



Research Article

Open Access

Kheng Lee Koay*, Dag Sverre Syrdal, Kerstin Dautenhahn, and Michael L. Walters

A narrative approach to human-robot interaction prototyping for companion robots

<https://doi.org/10.1515/pjbr-2020-0003>

Received April 17, 2019; accepted November 6, 2019

Abstract: This paper presents a proof of concept prototype study for domestic home robot companions, using a narrative-based methodology based on the principles of immersive engagement and fictional enquiry, creating scenarios which are inter-connected through a coherent narrative arc, to encourage participant immersion within a realistic setting. The aim was to ground human interactions with this technology in a coherent, meaningful experience. Nine participants interacted with a robotic agent in a smart home environment twice a week over a month, with each interaction framed within a greater narrative arc. Participant responses, both to the scenarios and the robotic agents used within them are discussed, suggesting that the prototyping methodology was successful in conveying a meaningful interaction experience.

Keywords: prototyping, human-robot interactions, user studies, social robotics, robot companions

1 Introduction

The ongoing challenge in developing complex, future and emerging technologies is that of eliciting *meaningful* information and feedback from the potential users of these technologies. From a user-centred design perspective, it is clear that the earlier potential end users can influence the development of a system, the more impact they will potentially have on the end result. As such, early input from

potential users of an emergent technology, such as multi-purpose home companions robots, is invaluable.

Some of these insights can be obtained through eliciting requirements from different stakeholders in the adoption of such technologies (see Bedaf et al. [1]). However, these endeavours often struggle with unrealistic expectations on the part of the stakeholders as to what the capabilities of the technologies are [2]. This concern can be met to a certain extent by actively guiding the expectations of the participants.

Frennert et al. [3] for example, conducted a study to give participants a greater insight into what sharing a home environment with a robot could be like. They used pictures of robots and different kinds of materials (e.g. fabric, wood, metal) that senior citizens could use to create their ‘perfect’ robot. However, given the physicality of a companion robot, and the range of situations in which its presence could be impacting, there are limits to this approach.

It is also possible to obtain insights from laboratory-based experiments as recommended by Bethel and Murphy [4]. However, there is a danger that when investigating the use of a technology outside of its particular use-context, the results may lack in ecological validity compared to situations outside of the laboratory. One of the clearest examples of this is the 1990s Microsoft Office Assistant, ‘Mr Clippy’ which was a system developed on the basis of examining over 25,000 hours of usage of the products that it was intended to assist with [5]. However, as noted by Whitworth [6], the lack of examining the assistant’s impact on users’ experiences of these applications in a holistic, situated manner, led to an almost wholesale rejection of this assistant amongst Microsoft’s customer base.

However, constrained approaches may be of use, in particular when addressing very specific aspects of human-robot interaction. In our work in the University of Hertfordshire (UH) Robot House we have employed such approaches for investigating several specific developments of robot companion technology. For instance, the use of shared memory visualisations [7] as well as an interface for end-user personalisation [8]. Nevertheless, such

*Corresponding Author: Kheng Lee Koay: Department of Computer Science, University of Hertfordshire; E-mail: k.l.koay@herts.ac.uk

Dag Sverre Syrdal: Department of Computer Science, University of Hertfordshire; E-mail: dagsyrdal@gmail.com

Kerstin Dautenhahn: University of Waterloo, Canada; E-mail: kerstin.dautenhahn@uwaterloo.ca

Michael L. Walters: Department of Computer Science, University of Hertfordshire; E-mail: m.l.walters@herts.ac.uk

reductionist approaches often fail to take into account the complexities of the human-robot companion experience.

This issue is particularly important when examining the role and use of companion robots [9] in environments that belong to the private and domestic life of their users. Here, the scope for the use-context is very wide and complex.

A third approach is to deploy prototype robots into the homes of potential users. This is an approach which is becoming increasingly viable as domestic robot platforms for consumer use are becoming more common. This does allow for a close to perfect degree of ecological validity in terms of participants' interactions with the robots [10, 11]. It also allows for investigations into the dynamics that may occur in such interactions and can inform the design of future systems [12, 13]. However, as a means of prototyping, these efforts are performed at a late stage of the design process when typically much of the development work has already been completed, and thus leaves little scope for changing the system based on the users' feedback.

Rather than obtaining user feedback on a system that is already close to a final product, it would be beneficial to allow for the most salient information being given to participants at the earliest stage of the design cycle as possible as part of the prototyping process. The development and evaluation of prototyping techniques that facilitate this process is important to the field of human-robot interaction.

The research reported in this article is complementary to other related projects involving end-users in care and assistance scenarios. A recent example is the GrowMeUp project which investigated how robots can adapt to older users' changing needs and preferences using machine learning approaches [14, 15]. The InStance project investigated fundamental issues of social attunement when people interact with robots, including the sense of agency and gaze cueing effects. Experiments are conducted in very constrained setups, using neuroscientific methods [16, 17]. In this article we follow a narrative framing approach towards gaining feedback from participants on home companion robots. The robots used are complex prototypes with multiple functionalities, including the ability for the (robot) agent's 'mind' to migrate between different embodiments.

The remainder of this article is structured as follows. Section 1.1 discusses engagement with home companion robots. Sections 1.2 and 1.3 motivate our approach to use narrative framing in order to facilitate immersive experiences, leading us to identify and adopt principles for engaging prototyping methods that are presented in section 1.4. Based on this context, we derived the research ques-

tions for the present study (section 1.5). In section 2 we describe the companion robots used in the study and the environment in which they were situated. Section 3 describes how we created the narrative scenarios and the overall narrative arc with the aim to immerse the participant, and to transform their exposure to the system into an engaging user-experience. Section 4 describes how the scenarios were used in our study. Section 5 presents quantitative and qualitative results from our long-term study. A discussion of these findings, limitations of the study and future work concludes the article (section 6).

1.1 The personal robot companion — socially acceptable and relationship-building

Getting informed feedback at all stages from potential users is important when developing technologies intended to be used as companion robots in domestic environments. Companion robots [9], intended to be able to provide assistance and companionship in different contexts in a domestic environment could potentially impact a wide range of interactions over a period of time, which might span years. As such, the impact on the user's experience could be severed if a companion robot performs in a way that is not socially acceptable, or if some of its behaviours impede certain tasks. Moreover, complex home companion robots are likely to be expensive and thus likely to be offered as part of a health or social care intervention, to be used by both the residents of a home as well as formal and informal carers [18]. In these situations, the resident may not have as much say in the specifics of how a robot is deployed in their home. This makes prototyping for social acceptability by the end-user at *all* stages of the technology development process even more important.

Home robot companions, as suggested by our previous work in the UH Robot House [19], while offering some physical assistance (i.e. fetch and carry), can primarily be classified as *socially assistive*. Socially assistive robots are defined by Feil-Seifer and Matarić [20] as robots that *provide assistance through interaction, without physical contact* (*ibid.* p.46). This assistance can be realised, for example by encouraging healthy behaviours, such as engaging in physical activity or performing specific exercises. It can also involve reminding the user to perform certain tasks such as self-care or preparing food. Several studies have suggested that the ability of a socially assistive agent to form a relationship with its user is important to maintain its continued use [21]. There is also some evidence suggesting that this applies to socially assistive robots as well [10].

Based on this evidence, a prototyping process for companion robots should attempt to investigate the role of such human-robot relationships within interactions.

For both the acceptability and relationship dimensions of home robot companions, there is also a strong temporal element. Relationships, by their very nature, change over time. Behaviours that seem engaging and interesting at first may soon become frustrating and annoying. Conversely, initial difficulties in interaction with the robot may also be smoothed over through the user adapting over time to the robot's idiosyncrasies.

1.2 Narrative – the story of the companion

Previously, we have adopted a narrative approach to prototyping interactions with robot companions in a domestic environment [19]. Fidelity in prototyping is often considered as a function of how closely the physical prototype resembles a completed, market-ready, product. Bartneck and Hu argue that due to the relative novelty of robots, such robot prototypes should have as high a fidelity as possible [22]. Participants often do not know enough about how they will respond to any given robot based on a low-fidelity prototype (e.g. a cardboard mock-up or a written vignette). Vlachos et al. [23] found that participants often changed their minds about their preferences regarding a particular robot after a brief physical interaction, compared to their stated pre-interaction preferences.

However, fidelity should not necessarily be constrained to the physical nature of the robot. It should also reflect the experience of using the robot within its intended setting as much as possible. Dindler and Iversen proposed Fictional Inquiry [24] where they suggest an approach to "create partially fictional situations, artifacts, and narratives that mediate collaborative design activities". Our experimental approach is based on this idea of utilising a *narrative framing technique*, which incorporates a spoken or textual narrative at the beginning of each session to mediate and introduce the intended setting and context of use for the specific forthcoming human-robot interaction [19]. We have previously created a set of scenarios in which participants were invited to engage in a set of episodic play sessions [25] in which they were asked to *play-act* their interaction with the robots as if they were the owner.

