

Caring About the User's View

The Joys and Sorrows of Experiments with People

Amedeo Cesta, Gabriella Cortellessa,
Federico Pecora

Planning and Scheduling Team
ISTC-CNR, National Research Council of Italy
Institute for Cognitive Science and Technology
{name.surname}@istc.cnr.it

Vittoria Giuliani, Massimiliano Scopelliti,
Lorenza Tiberio

Environmental Psychology Unit
ISTC-CNR, National Research Council of Italy
Institute for Cognitive Science and Technology
{name.surname}@istc.cnr.it

Abstract

Deploying AI, and specifically P&S, technology into the real world entails many stimulating problems for researchers and system designers. Various smart solutions have been produced both for specialized problems and, in fewer cases, for more general purpose domains. An important aspect which is fundamental for any successful application, and more specifically for those which address a broad audience, concerns the users' perception and acceptance of technology. This issue is often either neglected or underestimated.

We have been working for two years on the issue of importing experimental techniques from HCI and experimental psychology into smart system development. This approach has both *pros* – interesting features from the user perspective are discovered and can be used to bias design and research activities – and *cons* – experimenting with humans adds additional difficulty to the project, and applying a correct methodology is very time consuming.

This paper describes a fielded experimental investigation of a fully implemented AI system named ROBOCARE. The system uses constraint-based scheduling technology to actively monitor a pattern of activities executed by an assisted person and uses detected temporal constraint violations to trigger meaningful and contextualized proactive interaction. Dialogue with the users is managed by a robotic mediator who acts as the main communication channel between the users and the intelligent domestic environment. The paper presents an evaluation of elderly people's perception of the intelligent system, focusing on aspects related to the robot's aspect, interaction modalities, content and timing of interaction, providing suggestions and hints for system designers.

Introduction

The development of real applications of intelligent systems is often driven by the need to interact with users. Several projects have dedicated effort to develop front-ends for user interaction and/or more sophisticated mixed-initiative systems that more deeply involve human beings in the activity of the intelligent system. The connection with users is very often a very important enabler for the actual use of intelligent systems. This has been demonstrated by experiences in systems developed for several tasks in specialized domains, see as examples MAPGEN (Ai-Chang *et al.* 2004) and MEXAR2 (Cesta *et al.* 2007a) in the space domain, and AMC BARREL ALLOCATOR (Becker & Smith 2000) and

PASSAT (Myers *et al.* 2003) in military domains. A commonality among these very different systems is represented by their being dedicated to specialized users in restricted work environments, that somehow are aware of the role of an intelligent system and its ability to provide precious assistance in decision making. Nevertheless, the interaction is somehow a key issue for acceptability.

Further work attempts to address more broad application areas very close to consumer markets. In applications for everyday life, e.g., (McCarthy & Pollack 2002; Refanidis 2007) or more specifically dedicated to older people, e.g., (Levinson 1997; Pollack 2005; Cesta *et al.* 2007c), the acceptance of advanced services by users is paramount. How can we rely on the fact that the advancement in flexibility and partial automation offered by our technology will become an added value also for users that are not as specialized as in other cases (e.g., space, military) but rather are not even well acquainted with modern information technology? In light of the failures of AI in the 70-80s it is worth asking ourselves additional questions w.r.t. user acceptability.

For the purpose of evaluating user response to mixed-initiative systems we have been importing experimental techniques from HCI and experimental psychology (Cortellessa & Cesta 2006). In the ROBOCARE project we have also applied similar methodology and this paper specifically reports on that part of our work. ROBOCARE (Cesta & Pecora 2006) has involved research groups with different background with the goal of investigating how the use of intelligent technology for supporting elderly people can improve their quality of life, prolonging their independence at home. The focus on the domestic setting is not only motivated by the aim of improving home technology personalization, but also by recent studies, e.g., (Giuliani, Scopelliti, & Fornara 2005), that underscore the relevance of the attachment of elderly people to their home and the beneficial effects of increasing their independence at home. This paper shows the use of experimental methodology to judge the technical results of the project. In particular after three years of developing different advanced functionalities we have characterized a fixed point in our achievement, and asked ourselves the question "we have understood what we are able to do: but what exactly do potential users think of the system we have achieved?". We have set up a quite complex experiment and found 40 people to take part in it. The rest of the paper de-

scribes the experiment and its outcome, showing a number of interesting considerations that emerge from the data.

It may seem that we have found a sort of *panacea* for all the questions we have with respect to users. Indeed there is still a lot of work to be done. In particular, as we will comment at the end of the paper, setting up an experiment with people is an extremely demanding and specialized task.

The ROBOCARE Assistive Domain

The ROBOCARE Domestic Environment (RDE) is the result of a three year project aimed at developing cognitive support technology for elderly people. Our focus on the domestic scenario stems from a series of studies of different physical environments for elderly people (Cesta & Pecora 2006). The main “actor” in our smart home environment is a robotic agent with verbal interaction capabilities. The interaction content is generated through the combination of different intelligent systems and environmental sensors. A complete description of the technology is given in (Cesta *et al.* 2007b). Here we present a minimal description that should allow the reader to understand the experimental setup.

