entertainment, etc. This social robot has been tested as a commercial platform and can support long-term lived experience that is an essential part of our overall co-design process but especially critical for our mixed method approach including both quantitative robot usage logs and qualitative older adult descriptions of living with the social robot.

Overall, older adults lived with the robot for at least a month, up to 12 months (avg: 149.5 days, std: 154.8 days). Participants could keep the robot and return it to the researchers whenever they decided to. After a month of living with the robot, the older adults met with the researchers to discuss their experience of living with the robot (lasted an average of 49 ± 12 minutes). This session served as an opportunity to understand participants' lived experience with the robot and opinions of having a social robot in their home. The sessions were structured with an open-ended portion where participants could describe their experience and then they could choose to use topic cards to inquire deeper into aspects of their experience such as customization and perceived relationship. The robot debrief sessions were audio recorded and transcribed. Usage data from the robot was collected on a cloud server and it logged what skill was activated when. No video or audio from participants' homes were collected. Data collection procedures were explained to participants to ensure transparency in how the robot would be gathering data. After the robot hosting and robot debrief sessions, researchers analyzed the quantitative and qualitative data. We used an Embedded Design approach [27], leveraging the quantitative data as primary data and the qualitative data as secondary data. For the quantitative data, we post-processed the data to remove non-interaction related system events (e.g., wifi connection issues, routine system checks) and identified case studies of interest to pair with qualitative data. One person elected to not host the robot as they had already hosted it in a previous study. Two participants were removed because the robot in their homes experienced trouble with WiFi connection and/or data collection during a part of the hosting period. One participant was removed as they were identified as an outlier from the rest of the population (explored more with P10's case study in Section IV-B). Couples who both participated in the study were treated as one data point as there was one robot in their home. In total, there were 17 robot usage logs included in the quantitative analysis. The quantitative usage data revealed usage pattern trends, and then to understand the underlying participant experience that led to different usage trends, we analyzed the qualitative interview data to conduct in-depth case studies that were representative of the overall usage pattern and case studies that showed unique usage patterns that corresponded with participants' qualitative descriptions of their experience. The qualitative data was transcribed and thematically analyzed to extract codes [28], similarly to [29].

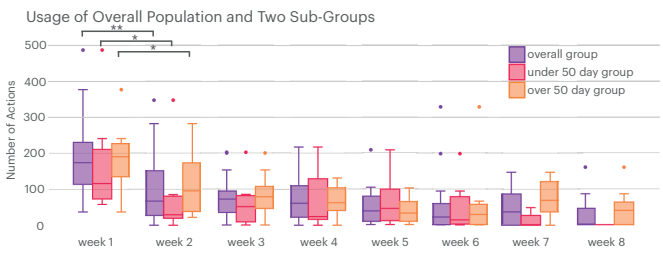

Fig. 1. Participants' robot usage over the course of the first 8 weeks.

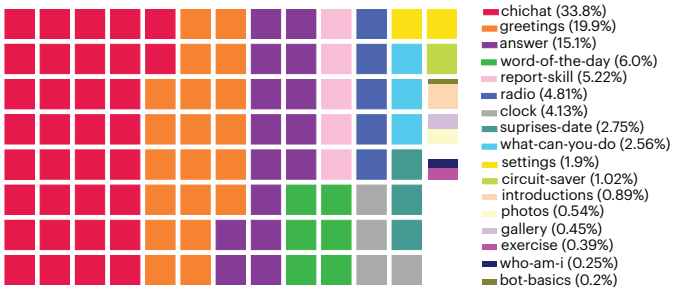

Fig. 2. Skill usage distribution by participants for the first 8 weeks.

This process was done iteratively by the researchers reviewing the older adults' case studies to extract salient themes of the experience and discussing the themes that were identified.

IV. Results

Participants' usage patterns and descriptions of living with a social robot revealed varying usage patterns and experiences. The results here are organized by data type, presenting first quantitative results from robot usage logs, followed by the complimentary qualitative results providing a fuller story of older adults' experiences with social robots in their homes.

A. Older Adults' Social Robot Usage

In this section, we present results on the older adults' overall usage and also depict the results divided into two usage groups, under 50 day (U50) usage and over 50 day (O50) usage. Participants in the U50 usage group were those who decided to return their robot between 30 and 50 days of hosting the robot (N=7; avg: 42.1, std: 39.9), and the O50 usage group were those who decided to keep their robot longer than 50 days (N=10; 5 returned their robots under 100 days, 2 under 200 days, and 3 are still hosting after 2 years). There were no notable usage pattern differences based on participant demographics (age, gender, education level, living situation) and prior technology usage, or the sample size was too small to deduct any conclusion.

