Grant Agreement number: 288899

Project acronym: Robot-Era

Project title: Implementation and integration of advanced Robotic systems and intelligent Environments in real scenarios for ageing population

Funding scheme: Large-scale integrating project (IP)

Call identifier: FP7-ICT-2011.7

Challenge: 5 – ICT for Health, Ageing Well, Inclusion and Governance

Objective: ICT-2011.5.4 ICT for Ageing and Wellbeing

Project website address: www.robot-era.eu

D8.2

Report on results obtained in the first cycle of the Robot-Era experimentation

Due date of deliverable: 31/12/2013
Actual submission date: 21/02/2014

Start date of project: 01/01/2012 Duration: 48 months

Organisation name of lead contractor for this deliverable: SSSA
Deliverable author: [NAME AND SURNAME OF THE AUTHOR]

Version: [1, 2, …]

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	**x**
PP	Restricted to other programme participants (including the Commission Service)	
RE	Restricted to a group specified by the consortium (including the Commission Service)	
CO	Confidential, only for members of the consortium (including the Commission Service)	

Document History

Version	Date	Author	Summary of Main Changes
1.0	10-12-2013	Filippo Cavallo (SSSA)	First version of the template for Robot-Era Deliverables
1.1	13-12-2013	Raffaele Esposito (SSSA)	Structure and preliminary contents
1.2	17-12-2013	Roberta Bevilacqua/ Elisa Felici (INRCA)	Contribution in chapter 2
1.3	20-12-2013	Susann Klemcke (YOUSE)	Contribution in chapter 2
1.4	23-12-2013	Raffaele Esposito (SSSA)	Contribution in chapter 2
1.5	3-01-2014	Filippo Cavallo (SSSA)	Revision of chapter 2
1.6	7-01-2014	Raffaele Esposito (SSSA)	Contribution in chapters 3 and 4
1.8	10-01-2014	Raffaele Esposito (SSSA)	Contribution in chapters 5, 6 and 7
1.9	13-01-2014	Susann Klemcke (YOUSE)	Contribution in chapters 5, 6 and 7
1.10	17-01-2014	Roberta Bevilacqua/ Elisa Felici (INRCA)	Contribution in chapters 5, 6 an 7
1.11	19-01-2014	Filippo Cavallo (SSSA)	Revision of chapter 3,4,5,6 and 7
1.12	22-01-2014	Raffaele Esposito (SSSA)	Contribution in chapters 8, 9 and 10
1.13	23-01-2014	Susann Klemcke (YOUSE)	Contribution in chapters 8, 9 and 10
1.14	24-01-2014	Roberta Bevilacqua/ Elisa Felici (INRCA)	Contribution in chapters 8 and 9
1.15	27-01-2014	Frank Broz (UOP)	Contribution in chapters 10
1.16	28-01-2014	Filippo Cavallo (SSSA)	Revision of chapter 8,9 and 10
1.17	29-01-2014	Raffaele Esposito (SSSA)	Contribution in chapters 11 and 12
1.18	31-01-2014	Raffaele Esposito (SSSA)	Contribution in chapters 13, 14 and 15

1.19	31-01-2014	Raffaele Esposito (SSSA)	Contribution in chapters 16 and 17
1.18	3-02-2014	Roberta Bevilacqua/ Elisa Felici (INRCA)	Contribution in chapters 13,14 and 15
1.19	4-02-2014	Susann Klemcke (YOUSE)	Contribution in chapters 13 and 14
1.20	5-02-2014	Filippo Cavallo (SSSA)	Revision of chapters 13,14 and 15
1.21	10-02-2014	Roberta Bevilacqua/ Elisa Felici (INRCA)	Contribution in chapters 16 and 17
1.22	14-02-2014	Filippo Cavallo (SSSA)	Revision of chapters 13,14 and 15
1.23	17-02-2014	Filippo Cavallo (SSSA)	Revision of chapters 16 and 17
2.0	20-02-2014	Filippo Cavallo (SSSA)	Contribution in chapters 18 and final revision

Table of Contents

Executive summary

The aim of the evaluation phase with elderly users is to investigate and measure effectiveness, usability and acceptability of the Robot-Era robotic services to generate feedbacks for improvement.
This document reports the results of the experimentation of six robotic services integrated in smart environments in the Peccioli pilot site and five services in the Angen pilot site.

1 Introduction

After designing and developing the Domestic, Condominium and Outdoor robotic platforms and defining the Robot-Era services, the first experimental loop started in Peccioli (Italy) on July and in Angen (Sweden) on September and it is still going on. 44, in Italy, and 12 , in Sweden, senior citizens 65+ years old were recruited to use and test the Robot-Era system, composed of the three robots and the AmI infrastructure, in the realistic indoor and outdoor environments of the Peccioli Town Lab and Angen. During the experimentation elderly volunteers were invited to interact with the robots in order to evaluate the technical effectiveness and acceptability of the Robot-Era robotic services.

In the Peccioli pilot site the following services were tested:

- Shopping and drug delivery by 35 elderly persons
- Communication by 35 elderly persons
- Garbage collection by 35 elderly persons
- Reminding by 27 elderly persons
- Indoor walking support by 27 elderly persons
- Outdoor walking support by 24 elderly persons

In the Angen pilot site the following services were tested:

- Food delivery by 11 elderly persons
- Laundry delivery by 10 elderly persons
- Reminding by 9 elderly persons
- Escort at night by 6 elderly persons
- Communication by 12 elderly persons

2 Analysis methodology

2.1 Usability analysis: methodology

For analysing the usability of the Robot-Era system a mixed method approach was set up. That means a mix of qualitative and quantitative data was used to select important issues with regard to usability.
The combination of the qualitative and quantitative data generated during and after the experimentation makes it possible to give an overall statement on the usability of the different Robot-Era services and specific issues (e.g. related to the Robot-Era interface). The following paragraphs will describe the specific results generated from the different data sources.

2.1.1 Usability by SUS scale analysis: methodology

The SUS consists of 10 items, with odd-numbered items worded positively and even-numbered items worded negatively[1]. To use the SUS, present the items to participants as 5-point scales numbered from 1 (anchored with "Strongly disagree") to 5 (anchored with "Strongly agree"). If a participant fails to respond to an item, assign it a 3 (the center of the rating scale). After completion, determine each item's score contribution, which will range from 0 to 4. For positively-worded items (1, 3, 5, 7 and 9), the score contribution is the scale position minus 1. For negatively-worded items (2, 4, 6, 8 and 10), it is 5 minus the scale position. To get the overall SUS score, multiply the sum of the item score contributions by 2.5. Thus, SUS scores range from 0 to 100 in 2.5-point increments. The interpretation of the score is [2]:

- 0-64 points: not usable
- 65-84 points: usable
- 85-100 points: excellent

2.1.2 Usability by video analysis: methodology

During the experimentation videos were taken from (at least) three different perspectives, which show the test person, the robot and the interaction on the tablet interface. The evaluation of the usability of the Robot-Era System is mainly based on a video-interaction analysis [3] of the experimentation. All videos taken during the experimental loop were analysed service by service. The analysis was conducted as an explorative content analysis [4] and was validated through multiple experts in the field of usability studies (at least two persons discussed each problem occurred during the analysis). After the selection of usability problems was made, the problems were classified and weighted with regard to international usability standards and norms (e.g. EN ISO 9241-110). Related to the weight of a usability problem specific recommendations were developed to optimize the usability perception from the users perspective and to improve the Robot-Era System and its system components.

To weight a usability problem the following benchmarks were used:

- Frequency of occurrence of the problem – percentage of test persons who had the problem during the test?

- Influence of the problem on fulfilling the task – Was it possible for the user to complete a task or not?
- Potential of learning from the users side– Does the problem disappear after some time of training? Or does it still occur?

The problems then were classified from 1 (no usability problem) to 5 (major usability problem – high priority/ changes on the system component are indispensable). The following paragraphs will give an overview of the main usability problems (weighted as important to change (Level 4/5)) and recommendations for the first experimental loop. Please find the detailed analysis in the annex.

2.2 Acceptance analysis: methodology

The acceptance is the sum of many factors as acceptability, attitude, HRI and etc. For analysing the acceptance of the Robot-Era system a mixed method approach was set up. That means a mix of qualitative and quantitative data was used to select important issues with regard to acceptance.
The combination of the qualitative and quantitative data generated during and after the experimentation makes it possible to give an overall statement on the acceptance of the different Robot-Era services and specific issues (e.g. related to the Robot-Era interface). The following paragraphs will describe the specific results generated from the different data sources.

2.2.1 Acceptance by ad hoc-questionnaire analysis: methodology

The acceptance ad-hoc questionnaire consists of several items and the users can reply to these statements on a 5 points Likert type scale. The negative items are converted in positive ones to be compared. In order to get the overall Acceptance score ranged from 0 to 100, we simply need to sum the score for each item and multiply it by a factor which depends by the number of items. The interpretation of the score is:

- 0-64 points: not usable
- 65-84 points: usable
- 85-100 points: excellent

2.2.2 Acceptability and HRI by video analysis: methodology

As reported in D8.1, a detailed checklist was developed for collecting and recording all the relevant information during the video analysis. The observation was conducted by two different trainee psychologists and one expert in the field of HRI, with the aim of capturing the information related to each tested service.
Before starting with the analysis, a set of indicators for acceptability and HRI was selected for each service, starting from:

- the definition of acceptability from the Unified Theory of Acceptance and Use of Technology [5] based on 13 core constructs, whose anxiety, attitude and perceived enjoyment were selected as the most appropriate ones to be retrieved;
- the definition of social robot as "an autonomous or semi-autonomous robot that interacts and communicates with humans by following the behavioral norms

expected by the people with whom the robot is intended to interact" [6]. From this definition, we have selected the following dimensions to be evaluated:

- communication, as it represents the basis of the interaction;
- physical interaction/not verbal communication, as set of specific indicators for understanding the emotional reaction during the cooperation with the robot.

In particular, for assessing the perceived enjoyment, the attitude, and the anxiety (= Acceptability), it was decided to collect free statements of the users that express enjoyment/disappointment with the task and the overall service, predisposition to the interface use, and feeling of being/not being competent, as well as, emotional reaction during the task execution, through the non verbal communication – i.e. smile, laugh, playing, scaring, etc…-.

Regarding HRI, it was decided to guide the observation on the physical reaction/interaction with the robot, through the analysis of the:

- Head orientation: robot head, robot body, robot eyes, experimenter, tablet
- Gaze: robot, experimenter, default
- Body orientation: robot, experimenter, default (straight ahead)
- Proxemics: approach robot, touch robot
- Difficulty: avoiding obstacle, turning, moving forward for the Indoor/Outdoor escort service.

At the beginning, it was necessary to distinguish among the indicators to be taken into consideration, in order to avoid the overlapping among the acceptability and HRI dimensions. After consultation among the partners involved in this phase of work, it was decided to report as "acceptability" all the contents oriented to the task performance/service and as "HRI" all the contents oriented on the robot itself or its features (i.e. I think I can enjoy shopping using the robot = acceptance; I think the robot is very nice = HRI; I think the robot it is easy to use = acceptance; I like the presence of the robot inside the home = HRI).

Results of the first experimental loop in Peccioli

3 Participants

3.1 Participants recruited for Shopping and Drug delivery, Communication and Garbage collection services

For the first three services, Shopping and Drug delivery, Communication and Garbage collection, we recruited 35 elderly persons. They were 22 woman and 13 men over 65 years old Table 3.1. Their mean age was 73,80±5,81 and their reached educational level was primary education for 14,3%, junior high school for 14,3%, secondary education for 57,1% and university for 14,3% Table 3.2. About their marital status, 49% of the participants was married and 31% was widowed Table 3.3. 34% of them lived alone in their own home and 66% lived with the spouse or the relatives. All recruited people were retired and the majority of them had a monthly economic income from 500€ to 2000€ Table 3.4. About their levels of autonomy, 89% of them had a High Level of User Autonomy (HLUA) and 11% had a Middle Level of User Autonomy (MLUA) according to the Instrumental Activities of Daily Living (IADL) test. Nobody had cognitive problems according to the Short Portable Mental Status Questionnaire.
About their skills regarding the everyday technological devices as expected, most of them knew very well : TV (~91%), telephone (~89%) / mobile phone (~74%) and washing machines (~74%) Table 3.5. An encourage data was that ~40% of them use a PC and ~49% internet, but most of the participants knew nothing about smartphone (~66%) and tablet (~71%) Table 3.5.

Table 3.1 Elderly people's characteristics

Sex	#	%
Male	13	37%
Female	22	63%

Table 3.2 Elderly people's education

Education	#	%
No education	0	0,0%
Primary education	5	14,3%
Junior high School	5	14,3%
Secondary Education	20	57,1%
University	5	14,3%

Table 3.3 Elderly people's Marital status

Marital status	#	%
Married	17	48,6%
Full time relationship	1	2,9%
Separated	2	5,7%

Divorced	2	5,7%
Single	2	5,7%
Widowed	11	31,4%

Table 3.4 Elderly people's income

Monthly family or personal income	#	%
0 - 500 €	0	0,00%
501 - 1000 €	7	20,59%
1001 - 1500 €	10	29,41%
1501 - 2000 €	7	20,59%
2001 - 2500 €	3	8,82%
2501 - 3000 €	5	14,71%
3001 - 3500 €	1	2,94%
3501 - 4000 €	0	0,00%
more than 4000 €	1	2,94%

Table 3.5 Attitude: Use of technology

Attitude: Use of technology							
	1	**2**	**3**	**4**	**5**	**not applicable**	**don't know**
Phone	0,00%	0,00%	2,86%	8,57%	88,57%	0,00%	0,00%
Mobile phone	0,00%	5,71%	8,57%	11,43%	74,29%	0,00%	0,00%
Smart phone	14,29%	0,00%	0,00%	0,00%	20,00%	65,71%	0,00%
TV/ Remote control	0,00%	0,00%	0,00%	8,57%	91,43%	0,00%	0,00%
Remote control for TV	0,00%	0,00%	0,00%	5,71%	94,29%	0,00%	0,00%
CD/DVD player	11,43%	2,86%	5,71%	11,43%	51,43%	17,14%	0,00%
PC	17,14%	5,71%	2,86%	14,29%	40,00%	20,00%	0,00%
Internet	17,14%	2,86%	5,71%	5,71%	48,57%	20,00%	0,00%
Tablet	20,00%	0,00%	2,86%	0,00%	5,71%	71,43%	0,00%
Bank ATM	11,43%	0,00%	2,86%	2,86%	71,43%	11,43%	0,00%
Devices to increase your safety	0,00%	0,00%	8,57%	8,57%	60,00%	22,86%	0,00%
Washing machine	14,29%	2,86%	0,00%	8,57%	74,29%	0,00%	0,00%
Telemedicine	11,43%	0,00%	5,71%	0,00%	2,86%	74,29%	5,71%
Alarm system	8,57%	2,86%	0,00%	0,00%	20,00%	68,57%	0,00%

3.2 Participants recruited for Indoor walking support, Outdoor walking support and Reminding services

For the second three services, Indoor walking support, Outdoor walking support and Reminding, we recruited 27 elderly persons. They were 20 woman and 7 men over 65 years old Table 3.6. Their mean age was 73,00±6,60 and their reached educational level was

primary education for 33,3%, junior high school for 11,1%, secondary education for 40,7% and university for 14,8% Table 3.7. About their marital status, 48% of the participants was married and 37% was widowed Table 3.8. 33% of them lived alone in their own home and 67% lived with the spouse or the relatives. All recruited people were retired and the majority of them had a monthly economic income from 500€ to 2000€ Table 3.9. About their levels of autonomy, 89% of them had a High Level of User Autonomy (HLUA) and 11% had a Middle Level of User Autonomy (MLUA) according to the Instrumental Activities of Daily Living (IADL) test. Nobody had cognitive problems according to the Short Portable Mental Status Questionnaire.
About their skills regarding the everyday technological devices as expected, most of them knew very well : TV (~71%), telephone (~96%) / mobile phone (~59%) and washing machines (~81%). A goodly data was that ~33% of them use a PC and ~41% internet, but most of the participants knew nothing about smartphone (~81%) and tablet (~70%) Table 3.10.