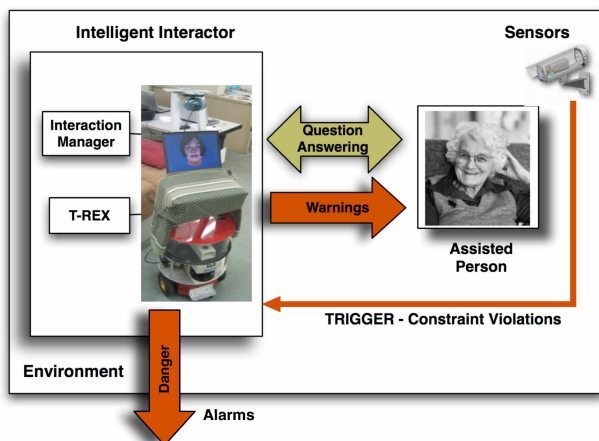


Figure 1: Interactions in the ROBOCARE environment.

The objective of the RDE is to provide support in the management of an elderly person’s daily activities. To this end, the RDE, sketched in Fig. 1, is composed of two fundamental subsystems. On one hand, an “intelligent observer” of the assisted person: information coming from environmental sensors¹ is used for maintaining an updated representation of what is happening in the environment. The sequence of observations from the artificial vision sensors allows to follow the evolution of the activities of the observed person. Based on the synthesis of these observations, the system is able to generate a report that reveals when the person’s activities have been performed within “reasonable” temporal boundaries or when important anomalies or even violations on their execution have been detected. In this light, the RDE’s basic functionality is an example of home *Activity Monitor* grounded on scheduling technology. Notice that,

¹At the moment sensors are cameras whose observation are elaborated by artificial vision algorithms to extract useful features.

on its own, the domestic activity monitor acts as a “silent observer” and does not take initiative with respect to the elder person in any way. On the other hand, the RDE also provides an interface with the assisted elder through an interactive subsystem. This subsystem is essentially a “proactive assistant” which closes the loop between the elder user and the intelligent environment, enabling the system to take initiatives based on Activity Monitor inference.

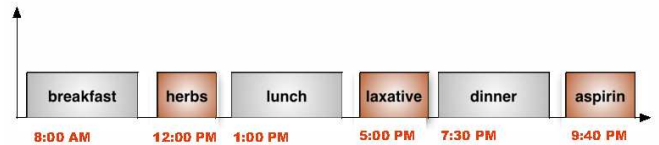


Figure 2: Example of desired behavior specified by the care giver for the assisted person in form of a *schedule*.

As a central component for activity management we have employed an AI-based schedule management environment called T-REX – Tool for schedule Representation and Execution (Pecora *et al.* 2006). T-REX allows to represent a set of activities and their quantitative temporal connections (i.e., a schedule of activities that the user is expected to carry out). These temporal constraints represent the behavioral requirements to which the assisted person should adhere. To be more concrete, let us consider a behavioral pattern described by a schedule composed of 6 different activities (*breakfast, lunch, dinner*, as well as taking three different medicines). Due to medical requirements, let us also suppose that such activities must satisfy certain temporal requirements, such as “dinner should not begin before 7:30 PM, nor should it occur less than 5 hours after lunch” and “aspirin should only be taken after dinner, but not too late”, and so on.

An “ideal schedule”, i.e., an enactment of these activities which does not violate any temporal constraint, is shown in Fig. 2. Broadly speaking, the objective of the Activity Monitor is to recognize deviations from this ideal situation. Specifically, the system should assess the extent to which the user’s behavior deviates from this situation. This equates to assessing which temporal constraints are progressively violated during the day. In a nutshell, system interventions are driven by constraint violations: warnings, alarms and suggestions result from violated constraints. System interventions are processed by the interactive subsystem on board the robotic mediator.

Managing Interaction with the User

As already mentioned, interaction within ROBOCARE relies on an embodied robotic assistant as the focal point between the user and the system. Communication between the user and the robotic mediator occurs verbally. For the purposes of this study, we distinguish two forms of interaction based on *who takes the initiative* to start a dialogue:

On-Demand interaction in which the user takes the initiative first. The assisted person commences interaction, for instance, by querying the system’s knowledge base: “have I taken my pills?”, or “can I make an appointment for tomorrow at 5 PM?”.

Proactive interaction in which the intelligent environment commences interaction guided by its internal reasoning. Within ROBOCARE, constraint violations have been considered as a *trigger* for the system to take the initiative and perform some actions: issue an alarm in case of illness, or verbalize warnings and suggestions.

Our work explicitly focuses on the development of active and, at the same time, unobtrusive services to integrate within the artificial assistant. All interaction services rely on the Interaction Manager. This module essentially consists in a rule-based system that fires *situation-action* rules. In other words, it continuously assesses the situation and activates a particular submodule as an action.