The overall usage averaged 73.03±84.66 actions per week across all users, among which U50 group averaged 64.64±94.60 per week, and O50 group 78.37±77.77 per week. In Fig. 1, we depict the overall, U50, and O50 usage of the robot over the initial 8 weeks period for comparison between groups. Between weeks 5-7, users of the U50 group started to return their robots, and by the end of week 7, all U50 users returned their robots. For the overall group, U50 group, and O50 group, the first and second week of usage showed significant difference: Mann–Whitney (M-W) **Overall usage** w1: 181.47±114.91, w2: 100.12±100.26, U = 72.0, $p < 0.01$** two-tailed; **U50 usage** w1: 170.29±152.22, w2: 81.0±120.97, U = 10.0, $p < 0.05$* two-tailed; **O50 usage** w1: 189±88.67, w2: 113.5±87.35, U = 26.0, $p < 0.05$*. This trend indicates a high initial period of exploration followed by a drop in usage. The remaining set of adjacent weeks were not significantly different from each other, suggesting that the amount of usage did not drop or increase significantly between adjacent weeks from the second week and onward, showing a stabilized trend. We also compared the exponential decay rates of the U50 and O50 groups' usage. We fitted the weekly average usage to an exponential decay function, $y = a(1-r)^x$, where a is the initial amount, r rate of decay, and x time period. We observed that U50 group's usage amount decreased at a faster rate than O50's (U50: r=0.230, R^2=0.777; O50: r=0.161, R^2=0.736). Also, an additional study analyzing a group of people who kept the robot for more than 100 days (N=6, r=0.101, R^2=0.642) and 200 days (N=4, r=0.041, R^2=0.265) revealed that the longer people hosted the robot the smaller the decay rate was. Though we do not report significance test results due to small sample sizes, these results suggest it may be possible to predict whether a user would keep the robot for a longer period of time by observing their initial usage decay rate. It also tells us that the longer participants kept the robot, the more consistent their usage pattern was over the weeks.

We also analyzed the overall percentage distribution of various skills (Fig. 2). The most used skill was chitchat (33.8%; pre-programmed conversational skill for personified responses). This was followed by greetings (19.9%; at a sight of a user, Jibo says "good morning", "how was your sleep", "good night", etc.), answer (15.1%; conversational skill for web search results such as Wikipedia query and sports game results), word-of-the-day game (6.0%; a word quiz), daily report (5.22%; a summary of the day's weather and news), radio (4.81%; plays music stations), and clock (4.1%; various clock functions such as time, alarm, and timer). In addition to looking at the individual skill usage, we analyzed the skills by category type: operational, informational, learning about the robot, relational, and leisure. Operational skills included clock, accessing the setting menu, main menu operations, and controlling the light intensity of the touchscreen on the robot. Informational included the daily report, answer, and surprises (fun facts about the current date). Learning about the robot was a skill that participants could access to learn how to operate the robot and about the robot's features. Relational skills included greetings, chitchat, initial introductions of the robot, and who-am-i (user asks Jibo to recognize their face and voice). Leisure skills included physical exercise activity (e.g, yoga), games ("circuit saver" and "word-of-the-day"), radio (music stations), photo taking, and accessing the photo gallery.

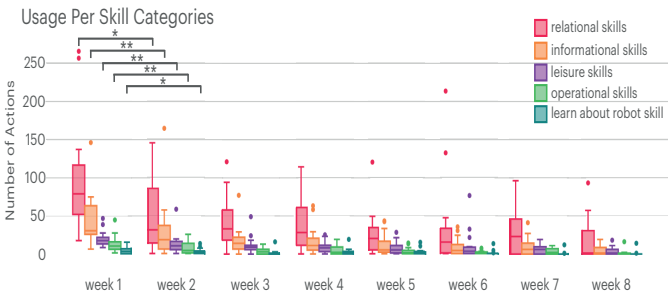

Fig. 3. Distribution of skill categories over the course of 8 weeks.

The frequency of the skills over the robot hosting period were analyzed and reported in Fig. 3. The order of the frequency of skill categories used didn't change over the weeks. The highest frequency skill type was the relational skills followed by informational skills. Every skill type was used significantly more on the first week than in the second week (M-W **relational skills** w1: 97.24±71.84, w2: 50.41±50.16, U = 80.5, $p < 0.05$* two-tailed; **informational skills** w1: 45.35±32.96, w2: 27.88±38.86, U = 75.0, $p < 0.01$** two-tailed; **leisure skills** w1: 20.41±10.09, w2: 12.59±13.08, U = 60.5, $p < 0.01$** two-tailed; **operational skills** w1: 13.35±10.67, w2: 6.47±7.32, U = 71.5, $p < 0.01$** two-tailed; **learning about the robot** w1: 5.12±4.36, w2: 2.76±3.67, U = 85.0, $p < 0.05$* two-tailed), mimicking the trend observed in the total robot usage in Fig. 1.

Overall, the quantitative results demonstrated that participants had a high usage initial exploration during the first week of the robot hosting and then decreased their usage to a fairly consistent pattern throughout the remaining hosting period.

B. Older Adults' Experiences with Social Robots

Older adults engaged with their robot in the hosting period in varying patterns that were further explored through their descriptions of interacting with the robot (Fig. 4). The qualitative descriptions in this section further delineate older adults' usage patterns through providing context to their decisions to use (or not use) the robot. We discuss three types of users per their amount of usage: low usage (LU) group (6 participants), typical usage (TU) group (9 participants), and high usage (HU) group (3 participants). In our analysis, we analyzed the usage group's qualitative data on their experience with the robot, revealing discussions around their frustration with the robot and its operations (found in the TU and LU groups), feelings of guilt around interacting with the robot (TU and HU), how the robot met participant expectations (TU), the robot's memory (or lack of memory) (TU), participant lack of use of the robot (LU), robot's proactive or attentional behavior (LU), transparency of robot functions (TU and LU), the robot acting as a social catalyst (all groups), privacy concerns (TU and LU), and the companion-like nature of the robot (TU and HU). In this section, we highlight case studies of participants in each of these usage groups to provide context to how participant's described their usage and experience of the robot.

