

Exploiting Ability for Human Adaptation to Facilitate Improved Human-Robot Interaction and Acceptance

Praminda Caleb-Solly^a, Sanja Dogramadzi^a, Claire A.G.J. Huijnen^b, Herjan van den Heuvel^b

^a Bristol Robotics Laboratory, University of the West of England, Bristol, United Kingdom

^b Smart Homes, Eindhoven, The Netherlands

RUNNING HEAD: Human Adaptation in Robot Interaction

CONTACT INFORMATION

Please provide mailing (postal) address and email address of the corresponding author.

Dr. Praminda Caleb-Solly, Bristol Robotics Laboratory, Faculty of Environment and Technology, University of the West of England, Bristol BS16 1QY, UK

praminda.caleb-solly@uwe.ac.uk

ABSTRACT

This article reports findings from the third phase – usability and user experience evaluation – in the last two years of a four-year evaluation project of the Kompaï robot. It focuses on the evaluations that were conducted in a robotic ambient assisted living studio in the UK (which was arranged as an open-plan studio apartment), a UK residential care home, and an older couple’s own home in The Netherlands over two days. It examines emergent adaptive human behaviour in human-robot interaction (HRI) to consider whether we are approaching the embodiment and functionality of service robots correctly. It discusses possible improvements that could be made at the systems level that better exploit people’s natural ability to adapt and find workarounds to technologies and their limitations.

INTRODUCTION

In order to make progress in the development of robust and safe assistive robots, it is imperative to conduct evaluation studies “in the wild.” However, this is a very complex and difficult endeavour, which involves technical challenges, unpredictable human behaviour, and ambiguous contextual information. These factors have stymied long-term realistic evaluation studies.

In this context, we need to develop alternate strategies that do not wholly rely on longer-term trials and real-world evaluations and yet yield insights that further the development of assistive robots. This study is one such effort. It studies emergent adaptive human behaviour in human-robot interaction (HRI) to consider whether we are approaching the embodiment and functionality of service robots correctly and provides input for the further development of assistive robots.

This article presents our findings from user-based evaluation studies of the Kōmpaï mobile robot platform from Robosoft. It proposes a more pragmatic approach for conceptualising and realising successful interactions with assistive robots when long-term trials are not possible in the early stages of development.

The next sections present a brief overview of related work, our methodology, findings of our evaluation studies, and a discussion of possible improvements that could be made at the systems level that better exploit people’s natural ability to adapt and find workarounds for technologies in the face of their limitations.

RELATED WORK

This section provides of a brief review of related studies on human-robot interaction and user

acceptance.

People's personal experiences of interacting with a robot are influenced by their expectations, which in turn are shaped by their previous experience and attitudes (Bartneck, Suzuki, Kanda, and Nomura 2007). Here exposure to popular culture, science fiction, and media has a formative influence. It tends to imbue unrealistic expectations of the interaction capability and functionality of robots. In their article examining portrayals of robots in movies and science fiction, Sandoval, Mubin, and Obaid (2014) identify the mismatch between what robots can actually do now and what the movies and science fiction portray them to be capable of. They then discuss the implications of this for HRI designers, especially the characteristics the robot should have to become socially acceptable.

A large proportion of the public is still ambivalent about the use of robots to support older adults (EC 2012). Familiarity and interaction with a product over a period of time can influence people's expectations and experiences, helping them understand its scope and utility, and develop an attachment to it. As one of the objectives of assistive robots is to support people in living healthier and more active lives, developing a positive attitude towards the robot needs to be encouraged. In his article on *Designing for the Self*, Zimmerman (2009) considers product attachment as a process of meaning making and identity construction. Here again we see the need for studies spanning longer periods of interaction with a robot in their daily lives. Such research could also help develop a more realistic and detailed understanding of user needs and requirements.

While longer-term studies help determine the functional effectiveness of the assistive robot, we also

need to understand which aspects of the user experience are affected by continued use. This involves understanding of a range of factors such as changes in levels of attentiveness and responsiveness, which can occur when interaction tasks become assimilated as procedural knowledge (Sun, Merrill, and Peterson, 2001), and what impact this might have on human error and safety.

While research teams developing assistive service robots invariably include testing and evaluating with end-users as an integral component of their work, these tests are often limited to semi-Wizard of Oz (WoZ) type simulations (Green et al., 2004) or conducted over rather short periods of time and under controlled conditions, and therefore do not always allow for ecologically valid results. Researchers such as Syrdal, Dautenhahn, Koay, and Ho (2014) are starting to develop more promising frameworks for longer-term evaluations that could improve the ability of participants to assess and respond to robots with a clearer understanding of how they relate to their everyday lives. However these are still only applied as part of high-fidelity interaction prototyping, where the robot is developed to a level which enables realism of the interaction and behaviour from the participants' perspective without implementing actual autonomous functionality.

The longer-term trials in the literature rely on commercially available robots that have been certified for safety or have limited mobility, which does not present a significant safety hazard if left unattended. Based on their studies of the commercially available Roomba robot, Sung, Grinter, and Henrik (2010) proposed Domestic Robot Ecology – an initial framework to unpack long-term acceptance of robots at home. As part of this framework, they defined four temporal stages: pre-adoption, adoption, adaptation and use/retention. Of particular interest, is the adaptation phase, where people are willing to learn more about the robot, its technical limitations and affordances, which prompts them to start adapting their environment and behaviour (e.g. picking up clutter and moving items to provide the robot better access)

to attain the benefits they had witnessed in the previous two stages. The adaptation often extends to making social changes as well, for example in the Sung et al. study, the teenagers in the family became responsible for cleaning.

More recently, de Graaf, Allouch, and Klamer (2015), using the commercial Nabaztag robot, have specifically considered issues that would be affected by long-term usage such as relationship building and long-term acceptance, focussing on what factors play a role in domestic environments and what kind of relationships people build with social robots. Their findings confirm those of Sung et al. with regards to the adoption process. de Graaf et al. noted a ‘mere-exposure effect’ which they describe as ‘the tendency to evaluate novel stimuli more positively after repeated exposure’. They found that familiarity leads to greater appreciation of several aspects of the robot over time, where participants evaluated the robot as more useful, more intelligent and more sociable after each usage phase. This finding is of great significance as most of the trials with assistive robots still tend to be carried out over short-time frames.