The main “interaction occasions” managed in the current version of the intelligent assistant are also shown in Fig. 1. We categorize as *On-Demand* interaction the “Question/Answer” category of dialogues. This activity is triggered by a speech input from the assisted person. The generation of the answer is managed mostly internally to the manager that has information on activity history and/or on the current state of the environment, to answer questions like “Have I had lunch?” or “What time is it?”, etc.

Instances of *Proactive* interaction are “Danger” and “Warning” scenarios. Undoubtedly, one of the important tasks for assistance is to recognize emergencies for the monitored person. The emergency trigger is fired by particular combinations of the input provided by the sensors that monitor the environment and the assisted person. As an example we can discriminate as a dangerous situation the case in which a person is “laying down on the kitchen floor” or “laying down in bed half an hour after usual wake up”, rather than “laying down in bed within an expected period” which is recognized as a regular situation. The danger trigger is dealt with by a specific behavior of the multi-agent system that interrupts the usual flow of activities and undertakes an action: the robot is sent to the assisted person, a specific dialogue is attempted, and if no answer from the assisted person is obtained, an *Alarm* is immediately fired to the external world (call to a relative, to an emergency help desk, etc.).

A warning scenario is one in which constraint violations are detected by the T-REX activity monitor. Broadly speaking, the activity monitor decides the values for the variables that are used by the interaction manager to trigger a proactive dialogue with the assisted person. The content of the dialogue is synthesized on the basis of the monitor’s internal knowledge.

Overall the Interaction Manager in Fig. 1 is a quite simple planner that supervises the initiative of the “interactor” towards the assisted person. It is worth underscoring how the combination of this manager and the activity monitor endows the whole assistive environment with capabilities of proactive participation in a mixed-initiative interaction.

Experiments with Elder Users

The aim of the remainder of this article is to present experiments aimed at understanding the perception of older people towards the assistance that the robot (and thus the assistive

environment as a whole) is able to offer at the moment.

Previous Evaluations

A previous study (Scopelliti, Giuliani, & Fornara 2005), performed at the beginning of ROBOCARE, was aimed at drawing some preliminary desiderata and requirements for assistive robots. This evaluation analyzed laypeople’s representations of domestic robots with respect to a variety of topics: the users’ expectations with respect to the robot’s capabilities to perform different everyday activities at home; their emotional response to a domestic robot; the image of the robot, referring to shape, size, color, cover material, speed; preferences and expectancies about the robot’s personification (given name, etc.) and the modalities of human-robot communication and interaction.

Results showed that people overestimate manipulative abilities and underestimate cognitive capabilities of the robot, whose representation is somewhat unrealistic: a domestic robot is still too far away from everyday life experience of laypeople. In addition, people at different stages of their lifespan showed very divergent opinions and preferences. In particular, older people clearly indicated a preference for a small robot, hardly resembling a human being, which has to intrude as less as possible in personal and domestic life; a device which is not autonomously free to move in the domestic environment and simply responding to tasks to be performed. In fact, while its practical utility was recognized, the robot emerged as a potential source of danger and discomfort in private life, and the idea of a non-autonomous device seemed to be a way to ward off their anxiety. Another issue to be addressed has to do with the context in which the robot is expected to operate. The use of new technologies and domestic robots in the home environment is not only a matter of general human-technology interaction, but is also associated with the specific sphere of human life in which assistance is needed (Giuliani, Scopelliti, & Fornara 2005). Elderly people showed a rather positive attitude towards a technological modification in the domestic environment, yet the inclination to use technological devices is strongly associated to the problem they have to cope with. In some situations, a technological aid seemed to be unrealistic, or unpractical, or it would have better been replaced by a more common alternative. In other cases, concerning health and personal/environmental safety above all, it emerged as a suitable solution to cope with losses imposed by ageing.

The Present Study

The studies mentioned previously focus on users’ attitudes toward a purely imaginary robotic agent, with unspecified abilities and not operating in a real domestic environment. For this reason, differences in users’ reactions could have been related to both diverse knowledge and bias toward technologies.

The final prototype achieved by ROBOCARE allows us to overcome this limitation. The evaluation of a tangible robot allows us to eliminate pre-conceptions and other biases. Performing the evaluation on the RDE prototype allows us to draw specific conclusions on the prototype itself, and also



(a) Non anthropomorphic version of the robot.

(b) Robot showing a human speaking face.

Figure 3: The two experimental conditions of the robot.

to investigate some general issues relative to the challenges of assistive technology for elderly people. This analysis is in line with current recommendations for the evaluation of complex assistive technology. For instance, it is recognized in (Hutchins 1995) that human-robot interaction is to be evaluated on socio-culturally constituted activities outside the design laboratory. In this light, the aim of our research is to analyze the potential reactions of final users to real life interactions between elderly people and an assistive robot.