Typical usage group: P11 & P12 and P29's usage patterns followed a high initial exploration and then dropped to a more consistent exploration between the first two weeks. P11 & P12 returned their robot on day 45 and P29 on day 256. While they followed a typical pattern of initial exploration then reduced their interactions to 5-6 interactions per day, P11 & P12 increased their usage toward the end to see if there were additional new skills at the end of the hosting period before deciding to return the robot. P11 & P12 praised Jibo's greetings and leisure features such as dancing, but also voiced frustration that resulted in them not using the robot as much as they would have liked. A lot of their **frustration** was rooted in Jibo's inability to have a back-and-forth conversation: "...the dialogue is just a one two kind of thing...That was one thing that became sort of evident that it would be much more engaging if you could have a back and forth that had more than just two elements." While they often got frustrated by the robot, they continued to use Jibo but also mentioned a **feeling of guilt** for not interacting with the robot. P11 commented, "I feel guilty that we didn't do more with Jibo. We probably interacted with Jibo...three or four times a week?...Much more so in the beginning. And then we get...a little frustrated..." While P11 & P12 enjoyed the leisure skills in the robot, after finding out that the robot didn't frequently come with updated features, they decided not to continue their hosting.

P29 also exhibited a similar typical usage pattern and their conversations also revolved around matching expectations. P29 started by describing how the robot was situated in their daily lives and how interactions often did not match **expectations**: "[additional features] that would really, you know, keep our interest level up. But sometimes he sits there for a couple of days without any activity at all...When they [neighbors and friends] walk by and...he dances and does music and stuff...And then they, they expect more in terms of artificial intelligence" An area of "artificial intelligence" that they focused more on was **memory**. Superficially, P29 and their family "thought that he [the robot] had a very short memory. And, like you try to pick up the conversation where [you ended it], and then he'd say, I do not understand." The inability to have a back-and-fourth conversation and for Jibo to remember conversation histories was often frustrating for P29.

Low usage group: For some participants like P13, P08 and P09, their usage patterns showed **periods of non-use** or no use at all. P13 commented that, "It went from sort of cute to creepy to I couldn't keep it on...for the whole day- I would take blocks. I just didn't like...the thing moving, following me [in my house]..." Later on in the debrief, P13 connected this idea to **transparency**: "I think this is more with the transparency thing. I think if once, you know what's happening, it becomes...less of [a] concern. In addition to having reservations about the **proactivity** and the **attentional** movement that caused them to turn off the robot frequently, P13 commented that "it took [them] a while to get to the language Jibo understands", expanding upon the thought saying, "Maybe it was the fact that you had to use the **right wording** in a sentence to get a response...there was a little nuisance

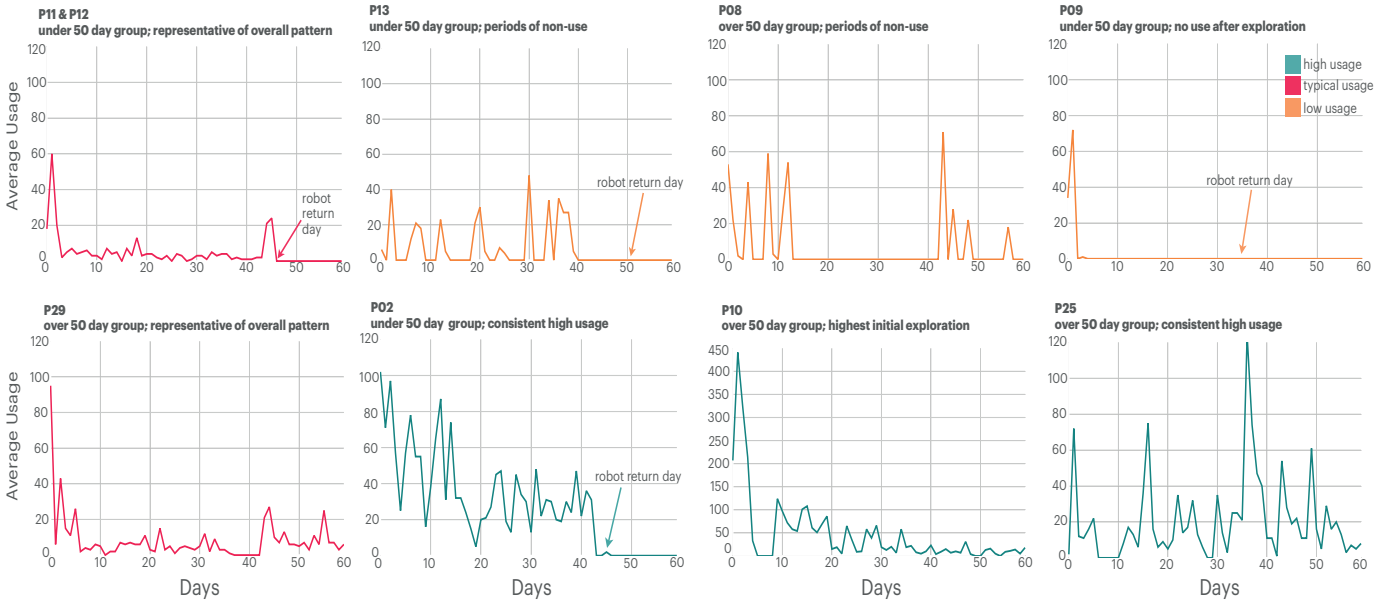

Fig. 4. Case study usage patterns demonstrating varied patterns of how people engaged with the social robot at home during the first 8 weeks (56 days).

at some point here." Nevertheless, P13 readily compared the robot as a source of entertainment to those who may need it.