While commercial robots offer a range of socially assistive functions, their utility is limited by their inability to reach the end-user in a pro-active manner. For larger and more mobile robots, such as the Kompaï robot used in our research, conducting safe and reliable long-term studies has proved problematic as there are considerable technical issues that have to be resolved in the first instance. Our research builds on the evaluation methods reported in the literature to identify usability and user experience issues and considers how we might be able to overcome some of the limitations of the technology during evaluation studies.¹



Figure 1 Kompaï mobile robot with ‘clothes’ as Molly

METHODOLOGY

The Kompaï physical robotic unit (Figure 1) can be programmed to move autonomously between mapped locations in a given space. Currently, its mobility is its primary advantage over a laptop or tablet.

The evaluation was conducted in three phases: a technical validation phase in the first year, an integration test phase in the second year, and usability and user experience evaluation phase in the last two years of the project. The third phase, which involved iterative prototyping, had 11 evaluation points (Table 1) was conducted with a total of 98 participants. These evaluations included laboratory based testing, field trials, expert usability evaluation and carer engagement. Based on this feedback, a new version of the Kompaï robot was developed.

Evaluation Point	Evaluation Type
1	Initial System Integration, Installation, and Heuristic Evaluations
2	Stage 1 Pilot Studies
	Usability and User Experience Testing
3	Stage 1 User Orientation Workshops
4	Stage 2 Individual Trials
	Field Trials
5	Dementia Unit User Experience Trial
6	Day Care Centre User Experience Trial
7	Residential Care Home User Experience Trial
8	Home Apartment User Experience Trial
	Domain Specific Studies
9	Hazard Analysis Study with Overnight Experience
10	Show and Tell Sessions in Care Centres
11	Consultation Session with Secondary and Tertiary Users

Table 1. Evaluation points in the third evaluation phase. This article reports findings related to evaluation points 3, 4, 7 and 8 (highlighted in green).

This paper describes the findings of evaluation points 3, 4, 7, 8 which were conducted one-to-one basis with older adults using a WoZ approach. Evaluation points 2, 3 and 4 were conducted in the robotic ambient assisted living studio in the UK (which was arranged as an open-plan studio apartment),

evaluation point 7 in a UK residential care home, and evaluation point 8 in an older couple's own home in The Netherlands over two days.

While we had hoped to conduct long-term experience trials, unfortunately, due to the lack of system autonomy and concerns regarding technical stability of the platform over continuous interactive use, we could only conduct these with two younger adults (evaluation point 9). We were limited to focus groups and interactive demonstrations (evaluation points 5, 6, 10 and 11). We have not included the findings from these studies here as we are focussing on specific experimental conditions in this article.

Experimental Set-up

The participants were exposed to the robot over two interaction sessions (evaluation point 3, 4) separated by a gap of 4.5 to 6 weeks (an average of 5 weeks for all participants).

In the user orientation workshops (evaluation point 3), the aim was to familiarize participants with the project and robot. The workshops featured live demonstrations of the robot, with one of the researchers play-acting the role of a user. Thereafter participants were provided an opportunity to have some interaction with the robot, copying some of the interaction that they had seen demonstrated. At the end of the session, participants were given a questionnaire with items on features and context of use. We hoped that this first session would reduce the novelty effect for subsequent interaction sessions.

In the second session, the individual trials (evaluation point 4), we used a specially designed WoZ graphical user interface (Figure 2) to control the robot's movement (send it to a specific location or turn it), speech (e.g. give prompts to the participant), and thereby generate different robot behaviour and user interaction situations.



Figure 2 The Wizard of Oz interface used to remotely control the robot

As per the Wizard-of-Oz framework defined by Green et al. (2004) our human-robot interaction trial session included the following elements:

a) User instruction:

The participants were given an instruction sheet to refer to if needed, as well as a task workbook with six scenarios and some additional questions. The scenarios were developed based on our previous research on user requirements to provide realistic situations wherein a person might interact with the assistive robot (Nani, Caleb-Solly, Dogramadzi, Fear, and van den Heuvel 2010). They were centred on following tasks and prompts: (1) viewing photos, (2) a carry and deliver task, (3) playing a game, (4) medication and appointment reminders, (5) exercise encouragement, and (6) social interaction via a video link. Figure 3 shows examples of filled in worksheets of the task workbook.

Scenario 1

- Imagine that you are feeling nostalgic and want to see some of the photos stored on Molly.
- Start by getting the robot to come to your armchair.
- You can put it into a comfortable viewing position by using the "TURN LEFT or TURN RIGHT" commands. You can also adjust the screen to a comfortable viewing angle for yourself.
- After you have viewed all the photos stored on Molly describe the one you found most interesting and tell us why

The picture of us on holiday. Very clear, and a great surprise.

- What photos you would like to store on Molly if you were using Molly at home

As many of the older photos of holidays and grandchildren and children. Mostly people, rather than just landscapes. Also well-kept places that I've never visited.

Scenario 2

- Imagine that you would like a drink, some fruit and a plate of biscuits which are all in the kitchen. You can get Molly to carry the drink and fruit for you from the Kitchen to your armchair. You can carry the plate of biscuits yourself.
- So try and do the following:
- Send her to the kitchen, put the thermos flask in the holder on Molly and then send her back to your armchair. Don't forget to bring the plate of biscuits yourself.
- When you are settled back down in your armchair again, please tell us:

	Please circle your choice				
How easy it was to interact with Molly for this scenario? 1 = Very Hard 5 = Very Easy	1	2	3	4	5
Did you feel that you were able to control Molly safely? 1 = Very Unsafe 5 = Very Safe	1	2	3	4	5
In the future, if you needed to, would you want a robot like Molly to help you transport items around the house? 1 = No, never 5 = Yes, definitely	1	2	3	4	5

No STOP COMMAND WHEN MOVING

Scenario 3

- Imagine that you are feeling bored and would like to play a game.
- Select a game of your choice and play it for a few minutes.
- Select another game and play that for a few minutes as well.
- After you have played the games please tell us whether you can see yourself playing these or other games if you were using Molly at home

robot recognize menu command should recognize parts of command e.g. play a game should recognize play or game

Yes No Maybe

- What games would you like to see added on Molly, and would you play them on your own, with a visitor or both?

Need to install games based on user choice e.g. Free cell. Some didn't seem to work all the time

Scenario 4

- Imagine that you have just woken up and are sitting up in bed. Molly is going to come and give you two reminders for the day.
- Please go and sit at the edge of the bed and wait for Molly.
- After Molly has given you the reminders, please send her to the dining room.
- How did you feel about receiving the reminders from Molly?

Great, because it would save me from setting alarms to take med. etc

- If you were using Molly at home, what reminders would you like her to give you?

Have you had a properly balanced meal. Have a good breakfast. Have you turned off the gas cooker.