The present analysis considered eight different scenarios, which were meant to be representative of daily situations in which elderly people may be involved. The situations were selected with reference to previous research on this topic (Giuliani, Scopelliti, & Fornara 2005), ranging from the most emotionally involving to less critical and emotionally neutral, with the aim of exploring elderly people’s evaluations of the potential role of a domestic robot as a useful support to ageing. Specifically, the study focuses on three main aspects.

First, we perform an evaluation of how meaningful each scenario is with respect to the respondents’ every day life. This allows us to understand how useful state-of-the-art assistive technology can be in real situations. Moreover, it provides a precious indication as to whether we are employing this technology to solve real needs. Scenarios were arranged in order to have evaluations of the robot in different typologies of interactive situations: we propose a main distinction between “On-demand” and “Proactive” scenarios. On-demand scenarios imply an explicit request for the robot’s activity by the final user; in proactive scenarios, the robot autonomously intervenes in the domestic environment, for both an emergency and a simple suggestion. The comparison between On-demand and Proactive situations is aimed to offer a suggestion as regards the preferred level of autonomy of the assistive device.

Second, we focus on the respondents assessment of our robotic mediator. The analysis focuses on aspects related to the physical aspect of the robot, its interaction capabilities,

and in general its suitability in the domestic context (e.g., size, mobility, integration with the environment).

Third, we observe user preferences with respect to robot’s features evoking a human being. Although our robot is not anthropomorphic, it is possible to deploy it in two slightly different versions: one in which the robot has a 3D facial representation (whose lip movement is synchronized with the speech synthesizer), and one without a facial representation. These variants were used to toggle the variable “Similarity to human beings”, which emerged as a key component in elderly people’s representation of domestic robots (Scopelliti, Giuliani, & Fornara 2005).

Materials. Eight short movies (ranging from about 30 seconds to little more than one minute) were developed showing potential interaction scenarios between an elderly person and the RDE’s robotic agent in a real domestic environment (see Fig. 4).



Figure 4: Sample frames from the videos depicting the eight scenarios.

The features of the robotic agent were manipulated according to two different experimental conditions: in the first condition (“Face”) a robot showing a human speaking face on a notebook monitor; in the second (“No-face”), a robot with no reference to human features (see Figure 3). The eight scenarios presented everyday life situations in which the robot provides cognitive support to the elderly person, and referred to critical areas, as highlighted by previous research: (a) management of personal/environmental safety, (b) healthcare, (c) reminding events/deadlines, (d) support to activity planning, (e) suggestions. In the following, the eight scenarios are shortly described.

Scenario 1 [Environmental safety] *The actor/actress is sitting on the sofa, watching TV. In the meantime, in the kitchen the sauce on the stove is overcooking. The sensors communicate this information to the robot. As a consequence, the robot moves toward the actor/actress and says: “The pot is burning. You should turn it off”. The actor/actress immediately goes to the kitchen and turns off the stove.*

Scenario 2 [Personal safety] *The actor/actress is sitting on the sofa, reading a magazine. Suddenly, he/she feels ill, and faints. The camera recognizes the situation and communicates this information to the robot. The robot approaches the actor/actress and says: “Are you all right?”. As it gets no answer, the robot calls the actor’s/actress’ son at work,*

who calls the medical emergency. The final scene shows the son and the doctor in the living room with the actor/actress, who feels fine.

Scenario 3 [Finding objects] The actor/actress is sitting on the sofa, and takes a magazine to read. Suddenly, he/she realizes that the glasses are not on the table in front of him/her. The actor/actress calls the robot and asks: "Where are my glasses?". The sensors in the rooms search for the glasses, and finally find them in the kitchen. The robot answers: "The glasses are on the table in the kitchen". The actor/actress goes to the kitchen and takes the glasses, then goes back to the sofa and starts reading the magazine.

Scenario 4 [Reminding analyses] The actor/actress is in the kitchen. He/she is about to have breakfast. When he/she puts the pot on the stove to warm up the milk, the robot says: "You cannot have breakfast now. You have an appointment for a medical analysis". The actor/actress answers: "You're right. I had forgotten all about it!".

Scenario 5 [Activity planning] The actor/actress is on the phone in the living room. He/she is speaking to the secretary of a clinical center to get an appointment for a medical examination. The secretary proposes an appointment for the next day, with two alternatives: either in the morning, or in the afternoon. The actor/actress asks the robot for eventual engagements the following day. The robot answers: "You have another engagement in the morning. In the afternoon, you do not have any appointments". The actor/actress accepts the appointment in the afternoon.

Scenario 6 [Reminding medication] The actor/actress is sleeping on the sofa, and suddenly wakes up. He/she does not realize what time is it, and thus he/she asks the robot. The robot answers: "It is four o'clock". The actor/actress does not remember whether or not he/she took his/her medicine after lunch, and asks the robot. The robot answers: "Yes, you took your medicine."

Scenario 7 [Suggestions] The actor/actress is watching TV on the sofa. It is five o'clock. The robot enters the living room and says: "You have been spending all the day at home. Why don't you go out and have a walk?". The actor/actress answers: "I really don't feel like it... I think I'll go water the plants in the garden".