In our present study, the spoken narrative is used to set the context of the interaction session and to draw on the usage scenario as the basis for the narrative, using the robots and the environment (Robot House) itself as props for the emergent interactions.

These scenarios were grounded in use-scenarios and originally developed through iterative considerations of typical user-personas ([26], [19]). This allowed the individual scenarios to address the technology being developed, but also the projected types of interactions prospective users could be expected to engage in.

The narrative allows participants to be gently guided by the robot's partially scripted responses into the desired interactions. Similar to an interactive novel, the participants were free to select or reject different options and actions, and to vary the order and time of these events occurring within the scenario.

1.3 From narrative to immersive engagement

The work we presented in Syrdal et al. [19] demonstrated how to successfully use personas to drive the creation of narrative scenarios which would allow for prototyping companion robot interactions in an ecologically appropriate environment. However, the episodic nature of the scenarios made it difficult to convey the long-term aspect of these interactions. Also, even though the episodes were narratively framed within themselves, the events in any single episode did not impact later episodes. This *lack of chronological coherence* made the break in the immersion between episodes more marked, as the transitions between episodes relied purely on the narrative frame provided by the researchers, rather than on the expectations and remembered events of the participants.

In order to mitigate this break, we here propose the concept of a *coherent narrative arc*, in which the interactions with the robots are conducted against the backdrop of a continuous interactive scenario that engages the participant. The narrative arc is the spine from which the individual scenario sessions (events and interactions) occur within an overall story-line and which follows Freytag's dramatic stages [27]. This provides a continuity within individual sessions and also between successive sessions.

In this manner, the episodic interactions have the potential to become *engagist* scenarios. John Tynes [28] defines engagist scenarios as those that provide '*...tools and opportunities for participants to explore and experiment ...in ways that real life may prohibit or discourage.*' This makes them very suitable for prototyping future and emerging technologies, as by their very nature they cannot yet be experienced in real life. Experiencing these interactions with a prototype system in this manner can be a powerful source of insight.

The way that engagist media operate, Tyne argues, is through *immersion*. Immersion is, according to Bowman and Standiford [29], a multifaceted construct, but can be described as sharing the experience of an imagined or fictional self in an imagined or fictional situation. However, unlike Syrdal et al. [19] who considered immersion as a function of fidelity, Standiford and Bowman [30], when considering medical simulations, address this issue separately. Fidelity, they argue, is exclusively a function of how closely the interaction *physically* resembles the situation that is being simulated. Immersion, on the other hand provides the participant with a deeper and more visceral response to the simulated situation. Basing themselves on the work of Harviainen [31], they define immersion as having three facets:

1. Reality Immersion

- *The extent to which the participant accepts the given scenario in the scenario space – i.e. the participant accepts and acts in accordance with the notion that they are the owner of the house in which the interactions with the robots take place for the duration of the interaction scenario.*

2. Character Immersion

- *The extent to which the participants experiences or ‘channels’ the responses or the feelings of the character they are portraying within the interaction – i.e. the participant experiences a changing relationship with their robot companion*

3. Narrative Immersion

- *The degree to which the participant accepts the narrative surrounding the interaction such as outside events, chronological changes, etc. – i.e. the participant accepts that events in any one session may be causally related to events in a previous session.*

This approach immerses the *participant* in the study to the extent that they share some of the experiences of that of a real *user* of the system, and is thus hoped to make their responses to the system within the scenario directly relevant to the development and the deployment of future and emerging technologies.

There is of course a danger that the effort to create an immersive, engaging scenario leads to including elements that are not supported by the technology. It is important for the purpose of prototyping that the immersive, engagist interaction scenarios are grounded in a realistic projection of the systems capabilities. Otherwise any results may be less a reflection of the potential of the technology being developed, but more a reflection of the pre-conceived no-

tions that the researchers and participants might have of what a robot *should be* [32].

1.4 Principles for engagist prototyping

Based on these considerations we adopted the following principles in the planning and execution of our prototyping study of a robot companion:

1. Interactions must be ...
 - (a) ...grounded in the actual technological development.
 - (b) ...motivated within the scenario.
 - (c) ...situated within a coherent timeline.
2. Participants should...
 - (a) ...treat the environment as their own.
 - (b) ...interact with the robot to achieve their own goal.
 - (c) ...be able to personalise the technology and its behaviours.
3. Technology should ...
 - (a) ...be based on a realistic projection of the system’s development.
 - (b) ...impact the narrative in which it is situated.

The current study investigated how these principles can be used to guide a Human-Robot Interaction (HRI) user study regarding the development of technologies to support a home robot companion.

1.5 Research questions

The research questions focused on two aspects: Scenario acceptability and human-agent relationships.

Research Question 1 – Do users accept scenarios inter-connected through narrative?

How acceptable do the participants find the overall scenario and narrative presented throughout the study? In order for the prototyping to be of use, participants should be able to draw on both the events in the individual scenarios and relate them to their everyday lives. In addition, do participants see such a system as being suitable to themselves and others? What is the reasoning behind these judgements?

Research Question 2 — Does the user-agent relationship change when the agent migrates to different embodiments?

As the participants interact with the agent across different robot embodiments in a series of inter-connected scenarios, how do they perceive their relationship with the agent? Would the relationship between user and agent improve or worsen? Also of interest is how participants reason about their feelings of closeness towards the agent during the long-term study.

2 The UH Robot House and the Sunflower robots

2.1 The UH Robot House

The study was performed in the UH Robot House, which is a typical British residential house located just outside the UH campus. It is mainly used for conducting research and user studies in the area of Smart Homes and Robotic Home Companions. The interior of the house is decorated with furniture, paintings and appliances found commonly in a typical home, in order to provide an ecologically valid environment for participants who take part in studies. For the purposes of this study, the UH Robot House was equipped as a smart home with two commercially available sensor systems; a Green Energy Options (GEO) System and a ZigBee Sensor Network. The set-up provided more than 50 sensors, embedded in the Dining Area, Living Room, Kitchen, Bedroom and Bathroom of the Robot House (see Figure 1), which supported detection of the user's activities relevant to daily living activities. Readers interested in the system that integrates the GEO System and ZigBee Sensor Network can refer to [33] and [34] for more information. This study relied primarily on the GEO System, which will be described briefly.

2.2 Robotic platforms

The participants interacted with an agent through several different robotic platforms. The agent used a process called *migration* to transfer its 'mind', that is, its memory, current task context and personalisation information, between the different platforms. For examples of migration, see e.g. Koay et al. [35] or Segura et al. [36]. Our motivation to include the mechanism of migration in our studies was based on the concept of one companion agent that may move between different robot embodi-



Figure 1: UH Robot House map showing the location of sensors (identified by numbers) and their states with green colour representing the sensor in open/on/free state, red colour representing the sensor in closed/off/occupied state and transparent representing the sensor in a not activated/unknown state.

ments while maintaining its memory and interaction history. Such changes in embodiment may be necessitated for example, by a breakdown of a specific robot platform or the need to access different functionalities and capabilities of another robot platform [37–40]. Moreover, future companion robots cannot be expected to have all functionalities that a user might want at that particular time or in future. We can envisage situations involving a range of robotic systems, some with a single functionality (e.g. vacuum-cleaning), others more complex with multiple functionalities (e.g. companion robots) that will be used in homes. Rather than requiring users to learn to interact with, use and personalise multiple different robotic systems, the concept of migration allows the user to only interact with one agent at a time; an agent that retains its 'mind' when migrating. Such complex home companion ecosystems are not yet widespread in the real world. However, technologies are developing at a fast pace, and so we can expect such systems in the near future. Prototyping such complex technologies thus allows us to engage and inform the development of such future systems.

The robotic platform used was the UH Sunflower Robot (see Figure 2). This robot was designed and developed to be a highly expressive robotic platform, in that it has four different non-verbal communicative channels: multi-coloured light signals, sound - MIDI (Musical Instrument Digital Interface) tunes, and the independent movements of its head and body. These modalities are used to create expressive multi-modal behaviours to communicate



Figure 2: The UH Sunflower Robot SF1 – (left) extends its GUI for interaction, (top-right) performing attention seeking behaviour to attract user's attention, (lower right) exhibits a non-verbal dog-inspired looking back behavior used in [41].

intent, such as to attract the attention of the user or to provide simple non-verbal feedback during interactions. The non-verbal communication signals used are prescribed sequences of concurrent actions. For example, the attention seeking behaviour was inspired by dog behaviours, and resulted from collaborative research involving ethologists [41]. Here, the robot (base) was scripted to move forwards and backwards repeatedly, while its head tilted up with panning left and right quickly, simultaneously with its LED displays blinking green and the robot audio system playing a tune via its MIDI sound system. The implementation of the non-verbal communication signal was similar to a notification concept used in some mobile phones, whereby different LED colours or sounds are associated with different types of notification (e.g. voice call or text message). This allows the robot to exhibit its awareness of the environment and its relationship with the user, hence fulfilling one of the requirements of social interaction. This compliments Lanillos, Ferreira and Dias [42] work which focused mainly on the implementation of automatic attention mechanisms to support social interaction.

