Robots in the Wild: Understanding Long-term Use

Ja-Young Sung, Henrik I. Christensen, Rebecca E. Grinter

School of Interactive Computing
Georgia Institute of Technology
85 5th Street, Atlanta GA 30332 USA

{jsung, hic, beki}@cc.gatech.edu

ABSTRACT

It has long been recognized that novelty effects exist in the interaction with technologies. Despite this recognition, we still know little about the novelty effects associated with domestic robotic appliances and more importantly, what occurs after the novelty wears off. To address this gap, we undertook a longitudinal field study with 30 households to which we gave Roomba vacuuming robots and then observed use over six months. During this study, which spans over 149 home visits, we encountered methodological challenges in understanding households' usage patterns. In this paper we report on our longitudinal research, focusing particularly on the methods that we used 1) to understand human-robot interaction over time despite the constraints of privacy and temporality in the home, and 2) to uncover information when routines became less conscious to the participants themselves.

Categories and Subject Descriptors

K.4.2 [**Computers and Society**]: *Social Issues*

General Terms

Design, Experiment

Keywords

Domestic Robot, User Study, Longitudinal Field Research

1. INTRODUCTION

It has long been recognized that novelty effects exist in the use of technology [18,21]. Novelty effects are the first responses to a technology, not the patterns of usage that will persist over time as the product ceases to be new. Despite the recognition of this empirical phenomenon, we still know little about novelty effects associated with domestic robotic appliances. Specifically, what happens to human-robot interaction when the novelty affect wears off? Studying the usage patterns beyond these novelty effects is crucial because it deepens our insights about what truly occurs when a robot becomes a part of people's everyday lives and therefore, can inform development of products that will remain useful beyond initial adoption [10].

To explore long-term adoption, many researchers have conducted long-term empirical studies of people's lived experiences in places such as offices [10,17], schools [11,24], and hospitals [15,16]. But a gap remains; there have been surprisingly few longitudinal studies of the use of robots in the home, or at least few studies where the results are publicly available (with exception of [5,13]). This is surprising considering the growing number of domestic robots and the desire to use them for purposes such as prolonging independent living at home. To address this gap, we undertook a longitudinal field study with 30 households who had never owned any kind of robotic appliance. We gave our participants Roomba™ vacuuming robots and then visited each household repeatedly over a six months period to better understand their evolving usage patterns. In total, the entire study spanned more than a year, from June 2007 to August 2008, and involved 149 household visits.

In this paper we report on our longitudinal research, focusing particularly on the *methods* we used. Conducting research in domestic environments raises empirical challenges stemming from the nature of domestic routines. Specifically, we learned that it was hard to observe and document (as explained in [3]). Our goal is to explain 1) how we did manage to capture evolving human-robot interaction despite the constraints of privacy and temporality in the home, and 2) how we uncovered information when routines of its usage became less conscious to the participants themselves. We begin this by reviewing the related work. Then, we discuss our study procedure in detail. Finally, we complete this paper by discussing our longitudinal research and the techniques it comprised. We conclude with a discussion of how longitudinal research can take place in domestic settings, and has an important place in HRI as we continue to design personal robotic appliances for real-world long-term use.

2. RELATED WORK

This section is divided into three parts. First, we discuss research that has observed human-robot interaction over longer periods of time in non-domestic settings. Second, we review domestic technology research that explains the unique characteristics of homes and how they require different study approaches from non-domestic places. Lastly, we describe domestic robotic research and argue how our study can contribute to this body of knowledge.

A growing body of HRI work reports findings from long-term observation of human-robot interaction in natural settings. Many report that novelty effects typically fade in over time. For example, Kanda et al. [11] deployed a robot for 9 weeks in a classroom and observed elementary school children's engagement. At first, children were excited and wanted to play

with Robovie, but over time the frequency of interaction and the number of playful interactions decreased. In their three month study, Tanaka et al. [24] tried robot dancing as a way to stimulate longer sustained interaction between the children and the robot, but they also saw the children's interest decreasing over time. Other studies of human-robot interaction in the wild also show that novelty effects that promote initial engagement typically wear off after a short period of time. For example, people begin to ignore a large robotic guide in a hallway after just days of interaction [17] and even forgot about the robot on a mission in a three-month-long field trial [10].

Collectively, these studies reveal intriguing novelty effects and show that interaction changes over time. Due to the time-length of these studies, much less is clear about how human-robot engagement changes over time (i.e., more than three months). And yet, some studies do hint at interesting interactions after the novelty effects have worn off. For example, Kanda et al. [11] report that the children voluntarily created a collective description of Robovie's personality towards the end of the 9 week study. Also, Forlizzi [5] reported that study participants had given names to their vacuuming robots.