Table 3.6 Elderly people's characteristics

sex	#	%
Male	7	**26%**
Female	20	**74%**

Table 3.7 Elderly people's education

Education	#	%
No education	0	0,0%
Primary education	5	14,3%
Junior high School	5	14,3%
Secondary Education	20	57,1%
University	5	14,3%

Table 3.8 Elderly people's marital status

Marital status	#	%
Married	13	48%
Full time relationship	0	0%
Separated	1	4%
Divorced	2	7%
Single	1	4%
Widowed	10	37%

Table 3.9 Elderly people's income

Monthly family or personal income	#	%
0 - 500 €	2	7,41%
501 - 1000 €	7	25,93%
1001 - 1500 €	5	18,52%
1501 - 2000 €	4	14,81%

2001 - 2500 €	2	7,41%
2501 - 3000 €	4	14,81%
3001 - 3500 €	1	3,70%
3501 - 4000 €	0	0,00%
more than 4000 €	1	3,70%
Don't know	1	3,70%

Table 3.10 Attitude: Use of technology

Attitude: Use of technology							
	1	**2**	**3**	**4**	**5**	**not appli cable**	**don't know**
Phone	0,00%	0,00%	0,00%	3,70%	96,30%	0,00%	0,00%
Mobile phone	0,00%	7,41%	25,93%	7,41%	59,26%	0,00%	0,00%
Smart phone	11,11%	0,00%	0,00%	0,00%	7,41%	81,48%	0,00%
TV/ Remote control	0,00%	0,00%	0,00%	5,71%	71,43%	0,00%	0,00%
Remote control for TV	0,00%	0,00%	0,00%	0,00%	77,14%	0,00%	0,00%
CD/DVD player	11,43%	0,00%	0,00%	8,57%	40,00%	17,14%	0,00%
PC	25,93%	7,41%	7,41%	11,11%	33,33%	14,81%	0,00%
Internet	25,93%	3,70%	7,41%	7,41%	40,74%	14,81%	0,00%
Tablet	22,22%	0,00%	0,00%	0,00%	7,41%	70,37%	0,00%
Bank ATM	18,52%	0,00%	3,70%	0,00%	62,96%	14,81%	0,00%
Devices to increase your safety	0,00%	0,00%	3,70%	11,11%	55,56%	29,63%	0,00%
Washing machine	7,41%	3,70%	3,70%	3,70%	81,48%	0,00%	0,00%
Telemedicine	14,81%	0,00%	0,00%	0,00%	0,00%	81,48%	3,70%
Alarm system	7,41%	3,70%	0,00%	0,00%	25,93%	62,96%	0,00%

4 Acceptance of Robot-Era Robots Appearance Results

An ad hoc questionnaire about the appearance of Robot-Era platforms, that could be replied to on a five point Likert type scale, was developed. In order to evaluate the acceptance of Robot-Era platforms appearance, the constructs were classified into six categories:

- Anxiety: Evoking anxious when the Robot-Era platforms were seen
- Aesthetics: Positive or negative judgements about Robot-Era platforms aesthetics
- Intention to interact: The outspoken intention to interact with the Robot-Era platforms when they were seen
- Trust: Inspired confidence when the Robot-Era platforms were seen
- Safety: The belief that the Robot-Era platforms perform reliability
- Facility Conditions: Factors of the Robot-Era platforms that could facilitate their use
- Perceived functions: The degree to which a person believes that Robot-Era platforms communicate their functions

Before starting the experimental test, the users observed the Robot-Era platforms for some minutes and then they replied to the questionnaire items. The questionnaire was completed by 44 participants and in the following we present the results concerning the proposed constructs about DORO, CORO and ORO Table 4.1.

The appearance of the Robot-Era platforms didn't evoke anxiety (1,61±1.06 DORO, 1,40±0,89 CORO and 1,58±1,05 ORO) in elderly people because according to their opinion the robots didn't look dangerous. Also seniors said they didn't feel overpowered by the three platforms because their appearance was evaluated proportionate and symmetric eliciting a sense of "order" and "elegance". These data were also confirmed by a positive judgment about the aesthetics (4,17±0,87 DORO, 4,40±0,76 CORO and 4,29 ±0,87 ORO) in fact according to the participants the colours of the robots were appropriate and the appearance was in good agreement with a domestic environment for DORO, a condominium one for CORO and an urban area for ORO. A high degree of confidence (4,32±0,88 DORO, 4,39±0,81 CORO and 4,23±0,96 ORO) was inspired after the Robot-Era platforms observation in fact many elderly people were well-disposed (4,15±1,22 DORO, 4,17±1,18 CORO and 4,14±1,17 ORO) to interact with robots and the interaction was also encouraged by the presence of an head on them.

After the first observation elderly people perceived the reliability of the Robot-Era platforms (3,92±0.99 DORO, 4,01±0,90 CORO and 4,10±0,94 ORO) because the robots and their components seemed robust and the materials used for their development were evaluated appropriate. Also the domestic, condominium and outdoor robots were developed considering several factors that could facilitate their use as position of the touch-screen and the dimensions compared to the three different environments and these facility factors got a positive feedbacks from users (4,12±1,23 DORO, 4,02±1,07 CORO and 4,19±1,10 ORO).

In spite of the good impression made on elderly people, the appearance of Robot-Era platforms didn't communicate their possible functions explicitly, in fact elderly people were unable to express a judgment (3,37±1,55 DORO, 3,33±1,51 CORO and 3,21±1,57 ORO).

Table 4.1 Descriptive statistics

	DORO		CORO		ORO	
	mean	std	mean	std	mean	std
Anxiety	1,61	1,06	1,40	0,89	1,58	1,05
Aesthetics	4,17	0,87	4,40	0,76	4,29	0,87
Intention to interact	4,15	1,22	4,17	1,18	4,14	1,17
Trust	4,32	0,88	4,39	0,81	4,23	0,96
Safety	3,92	0,99	4,01	0,90	4,10	0,94
Facility Conditions	4,12	1,23	4,02	1,07	4,19	1,10
Perceived functions	3,37	1,55	3,33	1,51	3,21	1,57

For reporting the results, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the Robot-Era appearance platform was not acceptable, from 65-84 was acceptable and from 85-100 the result was excellent. At last the Robot-Era platforms were acceptable (80,20±8,53 DORO Table 4.2, 81,58±9,70 CORO Table 4.3 and 80,58±9,38 ORO Table 4.4) by elderly participants (over 60%) and also we can note that over 30% of the users thought excellent the appearance of the three robots.

Table 4.2 Acceptance of DORO Descriptive statistics

DORO	mean	std	#	%
Not Acceptable 0-64	63,33	-	1	2%
Acceptable 65-84	75,82	4,39	29	66%
Excellent 85-100	90,48	4,56	14	32%
Total	80,20	8,53	44	100%

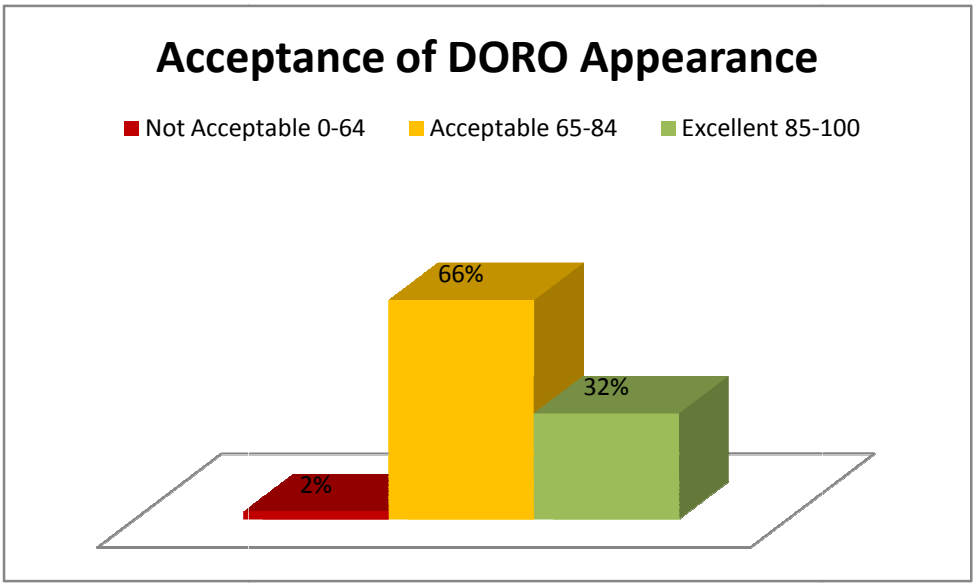


Figure 4.1 Acceptance of DORO result

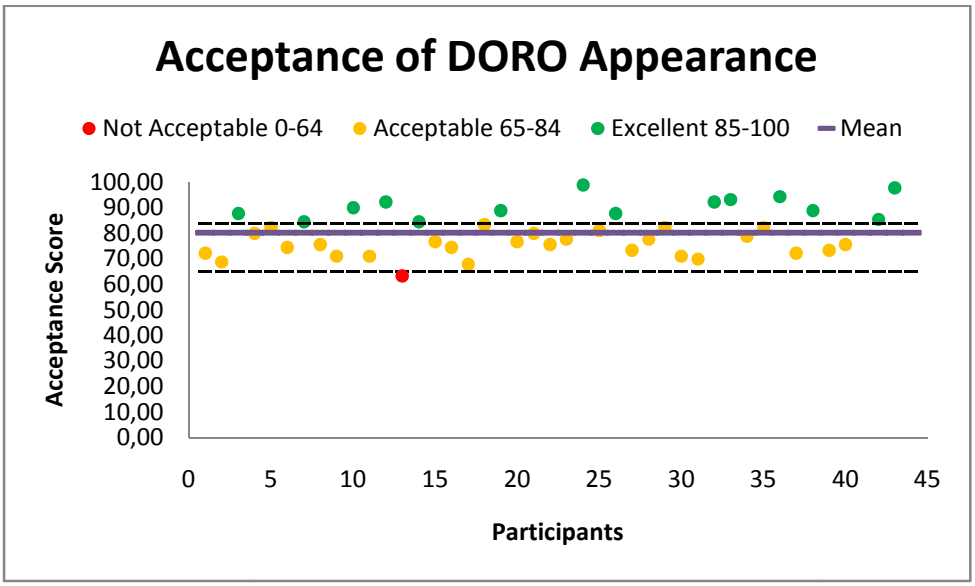


Figure 4.2 Acceptance of DORO result

Table 4.3 Acceptance of CORO Descriptive statistics

CORO	mean	std	#	%
Not Acceptable 0-64	58,67	-	1	2%
Acceptable 65-84	76,64	4,8	28	64%
Excellent 85-100	92,62	4,94	15	34%
Total	81,58	9,70	44	100%

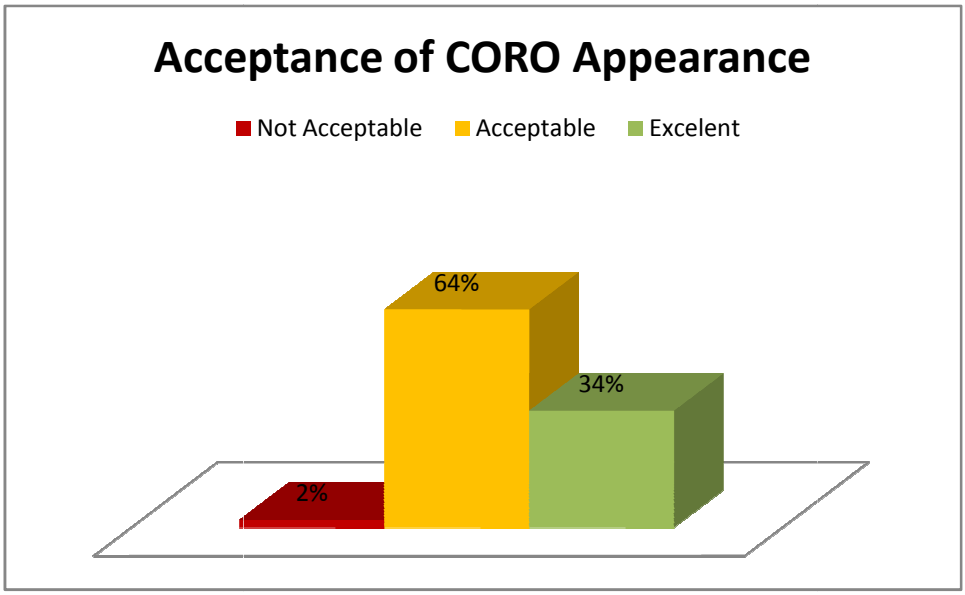


Figure 4.3Acceptance of CORO result

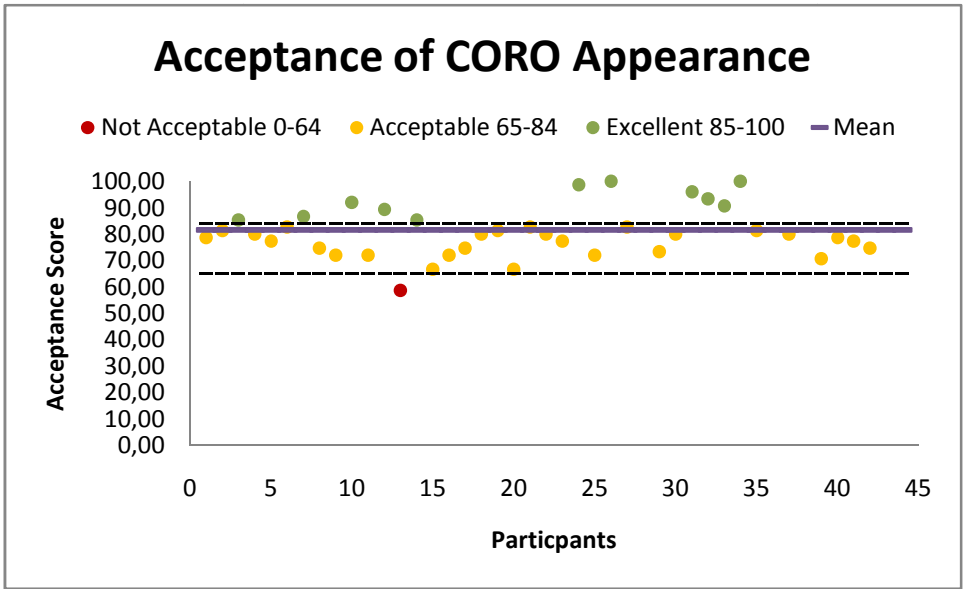


Figure 4.4 Acceptance of CORO result

Table 4.4 Table 4.2 Acceptance of ORO Descriptive statistics

ORO	mean	std	#	%
Not Acceptable 0-64	-	-	-	0%
Acceptable 65-84	75,08	5,49	29	66%
Excellent 85-100	91,02	5,13	15	34%
Total	80,58	9,38	44	100%

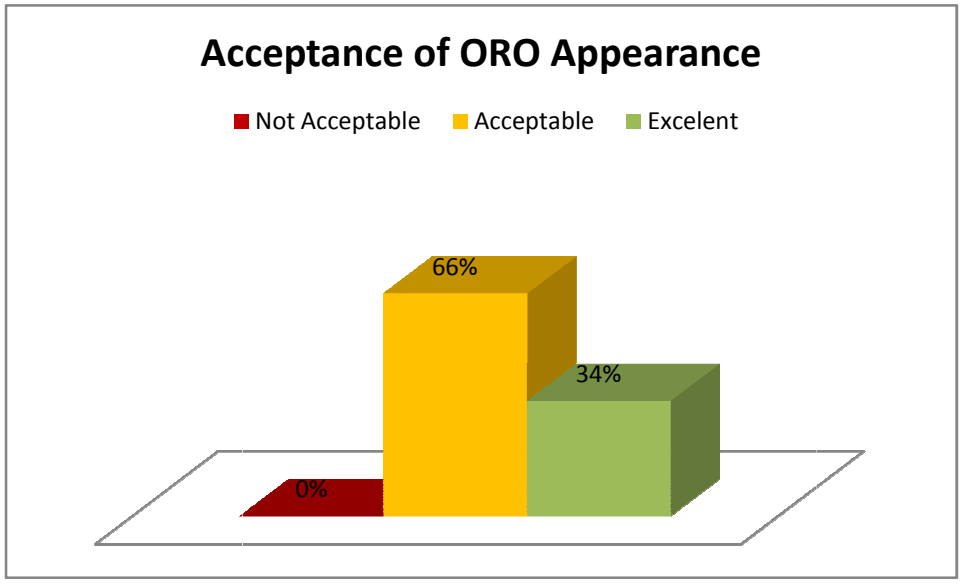


Figure 4.5 Acceptance of ORO result

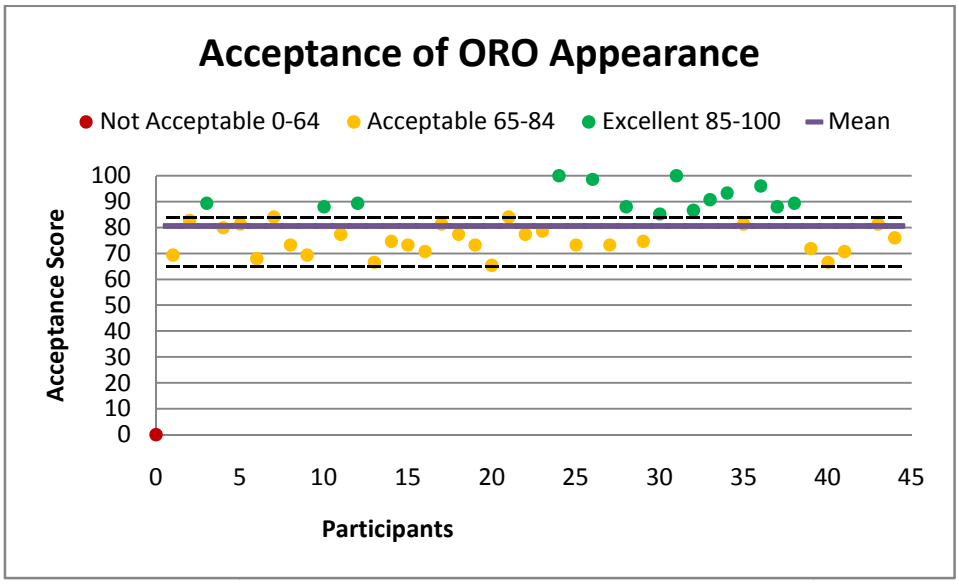


Figure 4.6 Acceptance of ORO result

5 Shopping and Drug delivery service results

In this scenario the elderly user had to imagine he was sick and he could not leave his home, but he needed several things to eat and drinks. Bearing in mind this presupposition the participant had to create a shopping list with 5 products, to send it and to wait for the shopping delivery. The speech interaction and interface interaction were necessary to perform this service. This service was composed by several steps that are shown in Figure5.1.

