SRS

Multi-Role Shadow Robotic System for Independent Living

Small or medium scale focused research project (STREP)

DELIVERABLE D6.1b

Validation of the concept, testing site preparation and protocol development

Contract number :	247772
Project acronym :	SRS
Project title :	Multi-Role Shadow Robotic System for Independent Living

Deliverable number :	**D6.1b**
Nature :	R – Report
Dissemination level :	PU – PUBLIC
Delivery date :	30-07-11

Author(s) :	Lucia Pigini, David Facal, Marcus Mast, Lorenzo Blasi, Rafa López Tarazón, Georg Arbeiter
Partners contributed :	INGEMA, FDCGO, HdM, HP, IPA, ROBOTNIK, CU
Contact :	lpigini@dongnocchi.it

The SRS project is funded by the European Commission under the 7th Framework Programme (FP7) – Challenges 7: Independent living, Inclusion and Governance

Document History:

Version	Author(s)	Date	Changes
ToC	Lucia Pigini; Lorenzo Blasi	11-05-11	
Toc Revised	Lucia Pigini, Lorenzo Blasi, Marcus Mast, David Facal, Georg Arbeiter, Anthony Soroka	25/05/11	Added part 3: RESEARCH QUESTIONS *Removed part about TESTS INTO CONTROLLED LABORATORY (described into deliverable of wp4)*
V1	Lucia Pigini, Lorenzo Blasi	01/06/11	First draft of each chapter presented
V2	Lucia Pigini, Lorenzo Blasi,	10/06/11	Introductory section added First revision of chapters: • Research questions • Advanced prototype tests
V3	Lucia Pigini, Lorenzo Blasi, David Facal, Rafael Lopez Tarazon	22/06/11	Executive summary added, Manipulation test chapter added
V4	Lucia Pigini, David Facal; Marcus Mast, Lorenzo Blasi	29/06/11	First complete version: • Chapter method completed, • Chapter "first case test completed"
V5	Lucia Pigini, David Facal; Marcus Mast, Rafa Lopez Tarazon	08/07/11	Added draft chapter 7: FIRST NOTES ABOUT SRS COST EFFECTIVENESS ASSESSMENT & SOCIO-ECONOMIC IMPLICATIONS
V6	Marcus Mast, Lucia Pigini,	08/07/11	Objective indicators
V7	Marcus Mast, CU, David Facal, Lucia Pigini	4/08/11	Document integration and revision
Final		(expected December 2011)	

EXECUTIVE SUMMARY

The aim of the evaluation phase with potential users is to investigate and measure effectiveness, usability and acceptability of the advanced prototype to generate feedback for improvement. The research goals are to elicit the participants' acceptance and intention to adopt the new assistive solution, and determine if an effective system enhancing the feeling of autonomy and security at home has been delivered.

This document begins with a summary of the iterative steps conducted so far, involving the stakeholders into the user centered design approach of SRS project. The main research questions are then defined, based on reconsideration of the results achieved in each of the iterative steps, and considering the high priority identified user requirements. Finally, the literature about validation methods is critically analyzed, to find out suitable indicators to assess the defined research questions. The final outcome of this document is a first user validation plan - **first draft.**

Considering that the SRS prototype is not yet fully developed, a final and definitive plan should be likely expected by December 2011.

However, at this stage, the document already reports an SRS prototype draft validation plan, composed of seven main incremental and complementary stages aiming to address the targeted research goal; each stage concentrating on specific aspects of the evaluation of the prototype, with a specific set of tests, experimental protocols and validation methods.

The stages are designed to be incremental in several respects: complexity of tested functionalities, number of people involved, and maturity of the system.

The document ends showing some preliminary information about measures of social costs, considering that cost-effectiveness assessments need to be performed once the validation process is ended, and the purchase cost of the robot will be more concretely identifiable (third year)

TABLE OF CONTENT

TABLE OF FIGURES

1 INTRODUCTORY SECTION: PURPOSE OF DOCUMENT AND CONTENTS

The present document reports a detailed user validation plan - **first draft**. The final and definitive plan is expected by **December 2011**, when the SRS prototype will be in an advanced development state, and the experimental sessions for evaluations with users can be completed and refined. The validation phase with the potential users should be designed in a way that allows investigating the effectiveness, usability, and acceptability of the advanced prototype, so as to generate feedback for improvement.

By now, this document shows the iterative and user centered design approach adopted during the entire project, summarizing the main steps done so far and the steps that still need to be addressed. The document will include a sketch of the validation plan with the users, comprising the test sites locations, main aim of tests, time schedule, and number of users involved in the tests (chapter 2).

The user requirements identified during the whole project through the users' studies are re-considered, in order to identify the main research questions which should be addressed to evaluate the specific SRS prototype developed. Moreover, other research questions are defined in order to evaluate also the usability and social acceptance of the system (chapter 3).

A critical overview of assessment methods is considered essential to allow selecting or helping the design of appropriate evaluation indicators. These indicators will be useful to address the identified research questions, and therefore properly design the final validation plan (chapter 4).

The last chapter reports a first detailed draft of the validation plan divided in three subparts. The plan was conceived to start as soon as possible with the user tests even in the case that the prototype is not yet completely integrated. The chapter starts briefly describing the whole set of experimental sessions to be held in the three test sites, then summarily explains the different aims of each set of tests, and the need of designing different experimental protocols and validation methods (Chapter 5).

The test sites settings predisposition, plan for ethical and safety issues, research questions addressed, experimental protocol description, and validation methods adopted are reported completely and in detail in the following three separated sub-sections:

- Interface usability tests - second iterations - aiming to go one step further towards the realism of the interaction (section 5.1).
- Real home environment case test, to be held in Stuttgart, to address first SRS experience outside the lab and the first users impressions (section 5.2)
- Manipulation and visualization tests,–to be held in San Sebastian, to address the main behaviour of the robot with real users (section 5.3)
- Advanced SRS prototype tests, to be held in Milan, to address the integrated SRS functionalities and the scenarios effectiveness with real users in a home environment (section 5.4)

To guide each participant during the scenario execution in the Milan tests, a detailed screenplay will have to to be produced for each scenario. A first draft version of it is available in Appendix 1.

The complete questionnaires adopted will be reported in Appendix 2, which, at this stage, is still under discussion.

The last chapter reports preliminary information about the SRS cost effectiveness assessment and Socio Economic implications. However, it should be considered that this kind of analysis should be conducted in conjunction with the dissemination and exploitation tasks, and can really start just once the validation process is concluded and a purchasing price of the robot is concretely estimable (third year) (chapter 7).

Appendix 3, reports an example of usage of the proposed method for assessing the social costs, SRS cost effectiveness assessment, and Socio Economic implications.

2 ITERATIVE VALIDATION PROCESS WITH REAL USERS IN SRS PROJECT

The SRS project is based on a user centered design (figure 1), involving potential stakeholders from the beginning of the study and at each step of the development of the prototype.

Therefore, from the beginning of the project an iterative approach has been adopted. Before the explanation of the next steps of the SRS evaluation plan, a short summary is provided of what done so far to involve the final users in this user centered and iterative approach to the project.

- The project started involving potential users in the focus group, aiming at finding the general features of stakeholders, their predisposition to new technologies, and their needs (not only physical but also social and related to privacy).
- The first results (see D1.1a and Mast M. 2010), provided to the researchers enough information to design ad-hoc questionnaires and to select appropriate validation methodologies (see for example SOTU questionnaire), in order to achieve quantitative results (see D1.1, D2.1 and Pigini L. 2011) about the users' needs and expectation from a service robot.
- These results enabled the researcher to define specific user requirements, to translate them into technical requirements (D1.2, D1.3), and to hypothesize the first list of SRS scenarios.
- These first list of SRS scenarios were presented again to the potential users through the method of visual simulation, providing a final validation of SRS concept in term of scenario selection, human robot interaction devices, target population (both local user and remote operators), and robot aspect (D6.1-interim report).
- In the mean while an ethnographic research allowed to find out the social, economic, and environmental context of the already defined target population: the elderly people and the informal caregivers (ethnographic research report and Facal D. 2011). The research also described the new identified stakeholders: the 24 hour service operators (D2.2), and with them, the need for another more specific interaction device.
- The achieved results enabled to conduct first usability tests based on a mockup of user interfaces, which generated feedback to improve the next step of interfaces development (D2.2).
- In the mean while, requirements referring to users were taken in consideration and selected according to their importance, as shown in (Table 20 - Full prioritized requirements list- D1.1a)

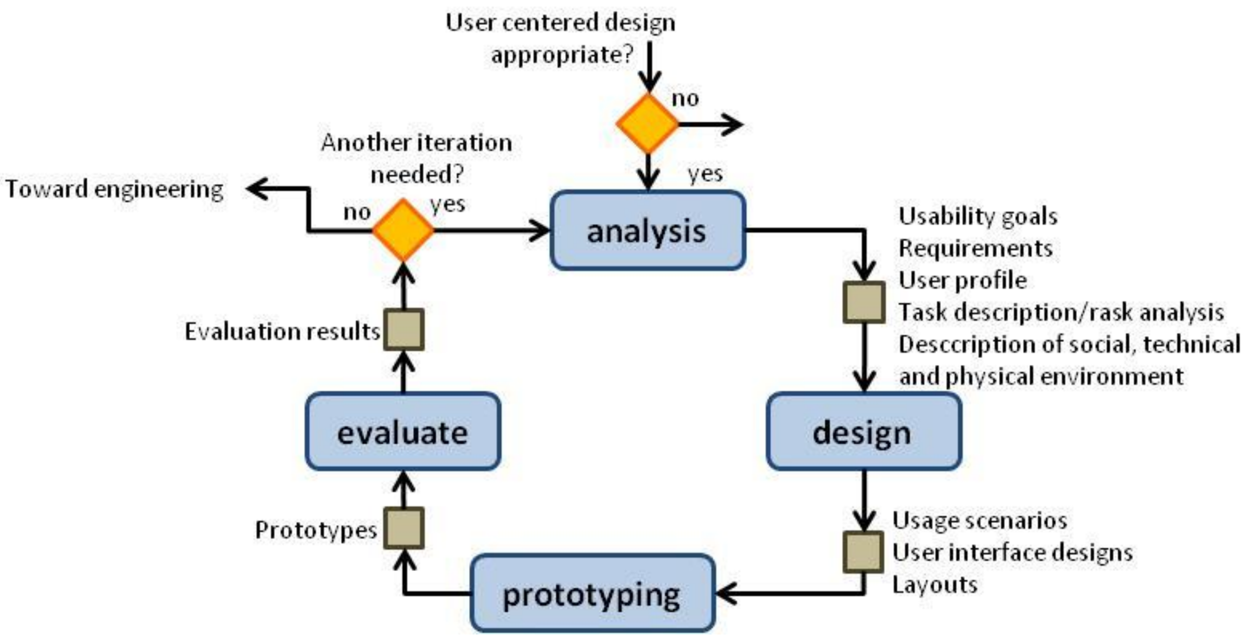


FIGURE 1: SRS PROJECT LIFECYCLE FROM AN USER-CENTRE DESIGN PROCEDURE (ADAPTED FROM BURMESTER, 2007)

The results achieved so far allowed the partners involved in the technical tasks to proceed with the development of advanced technology (together with technical tests continuously performed while developing).

This will lead, before the validation phase with real users in a real home environment, to the validation phase of the technical and functional requirements, which will be performed in a controlled laboratory. This is considered of primary importance in order to avoid the likely failure of the tests with users. The technical validation phase will be part of the integration meetings (Wp5). The outcome will help to design the final scenarios in an appropriate way. In particular the technical evaluation for the perception components is expected to answer to these questions:

- Possible locations where the objects can be placed
- Which objects can be detected, which not
- Robustness of detection, special focus on object selection by user
- Identification of exceptional cases: obstacles, occlusions, cluttered scenes
- Evaluation of the mapping pipeline for environment modelling

Of course, also manipulation, navigation, user interfaces, decision making and learning has to be evaluated and also the integration of all the components to a fully functional system. As soon as the prototype is technically tested, the user validation tests will start, with the research goal of determining the participants' acceptance and intention to adopt the new assistive solution, and to determine if the developed solution would:

- **For elderly people: enhance the feeling of autonomy and security at home, without making them feel a sort of control over their own life.**
- **For family members or other private caregivers: provide a less time- and effort-demanding solution to elderly care.**

This macroscopic target will be addressed into a validation process which is composed of seven main incremental and complementary stages; each one concentrating on specific aspects of the user's evaluation of the prototype, with a specific set of tests, experimental protocols and validation methods.

Table 1 shows the complete time schedule of the iterative user's evaluation plan, including the main steps completed so far (light blue) and the foreseen steps for the user validation plan (dark blue). Table 1 includes a short description of test sites, the main aim of tests, and the number of users involved in the tests. The table also shows for each WP the number of the task in which the main part of the validation phase will be developed, and the number of the deliverable in which the results of each part of the evaluation plan is (will be) described. In particular, a detailed user **validation draft plan** for provided evaluations - no. (3a), (3b), (4), (5), (6) and (7) is provided in chapter 5 of the present document.

Because the **final usability tests (task 2.6 - WP2) will be conducted in conjunction with acceptability and effectiveness tests (task 6.3 - WP6)**, the final results will be reported into two deliverables: D2.2b specifically dedicated to usability results, and D6.2b dedicated to overall results and conclusions about the user validation results (month 36).

TABLE 1 - USERS' VALIDATION PLAN AND SCHEDULE (UI-PRO=PROFESSIONAL INTERFACE, UI-PRI =PRIVATE INTERFACE, UI-LOC= LOCAL USER INTERFACE)

No., Time, Duration	Evaluation Description	Participants	Site, Leader	Task, Report, Report Date
(1) 2011-01 1 week	Initial usability test of user interfaces UI_LOC and UI_PRI using clickdummies and video-simulated robot behavior	7 elders 5 informal caregivers	Usability Lab HDM, Stuttgart Lead: HDM	Task 2.6 D2.2a 31/07/2011
(2) 2011-01 3 weeks	User evaluation of SRS concept (questionnaire-based survey)	30 elders 23 informal caregivers	3 sites: Italy, Spain, Germany Lead: FDGCO	Task 6.1 D6.1a 31/01/2011
(3a) 2011-12-05 2 weeks	Usability test of initial version of real user interfaces (no longer clickdummies) of UI_PRI and UI_LOC (basic functions present, connected to Care-o-Bot simulation; non-implemented functions will be simulated) & usability test of first version of UI_PRO (clickdummy). Focus on real interactive behavior, i.e. users will act simultaneously.	10 elders 10 informal caregivers 5 tele-assistance staff	Usability Lab HDM, Stuttgart Lead: HDM	Task 2.6 D2.2b 31/01/2013
(3b) 2011-12-05 2 days	Usability test of initial version of real user interfaces of UI_PRI and UI_LOC. This is a fork of test 3a, with the same goals but using the real robot (not the simulation). The test has a shorter duration and less participants due to restrictions in the availability of the Care-o-Bot.	2 elders 2 informal caregivers	IPA kitchen environment, Stuttgart Lead: HDM	Task 2.6 D2.2b 31/01/2013
(4) 2012-01-16 2 days	Real home pre-test with focus on technical performance in real apartments (informing the developers and task 6.3), necessary adaptations, and users' perception of robot	2 elders 2 informal caregivers	Private home, Stuttgart area Lead: HDM	Task 6.3 **D6.2** 31/01/2012

(5) 2012-02 4 weeks	Local manipulation test	10 elders	IZA Care Center, San Sebastián, Spain Lead: ING	Task 6.3 **D6.2** 31/01/2013
(6) 2012-04 4weeks	Remote manipulation and visualization test with UI_PRO	10 tele-assistance staff	IZA Care Center, San Sebastián, Spain Lead: ING	Task 6.3 **D6.2** 31/01/2013
(7) 2012-05 4 weeks	Final SRS entire-system evaluation and final user interface usability test (leaving some time for addressing issues before final demonstration in 2013-02)	10 elders 10 informal caregivers 3 tele-assistance staff	FDGCO real home environment, Milan Lead: FDGCO & HDM	Task 6.3 & Task 2.6 **D6.2** 31/01/2013 **D2.2b** 31/01/2013

3 RESEARCH QUESTIONS

As already mentioned in the conclusions of D1.1, there are two main aspects that must be taken into account when working in a user-centered design framework: usability and acceptability.