Our longitudinal research builds on the evidence that long-term patterns are worthy of study (for instance [8,10]), and that getting past novelty effects is critical for designing robots that people will interact with at home over time. However, longitudinal research brings new challenges, particularly when conducted in the home setting. For example, previous field studies outside the home have made use of video recording [14,15], participant observation [16] and logging mechanisms pre-programmed into the robot [7,11]. However, the private nature of the home, the potential for cleaning to take place any time during the day, and the fact that Roomba is designed to be moved around rooms all make video recording or extensive on-site observation socially and logistically complex. These difficulties are shared by the HCI community, who have sometimes used smart homes equipped with sensor and camera networks or scaled down the number of participant homes (to one or two) that require extensive "wiring up" to capture information [12,19]. But moving people into non-natural settings or scaling back participants did not meet our study goal. Ultimately we decided to use a variety of qualitative methods including observation and interviews, some of which have been seen in previous studies of domestic robotics [4,5,13] and other domestic technologies [1,3].

However, in addition to using these techniques, we also wanted to address the following challenge of longitudinal research: how do you get at routines that have ceased to be conscious in the minds of participants after novelty effect fades out? Research suggests that successful technology adoption involves ceasing to think about the item as a technological artifact, but seeing it as a part of a household routine [25]. Researchers faced with wanting to understand human engagement with these invisible-in-use objects have developed methods to answer questions about use. For example, Cultural Probes surface routines in the home through giving participants postcards and disposable cameras and encouraging them document their own activities [6]. The Generative Toolkit was also designed to understand tacit needs and desires that are hard to express in words or observe [20]. In our study, we incorporated these approaches to uncover users' routine interactions with Roomba because we found that participants were struggling to tell us about their vacuuming habits particularly towards the end of our study. In the following section, we provide detailed accounts of how we conducted our longitudinal study.

3. STRUCTURE OF LONG-TERM STUDY

We recruited 30 households in the Metro-Atlanta area to participate in our study. Each household participated for six months and we visited them five times during the study period, except for one house that dropped out after fourth interview and did not want us to return. In total, we conducted 149 home visits.

We recruited households using various methods, such as word of mouth, mailing lists, Craig's list (craigslist.com) and snowball sampling which involves asking study participants for referrals to other people who might participate. During recruitment, we tried to balance out the households' demographic profiles seeking diversity. For example, we wanted a balance in presence of children, income level, age range and technical knowledge. We characterized technical knowledge as having technological education, employment and/or hobbies, trying to understand whether this background influenced robotic use. We limited the participating households 1) to have at least one adult member (18-year-old and above) and 2) not to have plans to move in the six months since prior research shows that room layouts can impact usage patterns [23]. We compensated our participants by allowing them keep their Roomba after study completion.

We had 48 participants (22 men and 26 women, mean age=42) across the 30 households. More specifically, we had 17 households with married couples, of them, 13 households had at least one child living with them. For the 13 single households, two of them had children while the rest lived alone or with other adults. Whether single of married, we had 15 families who in total had 23 children (mean age-=9). Our participants were skewed towards those with higher-levels of education: 26 participants with graduate degrees, 20 with college degrees and 2 with high school degrees. Further, we consider 19 of our participants technical. We defined a person as being technical if they had received professional or academic training, or self-reported technologically-oriented hobbies, such as hacking. None had robotic appliances (i.e., Scooba) as it was one of our screening conditions. Given a focus on cleaning, we also recruited families with pets (n=16), and families with cleaning services (n=7) to compare the experience across these different households. Also, knowing that physical settings impact the Roomba usage [5,23], we recruited households who dwelled in various types of buildings, including multi-story houses, lofts and apartments. Our houses also varied in floor types, including hardwood, linoleum, tile, carpet and stained concrete. This variety in floor types sometimes revealed information about our participants. For instance, two households (P5 and P20) had no carpeting due to severe allergies that family members had and associated with extreme sensitivity to dust.

Next, we describe our study procedures and techniques we used in detail. To give an overview, we visited each house five times:

1. A*pproximately one week prior* to giving out the Roomba, we assessed their existing cleaning routines. This served as a baseline to compare Roomba usage against.

2. We brought a Roomba, and ask families to unpack it and use it in our presence. We watched their initial reaction.
3. T*wo weeks later* we returned to see what had changed in terms of how the household used Roomba, and cleaned.
4. T*wo months later* we visited again to see what has changed with Roomba usage.
5. A*pproximately six months* after the arrival of Roomba, we completed the study by asking for the householders' overall impressions of robot.

3.1 First Interview: One Week Before

Our first visit to each home took place *a week prior* to delivering Roomba. This visit had two objectives. First, we wanted to understand their attitude toward robots since it has been shown that expectations can influence how robots are used [5,9]. Second, we wanted to document the overall domestic environment including home layout and each householder's domestic responsibilities, including floor cleaning. We sought this holistic understanding of the home—because previous research [22,23] shows that the use of Roomba may cause physical changes in the home (i.e., Roombarization) and also in household divisions of labor associated with floor cleaning.