Scenario 8 [Reminding events] The actor/actress is having breakfast in the kitchen. The robot reminds him/her: "Today it's your friend Giovanni's birthday. Remember to call him". The actor/actress answers: "You are right. I will do it right away". Then he/she goes to the living room and calls Giovanni.

Scenarios 1, 2 and 4 showed proactive situations referring to domestic healthcare and emergencies; scenarios 7 and 8 showed proactive situations implying suggestions; scenarios 3, 5 and 6 showed on-demand interactions.

Tools. A questionnaire was developed for data collection. It consisted of three sections, plus a final part for socio-demographics. The sections were arranged as follows:

Section 1. Eight fill-in papers, each of them referring to one of the eight scenarios, were presented. For each scenario, questions about the likelihood of the situation for the elderly person, the utility and acceptability of the robot were asked.

Section 2. An attitude scale, consisting of 45 Likert-type items, referring to the physical aspect of the robot, its behavior and communication modalities; the level of integration with the domestic environment; the degree of perceived intrusion/disturbance of the robot in everyday life and routines; the personal advantages and disadvantages of having such a device at home.

Section 3. An emotional scale, consisting of sixteen adjectives through which respondents have to evaluate the possible presence of the robot in their home.

In the Likert-type items, the respondents had to express their level of agreement/disagreement on a scale ranging from 0 ("I totally disagree") to 4 ("I completely agree").

Participants and procedure. Subjects recruited for this exploratory study were forty elderly people (aged 56-88; mean age = 70.3 years). Participants were 13 males and 27 females; as for their educational level, 17.9% attended primary school, 43.6% attended middle school, 25.6% attended high school, 12.9% have a degree. Most of them (82.5%) are retired. Before retirement, 22.5% were teachers, 15% were office workers. Subjects were randomly assigned to one of the two experimental conditions (Face/No-face). The movies were either projected on a notebook monitor, in a face-to-face administration, or on a larger screen, in a small-group administration. Two different sequences of scenario presentation were used, in order to avoid the potential influence of an order effect of episodes on results. After viewing each scenario, participants were asked to fill the paper referring to it (Section 1 of the questionnaire). At the end of the whole presentation, subjects were asked to give general evaluations of the robot (Sections 2-3 of the questionnaire), and to fill the final part of the questionnaire, referring to socio-demographics.

Results

The results described in the following paragraphs, are obtained from a combination of quantitative (ANOVA, χ^2 and Pearson's correlation) and qualitative analyses of the user evaluation questionnaires.

On-demand vs. Proactive scenarios. An analysis of meaningfulness of scenarios shows that our selection was effective in identifying typical everyday situations. On the whole, both On-demand and Proactive situations involving emergency and healthcare were evaluated as significantly more common than Proactive situations referring to suggestions; ³ in Proactive situations involving emergency and healthcare the robot was evaluated as significantly more useful than in On-demand and Proactive situations referring to

²The letters (a,b,c) indicate significant differences between typologies of situations.

³($F(2, 78) = 15.00, p < .001$)

Table 1: Evaluation of the different types of situations. Users were asked to evaluate on a scale from 0 to 4.

Type of situation	Meaningfulness		Usefulness		Preference	
	Mean ²	St. dev.	Mean ²	St. dev.	Mean ²	St. dev.
Proactive (Emergency)	2.51 ^a	.59	2.74 ^a	.73	2.48 ^a	.87
On-demand	2.53 ^a	.67	2.44 ^b	.85	2.13 ^b	.97
Proactive (Suggestions)	1.99 ^b	.83	1.94 ^c	.98	1.76 ^c	1.13

suggestions; ⁴ finally, the preference for the robot's support was significantly higher in Proactive situations involving emergency and healthcare than in On-demand and Proactive situations referring to suggestions ⁵ (see Tab. 1).

A global picture of the robotic mediator reveals a rather positive perception. In particular, the robot emerged as a very useful device for Personal and Environmental safety, Reminding medications, and Finding objects; conversely, not particularly useful in case of Suggestions (see Fig. 5).

In addition to utility, the robot was also indicated as a solution users would accept when difficulties arise, again with specific reference to Personal ($M = 2.95$, $sd = 1.06$) and Environmental safety ($M = 2.55$, $sd = 1.01$). In general, a significant correlation emerged (Pearson's r) between meaningfulness of a specific scenario, utility of and preference for the robot in that scenario (i.e., the higher the meaningfulness of the scenario, the higher the users' perceived utility and the probability they would accept such a device at home).

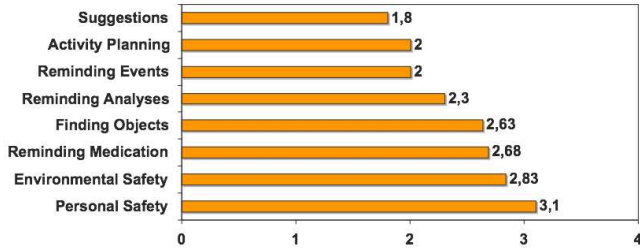


Figure 5: Utility of the domestic robot for everyday situations.