- Would you allow a relative or a carer to set-up reminders for you on Molly? Would you feel it was bossy or helpful?

HELPFUL.

Scenario 5

- Imagine that you have been sitting for a long while watching the television. Movement is important to keep your circulation going so Molly is programmed to encourage you to do a physical activity to keep you fit and well.
- Please go and sit in the OTHER CHAIR and wait for Molly to come to you and suggest you choose an exercise. Choose an exercise you would like to do and copy the movements if you are happy to do so.
- If you were using Molly at home, how would you feel about receiving exercise reminders and suggestions from Molly?

I think it would be very good to be motivated otherwise you may just be nothing

- Do you think that you would exercise more often if you were given these video guides?

Yes even for the fun of it to break the boredom of sitting doing nothing and not using your mind

- Can you imagine Molly as a guide to other physical activities? What would these be?

Rotating Ankles Raising legs up & down turning the head side to side facial muscles stretching fingers Rotating wrists

Scenario 6

- Imagine that you are having lunch and your friend calls you via Molly's video link.
- Please go and sit at the dining table on the marked chair and wait for Molly to come to you when the call is received. As you are in the middle of lunch, please answer the call and tell your friend that you will call them back after you have finished lunch.
- After you have hung up, have a drink of water, wait a few minutes and then call them back.

No indication that the call was connecting

	Please circle your choice				
How easy it was to interact with Molly for this scenario? 1 = Very Hard 5 = Very Easy	1	2	3	4	5
Would you like to use the video link while talking to your friends and relatives?	Yes	No	Maybe		
Do you use a phone frequently at the moment? Landline/Mobile/Skype/Other () (Delete ones you don't use)	Yes	No			

A few more questions

- How was your experience of using Molly today compared with the last time you were here?

I felt that I wanted to chat to Molly - doing a running commentary. However she responds when you just want to chat to yourself - wants to be given a command. I think I could see her more as a 'person' this time. This surprised me. I had to learn to pause after 'more to robot' before saying when to!

- What suggestions do you have for developing Molly further?

The options Molly offers are useful. I suppose in due course it would be useful for her to have arms to enlarge the repertoire of what she can do.

Thank you for your time!

Figure 3 Task Worksheets completed by participants

Following the completion of the scenarios by the participants, a post-task structured interview was conducted, which included questions such as: What is it like being with a robot? What they liked and disliked about the robot? Whether they could imagine a robot in their home? What would the robot have to do/not do to become a useful companion? What were their privacy concerns with regard to the data

collected by the robot? There was also a follow-up discussion on their responses in the original questionnaire to see whether and how their perceptions had changed.

b) Behaviour Hypotheses:

Informed by Salem and Dautenhahn (2015) and Weiss, Bernhaupt, Lankes, and Tscheligi (2009), we investigated a number of issues with regards to HRI such as impact of conversation style, voice, and specific behaviour patterns on situated user response on different populations (gender, age, disability, cultural, and educational background). Specifically, we assessed levels of user satisfaction, functions used, ease of use (ergonomic) of overall device, ease of use (ergonomic) of the input devices, and user rating of functions performed and feedback provided (in relation to quality, utility, and comprehensibility). In addition, other factors such as response time and time for error recovery were observed.

c) Robot behaviour:

For this trial, the participants were left alone with the robot in a studio apartment after a safety briefing and a reminder of the voice and touch interfaces and interaction process. The facilitators observed the interaction via the live camera feeds and a glass door. The WoZ interface was operated by a researcher to make the robot's behaviour appear autonomous as the user worked through each of the scenarios on the supplied task worksheets. Figure 4 shows the view from one of the two cameras of the user working through a scenario and an inset view from another camera.



Figure 4 Participant performing scenario 2

Based on our findings from previous studies (Caleb-Solly, Dogramadzi, Ellender, Fear, and van den Heuvel 2014) and (Huijnen, Badii, van den Heuvel, Caleb-Solly, and Thiemert 2011), which employed mixed-method approaches to investigate embodiments and acceptance of assistive robots, we decided to explore the impact of some changes to the physical appearance of the robot. Since the two previous studies found that people preferred softer shapes and forms and did not want to see exposed metal parts, cables and wires, some ‘clothes’ were custom-made to cover up the metal back and the front where cables connect the speakers and USB hub. In the UK such covering took the form of a ‘cape’ and ‘skirt’ and in the Netherlands a green and yellow ‘scarf’ was donned on an unusually long metallic ‘neck.’ In the stage 1 workshops in the UK the robot was called ‘Molly’ and in the Netherlands the robot was called ‘Max’ and later ‘Charley.’ These names were found to be easier for people to remember and relate to than the Robosoft’s generic name for the robot Kompai. The ‘clothes’ also helped the users visualize the ways in which they could customize e.g. choose ‘clothes’ that match their home décor. For the evaluation point 7, the orientation and trial sessions were combined due to the participant having early dementia.

Study participants

For evaluation points 3 and 4, there were eight participants aged between 64 and 78 years, 5 males (Coded as 1M, 2M, 3M, 4M and 5M) and 3 females (1F, 2F, 3F).

Evaluation point 7 only involved 1 male participant (Coded 6M), who had early dementia. The study was conducted in his private room in a residential care home in the UK and the scenarios were adapted accordingly.

Evaluation point 8 included a wife and husband, aged 84 (4F) and 90 (7M) respectively, who were living in their own apartment in the Netherlands in a building that offered some care services.

The characteristics of all the participants are detailed in Table 2. The recruitment criteria included people aged over 60 suffering from some ageing-related impairments but with stamina to participate in 2-3 hour studies over a 5-6 week period.

Characteristics	Participants
Living Status	Alone (1F, 2F, 2M, 4M, 5M, 6M); With Spouse (3F, 4F, 1M, 3M, 7M).
Mobility	Wheelchair/Walking Aid (1F, 5M, 7M); No Aid (All the rest).
Eyesight	Reading Glasses/Contact Lenses (All)
Hearing	Hearing Aid (1M, 7M)
Memory	Dementia (6M); Some Short-term Memory Loss (1M, 2M)

Existing Use of Technology	Email (all except 3F, 3M, 4M, 6M, 7M); Web Browsing (all except 3F, 3M, 4M, 6M, 7M); Internet Shopping (all except 1F, 3F, 3M, 4M, 6M, 7M); Word Processing (3F, 3M, 4M, 6M, 7M); Computer Games (2F, 5M); No Technology Use (3F, 3M, 7M)
-----------------------------------	---

Table 2 Participant characteristics (Evaluation points 3, 4, 7 and 8)

Analysis

The participants' responses on paper questionnaires and worksheets were entered into a digital system to enable easier analysis. Each trial session was recorded with a set of 5 cameras. MORAE², a user experience analysis software tool, was used to capture data on touch screen interactions, video recording of faces, and audio recording of voices. A video camera on the top of the robot's head recorded the participants' behaviour and speech. Another video camera recorded the interaction from behind the user. Two additional cameras were set up to record the kitchen and the bedroom areas of the studio apartment.