To achieve this understanding we broke the interview down into two parts: expectations of robotics and domestic routines. We began by focusing on people's expectations of robots. First, we asked them to provide a layman's description of a generic robot. Next we asked them to describe their familiarity with robotics since we expected that to impact their usage. To do this, we presented participants with images of 14 robots from the robot hall of fame (www.robothalloffame.org) and asked them to rate their knowledge of them on a seven point Likert scale. Finally, we asked participants to provide us with a visual depiction (Figure 1) of their ideal home robots. To help them with this activity, we provided magazines, paper, pens, scissors and glue. Following Sanders [20], this generative exercise helps participants articulate thoughts that are unusual or difficult to express as we thought thinking about robotics and ideation of one might be difficult. As shown in Figure 2, P17 was not only able to convey desirable functions (doing house chores) but also aesthetic (futuristic metal material) and interactive qualities (LCD communications) of their ideal domestic robot. An additional advantage of this exercise was that while we watched and listened, we learned about their individual contexts. For example, P4 wanted a security robot because she was a young single woman living in an unsafe neighborhood.

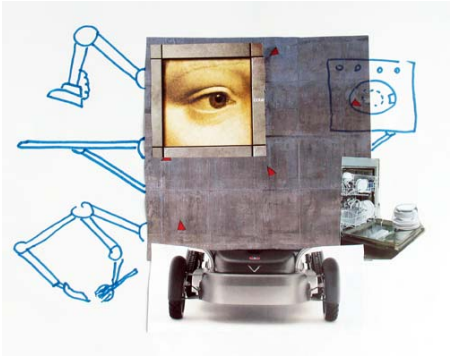

Figure 1. P17's ideal domestic robot

The second half of this interview focused on participants' domestic routines with an emphasis on cleaning practices. We began by asking how the household divided cleaning tasks (including children), how often they cleaned, whether they had a regular time, and what tools they used. At the end, we asked participants to draw a home blueprint (Figure 2). They included floor type and highlighted areas that they would run Roomba (yellow and largely colored area on Figure 3). We took the blueprint back to later interviews to ask the household to highlight again the areas that they had actually run Roomba (green and spotted area on Figure 2 at the final interview), thus allowing us to capture the difference between intended and actual usage. As shown in Figure 2, P20 clearly used Roomba as a spot cleaning tool despite initially thinking it would largely replace household vacuuming.

Figure 2. Blueprint of P20's Home

The first interview concluded with going on a home tour, where participants showed us the places that they intended to run Roomba. We took pictures of rooms and floors as a baseline from which to assess changes. In some houses we were easily able to see visual differences in rooms and floors over time. One household was not comfortable allowing us to see their upstairs bedrooms or offices, despite it being the place where they used Roomba the most. The blueprint in that household became essential for discussing how they used the robotic vacuum cleaner.

3.2 Second Interview: Roomba Introduction

Approximately one week later, we brought a Roomba to each household. For this visit, we asked everyone in the household to be present, including children and pets, because research shows that how a robot is introduced influences its usage [5]. Consequently, we wanted all householders to be introduced to Roomba at the same time and interact with each other in the process of exploring the new technology. To recreate the experience of making the purchase, we asked householders to examine the box as they might have done if they had bought it at a store. We also asked open-ended questions about their knowledge of Roomba, and asked them to rate their expectations about this product on a seven point Likert scale. The categories of Likert scale included, intelligence, ease of use, usefulness, emotional attachment, entertainment value, and degree of impact on household. Next, we asked participants to unpack the box and do what they would do if they had purchased it. We wanted to observe their interactions and listen to their dialog, thus we did not offer any advice on how to do this. After unpacking, most households voluntarily ran Roomba between five and 15 minutes. We asked the few households who did not run the device

unprompted, to turn it on so that we could see the participants' very first reaction to Roomba. After this brief first impression of Roomba's performance, we conducted a post-running debrief session about how it differed from their expectations. During this time, we asked them to complete again the seven point Likert scale questions that we had asked them earlier to see whether this actual use had changed their impression.

3.3 Third Interview: Two Weeks Later

Our third visit took place approximately two weeks after the deployment while the relationship was still fresh. To capture the evolution of their initial adoption, we had originally planned that participants would self-log using a disposable camera and/or a scrapbook. However, after four households we only had one person log Roomba experiences and that was in their own, not our provided diary. Others simply forgot, as they explained it was hard to remember to take a camera to every cleaning event despite Roomba's novelty. We abandoned this approach and instead asked people to send us email if anything unusual happened including accidents and memorable events. Seven participants responded by sending us emails and photos voluntarily talking about their Roomba experience as follows:

In case you are interested, we (I) have changed the name of the Roomba from Andreas to Ricky (as in Ricardo or Martin). He just seems more like a Ricky to me… I just thought you should know. — P9, day after Roomba delivery.

P9 sent us two additional emails about Roomba related activities, and three other robotics related news articles reflecting her growing interest in domestic robots. Some (like P4) did not send us email but took photos of Roomba documenting moments of fun and trouble, and then walked us through scenes captured in these images during our visits. Although we saw this type of engagement more during the earlier (novelty) phases of the study, some participants continued to send us email, photos, and videos throughout the entire study period to keep us up-to-date.