General evaluation of the robot. As to the different characteristics of the robot, some Positive and Negative aspects emerged (see Tab. 2).

Positive Aspects. Both face-to-face interaction with people and communication modality were on average positively assessed; in addition, elderly people favorably evaluate the possibility to interact with the robot for a training to reduce the loss in cognitive functioning. The overall integration with the home environment is good, even though a total freedom of movement is not completely appreciated. Among advantages given by the robot in the domestic environment, it can make people living alone feel safer, it can provide a support for cognitive functioning and, in general, in the organization of everyday activities.

⁴($F(2, 78) = 27.84$, $p < .001$)

⁵($F(2, 78) = 20.83$, $p < .001$)

Table 2: General evaluation of the robot

Positive Aspects	Mean	St. dev.
Face-to-face interaction	2,60	1,23
Communication modality	2,33	,62
Cognitive Training	2,53	1,24
Integration	2,34	,91
People feel safer	3,23	1,14
Support for activities	2,98	1,03
Negative Aspects	Mean	St. dev.
Management difficulty	2,95	1,11
Economic costs	3,25	,84
Possible Dependence	2,48	1,45

Negative Aspects. Some troubles with the management of the device (repairs, etc.) and the possible economic costs are expected. Also, users seemed to worry about potentially becoming dependent on the robot in certain cognitive tasks. Finally, the emotional reaction of elderly people to the robot was very good, scoring high on the positive adjectives *useful*, *interesting* and *relaxing*, and scoring very low on the negative adjectives *scary*, *overwhelming*, *gloomy*, *dangerous*, *uncontrollable* (see Fig. 6).

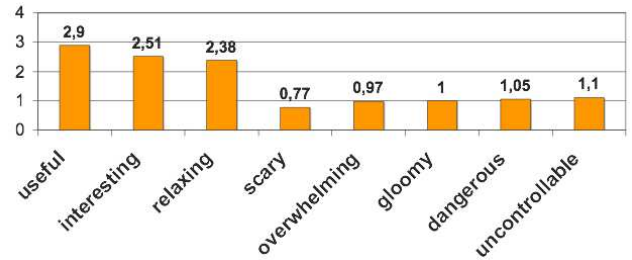


Figure 6: Emotional reaction of elder people to the robot (Means)

Similarity to human beings. As to this issue, our manipulation emerged to be effective, being the No-face version significantly preferred on the whole, specifically appearing both *less mechanical* and *less cold*. The No-face version was also evaluated as having a significantly higher level of integration with the domestic environment and a larger variety of advantages than the Face version, referring to *ease of use* and a *low need for repair*⁶. In addition, elderly people seemed to be more likely to develop a psychological at-

⁶overall preference: ($F_{(1,38)} = 6.34$, $p < .05$); *less mechanical* ($F_{(1,38)} = 5.11$, $p < .05$); *less cold*: ($F_{(1,38)} = 7.25$,

tachment towards the No-face version than towards the Face version ⁷.

Discussion

This study addresses some general acceptability requirements for assistive robotic agents. The general framework depicted by the On-demand vs. Proactive situations emerged to be highly meaningful in elderly people's experience, and the evaluation of proposed scenarios plainly shows their main concerns in everyday life and the potential role of a domestic robot in supporting them. Elderly people perceive a clear distinction between important and unimportant activities to be performed at home. For those activities which are perceived of greatest relevance, mainly concerning personal and environmental safety, the autonomy of the robot in the management of the home environment and in taking decisions proved to be a very useful resource. The robot is also appreciated for its capability in responding to a specific need expressed by the user, especially when referring to a cognitive difficulty associated with ageing, and involving activities related to healthcare (e.g., remembering things to do or what has been already done, with particular reference to medications and analyses).

Conversely, a robot making suggestions regarding unimportant activities is perceived as a bit irritating. A well defined relationship between likelihood of situations, perceived utility of and preference for the robot emerged. In the situation involving an emergency the preference for the robotic support is higher than the perceived likelihood of the situation itself, and the perception of utility scores highest. Conversely, with respect to activities which are not considered to be essential in everyday life, elderly people show a tendency to assign a low score on likelihood of occurrence, and even lower scores on usefulness and preference.

Overall, even if emergencies are not likely to occur, their central role in elderly people's experience makes the perceived utility of and the expressed preference towards a proactive robot higher. This picture is in line with the model of successful aging put forward by (Baltes & Baltes 1990), which stresses the role of selection and optimization of activities with increasing age, and the importance of compensation strategies to manage the loss of personal resources.

A difference emerged when comparing our results with other studies concerning evaluations of a domestic robot (Scopelliti *et al.* 2004): our study highlighted that elderly people are not afraid of the robot's autonomy, when they can actually understand what a robot can do in the domestic environment. In other words, a representation grounded on unrealistic ideas (as the ones proposed by science fiction) may negatively bias attitudes and expectations.