Parameters such as time taken to complete the tasks, number of errors made, error correction process adopted, and types of errors were calculated for three scenarios, namely (1) viewing photos, (2) the carry and deliver task, and (5) exercise encouragement.

RESULTS

Initial questionnaire

Figure 5 shows the results with 10 denoting (Very good) and 1 (unsatisfactory).

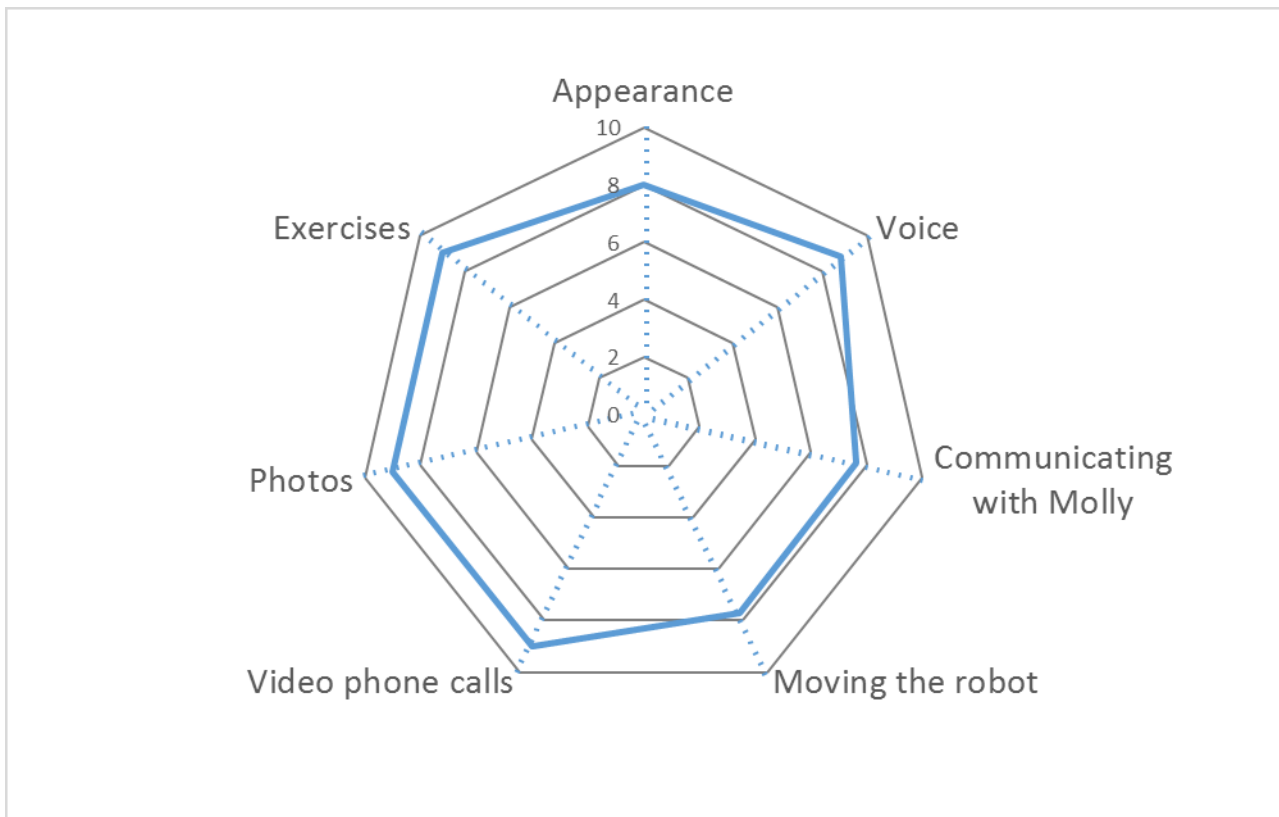


Figure 5 Questionnaire responses to robot features and functionality

Table 3 brings together some quotes excerpts made by participants with regard to the appearance of the robot.

Participant	Quote excerpts with regards to appearance
1F	If cover used may be match décor. (The participant is referring to the cape)
2F	Prefer the current face to a more humanoid one
3F	Very cute and non-threatening
4F	Quite funny
1M	Customization options, face and dress.
2M	I don't see the need for 'clothing'
3M	Easy on the eye

4M	Very good.
5M	It is a new concept.
7M	Robot is not complete yet as it doesn't have arms

Table 3 Quotes made with regards to appearance.

Participant	Viewing Photos (S1)	Carry and Deliver Task (S2)	Exercise Encouragement (S5)
1F	Wonderful for everyone.	Good again. I think you should think direction out.	Wonderful asset.
2F	Useful for reminiscence work.		Customisation needed; in current video the presenter talks too quickly.
3F	Really good for all people, especially with dementia.	A bit difficult at first but easy to pick up afterwards.	I really liked this one as a most useful programme for elderly.
4F			I would never do those exercises.
1M	Customize.	As above but good set of commands.	Must do some myself.
2M	Easy to view.	Formal	A little speedy and everything for older people could cope.

3M	Very worthwhile! Could a video be incorporated?	Confused me at the start with left/right, but I could learn.	Brilliant concept.
4M	Very good	Very good	Very good
7M	Good for reminding		

Table 4 Initial quote excerpts capturing the views with regards to the 3 scenarios.

Scenario interactions

The participants were asked to work through each of the scenarios in turn. All the scenarios could be completed either through voice or touch interaction or a combination of both. The only speech-based commands that did not have a touch-based equivalent were the ones to turn the robot to face right, left or 180 degrees round. These scenarios saw the highest number of errors.

The participants had complete freedom regarding which interaction modality they preferred to use.

Figure 6 shows the interactions for each participant at evaluation point 4.