Receiving multiple emails prior to a visit led us to expect that participants would have much to discuss when we visited. Surprisingly this was not always the case. So, we turned to a probing technique to help them talk about Roomba called Bubble Drawing [2]. It is designed to help facilitate discussion about subjects that are difficult to talk about. Although cleaning is not a subject that is socially awkward, it is something that people do not typically discuss, especially with strangers. We gave participants Figure 3 with empty bubbles and asked them to fill in while discussing the pros and cons of their Roomba experience.

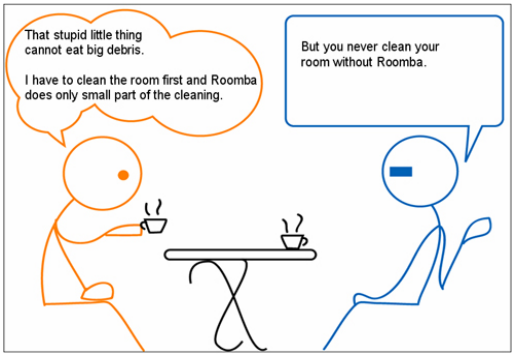

Figure 3. Bubble Drawing Technique (P9)

Participants created between one and four pages of conversation, which they explained to us during the visit. We found this a particularly good way to discuss negative experiences, because participants often talked about positive attributes —perhaps associating with Roomba with us. Next we asked questions about their "Roomba routine" such as how often it was used, who used it, and how, when and where they cleaned. We repeated these questions at the fourth and fifth interviews to continue to track stability and change in usage.

Additionally, we had participants do the following three activities to track their usage: highlighting their blueprints again to show us where Roomba had cleaned since the previous visit, ranking their perceptions of Roomba on the Likert scale, and finally checking the activities they did with Roomba off on an activity card (Figure 4). The activity card listed all the things that we heard about people doing with Roomba (e.g., hacking, naming, dressing, as well as cleaning) based on our own and others' previous research [6,7,28,29]. We asked them to check activities they had done. Interestingly, the number of activities was a good predictor of when households are in and move beyond the initial novelty effects. For example, we saw most of the households checking "watch it for fun" at two-week visit, which we barely saw it checked at the final interview. Also, we found the number of activities to be a helpful indication of whether our participants were adopting or rejecting Roomba. For example, P29 checked they had done 10 activities out of 11 that were on the list at the two-week visit. However, by the time of the last interview, they checked nothing—including using it for cleaning.

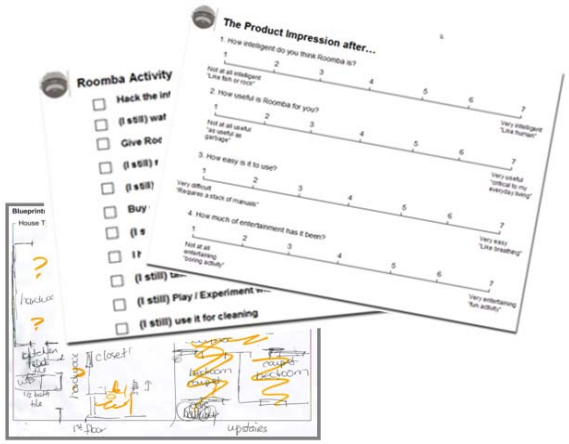

Figure 4. (Bottom to Top) Home BluePrint, Roomba Activity List, and Roomba Impression Likert Scale

3.4 Fourth Interview: Two Months Later

We decided to return for the fourth visit after two months of usage because prior research suggests that routinized (i.e., post novelty effects) human-robot interaction occurs after this time [11]. The interview format was similar to the third interview so that we could compare and see how adoption had evolved beyond novelty effects. We opened the interview by asking about Roomba routines and any special events or/and technical difficulties. Next, participants repeated the three exercises of blueprint, Likert scale and activity list in Figure 4.

One thing we noticed during this interview was that our participants found it harder to discuss their routines. The longer gap between interviews may have contributed to this difficulty as

it is easier to speak about what you did in two weeks than discuss activities taking place over two months. However, our participants also told us explicitly that they were having trouble remembering their Roomba usage because they thought less of how they used it on a daily basis than compared with at the beginning of the study. For example, P11 reported feeling that Roomba was such a natural part of the household that she did not really keep track of how she used it just as she did not think of how she used refrigerator or microwave. This experience highlights the complexity of documenting routines beyond novelty effects as they disappear into the lived experience of the home. For this reason in our next interview, we decided to use more generative activities that involved more than observing and interviewing, such as getting participants to create Bubble Drawings or checking the activity list again. We found these generative techniques helped our participants reflect on the activities that are difficult to express in words by allowing sufficient time to think and even discussing with other members of the household which we discuss further in section 4.3.

3.5 Fifth Interview: Six Months Later

The fifth and final interview occurred approximately six months after they began using Roomba. To maintain consistency with the previous interviews, we began with the Roomba routine questions followed by blueprint, ratings and activity list techniques. Furthermore, in our last interview, we focused on capturing Roomba routines in as much detail as possible, and therefore we incorporated two other generative activities that involved more than interviewing and observing.