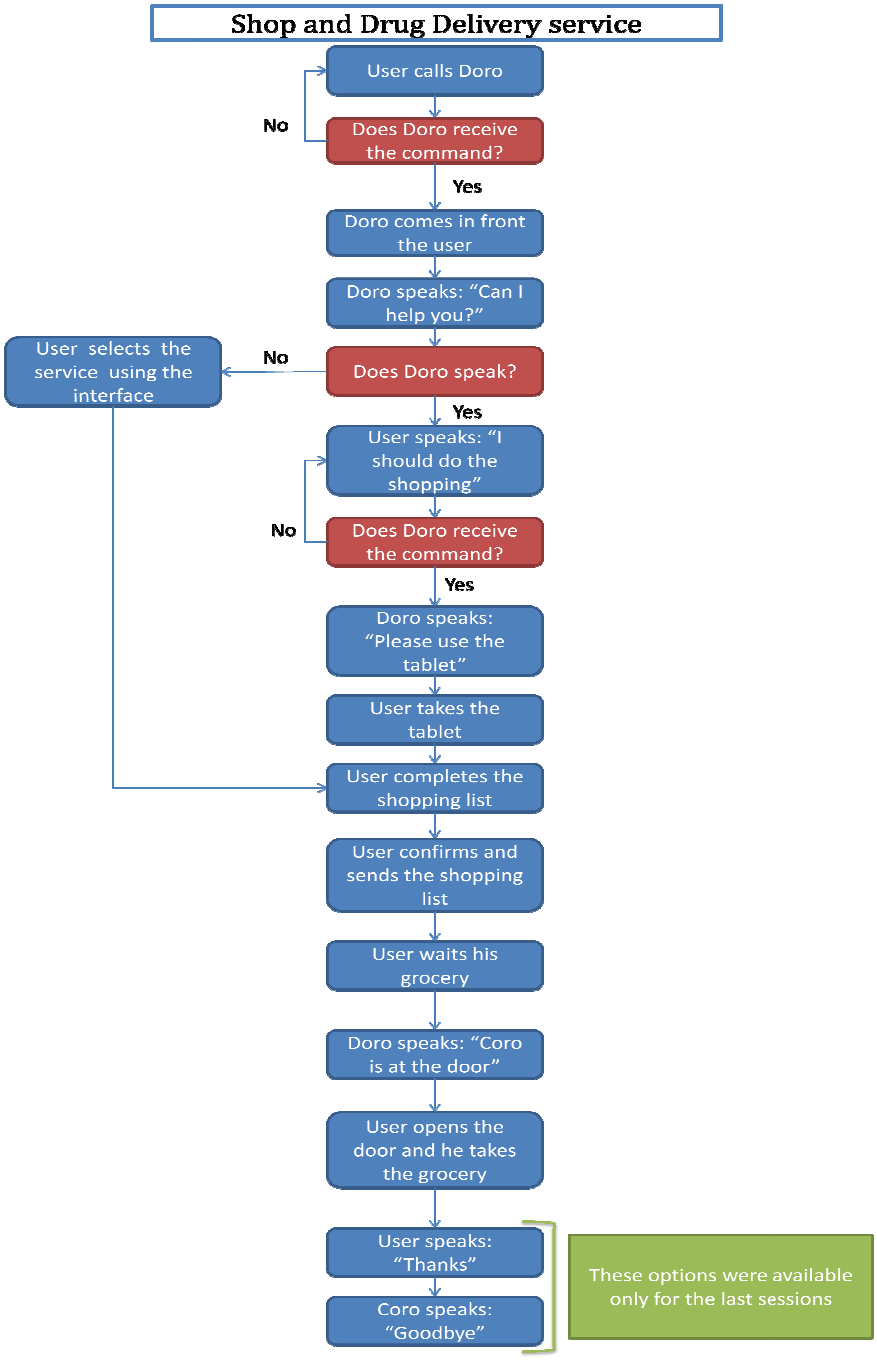
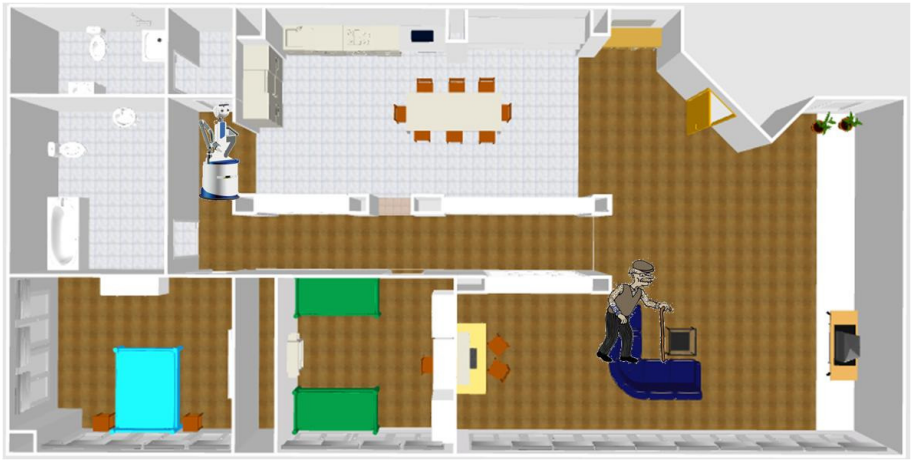


Figure 5.1 Shopping and Drug Delivery service flow chart

In this scenario all three Robot-Era platforms were involved working in three different environment: domestic, condominium and outdoor Figire5.2.



a

b

c

Figure 5.2 a: domestic environment b: condominium environment c: outdoor environment

Figure 5.3 Users during the experimentation

5.1 Usability

5.1.1 Usability by SUS Scale

In order to evaluate the usability of "Shopping and Drug delivery service" we proposed the system usability scale (SUS), a simple ten item Likert scale that gives a global view of subjective assessments of usability.

Table 5.1 SUS Score: Descriptive Statistics

Shopping and Drug Delivery	mean	std	#	%
Not Usable 0-64	50,83	7,85	6	17%
Usable 65-84	77,5	5,49	8	23%
Excellent 85-100	90,72	4,04	21	60%
Total	80,86	15,81	35	100%

The "Shopping and Drug delivery service" was perceived as usable (80,86±15,81) Table5.1 by elderly people and 60% of participants reported an excellent evaluation. In particular the easy use of the service was goodly estimated (4,03±1,32). Also older people thought that the actions performed by the Robot-Era platforms were well integrated (4,34±1,08) and usable.

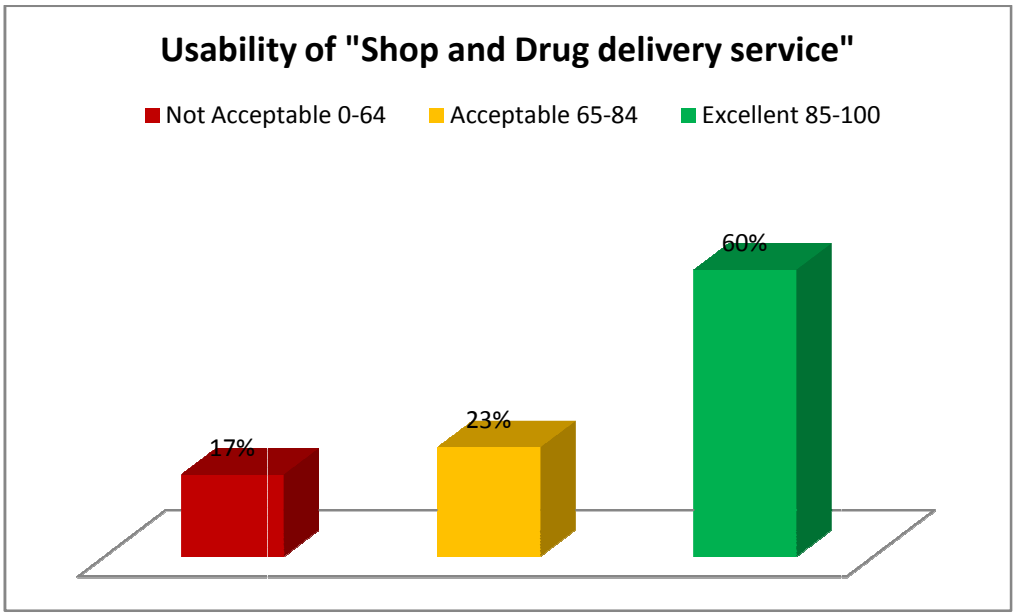


Figure 5.4 Usability result

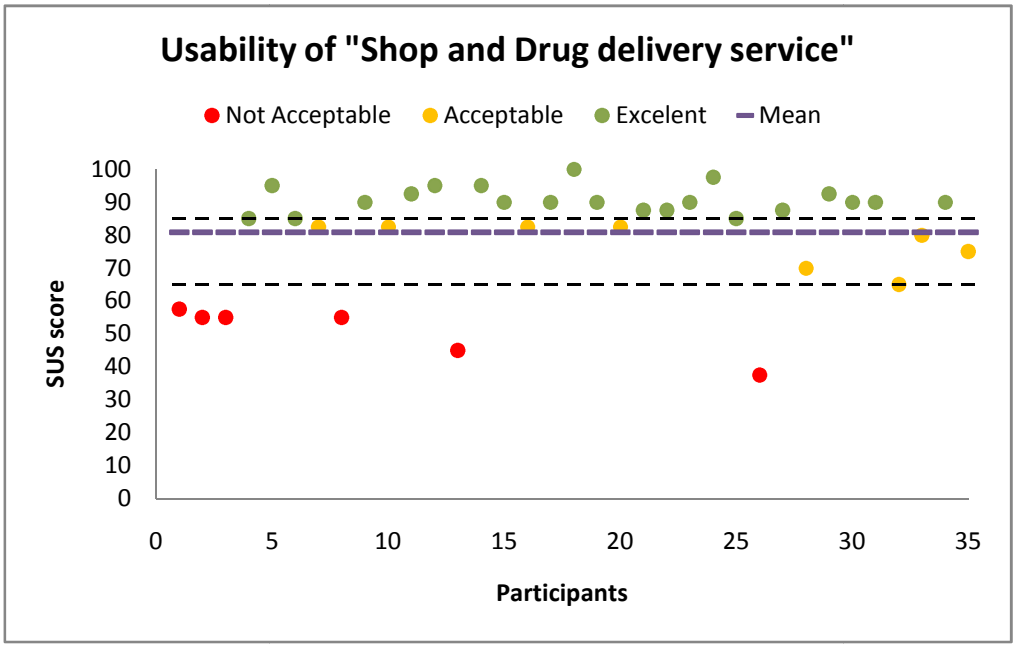


Figure 5.5 Usability result

5.1.2 Usability by video analysis

The test persons did not understand the concept to buy different items at the same time. They tried to buy each item after another. The interface did not inform the person enough. The task to "create a shopping list" as an action for "putting items into a shopping cart and then buy all the items at once" did not seem to be clear. The task "creating a shopping list" was not intuitive enough maybe because they expected a regular shopping list to be written on paper and not being part of the virtual purchasing process. The task needs a different name, instead of "creating a shopping list" e.g. "selecting the items to buy" and a more detailed request e.g. "you can know select all the items you would like to buy", "when you

are finished selecting the items you need to confirm your order", buttons: "confirm order" / "change selection".

The test persons did not understand the step "to enter the shopping list", because the interface did not inform the person enough that the list needs to be "entered" or confirmed. The task to enter the list did not seem to be intuitive to the participants. They did not expect the entering to be necessary. They expected the selected items already to be entered automatically.

The test persons did not see or recognize the food icons correctly, because the pictures did not show the food clearly enough due to the cuttings. The participants often did not recognize the "eggs" icon and mixed up the icons for "water" and "milk" because the pictures were too similar or not identifiable. Many of the participants confused the "pasta" icon with carrots, i.e. the picture was too indistinct.

The categorization of the items was unclear to the participants. The names of the categories were not self-descriptive enough. The participants did not understand the categorization that was given. They did not find the items they expected under the category names, e.g. they tried to find "milk" under the category "beverages". The categories need to be either clearer or omitted.

There is no information given that a selection of more than one item needs to be tapped as much as the item is wanted. The participants only tapped the food icon twice, when they were explicitly told to do so. They did not intuitively understand this function by themselves.

After fulfilling the task, the interface does not confirm or inform about the ending of the shopping process. The participants expected to get a confirmation of their order and an information that the task is finished.

5.2 Acceptance

5.2.1 Attitude

The elderly user demonstrated a medium attitude (3,40±1,35) towards this service maybe because they were autonomous for their daily activities. Also taking care of the shopping needs was not perceived as a burden but as a way to socialize and meeting friends. However the users reported positive responses if they would use the Robot-Era platforms for doing shopping in case of need (4,66±0,94) or if the robots could help the family/caregiver's work (4,49±1,07).

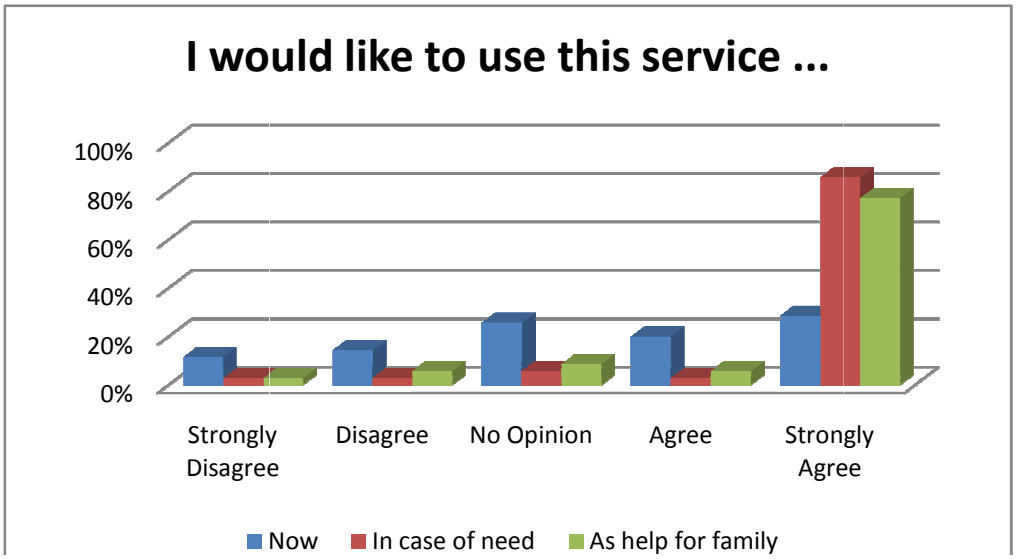


Figure 5.6

5.2.2 Acceptability

5.2.2.1 Acceptability by ad-hoc questionnaire

The questionnaire compiled by the old persons denoted a positive degree of acceptability (4,66±0,79) about the "Shopping and Drug delivery service". In particular using the Robot-Era system for shopping didn't evoke anxiety in the participants who reported they enjoyed during the experimentation and they would trust in the robot ability to perform shopping service.

5.2.2.2 Acceptability by video analysis

On the acceptability, it can be said that only one participant experienced anxiety during the performance, mostly shown by her emotional reaction during the execution of the task – her gaze was always on the experimenter for looking at support and on the tablet -. This is an encouraging result, even if it has to be compared with the quantitative data from the questionnaires. In general, the attitude towards technology was a driver factor for what concerns the execution of the performance. Nine of the subjects have shown to have a high predisposition to the tablet use and it seems that the past experience with technology has mediated the performance of the service in each phase. The users that have more confidence with technology have also provided an interesting corpus of statements for updating the tablet graphic interface and that can constitute part of the next usability and acceptability requirements to be implemented in D2.7 ["I prefer clearer icons for coffee, pasta and others. Maybe, the name of goods can be written under the icons", "The icons are too small", "It would be necessary to have an explanation of the goods under the icons", "What does it mean "essentials"? There could be differences among the people!", "You should change some icons, as for example pasta and eggs", "You should copy the supermarket categories"]. In some case, some subjects that reported to not have experience with the technology have shown a high predisposition too, even if some difficulty/negative evaluations were reported by two participants ["I do not understand the

icons on the tablet", "The tablet interface is ok on an aesthetical point of view, but the functioning is not good at all"].
Sixteen subjects have verbally expressed to be enjoyed by the use of the robot, one of them were mainly concentrated on the tablet use, so it is not clear whether this is motivated by the presence of the robot or not. This can be evinced by the free statements of the users, that in some cases expressed appreciation for the physical appearance of DORO ["From now, I think I will like more the robot, while tablet is still a little complicate for me. I need to exercise more", "I really like the robot with this face!",, "The robot is nice", "The robot is very nice", "I'm intrigued by the robot"] and emotional activation [i.e. by laughing and joking with DORO and ORO: "Thanks for coming here DORO!", "You are arrived very soon ORO, thank you!!!", "Let's take a picture with DORO"]. Other positive evaluations were given on the tasks performed by DORO, that can be related to a good level of acceptability of the service ["It could be useful if CORO would bring the shopping bags inside the home", "I do not think that the shopping service is too slow. If I was at home, I could do other things while he shops for me", "I'm happy that DORO does not open the door, because I would be afraid that he would open also to intruders"].
In general, it was observed a connection between the expressed enjoyment and the production of verbal statements on the robot, that suggest a relationship between enjoyment and communication or, eventually, differences in personality traits (i.e. novelty seeking, extraversion,…) that can facilitate the acceptability of the platforms, supporting its use. This issue should be better approached during the second loop enrollment, in order to understand possible strategies for the customization of the robot that can "attract" and facilitate the robot use by the individuals.