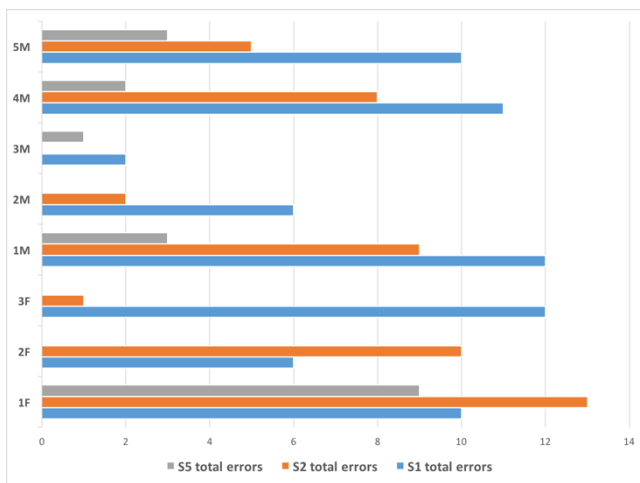


Figure 6: Interaction errors for the 3 selected scenarios (S1 Viewing Photos, S2 Carry and Deliver, S5 Exercise Encouragement)

The types of errors that were tagged included problems with synchronising voice interaction and speech recognition and the participant commands that were not recognised by the robot – speaking not clearly, speaking not loud enough, touching the wrong interaction element on the screen, touching the screen improperly. It should be noted that after experiencing difficulties with voice interaction for scenario 1, the participants resorted to using mostly touch interaction, which resulted in the reduction in errors for all participants apart from 1F and 2F for scenario 2. Both 1F and 2F attempted to chat and converse either with the robot (1F) or to themselves (2F), which resulted in errors in speech recognition by the robot. Participant 2F relied on just touch-based interaction for scenario 5 and did not create any errors.

Usability issues

Synchronization of voice interaction and speech recognition

Voice interaction tends to have problems with synchronization, which affects the accuracy of recognition. These problems usually occur when the user starts to speak before the robot goes into the listening mode. The ‘Not Listening’ and ‘Listening’ modes are flagged on the top right-hand corner of the screen in red and green respectively. Synchronisation errors occurred most when the screen was not in direct view of the participant and in the initial stages of the interaction when the user expectation was of a prompt response from the robot. Some participants did not seem to pay attention to the Listening/Non-Listening mode status and the facilitator had to intervene, this happened most often in the case of participant 1M and 5M. Poor speech recognition and long command sequences were found to result in frustration, particularly in the case of 1F. The touch screen was seen as a natural alternative by all the participants after an average of 4 to 5 failed attempts, apart from 1F and 1M who persisted more than the others in using voice interaction.

The problems with voice interaction for participants 3M and 2M were greatly reduced, as they worked through the scenarios. Understanding the synchronisation and clarity issues, they began to modulate their speech. This points to the potential for the users' to identify obvious incorrect or erroneous behaviour. The designers could use this understanding to develop system capabilities to alternative interaction modalities on encountering repeated undecipherable user actions. To test this potential, we simulated such a capability by giving a verbal message via WoZ advising participants 6M and 2M to wait for the green 'listening' sign. This verbal message resolved the problem.

For participant 6M, who had early dementia and was in a residential care home, we made an interesting observation where he picked up his television remote control and tried to control the robot. This is not altogether surprising as he was regressing to his old habits on how to control a screen. It has implications for design of new technologies that call for screen control via voice or touch interaction. This finding resonates with conclusions of earlier studies (Nygard and Starkhammar 2007, Dick et al.1988).

It was found that the voice of the robot was not always loud enough, even at full volume; participant 1M who used hearing aids in particular had difficulty. In such cases, the robot could be configured to broadcast to an installed loop system and a hearing aid user could switch their hearing aids to the loop. Also, the pitch of the robot's speech could be matched to the user's hearing capability.

With any background noise the robot failed to operate, such as participant 2F talking to herself when the robot was in the listening mode. Participant 2F wrote in the post-session questionnaire: "I felt that I wanted to chat to Molly – doing a running commentary!" In general, misinterpretation of the users'

commands resulted in the robot sometimes executing a wrong response or going into an unexpected state. This seemed to happen most frequently for participants 1F and 1M, who also had the most trouble in remembering to wait for the robot to go into listening mode, resulting in the robot only partially hearing their speech commands. When the robot makes an error in recognition or does something unexpected, some users are quick to blame themselves, participant 1F in particular was saying “Sorry about that” whenever the robot did something unexpected. When the robot does not understand a command it says “Sorry, I did not understand, could you repeat that please,” which invariably compounded the errors as the natural reaction for the participants was to apologise again or repeat the command in a different way immediately afterwards before the robot went into listening mode.

Command ambiguity and logic errors

All the participants had trouble remembering that to make the robot go to a different room one has to go to the “Move the robot” screen first. This seemed counterintuitive to them, particularly as the commands to turn the robot left, right or around are available at the top level. Participants 3M and 4M tried saying “Move forward” and 6M repeatedly said “Come Here,” which seemed obvious and natural to them as the robot seemed to be autonomous. Participants 1M and 2M tried to use more common actions, which should be logically possible but were not programmed, such as “move back.” This backward movement was not implemented due to safety concerns, as the robot did not have edge detection capability at the rear. Similarly, participant 6M, who could not remember specified/programmed location names, often said “Go to the ...[name of place]” – the robot could not understand such commands. Participants used some of the commands differently, such as “Move Right” or “Turn Right”. However, most participants (all except 1F, 4M and 6M) soon realised and learnt how the robot responded or did not respond and were able to use the correct command. The voice “stop” command did not always work, particularly if the user was some distance away from the robot. Consequently, all the participants were found to get

confused as to whether the stop command aborted the current action or not.

Participants 2F, 1M and 3M noted after the individual trial sessions that there should be alternative commands for the same action. Participant 1M suggested that they should be “structured so that they do not have to be accurate, e.g. ‘move the robot’ or ‘move robot’ or ‘move’ should all be acceptable”.

Wizard helping users problem-solve

If participants started getting something wrong repeatedly, the Wizard could get the robot to communicate to them what action they could take. Here are some examples. When the robot’s path was inadvertently obstructed by an object (e.g. participant’s legs in one situation), the Wizard got the robot to say: “I think that something is blocking my way.” When participant 6M repeatedly spoke before the robot switched from speaking to listening mode, the Wizard got the robot to say: “You have to wait till I am in listening mode before speaking to me.” When the participant used words that were not in the robot’s vocabulary, the Wizard got the robot to say: “I don’t understand that phrase.” This was found to greatly improve the smoothness of the interaction and forestall participants getting overly frustrated, with the users soon remembering, understanding, and resolving problems for themselves.

Personification

When the robot got stuck in a loop of listening and responding to itself when it was doing the text to speech conversion for the photo slide show, participant 3F said: “she’s having a nervous breakdown, I think.” Such a comment exemplifies the human tendency to personify robots. In fact, by evaluation point 4 sessions, all the UK participants had started to refer to the robot as “her” or “Molly.”