First, we wanted to understand who the primary Roomba users were and how they maintained the robot. Previously, we had difficulty in identifying householders who used and maintained the robot because some participants felt guilty with not using and maintaining the robot. For example, P20 had only run Roomba a couple of times over the last two months. Upon asking why, the household told us they assessed it unfairly (negatively) and that they should have used it more often than they did. They spoke of what they should have done instead of answering our question. Also, householders who used Roomba less, tended to report that they had used it but did not articulate their actions in detail. Therefore, in our fifth interview, we returned to a technique similar to Bubble Drawing to get people to talk about their actual Roomba usage and to explain why they used or not in those ways. This time, we provided participants with a photo of Roomba with blank boxes that had a question asking on what occasions the robot generates sound. Also, we gave out a check list of Roomba activities focused on maintenance (i.e., have replaced battery) and ask participants to mark those they had done. Both activities helped us calibrate their familiarity and determine whether or not they had used Roomba. For example, most users could tell us about hearing the start, stop and low battery sounds while others were unaware that the Roomba even produced noise. To ensure the validity of these exercises, we asked participants to elaborate on what they wrote or checked.

Second, we asked participants to visually compare the value of Roomba to other appliances. In previous interviews, many participants had compared Roomba positively and negatively to other technologies they owned. Inspired by this, we generated a list of potential domestic technologies each of which we put on to a sticker. We then asked them to group the appliances that they owned on axes that reflected pleasant-unpleasant and useful-useless (Figure 5). After they completed this activity, we asked them to add Roomba into their diagram. It allowed us to hear how people thought about it and compared it with other household appliances. Frequently, participants spoke about appliances in theme of the consequences of not having the item, or the joy felt using it. Most importantly, we learned whether households regarded Roomba as an essential appliance even having replaced the vacuum or whether it was a luxury that remained fun to use.

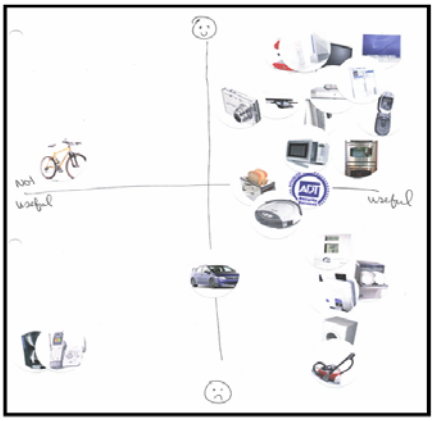

Figure 5. P2's Assessment of Appliance/Roomba Value

After these two generative activities, we asked participants to offer potential improvements that could make Roomba easier to incorporate into their households beyond these first six months. We wrapped up the study by asking the participants to discuss the overall study procedures with us, particularly regarding some of the generative activities to learn whether, and if so how, it helped them reflect on their Roomba routines and if it was burdensome to write and draw instead of talk.

To summarize, we conducted five interviews over six months in each participating household to capture how householders adopted Roomba from starting with it as a novelty to it becoming a part of their everyday lives. In each interview we used various methods to learn about participants' developing relationships with Roomba. Particularly, we used several generative methods that involved more than just either talking and observing, such as writing, rating and drawing. We did so because literature [2,20] suggests that these activities help participants more easily discuss topics that are difficult to express in words. In our study, our participants talked about robots which they had no prior experience of, while scaffolded by activities that involved rating, drawing and creating conversations around Roomba. In the later interviews such as those that occurred after 2 months usage, these generative activities helped us uncover Roomba routines that participants had difficulty in articulating. We incorporated activities including blueprints of the home to capture where Roomba was used, an activity checklist to identify what they have or have not done and a visual map of Roomba's value to understand how participants viewed it among their appliances as essential or a luxury, and as useful or pointless.

4. DISCUSSION

Learning from our own and our participants' experiences, here we discuss three implications of doing a long-term study in the home. First, we discuss the general logistics of conducting a

longitudinal study. Second, we describe the study methods we used to mine natural interaction and how we attempted to do this without interfering with privacy and temporality of domestic routines. Finally, we capture how we surfaced the conceptually less visible routines.

4.1 Conducting a Long-Term Study

A previous study [24] notes that long-term research is required to move beyond the novelty effects of adoption and understand natural human-robot interaction. We agree. However, longitudinal research on robots requires, as we learned, a variety of techniques. It also requires answering a question: how long is enough to uncover long-term interaction?

Our own research shows that about two months was the minimum time that we needed before participants settled into routines. For example, after two weeks we saw children excitedly turning on the vacuum, but they had stopped this by two months. Of course we saw individual differences in settling into routine patterns of use. For instance, P12 were still experimenting with the remote and scheduling functions at their 2nd month usage while P30 only began using Roomba regularly in their 5th month interaction when they took it to their cottage. However, we generally saw less change in their Roomba usage from the two to six months than from two weeks to two month visits, implying that participants' relationship with the robot whether positive or negative had stabilized. Further, our participants began to talk about Roomba as part of the family after two months. More generally, our point here is that whether the perceived robot experience was positive or negative, all of these human-robot interactions took time to play out. At a minimum we suggest that two months is essential.