Koay et al. [41] have successfully utilised the Sunflower Robot in their study which explored the effectiveness of dog-inspired non-verbal expressive behaviours as visual communication signals for robots to communicate intent. The Sunflower Robot uses a Pioneer P3-DX wheeled base, which was commercially available from Omron Adept Mobile Robots, for mobility, on which a square body and cylindrical head was mounted. The 'shoulder' (i.e. the top of the square body) has been equipped with a display of diffuse LEDs and there is a drawer that slides out to become a carrying tray. In addition, the front of this drawer has been mounted with a tablet computer running an integrated Graphical User Interface (GUI). This allows the

robot to slide out its drawer/tray to give the user a better access to the GUI when initiating explicit two-way communication with the user through its menu system. The robot's head is articulated with four degrees of freedom (roll, pitch, yaw, and extension/contraction movement) and its 'face' is non-animated, with two static white 'eyes' and a web cam appearing as its nose. Four degrees of freedom afford the robot head the ability to perform a variety of realistic head and gaze based gestures such as head nods and shakes, gaze alternation between user and target object, and in combination with body motion to produce spatio-temporal gestures such as the looking back behaviour (see Figure 2, lower right) which illustrates the pose used by the robot to express a follow-me intention.

This study used three different Sunflower variants, and one or two variants were used for each session. The standard Sunflower was designated SF1. SF2, a stationary Sunflower robot, and was identical to SF1 except that its base was not mobile and had a Skype compatible handset in the automated slide out drawer. SF3, the replacement Sunflower, was identical to SF1, but did not have an articulated head, and the tablet GUI was mounted on its 'shoulder' (See Figure 3 and 4). The companion agent would migrate between these different embodiments within the sessions. See Table 1 for a summary of the differences between the three embodiments of the companion robot.

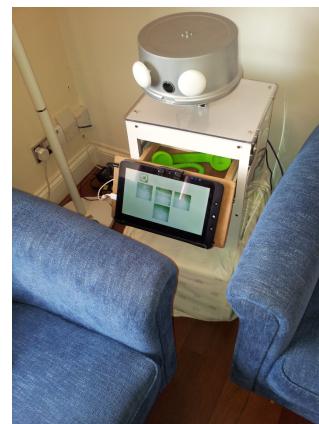


Figure 3: The Stationary Sunflower Robot (SF2) extends its tray for the user to answer a Skype call.

2.2.1 Capabilities within the scenario

The Sunflower robots are integrated into the Robot House's computational infrastructure and as such have the competencies required to navigate autonomously and detect user



Figure 4: The Replacement Sunflower Robot (SF3) extends its tray and offers to carry an object for the user.

Table 1: Differences between the robots.

Robot	Mobile	Stationary	Replacement
	SF1	SF2	SF3
Navigation	Yes	No	Yes
Expression: Body movement	Yes	No	Yes
Expression: Head movement	Yes	Yes	No
Expression: Flashing lights	Yes	Yes	Yes
Expression: Sounds (MIDI tunes)	Yes	Yes	Yes
Tray movement (extent and retract)	Yes	Yes	Yes
Interaction GUI	Yes	Yes	Yes
Skype calling	No	Yes	No

activities and other events based on the sensors used in the Robot House. This allows them to provide cognitive help (i.e. inform the user of events occurring, remind the user of plans, display messages) as well as physical assistance (carry things in their trays). For example, they would be able to detect the user switching on the kettle or opening a refrigerator door via the GEO system. They would also be able to detect the doorbell ringing, and alert the user if there was an incoming Skype call. Based on these detected sensor events, the robots would perform the appropriate task associated with the event.

To give an example of assistive behaviour, if the doorbell was ringing, the robot would then approach the user and first perform an attention seeking behaviour. This normally involved the blinking of the diffuse LED in a particular color, accompanied by MIDI sounds and biologically inspired movements of both head and body [41]. Next, the robot would extend its tray and display a message stating that the doorbell had been rung, with options for the user

to request that the robot accompanies them to the front door, or that the doorbell is to be ignored.

In addition, the user can also initiate interactions with the agent (robot) via the tablet GUI, enter their preferences (for preferred food, drink and activities), as well as personalise the expressive behaviours of the robot. These preferences and personalisation settings are retained by the agent across all its embodiments. Depending on the task that the agent has been asked to perform, it will migrate to a more appropriate embodiment to perform its task. For instance, if the agent was asked to move to the kitchen with the user while in the Stationary Embodiment (SF2), it would ask for permission to migrate into the Mobile Sunflower (SF1). Likewise, if the user received a Skype call, while the agent was in embodiment SF1, it would ask for permission to migrate into SF2 to assist the user in answering the call.

3 Creating an immersive scenario

3.1 Overview

The contents of the scenarios created to explore the issues of relationship and acceptability could be separated into two categories, both of these categories would support *immersion*. Based on the purpose they served they could be ...

1. ... drawn from the technology and its use.
2. ... created to support the narrative arc.

3.2 Technology-driven aspects

These aspects concern the technical functionalities and capabilities of the particular technology used which are limiting the choice of scenarios that can be implemented. The technical capabilities discussed in section 2 were deployed in two use-scenarios. The first was a morning scenario, in which the robots would remind the user of their preferences for breakfast and accompany them to the kitchen. The second was a lunch-scenario, where the participant would be reminded of their preferences for lunch. Within these, the robot would respond to events, such as the kettle having finished boiling, the toaster finishing, newspaper deliveries and phone-calls, inspired by the content of the open-ended scenarios described in Syrdal et al. [19].

3.3 The narrative arc

The narrative arc through which the participants were exposed to the use-scenarios of the robots was intended to give the participants an overall context for the use of the robots. This narrative set the participant in the role of a recent owner of two robots, who learns how to use the robots in their own daily life. Once they get used to the robots, however, the mobile robot starts to act erratically and breaks down. The robot is taken away and a new replacement robot is provided to the participant. After getting used to interacting with the new robot, the original robot is repaired and returned to the owner, who is reunited with their original robot companion. In the context of our multi-functional, complex robot home companion, the choice of a narrative around how users responded to a possible breakdown of the hardware, seemed realistic for early adopters of such technology.

The structure of the narrative arc was loosely based on Gustav Freytag's 5-act structure [27]. This approach argues that successful narratives have 5 parts, *Exposition, Rising Action, Climax, Falling Action* and *Conclusion*.

3.3.1 The exposition phase

This phase is intended to set the scene for the audience, to provide them with the information that they need to understand the subsequent dramatic tension in the narrative. In addition, when considering interactive media such as games or interactive fiction, this phase is also intended to provide the audience with the understanding of how to interact with the narrative space and the interfaces for doing so [43]. Due to the complexity of interacting with novel robot technologies, this phase lasted for four sessions. Participants interacted with the system during these sessions without any complications in order to get used to the system in daily life activities. Freytag's structure ends this phase with what is termed the *inciting moment*. In this arc, this phase culminates with the mobile robot breaking down.

3.3.2 The rising action phase

Here, the participants were still interacting with the system and engaging in the breakfast/lunch scenarios. The complications caused by the lack of a mobile robot would have to be dealt with. However, in addition they had to liaise with a technician (a researcher took the part of this

role) in order to negotiate the removal of the faulty robot and the delivery of their temporary replacement robot.

3.3.3 Climax

The Climax of this part of the narrative was the arrival of the replacement robot. At this point the user had to transfer their preferences to the new robot, and then start interacting with the replacement.

3.3.4 The falling action phase

This phase consisted of the participant interacting with the new robot within the scenarios. Once the participant was informed that the original robot was repaired, they then organised the removal of the replacement robot and the delivery of the original.

3.3.5 The conclusion

The conclusion was given as the return of the original robot. The participant would then transfer their preferences back into the original robot and have a final interaction session with their original robot. We also considered the debrief session to be part of this phase as this session allowed for an exploration of the participants' reaction to the narrative.

3.4 Immersion within the scenarios

These scenarios ensured that the three tiers of Harvinainen's [31] model of immersion were supported.

3.4.1 Reality immersion

Reality Immersion was supported by several aspects of the scenarios. Care was taken to ensure that the three rooms used in the scenario (Living room, kitchen and front hall) were all exclusively used for the interaction scenarios. Debriefing was conducted in a separate room. All events in the scenarios were supported by props. There was real food in the kitchen, deliveries were pushed through the letter box, phone calls were channeled through the Sunflower robot etc. Exceptions to this were made explicit and delineated clearly (see Figure 5) to the participant.



Figure 5: Sign instructing participants what to do in case they needed help.

3.4.2 Character immersion

Character Immersion was supported by ensuring that the participants had an understanding as to why they were engaging in a given situation. The food and entertainment options suggested by the robots were based on the preferences of the participants. They ate and drank real food and beverages, and they engaged in intrinsically rewarding entertainment activities.