Personification seems to set in with familiarity with the robot. For instance, in the stage 1 orientation

questionnaire, participant 2F has stated ‘I am ambivalent! Currently I value my independence and find it hard to visualise tolerating a robot in the house.’ But then, in the stage 2 trial questionnaire, she noted: “I felt I could see her more as a ‘person’ this time. This surprised me.”

In the home apartment user experience trial in the Netherlands, when the robot kept repeating “sorry, I did not understand that”, participant 4F said “Why does it say this now? It has to go to the doctor” and laughed.

This personification can have ethical implications, as it could lead to attachment and emotional dependence. Huber et al. note how every single human-robot interaction carries potential ethical risk and can have major impact over the long run. They present a model to help identify ethical risks involved in the development of robots as social companion robots which can be used as a guide to minimise potential negative impact of such interventions.

User Acceptance

The phased introduction approach and training seemed to enhance user acceptance. All the participants were more comfortable using the robot than in the orientation. For instance, participant 3F noted: “I’m getting quite used to her, and am able to correct my mistakes with her, a good sign of growing confidence.”

In the questionnaires completed in the stage 1 user orientation workshops, the participants had made a number of negative comments about what they would not like the robot to do and why. These included: “I wouldn’t wish to be ‘nagged’ about things I should be doing” (2F), “Wake me up early!” (3F), “Wake me up when not ready to wake. Criticise my mistakes – diplomatic robot needed” (3M). However, at

the end of the stage 2 sessions, the participants were happy to accept reminders from the robot but wanted to have some control over this: “need to be able to set reminder, e.g. remind me at 2 o’ clock” and “take a drink – meal times” (1M, after scenario 4); “a reminder is OK but not repeated too often” (2F, after scenario 5).

It was also noted that despite having considerable problems, participants on the whole were quite forgiving of the difficulties experienced in their interactions with the robot (Figure 7). It was particularly interesting to note that after completing scenario 5 and experiencing a high number of interaction errors, participant 1F gave the maximum scores of 5 for ease of interaction, feeling she could control the robot, which helped her transport things around the house, safely. One of her later entries in her workbook post-task questionnaire said: “Wonderful today to activate Molly myself the first time was unbelievable (sic) as I had really not much knowledge of this.” This finding corresponds to the results reported by Rosenthal-von der Pütten, Bock, and Brockmann (2017) who found that any interaction with the robot in their study increased self-efficacy.

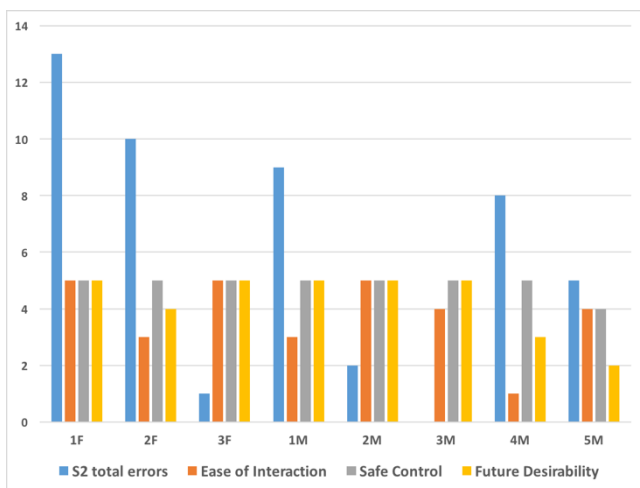


Figure 7 User assigned scores for scenario 2 (Ease of Interaction: 1 Very Hard, 5 Very Easy; Safe Control: 1 Very Unsafe, 5 Very Safe; Future Desirability: 1 No Never, 5 Yes Definitely) in comparison to number of interaction errors for the task.

DISCUSSION

User Learning and Adaptation

Our findings show that even over short trial periods learning and adaptation takes place and users find a way to complete a given task. In our study participants adapted the volume and pace of their voice, as well as their articulation, to improve voice recognition by the robot. They did so on their own, i.e. without prompts from the research team. This finding resonates with earlier research that indicates we humans naturally seek to understand how things behave and respond in different situations and to different stimuli (Mazur, 2015). We are constantly assimilating and building up models of things we interact with, by observations and a trial and error process. Being able to reliably predict responses is reassuring.

In the specific case of interactive systems, Donald Norman (2013) emphasized the importance of the user developing a good mental model for them. To facilitate the development of a mental model, it is critical to make it is clear to the user how a system will respond to the user's explicit or implicit inputs by providing unambiguous and timely feedback. Due to a processing time lag, there were a number of instances where the user spoke too quickly, or the system misheard and responded to something else. Such occurrences prevent the construction of a robust mental model and that negatively impacts learnability. We develop a mental model of how a system functions by identifying how the system changes and behaves in response to our interactions. Here the system's physical characteristics are important, as they give us clues on how to interact with it. When the mode of interaction is not intuitive,

the ability of the user to learn how to use the system is eroded. The more feedback the robot provides to the user on its present status in a timely manner, the better informed the user can be of what is going on, and thereby can pre-empt possible misinterpretations and be more responsible and alert during an interaction.

Taking the blame for errors

Our participants generally tended to take the blame for an unexpected action by the robot and usually thought that they had done something wrong. Low self-efficacy was noted to be higher in the participants who were relatively inexperienced computer users – particularly participants 1F, 3F and 4M.

Better understanding of the constraints of the robot and what it can and cannot do and what it gets wrong could improve recovery from errors. If the robot was able to actively communicate with the user that it was in an error state and perhaps provide clear instructions or have a single button recovery process, it might make it easier to carry out longer trials without the need for an expert to be present. Admittance and recognition of faults and shortcomings, however, might be run counter to the need for building trust in the robot. User acceptance is closely linked to trust and trust is strongly linked to people's willingness to accept the robots recommendations (Salem and Dautenhahn, 2015).

Learning from the user and adapting to the user

The capacity for system flexibility and personalisation are well recognised features of any sophisticated software system. Here our observations in the Wizard mode provided valuable insights. For example, older peoples' voices commonly change and this affects the robot's voice recognition accuracy, even if they try to adapt as well. To forestall such problems, their speech data could be monitored for changes over time and protocols developed for corresponding corrective action.