At the final interview we asked householders whether they had felt obligated to use Roomba because of their participation in our study. Fortunately, none reported feeling pressure, although a number commented that this was particularly true in the later half of the study perhaps as they became more familiar with us, or as they realized that we were genuinely looking for their own responses. One participant, P24, even said that she had forgotten that Roomba came from the research team until we scheduled the last interview. P30, who had not used Roomba in the first four months of our study, also said that she felt comfortable with her non-use since we had stressed the importance of letting participants decide whether they wanted to use it or not.

Overall, our study findings suggest that six months was sufficient for participants to develop stable routines. Comparing the last two visits, we did not see much change in usage between months two and six (i.e., frequency of use and primary user). What did appear to change was an increase in the number of technical problems people reported with Roomba and how they coped with those difficulties. Consequently, we suggest that for routines in the home, approximately two months may suffice to see stability although longer timeframe might yield information about reliability. We note that our study case specifically speaks to Roomba and cleaning routines, and hence different time frames may be necessary for other robots or routines. Yet, other studies such as [11] also suggested that two months allowed researchers to identify routine interactions. Hence, we believe that a minimum two months provides a good baseline for upcoming long-term studies on Human-Robot Interaction.

4.2 Mining Natural Interaction in Homes

In this study we constantly sought to balance interventions against creating a natural adoption experience. Where we intervened, such as asking all family members to be present at opening Roomba's initial arrival in the homes, it was to ensure that people did not reject it just because they were not present at the time of arrival. But, we also chose not to make certain types of interventions, in particular not remotely recording video/audio or logging that was possible by modifying Roomba. Our initial motivation had been focused on ensuring that the units were all "as is" so that participants did not get confused between our modifications and the product itself.

However during the study, we also learned that logging could have caused us serious difficulty. Although participants spoke of not wanting the device to log due to a sense of that being invasive, it was particularly acute in the case of P9. P9 sent us a help request saying that her Roomba (named Ricky) was broken and asked us for help. Given our interest in how householders coped with trouble, we asked her to first troubleshoot. This resulted in the following angry response to us:

It is cruel. I just realized that each of your subjects has been given a Roomba preprogrammed to fail. If you want me to further participate in your study, just fix this machine. I am disabled, busy and don't know where my product manual is.

Of course, we responded by again explaining our motivations, but also offering help. But we learnt that had we tweaked the system, we would have been far more open to the perception that this was a deliberate part of the study itself.

Most of our choices about interventions (writing, drawing) were made in response to routines. For example, our structured exercises (i.e., stickers, blueprints) sought to mine the natural interaction by surfacing data about increasingly invisible routines. We return to this in our next section. But, we also learned that such a structured exercise would not work particularly well when it interfered with everyday activities, most notably in the case of the log tools we tried to use early on. Logging was unsuccessful, but by contrast, inviting participants to keep in touch with us through the means they use on a daily basis such as email and photo-sharing generated much more data that we could use. Indeed, some participants still contact us even though the study has finished, just to let us know how their Roomba is being used. For example, P27 emailed us with photos two months after completing the study because they wanted to let us know how their baby had continue to grow attached to the robot:

(our baby) started crawling this month and his favorite thing to do is to crawl straight to the Roomba and sit on it! He will also try to follow it around if it is moving and cleaning. Some pictures are attached. Thought that might be an interesting side note for the study :)

Sharing pictures, particularly of children is a relatively natural phenomenon, and we suggest that it has practical (although variable—not all families will participate in this) application to long-term research studies. Another downside of moving to these self-reports of course focuses on assessing their accuracy. Where possible, we sought to triangulate our data with multiple methods. For example, we were consistent about asking questions via multiple techniques such as using the blueprint to reinforce the interview. Because of this desire to triangulate, our exercises also

shared some common questions, for example, asking people to rank Roomba impressions on a Likert scale (Figure 4) and on a visual map (Figure 5). Again, we incorporated redundancy so that we could improve the veracity of our participants' responses.

Another important means of triangulation was the participants themselves. By interviewing people together we were able to hear the participants talk with each other and comment on whether they thought their spouse's or children's response was true. Further, given that we had also used snowball sampling to recruit some of our households, it was also the case that they knew what was going on in other people's homes. For example, P15 and P16 were close friends, and P16 characterized P15's adoption:

P15 said (Roomba) changed her life. She just loves it. I was like "are you talking about the bible or God, or are you talking about Roomba?" and she talked about Roomba.

To summarize, longitudinal studies of domestic settings require different approaches from field research in public spaces. Interventions likely will need to be made, in order to combat the problem of invisibility in use. That said, leveraging natural routines such as photo sharing and emailing may be more advantageous than attempting to create even more new routines such as logging around the already new routines under investigation (robotic cleaning). We also incorporated as much triangulation and consistency over time into our research. That said, the five study phases were intertwined with each other, which made us especially judicious about piloting all our techniques, minimizing the risk of having a household stuck for as much as six months in a poorly designed study.