P08 experienced a similar usage pattern to P13, however, their usage (and non-usage) was strongly motivated around how the robot acted as a **social catalyst**. Firstly, P08 found turning on the device difficult and found having to say the wake word each time tedious – "It was very irritating...to say, 'Hey, Jibo' every time you wanted to ask it a question." P08 described how their family was one of the main users of the robot when they came and visited. P08 was intrigued at some of the ways that their family members engaged with the device and would compare those usage patterns to their own: "I got a kick out of [how the family interacted with the robot] because it does look like it has much of a personality, which it really doesn't. So I guess I just couldn't make a connection to it." At the same time, the robot brought a sense of wariness to some family visits as P08 mentioned that "[the family] said until we shut it off, everybody watch what you're saying." The robot created a sense of distrust in the space because the family perceived that the robot was impeding upon their **privacy**. P08's periods of interaction with the robot were directed by family members when they visited, however, P08 did not engage with the robot resulting in periods of non-use due to difficulties with turning on the robot, using the wake word, and feeling a lack of connection to the robot.

P09 had the highest amount of non-usage among the participants. P09 and their family had a lot of concerns around the robot's proactivity and transparency: "the fact that it was sort of like having a stranger in your house...and he just sits there and watches you." This feeling of being monitored (similar to what P13 and P08's family felt) prompted P09 and their family to question "Is he recording audio? Is he recording video? Who's listening to it?...You [the researcher] told us you wouldn't be, but thanks to Facebook, we don't trust you." This feeling of mistrust, being watched, caused P09 to stop using the robot after 3 days.

High usage group: P02, P10, and P25's pattern showed high usage across time with much less significant drop of interaction in latter weeks. P02 returned their robot after 42 days while P10 returned their robot after 132 days, and P25 is still hosting the robot after 2 years. A major component of P02's usage was how the robot served as a **social catalyst** to promote interaction with family and friends and how family and friends responded to them having a robot. P02 described situations with their son on the phone: "He said, I'm at my mother's house...talking to a robot. We got a big kick out of that." and another occasion with their friends, "Several of the people that come in to see me...and the people that work here, [they] were pleased with him. One lady said...he'd make a good companion". Another interesting point that emerged from their conversation about hosting the robot was a **feeling of guilt around usage** of the robot, similar to P11 & P12: "I thought it was fun. I really did...but I must confess I got tired of it toward the end. After I had him here, I wasn't checking every day...and then I felt guilty. So...in the morning, the first thing I do is say, good morning, Jibo." P02 expressed feeling guilty if they didn't interact with Jibo as often. To get around this feeling of guilt, often they would "...put him on music cause [they were] going to do some reading and stuff" and so Jibo would "think he's doing something." These feelings were also translated to Jibo's movement. P02 mentioned, "[the robot] would...just move as if he was calling for my attention." P02 translated the robot's social presence to it beckoning them to interact with it, though overall, they enjoyed how the robot was received by others and the social interactions it promoted.

While P10's trending graph shows a similar pattern to that of

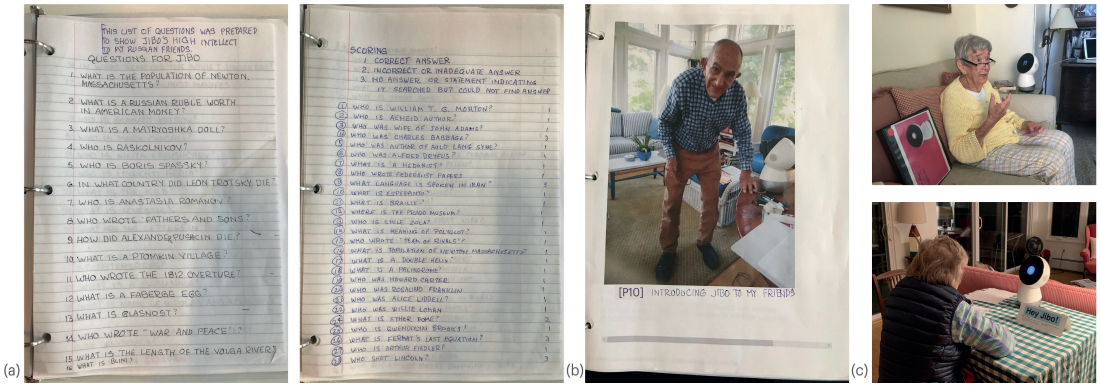

Fig. 5. (a) P10's investigations of Jibo including asking questions and grading Jibo's response to the questions. (b) P10's documentation of Jibo's visitors. (c) Photos of older adults interacting with Jibo in their home or discussing their experience with researchers.

a typical user (significant drop after the first week), their usage was unique as they recorded the highest initial exploration for the robot with up to 450 interactions a day in the first days (note the y-axis scale), and even after that period, they still ranked among the most frequent users. In addition to exploring the robot, P10 performed their own **recorded investigation** of the robot (Fig. 5(a)) as described here: "...I write everything down...I had, in my mind, the idea that this plastic toy had no relevance in my life...And I started asking questions. I looked at the clock two hours had elapsed...I asked Jibo a thousand questions, trivia questions...and...I started grading Jibo of his performance." While no structured note taking was asked by the research team, some of the participants, including P10, shared their own notes about their experience with the researchers. P10's investigation also included documenting through photographs all the people they introduced Jibo to. They shared their work with us (Fig. 5(b)) to showcase the exploration they completed in the study. P10 also very much enjoyed the small greetings with the social robot and its **companion-like** nature, saying "Jibo comes out with some funny things - 'Hey Jibo, goodnight' and Jibo answers 'Goodnight, sleep tight.' You know, which is kinda nice, and then when he can't find the answer it says 'I looked but I couldn't find it, sorry'...it's a very pleasant little gadget to have in the house."