3.4.3 Narrative immersion

Narrative Immersion was supported by providing the participant with a briefing outlining the situation prior to the beginning of each session. In addition, there was a continuity of events, so that events in one session would impact events in later sessions (for instance, the robot breaking down in Session 4, led to it being taken away in Session 5), and plans were made and carried across sessions. This supported the participants' immersion into the narrative flow of the scenarios.

4 Methodology

4.1 Participants

Nine participants took part in this study, 6 female and 3 male. They were recruited via advertisements on the

Table 2: Sessions in the study.

Freytag's 5-act structure [27]	Session theme	Robot involved
Exposition	1 Tutorial	SF1, SF2
	2 Setting the Scene: Habituation	SF1, SF2
	3 Setting the Scene: Habituation	SF1, SF2
	4 Inciting incident: Robot breaks	SF1, SF2
Rising Action	5 Mobile robot removed	SF1, SF2
	6 Interaction with only stationary robot	SF2
Climax	7 Replacement arrives	SF2, SF3
Falling Action	8 Interacting with replacement robot	SF2, SF3
Conclusion	9 Original mobile robot returned	SF1, SF2
	10 Evaluation	

University of Hertfordshire Intranet[†]. The participants were between 21-32 years of age with a median age of 25 years. One of the participants dropped out of the study after session 6, but their responses up until this point were retained in the analysis.

Participants attended two sessions per week over a month. Since each session lasted about one hour, with additional time beforehand and afterwards to set up the system, charging the robot etc., accommodating 18 experimental sessions during a working week was stretching the available resources to the maximum, which shows the limitations of carrying out long-term HRI studies with complex scenarios.

The schedule for the sessions is shown in Table 2.

4.2 Procedure within the sessions

In the first session, the experimenter who acted as facilitator welcomed the participants to the Robot House, introduce himself and a second experimenter whose responsibility was to monitor the systems during the trials from a small adjoining office (a converted bedroom not used in the study). This second experimenter also took on the role of the technician during the scenarios when needed. The participants were then introduced to the Robot House and

[†] The use of human participants had been approved by the University of Hertfordshire Ethics Committee under protocol number 1112/39.

shown how to use the house's electrical appliances, where the food was kept, where the drawers and cupboards for cutlery and plates were and so on. After this, the participant completed a consent form and brief demographic questionnaire.

Each interaction session began with an introduction to the session, which was intended to ground the interaction within the overall narrative and provide a context for the participant. An example narrative is provided below:

In the introductory session you gave us some preferences for what you like to do in the early morning. Your robotic companion has these preferences and will apply them when interacting with you.

Now imagine that you have woken up in your bedroom. When you are ready, you will come out of your bedroom, sit down on the sofa, and log in to the robot with your user account and password. The robot will then begin today's session.

The facilitator would then ask if the participant had any questions about the session, and after answering any questions in an appropriate manner, he would leave and go to the facilitator room, allowing the participant to conduct the interaction alone with the robots. Throughout the interaction, the technician monitored the interaction through networked cameras to ensure the safety of the participant.

The human-robot interactions then took place. The robot and the participant would interact with each other throughout the scenario without any involvement from the researchers. After the above briefing, the interaction began with the agent using SF1 to approach the participant and suggest breakfast and a hot drink. It reminded them of the toaster and kettle having finished, and also alert them to a newspaper delivery.

After the interaction, the participant then met with the facilitator in order to complete a series of post-interaction questionnaires. They also had the opportunity to discuss their experience with the facilitator. The session would then end with the facilitator and participant arranging a time for the next session.

4.3 Measures

There were several measures used to address the research objectives.

4.3.1 Research Question 1 — Do users accept scenarios inter-connected through narrative?

A Scenario Acceptance Scale was used to measure the participants acceptance of the scenario as well as the role of the companion within it [19]. It consists of ten different 10 point Likert scale ratings which were combined in order to allow for scores between 0 and 100.

The participants were also given the opportunity to respond to Likert scale ratings for the suitability of the companion for themselves, as well as for someone else who was elderly and/or disabled. The qualification regarding the use of the companion for someone that is disabled or elderly, was made to reflect the fact that decisions regarding the deployment of technologies as part of a medical or assistive intervention is typically made by third parties in order to address specific needs [18].

Finally, participants were given the opportunity to provide an open-ended response as to why they felt the robot was, or was not suitable.

4.3.2 Research Question 2 — Does the user-agent relationship change when the agent migrates to different embodiments?

The participants' feelings of closeness to the companion was measured using the Inclusion of Other in a Self questionnaire [44]. This is a pictorial scale of closeness which allows respondents to describe their relationship with an 'other' by selecting one from a series of Venn-like diagrams that overlap to varying degrees, with feelings of closeness being associated with greater overlap. It has previously been used in HRI to gauge affective reactions between children and a migrating agent [36].

Participants were also asked to contrast their feelings towards the companion in the current session with how they viewed their feelings in the previous session, by marking their ratings on a 5-point semantic differential scale.

Finally, participants were asked to explain their reasoning behind their responses to the semantic differential scale in an open-ended question.

5 Results

5.1 Quantitative results

5.1.1 Acceptability

Acceptability was assessed using the Scenario Acceptance scale in addition to two single-item Likert scales, which rated their desire to own such an agent in their own life, as well as whether or not they thought these agents might be suitable for others.

Table 3 and Figure 6 show how the non-standardised scenario acceptability scores changes over time. It suggests that there is a small increase across the sessions, which may be an indicator of increased acceptance of the scenario as the study progressed. However, there was a large variability between participants, and this trend was not significant (Friedman's $\chi^2(7) = 3.10, p = 0.88$). As such, this trend cannot be taken to support the notion that participants became better able to relate the experienced scenarios to their own everyday lives. Overall however, the participants did rate the scenario quite highly along this dimension for *all* sessions. The difference between the observed scores, and the score of 50 (which is what one would have seen if a participant were to respond with a 'neutral' response to every question in the scale). It can be seen that responses were significantly higher than 50 along all the sessions (Wilcoxon's $p < 0.05$), which suggests an acceptance of the scenarios inter-connected through narrative.

Table 3: Scenario Acceptance Scores.

Session	Mean	SD	Median	25th Perc.	75th Perc.
Session 2	71.88	21.41	76.25	57.50	86.25
Session 3	73.12	23.37	76.25	63.75	88.12
Session 4	71.25	26.49	77.50	58.75	88.12
Session 5	73.12	20.47	73.75	62.50	89.38
Session 6	75.31	23.05	76.25	67.50	93.75
Session 7	76.56	27.15	83.75	62.50	100.00
Session 8	76.25	28.13	81.25	66.88	98.12
Session 9	78.44	24.13	81.25	67.50	100.00

Table 4 and Figure 7 show the unstandardised descriptives for the responses to the item asking participants whether or not they wanted a robot across the sessions. It suggests that overall there were small differences between sessions, but which were not significant (Fried-

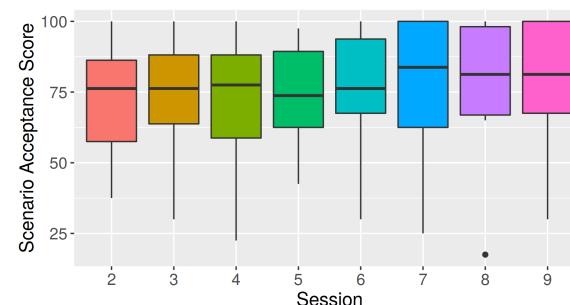


Figure 6: Responses to Global Evaluation measures across sessions for Scenario Acceptance.

Table 4: Robot for Self.

Session	Mean	SD	Median	25th Perc.	75th Perc.
Session 1	4.00	1.20	4.5	3.00	5.00
Session 2	3.62	1.41	3.5	3.00	5.00
Session 3	3.88	1.36	4.0	3.75	5.00
Session 4	3.75	1.39	4.0	3.00	5.00
Session 5	3.38	1.41	3.5	2.75	4.25
Session 6	3.88	1.36	4.0	3.75	5.00
Session 7	4.00	1.41	4.5	3.75	5.00
Session 8	3.75	1.39	4.0	3.00	5.00
Session 9	3.88	1.46	4.5	3.00	5.00

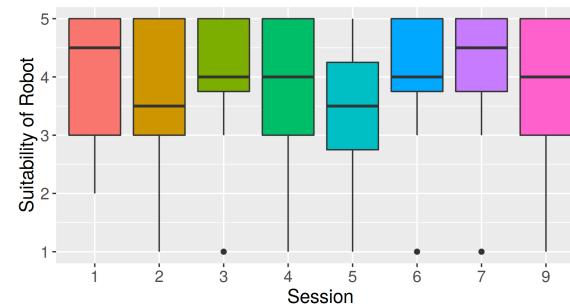


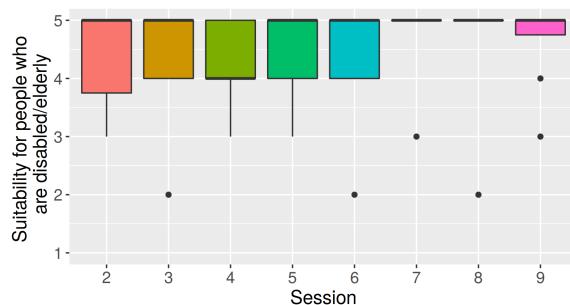
Figure 7: Responses to Global Evaluation measures across sessions for Robot for Self.

man's $\chi^2(8) = 6.53, p = 0.59$). In addition, the overall scores were only significantly higher than the 'neutral' score of 3 in Session 1.