4.3 Dealing with Invisibility in Use

Several scholarly works on domestic technologies have confirmed that routines are challenging for participants to identify clearly [3,26]. As a way to overcome, we incorporated generative activities inspired by Sanders [20]. She suggests that the making process (i.e., the process of active construction such as the ideal robot, and Bubble Drawing) helps participants surface their unconscious needs and desires that otherwise remain difficult to express in words.

During our debrief on the overall study procedures at the end of the final interview, several participants noted how the sticker exercise of placing Roomba along with other home technologies (figure 5) helped reveal parts of their lives that they were not aware of. Some householders even reported that they learned new information about their spouse or children by hearing how they valued or thought about Roomba. A 17-year-old son from P24 remarked on the sticker exercise as:

(p24 son) I like stickers. It makes you think about it. You don't really know what you are doing (in your daily activities) until you start twining it (through stickers of different home technologies) and you learn more about it, about how you really use stuff and how you value different things.

Our participants mostly responded positively toward generative practices including, writing, rating, drawing and making a collage by referring to the workload as "reasonable". Of course, a few participants clearly noted they preferred the convenience of speaking. However, our participants discussed how adding fun activities got them more engaged into the lengthy process. P4 reflected upon the Bubble Drawing technique (figure 3) and explained how the cartoony drawings and the scenarios of two friends conversing made her much more willing to write pros and cons of Roomba. These fun activities also played a crucial role in engaging people in the study. P2 (male) enjoyed the collage exercise of making an ideal robot (Figure 2). He went as far as to take a photo of it for his personal use. When we returned for later interviews, his wife told us that he excitedly awaited each interview. Fun and engagement of participants, particularly over the long duration of study is, we suggest, a key to successfully conducting such research, and yet not frequently highlighted. Besides having fun, some participants told us that they were much more inclined to go through the lengthy process of generative exercises if they saw good reasons behind it. For example, P24 (mom) found it difficult to draw the blueprint of the house but she tried her best since she felt that it "was necessary to look at how we use Roomba". Later, she reworked the blueprint on her computer to help us better. In her words,

Okay...I put this floor plan together (in Power Point). It should be a lot more helpful than the pitiful :-) replication of a drawing I drew today. Hope you find it much more useful than the other.

Of course, this creates another challenge. We needed to explain the purpose of the tools that we used, but sought to do this without biasing their use. More generally, we found that by sharing the rationale for our methods, rather than just administering them, built up the rapport that we needed with our participants. The fact that only one house chose not to complete, and that was just one final interview, we think in no small part was due to our openness about the approaches we took.

In summary, we find generative exercises—anything that involves writing, drawing and making—useful in capturing behavior that have become routine in everyday activity. Our participants supported this by saying it permitted them to think deeper. Although some reported difficulties of expressing their thought with their hands, they also said that both adding fun and sufficient justification helped engage with the exercises.

5. CONCLUDING REMARKS

In this paper, we reported on our approaches to conducting a long-term field study *in the home*, a site that requires different means of investigation from public spaces such as hospitals and schools due to the private and temporal nature of domestic routines [3]. Our goal was to understand householders' interactions with a robot "in the wild". Methodologically this involved mining natural activities without interfering with their daily routines and also surfacing the less visible routines particularly in later phases of the study. We described our study methods in detail by explaining the fives phases of our longitudinal investigation. The five phases can be identified as: first, a*pproximately one week prior* to giving out the Roomba, we assessed their existing cleaning routines, and second, we brought a Roomba and observed householders' first responses, and third, t*wo weeks later* we returned to see how they used the robot, and fourth, t*wo months later* we visited again to see what has changed with Roomba usage, and fifth and finally a*pproximately six months* after the arrival of Roomba, we completed the study by asking for the householders' overall impressions of robot.

We concluded by discussing three points. First, we learned that at least two months is desirable for seeing stable interactions between robots and householders emerge. Second, we argued that

some interventions such as logging may be necessary to mine natural interactions. For the interventions that require participants doing tasks without the interviewer's presence, we suggest that they are more willing to cooperate if the activity fits into their routines (e.g., emailing photos rather than keeping a scrapbook). Third and finally, we reported on how we coped with the difficulties participants had in expressing their daily routines such as cleaning, which we addressed through the utilization of generative activities such as creating Bubble Drawings that allowed our participants to surface the unconscious.

More generally, we offer our findings as a starting place for future researchers interested in conducting longitudinal studies, particularly in the home. As personal domestic robotics become increasingly posited as the solution to a variety of challenges, such as those faced by an aging population who wish to live independently for longer, so we think that HRI will be challenged and find exciting empirical and design questions focused on what it means to build robots for long-term use. We offer our experiences as one approach for engaging in this type of research agenda, and look forward to learning from others.

6. ACKNOWLEDGMENTS

We thank all of our participants for their time and engagement in this study. We also thank iRobot for their donation of Roombas.