If it is accompanied by cognitive decline and reduced voice recognition accuracy, particularly after a long period of good performance, it could be a very de-motivating and confusing experience for the user. As we are still quite a few years away from a fully safe and autonomous assistive robot that can be used without expert supervision in a home environment, the more the technology can be made to work collaboratively with the user, the more hope there is of real deployment. While the present research is based on a single case study, it highlights some mismatches between what the participant believes was occurring and what actually occurred. Therefore, in regards to feedback to and from the robot, there are several issues to consider such as the clarity of the feedback by the robot regarding the likely cause of the error and the cognitive load on the user incurred by the process of interpreting and paying attention to what the robot is saying.

As part of the Accompany project, Saunders Burke, Koay, and Dautenhahn (2013) looked at how a computational memory architecture could be used to enable a companion robot to co-learn. In a more recent study as part of the GrowMeUp project, where the robot is being developed as a self-training system, Martins et al. develop a context-aware service selection model. Herein the system has the ability to estimate the user's expectation, assess the degree of satisfaction, and use it as feedback to improve subsequent interactions.

Learning the user's daily routine to understand and recognise unusual behaviours could be extremely valuable for prompting adaptation. Fiorini, Caleb-Solly, Tsanaka, Cavallo, Dario, and Melhuish (2015) show how simple ambient sensor data, from wireless contact sensors on fridge and cupboard doors and power sensors on the kettle, could be used by adaptive machine learning algorithms to develop a personalised model of user behaviour that recognises a specific activity, such as brewing a hot drink. Contextualization of help and guidance messages and reminders communicated to the user by the robot

vis-à-vis current, previous and missed activities could greatly enhance the utility of the assistive support provided.

Improvements in semantic scene analysis (Koppula, Gupta, and Saxena 2013) which can aid in recognising activities of daily living and unsupervised learning of temporal sequential actions (Wu, Zhang, Savarese, and Saxena 2015) are key research developments to support intelligent learning and adaptation.

Enhancing user acceptance

Building familiarity with the robot is key to enhancing user acceptance. Here training sessions for users are essential, starting with a short introduction session and building up slowly to longer sessions. Peoples' attitude regarding acceptance of computers and related technologies has been found to be a good indicator of their ease at learning interactions with robots. The elements of successful training sessions include – a reassuring and relaxed environment wherein the user does not feel anxious or intimidated if the robot does not respond as expected, effective strategies for recovering from unintended states without fuss, upfront demonstration of the existing problems with the technology, and clear instructions for controlling different functions.

Our findings correlate with the conceptual design proposed by Bajones, Huber, Weiss, and Vincze (2014) wherein they considered two key parameters for successful interaction and acceptance – robot pro-activity level and robot presence level, and related these to extrovert or introvert personality types. There have been numerous studies for determining the criteria and factors for enhancing user acceptance of assistive robots such as those by Pino, Boulay, Jouen, and Rigaud (2015) and Frennert and Ostlund (2015), however to truly understand whether these criteria also determine long-term acceptance and

continued use given the pragmatic challenges of everyday life and end-user environments, more longitudinal studies are needed.

CONCLUSIONS

We need to consider pragmatic approaches that would enable longitudinal studies carried out over a period of up to 2 years. They would aid understanding of how familiarity and continued interaction over a period of time influences people's expectations, experiences, and attachment to assistive robots. However, before we are able to do so cost-effectively, safely, ethically, and reliably, we need to be open to more pragmatic approaches that provide a range of perspectives from across a spectrum of users. We need to start by identifying and validating new qualitative and quantitative criteria and evaluation instruments that provide reliable and repeatable multidimensional assessment.

Our multi-point evaluations show that more design consideration could be given to assist people in understanding and working around system constraints. More emphasis could be put on more realistically communicating the specific shortcomings and failures of the system to users, managing user expectations, and providing the means for people to overcome the system's technical constraints. The potential for human adaptation in face of the constraints of the technology needs to be better understood. This potential for adaptation can be tapped into to bypass or mitigate some system software and hardware limitations.

As we work towards designing assistive robots that provide higher levels of support, we are more cognizant of the necessity of mutual awareness by both the user and robot of their current and anticipated status. The robot requires an understanding of the needs and status of the user to initiate appropriate action. Correspondingly, the user also needs to be aware of the status and intentions of the

robot. So, while the robot helps the user, the user can also help the robot (e.g. by speaking clearly and precisely, explaining when there is a fault, inability to proceed with a command or problem with interpreting the user's command with a high degree of confidence,). This also potentially empowers the users, with more cognitive input resulting in a potentially more engaging and pro-active interactive experience. More research effort should be put into developing a robot that could work with the support of the person it was assisting; taking a more cooperative interaction approach, with the person providing as much of the contextual intelligence as she can. Such a co-dependent relationship will compensate for what current robot systems are lacking or could always have difficulties with – dealing with ambiguous environments. Developing more creative and engaging ways for this cooperative interaction are needed, which could result in an interesting journey for both the designers and end-users towards more meaningful, robust, and longer-term assistive robotic trials and applications.

ACKNOWLEDGMENTS

This work was developed within the MOBISERV project partially funded by the European Commission under the 7th Framework Programme (FP7-248434). The authors wish to thank all project members, particularly Anne Allardice and Daniel Davies for their work in the trials, and all the study participants.

REFERENCES

Bajones, M., Huber, A., Weiss, A. and Vincze, M. (2014). Towards more flexible HRI: How to adapt to the user? Paper presented at Cognitive Architectures for Human-Robot Interaction Workshop, Human-Robot Interaction, Bielefeld, Germany, March.

Bartneck, C., Suzuki, T., Kanda, T., and Nomura, T. (2007). The influence of people's culture and prior experiences with Aibo on their attitude towards robots. *AI & Society*, 21(1-2), 217-230.

Caleb-Solly, P., Dogramadzi, S., Ellender, D., Fear, T. and van den Heuvel, H., 2014, March. 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* (pp. 374-381). New York: ACM.

Dick, M.B., Kean, M.L. and Sands, D., 1988. The preselection effect on the recall facilitation of motor movements in Alzheimer-type dementia. *Journal of gerontology*, 43(5), pp.P127-P135.

Ekman, P., and Rosenberg, E.L., (1997) *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. New York: Oxford University Press

European Commission, 2012, Special Eurobarometer 382 / Wave EB77.1 – *TNS Opinion & Social, Public Attitudes Towards Robots Report*. Brussels: European Commission.

Feil-Seifer, D., and Matarić, M. J. (2005). Defining socially assistive robotics. In *Rehabilitation Robotics, 2005* (pp. 465-468). New York: IEEE.