5.2.3 HRI

5.2.3.1 HRI by questionnaire: Speech and GUI interaction

Regarding the speech interaction elderly users reported positive responses (4,75±0,66), asserting that it was easy to speak to the domestic robot to perform the service, using simple "key words".
About the GUI interaction the old persons had a more neutral approach (3,71±1,25) because many of them never used a GUI on a tablet and some problems occurred. However they believed they could improve by practice.

5.2.3.2 HRI by video analysis

On the HRI, interesting input should be evinced by the non verbal communication of the subjects. Twelve subjects are focused mainly on the robot, with the head, eyes and body orientation. In one case the user started walking around the robot for looking better at him, moving close around DORO. At the same time, twelve subjects were more oriented to the tablet, suggesting that the use of this interface would have distracted the user from the interaction with DORO. In five cases, it was not possible to collect the information, for the technical failure of the system – DORO did not talk with the user – and for the impossibility of evaluating the dimension through the videos. Just one user touched the fingers of the robot, as she was looking for more interaction with it, also shown by her statements ["I'd like the robot has a human-like arm, to be adapted to my apartment"; "The robot is very

nice", "Does he talks if I touch him?"]. From an analysis of the communication aspects related to the HRI, the majority of the users seems to like DORO and its features ["I thinks DORO speed is appropriate for the domestic environment", "I like DORO voice because it is not too metallic", "I think that the robot is a game"] and look for more speech communication with it ["Is it possible to use a courtesy communication?", "I think that it is better if the robot has the vocal command, because if I have to stay in bed, it is difficult to take the tablet", "I think I like more the robot, while tablet is still a little complicate for me. I need to exercise more"], while no one expressed preferences for the tablet interface. The only negative aspect highlighted by two subjects is the too slow velocity of DORO in moving.

5.2.4 Quality of life

Regarding quality of life, elderly users reported neutral responses (3,69±1,53) to item *"I think my independence would be improved by the use of the robot for shopping"* because, as said before, performing shopping was not perceived as a burden, however they perceived the utility of this service in case of need.

Table 5.2 Acceptance attributes: descriptive statistics

Shopping and Drug delivery		mean	std	Acceptance Score	std
ACCEPTANCE	Attitude	4,57	1,00	87,69	8,23
	Acceptability	4,66	0,79		
	Speech interaction	4,75	0,66		
	GUI interaction	3,71	1,25		
	QoL	3,69	1,53		

For reporting the final acceptance result, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the "Shopping and Drug delivery service" was not acceptable, from 65-84 was acceptable and from 85-100 the result was excellent. At last the "Shopping and Drug delivery service" was acceptable (87,69±8,23) by elderly participants. However we can note that 66% of the users evaluated excellent the "Shopping and Drug delivery service".

Table 5.3 Acceptance Score: Descriptive Statistics

Shopping and Drug Delivery	**mean**	**std**	**#**	**%**
Not Acceptable 0-64	-	-	0	0%
Acceptable 65-84	78,59	6,20	12	34%
Excellent 85-100	92,44	4,09	23	66%
Total	87,69	8,23	35	100%

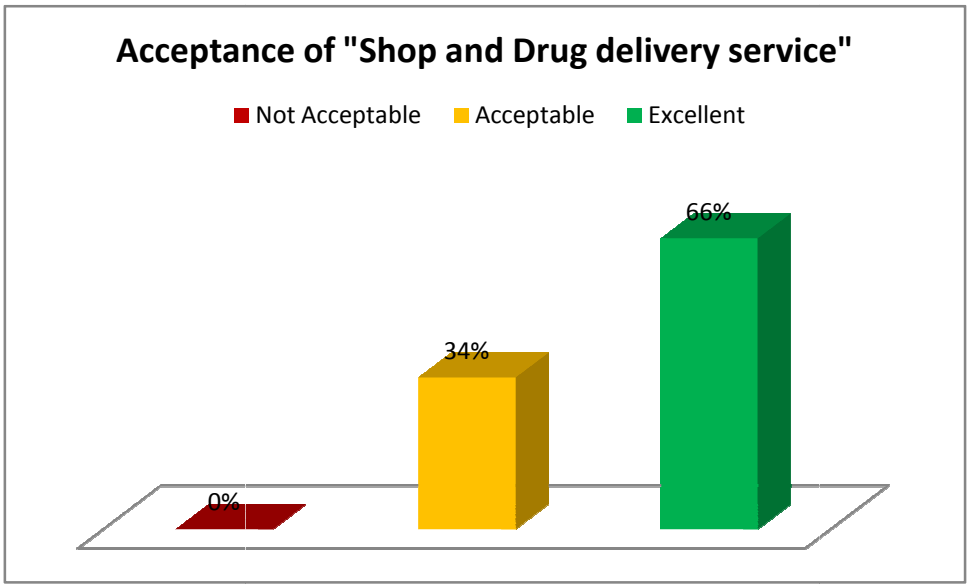


Figure 5.7 Acceptance result

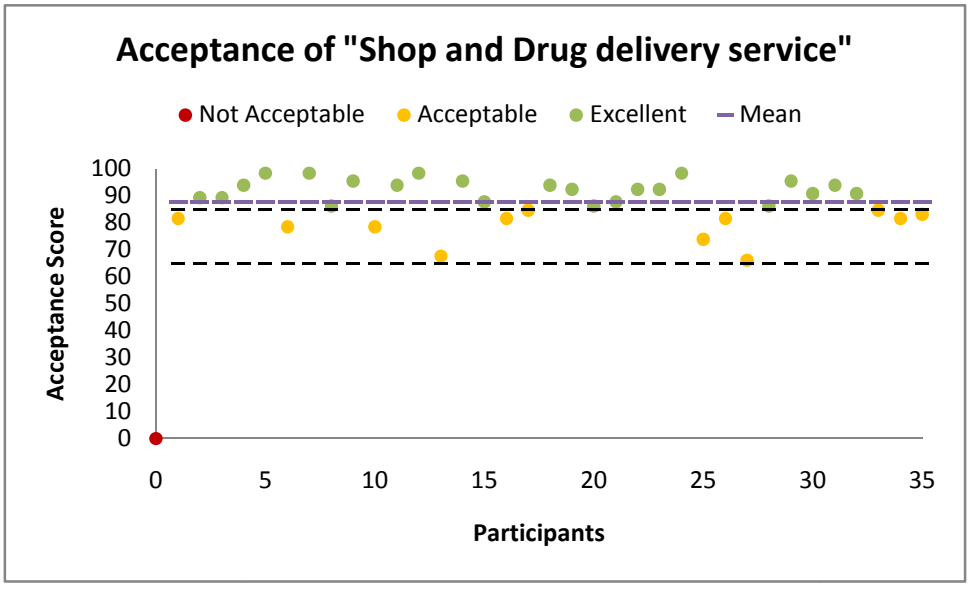


Figure 5.8 Acceptance result

5.3 Technical data

5.3.1 Docking and exchange task

Taking in account the technical evaluation of the service experimented, this mainly focused on the most critical issues which were the navigation in the elevator and the exchange good phase between condominium robot and outdoor platform.
About the docking and exchange task, a lot of failures occurred and the procedure needs a deep improvement: causes can be found both in the procedure, since the used docking

function needs a more adaptive behavior to be successfully used in this task, and in the environment, e.g. the height misalignment between indoor and outdoor areas.

Table 5.4 Docking/Exchange

	Success	Failure	No evaluated
Docking	16,67%	54,17%	29,17%
Exchange	8,33%	62,50%	29,17%

Figure 5.9 Exchange

5.3.2 Interaction with elevator

In the Shopping and Drug delivery service the success rate for the interaction of elevator can be considered acceptable; however, localization errors occurred and increased the difficulties in the navigation planner.

Table 5.5 Entering elevator

	Success	Failure	No evaluated
Entering elevator at floor 0	87,50%	0,00%	12,50%
Exiting elevator at floor 1	70,83%	16,67%	12,50%

Figure 5.10 Interaction with elevator

5.3.3 Speech recognition

With reference to Figure5.1, we analyze the number of success about the speech interaction. As shown in the table when the user spoke to the domestic robot, it received the command at the first the time in the majority of cases. Also DORO gave a notification to the elderly person without problem.

Table 5.6 Speech interaction

User Calls DORO					
# Call				**Recognized**	
1	**2**	**3**	**>3**	**Yes**	**No**
27/35 (~77%)	3/35 (~9%)	3/35 (~9%)	2/35 (~6%)	35/35 (100%)	-

Did DORO talk? "Can I help you?"			
Yes	***No***	***Other***	***Out of order***
29/32 (~91%)	*1/32 (~3%)*	*2/32 (~6%)*	*3/35 (~6%)*

User selects the service						
# Call				***Recognized***		***Out of order***
1	***2***	***3***	***>3***	***Yes***	***No***	
22/32 (~69%)	*6/32 (~18%)*	*1/32 (~3%)*	*3/32 (~9%)*	*31/32 (97%)*	*1/32 (3%)*	*3/35 (~6%)*

Did DORO talk? "Please use the tablet"

Yes	No	Out of order
32/33 (~97%)	*1/33 (~3%)*	*2/35 (~6%)*

Did DORO talk? "CORO is at the door"	
Yes	**No**
35/35 (100%)	-

5.3.4 Time to perform the service

After DORO was called by the user, the response time of the domestic robot was by 10 sec (7,13±1,99) in ~43% of cases and by 20 sec (12,15±2,3) in ~37% of cases. Table5.7

Table 5.7

Time (sec) between the user calls DORO and DORO starts

	0-9	10-19	20-29	30-39	40-49	50-60	>60
#	*15 (~43%)*	*13 (~37%)*	*2 (~6%)*	*2 (~6%)*	*1 (~3%)*	*1 (~3%)*	*1 (~3%)*
Mean	*7,13±1,99*	*12,15±2,3*	*20±0*	*33,5±3,53*	*49*	*54*	*66*

After DORO stopped in front the user, it talked "Can I help you?" by 10 sec (5,52±1,91) in (~84%) of cases. Table 5.8

Table 5.8

Time (sec) between DORO arrives in front the user and DORO starts to speak

	0-9	10-19	20-29	30-39	40-49	50-60	>60
#	*21 (~84%)*	*3 (~12%)*	-	-	-	*1 (~4%)*	-
Mean	*5,52±1,91*	*12±2,65*	-	-	-	*54*	-

After the user selected the service, the domestic robot talked "Please, use the table to complete your shopping list" by 10 sec (*6,95±1,72*) in ~63% of cases and by 20 sec (*15,00±2,24*) in ~37% of cases. Table5.9

Table 5.9

Time (sec) between service selection and DORO starts to speak

	0-10	11-20	21-40	41-60	>60	No evaluated
#	*19/30 ~63%*	*5/30 ~17%*	*3/30 ~10%*	*1/30 ~3%*	*2/30 ~7%*	*5/35 ~14%*
means	*6,95±1,72*	*15,00±2,24*	*30,67±8,08*	*55*	*78,50±3,54*	

The "Shopping and Drug delivery service" was performed by 20 minutes (16,77±1,54)in the 54% of cases. Table5.10

Table 5.10

Time (min) to perform the service						
	0-5	***6-10***	***11-15***	***16-20***	***21-25***	***>25***
#	*-*	*-*	*6(~17%)*	*19(~54%)*	*6(~17%)*	*4(~11%)*
means	*-*	*-*	*12,39±0,80*	*16,77±1,54*	*21,88±1,55*	*27,14±1,24*

6 Communication service results

In this scenario consisted in two parts:

- Warning alert case
- Phone call case

A gas leak inside the home was simulated and the AmI detected it. The domestic robot went to the user who was informed about this dangerous situation. Immediately after an incoming call by a possible caregiver was visualized on the tablet and the user had to accept it.

Then the users you wanted to use to robot to call a family member via Skype. This service could be performed by speech interaction and GUI use. Even if the communication service was composed by two parts, it was analyzed as a single service. The service was composed by several steps that are shown in Figure6.1