Usability is the perception of the ease of using and learning to use the new devices developed to control the robot. Usability tests have the aim to detect problems related to the use of the systems, in order to improve the subsequent development stages until the feedback from the users becomes satisfactory. For this reason, usability tests of the human-robot interface devices have been carried out from the early stage of the project, with the first iteration in month 12. A second iteration will be carried out when the devices development has reached an intermediate level (month 23), and a third iteration will be carried out towards the end of the project. The third iteration will focus on the usability of the entire robotic system rather than on singular interaction devices.

Social Acceptance instead is defined (Dillon, 2001) as “the demonstrable willingness within a user group to employ technology”. Therefore, the task of assessing the social acceptance in this project could be considered as the evaluation of “the satisfaction and the intention to adopt the proposed robotic solution to solve the identified user needs fulfilling the user requirements”.

Overall, the aim of the evaluations is to generate feedback for improvement in terms of effectiveness, fulfilling of user’s expectations, usability and acceptability of the advanced prototype.

In order to address these main validation goals, appropriate research questions for evaluating the SRS system have to be identified. Considering first of all the evaluation of the fulfillment of the user needs and requirements, table 2 shows the research questions identified to evaluate the effectiveness, usability and acceptance of the system. These questions focus on the user requirements ultimately considered of high importance, as extracted from table 20 - Full prioritized requirements list - of D1.1a.

TABLE 2 - HIGH IMPORTANCE USER REQUIREMENTS AND CORRESPONDING RESEARCH QUESTIONS

Requirements	***Research question for system evaluation***
R1 The system should be understood to be usable and acceptable	*Learnability*
R08-R09 The users’ objects selections is translated into the correct system	*Efficiency in ADL task completion*

actions sequence (the system recognized and identifies the selected objects (shapes, colors, letters on food boxes, numbers on microwave display) and is is able to firmly grasp objects without damage them (i.e. bottle, books).	
R34 Communication of action outcomes during performance of the robot, in order to maximize the awareness of the elderly user. **R35** No robot movement should happen without initial confirmation by the user who is in direct physical contact with the robot. **R36** There should be a clear indication on the robot side if the robot is in autonomous mode or in remote controlled operation	*Secure in ADL task completion*
R07A The system should help elderly people with mobility issues such as reaching objects.	*Improving the autonomy*
R02 A flexible system of communication and advice sending should be designed, because family caregivers like the system but they do not want to be on-line 24 hours-a-day (related to psychological burden).	*Improving of communication and interaction modality*
R14 The system should help with coping with unexpected, emergency situations such as falling.	*Improving safety*
R22 The system allows communication between user and remote operator, so providing the user with help in housekeeping and mobility could be an indirect way of making him/her able to use more spare time for social contacts	*Acceptable from Psychological/emotional point of view*
R23 Only authorized persons have access to the remote control of the system **R24** Authentication procedure as a protection of the access to be included for both family caregivers and professionals. **R26** Avoid possibility of access to the system without explicit consent of the elderly, including non authorized access of authorized remote operators **R27** If remote operator changes within one session, the elderly user must be informed **R28-R29-30-31** Personal information data protection managed in a safe way **R32** An —on/off|| mode to be implemented in order to protect privacy in very personal moments. The access to the —on/off|| mode could be adaptable attending to the specific frailty of the elderly user. **R33** Verification of the plans of action by asking the elderly user before it starts acting.	*Safeguarding of Privacy and Ethics* *(e.g. Avoiding sense of control over one's life)*
R03 The system is able to maneuver in narrow spaces: usually elderly lives in small apartments full of furniture.	*Easy Integration in the private home*

Other research questions however have also to be investigated in order to assess the global usability and acceptability of every new product under development, including:

- *Advantages/disadvantages perceived*
- *Attractiveness*
- *Comfort perception*
- *Eligibility (Intention to adopt)*
- *Usefulness*

Finally, from a technical point of view, the effectiveness of the SRS scenarios execution will be addressed in terms of:

- *evaluating the success of each single task execution*
- *evaluating the time needed to complete the tasks*
- *describing the eventual problems occurred in tasks completions*

The complete list of research questions, which need to be investigated to assess the macroscopic research goals, is reported in table 3. To find measurable and standard indicators in order to address all the research questions related to the evaluation of the prototype, suitable validation methods have been selected and developed for the purpose. The next chapter discusses the methods which could be adopted to better answer these goals.

TABLE 3 COMPLETE LIST OF RESEARCH QUESTIONS, ADDRESSING THE MAIN RESEARCH GOALS

Main research goals	Complete list of research questions	Validation methods to address research questions
Effectiveness	*evaluating success of each single task execution*	?
	evaluating time needed to complete the tasks	
	describing eventual problems occurred in tasks completions	
SRS peculiar requirements	*Efficiency in ADL task completion*	
	Secure in task completion	
	Improving the autonomy	
	Improving of communication and interaction modality	
	Improving safety	
	Easy Integration in the private home	
Usability/ learnability	*Easy to Learn*	
	Comfort perception	
	Attractiveness	
Acceptability/ intention to adopt	*advantages/disadvantages perceived*	
	Acceptable from Psychological/emotional point of view	
	Safeguarding of Privacy and Ethics	
	Usefulness	
	Eligibility (Intention to adopt)	

4 VALIDATION METHODS

In this section, a critical **overview of assessment methods is provided**, allowing selecting and properly designing the final validation plan detailed in chapter 5.

The main focus of the validation process with the users is to generate a set of recommendations for improvement, in order to obtain user's feedback including views, feelings, critical suggestions, etc. In this sense, the best approach to get this kind of feedback should be the qualitative investigation of the user's perception once the SRS system's potential is tested. Suitable qualitative methods to achieve these kind of results are represented by methods such as "think aloud" (Lewis C. H., 1982), "behavior observation" (Altmann J., 1974), and open questions administered for example through focus group methodology (Krueger & Casey, 2000).

However, quantitative methods are helpful to quantify the overall feedback on usability, acceptability, satisfaction, intention to use, and on more specific features of the developed system. In this sense, suitable measurable standard parameters are needed. This kind of results can bring to a final evaluation enabling also to analytically compare results between groups of stakeholders, or to compare results achieved in following evaluation stages. Table 4 reports a short analysis of quantitative methods.

TABLE 4 - QUANTITATIVE METHODS ANALYSIS

Method	What measures	Critical considerations	Fitting with the research questions
AttrakDiff	A method to measure	Yes,	• Acceptable from

(Hassenzahl, M., 2003)	attractiveness, hedonic, and pragmatic quality of interactive systems	The AttrakDiff™ provided valuable input regarding emotional perceptions of the users in firsts usability SRS tests and in other projects such as the HERMES project (Cognitive Care and Guidance for Active Aging- http://www.fp7-hermes.eu/), Information provided by the AttrakDiff was coincident with qualitative data. Questions adaptable for every kind of product under development. Potential users can easily indicate their perception of the system.	Psychological/emotional point of view (R22) • Easy to learn (R1) • Improving of communication (R02) • Outcome expectations • Advantages/disadvantages perceived • Attractiveness • Comfort perception • Eligibility (Intention to adopt)
NARS (Bartneck, 2005)	Negative Attitude toward Robot Scale is a method which allows obtaining a psychological Index about attitude toward Robots.	No, attitude toward robot already investigated at the beginning of the project	• Psychological/emotional point of view R22
PANAS (Watson et al, 1998)	The Positive and Negative EffectSchedule: a self report schedule to measure the positive and negative effect.	No, In order to asses effects related to the interaction with the robotics system, the AttrakDiff matches better the research aims outlined in the previous chapter	• Psychological/emotional point of view R22
UX Questionnaire (Laugwitz et al, 2008).	The user-experience questionnaire enables to measure user experience evaluation factors: embodiment, emotion, human-centred perception, feeling of security, and co-experience.	Partially, it can help to develop ad-hoc questions It contains some suitable question for the evaluation purpose like "I felt afraid with the robot" or "I liked that the robot understood my command". However, other questions are not suitable for this project, e.g. UX questionnaire: "I enjoy talking with the robot", "I liked that the robot has human-like: face, ears, eyes".	• Improving safety R14 • Acceptable from Psychological/emotional point of view R22 • Secure in task completion (R34, R35, R36) • Attractiveness
SCAI (Andrich R.,2007)	instrument: Siva Cost Assessment Instrument: Analysing the cost of assistive technology programmes	Yes, It helps operators and users to estimate the cost of choosing a solution for autonomy (aid, personal care, environmental adaptations, etc.) and to compare the various possible solutions in terms of economic cost. It needs to be adapted to this project in order to be applied in a prospective way with a technology prototype	• Methods for the system economic assessment (as input to the economic study of task 6.4).
PIADS: The Psychosocial Impact of Assistive Devices Scale (Demers L,2002)	It measures the quality of life (QoL) impact related to the use of assistive technologies from the disabled's point of view.	Partially, it can help to develop ad-hoc questions It (EDWARD M. GIESBRECHT 2008) measures quality of life using three component subscales : Adaptability, competence, and Self-esteem. It is a self-completion questionnaire to be filled in by the user after he/she has acquired a certain familiarity and competence for the proper use of an assistive device. However, it could be adapted to our project by asking the questions to people in a predictive way.	• Acceptable from Psychological/emotional point of view • Comfort perception • Outcome expectations • Improving the autonomy • Improving safety • Usefulness • Advantages • Efficiency in ADL task completion
UTAUT model (Venkatesh et al, 2003)	Model developed to evaluate technology acceptance in term of performance expectancy, effort expectancy, attitude towards using technology, self efficacy, forms of grouping, attachment, and reciprocity.	Partially, it can help to develop ad-hoc questions The UTAUT scale is based in the well-established Technology Acceptance Model (TAM), which is the most commonly used model in the field of technology acceptance On the other hand, one of the main criticisms of the UTAUT scale is that it has its origins in a work-related context, and is focused on the acceptance and use of work-related ICT and software. Therefore, the utilization context is assumed to play a major role, and the related motivation concerning technology use and perceived benefits vary. Further, the heterogeneity of elderly users (gender, age, experience, and voluntariness) could have an even stronger impact than the	• Acceptable from Psychological/emotional point of view R22 • Willing to use the technology - Eligibility (Intention to adopt). • Outcome expectations (performance expectancy, effort expectancy) • Advantages/disadvantages perceived • Usefullness • Facilitating conditions

		individual factors on the acceptance. In the HERMES project, difficulties were found when the scale was used with non-existing technology through a narrative approach.	
Robot-centric measures (Steinfeld et al, 2006,	**Methods focusing on system performance; e.g. does everything work as it should?**	**No,** **They do not focus on whether or not the interaction with humans is appropriate, easy, enjoyable, etc.** **They should be adopted to evaluate the success of single technological innovations, where they would be suitable for usability and technical evaluations of the devices under development (Wp2-Wp4)**	
The SUS (Brook, 1996).	**System Usability Scale) is a standardized questionnaire addressing the effectiveness, efficiency, and satisfaction with a system**	**Yes** **It consists of 10 items and yields a single number representing a composite measure of the overall usability of the system being studied.** **Questions like "e.g. "I think that I would like to use this system frequently" or "I found the system unnecessarily Complex"…can be answered on a 5 scale ranging from "strongly disagree" to "strongly agree"** **Very generic, adaptable to each new technical system, it does not investigate all the aspect but is very short and easy to complete .**	• **efficiency** • **satisfaction with a system**
AmI Appliances Questionnaire (Allouch, 2009)	**Questionnaire designed to examine Ambient intelligence appliances for domestic settings perception**	**Partially, it can help to develop ad-hoc questions** **Based on other acceptance theories and models of technology such as the technology acceptance model (TAM) and unified model of acceptance and use of technology (UTAUT).** **No need of having experience with the product, a questionnaire dedicated to ambient intelligent systems, such as intelligent fridges and mirrors, but easily adaptable to every technology for the home.**	• **Attitudes** • **Intentions** • **Investigate adoption** • **Outcome expectations**
Perceived control in ubiquitous computing (Spiekerman, 2005)	**Ease-of-use of the PET, Information Control and Helplessness scales**	**Partially, it can help to develop had hoc questions about perceived control.** **This construct complement perceptions about improvements in independency – autonomy, which are not always transparent for frail elderly people.** **Perceived control in a UC environment is the belief of a person in the electronic environment acting only in such ways as explicitly allowed for by the individual. User-friendly technology design are actually targeted to increase the perceived control of the users, improving perceived control over information use and maintenance** **Example of statements : "I perceive perfect control over the activity of the system"; . "I perceive the system can help me to control over the things that happen to me / the difficulties I have"**	• **Perceived control in daily life** • **Perceived control in using technologies**
Scenario-based questionnaires (Gonzalez et al., 2011)	**Acceptance of a scenario – task trough specific questions related to the actions performed**	**Yes,** **Scenario-based assessments have been successfully applied in ICT projects for elderly people, alone (iWard) or combined with other methodologies (HERMES, Companionable, Soprano).** **Target-oriented questions better fit with scenarios/tasks to be assessed when compared to standardized questionnaires.**	• **Flexibility** • **Adaptability** • **Within context** • **Within the aim of the project**
RACER methodology **(Wiedmann, T., 2009)**	**Method developed for evaluation of methodologies and indicators**	**Yes, to verify that the general approach of the evaluation take into account RACER CRITERIA**	**Referring to the project:** • **Relevant** • **Accepted** • **Credible** • **Easy** • **Robust**

Finally, performance evaluation should also be measured through objective indicators regarding task execution, errors rates, and time required to complete the tasks. As an example of this kind of measures the systematic procedure developed by Parson, White, Warner, & Hill (2006) can be taken. The aim of this procedure is to obtain numerical estimates of the effectiveness of task analysis for a wheelchair mounted manipulator for use by severely disabled persons, measuring indexes of the use of various input devices, such as the style of interaction selected, the nature and number of user tasks addressed, task completion times, and the number of available or selected control modes.