The overall evaluation of the robot emerged to be very positive, with reference to many specific features, ranging from interaction modalities to the degree of integration in the domestic environment. In this respect, however, the issue

of safety confirmed to play a key role in elderly people's experience and, though not anxious about it, they would like the robot to move in the domestic environment only when a specific task has to be performed.

The most distinctive feature of the robot was undoubtedly associated to its practical utility, as emerged from both a cognitive and an affective evaluation. The robot can help people in the management of everyday activities requiring an efficient cognitive functioning, which is likely to be defective with increasing age. In addition, the presence of such a device in the domestic environment appears to be fundamental in making elderly people feel safer, especially when they live alone. On the other hand, elderly people also showed to be aware of potential troubles with the robot, for both practical and psychological reasons. The practical difficulties are mainly expected with reference to the price they have to pay, both to acquire the assistive robot and to keep it efficient. More importantly from a psychological point of view, elderly people seem to forecast a potential loss in personal autonomy depending on the robot, which may lead them to reduce perceived competence and self-efficacy (Bandura 1977), key factors for a successful ageing of people (McAvay, Seeman, & Rodin 1996; Willis 1996; Lawton 1982). In this respect, they showed to appreciate the possibility to interact with the robot not only passively relying on its capabilities, but also through an active training to enhance their cognitive functioning. Beyond the cognitive component of their attitude, also the affective one emerged to be definitely positive, being the robotic agent depicted in terms of relaxation and interest, and hardly recognized as a source of danger, fear and other negative affects.

The physical aspect of the robot emerged to be an important feature which can help support acceptability. Any allusion to human beings seemed to have an impact on the relationship between elderly people and their domestic environment. In particular, the No-face version of the robot was definitely preferred, and the physical aspect proved to affect also the evaluation of other features which are apparently unrelated. In fact, the No-face version was perceived as less artificial and psychologically distant from the user, better integrated in the home setting and easier to manage. In other words, the better the aspect, the stronger the perception of positive qualities attributed to the robot. This suggests the occurrence of a halo effect, consistently emerging in social sciences with reference to personality judgements (e.g., (Asch 1946)).

Open Issues

Given its exploratory purposes, some shortcomings of the present study should be recognized. First of all, our preliminary results emerged from a small sample, and a stronger statistical robustness is indeed needed. In addition, our study presumably lacks external validity, in that our respondents were rather well-educated and in general in sufficiently good health conditions: the evaluation of a robotic agent which has to be a support for impairments related to ageing may be different when people are in a condition of critical need. Nonetheless, our findings can be considered an intriguing starting point to address the issue of acceptability of robotic

$p < .05$); *better integrated*: ($F_{(1,38)} = 5.65, p < .05$); *ease of use*: ($F_{(1,38)} = 9.36, p < .01$); *low need for repair*: ($F_{(1,38)} = 4.33, p < .05$)

⁷($\chi^2 = 6.11, df = 2, p < .05$)

agents in everyday life of elderly people. One concern has to do with the general role of a domestic robot in the everyday experience of elderly people. In their eyes, the robot is perceived as a practical device: they do not seem to be particularly interested in matters of aspect, shape, cover materials; and they would like it not to resemble a human being. On the other hand, interaction which involves a face-to-face relationship seemed to reduce a feeling of emotional distance from this device. In this respect, it would be interesting to evaluate in further research a possible difference in response to a domestic assistive device which cannot move about in the environment. An environmental system equipped with software, sensory and speaking services would probably be able to perform the same activities provided by the mobile device shown in this study. At the time of writing, we have started experimentation with a system which does not employ a robotic mediator for interaction rather the user communicates with the domestic environment through pervasive microphones and speakers. The aim of this further study is to understand how acceptability might be affected by such a difference.

Conclusions

We have presented here a complete example of experimental methodology applied to a project integrating intelligent systems. The contribution with respect to the workshop topics is to point out the extreme difficulty that is connected to developing intelligent applications, that include P&S technology, for the mass market. On one hand it is worth underscoring that these experiments are able to either reinforce or contradict common beliefs on the utility of general features of these systems. It is also worth noting the amount of work involved in performing this type of experiments. Again the conclusions mentioned in (Cortellessa & Cesta 2006) hold: the effort to apply established methodologies such as the above is rather time consuming but extremely precise and useful. Some work still remains to be done in order to better understand how to speed up and facilitate the application of this kind of methodology in the specific context of mixed-initiative system evaluation, and to understand further the generality of the outcomes.

Acknowledgements. This research is partially supported by MIUR (Italian Ministry of Education, University and Research) under project ROBOCARE (A Multi-Agent System with Intelligent Fixed and Mobile Robotic Components). Special thanks to the colleagues of the Dept. of Computer and Systems Science (DIS) of the University of Rome “La Sapienza” for joint work on the ROBOCARE intelligent environment.