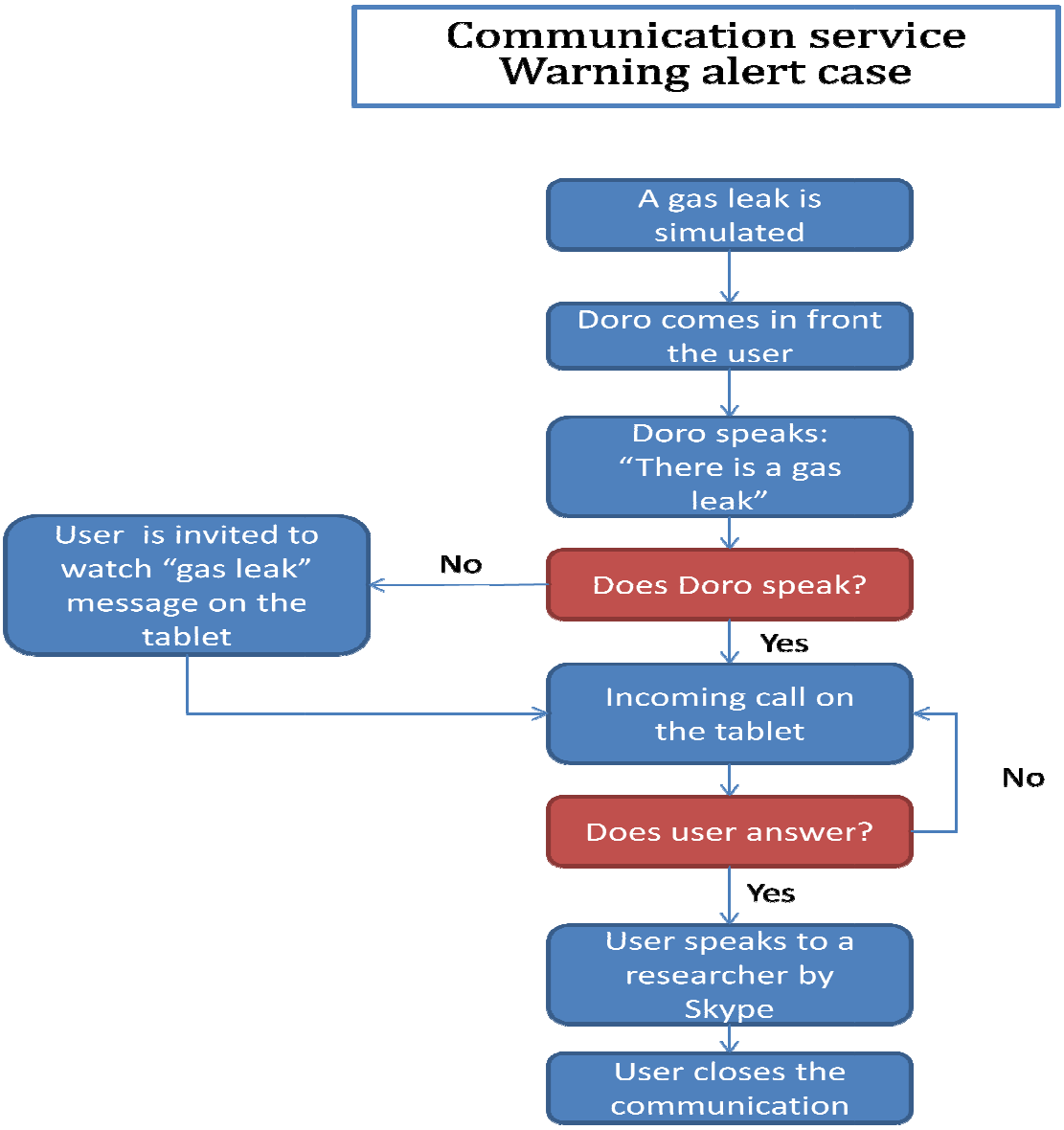


Figure 6.1Warning case flow chart

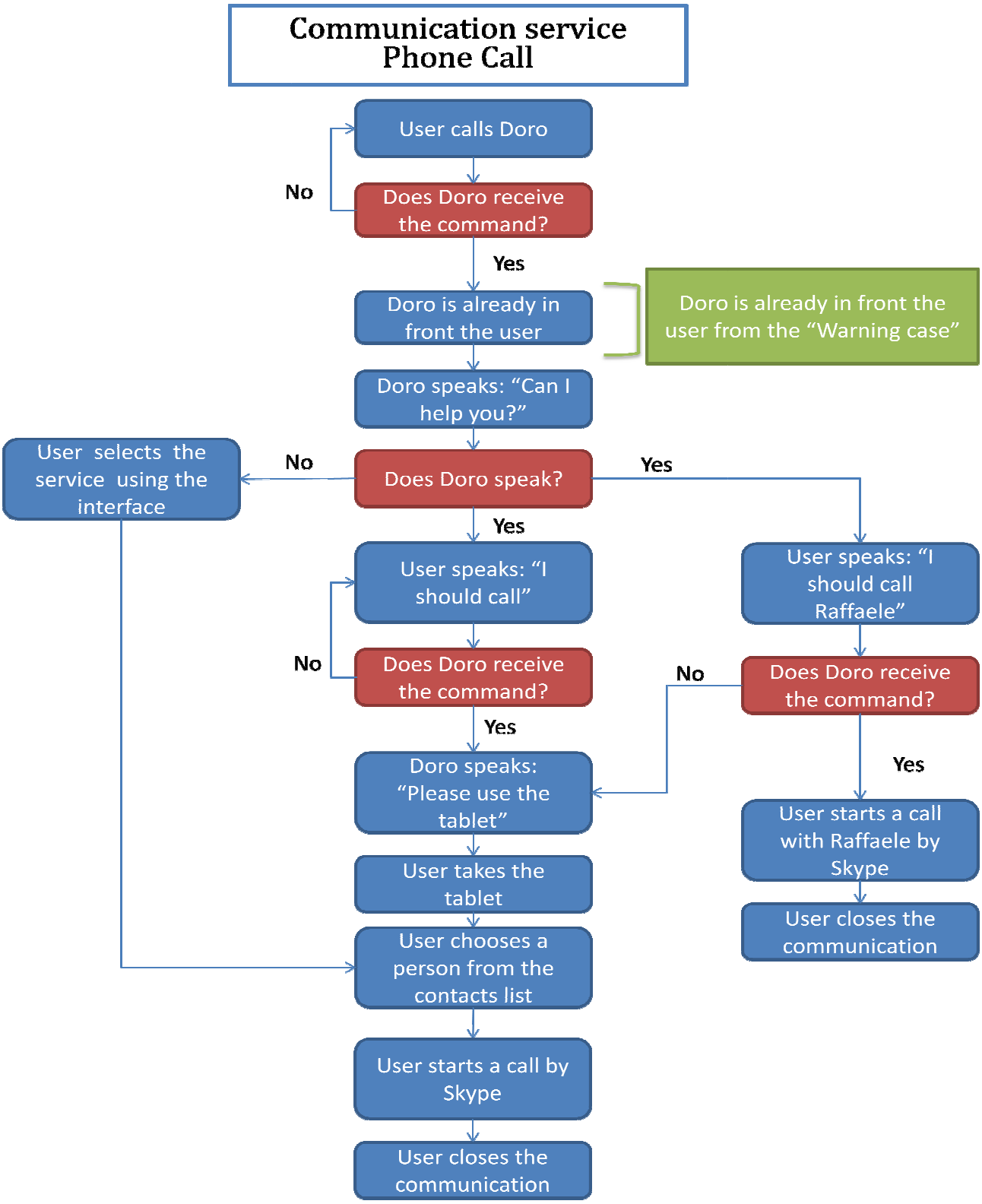


Figure 6.2 Phone call case flow chart

In this scenario only DORO was involved working in the domestic environment.

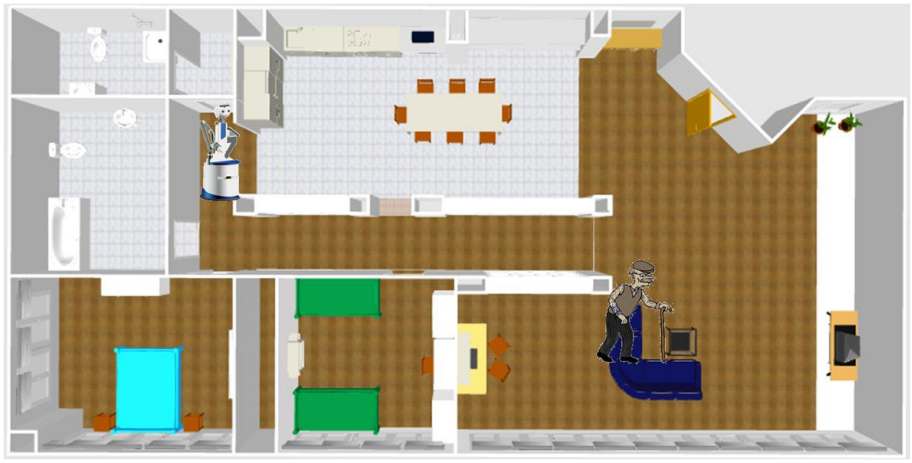


Figure 6.3 Domestic environment

Figure 6.4 User during the experimentation

6.1 Usability

6.1.1 Usability by SUS Scale

In order to evaluate the usability of "Communication service" we proposed the system usability scale (SUS), a simple ten item Likert scale that gives a global view of subjective assessments of usability.

Table 6.1Score: Descriptive Statistics

Communication	mean	std	#	%
Not Usable 0-64	41,25	15,91	2	5,7%
Usable 65-84	77,75	4,92	10	28,6%
Excellent 85-100	92,83	4,90	23	65,7%
Total	85,57	14,09	35	100%

The "Communication service" was perceived as usable (85,57±14,09) by elderly people and 66% of participants reported an excellent evaluation. In particular the easy use of the service was goodly estimated (4,49±0,92), demonstrating that the actions performed by the Robot-Era platforms were well integrated (4,63±0,69).

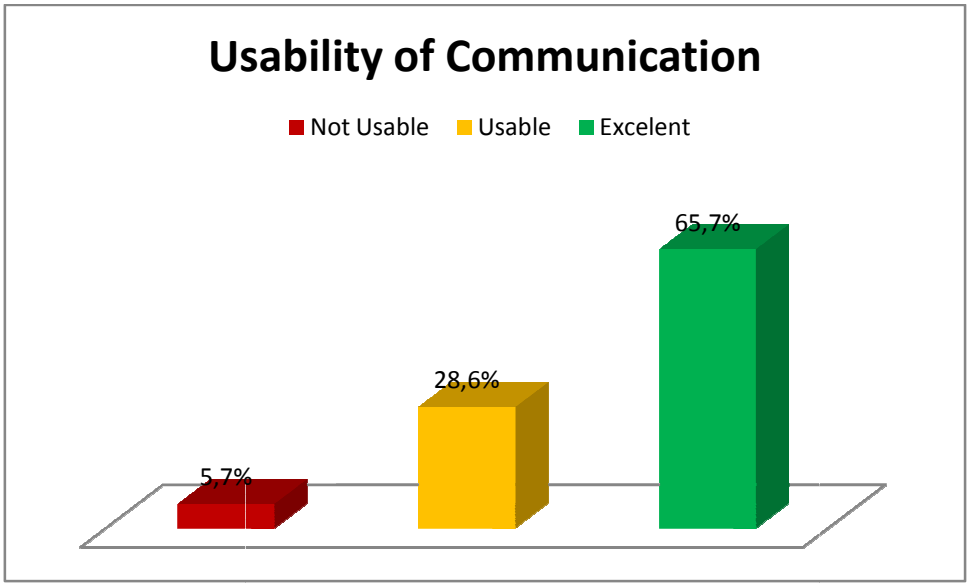


Figure 6.5 Usability result

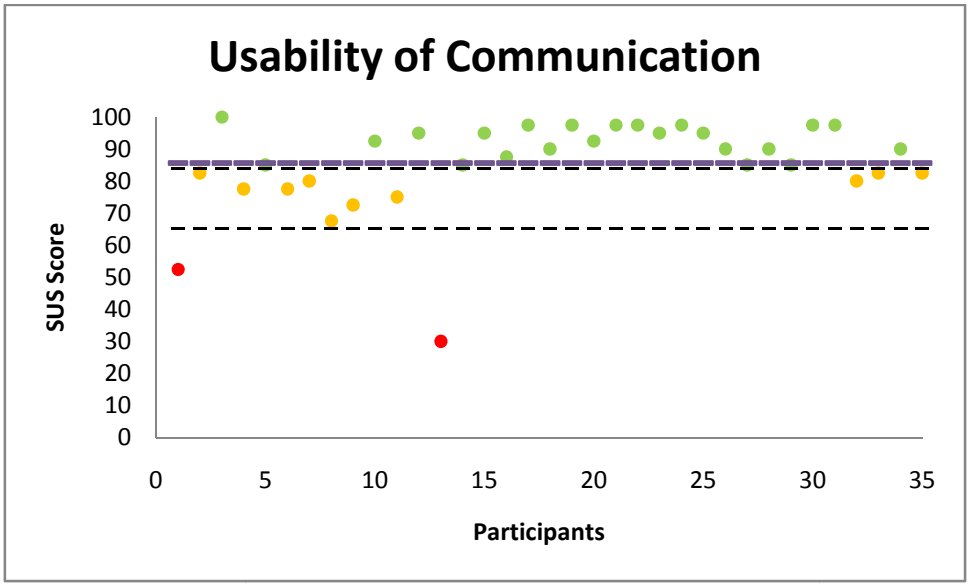


Figure 6.6 Usability result

6.1.2 Usability by video analysis

No strong usability problems can be identified during the test of the communication service.

6.2 Acceptance

6.2.1 Attitude

The elderly user demonstrated a high attitude (4,51±0,93) towards this service because they
perceived the utility of being advised in case of a dangerous situation as a gas leak. The warning alert case was judged more important than the phone call case because according to participants' opinion, at the present-day there are many devices and way to communicate and robot is not necessary.

6.2.1.1 Acceptability by ad-hoc questionnaire

The questionnaire compiled by the old persons denoted a positive degree of acceptability (4,64±0,89) about the "Communication service". In particular using the Robot-Era system for the communication didn't evoke anxiety in the participants who reported they enjoyed during the experimentation and they would trust in the robot ability to perform the service.

6.2.1.2 Acceptability by video analysis

On the acceptability, it can be said that the opportunity of receiving warnings on time is considered one of the favorite functionalities of the entire system by almost all the sample ["I think it is useful to receive warnings on that - gas leak -", "DORO should go faster, if there is a gas leak!", "The robot should give advice to someone, for example to my son", "I hope this service will be improved, because it could be useful", "I'd like to know how to use this service", "It's really nice!", "This service is interesting!", "It is interesting, personally I'm feeling more skilled with technology now!"]. No one has experienced anxiety during the use, mostly because the support of the interviewer was required all along the service duration for the system failures. Only five persons have shown a feeling of being not confident with performing this service, before starting to test it ["Technically, I'm not good!", "I think I should learn more to use technology", "I don't understand what I have to do!", "I've never used Skype, don't laugh at me, because they do!", "I don't have the mobile phone, I don't know if I will be able to use it!"]. On the service side, the users expressed the wishes of having more complex functionalities associated to the warnings ["The robot should also switch off, if there is a gas lick!", "Is it possible to turn on/off the TV or other appliances from another room, if DORO said they are switched on?"].
One of the most common difficulty encountered by the subjects is represented by how to deal with the warnings (specifically observed with three users): after a warning is given, they did not know what to say/do with DORO. This problem can be interpreted as a bias caused by the double tasks required by the use case – to understand the warning received and to answer to the incoming call – or to the lack of control over the event/process – the user do not receive any feedback on what the robot is going to do (calling someone), for example -. This aspect should be improved for the second experimental loop.
Some users have suggested more support from DORO, for what concerns the security issue ["If there are intruders will DORO advice me?", "Can the robot help in case of intruders in the home?"]. Other overall improvements were suggested by the older people statements ["We should have more people as possible to phone to!", "The tablet should be put under the face of the robot, so you can have the tablet in front when he is coming with a call!",

"The robot should call directly the services for helping in case of gas leak, for example, not familiars or friends, in order to avoid the loss of time in case of need!"].
This input can be useful mostly for improving the surveillance service and the connection between DORO and CORO, for example. Finally, emotional expressions of enjoyment during the service was detected in thirteen subjects, analyzing their facial expression and their posture. The activity of starting a call with the robot was perceived as easier to perform, even if a clear quantification of this impression is too difficult to give, due to the technical failures of the system, that many times obliged the interviewer to support the users, even if they seemed to be able to perform the task autonomously.

6.2.2 HRI

6.2.2.1 HRI by questionnaire: Speech and GUI interaction

Regarding the speech interaction elderly users reported positive responses (4,73±0,66), asserting that it was easy to speak to the domestic robot to perform the service, using simple "key words". However this result was not truthful there was a mismatch between the real working of speech recognition, many problems occurred during the tests, and elderly evaluation. Maybe the elderly evaluated the potential service than the tested service.

6.2.2.2 HRI by video analysis

On the HRI, it can be said that an overall low HRI was detected during this services. For what concerns HRI, in fact, all the chosen indicators were influenced by the technical failures as for example the head/gaze orientation that were directed mostly to the tablet to perform the tasks, and to the experimenter, to ask for/receive assistance. Also the communication with the robot, as part of the interaction functionality, was negatively influenced by the technical failures: the users repeats many times the commands or they did not understand properly the instruction, as mentioned above.

6.2.3 Quality of life

Regarding quality of life, elderly users reported positive enough responses (3,91±1,34) to item *"I think my independence would be improved by the use of the robot for this service"* . As said before, the warning alert case was judged more important than the phone call, so the participants asked to improve the warning notification service in order to increase their safety and security inside home.

Table 6.2 Acceptance attributes: descriptive statistics

Communication		mean	std	Acceptance Score	std
ACCEPTANCE	Attitude	4,51	0,93	89,97	8,33
	Acceptability	4,64	0,89		
	Speech interaction	4,73	0,66		

GUI interaction	4,18	1,16
QoL	3,91	1,34

For reporting the final acceptance result, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the service was not acceptable, from 65-84 was acceptable and from 85-100 the result was excellent. At last the "Communication service" was acceptable (88,54±7,94) by elderly participants. We can note that 61% of the users the acceptance as excellent (92,59±7,94).

Table 6.3 Acceptance score: descriptive statistics

Communication	mean	std	#	%
Not Acceptable 0-64	-	-	0	0%
Acceptable 65-84	78,42	4,99	11	31%
Excellent 85-100	92,59	4,50	24	69%
Total	88,54	7,94	35	100%

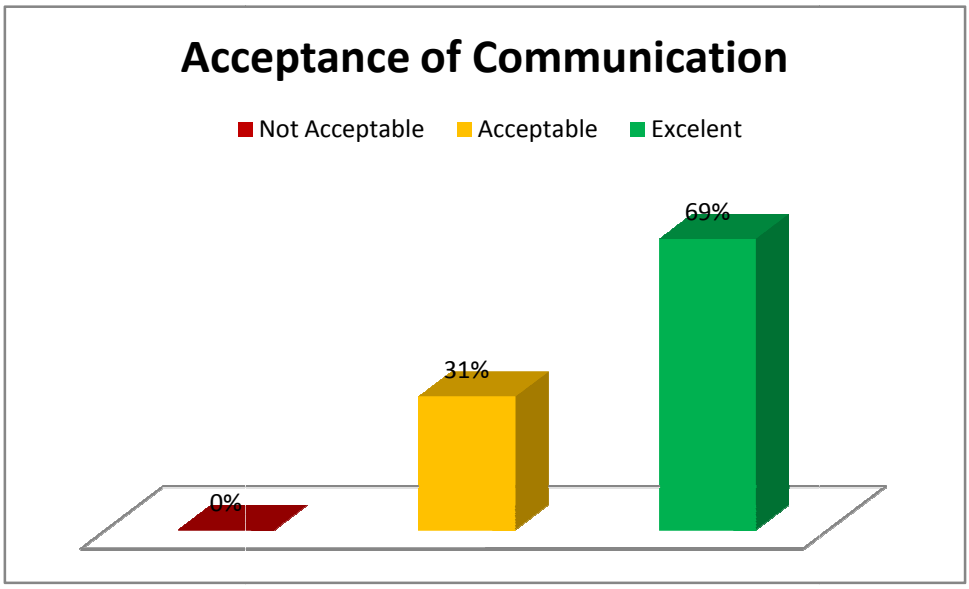


Figure 6.7 Acceptance result

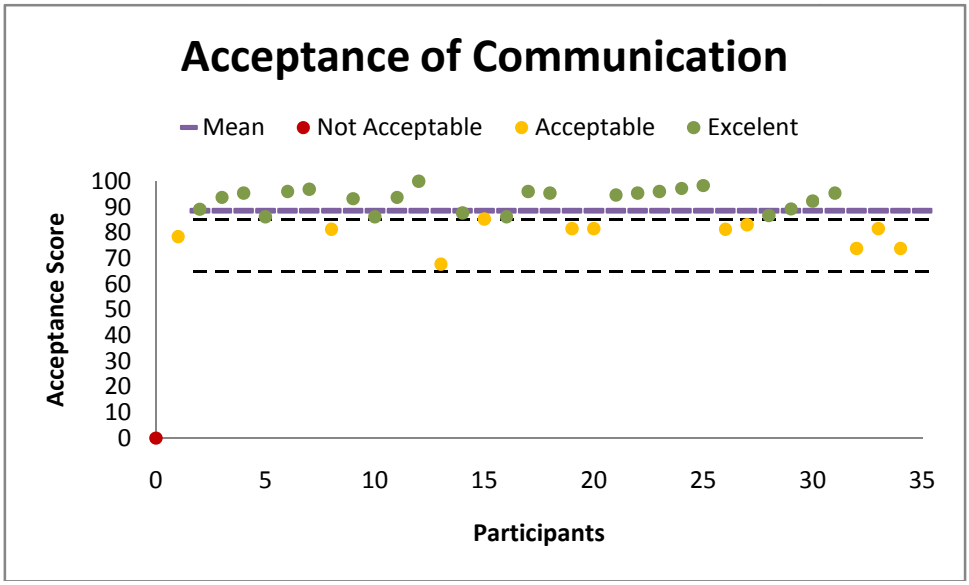


Figure 6.8 Acceptance result

6.3 Technical data

6.3.1 Speech recognition

With reference to Figure6.1, we analyze the number of success about the speech interaction.
About warning alert case the domestic robot gave the notification to the elderly person without problem Table6.4.