In order to take care of the specific research questions linked to the main goal concerning the "SRS peculiar requirements", a set of ad-hoc questions need to be developed for the purpose. This is because the SRS system is an innovative product, and there are no existing evaluation methodologies appropriate to address some of its features. These ad-hoc questions can be defined at this stage of the project just considering the user requirements specifications that are under technical development, and considering possible difference between the initial expectation and the final realization (due to technical limitations, feedback achieved during usability, and technical tests…).The final and definitive questions need to be prepared just when the real prototype is at an enough advanced state of development.

Ad-hoc questionnaires based on scenario development (Gonzalez et al., 2011) have been successfully applied in ICT projects for elderly people, alone (iWard), or combined with other methodologies (HERMES, Companionable, Soprano). As an example, the validation procedure of the iWard project has been based on specific questions (i.e. "The user personalizes the robotic creature 0 - 1 - 2 - 3 – 4"), evaluation criteria (i.e. "Patients feel comfortable with the robot Yes/No"), and concrete evaluation criteria for each scenario (Guidance, Cleaning, Delivery, Monitoring, and Surveillance), each presented in separate evaluation sheets (Oztemel et al., 2008). The self-organizing swarm of service robots, modular design of robot equipment, and usability and unique user interface were also assessed separately. iWard is a project close to the technology to be developed and to the validation aims to be covered in SRS. Although iWard's context of application is a hospital environment, which is clearly different to the home environment targeted in the SRS project, the basic approach can be similar, taking advantage of the flexibility, adaptability, and within-context possibilities of the scenario-based questionnaires.

Ad-hoc questionnaires have also been widely used in the scientific literature, (Caulfield, 2010; Cherubini, Oriolo, Macri, Aloise, Babiloni, & Cincotti, 2007; Mataric, Eriksson, Feil-Seifer, & Winstein, 2007; Parson, White, Warner, & Hill, 2006; Tapus, Tapus, & Mataric, 2008). Cherubini et al. (2007) tested a prototype system that provides remote control of home-installed appliances, including AIBO, through single-step, semi-autonomous and autonomous operating modes with different levels of interaction. The performance of the navigation system was shown by experiments (a comparison between the navigation modes and a autonomous battery charging operation) and, complementarily, the system underwent clinical validation conducted in Fondazione Santa Lucia in Rome with eight subjects suffering from Spinal Muscular Atrophy type II, and six subjects suffering from Duchenne Muscular Dystrophy, in order to obtain the assessment through patient feedback.

The appliances used were installed in the experimental apartment and, as a result, all of the patients were able to master the system and control AIBO within 5 sessions. Most of the patients reported in an ad-hoc questionnaire to have experienced 'the possibility to interact with the environment by myself' (Cincotti, Mattia, Aloise, Bufalari, Schalk, Oriolo, Cherubini, Marciani, & Babiloni, 2008). Matarić et al. (2007) conducted a study involving an autonomous assistive mobile robot that aids stroke patient rehabilitation, by using exit interviews and ad-hoc questionnaires about their impressions of the robot that the participants filled-out after each session. Sessions took place in rehabilitation research labs at the University of Southern California Health Sciences campus, with stroke patients with sufficiently mobile to perform the activities in the experiments. Every evaluation session comprised six experimental runs; in all experiments the robot asked the participant to perform one of two activities: the first activity was to shelve magazines; the number of magazines shelved was used as the final evaluation; the second activity consisted of any voluntary activity that involved the movement of an affected arm. Authors used questionnaire data to show that the robot was well-received by both patients and physical therapists and has a positive impact on their willingness to perform prescribed rehabilitation exercises. Parson et al. (2006), after a familiarization stage, a feeding task and a pick and place task, conducted an interview stage using a semi-structured questionnaire to allow a more formal recording of user impressions. These authors recognized that, although questionnaires are of limited value for single-user studies, the approach allows them to structure the interview, ensuring that issues addressed by similar studies were included and facilitating future acceptance evaluations.

From the user side, variables such as perceived control have also been taken into account in design and validation procedures. About perceived control, gerontologists have been increasingly interested in this notion because such perception, which is closely linked with the concept of successful aging, is threatened by age-related changes such as declining health and functional losses (Chipperfield & Greenslade, 1999). Perceived control interacted with functional health and functional status for frailty older adults (Infurna, Gerstorf & Zarit, 2011; Levy, Slade & Kasl, 2002; Menec & Chipperfield, 1997) in terms of perceived health, hospitalization, or mortality. Older adults are less likely than the young to believe there are things that can be done to control aging-related declines because of changes in cognitive performance, health, and well-being (Lachman, 2006). Mechanisms linking perceived control and positive outcomes include adaptive behaviors such as strategy use, physical activity and physical aids. Individuals who perceived low levels of control have been found to need more care and to use more health services (Chipperfield & Greenslade, 1999; Menec & Chipperfield, 1997). In technology projects, perceived control has also been related to privacy issues (Spierkermann, 2005).

In short, we have analyzed different evaluation methods:

- Qualitative methods, such as think-aloud and behavioral observation.
- Quantitative methods, including standard pre-existing questionnaires and ad-hoc questionnaires developed for the SRS purpose
- Objective and measurable indicators for performance evaluation.

All these methods can be combined in order to achieve a comprehensive evaluation of user acceptance, user experience and intentions to use the technology.

A multiple approach so appear to be the more suitable:

- providing measurable feedback, to obtain objective and comparable results moreover answering to effectiveness and learnability research goals;
- providing quantitative feedback through ad-hoc developed questions, to obtain specific results mostly related to those features of the SRS that are impossible to be analyzed using pre-existing methodologies,
- providing quantitative feedback about acceptability, attractiveness, usability, intention to adopt, etc., using standard questionnaires already available in literature and suitable to our case,
- finally providing also qualitative feedback, respect to all research questions related to the social and psychological impact (that is, almost all except those related to effectiveness), in order to understand in deep the mechanisms affecting the quantitative results.

The following Venn diagram (figure 2) shows how the mixed approach evaluation methods would fit with the research goals, which can be summarized as *“to determine the participants’* ***acceptance*** *and* ***intention to adopt*** *the* *new assistive solution, which will enhance the **feeling of autonomy** and **security** at home **without** being perceived as a sort of **control over one's own life**".*

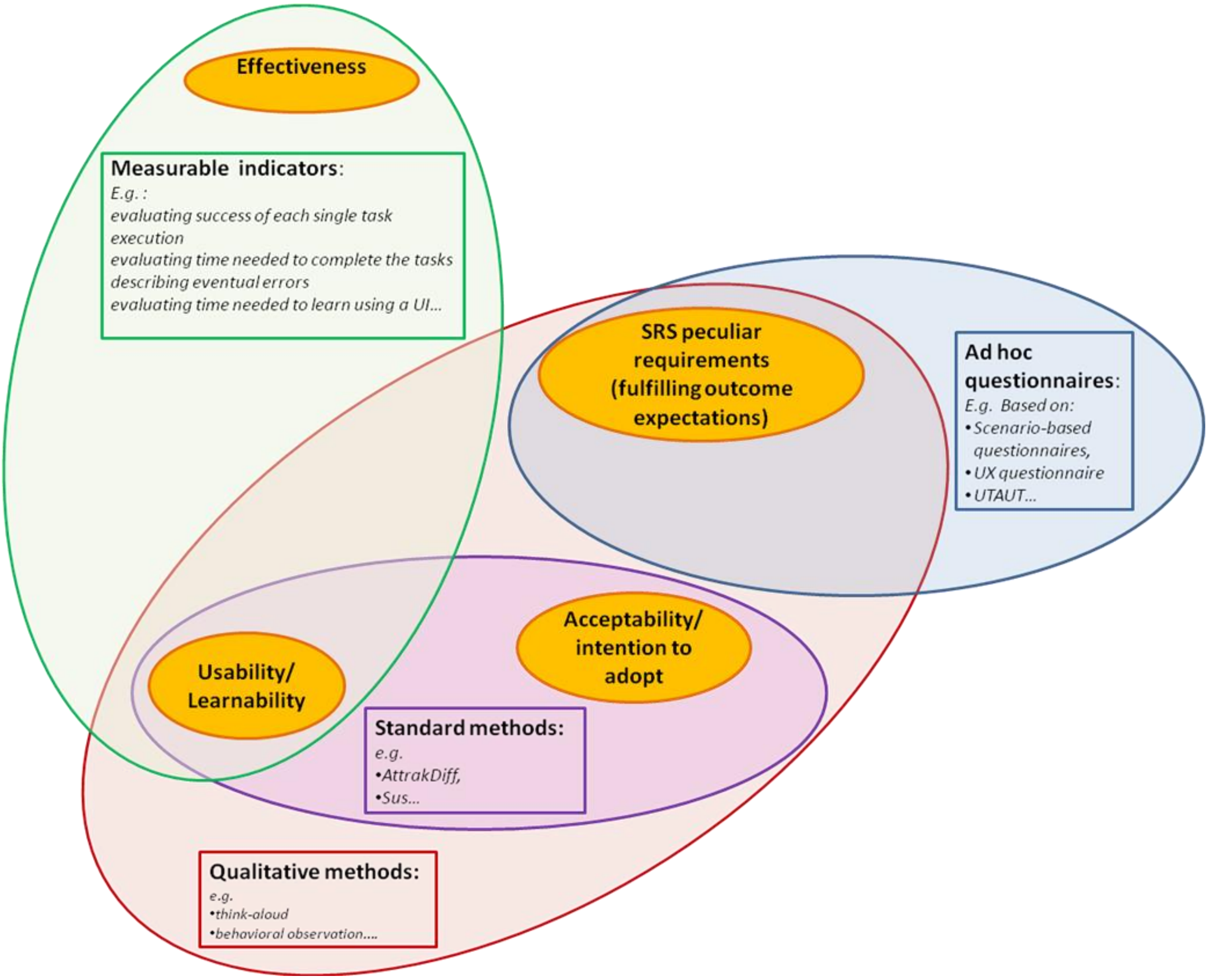


FIGURE 2: THE PROPOSED MIXED EVALUATION METHODS APPROACH COVERING THE RESEARCH GOALS (IN THE YELLOW BALLOONS)

5 PROTOTYPE VALIDATION PLAN (FIRST DRAFT)

This chapter reports a first draft of the validation plan; the final plan is expected in December 2011.

The SRS prototype validation plan is composed of several incremental and complementary stages, as indicated in table 1, chapter 3. The stages are complementary because each one will concentrate on specific aspects of the evaluation of the prototype, with a specific set of tests, experimental protocols, and validation methods. The stages are also incremental in several respects: complexity of tested functionalities, number of people involved, and maturity of the system.

Execution of the evaluation tests will involve different test sites; they are located in Germany (Stuttgart), in Spain (San Sebastián), and in Italy (Milan), where the final SRS demonstration will also be shown.

The first set of tests (December 2011), which will be performed at the Fraunhofer IPA model kitchen and at the Usability Lab of Stuttgart Media University, regard the second iteration of usability tests of the SRS user interfaces.

A whole system pre tests will be performed at a private home located in Stuttgart (January 2011), and will be used to determine which specific difficulties the SRS prototype could encounter in a real home environment, with respect to the controlled laboratory environment in which it has been tested so far. Some possible tests in this first stage will be for example: check the time needed to learn a new map for a real house, check functionalities under real lighting conditions, verify reception of WiFi antennas, etc. All the identified difficulties will be noted and this list will be the basis for preparing the Milan testing site, which could be considered as another unknown and new home environment. Another focus of these first field tests will be the possible impacts on the apartment and the restrictions created by the presence of the robot, concentrating thus on user requirements related to the integration in a private home and the safeguarding of privacy.

Two set of tests will be performed at IZA Care Centre of San Sebastian (February 2012 and April 2012). These tests will have to determine effectiveness, acceptability and usability of the arm subcomponent. They will involve a greater and more various set of users, both Local Users (LU, the elderly) and Remote Operators (RO). Evaluation in this phase will be related to the functionalities satisfying requirements such as safe manipulation, secure grasping, and learning.

Finally, the more complex set of tests will be performed in the Milan apartment (May 2012), and will finally focus on the effectiveness, usability and acceptability of the whole prototype in an advanced stage of development. Evaluation will be related to the major user requirements, and especially on the improvements in ADL performance through the execution of real-life scenarios experienced by real users.

A complete and detailed description of each set of tests including objectives, preparation work, method, and evaluation procedure, is provided in the following sections.

5.1 INTERFACE USABILITY TESTS - SECOND ITERATION

5.1.1 OBJECTIVES

The second usability test of the SRS user interfaces aims to go one step further towards the realism of the interaction. These tests builds on:

- First version of clickdummies of UI_LOC and UI_PRI
- Usability test results of the first version clickdummies
- Second, improved versions of the two clickdummies having been passed on to development

The test artifacts will be:

- First version of the real (non-clickdummy) interactive user interfaces of UI_LOC and UI_PRI (mobile devices for elderly and informal caregivers)
- First version of gesture-based local interaction
- First version of UI_PRO clickdummy

As in every usability test, the overall objective is to find usability problems and improvement areas in the tested artifacts. Furthermore, specific interaction solutions that are considered candidates for interaction design patterns will be evaluated for their appropriateness and ability to answer the underlying problem.

5.1.2 Method

This test will focus on the real interactive behavior, i.e. informal caregivers and elderly users will interact simultaneously in the same test session with the robot. The degree to which realistic interaction will be possible depends largely on the state of the SRS development. Functionality that will be not yet available will be improvised (for example with a wizard-of-oz approach). Ideally, this test should be carried out entirely using the real robot. However, this is not feasible due to availability restrictions. Therefore, there will be two parts of this test, carried out at two test locations:

Fraunhofer IPA model kitchen, Stuttgart: 2 elders and 2 informal caregivers will interact with the SRS system through gestures, UI_LOC, and UI_PRI.

Usability Lab at Stuttgart Media University: 10 elders, 10 informal caregivers, and 5 tele-assistance staff will participate. The Care-o-Bot simulation will be used instead of the real robot. In addition to UI_LOC and UI_PRI, the initial clickdummy version of UI_PRO will be tested with a similar approach to the evaluation of the first iterations of UI_LOC and UI_PRI clickdummies, i.e. using videos and improvised interaction to substitute real interaction.

The method will involve task-based interaction, thinking-aloud, structured interviews or a questionnaire on the appropriateness of potential interaction patterns, and the AttrakDiff questionnaire to obtain quantitative data on the overall user experience. Exact methodology depends largely on the state of the development at the time of testing and will therefore be specified when test preparation commences.