Table 5 and Figure 8 show the non-standardised descriptives for the responses to the item asking participants whether or not such a robot was suitable to help someone who was disabled or frail. There were small differences between sessions, and these were not significant (Friedman's $\chi^2(7) = 9.15, p = 0.24$). Across all sessions, however, responses to this item scored significantly higher than would

Table 5: Robot for Others.

Session	Mean	SD	Median	25th Perc.	75th Perc.
Session 2	4.38	0.92	5	3.75	5
Session 3	4.38	1.06	5	4.00	5
Session 4	4.25	0.71	4	4.00	5
Session 5	4.50	0.76	5	4.00	5
Session 6	4.38	1.06	5	4.00	5
Session 7	4.75	0.71	5	5.00	5
Session 8	4.62	1.06	5	5.00	5
Session 9	4.62	0.74	5	4.75	5

**Figure 8:** Responses to Global Evaluation measures across sessions for Robot for Others.

be expected than if participants were to respond with a ‘neutral’ response of 3 (Wilcoxon’s $p < 0.05$).

It appears that from the start, the participants did seem to accept the scenarios they engaged with as meaningful to their own experience.

5.1.2 Relationship

5.1.2.1 IOS Scores

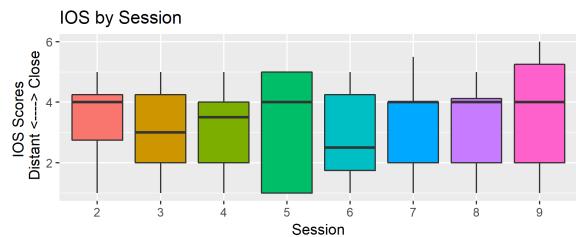
The IOS ratings presented in Table 6 and Figure 9 suggest that there were no significant differences between the absolute IOS ratings across the different sessions (Friedman’s $\chi^2(7) = 7.33, p = 0.39$).

5.1.2.2 Relative Closeness

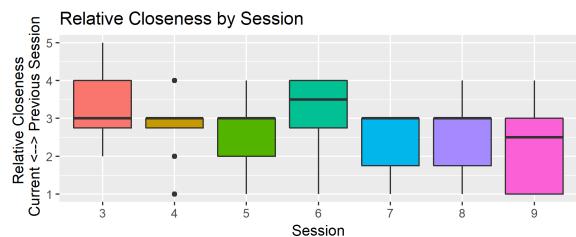
Participants were also invited to directly compare their experienced closeness to the agent on a semantic differential scale which had 1 as closest to the agent in the current session, and 5 as closest to the agent in the previous session. Table 7 and Figure 10 suggest that overall the central tendency in these scores was close to the neutral score of 3 (Wilcoxon’s $p > 0.25$). There were also no significant differences between the sessions (Friedman’s

Table 6: Closeness to Agent across Sessions.

Session	Mean	SD	Median	25th Perc.	75th Perc.
Session 2	3.50	1.41	4.0	2.75	4.25
Session 3	3.12	1.55	3.0	2.00	4.25
Session 4	3.12	1.36	3.5	2.00	4.00
Session 5	3.25	1.91	4.0	1.00	5.00
Session 6	2.88	1.64	2.5	1.75	4.25
Session 7	3.31	1.49	4.0	2.00	4.00
Session 8	3.31	1.44	4.0	2.00	4.12
Session 9	3.75	1.91	4.0	2.00	5.25

**Figure 9:** Relationship across sessions measured by IOS Scores by Session.**Table 7:** Relative Closeness by Session.

Session	Mean	SD	Median	25th Perc.	75th Perc.
Session 3	3.25	1.04	3.0	2.75	4
Session 4	2.75	0.89	3.0	2.75	3
Session 5	2.62	0.92	3.0	2.00	3
Session 6	3.12	1.13	3.5	2.75	4
Session 7	2.38	0.92	3.0	1.75	3
Session 8	2.50	1.07	3.0	1.75	3
Session 9	2.25	1.16	2.5	1.00	3

**Figure 10:** Relationship across sessions measured by Relative Closeness by Session.

$\chi^2(6) = 4.46, p = 0.62$). This suggests that participants did not see a session-by-session progression in terms of how they viewed their relationship with the agent.

Note, the participants' responses, while not showing clear changes across the sessions, showed considerable individual variation, suggesting that idiosyncratic factors in terms of how they responded to the robot might be more appropriately be examined in a qualitative manner.

5.2 Qualitative responses

Qualitative Responses were analysed in a descriptive manner. Responses were initially sorted into a set of emergent categories. These categories, and their application to the different statements, were created in an iterative process. First, one of the researchers would read through the statements and create an initial set of categories in an ad-hoc manner. Second, these categories would then be systematically applied to every statement. This process allowed for an initial testing of the validity of the category-test, leading to an iterative refinement of the categories. This was performed for both sets of open-ended questions. Finally a set of categories was arrived at that could be used to describe statements for both questions. At each step, another researcher would independently categorise the statements and discrepancies were resolved.

5.2.1 Research Question 1 — Do users accept scenarios inter-connected through narrative?

In addition to the Scenario Acceptability Scale and the two Likert scale items assessing whether or not participants wanted the robots for themselves, or thought them suitable for an elderly or disabled person, there were two open-ended questions asking for their reasoning for their responses to the two Likert scale items. A descriptive analysis was performed on these responses in order to categorise these statements, present the relationships between the different categories, and also the relationship between the categories and responses to the Likert scale items.

The Following categories were arrived at:

1. Negative usefulness
 - References to the Robots making tasks more difficult or the robots being difficult to use.
 - *I think it may slow down household activities.*
 - *It hindered the tasks rather than helped.*
2. Positive usefulness
 - References to the Robots making tasks easier, or specific aspects of the robots being easy to use.
 - *It helps to make things easier, like accessing the remote control and answer to calls quickly*

- *It can help transporting things and alerting a person to phone calls or doorbell*
3. Every Day Experience
 - References to the robots being used outside of the experimental context.
 - *It will be handy to have a companion at home to help with some activities. Like the music for relaxation*
 - *It would be good to show off, but on a practical basis, the activities took more time to do due to the robot which should be actually faster with the robot.*
 4. Scenario Capability
 - References to specific robots behaviours or capabilities displayed in the preceding scenario.
 - *It can help with transporting things and alerting a person to phone calls or doorbell.*
 - *It is able to migrate, and give alerts.*
 5. Companionship
 - References to the robots providing social interactions or companionship.
 - *It is very friendly and helpful and assists very well. Also can be a good companion.*
 - *For companionship and assisting with house chores and simple tasks*
 6. Specific Needs
 - References to the robots filling needs caused by disability or age.
 - *Definitely elderly people and disabled people. Especially having dementia can use this kind of robots for their daily life activities.*
 - *It could notify the user of sound if they have a hearing impairment.*
 7. Specific difficulty
 - References to aspects of the robots being particularly difficult due to age or disability.
 - *Elderly people might find it taxing to use the keypad for instructing the robot. Voice recognition and verbal commands will be better for elderly and disabled people.*
 - *Suitable for deaf people, but not for the blind.*

Categories were not mutually exclusive, and a given statement could be assigned to more than one category. For example, the following statement:

It will be handy to have a companion at home to help with some activities. Like the music for relaxation.

This statement was categorised as *Positive Usefulness* as the participant's comment reflects, *Every Day Experi-*

ence as the participant refers to their home, *Scenario Capability* as the robot playing music for the participant in this particular session was mentioned, and *Companionship* as the participant references the robot as a companion.

5.2.1.1 Overall responses categories

The overall responses to the open-ended questions are shown in Table 8.

Table 8: Responses by Category to open-ended items.

	For Self	For Others
Negative use	22	9
Positive use	30	41
Everyday	32	38
Capability	29	26
Companionship	18	3
Specific need	2	28
Specific difficulty	0	8

Table 8 shows that there were differences between the two items in terms of how participants' reasoning could be categorised. In terms of *negative usefulness*, there were 22 statements in response to the 'For Self' item, but only 9 to the 'For Others' item ($\chi^2(1) = 5.46, p = 0.02$). The number of statements referencing *positive usefulness*, *everyday experience*, and *scenario capabilities*, was comparable across both items, and comparatively numerous when compared to the other types of statements. *Specific needs* and *specific difficulties* were naturally represented to a larger extent in the responses to the 'For Other' statements.