Table 6.4

Did DORO talk? "Gas leak"		
Yes	**No**	**Out of order**
33/35 (~94%)	**2/35 (~6%)**	**-**

About the call phone case, the domestic robot received the command at the first time only in the 56% of cases and the robot answered in 88% of cases Table6.5/6.6.

Table 6.5

User Calls DORO					
# Call				Recognized	
1	**2**	**3**	**>3**	**Yes**	**No**
9/16 (~56%)	**6/16 (~38%)**	**-**	**1/16 (~6%)**	**16/16 (100%)**	**-**

Table 6.6

Did DORO talk? "Can I help you?"			
Yes	**No**	**Other**	**Out of order**
14/16 (~88%)	**1/16 (~6%)**	**-**	**1/16 (~6%)**

The user tried to perform a direct call using the vocal command, but one success occurred Table6.7.

Table 6.7

Direct Call					
# Call				**Recognized**	
1	**2**	**3**	**>3**	**Yes**	**No**
9/16 (~56%)	**3/16 (~18%)**	**2/16 (~13%)**	**2/16 (~13%)**	**1/16 (~6%)**	**15/16 (~94%)**

About the fifteen failures:

- The domestic robot invited the elderly person to use the tablet to perform the task in 7 cases
- The user tried to select the communication service using the vocal command in 4 cases Table6.8-6.9
- The user completed the task using the GUI in 4 cases

Table 6.8

User selects the service		
# Call	**Recognized**	
1	**Yes**	**No**
4	3	1

Table 6.9

Did DORO talk? "Please use the tablet"	
Yes	**No**
3	-

6.3.2 Time to perform the service

After the AmI detected the gas leak, the response time of the domestic robot was by 20sec (15,67±3,05) in ~48% of cases, by 30 sec (22,89±3,02) in ~36% of cases. Table6.10

Table 6.10

Time (sec) between the user calls DORO and DORO starts						
0-9	**10-19**	**20-29**	**30-39**	**40-49**	**50-60**	**>60**

#	3 (~12%)	12 (~48%)	9 (~36%)	2 (~6%)	-	-	1 (~4%)
Mean	4±1,73	15,67±3,05	22,89±3,02	33,5±3,53	-	-	117

After DORO stopped in front the user, it talked "There is gas leak" by 10 sec (4,8±2,4) in (~71%) of cases. Table6.11

Table 6.11

Time (sec) between DORO arrives in front the user and DORO starts to speak

	0-9	**10-19**	**20-29**	**30-39**	**40-49**	**50-60**	**>60**
#	21 (~86%)	2 (~7%)	-	-	-	-	2 (~7%)
Mean	5,17±1,76	12,5±0,71	-	-	-	-	101

7 Garbage collection service results

In this scenario the elderly user wanted to dispose of garbage. So he had to imagine he could not leave his home because he was sick or the weather was not nice. Bearing in mind this presupposition the participant had to call the domestic robot in order to select and active the "Garbage collection service". This service could be performed by speech interaction or GUI use. The service was composed by several steps that are shown in Figure7.1

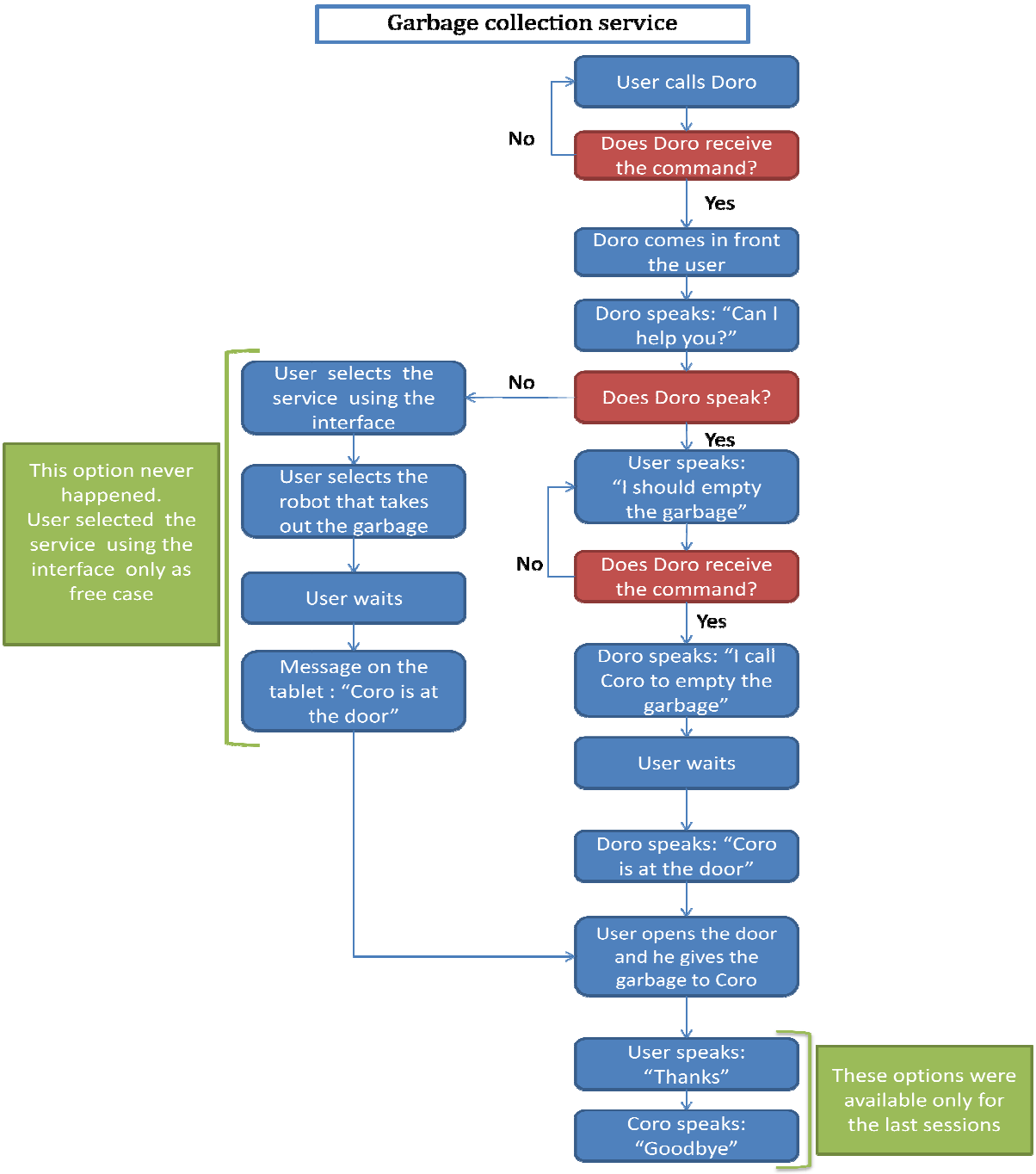
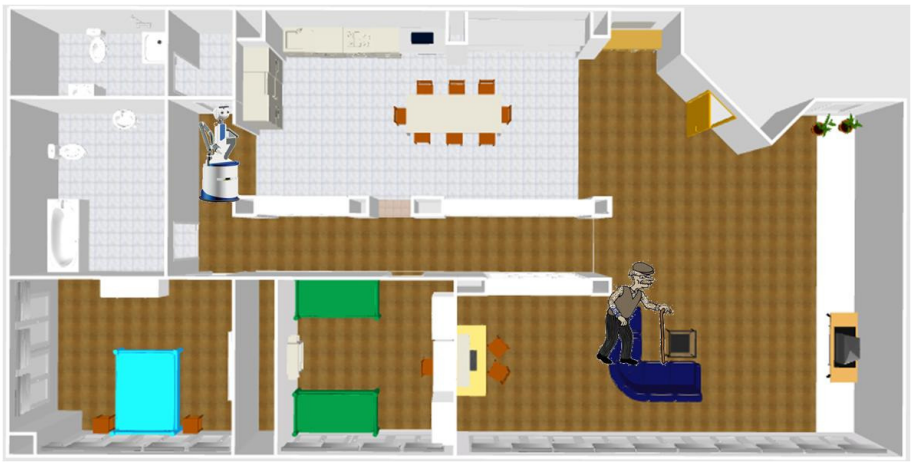


Figure 7.1 Garbage collection service flow chart

In this scenario all three Robot-Era platforms were involved working in three different environment: domestic, condominium and outdoor.



a

b

c

Figure 7.2 a: domestic environment b: condominium environment c: outdoor environment

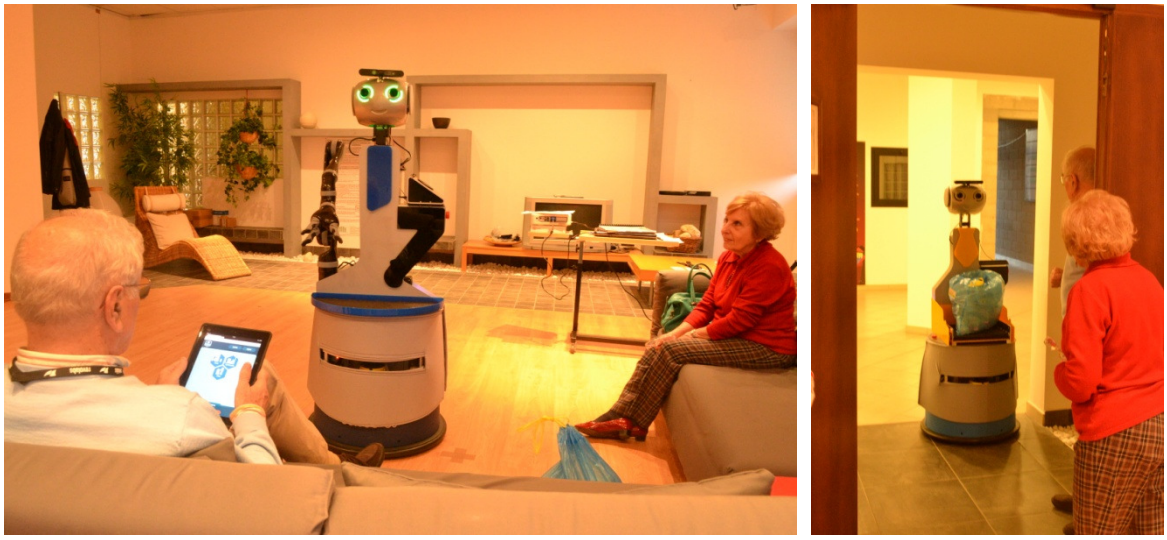


Figure 7.3 Users during the experimentation

7.1 Usability

7.1.1 Usability SUS Scale

Immediately after the end of the "Garbage collection service" , we proposed the system usability scale (SUS) referred to the service.

Table 7.1 SUS Score: descriptive statistics

Garbage Collection	mean	std	#	%
Not Acceptable 0-64	49,17	3,81	3	8,57%
Acceptable 65-84	74,17	8,04	3	8,57%
Excellent 85-100	96,03	3,98	29	82,86%
Total	90,14	14,76	35	100%

The usability of the "Garbage collection service" was evaluated as excellent (90,14±14,76) by elderly people. The actions performed by the Robot-Era platforms were perceived as well integrated (4,74±0,61) and easy to use (4,54±1,09), in fact nobody thought that he would need the support of a technical person to be able to use this garbage collection service.

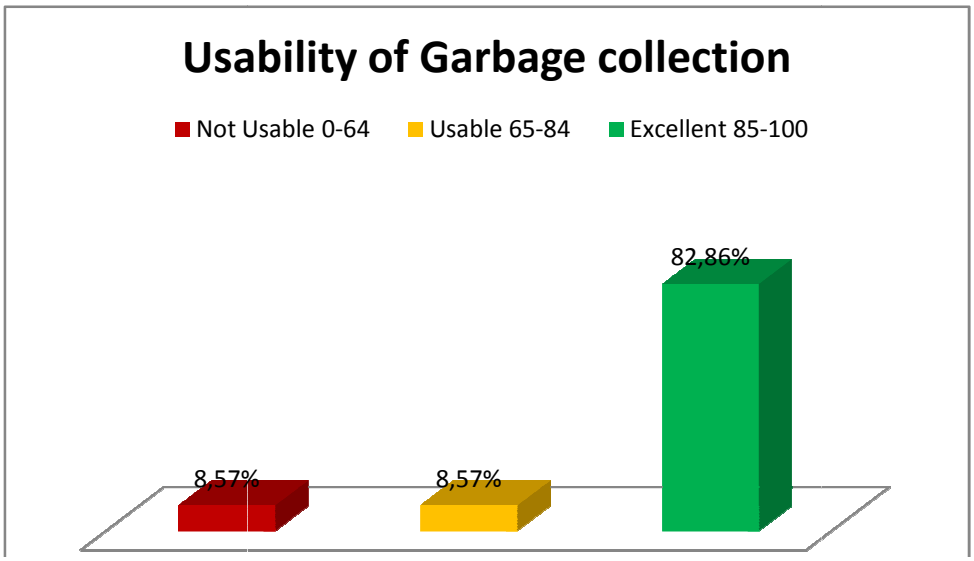


Figure 7.4 Usability result

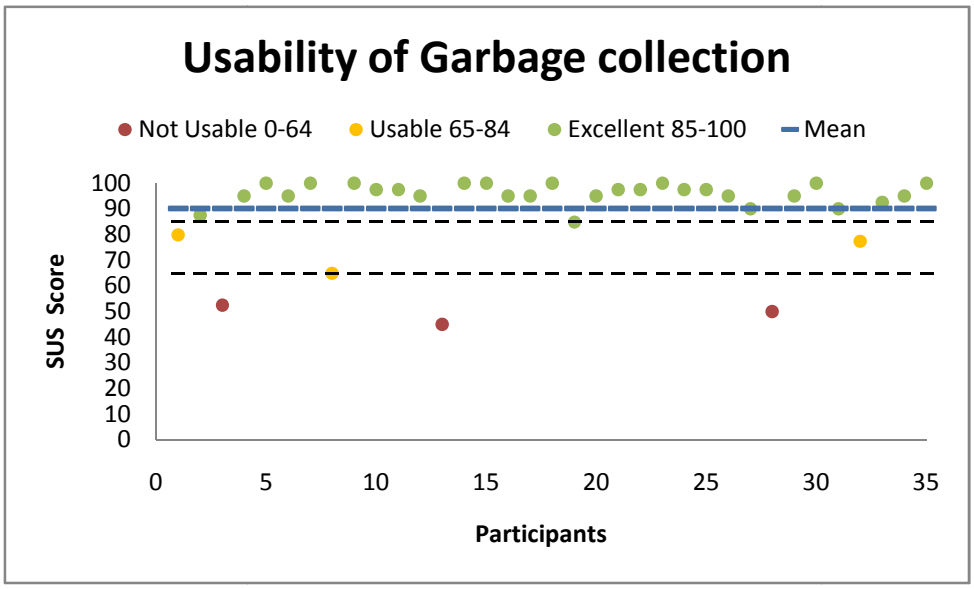


Figure 7.5 Usability result

7.1.2 Usability by video analysis

The persons didn't understand the meaning of the logos "Outdoor" "Indoor" and "Condominium". The test persons were insecure about the selection of the right icon to call the robot to the door of the flat. It was not clear for the test persons which perspective was meant (e.g." indoor" -because the person was located indoor? Or "outdoor" because he/she wanted to transport the garbage outdoor?). The test persons were confused about the perspective. Outdoor could be set into relation to the robots position or to the users position, or to the position where the garbage should be taken.

7.2 Acceptance

7.2.1 Attitude

The responses to attitude questions showed a high value (4,61±0,94) towards the "Garbage collection service". We can note that the participants would like to use this garbage collection service frequently (4,60±0,74) and obviously they would use it even in case of need (4,69±0,90) or if the family/caregiver's work could be helped (4,54±0,98). The high attitude was determined by the fact that throwing away the garbage was seen as a boring task and a burden especially during the winter season.