5.2 WHOLE-SYSTEM PRE-TEST IN REAL HOME

The main goal of this test is to determine the specific difficulties that the SRS is expected to encounter in a real elderly home environment as opposed to the lab environment where it has been used so far. This could for example be due to door widths, lighting conditions, carpets on the floor, objects, small corners that cannot be reached, weight of the robot, reception of WiFi antennas, etc. The robot will be used for navigation and manipulation tasks taken from the SRS scenarios. This test can be considered a pre-test for the final SRS evaluation (task 6.3) with focus on technical aspects. All difficulties that emerge will be noted. The results of this test will be the basis for preparing the final testing site in Milano. The Milano apartment will be prepared in such a way that it encounters all the difficulties that the robot would encounter in a home, with the exception of such obstacles that would hinder the operation of the robot. Further objectives are to determine the impact on, and restrictions to, elderly users due to the robot (e.g. furniture that needs to be removed or moved), perception of the robot by the users, potential safety hazards, and privacy-related issues.

5.2.1 Objectives

In particular, the objectives are:

- Determine the preparation of the final SRS evaluation: determine the robot's operational limits in a real home environment (e.g. problems with small room sizes for object detection, with floor types, maximum height of an object, maximum weight of an object, required light necessary for visual system, etc.)
- Determine the development: determine the current progress of SRS system components (generate a list of all the non-working components and a list of all issues with working components), determine potential safety issues.
- Collect initial data on restrictions for elderly users due to the robot (e.g. furniture that needs to be removed or moved)
- User perception: determine potential psychological issues when using the robot in a home (e.g. privacy-related issues, ethical issues), determine the appropriateness of robot aspects like speed of movement

(too slow or acceptable?), approach behavior (e.g. scary when approaching from the front?), adequacy of volume of operation (e.g. should the robot announce itself when entering a room because it is too quiet?), any other unexpected or unpleasant behavior.

5.2.2 Preparations

5.2.2.1 Apartment selection requirements

Ideally two adjacent apartments (1 for the elder, 1 for the remote operators) will be used. If this is not possible, the remote operators will be situated in a separate room of the same apartment.

The relevant data and limits of the robot will be determined in preparation of the test. The selection and adaptation of the apartment will consider these restrictions. Since the robot has never been tested in a real home, some of the figures are necessarily estimates, or have been determined experimentally, and not by practical application:

- Robot weight (relevant for transport): 180kg; this prevents the use of the robot in any apartments not located at the ground floor, unless there is an elevator
- Minimum width of a narrow passage (door) to pass: 0.83 m (robot footprint is 0.6m x 0.8m)
- Maximum height of a door sill: 0.5 cm to 1 cm in autonomous mode; 2 cm in manual mode
- Maximum climbable slope: 5° to 10° in autonomous mode; max. 30° in manual mode
- Permissible floors: any flat and hard ground works fine (e.g. parquet, linoleum); with some limitations also (a well-fixed) carpet or (not too uneven) tile floor (navigation will not work properly on uneven floors)
- For turning around its own axis, the robot needs a radius of at least 0.8m
- Load on floor: it should be considered that the robot's 180kg are distributed over four wheels which nearly have only a point of contact with the ground

5.2.2.2 Object grasping requirements

For grasping objects, the following limits have been determined beforehand. Again, these are approximations and the precise limits will be determined during the test.

- Types of objects suitable for grasping:
    - Only solid, not deformable objects
    - Nearly box-like or cylinder-like shape
    - With current (not anticipated) object detection algorithm, objects would have to be textured (not uniformly colored); if new algorithms are already in place there is no such restriction
    - Object size should not exceed approx. 20 cm in the smallest dimension; however door handles can still be dealt with as an exceptional case
    - Examples of permissible objects: bottles (0.33-0.5 liters, e.g. Coke), cups and glasses (ideally without handle), books (not too heavy or too large, normal paperback books can be dealt with), grocery packages: boxes, cans (e.g. Pringles)

- Maximum vertical reach: 2m. However, a height between 0.5m and 1.5m is the ideal range to have a reasonable workspace for manipulation
- Maximum workspace for reaching objects in front and back of the robot: in front approximately an area equal to the size of the tray; on the backside the arm can reach objects placed approximately 1m from the centre of the robot, that is, approximately 0.5m from the rear wheels
- Maximum weight of an object: the end effector's tactile sensors are quite sensitive, therefore the maximum weight of objects with small bearing surfaces is likely to be around 2 kg (e.g. a bag that "cuts" into the gripper due to its small bearing surface); objects with a large bearing surface (e.g. a bottle) might weigh up to 5kg; however these numbers have to be verified

5.2.2.3 Technical and other necessary equipment

The following equipment will be required on the test site:

- Wireless access point with 802.11n, and large antennas for good signal (Internet connection not required)
- Laptop computer for SRS communication server
- Wireless emergency push-button for stopping the robot
- Wireless joystick for manual steering of CoB
- Laptop for control and scripting of CoB
- Ramp for pushing Care-o-Bot up the stairs
- CoB battery charger
- Mobile video recording equipment (cameras, tripod, etc.)
- Laptop for session transcript and electronic questionnaire
- Printed documents such as informed consent, interview guide, questionnaires

5.2.3 Method

5.2.3.1 Recruitment

The test will be carried out in a real apartment (in active use) of an elderly person (or an elderly couple). Participants will be two elders and two informal caregivers conforming to the SRS target group specification. The elderly person living in the apartment will be asked to refrain from making any changes to the apartment except for removing unique valuables (removed items will be noted).

5.2.3.2 Evaluation procedure

There will be two test trials with 1 elder and 1 informal caregiver each. Two interviewers will be attending to the two participants. The duration of each trial is scheduled to be about 90 minutes. However, since there is a high degree of technical uncertainties in this pre-test, participants will be cautioned beforehand about possible technical problems and associated wait times. The two sessions will be recorded on video with several cameras in the apartment as well as with the robot's cameras.

The general procedure will be the following: the robot carries out the SRS "fetch and carry" scenario initiated by an informal caregiver, fetching an object, navigating through several rooms, and delivering the object to the elderly person. This scenario will be supplemented with several central elements from the other SRS scenarios (e.g. emergency assistance). Thus, a “scenario mix” of all SRS scenarios with focus on the fetch and carry scenario will be used. The informal caregiver will not be able to see or hear the robot or the local elderly user. This person will be asked to imagine to be sitting at his home, on the other side of the city. Informal caregivers as well as the elderly person will interact simultaneously (to the degree possible). At the time of the test, not all SRS components will be fully functional. The approach is to improvise (e.g. by using a wizard-of-oz approach, showing a video, or the CoB simulation) any functionality not yet implemented.

The following evaluation methods will be used:

- Interactive think-aloud with moderators close to participants
- Structured interviews using prepared questions on robot perception
- Checklist of technical questions on performance of system components
- AttrakDiff questionnaire with direct subsequent interview on reasons for answers

5.3 MANIPULATION TESTS

5.3.1 OBJECTIVES

Manipulation and visualization tests in San Sebastian are planned to address the main peculiarities of the robotic arm with real users (elderly and potential professional remote operators), and to study the effectiveness, usability and user acceptance of its arm/manipulation. The experimental objectives of the tests are:

- Determine the accuracy and efficiency of the robotic arm manipulator in performing tasks (selected from scenarios).
- Evaluate the feedback provided to the professional operators through the visualization system of the UI-PRO device during tasks executions.
- Collect data about users’ subjective perception on the robotic arm and on tasks based on arm manipulation. Because of their peculiar appearance and technological restrictions, current robotic arms like those used within the SRS project could evoke a peculiar response in elderly users. By combining quantitative and qualitative responses to robotic arm movements with feedback-visualization, we expect to gain knowledge on how to improve elderly user-robot interactions based on robotic arm manipulation.
- Collect data about user acceptance and will to use the technology (both elderly and potential professional remote operators).

5.3.2 PREPARATIONS

In these trials, full assistance will be obtained from both the Rehabilitation Service of Birmingham Hospital and Ingema’s laboratories, placed very close to each other (see Figure 3). Depending on the requirements of the technology to be tested, and the mobility difficulties of the frail elderly users, the Rehabilitation facilities for in-site testing may have to be used.

Ingema’s laboratories will be used for technology installation and test making with remote operators, and eventually also with the elderly users.

Regarding ethical issues, a document will be submitted to the attention of the Matia/Ingema/Urkoa Ethics Committee to complete the documentation already approved about the first survey on users' needs. This documentation will be written in parallel to those described in section 5.4, and it will include experimental protocol description, informed consent, authorisation for video recording documents and, documentation about storage and exchange data procedures with other partners.

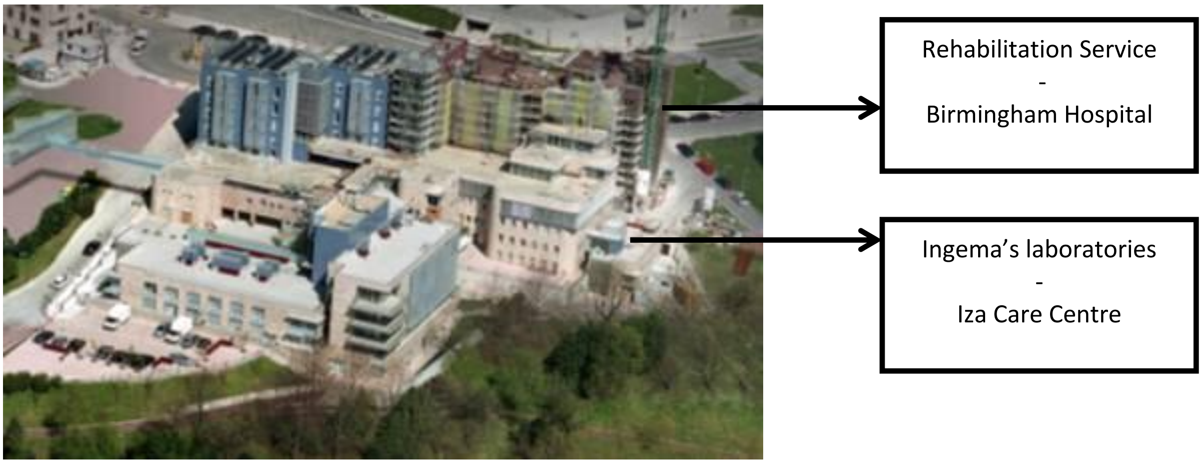


FIGURE 3: REHABILITATION SERVICE OF THE BIRMINGHAM HOSPITAL AND FROM THE INGEMA'S LABORATORIES AT IZA CARE CENTRE

5.3.3 METHOD

5.3.3.1 RECRUITMENT

Tests with advanced prototypes will be based on an experimental protocol involving 20 participants (recruitment criteria are defined into D6.1-1):

- 10 local users (LU): frail elderly people attending to the Rehabilitation Service of Birmingham Hospital (Matia Foundation)[1]. This building is located beside Iza Care Centre, where Ingema's laboratories are located. Elderly people attending the Rehabilitation Services are mainly involved in therapy for functional recovery and, if possible, autonomy. The profile of patients of the Rehabilitation Service match the definition of frailty included in D1.1R.

- 10 remote operators (RO-PRO): potential professional users of the SRS system.

5.3.3.2 EXPERIMENTAL PROTOCOL

These trials focus on grasping and fetching things, as well as on the visualization systems supporting these actions:

- Regarding the requirements, the main aim of manipulation and visualization tests is to answer to research questions concerning the ease of learning for the user (referring to user requirement R1 –"*The system should be usable and acceptable*" - and referring to requirement R09 –"*The system is able to firmly grasp objects*").

[1] http://www.matiaf.net/upload/doc/caste/centros/Hospital%20Ricardo%20Bermingham.pdf

- Regarding the acceptance, the typical dimensions will be observed including: outcome expectations, advantages/disadvantages perceived, attractiveness, comfort perception, eligibility (intention to adopt) and perceived usefulness.

- Regarding the performance evaluation, the success on task execution will be recorded, as well as the time required for its completion, and the errors and/or usability difficulties that may occur within the process.

In order to achieve these results, the final four scenarios selected within WP1 have been analyzed, and two situations/use cases where manipulation is the main part of the action have been selected. In this regard, an integrated assessment of a single but interrelated part of the scenarios can be conducted, and also an evaluation within a scenario-based framework.

The situations selected are:

- Month 25-Manipulation test. This test is going to be conducted with frail elderly people. The robotic arm will be locally controlled by an experienced user by using an already existing local application programming interface (that will imitate the UI_PRO controlling the COB), simulating the way in which a professional user would act and trying to simulate the use of the UI_PRO as much as possible.

  This test is going to be based on Scenario 1: "*The elderly person lies down. He/she wants some milk, but the carton is placed far from him/her on the table, so he/she uses the robotic arm to bring it closer (…).*"

- Month 27-Visualization test. This test is going to be conducted with potential professional operators by using the UI_PRO for robot visualization and grasping. In this test, the robotic arm is not directly needed, so it is going to be simulated on a laptop. The usage of UI_PRO will be investigated in order to determine if it provides the professional users a tool to control the SRS system. One of the main features of UI_PRO is in fact its capability to provide a "visualisation of environment". This visualisation will be perceived not only through the cameras of the COB, but will capture also a visualization of the simulated environment by means of the remaining sensors. For example, using only a webcam is very difficult to grasp an object, so 3D info is needed in addition to provide the user a better knowledge of the environment.

  This test is going to be directly based on Scenario 1: *"The LU asks for a glass of water. So the RO wants to fetch a bottle of water and a glass from the kitchen. He uses a room plan to specify that SRS should go to the kitchen. Having arrived in the kitchen, the RO switches to manual navigation mode to drive SRS to the specific place where the bottle and glass are located (…). Then (…), and then the RO directs SRS back to the bedroom of the LU (…)"*.

  This task will also be repeated several times in order to measure the learnability (time required in T_{n1} compared with time required in T_{n2}).

Although these trials mainly focus on the manipulation abilities of the system, they give also the chance to test the human-interface devices developed so far, concentrating on the specific context of the manipulation tasks.

Each test session will be supervised by at least two SRS Operators, one from Robotnik taking care of technological aspects, and one from Ingema coordinating the experimental procedure and registering qualitative data. In the manipulation test, in which a frail elderly person is involved as LU, a care professional, possibly already known to the elderly person, will also be involved in the process. This professional operator will assist the LU, monitoring his/her conditions and being available for intervention (i.e. if the LU desire to stop the test).