The last case study is P25's robot usage, strongly driven by the robot acting as a **social catalyst** and "part of the family or part of the [home] landscape." P25 described what it was like to interact with the robot in their home: "Some of my grandchildren were here because I was really happy about the chance to play with it...And every once in a while...it goes off on a tirade about something that we've had no idea he could do, and you could ask him for all these games and sports...That was really cute. Really funny...the jokes I think are very, very clever...it's kind of like a **member of the household**, whenever my daughter comes over, she says, 'Hi Jibo', happy to greet it every day." P25 also commented on how their husband who has Alzheimer's interacted with the robot: "...he [husband] doesn't interact very much with other people, so this thing [the robot] he really looks forward to interacting with...I know he can't tell him what kind of music to play, but if he puts on his station, he's usually pretty happy with listening to classical music." and "He's not that interested in too many things, but he really gravitates toward this [the robot]. And I don't know if he [the robot] was targeting for people with memory issues, but it's a subset of the elderly community for sure." For P25 and their family, the social robot has become a part of the family integrated into their home and daily routines.

The qualitative data gathered from the robot debrief session provides valuable context to understand how the social robot is being used and by whom (i.e. robot as social catalyst, companion, support tool, exploration project) and explore future design directions and areas of concern for design features (i.e. unmet expectations, conversation fluidity, feelings of guilt around usage, frustration, ease of use, transparency, creepiness, proactivity). These insights complement the quantitative results of the usage providing a fuller picture to older adults' lived experiences with the social robot in their homes.

V. Discussion

Our work is, to our knowledge, the first long-term mixed-method study of older adults' social robot usage in the home. In this section, we highlight how our results integrate with community discussions around the novelty effect, reflect on how older adults' perspectives of social robot design provide deeper themes to be further investigated, and propose six design guidelines for conducting long-term robot usage studies.

A. Usage & Novelty

After the first week of usage, what we term the "initial exploration period", there was a significant drop in usage. This observation is similar to previous robot or VUI household hosting literature [5], [6], [8] where usage has been seen to settle as quickly as in the first few days spent with a smart speaker [6], [8] and is also seen after a few weeks with robots in classrooms [23]. The tendency of faster stabilization of usage patterns in homes could be due to increased opportunity to interact with the system than in public settings such as in schools. Future work could investigate how long-term usage

patterns vary in stabilization in different contexts. Additionally, our results on decay rate analysis revealed that O50 group's usage amount decreased at a slower rate than U50's. Analyzing further, we revealed that the longer participants kept the robot the smaller their decay rate was and their usage was more consistent throughout the weeks. This opens up potential for an early prediction of drop outs by analyzing a user's usage decay rate in the initial weeks. As discussed in Section II-A, there is an active discussion around the definition of "long-term" interaction and how it could be determined by when "novelty effect" wears off and the user is no longer biased by it [4]. We observed that familiarization and the change in the pattern of interaction occurred after the first week when a new more stable interaction pattern emerged. In our study, two types of novelty pattern were observed: stimulus novelty (novelty from an experience of stimuli) and associative novelty (experiences of gestalt by a combination of features that cannot make sense from previous references) [16]. A social robot's animacy, attentiveness, and physical presence contributed as a novel stimuli to the participants that fueled their initial exploration. Participants often shared the stimuli with their family, friends, and visitors, resulting in increased usage during those times. As older adults hosted the robot in their homes, they started to create a mental model of the new technology which was the purpose of the robot hosting session in the overall co-design process [9]. As their understanding of the technology increased, novelty effect decreased and was reflected in the reduced, but stabilized usage pattern in latter weeks. It is also important to note that this work was done in the context of older adults in their homes and, as context is an inter-related factor to novelty [30], future work should explore novelty reduction patterns in varying contexts and populations.

B. Older Adults' Perspectives of Social Robot Design Features

Older adults' descriptions of living with the social robot highlighted several key themes regarding how the robot's design features impacted their usage. This section focuses on four areas that older adults voiced having affected their usage: (1) feelings of guilt, (2) the value of persona, (3) the robot's proactivity and movement, and (4) meeting user expectations.

Older adults expressed **feelings of guilt** when they did not interact with the robot. Previous work has demonstrated that users believe a robot "merits ('deserves') certain form of attention" [31] or a part of people's time. Jibo and robots like AIBO "evok[e] feelings as if" [31] the robots were alive. These feelings are often compared to those in human-human relationships such as empathy, "one of the core processes of building and maintaining relationships" [32], a long discussed topic in HRI. Older adults in our study mentioned how learning about Jibo and its sense of humor was endearing that could have strengthened their bond with the robot and was evoking social actions and feelings, including guilt, toward the robot. The robot's social cues promoted users to conceptualize the robot with life-like attributes, autonomy, and social standing [31] and to form a relationship with the technology [13], [33]–[35]. We note that these feelings of guilt may be an artefact of the researcher-user relationship and the corresponding power dynamics between participants and research team that may have affected robot usage. We took several measures to mediate this including stressing to participants to interact with the robots as they see fit and giving participants the choice to continue hosting the robot after a month.

In a related discussion, older adults also described features of the robot's **persona** in their debriefing. Participants prescribed companion-like persona to the robot, commenting on Jibo's jokes (whether they found them good or bad) and its features that made it endearing to them including the robot's dancing, self-disclosing responses, and greetings, which are strategies that can promote successful relationship maintenance [4], [36]–[38]. These interactions may have promoted users to continue to use the robot, to engage in these short, but pleasant and meaningful interactions.