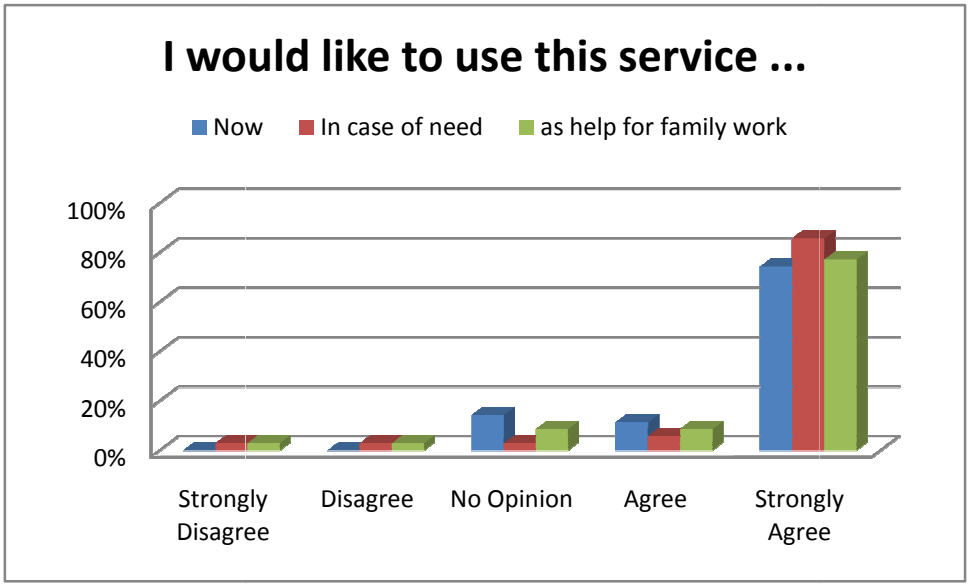


Figure 7.6

7.2.2 Acceptability

7.2.2.1 Acceptability by ad-hoc questionnaire

The questionnaire compiled by the old persons denoted a positive degree of acceptability (4,72±0,80) about the "Garbage collection service". In particular using the Robot-Era system for garbage collection didn't evoke irritability in the participants who reported they enjoyed during the experimentation and they would trust in the robot ability to perform this service.

7.2.2.2 Acceptability by video analysis

On the acceptability, many general statements of the users have confirmed the relevance of the service for supporting the users in their daily life ["I like this service", "I think that this service is very good!", "This service is interesting"]. Even if they liked the service, some users suggested improvements for more complex activities to be performed by the robot ["Can the robot differentiate the garbage?", "Can DORO takes the garbage and carry it to CORO?"]. As for the shopping case, the user requested more support to DORO for the transportation of heavy bags. It seems to be fundamental to find a technical solution for improving this transversal capability of DORO.

Additionally, this statement could suggest to improve also the connection between DORO and CORO for the second loop. No one of the users experienced anxiety during the performance.
Regarding the enjoyment, a high number of users seem to have fun in cooperating with DORO and CORO: at least twenty-two participants have shown fun from their expression and overall friendly communication.

7.2.3 HRI

7.2.3.1 HRI by questionnaire: Speech and GUI interaction

Regarding the speech interaction elderly users reported positive responses (4,38±1,01), asserting that it was easy to speak to the domestic robot to perform the service, using simple "key words".
About the GUI interaction the old persons had a positive approach (4,50±0,83), but this result was not truthful. In fact even if they said to understand what to do to perform the garbage collection, the usability video analysis demonstrated the contrary

7.2.3.2 HRI by video analysis

On the HRI, it has to be said that short duration of some observations has negatively influenced the opportunity of conducting an appropriate and complete analysis.
In general, many users have considered DORO as a nice robot and expressed appreciation on its features ["DORO is very nice", "The robot does not seem more only a machine", "DORO has a really nice head!", "DORO is joker!" "I think that the robot is fantastic"]. In one case, the robot evokes also an emotional reaction ["I fell tenderness for him"].
Even if the communication with the robot was "formal", however, many users demonstrated to have a overall good comprehension and a good management of the command to be given, as well as friendly communication. Some users declared of not understanding properly the name of the condominium robot – CORO -, when DORO advice he is at the door for taking the garbage. This problem should be fixed for the second experimental loop.

7.2.4 Quality of life

Regarding quality of life, elderly users reported positive responses (4,14±1,33) to item *"I think my independence would be improved by the use of the robot for this service"* because, as said before, throwing away the garbage is seen as a boring task and a burden especially during the winter season.

Table 7.2 Acceptance attributes: descriptive statistics

Garbage collection		mean	std	Acceptance Score	std
ACCEPTANCE	Attitude	4,61	0,94	89,97	8,33
	Acceptability	4,72	0,80		

Speech interaction	4,38	1,01
GUI interaction	4,50	0,83
QoL	4,14	1,33

For reporting the final acceptance result, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the "Garbage collection service" was not acceptable, from 65-84 was acceptable and from 85-100 the result was excellent. At last the "Garbage collection service" was very acceptable (89,97±8,33) by elderly participants. We can note that 77% of the users the acceptance as excellent (94,18±4,76).

Table 7.3 Acceptance Score: descriptive statistics

Garbage collection	mean	std	#	%
Not Acceptable 0-64	-	-	0	0%
Acceptable 65-84	79,45	5,42	8	23%
Excellent 85-100	94,18	4,76	27	77%
Total	89,97	8,33	35	100%

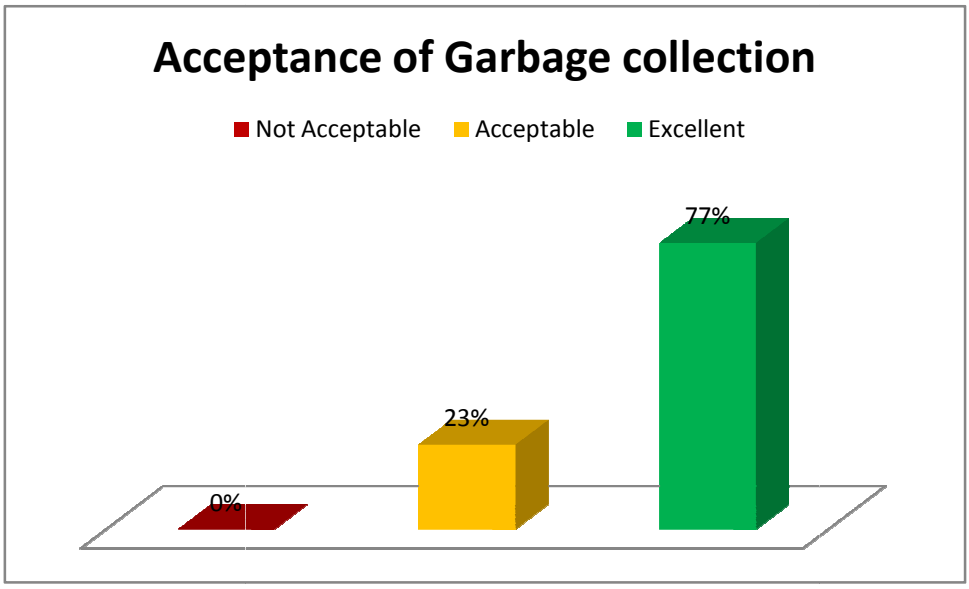


Figure 7.7 Acceptance result

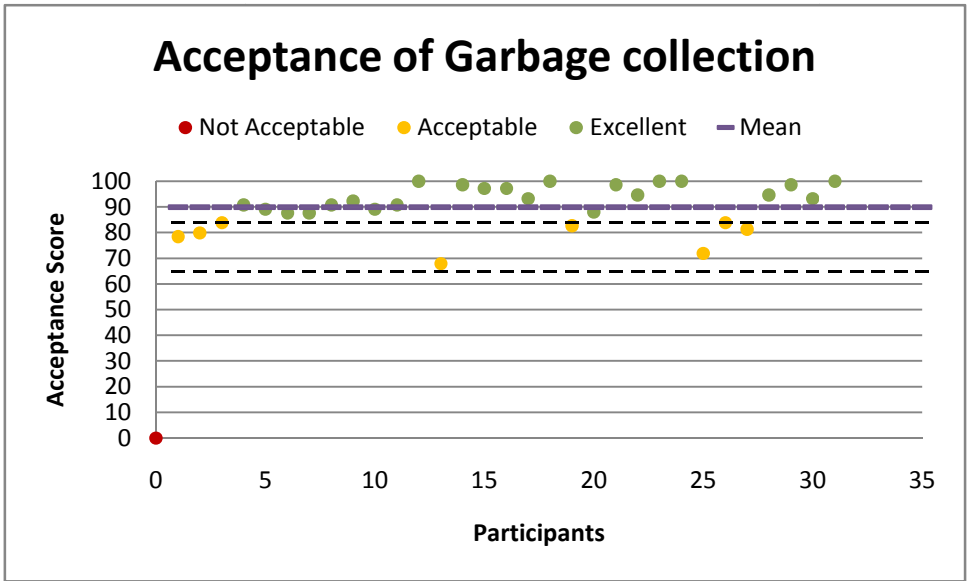


Figure 7.8 Acceptance result

7.3 Technical data

7.3.1 Docking and exchange task

About the docking and exchange task, a lot of failures occurred and the procedure needs a deep improvement: causes can be found both in the procedure, since the used docking function needs a more adaptive behavior to be successfully used in this task, and in the environment, e.g. the height misalignment between indoor and outdoor areas.

Table 7.4

	Success	Failure	No evaluated
Docking	20,83%	33,33%	45,83%
Exchange	21,74%	30,43%	47,83%

Figure 7.9 Exchange

7.3.2 Interaction with elevator

Analyzing the technical results come out during the Garbage Scenario experiments, these represent the same problem raised up in the Shopping service; furthermore, it is highlighted the navigation problem in the interaction with lift at first floor where the area in front of elevator is more narrow.

Table 7.5

	Success	Failure	No evaluated
Entering elevator at floor 0	75,00%	8,33%	16,67%
Exiting elevator at floor 1	83,33%	0,00%	16,67%
Entering elevator at floor 1	58,33%	25,00%	16,67%
Exiting elevator at floor 0	66,67%	4,17%	29,17%

7.3.3 Speech recognition

With reference to Figure 7.1, we analyze the number of success about the speech interaction. As shown in the tables when the user spoke to the domestic robot, it received the command at the first the time in the majority of cases. Also DORO gave a notification to the elderly person without problem.

Table 7.6 Speech interaction

User Calls DORO		
# Call	**Recognized**	
1	**Yes**	**No**
31/31 (100%)	**31/31 (100%)**	**-**

Did DORO talk? "Can I help you?"	
Yes	**No**
31/31 (100%)	-

User selects the service				
# Call			**Recognized**	
1	**2**	**3**	**Yes**	**No**
24/31 (~77%)	5/31 (~16%)	2/31 (~6%)	30/31 (97%)	1/31 (3%)

Did DORO talk? "I call CORO"	
Yes	**No**
29/33 (~94%)	2/33 (~6%)

Did DORO talk? "CORO is at the door"	
Yes	**No**
31/31 (100%)	-

7.3.4 Time to perform the service

After DORO was called by the user, the response time of the domestic robot was by 10 sec (4,9±2,6) in ~55% of cases, by 20 sec (12,8±3,3) in ~18% of cases and by 30 sec (23,4±3,8) in ~15% of cases. Table7.7

Table 7.7

Time (sec) between the user calls DORO and DORO starts

	0-9	**9-19**	**19-29**	**29-39**	**39-49**	**49-60**	**>60**	**No evaluated**
#	18/33 (~55%)	6/33 (~18%)	5/33 (~15%)	1/33 (~3%)	-	1/33 (~3%)	2/33 (~6%)	2/35
mean	4,9±2,6	12,8±3,3	23,4±3,8	33	-	59	80±21,2	

After DORO stopped in front the user, it talked "Can I help you?" by 10 sec (4,8±2,4) in (~71%) of cases. Table7.8

Table 7.8

Time (sec) between DORO arrives in front the user and DORO starts to speak

	0-10	***10-20***	***20-30***	***30-40***	***39-49***	***49-60***	***>60***	**No evaluated**
#	*20/28 (~71%)*	*2/28 (~7%)*	*1/28 (~4%)*	*1/28 (~4%)*	*-*	*1/28 (~4%)*	*3/28 (~11%)*	*7/35*
mean	*4,8±2,4*	*19,5±0,7*	*21*	*34*	*-*	*52*	*248,7±155*	

The "Garbage collection service" was performed by 9 minutes (*7,15±0,50*)in the 54% of cases and by 12 minutes (*10,44±0,42*)in the 21% of cases. Table7.9

Table 7.9

Time (min) to perform the service

	0-3	***3-6***	***6-9***	***9-12***	***12-15***	***>15***	***No evaluated***
#	*-*	*5/33 (~15%)*	*19/33 (~58%)*	*7/33 (~21%)*	*1/33 (~3%)*	*1/33 (~3%)*	*2/35*
mean	*-*	*5,19±0,18*	*7,15±0,50*	*10,44±0,42*	*13,21*	*16,36*	

8 Reminding service results

In this scenario the elderly user wanted to set a date on Robot-Era agenda. The user called the domestic robot to perform the task, then he moved to another room where he localized by AmI and the robot got through to him to remember the date. The speech interaction and interface interaction were necessary to perform this service. This service was composed by several steps that are shown in Figure8.1

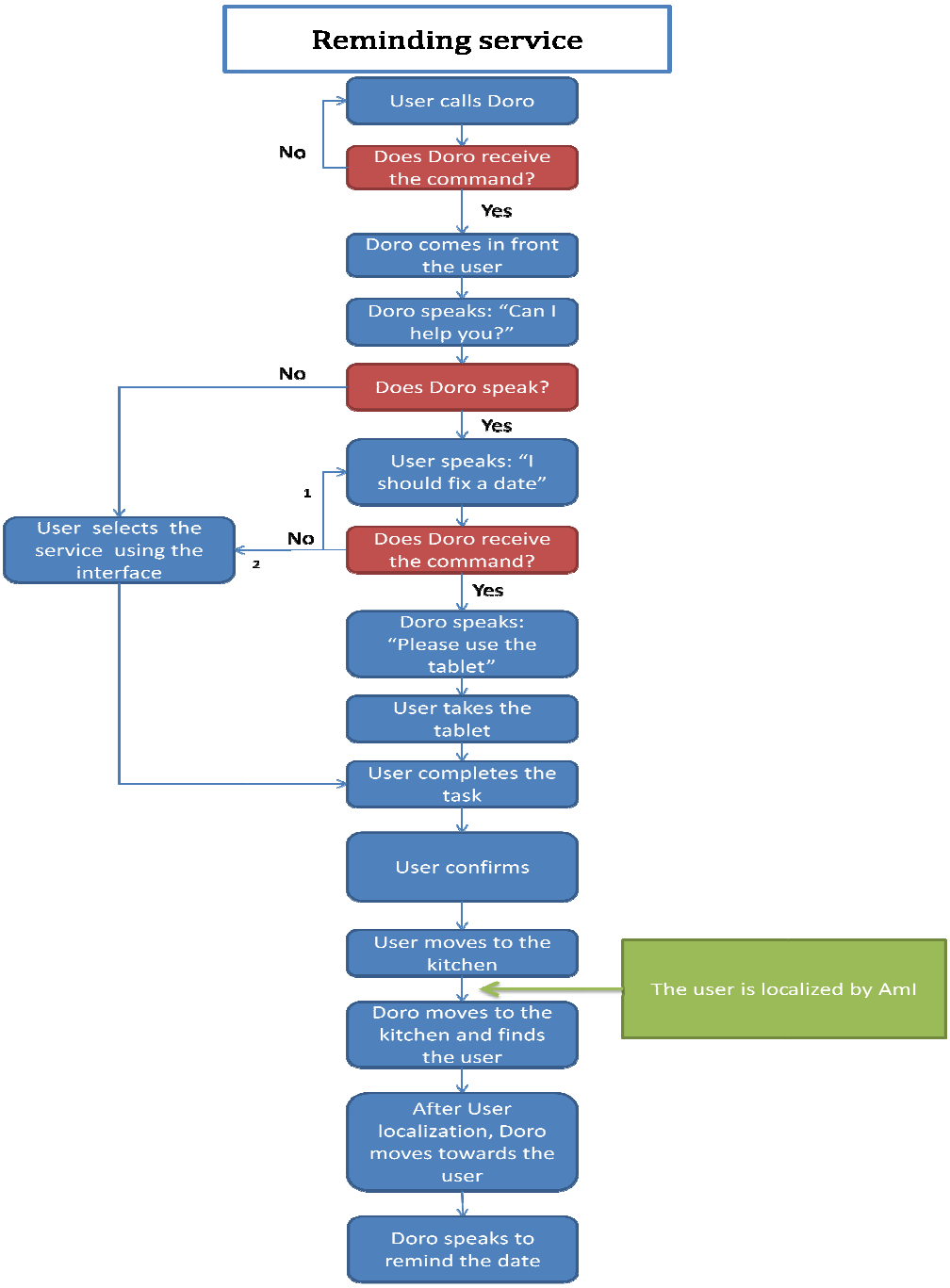


Figure 8.1 Reminding service flow chart

In this scenario only DORO was involved working in the domestic environment.