5.3.3.3 Evaluation procedures

The main parameters to be measured in these trials are:

- Secure grasping (Efficiency in ADL task completion)
- Usability + Learnability
- Acceptance (Comfort perception)
- Acceptability (from a Psychological/emotional point of view)
- Safe manipulation (the elderly person is out of the working area of the robotic arm during the test; subjective perception from the elderly person –*to be discussed with the consortium*-).
- Perceived control in daily life

At the end of the scenarios, a specific assessment will be conducted including:

Specific questions mainly related to the dimensions discussed above (including secure grasping, usability, improvement of communication, intention to use the system if applicable). For example (5 scale ranging from totally agree to totally disagree):

- Patients feel comfortable with the robotic arm - Agree/Disagree
- The robotic arm helps the patients to achieve their goals Agree/Disagree
- The robotic arm fits well with patient's routine Agree/Disagree
- The robotic arm delivers the objects to the target destination - Agree/Disagree
- The robotic arm completes the scheduled tasks Agree/Disagree
- I perceive perfect control over the robotic arm - Agree/Disagree
- I perceive the system can help me to control over the difficulties I have in my daily life. Agree/Disagree

Standard questionnaire, such as AttrakDiff, will be administered at the end of the experimental session, mainly to evaluate the acceptance from a psychological/emotional point of view, but also to evaluate comfort perception in the use of the manipulation-visualization system. Complementarily, some observational items adapted from the iWard Project's assessment protocol are also going to be included (i.e. "The user examines the robot with interest" or "The user gets scared by a movement of the robot").

Learnability will be objectively measured, recording the time required to carry out the task in several applications, and the errors. The aim of this part of the assessment is to check whether it is possible to learn how to operate with the robotic arm and its visualization system in an accurate, but also acceptable way. Since learnability refers to the experience of a new user when he/she is starting to use the system (it should be possible to learn quickly and error-free), it can be properly measured, because the characteristics of the tests are based on multiple repetitions of the same task.

Complementarily, specific items about learnability adapted from existing scales will be included in the experimental protocol (i.e. "It is easy to forget how to do things" or "It is easy to make the software do exactly what you want", but also "Working with this software is mentally stimulating").

Usability questions will be performed based on the UI_PRO development stage. At this point, it will be measured how easy/difficult is the interaction with the interface.

All the measures are going to be completed with information collected by Ingema's staff through different methods: a) participant observation in Test 1 - Manipulation test with elderly users; b) think-aloud recording in Test 2 - Visualization test with potential caregivers. With these methods, qualitative information that complements and confirms the information collected directly from the users will be obtained.

5.4 ADVANCED SRS PROTOTYPE TESTS

5.4.1 OBJECTIVES

The final and most comprehensive set of tests will be performed in the Milan apartment, focusing on the evaluation of the whole prototype in an advanced stage of development. The evaluation will consider the major user requirements determined at the beginning of the project, expressed through real-life scenarios which will be experienced by real users. In these last tests, all the previously mentioned research goals will be assessed. The evaluation will run in conjunction with the final usability tests as part of task 2.6. Therefore, the work will be planned by the partners of both WP6 and WP2. **The final results will be reported separately in two deliverables; D2.2b specifically dedicated to the usability results, and D6.2 dedicated to the overall results and overall conclusions about user evaluation (month 36).**

5.4.2 PREPARATIONS

The acceptance tests of the advanced SRS prototype will be performed into an apartment located inside the hospital Santa Maria Nascente of Don Carlo Gnocchi Foundation. The selected test site, called "SMART HOME" (described in detail in D6.1-1) is part of the DAT service *(Ita: Domotica, Ausili e Terapia Ocupazionale – Eng: Occupational Therapy, Assistive Technology, Smart Home)*, a specialized service of Don Carlo Gnocchi Foundation which provides information, guidance, consultations and individual assessment in the field of assistive equipment for people with disabilities.

A formal collaboration with the DAT service has been agreed. In the context of this collaboration DAT will provide:

- Recruitment of the evaluation tests participants, chosen among the caregivers and the elderly patients (and their relatives) of the service;
- Involvement of health professional staff for the tests execution, in order to provide expert advice and support for the privacy and safety management of the elderly (clinical responsibility);
- Availability of the test site for the evaluation period.

Other contacts are going to be made with tele-assistance centers for elderly, in order to recruit also real 24 hour operators for the experiments. If no tele-assistance operator can be recruited, the plan is to involve DAT health professionals for this role, as their attitude and knowhow can be comparable to the one required for tele-assistance operators specialized on elderly support.

As soon as the draft of the experimental protocol will be approved by the SRS consortium, a document about ethical and safety issues will be prepared and submitted the Ethics Committee of Don Carlo Gnocchi Foundation, to complete the documentation already approved about the first survey on user needs. The document will include: experimental protocol description, informed consent, the authorization for video recording documents, the documentation about sensitive data storage and exchange procedures.

The apartment used for the evaluation tests will be prepared and "configured" taking into consideration the results of the ethnographic research and the final scenarios. The house rooms mentioned in the four planned scenarios are: kitchen, bedroom, living room, way to toilet, entrance. The current idea is to exploit the living room of the Milan apartment as the elderly location for all of the scenarios, mainly because in this place five recording cameras are installed that allow for easy remote monitoring and recording of each test (see section **Error! Reference source not found.**). The living room of the apartment has already a kitchen in one corner, and can be configured with a bed in another corner so as to provide all the needed places for the scenarios. The only place not covered by the recording cameras is the main entrance of the apartment. The office and the tele-operator centre places will be hosted in other rooms of the apartment, indicated in figure 2 as “Remote operators rooms”.

Other technical and logistic problems are now under examination, for the moment the following features where checked:

- internet access available inside the house
- door passages accessible to the SRS platform (minimum door passage=82 cm)
- Easy access to the house allowing the first arrival of the robot
- Locked room to keep the robot secure when not in use.

The next version of this document will describe in more detail the initial setup in terms of furniture, ornaments, narrow passages, space division, etc. The setup of the apartment will be optimized to minimize reconfiguration between subsequent tests. Moreover, as soon as the Stuttgart tests in a real home will be performed, the results will be taken into account to finalize the apartment predisposition.

5.4.3 METHOD

5.4.3.1 RECRUITMENT

Tests with the advanced prototype will be based on a protocol involving 23 participants (recruitment criteria are defined in D6.1-1):

- 10 local users (LU): the frail elderly people
- 10 private remote operators (RO-PRI): relatives of the elderly people
- 3 professional remote operators (RO-PRO): potential tele-assistance operators of a 24 hour call centre service for elderly

5.4.3.2 EXPERIMENTAL PROTOCOL

The evaluation tests will be based on the four scenarios which have already been selected and validated by users in the course of the project:

- Basic fetch and carry and video communication
- Emergency assistance
- Teaching SRS complex action sequences (also called “preparing food”)
- Fetch and carry of difficult objects

Participants will perform the defined scenarios using the three human interface devices developed by the project:

- UI-LOC, the device dedicated to local users (elderly),
- UI-PRI, the device dedicated to relatives or private caregivers or also to those elderly particularly skilled with technology,
- UI-PRO, the device dedicated to 24 hour professional remote operators.

Using these devices both elderly and remote operators will cooperate in playing the four selected demonstration scenarios. Through the scenarios each participant will have the possibility to control the robot, see it in action, and appreciate a concrete result in order to give a feedback.

As indicated in the D1.3a (2.4.1 Phases of SRS System Usage) the lifecycle of an SRS system is composed of three main phases: Pre-Deployment (or Production), Deployment and Post-Deployment (or Operational). At the time of this evaluation test the Pre-Deployment phase will be considered finished, with the robot knowledge base already loaded with action sequences, household objects and 3D models. As soon as the robot will arrive at the Milan site the standard Deployment phase will be executed, in order to perform personalized setup actions such as loading a 2D map of the apartment, building a 3D map of the environment, recognizing and learning the position of useful objects, etc. The focus of the evaluation tests will be on the Post-Deployment / Operational phase, i.e. on the everyday use of the system where new objects and action sequences are still learned.

To guide each participant during the test execution, a detailed screenplay will be produced for each scenario. A first draft version of it is available in the Appendix 1. Details of each scenario are already available in the internal “SRS scenario revision process” living document. This document is continuously updated during the course of the project based on real technical development. The final screenplay will be produced as soon as the prototype is in a sufficiently advanced development stage so that the scenarios can be defined as “final”.

To create a detailed plan and assign the test participants to the scenarios, at least three things are to be considered: how many people will participate to each scenario, how much time the average elderly person will stand up while playing the scenarios and performing all the related activities, and how much time will be needed to run each scenario.

At the moment it can be estimated that each experimental session will involve up to three participants: an elderly (LU), and one or two ROs (just the relative RO for scenarios 1-3 and both the relative and the 24 hour operator for scenarios 2-4).

The duration of each test should not be a problem for ROs, but must be carefully considered for the involvement of elderly people, who should not get tired too much as a result of this activity. Based on previous experiences, and mediating between the need to conduct quite long and complex tests and the need to involve frail participants, it is currently estimated that each experimental session should not last more than two hour and a half. This time will be considered as an upper limit not to be exceeded and thus a constraint for scheduling the testing sessions.

In order to produce a first estimation of the time needed to run a single test session, a “wizard of oz” first experimental session made by actors will be performed before the prototype arrival. This experiment will be performed early, and will indicate if at least one test session can be run within the two hour and a half maximum limit; if not, a simplification of the session scripts will be done to reduce the running time. To obtain a more precise estimation, a set of trials with actors will be performed as soon as the prototype arrives in Milan and is configured. These “wizard of oz” sessions will also provide a useful learning phase dedicated to all the operators involved in managing the experiment.

Ideally, all local users should perform each of the four scenarios; but, if in the wizard of oz trial it will emerge that the time needed to complete a whole four-tests session is too demanding, the initial plan will be revised to reduce the number of tests assigned to each participant.

Each test session will be supervised by a DAT health professional and a number of SRS Operators (SRS OP):

- A DAT health professional, possibly already known to the elderly person, will assist the LU, monitor his/her conditions and will always be available for intervention or should the LU desire to stop the test
- An SRS OPERATOR will assist the RO-PRI, both for the use of the human interface device and to coordinate the test

- An SRS OPERATOR will assist the RO-PRO, if any is involved in the test, both for the use of the human interface device and to coordinate the test
- An SRS OPERATOR will be responsible for data collection, in particular video recording
- An SRS OPERATOR will coordinate the test and will be responsible for marking timestamps at each test step

Each test session ideally will be composed of the following steps:

1. The coordinator marks the session starting time and declares the current test session open
2. A DAT health professional will explain to the LU the general aim of the study and the particular aim of the current test. The SRS OPERSATORS in the mean time will also explain the same to the ROs
3. The LU and the ROs will read and sign the informed consent and the authorization for video recording documents
4. A DAT health professional will explain to the LU how to run the test and how to use the interaction device. The SRS OPERSATORS in the mean time will also explain the same to the ROs
5. The SRS OP acting as coordinator, after verifying that every participant is ready and that all the subsystems work properly, will order the startup of a scenario and mark the scenario starting time
6. The LU and the other ROs participant(s), assisted by their respective operators, will play one of the four scenarios, while
7. The coordinator supervises the smooth running of the test, marks a timestamp for each step and, if needed, suggests to any participant who is experiencing difficulties how to proceed
8. At the end of the test the coordinator marks the ending time, and instructs each operator to start the data collection and rearranging phases
9. Each assistant will ask to their respective participant some evaluation questions and will record the answers, while
10. The coordinator eventually re-arranges the set up before the execution of the following scenario and
11. The data collection operator archives the recorded videos and prepares for a new recording session
12. When the data collection/rearranging phase is terminated, the coordinator verifies that enough time is available for testing another scenario and orders to prepare for the startup of a new scenario
13. The session continues at step 4
14. If no more time is available or if all of the four scenarios have been tested, the coordinator instructs each operator to start the final data collection phase
15. Each assistant will ask to their respective participant to fill in the final questionnaire
16. At the end the coordinator declares the testing session closed and marks the ending time

In the following figure 4 a possible position in the apartment is indicated for each actor during a testing session.

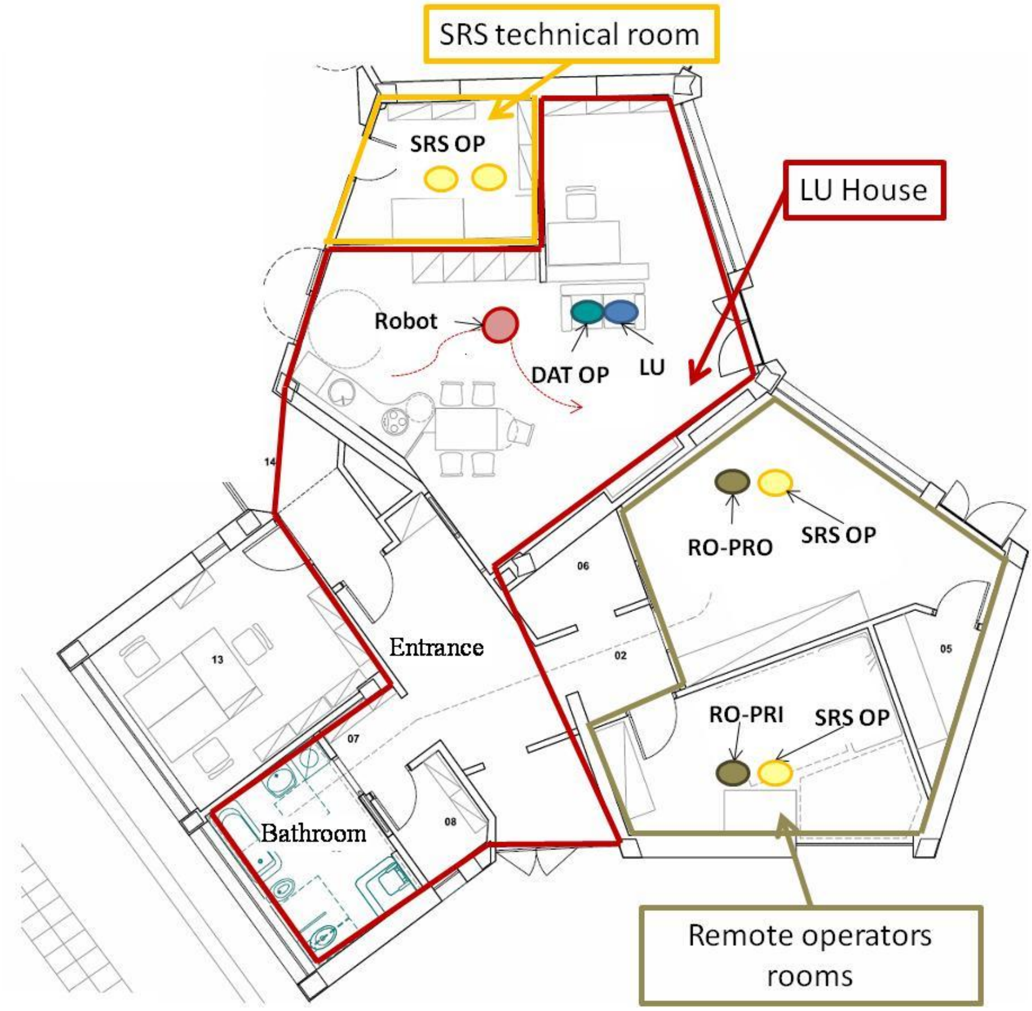


FIGURE 4: PLANNED LOCATION OF PARTICIPANTS DURING THE TESTING SESSIONS

5.4.3.3 Evaluation procedure

Each experimental session will be video recorded, using the 5 cameras (AXIS 212 PTZ) integrated into the walls of the kitchen corner and the remaining of the open space room, (configurable as living room, bedroom, office, etc., depending on scenarios). This camera disposition (see Figure 5) not interfering with the house appearance, allows to record tests with the minimum discomfort for the local users, thus fostering repeatability of tests, and both measurability and comparability of results.