5.2.1.2 Positive and Negative Statements

The open-ended responses to the questions are presented in Table 9 according to question and to whether or not participants answered positively to the relevant Likert scale.

Table 9 and Figure 11 suggests that there are some differences between how participants justify negative and positive responses to the Likert scale responses. Both reference capabilities they had just seen in the experimental scenario as well as in contexts outside of the study. The main difference between positive and negative responses overall lie in that participants who responded positively generally referenced the robot making certain tasks easier, while participants who responded negatively, tended to reference the robot making tasks more difficult (Fisher's

Table 9: Response by Category and Likert scale responses.

	+ Self	+ Others	- Self	- Others
Negative use	0	1	22	8
Positive use	28	39	2	2
Everyday	20	35	12	3
Capability	19	23	10	3
Companionship	17	3	1	0
Specific need	1	23	1	5
Specific difficulty	0	2	0	6

Exact $p < 0.01$). Another salient result is that participants who responded positively to questions about the robots for themselves, also tended to mention '*Companionship*' to a much larger extent than participants who responding negatively to the robot (Fisher's Exact $p = 0.002$). There were no significant differences between participants' referencing of their own everyday lives between negative and positive responses to the 'For Self' item (Fisher's Exact $p = 0.81$). However, the difference between the two in responses for the 'For Other' responses were approaching significance (Fisher's Exact $p = 0.07$).

5.2.1.3 Everyday Experience

Thirty-two statements related to reasoning on the everyday experience of the participants in the Robots for Self item, and 38 statements did so for the Robots for Other item. Table 10 show how these references co-occur with other categories.

Table 10: Co-occurrences of Everyday experience by Category.

	For Self	For Others
Negative use	8	2
Positive use	15	27
Capability	7	22
Companionship	4	2
Specific need	1	15
Specific difficulty	0	0

Table 10 suggests some differences in terms of references that co-occurred with references to "Everyday Experience" between the justifications "For Self" ratings and justifications "For Other". First of all, there are fewer co-occurrences overall in the "For Self" responses. In addition, there are relatively fewer references to '*Negative use*' in the "For Other" responses. This suggests that partici-

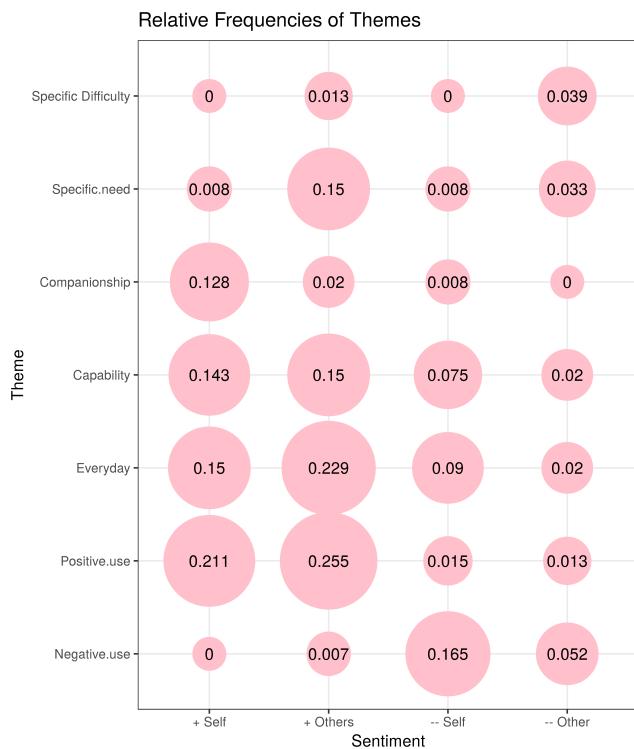


Figure 11: Relative frequencies of Themes referenced when discussing Sentiments by Category.

pants were more likely to consider the use of the robot as something that *would* be of use to others.

5.2.1.4 Scenario Capability

There were twenty-nine statements related to reasoning about the capabilities exhibited by the robot in the preceding scenario in the Robots "For Self" and 26 statements in the Robots "For Others" items.

Table 11 suggests a similar pattern as for the "Everyday Experience" co-occurrences. There were overall fewer co-occurrences in the "For Self" responses than in the "For Others" responses. In addition there were comparatively fewer 'Negative use' references in the "For Other" Responses.

5.2.1.5 Comparisons of co-occurrences

Note that there were more co-occurrences overall for the "For Others" responses, suggesting that participants were more likely to provide more context for their answers when discussing the robots' usefulness to others than when discussing it for themselves. Another interesting issue was that participants were more likely to reference *Negative use* when referencing both *Everyday Experience* as well as *Sce-*

Table 11: Co-occurrences of Scenario Capability by Category.

	For Self	For Others
Negative use	8	2
Positive use	8	18
Everyday	7	22
Companionship	4	1
Specific need	1	12
Specific difficulty	0	0

nario Capability when discussing the possibility of owning such robots themselves, compared to when discussing their usefulness for others. This was primarily caused by an underlying assumption that a disability or frailty due to old age would make the tasks more difficult for others in these groups.

5.2.2 Relationship

Participant responses to the open-ended questions as to their justification for which agent they felt closer to was also examined, and categorised using the same process as in the previous section. The following categories were created:

- Practicality
 - Statements referencing the functionality of the robots both in terms of tasks they could assist with as well as the presence of technical flaws.
 - * *Communication was more successful and straightforward in the previous session*
 - * *The robot worked better today*
- Familiarity
 - Statements referencing changes in the participant's perceived relationship to the robot.
 - * *Getting used to the robots may be the reason*
 - * *I felt more relaxed than last session because I got used to it*
- Context
 - Statements referencing responses to the context of the interaction.
 - * *I interacted more with it in the last session*
 - * *Because I had more interaction with the agent in this session than in the last interaction*
- Narrative
 - Statements referencing specific events that occurred as part of the over-arching narrative described in the methodology section.
 - * *The normal robot with the face was working again.*

- * There was only the stationary robot in this session.

As suggested by Table 12 and Figure 12, overall the three categories Practicality, Familiarity and Contextual were equally distributed within the participants' responses for their reasoning as to which session(s) they felt closer to the agent.

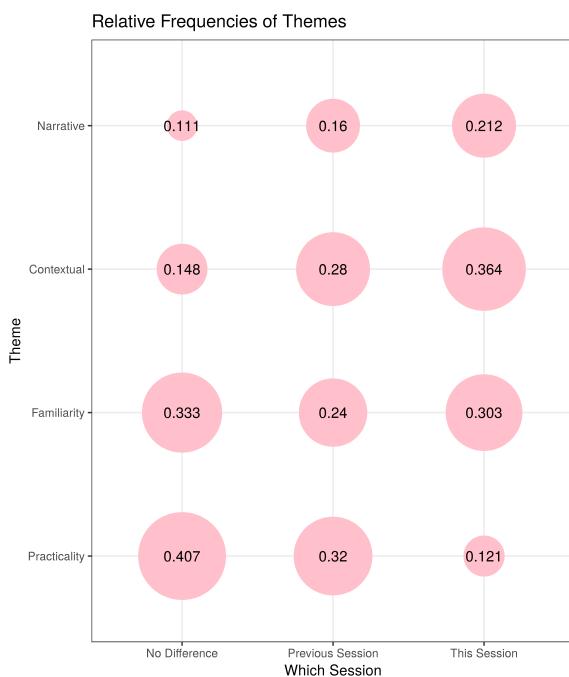


Figure 12: Relative frequencies of Themes referenced when discussing Relative Closeness by Preferred Session

Table 12: Overall References to Themes.

Theme	Number of instances
Practicality	23
Familiarity	25
Contextual	23
Narrative	14

5.2.2.1 Open-ended Responses and Preferred Session

Table 13 and Figure 12 suggests that there were some differences between the preferred sessions in terms of which themes were being referenced. This was assessed using Fisher's exact tests, and found no significant differences for Familiarity and Narrative (Fisher's Exact $p > 0.13$),

but approached significance for Practicality (Fisher's Exact $p = 0.07$) and Contextual (Fisher's Exact $p = 0.004$). The participants were more likely to reference practicality when justifying why they felt there was no difference or a preference for the previous session, than when they were stating a preference for the current session.

Table 13: Referenced Theme and Preferences.

Theme	No difference	Previous session	This session
Practicality	11	8	4
Familiarity	9	6	10
Contextual	4	7	12
Narrative	3	4	7

5.2.2.2 Open-ended Responses and Scenario Phase

As suggested by Table 14 and Figure 13, there were no differences between the phases for any of the themes, with the exception of the practicality being rated as comparatively less important *after* the exposition phase.

Table 14: Exposition Phase vs Later Phases and Referenced Theme.

	Initial Exposition	Later Phases
Practicality	10	13
Familiarity	7	18
Contextual	4	19
Narrative	0	14

5.2.2.3 Co-occurrences

Co-occurrences are presented in Table 15, which suggests no salient pattern in the co-occurrences.