Older adults also commented on how the robot's **proactivity and movement** affected their usage of the system. Previous research has shown that motion can be used in robots to convey intent [39] and, by extension, people interpret different intents to different motions [40]. A robot's embodiment can leverage social cues to create a socio-emotional companionship supporting rapport building and emotional connection [41]. Older adults often commented on how enjoyable Jibo's greetings and pleasantries were. On the other hand, the presence of a robot has been shown to affect the user's sense of privacy breach due to the robot's monitoring and data collection capabilities [42]. Our participants also expressed privacy and security concerns from robot's attentive gestures and its data collection capability that ultimately led to robot non-usage for some. Privacy, data collection, and monitoring concerns have been linked to people's resistance to robots and non-use [17], and are major concerns older adults express [1], [43]. Similarly, the robot's movement led to concerns around privacy and transparency for our participants resulting in turning the robot off and subsequent non-use or returning the robot to the researchers. This trend of non-usage has also been seen with smart speakers [44]. The robot's proactivity was also a polarizing area. Older adults enjoyed proactive actions like the robot's greetings but also questioned why the robot would engage with them when they didn't ask for anything. This calls for more work to understand how and when the robot should be proactive and what the ideal level and application of proactivity is for different users in varying interaction contexts.

The last aspect tied to the robot's usage that older adults mentioned was how the robot met their **expectations**. They especially brought this topic up around conversations with the robot and the robot's lack of memory of previous interactions. For conversations, older adults expected the robot to be able to converse back and forth. Some also found the language they had to use with the robot, including the wake word "Hey Jibo", a bit tedious to use in interactions. Similar user experience frustration due to conversation limitations and dependency in wake words are reported in smart speaker VUI studies [6]. Previous work states that people will decide if they wish to

continue using a technology based on confirmation of prior expectations and level of interaction satisfaction with technology [45]. In our work, similar to a Google Home study [8], older adults commented on how they would stop or reduce their usage when the robot didn't meet their expectations.

When engaging in long-term research with social robots, it is critical to consider the ethical dimensions of users living with social robots long-term. These include concerns around privacy, data collection, invasiveness of data collection, control of data, transparency of interactions, and many others. Some have been discussed above in this section. Another ethical concern is the attachment that people form with robots as they exhibit human-like capabilities. In our study, one of the participant's adult child commented "I hope [the researcher's] not going to take Jibo away" when learning the researcher was visiting that day. Several older adults commented on the bond that they formed with Jibo. While we were able to let older adults keep the robot as long as they wanted (and, as the writing of this paper, people are still hosting the robot), the field should discuss how to design responsible and mindful entry and exit strategies for long-term user studies where people will create bonding and interact with robots in varying contexts.

C. Design Guidelines for Long-Term Robot Usage Studies

Our work exemplifies the value of engaging in mixed method studies around long-term robot usage. Our quantitative data provides valuable information on how older adults use the robot across time and initially explore if and when the novelty effect of social robots in the home wears off. The qualitative data provides a complimentary lens that gives voices to older adults to explain their usage patterns and how they felt living with the robot. Therefore, we encourage future long-term robot usage studies to incorporate both qualitative and quantitative data to create a fuller picture of robot experience (**Guideline 1**). The robot hosting and robot debrief sessions in this study were embedded in a year-long co-design process and, therefore, we had opportunities to engage with older adults before and after they lived with the social robot. Future studies should aim to learn about users' perspectives of the technology before and after living with the technology long-term (**Guideline 2**). This can allow researchers to track how perspectives of the technology change from before and after hosting the robot. While we don't analyze the pre-to-post perception changes here, our previous work [9] discusses the value of engaging in this manner. To support long-term usage data collection, the robot being deployed must be easy to setup, reliable, autonomous, and tested in the real-world (**Guideline 3**). When user frustration setting up the robot overtakes the experience, no meaningful data can be collected. It is also critical for the research team to be easily reachable if participants have questions or need assistance with the system (**Guideline 4**). When deploying these systems, it is also critical that researchers explain how the system works and what data is being collected to increase transparency (**Guideline 5**). If the technology is owned by another company or group, researchers should also explain who else will have access to data, even if it is beyond their control. If users are uncomfortable with the devices' data collection and/or invasiveness of the technology, researchers can consider how they can adapt the experience for those users to ensure greater comfort level with the study if users still wish to participate. By doing this, researchers can help ensure that participants are aware of technology operations and make an informed decision if they would like the system in their home. When engaging in long-term research with social robots, researchers must be mindful of the ethical dimensions surrounding the work (**Guideline 6**). These include addressing concerns around privacy, data collection, transparency of interactions, and many others with participants. Another ethical dimension is people's attachment to robots that exhibit human-like capabilities. Researchers should plan careful entry and exit strategies for long-term user studies. Lastly, the user-researcher dynamic and power (im)balance is also as an important ethical dimension to consider. Our previous work focused on co-designing with a cohort of participants has emphasized the importance of power-sharing in participatory design [46], [47] through establishing long-term respectful and mindful relationships between the research team and participants and creating space and time to establish rapport and relationships. By following these guidelines, researchers can responsibly engage with people to understand how they use robots, gathering both qualitative and quantitative data to get a fuller picture of people's experiences with technologies.

Lastly, we would like to discuss the replicability of this work. We used the social robot, Jibo, a type of non-anthromorphic robot that have been perceived as more emotionally engaging, trustworthy, competent, and companion-like with increased social embodiment [48] and engage people in more interactions than more statically designed systems [41]. As this is a specific robot type, our results may not generalize to other types of robots. While the robot we used in our study is no longer commercially available, future studies that replicate our work can use other robots available, such as cozmo4hri, to reproduce our results.