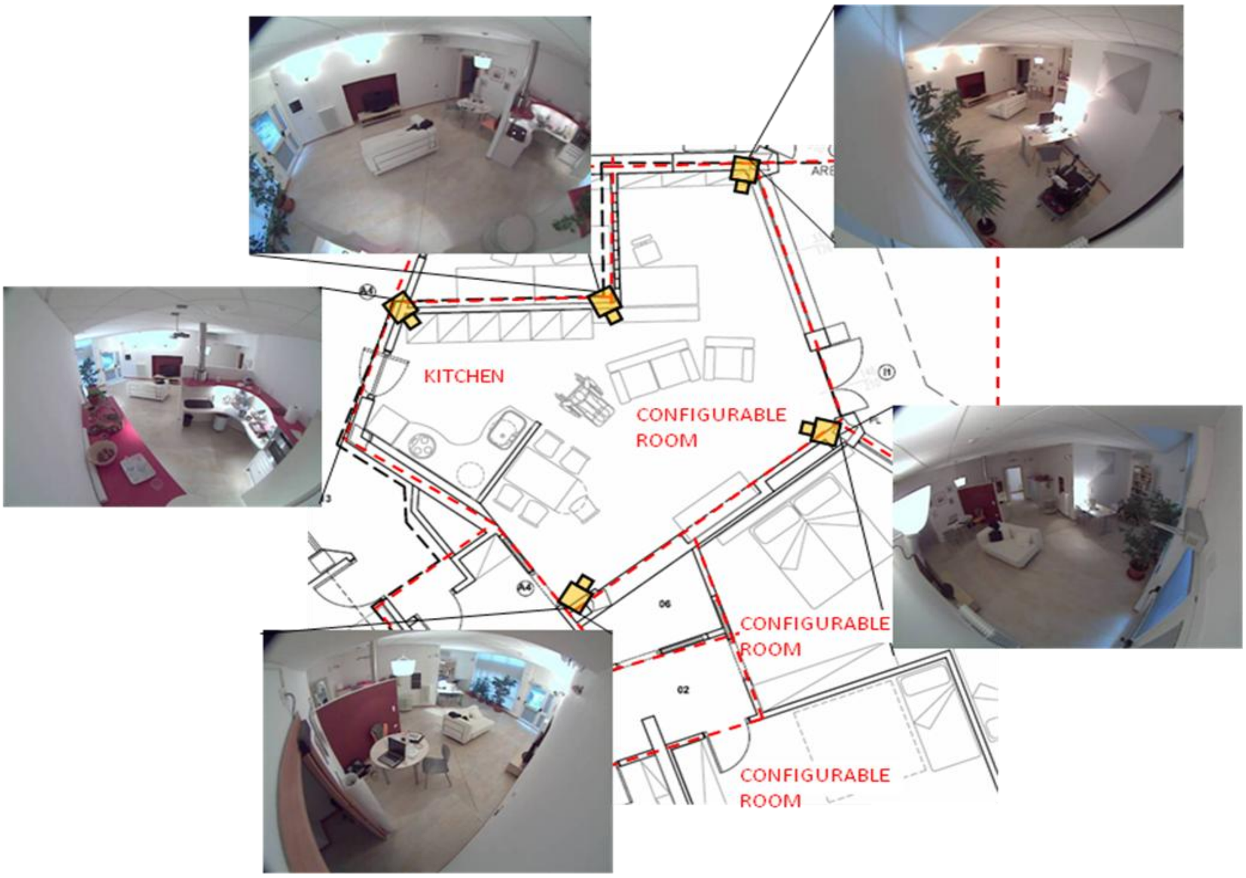


FIGURE 5: CAMERA DISPOSITION INTO MILAN TEST SITE

During the execution of the scenarios, the participants will be asked to “think aloud” so as two SRS operators, watching the tests from the technical room outside the test site, could take note of every particular behavior, reaction, comment, first impression while video and audio recording.

The SRS operators will also have to fulfill a data collection form (an example is shown in Table 3) to consider objective and measurable parameters about the technical **effectiveness** of each scenario (tasks completion, time needed, eventual problems occurred in tasks completion).

TABLE 3 – EXAMPLE DATA COLLECTION FORM FOR SRS OPERATORS- MEASURING “ EFFECTIVENESS” RESEARCH GOAL

Scenario 1	Time needed	solved successfully	Comments
The video communication established	*1 minute*	*Yes*	
the robot arrived into the kitchen (Semi-autonomous navigation by room plan)	....	*Yes*	
The robot recognized the already known object and correctly put in on the tray	....	*Yes*	
The robot learned the unknown object and correctly put in on the tray	......	*Yes*	
The robot correctly brought the objects to the elderly place	....	*Not completely*	*i.e. The robot passing too close to a table along the way has dropped an ornament*
Scenario 2	**Time needed**	**solved successfully**	**Comments**
The emergency device correctly made the emergency call specifying also the room in which the elderly is located			
The robot Localized the elderly person through SRS autonomous navigation (after elderly call for emergency in which specifies his/her position)			
The robot through manual navigation and camera perspective allows to evaluate the health state of the elderly			
The video communication between three people is established (elderly, relative			

and 24hour caregiver simultaneously)			
The robot goes to the entrance door			
The robot open the doors to rescuers			
Scenario 3	**Time needed**	**solved successfully**	**Comments**
The Video communication was established			
The robot went to the kitchen			
The SRS executed the sequence of fetch-and-carry tasks through remote control (open microwave open fridge, bring food, put it into the microwave….)			
The robot is taught about the actions sequence through the User editing of action sequences			
The robot demonstrated the learned behavior, performing the complete sequence autonomously			
Scenario 4	**Time needed**	**solved successfully**	**Comments**
SRS suggested the correct chain of call requests (first to the son, then , cause the son is not available to the 24 hour service)			
The robot correctly fetched the book (through the professional manual mode with the force-feedback interaction device to grasp the book, moving aside the other books).			
The robot was taught about the taught about the 3d model library and drawing a line around the object in video image approaches by standard semi-autonomous grasping mode to return the book on the shelf			
The robot demonstrated to have learned the "rotate-on-gripper" approach and so return the book on its place (controlled by LU using UI-PRI)			

Objective and subjective indicators for measuring SRS's success with users

Taking into account SRS's goals,

- elderly should be more autonomous (A)
- elderly should be safer (S)
- family should be able to provide care (C) where it previously wasn't possible (e.g. due to geographical distance)
- family should spend less time (T) and effort (E) for care

and considering that:

- elderly autonomy (A) is defined as: being able to carry out a necessary action that they could previously only carry out with either help by someone or under substantial effort/difficulty (e.g. difficult to move, etc.) without needing help by their caregivers or with less help; perceiving a higher sense of self-control and not conversely perceiving a sense of being controlled.
- elderly safety (S) is defined as the ability to receive appropriate help in an emergency situation
- family care enabling (C) is defined as making care possible in a situation where it previously wasn't (this applies mainly to potential caregivers living far away or when someone is busy and can't come);
- family time (T) saving is defined as spending less time and being needed less often for the care
- family effort (E) saving is defined as a subjective perception of the amount of effort needed to resolve a care situation

to operationalize the above expressed concepts, some objective and subjective indicators are now under consortium discussion. The proposal under examination is the following:

Measuring the increase in autonomy:

Subjective autonomy indicator (A1): elderly person is asked "How would you rate your autonomy / dependence on your typical caregiver in this situation with the robot compared to your previous situation without the robot?"

scale: -2 much worse with robot, -1 worse with robot, 0 same, +1 better with robot, +2 much better with robot.

Subjective Perceived sense of control indicator (A2): elderly person is asked “How would you rate your control in this situation with the robot compared to your previous situation without the robot?” / “How would you rate the impact of the robot in your sense of control?”

scale:

-2 much lower with robot, -1 lower with robot, 0 same, +1 higher with robot, +2 much higher with robot.

Objective time saving indicator (T): how long it takes to the elderly’s caregiver on average to resolve actually the situations compared with the use of SRS. For example, if currently the time needed to call the son, waiting him coming at home and performing the requested tasks is 40 min, (this includes driving time); this could be the reference (different for each participant like in real life).

It then can be measured how often and how long a caregiver had to be consulted using SRS instead (e.g. 2 times, 10 min each). Obtained values could be put in relation to obtain a metric. In the example above, we would have had 40 min in current situation and 20 min with robot.

So that would be an improvement in time saving of (50%).

Measuring increase of safety:

Elderly subjective safety indicator (S1): elderly are asked "How would you rate your sense of safety having the robot compared to your previous situation without the robot?"

Scale: -2 much worse with robot, -1 worse with robot, 0 same, +1 better with robot, +2 much better with robot.

Professional users’ subjective safety indicator (S2): “How would you rate your current abilities of remotely ensuring safety for elderly customers versus the new situation with the robot?”

Scale: -2 much worse with robot, -1 worse with robot, 0 same, +1 better with robot, +2 much better with robot.

Objective safety indicator (S3)

how long it takes to the elderly’s caregiver on average to resolve actually the situations compared with the use of SRS. Same methodology of time saving referred to ADL activities

Measuring SRS’s ability to enable care as well as time and effort saving during care

Objective care enabling indicator (C): Care is enabled if it was previously not possible to be carried out (for any reason: distance, work commitments…). We ask informal caregivers: "Did you previously provide this kind of help in the situation just experienced by going to your assisted elderly person?" (YES/NO).

If answer is yes, the “objective time saving indicator (T)” (see below) will be used.

If answer is no, we ask (C): "How do you rate the suitability of the solution with the robot in terms of improving your current situation?"

Scale:-2 much worse than current, -1 worse, 0 same, +1 better, +2 much better. And Why?

Subjective effort indicator (E): "How do you perceive the resolution of the situation with the robot compared to your current situation without the robot?"

Scale: -2 requires much more effort, -1 requires more effort, 0 same, +1 less, +2 require much less effort

Success criteria for reaching the goals

1. Autonomy improvement of elderly with SRS of +1.0 scores on average (subjective) over their current situation
2. Time saving improvement of 25%over the current situation
3. Safety improvement of elderly with SRS of 25% (objective) and +1.0 scores on average (subjective) over the current situation
4. Care enabling improvement for caregivers of +1.0 scores on average over the current situation
5. Effort saving improvement for caregivers of +1.0 over the current situation

In addition to the indicators presented in the previous section, specific questions developed ad hoc for the purpose, adapted from based on reviewed methods of chapter 4, mainly adapted from Ami Questionnaire (Ben AllouchB.2009) would also provide feedback about **usefulness** of the entire system, **outcome expectations**, **perceived benefits** and **intention to adopt**. A first draft of the questions is shown in Table 6.

The goal could be to achieve at least “+1.0” on average for each success indicator group.

e.g. Perceived benefits= mean (Enjoinment, Ease of use, personalization, Usefullness) ≥ +1.0

TABLE 6 – EXAMPLE OF QUANTITATIVE INDICATORS INVESTIGATING SRS MAIN USER REQUIREMENT FULFILLING ACCEPTABILITY USEFULLNESS, OUTCOME EXPECTATIONS, PERCEIVED BENEFITS AND INTENTION TO ADOPT (COVERING TOPICS IN TABLES 2, … AND, PARTIALLY, IN TABLES 3 AND 4).

Quantitative indicators	Measure Not at all O O O O O Very Much -2 -1 0 +1 +2	Why question
Success Indicators- 1: Perceived benefits		
Enjoyment	Not at all O O O O O Very Much	*e.g. The robot is funny…*

Ease of use	Not at all O O O O O Very Much	*It is difficult to learn how to control it..*
Personalizations (adaptative to changing needs of users)	Not at all O O O O O Very Much	…………….
Usefulness	Not at all O O O O O Very Much	………………….
Success Indicators- 2: Activities outcomes		
To make ADL activities executions **possible/easier** for you	Not at all O O O O O Very Much	
To have more **control** over your ADL activities		
To make ADL activities executions **safer** for you	Not at all O O O O O Very Much	
To be able to carry on tasks difficult for me **without waiting for someone** coming to my home	Not at all O O O O O Very Much	
Because it offers you **more autonomy** (for elderly)	Not at all O O O O O Very Much	
Because it offers you **more freedom** (for RO-PRI)	Not at all O O O O O Very Much	
Because it offers you a a system **improving health care assistance of elderly compared with the ones already existing** (for RO –PRI and PRO)	Not at all O O O O O Very Much	
Because it offers you a system which make you feeling **less lonely** entertain you by video communication during activities executions	Not at all O O O O O Very Much	
Success Indicators- 3: Monetary outcomes		
To be able to do different things at once (For RO-PRI)	Not at all O O O O O Very Much	
To facilitate the monitoring of the health status of my assisted	Not at all O O O O O Very Much	
Not to have to do everything yourself	Not at all O O O O O Very Much	
To make your everyday life easier	Not at all O O O O O Very Much	
Because it is convenient that you do not have to carry out certain tasks yourself	Not at all O O O O O Very Much	
To save time	Not at all O O O O O Very Much	
Success Indicators- 4: Social outcomes		
To strengthen my relationship with family and friends	Not at all O O O O O Very Much	
To be able to communicate with family and friends	Not at all O O O O O Very Much	
To have something to talk about with others	Not at all O O O O O Very Much	
Success Indicators- 5: Self-reactive outcomes		
To feel less lonely	Not at all O O O O O Very Much	
to fell less scared of being at home alone	Not at all O O O O O Very Much	
Success Indicators- 6: Intention to adopt		
I intend to use the SRS system if it is available	Not at all O O O O O Very Much	
I plan to buy the SRS system as soon as it is available	Not at all O O O O O Very Much	
I will use the SRS system if it is available	Not at all O O O O O Very Much	

At the end the standard Attrakdiff questionnaire (see Appendix 2) could be filled in by each participant, mainly to complete the evaluation of the **attitude, usability** and **acceptability** of the entire system.

Usability success criterion

Attrakdiff results should belong to the area called "desired"

Safety/ Ethics qualitative evaluation

At the end of the entire set of experimental sessions, a focus group with the health professionals who conducted and assisted to the experimental session will provide also an **expert opinion mainly focusing on safety, ethical and privacy issues**.

6 FIRST NOTES ABOUT SRS COST EFFECTIVENESS ASSESSMENT & SOCIO-ECONOMIC IMPLICATIONS

Answering to the question: *"Is the outcome worth the investment? ",* means to take into consideration not only the financial aspect of a product developed for people assistance.