6 Discussion

6.1 Research questions

6.1.1 Research Question 1 – Do users accept scenarios inter-connected through narrative?

The results of our proof of concept study suggest that the engagist immersive approach, using scenarios interconnected through narrative, was able to elicit responses from participants that were meaningful in terms of relating

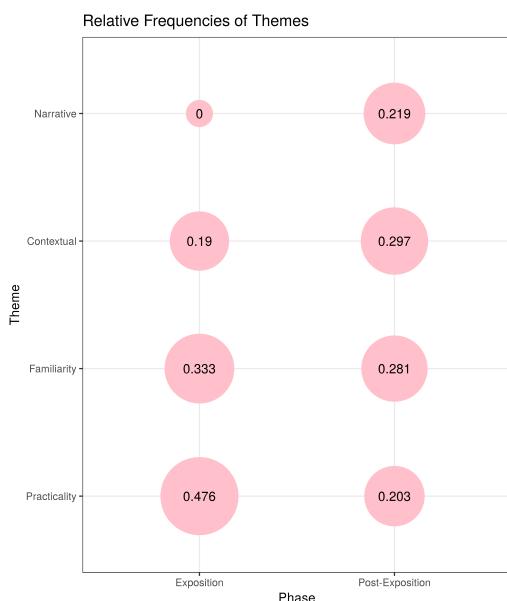


Figure 13: Relative frequencies of Themes referenced when discussing Relative Closeness by Scenario Phase.

Table 15: Co-occurrences in perceived closeness reasoning.

	Practicality	Familiarity	Contextual	Narrative
Practicality				
Familiarity	9			
Contextual	6	10		
Narrative	3	9	8	

to the behaviour of the robot and also to the actual interaction that they had had with it. It also was meaningful to the wider contexts of their everyday lives. This is encouraging as it lends greater validity to the prototyping approach used in this study.

When contrasting participants' reasoning about whether the robots were suitable for themselves or for others, participants were far more likely to reference companionship as a reason for adopting the robots themselves, while the adoption of the robots for others was considered to be more a matter of the utility that they could provide.

6.1.2 Research Question 2 – Does the user-agent relationship change when the agent migrates to different embodiments?

Overall, the participants did not really change their views of the agent in terms of measures on the IOS scale, and there were no overall preferences for the current session when asked during which session the participants felt

closest to the agent. However, there were differences in terms of how the participants reasoned about these ratings. Participants were more likely to reference practicality when justifying why they preferred the robot in the last session or when they considered their feelings towards the robot to be no different. Statements that highlighted that the robot performed similar tasks across the sessions were used to explain any lack of differences in the participants' feelings of closeness.

Those participants who reported that they felt closer to the robot in the current session (as compared to the previous sessions) tended to highlight contextual factors, such as the way that the scenario was structured to allow for more interactions. In addition, they tended to highlight the impact of the narrative intervention. These participants were also more likely to reference aspects related to the narrative intervention, such as the departure of the mobile robot and its return when reasoning about their feelings of closeness. This also suggests that the narrative intervention did succeed in giving the participants further insight into how interacting with robot companions over time would be like. It also suggests that the narrative intervention had a more pronounced effect on the immediate affective response to the scenarios.

6.2 Conclusions

Taken together, the results paint a complex picture of the participants' experiences of the robots and the scenarios in which they were presented. Participants consistently rated the acceptability of the scenarios in which the robot was being used quite highly, and would often reference their daily lives when discussing the possibility of similar robots being used outside of the experimental setting. What is of particular interest is that the participants viewed the decision of having a robot for themselves and having a robot for others differently. When considering a robot for themselves, they would consider the companionship aspects (emotional and hedonic) qualities of the interaction. However, when considering it for others, the main concerns would be the utility and practicality. While interesting in itself, this phenomenon highlights the possibility of a tension between different users of a robot intended for care. As suggested by Bedaf et al. [18], the primary user (the person whose home the robot operates in, and who will have the most interactions with the robot) may not be the person that commissions or organises the deployment of a robot companion in a care-scenario. Our findings suggest that even when carers or care professionals have a strong idea of the capabilities and interactions

provided by a robot companion, they may still not share the perspective of the primary user. This suggests that while functional aspects of a robot companion can be decided by third parties, interactional aspects, such as expressive or other behaviours supporting companionship, may be best left to the primary user. Given that we emulated this aspect of decision making by making the care-aspect explicit in our questionnaire when assessing suitability for others, it becomes even more interesting that what participants deem important in *their own* acceptance of the robot is not what necessarily matters when deploying it for others (e.g. to fulfil medical needs). As such, this exploration of robot companions, echo the dichotomy raised by Sharkey and Sharkey [45].

6.2.1 Limitations

The main limitation of the current study is the relatively low participant numbers, but this was a natural consequence of the large amount of resources that are required to maintain both the prototype technologies and to structure the narrative itself. In addition, the need for participants to have at least two one hour interactions sessions per week limited the number of participants that could practically be accommodated in the UH Robot House during a working week.

This may limit the generalisation of results, but even so these findings suggest that studies such as these can be a rich source of insight into human-robot interactions in domestic environments. Also that complex, meaningful and structured scenarios inter-linked through narrative are acceptable to naive users. Future work can expand on this work through applying the narrative and immersive techniques presented here to other functional interactions. In particular, expanding the scope for choice in interactions which allows for a wider range of narrative outcomes, while still retaining enough similarity in the immersive experience for them to be relatable and comparable to each other.

6.2.2 Future work

In future work, ideally, comparative studies would be performed with separate long-term studies using different scenario prototyping approaches. Such comparisons could illuminate whether the narrative framing or other factors were responsible for participants' high acceptability scores in our study. However, such comparative studies will have to address many methodological chal-

lenges, given that: a) participants' responses in our study showed large individual variation; and b) given that scenarios not being inter-connected through narrative will be qualitatively different from those using narrative framing. Future work could also compare different robotic platforms; for example android, humanoid, zoomorphic and mechanoid, and different combinations of those in a given long-term study. Finally, different themes could be explored within the scenarios, including therapeutic, educational, health and wellness, lifestyle or rehabilitation elements.

Our work has provided a first step towards prototyping home companion robots in long-term studies, adopting principles from diverse areas such as immersive engagement and fictional enquiry, creating scenarios which are inter-connected through a temporally linked episodic narrative, and our results, while limited, are hoped to inspire future research in this domain.

Acknowledgements: We would like to thank our colleague Wan Ching (Steve) Ho, for his efforts in implementing and running the robots in the scenarios. We would also like to thank our colleague Joe Saunders for his proof-reading and sense-checking of the paper. The work described in this paper was conducted within the EU Integrated Projects LIREC (LIVING with Robots and intEractive Companions), funded by the European Commission under contract numbers FP7 215554 and partly funded by the ACCOMPANY project, a part of the European Union's Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 287624.

References

- [1] S. Bedaf, G. J. Gelderblom, D. S. Syrdal, H. Lehmann, H. Michel, D. Hewson, F. Amirabdollahian, K. Dautenhahn, and L. de Witte, "Which activities threaten independent living of elderly when becoming problematic: inspiration for meaningful service robot functionality," *Disability and Rehabilitation: Assistive Technology*, vol. 9, no. 6, pp. 445–452, 2014.
- [2] M. Scopelliti, M. Giuliani, A. D'amico, and F. Fornara, "If I had a robot at home. Peoples' representation of domestic robots," in *Designing a More Inclusive World* (S. Keates, P. J. Clarkson, P. Langdon, and P. Robinson, eds.), pp. 257–266, London: Springer, 2004.
- [3] S. Frennert, B. Östlund, and H. Eftring, "Would granny let an assistive robot into her home?," *Social Robotics*, pp. 128–137, 2012.
- [4] C. L. Bethel and R. R. Murphy, "Review of human studies methods in HRI and recommendations," *International Journal of Social Robotics*, vol. 2, no. 4, pp. 347–359, 2010.