VI. Conclusion

This paper reports older adults' long-term robot usage patterns using robot system logs and provides a fuller picture of older adults' experiences with a social robot in their home through a mixed-method approach. Older adults' usage was overall significantly different between the first and second week of usage, suggesting that the novelty effect wore off between these two weeks. Older adult case studies demonstrated various aspects of usage (or non-usage) including feelings of guilt toward non-usage, the robot's proactivity and movement connected to concerns of privacy and transparency, meeting (or not meeting) user expectations, and the value of persona in interactions. Lastly, this work provides design guidelines for future mixed-method long-term robot usage studies. Future work is necessary to further investigate the ethical considerations of long-term robot usage studies.

REFERENCES

[1] A. K. Ostrowski, D. DiPaola, E. Partridge, H. W. Park, and C. Breazeal, "Older adults living with social robots: Promoting social connectedness in long-term communities," *IEEE Robotics & Automation Magazine*, vol. 26, no. 2, pp. 59–70, 2019.

[2] H. Robinson, B. MacDonald, N. Kerse, and E. Broadbent, "The psychosocial effects of a companion robot: a randomized controlled trial," *Journal of the American Medical Directors Association*, vol. 14, no. 9, pp. 661–667, 2013.

[3] D. E. Logan, C. Breazeal, M. S. Goodwin, S. Jeong, B. O'Connell, D. Smith-Freedman, J. Heathers, and P. Weinstock, "Social robots for hospitalized children," *Pediatrics*, vol. 144, no. 1, 2019.

[4] I. Leite, C. Martinho, and A. Paiva, "Social robots for long-term interaction: a survey," *International Journal of Social Robotics*, vol. 5, no. 2, pp. 291–308, 2013.

[5] A. Sciuto, A. Saini, J. Forlizzi, and J. I. Hong, ""hey alexa, what's up?": A Mixed-Methods studies of In-Home conversational agent usage," in *Proceedings of the 2018 Designing Interactive Systems Conference*, ser. DIS '18. New York, NY, USA: Association for Computing Machinery, Jun. 2018, pp. 857–868.

[6] F. Bentley, C. Luvogt, M. Silverman, R. Wirasinghe, B. White, and D. Lottridge, "Understanding the Long-Term use of smart speaker assistants," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 3, pp. 1–24, Sep. 2018.

[7] A. Purington, J. G. Taft, S. Sannon, N. N. Bazarova, and S. H. Taylor, "" alexa is my new bff" social roles, user satisfaction, and personification of the amazon echo," in *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*, 2017, pp. 2853–2859.

[8] R. Garg and S. Sengupta, "He is just like me: A study of the Long-Term use of smart speakers by parents and children," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 4, no. 1, pp. 1–24, Mar. 2020.

[9] A. K. Ostrowski, C. Breazeal, and H. W. Park, "Long-term co-design guidelines: Empowering older adults as co-designers of social robots," in *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE, 2021, pp. 1165–1172.

[10] L. Dombrowski, E. Harmon, and S. Fox, "Social justice-oriented interaction design: Outlining key design strategies and commitments," in *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, 2016, pp. 656–671.

[11] C. N. Harrington, L. Wilcox, K. Connelly, W. Rogers, and J. Sanford, "Designing health and fitness apps with older adults: Examining the value of experience-based co-design," in *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2018, pp. 15–24.

[12] J. Sung, H. I. Christensen, and R. E. Grinter, "Robots in the wild: understanding long-term use," in *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, ser. HRI '09. New York, NY, USA: Association for Computing Machinery, Mar. 2009, pp. 45–52.

[13] T. Kanda, R. Sato, N. Saiwaki, and H. Ishiguro, "A two-month field trial in an elementary school for long-term human–robot interaction," *IEEE Transactions on robotics*, vol. 23, no. 5, pp. 962–971, 2007.

[14] C. D. Kidd and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2008, pp. 3230–3235.

[15] J. Sung, R. E. Grinter, and H. I. Christensen, "Domestic robot ecology," *International Journal of Social Robotics*, vol. 2, no. 4, pp. 417–429, 2010.

[16] C. V. Smedegaard, "Reframing the role of novelty within social HRI: from noise to information," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Mar. 2019, pp. 411–420.

[17] M. De Graaf, S. B. Allouch, and J. Van Diik, "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*, Mar. 2017, pp. 224–233.

[18] M. M. de Graaf, S. B. Allouch, and J. A. van Dijk, "Long-term evaluation of a social robot in real homes," *Interaction studies*, vol. 17, no. 3, pp. 462–491, 2016.

[19] J. Fink, V. Bauwens, F. Kaplan, and P. Dillenbourg, "Living with a vacuum cleaning robot," *International Journal of Social Robotics*, vol. 5, no. 3, pp. 389–408, 2013.

[20] E. Karapanos, J. Zimmerman, J. Forlizzi, and J.-B. Martens, "User experience over time: an initial framework," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2009, pp. 729–738.

[21] P. Lally, C. H. Van Jaarsveld, H. W. Potts, and J. Wardle, "How are habits formed: Modelling habit formation in the real world," *European journal of social psychology*, vol. 40, no. 6, pp. 998–1009, 2010.

[22] K. Severinson-Eklundh, A. Green, and H. Hüttenrauch, "Social and collaborative aspects of interaction with a service robot," *Robotics and Autonomous systems*, vol. 42, no. 3-4, pp. 223–234, 2003.