The **purchase price is not** a meaningful indicator of the social cost. The social cost depends to a large extent on **how to use** the aid and on the **environment**, as well as the role of the aid within the **whole assistive program.** The triad **person / activity / environment**, as well as determining the criteria for choosing a particular assistive solution, influences the overall social cost

The most appropriate indicator of the cost of an intervention should take into account:

- The costs are distributed among several actors: social cost could be seen as the sum of costs incurred by all players
- Some costs have to be considered fixed (independent from the specific choosen product), some other costs insteand are marginal (from the specific chosen system for care)
- The cost of the intervention has to be compared with the cost of "non-intervention" : what matters is the additional cost

6.1 IMPLEMENTATION OF THE SIVA COST ASSESSMENT INSTRUMENT (SCAI) IN A MULTI-ROLE ROBOTIC-SYSTEM RESEARCH PROJECT.

SCAI is a specific instrument for social cost analysis designed to help clinicians estimate the economic aspects of providing individual users with assistive technology solutions. It is an informative instrument that, used during clinical assessment, makes clinicians and users aware of the economic consequences of their decisions (Andrich R., 2007). It is designed to help to **estimate** the cost of choosing a **solution for autonomy** (aid, personal care, environmental adaptations ...) and to economically compare the various possible solutions.

In most cases, Service Delivery Systems consider just the purchase price of the assistive device, which would seem the most logical indicator to describe whether an AT solution is cheap or expensive. This is not so, since this view often leads to severe distortion of the cost-outcome analysis. SCAI estimates the additional social cost involved by the chosen solution over a certain period of time. This basically includes four cost categories:

- Investment: cost of purchasing the equipment and having it installed, personalised and ready-to-use. This also includes the provision of adequate training for the client.
- Maintenance: running costs of technical maintenance; depending on the case, this may include repairs, insurance, power supply, etc.
- Services: other services that may be needed in relation to the chosen AT solution (e.g., a bulky powered wheelchair might require specialized minibus transport instead of a cheaper ordinary bus).
- Assistance: the amount of human assistance needed in relation to the device (e.g., a pushchair works only if a personal assistant is there to push), independently of whether that manpower is paid for, or offered for free by relatives or friends or volunteers.

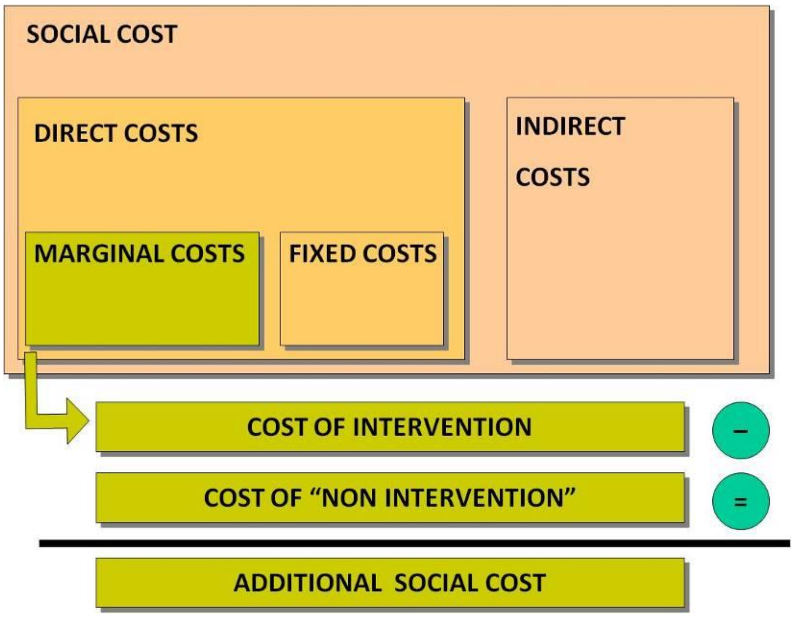


FIGURE 6 SCAI SCHEME: TYPES OF COSTS CONSIDERED IN THE ANALYSIS.

In order to better explain the way in which SCAI instruments analyzes social costs, some definitions are needed:

- **Social Cost** The set of all resources used in a certain period of time by all actors involved (eg family, National health system facilities, City, volunteering, etc. ..)
- **Direct Social Cost** The total costs that can be recognized as directly related to the choice of that particular solution
- **Additional Social Cost** Difference between the social cost of intervention, and that in the absence of intervention. This difference can be>0 (investment), zero (moving resources) or <0 (savings)
- **Expenditure** The actual outlay of money by the different ... "co-financing" actors
- **Time horizon** For how many years the costs need to be accounted for:
  - Clinical duration: Within the time horizon, how many years that type of help will be useful to the user
  - Technical duration: lifespan of the system

But what does "non-intervention" with respect to SRS adoption mean? Different options are possible; first of all, simply no help at home, which implies autonomy and ethics consequences; second, a human caregiver at home, which however implies lost of privacy and autonomy all day long

Deliverable 6.2 (expected in month 36), will report the results about social cost analysis at this starting stage. Appendix 3 shows an exemplification of the outcome that such kind of analysis could generate, comparing the social cost of three different kinds of "SRS intervention", compared with the "non SRS intervention" (i.e. the human 24 hour caregiver).

7 REFERENCES

Allouch, S.B., van Dijk, J.A.G.M., & Peters, O. (2009). The Acceptance of Domestic Ambient Intelligence Appliances by Prospective Users. Lecture Notes in Computer Science, 5538/2009, 77-94, DOI: 10.1007/978-3-642-01516-8_7

Altmann J. (1974) Observational study of behavior: sampling methods. Behaviour. 1974;49(3):227-67

Andrich R, Caracciolo A (2007). Analysing the cost of individual assistive technology programmes. Disability and Rehabilitation Assistive Technology 2 (4), 207-234.

Brooke, J. (1996) SUS: a "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A.Weerdmeester & A. L. McClelland (eds.) Usability Evaluation in Industry. London: Taylor and Francis.

Caulfield, B. (2010). Technology Enabled Assessment and Intervention Protocols in a Community and Home Care Setting for Independent Living. IEEE32nd Annual International Conference of the IEEE EMBS, Buenos Aires, Argentina, August 31 - September 4, 2010.

Chipperfield, J.G. & Greenslade, L.G. (1999). Perceived Control as a Buffer in the Use of Health Care Services. Journals of Gerontoly B, Psychological Sciences and Social Sciences 54B, 3, P146-P154. doi: 10.1093/geronb/54B.3.P146

Cincotti F, Mattia D, Aloise F, Bufalari S, Schalk G, Oriolo G, Cherubini A, Marciani MG, Babiloni F. [2008]. Non-invasive brain-computer interface system: towards its application as assistive technology. Brain Research Bulletin, 15,75(6), 796-803.

Cherubini, A., Oriolo, G., Macrı, F., Aloise, F., Babiloni, F., Cincotti, F., & Mattia, D. [2007]. Development of a multimode navigation system for an assistive robotics Project. Proceedings of the IEEE International Conference on Robotics and Automation, Roma, Italy, 10-14 April 2007.

Crandall, J.W., Goodrich, M.A., Olsen, D.R. & Nielsen, C.V. (2005). Validating Human–Robot Interaction Schemes in Multitasking Environments. IEEE Transactions on systems, man, and cybernetics – Part A: Systems and humans, 4, 438-449.

Dillon A. (2001) User acceptance of information technology. I: Karwowski W. (ed) Encyclopedia of Human Factors and Ergonomics, Taylor and Francis, London

Demers, L., Monette, M., Descent, M. Jutai, J., & Wolfson, C.2002) The Psychosocial Impact of Assistive Devices Scale (PIADS): Translation and preliminary psychometric evaluation of a Canadian–French version. Quality of Life Research, 11, 583–592.

Giesbrecht, I.M., Ripat, J.D., Quanby, A.O., & Cooper, J.E. (2009). Participation in community-based activities of daily living: Comparison of a pushrim-activated, power-assisted wheelchair and a power wheelchair. Disability and Rehabilitation: Assistive Technology, 4(3), 198–207

Gonzalez, M.F., Facal, D., Navarro, A.B., Geven, A., Tscheligi, M., Urdaneta, E., & Yanguas, J. (2011). Analysis of Older Users' Perceived Requests and Opportunities with Technologies: A Scenario-Based Assessment. International Journal of Ambient Computing and Intelligence, 3(1), 42-52

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität [AttracDiff: A questionnaire to measure perceived hedonic and pragmatic quality]. In J. Ziegler & G. Szwillus (Eds.), Mensch&Computer 2003. Interaktion in Bewegung (pp. 187–196). Stuttgart, Leipzig: B. G. Teubner.

Infurna, F.J., Gerstorf, D., Zarit, S.H. (2011). Examining dynamic links between perceived control and health: longitudinal evidence for differential effects in midlife and old age. Developmental Psychology, 47, 1, 9-18.

Lachman, M.E. (2006). Perceived Control Over Aging-Related Declines. Current Directions in Psychological Science, 15, 6, 282-286. Doi: 10.1111/j.1467-8721.2006.00453.x

Lange J. K., (2000) Focus Groups. A Practical Guide for Applied Research (3rd Edition). Thousand Oaks.

Levy, B.R., Slade, M.D. & Kasl, S.V. (2002). Longitudinal Benefit of Positive Self-Perceptions of Aging on Functional Health. Journals of Gerontoly B, Psychological Sciences and Social Sciences 57, 5, P409-P417. Doi: 10.1093/geronb/57.5.P409

Lewis, C. H. (1982). Using the "Thinking Aloud" Method In Cognitive Interface Design. Technical Report IBM RC-9265.

Matarić, M.J., Eriksson, J., Feil-Seifer, D.J., & Winstein, C.J. (2007). Socially assistive robotics for post-stroke rehabilitation. Journal of NeuroEngineering and Rehabilitation, 4, 5. Doi:10.1186/1743-0003-4-5

Menec, V.H. & Chipperfield J.G. (1997). The interactive effect of perceived control and functional status on health and mortality among young-old and old-old adults. Journals of Gerontoly B, Psychological Sciences and Social Sciences 5*2, 3, P118-26.*

Oztemel, E. et al. (2008). Evaluation criteria, proof of concept. IWARD Intelligent Robot Swarm for Attendance, Recognition, Cleaning and Delivery, D7.1.1.

Parsons, B., White, E., Warner, P., & Gill, R. (2006). Validation methods for an accessible user interface for a rehabilitation robot. Univ Access Inf Soc, 5, 306–324. DOI 10.1007/s10209-006-0051-y

Spiekermann, S. (2005). Perceived control: Scales for privacy in ubiquitous computing. In 10th International Conference on User Modelling, July 2005. Doi=10.1.1.67.4989

Tapus, A., Tapus, C., Mataric, M.J. (2008). User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. Intel Serv Robotics, 1, 169–183. DOI 10.1007/s11370-008-0017-4

Watson D., Clark L., Tellegen A. (1998) Development and validation of brief measures of positive and negative affect: The panas scales. J. Pers. Soc. Psychol. 54(6):1063-1070

Wiedmann, T., Wilting H., Lutterc S., Palmd V., Giljumc S., Wadeskog A., Nijdamb D. (2009) Development of a methodology for the assessment of global environmental impacts of traded goods and services. SKEP ERA-NET Project EIPOT (www.eipot.eu)

8 APPENDIX 1- SCENARIOS SCREENPLAY

8.1 FETCH AND CARRY + VIDEO CALL (BASE SCENARIOS)

Aim of this scenario: to demonstrate basic fetch and carry functionality and video communication ….
Actors: LU and RO-PRI
Places:

- Into the house: kitchen, bedroom,
- Outside the house: office


Scenario screenplay:

LU: Elisabeth Baker (84) lies at home in bed due to a cold.

RO-PRI: To check if everything is alright, her son Martin initiates a request for a remote session from his workplace.

LU: Elisabeth accepts the request on her portable communication device and a video communication is established.

RO-PRI: Martin asks if he could do anything for his mother.

LU: Elisabeth answers that she feels a bit thirsty.

RO-PRI: Martin therefore wants to fetch a bottle of water and a glass from the kitchen. He uses a room plan to specify that SRS should go to the kitchen. Having arrived in the kitchen, Martin switches to manual navigation mode to drive SRS to the specific place where the bottle and glass are located. SRS indicates by a rectangle that it recognizes the bottle. This bottle has been previously taught to SRS. However, the glass is not indicated as recognized. It is a new glass that SRS has not been taught before. Martin clicks on the bottle and SRS puts it on the tray. Because the glass is not recognized, Martin switches to user-assisted grasping mode. From a library of 3D object models, Martin selects from the category “glasses” a cylinder-shaped glass similar to the one to be grasped. He adjusts its shape (height, width, position) so that it matches what he sees on the video picture. He then clicks “GO” and SRS grasps the object and puts it next to the glass on its tray. Having finished the grasping, SRS asks Martin if this object should be saved for future grasping. SRS suggests to save it in the category “glasses”. Martin confirms and assigns a name: “long IKEA glass”. Martin directs SRS back to the bedroom of his mother. While SRS drives back, Elisabeth asks Martin what he is doing and why it takes so long. Martin speaks with his mother telling her that SRS will be there soon.

LU and RO-PRI: Martin and his mother agree to end the conversation and speak again tomorrow. After ending the call, SRS autonomously drives back to its charging place.

RO-PRI: Martin calls again and again wants to get his mother a glass and bottle. Today, Martin can just click on the glass to grasp it. However, today grasping the bottle fails even though this is an object taught to SRS. There are many other objects in the scene. Martin uses the “reduce search space” approach and SRS successfully grasps the bottle.

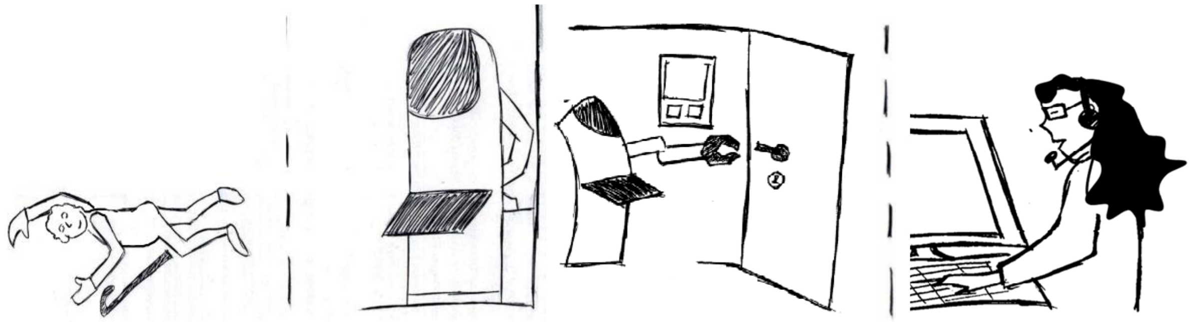
8.2 EMERGENCY ASSISTANCE

Aim of this scenario: to demonstrate emergency intervention functionality and multi video communication

Actors: LU , RO-PRI, RO-PRO

Places:

- Into the house: living room, way to toilet, entrance
- Outside the house: office, teleoperator centre



Scenario screenplay

LU: Elisabeth Baker (84) watches TV. In the commercial break, she wants to go to the bathroom but falls on her way (*sit on the ground in the scenario execution*), unable to get up again. With a device she always carries attached to her belt, Elisabeth presses a button "emergency". Right away, a call is placed to her son and daughter as well as to the 24-hour teleassistance centre.
The device asks Elisabeth for her current position and she selects the room from a list. SRS starts moving from its charging station to the room where Elisabeth fell.