References

- Ai-Chang, M.; Bresina, J.; Charest, L.; Chase, A.; Hsu, J.; Jonsson, A.; Kanefsky, B.; Morris, P.; Rajan, K.; Yglesias, J.; Chafin, B.; Dias, W.; and Maldague, P. 2004. MAPGEN: Mixed-Initiative Planning and Scheduling for the Mars Exploration Rover Mission. *IEEE Intelligent Systems* 19:8–12.
- Asch, S. 1946. Forming impressions of personality. *Journal of Abnormal and Social Psychology* (41):258–290.
- Baltes, P. B., and Baltes, M. M. 1990. *Psychological Perspectives on Successful Aging: The Model of Selective Optimization with Compensation*. New York: Cambridge Univ. Press. 1–34.
- Bandura, A. 1977. Self-efficacy: Toward a unifying theory of behavioural change. *Psychological Review* 84:191–215.
- Becker, M., and Smith, S. 2000. Mixed-Initiative Resource Management: The AMC Barrel Allocator. In *Proceedings 5th International Conference on Artificial Intelligence Planning and Scheduling (AIPS-2000)*.
- Cesta, A., and Pecora, F. 2006. Integrating Intelligent Systems for Elder Care in RoboCare. In Mann, W., and Helal, A., eds., *Promoting Independence for Older Persons with Disabilities*. IOS Press. 65–73.
- Cesta, A.; Cortellessa, G.; Fratini, S.; Oddi, A.; and Policella, N. 2007a. An innovative product for space mission planning: an *a posteriori* evaluation. In *Proceedings of the 17th International Conference on Automated Planning & Scheduling (ICAPS-07)*.
- Cesta, A.; Cortellessa, G.; Pecora, F.; and Rasconi, R. 2007b. Coordinating Heterogeneous Agents to Synthesize *Proactive Monitoring*. In *ICAPS Workshop on Planning and Plan Execution for Real-World Systems, Providence, RI, USA*.
- Cesta, A.; Cortellessa, G.; Pecora, F.; and Rasconi, R. 2007c. Supporting Interaction in the ROBOCARE Intelligent Assistive Environment. In *Proceedings of AAAI Spring Symposium on Interaction Challenges for Intelligent Assistants*.
- Cortellessa, G., and Cesta, A. 2006. Evaluating Mixed-Initiative Systems: An Experimental Approach. In *Proceedings of the 16th International Conference on Automated Planning & Scheduling (ICAPS-06)*.
- Giuliani, M.; Scopelliti, M.; and Fornara, F. 2005. Elderly people at home: technological help in everyday activities. In *IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 2005)*, 365–370.
- Hutchins, E. 1995. *Cognition in the Wild*. MIT Press.
- Lawton, M. P. 1982. Time budgets of older people: A window on four lifestyles. *Journal of Gerontology* 37:115–123.
- Levinson, R. 1997. PEAT - The Planning and Execution Assistant and Trainer. *Journal of Head Trauma Rehabilitation*.
- McAvay, G. J.; Seeman, T. E.; and Rodin, J. 1996. A longitudinal study of change in domain-specific self-efficacy among older adults. *Journal of Gerontology* 51:243–253.
- McCarthy, C., and Pollack, M. 2002. A Plan-Based Personalized Cognitive Orthotic. In *6th International Conference on AI Planning and Scheduling (AIPS-02)*.
- Myers, L. K.; Jarvis, P. A.; Tyson, W. M.; and Wolverton, M. J. 2003. A mixed-initiative framework for robust plan sketching. In *Proceedings of the 2003 International Conference on Automated Planning and Scheduling (ICAPS-03)*.
- Pecora, F.; Rasconi, R.; Cortellessa, G.; and Cesta, A. 2006. User-Oriented Problem Abstractions in Scheduling, Customization and Reuse in Scheduling Software Architectures. *Innovations in Systems and Software Engineering* 2(1):1–16.
- Pollack, M. 2005. Intelligent Technology for an Aging Population: The Use of AI to Assist Elders with Cognitive Impairment. *AI Magazine* 26(2):9–24.
- Refanidis, I. 2007. Managing personal tasks with time constraints and preferences. In *Proceedings of the 17th International Conference on Automated Planning & Scheduling (ICAPS-07)*.
- Scopelliti, M.; Giuliani, M. V.; D’Amico, A. M.; and Fornara, F. 2004. If I had a robot ... peoples’ representation of domestic robots. In Keates, S.; Clarkson, P. J.; Langdon, P. M.; and Robinson, P., eds., *Design for a more inclusive world*. London: Springer-Verlag. 257–266.
- Scopelliti, M.; Giuliani, M. V.; and Fornara, F. 2005. Robots in a domestic setting: A psychological approach. *Universal Access in the Information Society* 4(2):146–155.
- Willis, S. L. 1996. Everyday cognitive competence in elderly persons: Conceptual issues and empirical findings. *The Gerontologist* 36:595–601.