7. REFERENCES

[1] Bernhaupt, R., Obrist, M., Weiss, A., Beck, E. and Tscheligi, M. Trends in the living room and beyond: results from ethnographic studies using creative and playful probing. Comput. Entertain. 6, 1 (2008), 1-23.

[2] Boddy, C. Projective techniques in market research: valueless subjectivity or insightful reality? International Journal of Market Research. 47, 3 (2005), 239-254.

[3] Crabtree, A., Hemmings, T. and Rodden, T., Pattern-based support for interactive design in domestic settings. in DIS'02.

[4] Forlizzi, J. 2007. How robotic products become social products: an ethnographic study of cleaning in the home. in HRI '07. 2007, Arlington, Virginia, USA, ACM Press

[5] Forlizzi, J. and DiSalvo, C. 2006. Service robots in the domestic environment: a study of the roomba vacuum in the home. in HRI' 06. Salt Lake City, Utah, USA, ACM Press.

[6] Gaver, B., Dunne, T. and Pacenti, E. Cultural Probes. Interactions. 6, 1 (1999), 21-29.

[7] Gockley, R., Forlizzi, J. and Simmons. 2006. Interactions with a moody robot. in HRI' 06. 2006, Salt Lake City, USA.

[8] Gold, K., Fasel, I., Freier, N.G., and Torrey, C. 2007. Young researchers' views on the current and future state of HRI. in HRI '07. 2007, Arlington, Virginia, USA, ACM Press.

[9] Green, A., Huttenrauch, H., Norman, M., Oestreicher, L. and Eklundh, K.S., User centered design for intelligent service robots. in RO-MAN 2000. 161-166.

[10] Huttenrauch, H. and Eklundh, K.S., Fetch-and-carry with CERO: observations from a long-term user study with a service robot. in RO-MAN 2002. 158-163.

[11] Kanda, T., Sato, R., Saiwaki, N. and Ishiguro, H. A Two-Month Field Trial in an Elementary School for Long-Term Human-Robot Interaction. IEEE Transactions on Robotics. 23, 5 (2007), 962-971.

[12] Kientz, J.A., Patel, S.N., Jones, B., Price, E., Mynatt, E.D. and Abowd, G.D., The Georgia Tech aware home. in CHI '08 extended abstracts on Human factors in computing systems.

[13] Kim, H., Lee, H., Chung, S., and Kim, C. 2007. User-centered approach to path planning of cleaning robots: analyzing user's cleaning behavior. in HRI '07. 2007, Arlington, Virginia, USA, ACM Press.

[14] Kotaro, H., Daisuke, S., Takayuki, K., Masahiro, S., Satoshi, K., Hiroshi, I., Tsukasa, O., and Norihiro, H. 2007. Humanoid robots as a passive-social medium: a field experiment at a train station. in HRI '07. 2007, Arlington, Virginia, USA, ACM Press.

[15] Marti, P., Pollini, A., Rullo, A., and Shibata, T. 2005. Engaging with Artificial Pets. in Proceedings of the 2005 annual conference on European association of cognitive ergonomics. 2005, University of Athens, Chania, Greece.

[16] Mutlu, B. and Forlizzi, J., Robots in organizations: the role of workflow, social, and environmental factors in human-robot interaction. in HRI'08, ACM, 287-294.

[17] Pacchierotti, E., Christense, H.I. and Jensfelt, P., Design of an Office-Guide Robot for Social Interaction Studies. in International Conference on Intelligent Robots and Systems (IEEE/RJS) 2006, Beijing, China.

[18] Rogers, E.M. 1995. Diffusion of Innovations. Free Press.

[19] Rowan, J. and Mynatt, E.D., Digital Family Portrait Field Trial: Support for Aging in Place. in the SIGCHI conference on Human factors in computing systems, (2005).

[20] Sanders, E. Converging Perspectives: Product Development Research for the 1990. Design Management Journal (1999).

[21] Shove, E. and Southerton, D. Defrosting the Freezer: From Novelty to Convenience. Journal of Material Culture. 5, 3 (2000), 301-319.

[22] Sung, J.-Y., Grinter, R.E., Christensen, H.I. and Guo, L., Housewives or technophiles?: understanding domestic robot owners. in HRI'08, Amsterdam, Netherland, ACM Press.

[23] Sung, J., Guo, L., Grinter, R.E., and Christensen, H.I. 2007. "My Roomba is Rambo": Intimate Home Appliances. in UbiComp 2007. Innsbruk, Austria.

[24] Tanaka, F., Movellan, J.R., Fortenberry, B. and Aisaka, K. Daily HRI evaluation at a classroom environment: reports from dance interaction experiments. in HRI' 06. Salt Lake City, Utah, USA, ACM Press.

[25] Tolmie, P., Pycock, J., Diggins, T., MacLean, A. and Karsenty, A., Unremarkable Computing. in SIGCHI conference on Human factors in computing systems. 2002.

[26] Venkatesh, A. Computers and other interactive technologies for the home. Commun. ACM. 39, 12 (1996), 47-54.