- [5] E. Horvitz, J. Breese, D. Heckerman, D. Hovel, and K. Rommelse, "The lumiere project: Bayesian user modeling for inferring the goals and needs of software users," in *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, pp. 256–265, Morgan Kaufmann Publishers Inc., 1998.
- [6] B. Whitworth, "Polite computing," *Behaviour & Information Technology*, vol. 24, no. 5, pp. 353–363, 2005.
- [7] J. Saez-Pons, D. S. Syrdal, and K. Dautenhahn, "What has happened today? memory visualisation of a robot companion to assist user's memory," *Journal of Assistive Technologies*, vol. 9, no. 4, pp. 207–218, 2015.
- [8] J. Saunders, D. S. Syrdal, K. L. Koay, N. Burke, and K. Dautenhahn, "teach me—show me"—end-user personalization of a smart home and companion robot," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 1, pp. 27–40, 2016.
- [9] K. Dautenhahn, "Robots we like to live with?-a developmental perspective on a personalized, life-long robot companion," in *2004 ROMAN 13th IEEE International Workshop on Robot and Human Interactive Communication*, pp. 17–22, IEEE, 2004.
- [10] C. D. Kidd and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3230–3235, IEEE, 2008.
- [11] Y. Fernaeus, M. Häkansson, M. Jacobsson, and S. Ljungblad, "How do you play with a robotic toy animal?: A long-term study of pleo," in *Proceedings of the 9th International Conference on Interaction Design and Children*, pp. 39–48, ACM, 2010.
- [12] M. M. de Graaf, S. B. Allouch, and T. Klamer, "Sharing a life with harvey: Exploring the acceptance of and relationship-building with a social robot," *Computers in Human Behavior*, vol. 43, pp. 1–14, 2015.
- [13] S. Payr, "Virtual butlers and real people: Styles and practices in long-term use of a companion," in *Your Virtual Butler*, pp. 134–178, Springer, 2013.
- [14] S. Kleanthous, C. Christophorou, C. Tsioriti, C. Dantas, R. Wintjens, G. Samaras, and E. Christodoulou, "Analysis of elderly users' preferences and expectations on service robot's personality, appearance and interaction," in *Human Aspects of IT for the Aged Population. Healthy and Active Aging* (J. Zhou and G. Salvendy, eds.), (Cham), pp. 35–44, Springer International Publishing, 2016.
- [15] G. S. Martins, L. Santos, and J. Dias, "User-adaptive interaction in social robots: A survey focusing on non-physical interaction," *International Journal of Social Robotics*, vol. 11, pp. 185–205, Jan 2019.
- [16] K. Kompatsiari, F. Ciardo, V. Tikhanoff, G. Metta, and A. Wykowska, "On the role of eye contact in gaze cueing," *Scientific Reports*, vol. 8, p. 17842, Dec. 2018.
- [17] F. Ciardo, D. De Tommaso, F. Beyer, and A. Wykowska, "Reduced sense of agency in human-robot interaction," in *Social Robotics* (S. S. Ge, J.-J. Cabibihan, M. A. Salichs, E. Broadbent, H. He, A. R. Wagner, and Á. Castro-González, eds.), (Cham), pp. 441–450, Springer International Publishing, 2018.
- [18] S. Bedaf, H. Draper, G.-J. Gelderblom, T. Sorell, and L. de Witte, "Can a service robot which supports independent living of older people disobey a command? the views of older people, informal carers and professional caregivers on the acceptability of robots," *International Journal of Social Robotics*, pp. 1–12, 2016.
- [19] D. S. Syrdal, K. Dautenhahn, K. L. Koay, and W. C. Ho, "Views from within a narrative: Evaluating long-term human–robot interaction in a naturalistic environment using open-ended scenarios," *Cognitive Computation*, vol. 6, no. 4, pp. 741–759, 2014.
- [20] D. Feil-Seifer and M. J. Mataric, "Defining socially assistive robotics," in *9th International Conference on Rehabilitation Robotics, ICORR 2005*, pp. 465–468, IEEE, 2005.
- [21] T. Bickmore, D. Schulman, and L. Yin, "Maintaining engagement in long-term interventions with relational agents," *Applied Artificial Intelligence*, vol. 24, no. 6, pp. 648–666, 2010.
- [22] C. Bartneck and J. Hu, "Rapid prototyping for interactive robots," in *Conference on Intelligent Autonomous Systems*, pp. 136–145, 2004.
- [23] E. Vlachos, E. Jochum, and L.-P. Demers, "The effects of exposure to different social robots on attitudes toward preferences," *Interaction Studies*, vol. 17, no. 3, pp. 390–404, 2017.
- [24] C. Dindler and O. S. Iversen, "Fictional inquiry—design collaboration in a shared narrative space," *CoDesign*, vol. 3, no. 4, pp. 213–234, 2007.
- [25] G. Seland, "Empowering end users in design of mobile technology using role play as a method: reflections on the role-play conduction," *Human Centered Design*, pp. 912–921, 2009.
- [26] D. S. Syrdal, K. Dautenhahn, K. L. Koay, and W. C. Ho, "Integrating constrained experiments in long-term human–robot interaction using task-and scenario-based prototyping," *The Information Society*, vol. 31, no. 3, pp. 265–283, 2015.
- [27] G. Freytag, *Freytag's technique of the drama: an exposition of dramatic composition and art*. Scholarly Press, 1896.
- [28] J. S. Tynes, "Prismatic play: Games as windows to the real world," in *Second Person: Role-Playing and Story in Games and Playable Media* (P. Harrigan and N. Wardrip-Fruin, eds.), pp. 221–228, Cambridge, Mass.: MIT Press, 2007.
- [29] S. L. Bowman and A. Standiford, "Enhancing healthcare simulations and beyond: Immersion theory and practice," *International Journal of Role-playing*, vol. 6, pp. 12–19, 2016.
- [30] A. Standiford, "Lessons Learned from Larp: Promoting Social Realism in Nursing Simulation," in *The Wyrd Con Companion Book* (S. L. Bowman, ed.), (Los Angeles, CA), pp. 150–159, Wyrd Con, 2014.
- [31] J. T. Harviainen, "The multi-tier game immersion theory," *2003) As Larp Grows Up-Theory and Methods in Larp. Copenhagen, Projektgruppen KP03. The book for Knudepunkt*, 2003.
- [32] Y. Fernaeus, M. Jacobsson, S. Ljungblad, and L. E. Holmqvist, "Are we living in a robot cargo cult?," in *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 279–280, IEEE, 2009.
- [33] I. Duque, K. Dautenhahn, K. L. Koay, L. Willcock, and B. Christianson, "Knowledge-driven user activity recognition for a smart house—development and validation of a generic and low-cost, resource-efficient system," in *Proceedings of the 6th International Conference on Advances in Computer-Human Interactions*, 2013.
- [34] J. Saunders, N. Burke, K. L. Koay, and K. Dautenhahn, "A user friendly robot architecture for re-ablement and co-learning in a sensorised home," *Assistive Technology: From Research to Practice (Proc. of AAATE)*, vol. 33, pp. 49–58, 2013.
- [35] K. L. Koay, D. Syrdal, K. Dautenhahn, K. Arent, B. Kreczmer, et al., "Companion migration—initial participants' feedback from a video-based prototyping study," in *Mixed Reality and Human-Robot Interaction*, pp. 133–151, Springer, 2011.

- [36] E. M. Segura, H. Cramer, P. F. Gomes, S. Nylander, and A. Paiva, “Revive!: Reactions to migration between different embodiments when playing with robotic pets,” in *Proceedings of the 11th International Conference on Interaction Design and Children*, pp. 88–97, ACM, 2012.
- [37] M. Kriegel, R. Aylett, P. Cuba, M. Vala, and A. Paiva, “Robots meet ivas: a mind-body interface for migrating artificial intelligent agents,” in *International Workshop on Intelligent Virtual Agents*, pp. 282–295, Springer, 2011.
- [38] M. Kriegel, R. Aylett, K. L. Koay, K. Casse, K. Dautenhahn, P. Cuba, and K. Arent, “Digital body hopping-migrating artificial companions,” *Proceedings of Digital Futures’ 10*, 2010.
- [39] D. S. Syrdal, K. L. Koay, M. L. Walters, and K. Dautenhahn, “The boy-robot should bark!-children’s impressions of agent migration into diverse embodiments,” in *Proceedings: New Frontiers of Human-Robot Interaction, a symposium at AISB*, 2009.
- [40] K. L. Koay, D. Syrdal, W. Ho, and K. Dautenhahn, “Prototyping realistic long-term human-robot interaction for the study of agent migration,” in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 809–816, IEEE, 2016.
- [41] K. L. Koay, G. Lakatos, D. S. Syrdal, M. Gácsi, B. Bereczky, K. Dautenhahn, A. Miklósi, and M. L. Walters, “Hey! There is someone at your door. A hearing robot using visual communication signals of hearing dogs to communicate intent,” in *2013 IEEE Symposium on Artificial Life (ALife)*, pp. 90–97, IEEE, 2013.
- [42] P. Lanillos, J. F. Ferreira, and J. Dias, “Designing an artificial attention system for social robots,” in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4171–4178, Sep. 2015.
- [43] B. Rolfe, C. M. Jones, and H. Wallace, “Designing dramatic play: Story and game structure,” in *Proceedings of the 24th BCS Interaction Specialist Group Conference*, pp. 448–452, British Computer Society, 2010.
- [44] A. Aron, E. N. Aron, and D. Smollan, “Inclusion of other in the self scale and the structure of interpersonal closeness.,” *Journal of Personality and Social Psychology*, vol. 63, no. 4, p. 596, 1992.
- [45] A. Sharkey and N. Sharkey, “Granny and the robots: ethical issues in robot care for the elderly,” *Ethics and Information Technology*, vol. 14, no. 1, pp. 27–40, 2012.