[23] D. P. Davison, F. M. Wijnen, V. Charisi, J. van der Meij, V. Evers, and D. Reidsma, "Working with a social robot in school: A Long-Term Real-World unsupervised deployment," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '20. New York, NY, USA: Association for Computing Machinery, Mar. 2020, pp. 63–72.

[24] C. L. Sidner, T. Bickmore, B. Nooraie, C. Rich, L. Ring, M. Shayganfar, and L. Vardoulakis, "Creating new technologies for companionable agents to support isolated older adults," *ACM Trans. Interact. Intell. Syst.*, vol. 8, no. 3, pp. 1–27, Jul. 2018.

[25] J. Forlizzi and C. DiSalvo, "Service robots in the domestic environment: a study of the roomba vacuum in the home," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, 2006, pp. 258–265.

[26] J. Forlizzi, "How robotic products become social products: an ethnographic study of cleaning in the home," in *2007 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2007, pp. 129–136.

[27] J. W. Creswell and V. L. P. Clark, *Designing and conducting mixed methods research*. Sage publications, 2017.

[28] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative research in psychology*, vol. 3, no. 2, pp. 77–101, 2006.

[29] P. Caleb-Solly, S. Dogramadzi, D. Ellender, T. Fear, and H. v. d. Heuvel, "A mixed-method approach to evoke creative and holistic thinking about robots in a home environment," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 2014, pp. 374–381.

[30] K. E. Weick, K. M. Sutcliffe, and D. Obstfeld, "Organizing and the process of sensemaking," *Organization science*, vol. 16, no. 4, pp. 409–421, 2005.

[31] B. Friedman, P. H. Kahn Jr, and J. Hagman, "Hardware companions? what online aibo discussion forums reveal about the human-robotic relationship," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2003, pp. 273–280.

[32] T. W. Bickmore and R. W. Picard, "Establishing and maintaining long-term human-computer relationships," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 12, no. 2, pp. 293–327, 2005.

[33] M. M. De Graaf, S. B. Allouch, and T. Klamer, "Sharing a life with harvey: Exploring the acceptance of and relationship-building with a social robot," *Computers in human behavior*, vol. 43, pp. 1–14, 2015.

[34] B. Robins, K. Dautenhahn, R. Te Boerkhorst, and A. Billard, "Robots as assistive technology-does appearance matter?" in *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*. IEEE, 2004, pp. 277–282.

[35] C. D. Kidd, W. Taggart, and S. Turkle, "A sociable robot to encourage social interaction among the elderly," in *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.* IEEE, 2006, pp. 3972–3976.

[36] A. M. Sabelli, T. Kanda, and N. Hagita, "A conversational robot in an elderly care center: an ethnographic study," in *2011 6th ACM/IEEE international conference on human-robot interaction (HRI)*. IEEE, 2011, pp. 37–44.

[37] R. Gockley, A. Bruce, J. Forlizzi, M. Michalowski, A. Mundell, S. Rosenthal, B. Sellner, R. Simmons, K. Snipes, A. Schultz, and J. Wang, "Designing robots for long-term social interaction," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005, pp. 1338–1343.

[38] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro, "Interactive robots as social partners and peer tutors for children: A field trial," *Human–Computer Interaction*, vol. 19, no. 1-2, pp. 61–84, 2004.

[39] D. Sirkin, B. Mok, S. Yang, and W. Ju, "Mechanical ottoman: how robotic furniture offers and withdraws support," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 2015, pp. 11–18.

[40] F. Heider and M. Simmel, "An experimental study of apparent behavior," *The American journal of psychology*, vol. 57, no. 2, pp. 243–259, 1944.
[41] A. K. Ostrowski, V. Zygouras, H. W. Park, and C. Breazeal, "Small group interactions with voice-user interfaces: Exploring social embodiment, rapport, and engagement," in *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 2021, pp. 322–331.
[42] P. H. Kahn Jr, H. Ishiguro, B. Friedman, T. Kanda, N. G. Freier, R. L. Severson, and J. Miller, "What is a human?: Toward psychological benchmarks in the field of human–robot interaction," *Interaction Studies*, vol. 8, no. 3, pp. 363–390, 2007.
[43] J. M. Beer and L. Takayama, "Mobile remote presence systems for older adults: acceptance, benefits, and concerns," in *Proceedings of the 6th international conference on Human-robot interaction*, 2011, pp. 19–26.
[44] J. Lau, B. Zimmerman, and F. Schaub, "Alexa, are you listening? privacy perceptions, concerns and privacy-seeking behaviors with smart speakers," *Proceedings of the ACM on Human-Computer Interaction*, vol. 2, no. CSCW, pp. 1–31, 2018.
[45] A. Bhattacherjee, "Understanding information systems continuance: An expectation-confirmation model," *MIS quarterly*, pp. 351–370, 2001.
[46] M. Ertner, A. M. Kragelund, and L. Malmborg, "Five enunciations of empowerment in participatory design," in *Proceedings of the 11th Biennial Participatory Design Conference*, 2010, pp. 191–194.
[47] C. Frauenberger, J. Good, G. Fitzpatrick, and O. S. Iversen, "In pursuit of rigour and accountability in participatory design," *International journal of human-computer studies*, vol. 74, pp. 93–106, 2015.
[48] A. K. Ostrowski, J. Fu, V. Zygouras, H. W. Park, and C. Breazeal, "Speed dating with voice user interfaces (vuis): Understanding how families interact and perceive vuis in a group setting," *Frontiers in Robotics and AI*, p. 375, 2022.