RO-PRO and LU: The 24-hour centre first accepts the call. Through SRS's camera, Claudia, the teleoperator, can see Elisabeth on the floor and asks what happened. She uses manual navigation to further drive the robot to the place where Elisabeth lies and to point the robot's camera more downwards.

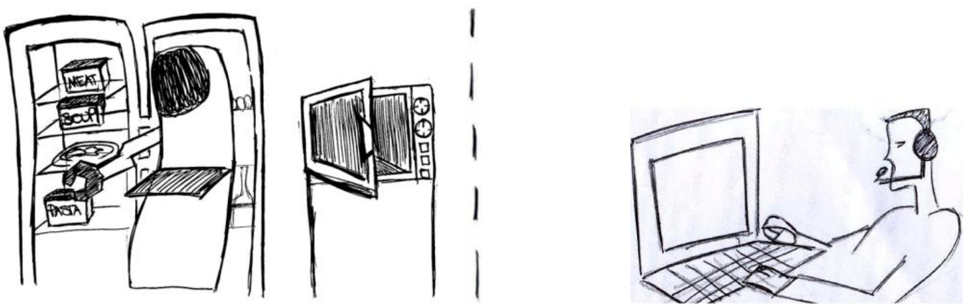
RO-PRI, RO-PRO, LU: Then Martin, Elisabeth's son joins the remote session. Because Elisabeth can no longer move her legs due to strong pain, the three decide to call an ambulance. Martin logs off to come over in person and Claudia from the 24-hour service keeps talking to Elisabeth.

RO-PRO: The ambulance arrives before Martin and rings the door bell. As Elisabeth cannot move, the teleoperator navigates SRS to the door to open it. *Note: In case door opening should turn out to be too risky from a safety perspective, SRS could fetch medication for Elisabeth.* SRS fails to find a suitable grasping point. Claudia tries to use user-assisted grasping mode (3D model approach) but it fails too. Therefore, she changes to professional manual mode and uses the force-feedback device to open the door. The ambulance personnel enters and helps Elisabeth.

8.3 COMPLEX TEACHING ACTION SEQUENCY (PREPARING FOOD)

Aim of this scenario: This scenario primarily aims to demonstrate SRS's ability to learn action sequences.
Actors: LU , RO-PRI
Places:

- Into the house: Kitchen
- Outside the house: office



RO-PRI: Because Elisabeth Baker (84) recently neglected to eat, stating she has no appetite, she and her son Martin have agreed that Martin would prepare lunch for her daily for a while and they would have a chat while he does it. Therefore, Martin today during his lunch break at work calls his mother.

LU: Elisabeth accepts the call and they talk about how her day went. Martin asks her what she would like to eat and Elisabeth chooses pasta.

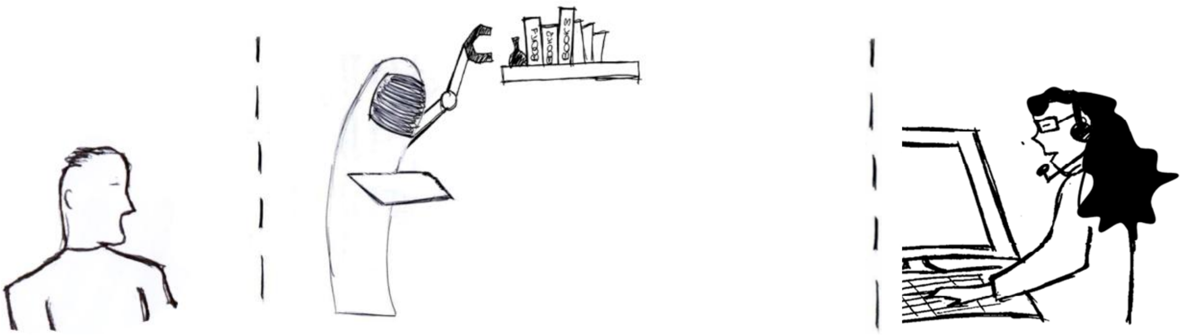
LU and RO-PRI: During the conversation, Martin directs SRS to the kitchen. Through SRS he opens the microwave oven, then the fridge, fetches the pasta microwave meal package, puts in the microwave, closes fridge and microwave oven, turns on the microwave by setting it to 5 min, fetches some water and puts it on the table, and after 5 min fetches the food and places it on the table. At the end of the process, Martin receives a message from SRS notifying him that a similar action sequence as today has been carried out 2 times before. SRS displays the recognized sequence and asks if it should be saved for future autonomous execution:
1. Open microwave
2. …
16. Place object *pasta microwave meal* on *living room table*.
Martin is also given the option to edit the action sequence before saving it. E.g. he can shorten it, delete elements, or define variable elements that SRS should ask for before executing the sequence. Martin thinks to himself “This is nice, so next time I can fully focus on my conversation with Mum and I will simply wait for SRS to finish preparing the meal, only intervening in case SRS encounters a problem.”
Martin cuts the segment “Fetch object *water bottle;* bring to location *living room table*” because his mother often has some water sitting there already. Also, Martin sets the sequence object *pasta microwave meal* a variable object so SRS will next time ask what kind of food to prepare.

RO PRI: Next day: Martin again calls his mother. However, today, SRS prepares the meal autonomously and Martin and his mother chat on how her day has been.

8.4 FETCHING AND CARRYING OF DIFFICULT OBJECTS

Aim of this scenario: fetch an object too high on a shelf or too heavy (still to be decided depending on what is most feasible with the robot prototype platform).
Actors: LU, RO-PRO
Places:

- Into the house: Kitchen
- Outside the house: teleoperator centre



Scenario screenplay

LU: Francesco Rossi (78) is mentally still quite fit. However, he does not feel safe climbing a ladder and has fallen before. He has an SRS system to help him with difficult objects. Since he has no cognitive deteriorations, he usually handles SRS himself, only falling back to a teleoperator in case it fails to execute an interaction with SRS.
Francesco wants to find some information in an old book located on a high shelf. He uses his interaction device to navigate SRS by map to the shelf.
Since Francesco knows that SRS has never before seen this object, he switches to 3D object model approach of grasping. However, after several failed attempts, he gives up (the book is surrounded by other books causing problems with the collision-free path planning for the arm).
Recognizing the failed attempts, SRS suggests to forward the interaction request to Gianni, his son. Francesco agrees. Gianni does not answer however, so SRS suggests forwarding the interaction request to the 24-hour service. Francesco agrees.

RO-PRO: Claudia from the 24-hour service answers the call and sees on her screen the steps that lead SRS to suggest to call him (failed manipulation attempts). She greets Francesco and asks him to explain what he would like to do.

LU: Francesco explains it and shows her the book.

RO-PRO: Claudia uses the professional manual mode with the force-feedback interaction device to grasp the book, moving aside the other books. Knowing that Francesco will later want to return the book on his own, Claudia teaches SRS the book by the “rotate-on-gripper” approach.

LU and RO-PRO: Francesco says thank you and the two agree to end the remote session.

LU: Francesco searches the book and finds what he was looking for. He now uses the standard semi-autonomous grasping mode to return the book. He simply taps the object on his device (it is highlighted by a rectangle) and places it back on the shelf by tapping the desired place on the shelf.

9 APPENDIX 2 - ADOPTED AND DEVELOPED VALIDATION METHODOLOGIES

9.1 THE SELECTED STANDARD METHOD: ATTRAKDIFF™

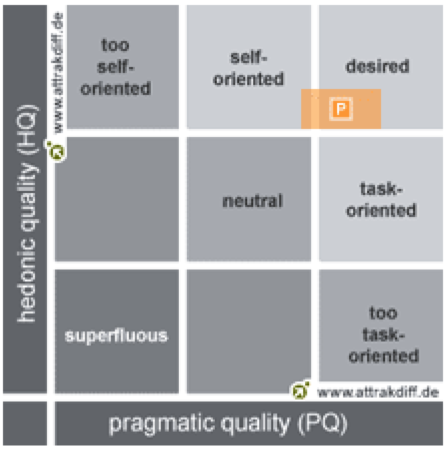
AttrakDiff™ (http://www.attrakdiff.de/) is the selected standard method trough which usability and acceptance of the system will be mainly investigated. It consists of word-pairs which simplify the rating procedure. It enables to gauge how the attractiveness of the product is experienced, in terms of usability and appearance and whether optimisation is necessary.

This type of evaluation lends itself to one-off evaluations.

Example: A software-prototype P was evaluated by future users using AttrakDiff. Ten users participated in the evaluation and the results were following:

The prototype was rated well in both hedonic and pragmatic quality. There was little room for optimisation.

The confidence rectangle shows that according to user consensus, the HQ is greater than the PQ. For prototype P the confidence rectangle extends from the desired area and into the self-oriented area. It can therefore not clearly be classified as desirable.



Medium value of the dimensions with prototype P

Confidence rectangle

ATTRAKDIFF WORS-PAIRS

human	○	○	○	○	○	○	○	technical
isolating	○	○	○	○	○	○	○	connective
pleasant	○	○	○	○	○	○	○	unpleasant
inventive	○	○	○	○	○	○	○	conventional
simple	○	○	○	○	○	○	○	complicated
professional	○	○	○	○	○	○	○	unprofessional
ugly	○	○	○	○	○	○	○	attractive
practical	○	○	○	○	○	○	○	impractical
likeable	○	○	○	○	○	○	○	disagreeable
cumbersome	○	○	○	○	○	○	○	straightforward
stylish	○	○	○	○	○	○	○	tacky
predictable	○	○	○	○	○	○	○	unpredictable
cheap	○	○	○	○	○	○	○	premium
alienating	○	○	○	○	○	○	○	integrating
brings me closer to people	○	○	○	○	○	○	○	separates me from people
unpresentable	○	○	○	○	○	○	○	presentable
rejecting	○	○	○	○	○	○	○	inviting
unimaginative	○	○	○	○	○	○	○	creative
good	○	○	○	○	○	○	○	bad
confusing	○	○	○	○	○	○	○	clearly structured
repelling	○	○	○	○	○	○	○	appealing
bold	○	○	○	○	○	○	○	cautious
innovative	○	○	○	○	○	○	○	conservative
dull	○	○	○	○	○	○	○	captivating
undemanding	○	○	○	○	○	○	○	challenging
motivating	○	○	○	○	○	○	○	discouraging
novel	○	○	○	○	○	○	○	ordinary
unruly	○	○	○	○	○	○	○	manageable

10 APPENDIX 3 – SCAI ANALISYS: AN APPLICATION EXAMPLE

SRS Objective: enable elderly people to continue to live at own home.

Costs include: equipment, maintenance, related services, human assistance

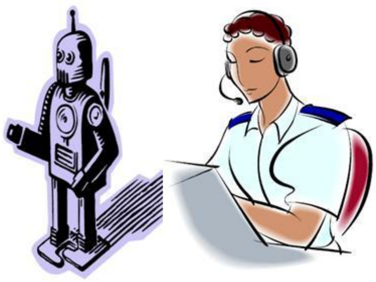
Costs are expressed in euro, SRS purchasing price and Governament contributions are completely invented just to show a concrete example of SCAI instrument usage.

Human assistance can be: Level A: anybody; Level B: strenght; Level C: professional

Costs of human assistance are based on Italian current values

Solution 1:

Elder living alone at his home, Robot purchased, and 24 hour service



Solution 1	SRS ROBOT and 24 hour service
Purchase Price	20000
Direct social cost in 10 years	86800
User expenditure	76.800
National health system expenditure	10000
Tax recovery	
Total expenditure in 10 years	86800

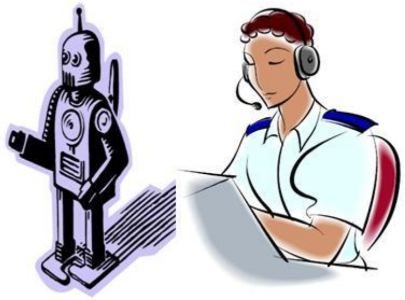
Ipothesys: technical duration 10 years
Assistance level B: 18 €/hour

Economic elements :
- Robot € 20000
- Energy/Maintenance 200 €/anno
- 24 hour service assistance: 6480 €/year

Financial elements :
- Robot purchasing (20000€) +
- 24hour service (6480) +
- Mainteinance -
- Governament Contribution

Solution 2

Elder living alone at his home, Robot for rent, and 24 hour service



Solution 3	SRS for rent
Purchase Price	
Direct social cost in 10 years	126800
User expenditure	116800
National health system expenditure	10000
Tax recovery	
Total expenditure in 10 years	126800

Ipothesys: technical duration 10 anni
Assistance level B: 18 €/hour

Economic elements :
- Robot for rent and 24 hour service (6000 €/year plus 6480 €/year)
- Manteinance and energy (200€/year)

Financial elements :
- Robot rent and service +
- Mainteinance -
- NHS Contribution

Solution 3:

Elder living alone at his home, Robot purchased, relative RO

Ipothesys: technical duration 10 anni
Assistance level B: 18 €/hour

Economic elements :
- Robot € 20000
- Energy/Maintenance 200 €/anno
- Relatives assistance: 6480 €/year

Solution 2	SRS , relative RO
Purchase Price	20000
Direct social cost in 10 years	**78.800**
User expenditure	12000
National health system expenditure	10000
Tax recovery	
Total expenditure in 10 years	22000

Financial elements :
- Robot purchasing (20000€) +
- Mainteinance -
- NHS Contribution -
- Governament Contribution =

Solution 4:

non-intervention with SRS robot: caregiver 24 hours at home with elderly

Hypothesis of non-intervention: caregiver 24 hours at home with elderly
(Asistance level A -12 €/hour)

Economic elements :
- Caregiver at home , 22 hours work every day (1200 € /month)

Solution 4	**24 hour caregiver**
Purchase Price	
Direct social cost in 10 years	**144.000**
User expenditure	**134000**
National health system expenditure	**10000**
Governament expenditure	
Tax recovery	
Total expenditure in 10 years	**144.000**

Financial elements :
- User expenditure -
- Governament contribution

A Comparative analysis between solutions considering the solution “caregiver at home” as “non intervention” would bring to this conclusion: all the proposed solution of SRS interventions appear to be less expensive in terms of socio economic impact compared with the “non intervention”.

Solutions	SRS ROBOT and 24 hour service	SRS and 24 hour service for rent	SRS , relative RO
Purchase Price	20000		20000
Direct social cost in 10 years	86800	126800	86800
User expenditure	76.800	116800	12000
National health system expenditure	10000	10000	10000
Tax recovery			
Total expenditure in 10 years	86800	126800	22000

Additional cost

=

Cost of “intervention”

-

cost of “non intervention”

86800-144000= -57200

126800-144000= -17200

22000-144000= -122000