Ferrús, R.M. and Somonte, M.D., 2016. Design in robotics based in the voice of the customer of household robots. *Robotics and Autonomous Systems*, 79, pp.99-107.

Fiorini L, Caleb-Solly P, Tsanaka A, Cavallo F, Dario P, and Melhuish C, (2015). The efficacy of “Busyness” as a measure for behaviour pattern analysis using unlabelled sensor data: A case study. Paper presented at the *IET International Conference on Technologies for Active and Assisted Living (TechAAL)*, London, November.

Frennert, S. and Östlund, B., 2014. Review: seven matters of concern of social robots and older people. *International Journal of Social Robotics*, 6(2), pp.299-310

de Graaf, M.M., Allouch, S.B. and Klamer, T., 2015. Sharing a life with Harvey: Exploring the acceptance of and relationship-building with a social robot. *Computers in Human Behavior*, 43, pp.1-14.

Green, A., Huttenrauch, H. and Eklundh, K.S., 2004. Applying the Wizard-of-Oz framework to cooperative service discovery and configuration. In *Robot and Human Interactive Communication*, 2004. ROMAN 2004. 13th IEEE International Workshop on (pp. 575-580). IEEE.

Huber, A., Lammer, L., Weiss, A. and Vincze, M., 2014. Designing adaptive roles for socially assistive robots: a new method to reduce technological determinism and role stereotypes. *Journal of Human-Robot Interaction*, 3(2), pp.100-115

Huber, A., Weiss, A. and Rauhala, M., 2016. The ethical risk of attachment how to identify, investigate and predict potential ethical risks in the development of social companion robots. In *Human-Robot Interaction (HRI)*, 2016 (pp. 367-374). New York: IEEE.

Huijnen, C., Badii, A., van den Heuvel, H., Caleb-Solly, P. and Thiemert, D., 2011. “Maybe it becomes a buddy, but do not call it a robot”– seamless cooperation between companion robotics and smart homes. In *Ambient Intelligence* (pp. 324-329). Berlin: Springer.

Janneck, M. (2009). Recontextualizing technology in appropriation processes. In B. Whitworth and A. Moor (eds), *Handbook of Research on Socio-Technical Design and Social Networking Systems*, pp. 153-166. Hershey, PA: IGI Global.

Kolb, D. A., Boyatzis, R. E., & Mainemelis, C. (2001). Experiential learning theory: Previous research and new directions. *Perspectives on Thinking, Learning, and Cognitive Styles, 1*, 227-247.

Koppula, H.S., Gupta, R. and Saxena, A., 2013. Learning human activities and object affordances from rgb-d videos. *The International Journal of Robotics Research, 32*(8), pp.951-970.

Martins, G.S., Ferreira, P., Santos, L. and Dias, J., A Context-Aware Adaptability Model for Service Robots. Available at: <http://ap.isr.uc.pt/archive/639.pdf> (accessed January 2, 2018)

Mazur, J.E., 2015. *Learning and behavior*. London: Psychology Press.

McColl, D., Zhang, Z., & Nejat, G. (2011). Human body pose interpretation and classification for social human-robot interaction. *International Journal of Social Robotics, 3*(3), 313-332

Nani, M., Caleb-Solly, P., Dogramadzi, S., Fear, T., & van den Heuvel, H. (2010). MOBISERV: an integrated intelligent home environment for the provision of health, nutrition and mobility services to the elderly. Paper presented at the *4th Companion Robotics Workshop*, Brussels, September.

Norman DA (2013). *The design of everyday things: Revised and expanded edition*. New York: Basic Books.

Nygård, L. and Starkhammar, S., 2007. The use of everyday technology by people with dementia living alone: Mapping out the difficulties. *Aging & Mental Health*, 11(2), pp.144-155.

Pino, M., Boulay, M., Jouen, F. and Rigaud, A.S., 2015. "Are we ready for robots that care for us?" Attitudes and opinions of older adults toward socially assistive robots. *Frontiers in aging neuroscience*, 7. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4512026/> (accessed January 2, 2018).

Rosenthal-von der Pütten, A.M., Bock, N. and Brockmann, K., 2017. Not Your Cup of Tea?: How Interacting With a Robot Can Increase Perceived Self-efficacy in HRI and Evaluation. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 483-492). New York: ACM.

Salem, M. and Dautenhahn, K., 2015. Evaluating trust and safety in HRI: Practical issues and ethical challenges. Paper presented at Emerging Policy and Ethics of Human-Robot Interaction: A Workshop at 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2015), Portland, OR, March.

Sandoval, E. B., Mubin, O., & Obaid, M. (2014). Human Robot Interaction and Fiction: A Contradiction. In *Social Robotics* (pp. 54-63). Berlin: Springer.

Saunders, J., Burke, N., Koay, K.L. and Dautenhahn, K., 2013. A user friendly robot architecture for re-ablement and co-learning in a sensorised home. In P. Encarnação, L. Azeedo, G. J. Gelderblom, A. Newell, and N-E Mathiassen (eds), *Assistive Technology: From Research to Practice*, pp. 49-58. Amsterdam: IOS Press.

Syrdal, D. S., Dautenhahn, K., Koay, K. L., & Ho, W. C. (2014). Views from within a narrative: Evaluating long-term human–robot interaction in a naturalistic environment using open-ended scenarios. *Cognitive Computation*, 6(4), 741-759.

Sun, Ron, Edward Merrill, and Todd Peterson (2001) From implicit skills to explicit knowledge: A bottom-up model of skill learning. *Cognitive science* 25 (2): 203-244.

Sung, JaYoung, Rebecca E. Grinter, and Henrik I. (2010). Christensen. "Domestic robot ecology." *International Journal of Social Robotics* 2(4): 417-429.

Weiss, A., Bernhaupt, R., Lankes, M. and Tscheligi, M., 2009. The USUS evaluation framework for human-robot interaction. In *AISB 2009: Proceedings of the Symposium on New Frontiers in Human-Robot Interaction* (Vol. 4), pp. 11-26. Red Hook, NY: Curran Associates

Wu, C., Zhang, J., Savarese, S. and Saxena, A., 2015. Watch-n-patch: Unsupervised understanding of actions and relations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4362-4370). New York: IEEE

Zimmerman, J. (2009). Designing for the *Self*: Making products that help people become the person they desire to be. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 395-404). New York: ACM.

NOTES

¹ Weiss et al. (2009) put forward a very comprehensive evaluation framework that has multiple dimensions, including usability, social acceptance, user experience and societal impact. The latter two dimensions in particular would be critical for longer-term trials.

² <https://www.techsmith.com/morae.html> MORAE User Experience Analysis Software