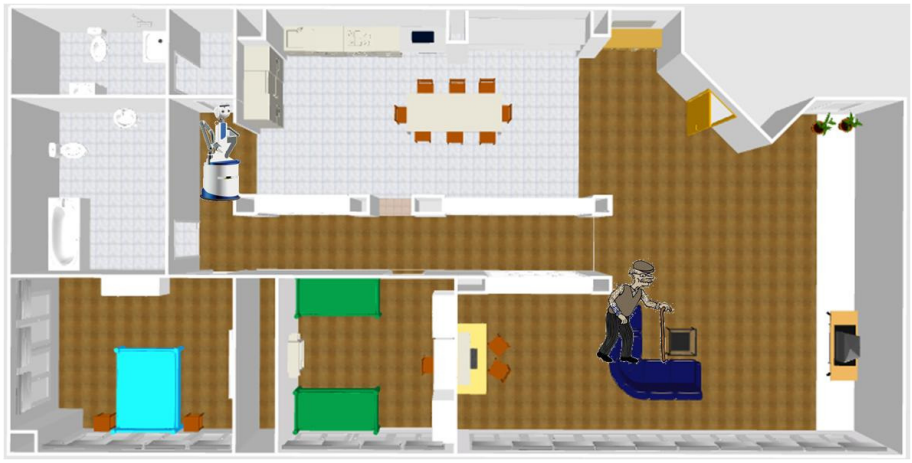


Figure 8.2 domestic environment

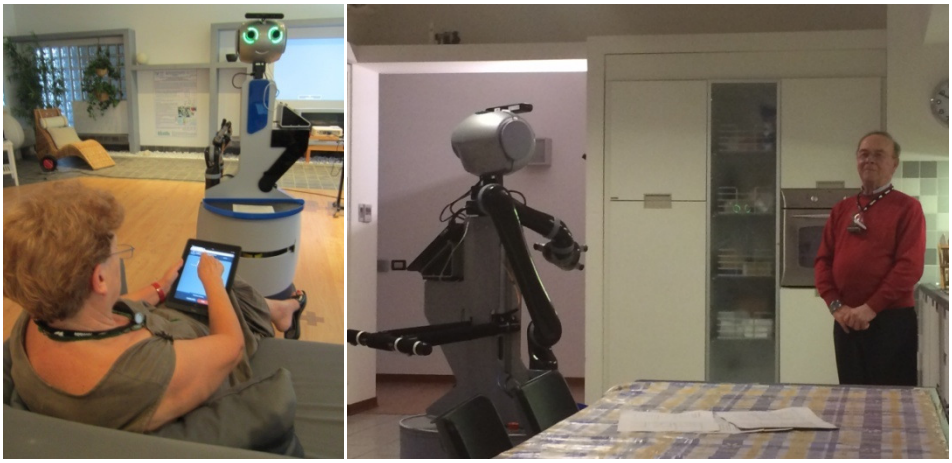


Figure 8.3 Users during experimentation

8.1 Usability

8.1.1 Usability by SUS Scale

In order to evaluate the usability of "Reminding service" we proposed the system usability scale (SUS), a simple ten item Likert scale that gives a global view of subjective assessments of usability.

Table 8.1 SUS Score: descriptive statistics

Reminding	mean	std	#	%
Not Usable 0-64	40,83	14,72	6	22,22%
Usable 65-84	70,56	3,49	9	33,33%
Excellent 85-100	96,04	4,45	12	44,44%
Total	75,28	23,12	27	100%

The "Reminding service" was perceived as medium usable (75,28±23,12) by elderly people. Even if the actions performed by the domestic platform and AmI were evaluated well integrated (4,59±0,89), many participants had problem using the GUI to set a date so the easy use of the service was medium estimated (3,78±1,40), reducing the final SUS score. We can note that the "Reminding service" was not usable for 22% of older persons.

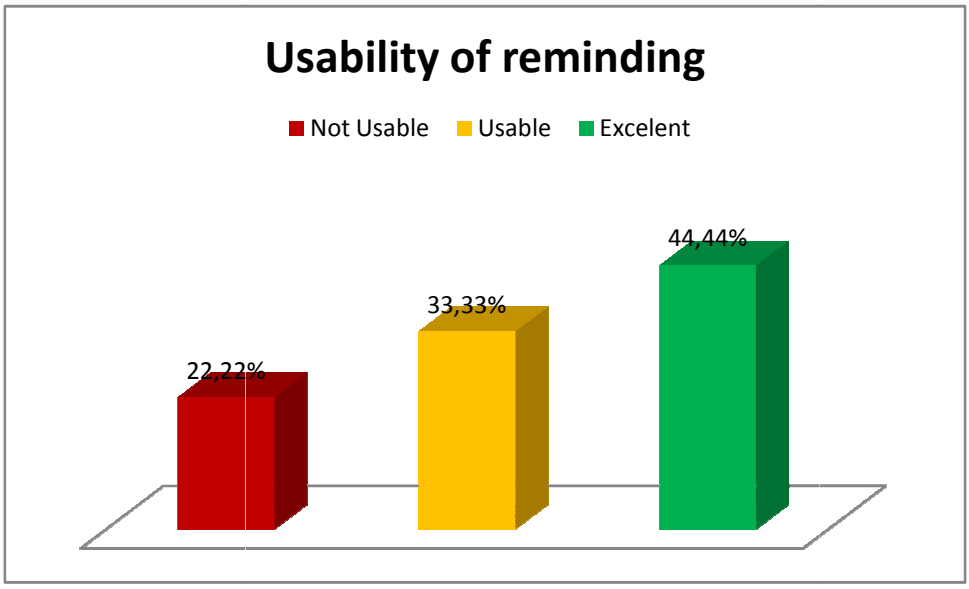


Figure 8.4 Usability result

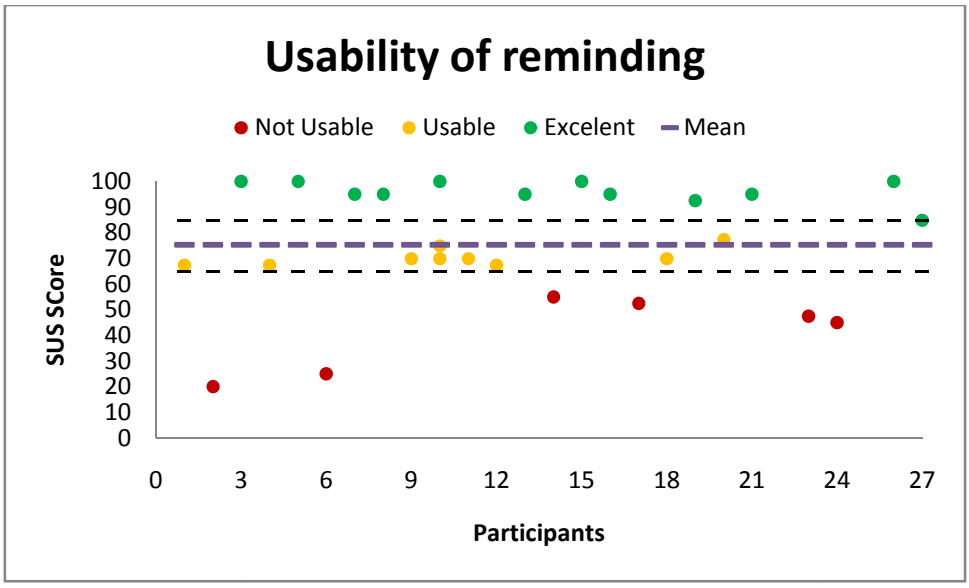


Figure 8.5 Usability result

8.1.2 Usability by video analysis

The title of the reminder is presented by a drop-down menu, where the other options are hidden. The standard title of the reminder is "Medicine". The other options "Telephone" and "Generic Alert" are only shown when the person opens it. The person does not expect that there are other options too (for example "Telephone").

In real life the option might be overlooked so that the person thinks that it is just a reminder for medicine or never change the title and use instead the description field.

8.2 Acceptance

8.2.1 Attitude

The elderly user demonstrated a high attitude (4,55±1,09) towards this service because they perceived it very useful. At the present the attitude was positive enough (3,85±1,51), (4,44±1,19) in case of need and (4,37±1,33) as help for family work.

8.2.2 Acceptability

8.2.2.1 Acceptability by ad-hoc questionnaire

The questionnaire compiled by the old persons denoted a positive degree of acceptability (4,51±1,06) about the "Reminding". In particular using the domestic robot for this service didn't evoke anxiety in the participants who reported that being localized inside home by robot was not too much invasive for their lifestyle.

8.2.2.2 Acceptability by video analysis

The principal mean of interaction for this service is constituted by the tablet. From the observation, it was found that just the users with more skills with technology are able to perform autonomously this task – nine participants –. This fact can influence the acceptability of the service, that sometimes is felt as too complicated for the users ["The menu really does not work for me!", "I don't know what to do!", "I don't understand this – tablet -","It's too difficult for me!", "I don't like this service: it is too complicate, not easy like garbage for example. You should do too many things for reminding something!" "I am not expert in technology"]. This issue needs to be more investigated through the comparison of the data on dependability and usability of the service, because it is not clear if the difficulty is due to technical aspects to improve. Anyway, from the observation can be understood that the majority of the sample have looked for the step by step support from the experimenter – in particular for fourteen users -. For this reason, it is not possible to evaluate the overall satisfaction with this service, also due technical failures that obliged a restart of the sessions and, consequently, the distraction from the interaction.
Anyway, the more skilled users have given input on what can be done to ameliorate the service satisfaction ["Aesthetically, it should be more similar of the shopping list, so you can choose what to remind trough icons"] and the robot communication capability for this service ["Why DORO does not say more things? For example, he can remind me also who I have to call to, not just that I should call"].

8.2.3 HRI

8.2.3.1 HRI by questionnaire: Speech and GUI interaction

Regarding the speech interaction elderly users reported positive responses (4,15±0,94), asserting that it was easy to speak to the domestic robot to perform the service, using simple "key words".
About the GUI interaction the old persons had a positive approach (4,09±1,24), but this result was not truthful. In fact even if they said to understand what to do to perform the garbage collection, the usability video analysis demonstrated the contrary.

8.2.3.2 HRI by video analysis

Even if they were more concentrated on the tablet for the task execution, many users have a general good HRI, as shown by the adoption of a very friendly communication directed to DORO ["DORO attention!", "Many thanks DORO!", "I really need you DORO!" "DORO, you are perfect. This service is very useful"] and free statements on its features, especially the eyes ["DORO has a really "intelligent" eyes", "DORO has a funny eyes", "The face is really funny! But he stops too distant to me!", "DORO is really fantastic"]. In addition, an affective reaction to the robot was detected in one user ["It seems a person, it is a companion!"].

8.2.4 Quality of life

Regarding quality of life, elderly users reported positive responses (4,19±1,27) to item *"I think my independence would be improved by the use of the robot for this service"* because they evaluated the system an efficient way to remember dates.

Table 8.2 Acceptance attribute: descriptive statistics

Shop and Drug delivery		mean	std	Acceptance Score	std
ACCEPTANCE	Attitude	4,55	1,09	87,26	12,32
	Acceptability	4,51	1,06		
	Speech interaction	4,15	0,94		
	GUI interaction	4,09	1,24		
	QoL	4,19	1,27		

For reporting the final acceptance result, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the service was not acceptable, from 65-84 was acceptable and from 85-100 the result was excellent. At last the "Reminding service" was acceptable (87,26±12,32) by elderly participants. However we can note that 74% of the users evaluated excellent the "Reminding service".

Table 8.3 Acceptance Score : descriptive statistics

Reminding	mean	std	#	%
Not Acceptable 0-64	52,00	-	1	4%
Acceptable 65-84	72,44	6,18	6	22%
Excellent 85-100	93,47	4,72	20	74%
Total	87,26	12,32	27	100%

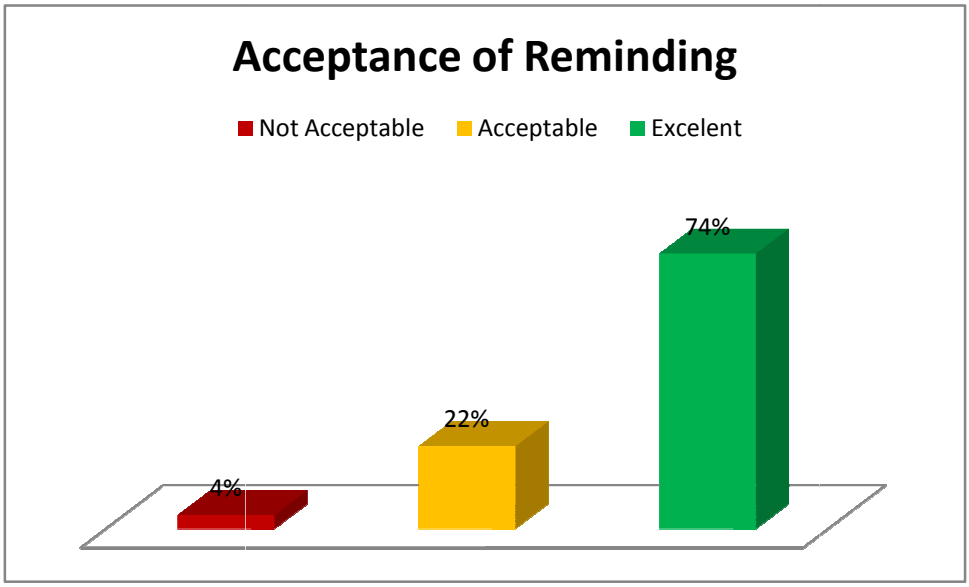


Figure 8.6 Acceptance result

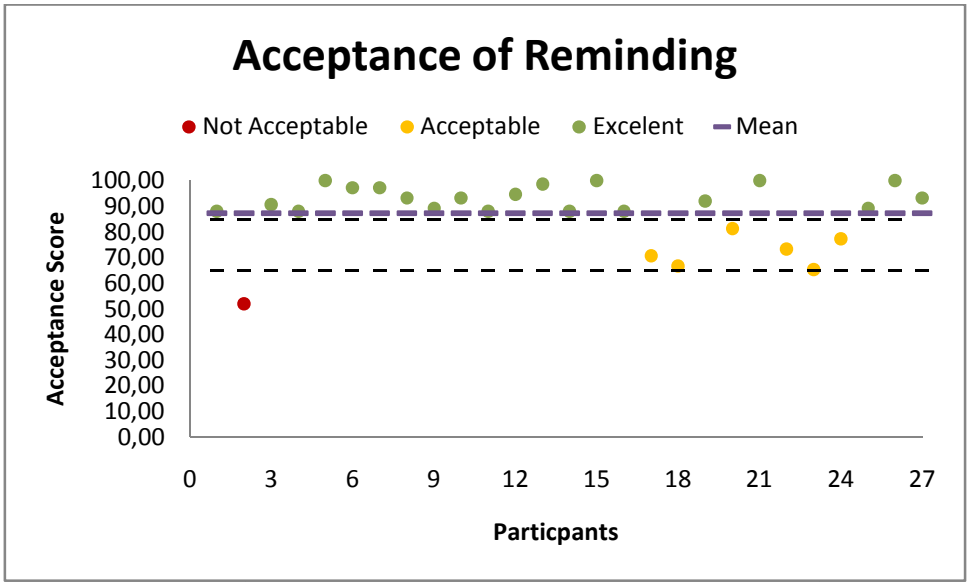


Figure 8.7 acceptance result

8.3 Technical data

8.3.1 Speech recognition

With reference to Figure 8.1, we analyze the number of success about the speech interaction. As shown in the tables when the user spoke to the domestic robot, it received the command at the first the time in the majority of cases. Also DORO gave a notification to the elderly person without problem.

Table 8.4 Speech interaction

User Calls DORO					
# Call				Recognized	
1	2	3	>3	Yes	No
21/27 (~78%)	2/27 (~7%)	1/27 (~4%)	3/27 (~11%)	27/27 (~100%)	

Did DORO talk? "Can I help you?"	
Yes	No
27/27 (~100%)	-

User selects the service						
# Call				Recognized		Other Service actived
1	2	3	>3	Yes	No	
23/27 (~85%)	3/27 (~11%)	1/27 (~4%)	-	20/27 (~74%)		7/27 (~26%)

Did DORO talk? "Please use the tablet"	
Yes	No
27/27 (~100%)	-

8.3.2 Time to perform the service

After DORO was called by the user, the response time of the domestic robot was by 10 sec (8±2) in ~70% of cases and by 20 sec (12±3) in ~30% of cases. Table8.5

Table 8.5

Time (sec) between the user calls DORO and DORO starts							
	0-9	10-19	20-29	30-39	40-49	50-60	>60
#	19/26(~70%)	7/26(~30%)	-	-	-	-	-
Mean	8±2	12±3	-	-	-	-	-

After the user selected the service, the domestic robot talked "Please, use the table to complete your shopping list" by 10 sec (8±1) in ~38% of cases and by 20 sec (13±3) in ~42% of cases. Table8.6

Table 8.6

Time (sec) between service selection and DORO starts to speak

	0-9	**10-19**	**20-29**	**30-39**	**40-49**	**50-60**	**>60**	No evaluated
#	9 /24 (~38%)	10/24 (~42%)	2/24 (~8%)	1/24 (~4%)	1/24 (~4%)	-	1/24 (~4%)	3/27
mean	8±1	13±3	24±6	30	45	-	94	

8.3.3 User localization

The localization accuracy was computed as the Manhattan distance between the micro area where the user was located by the localization system, and the micro area where he/she was effectively standing. The kitchen was divided in ten micro-areas of about 2 m^2 each Figure 8.8 and localization errors were provided in number of micro areas Table 8.7.

Figure 8.8 The kitchen map

Table 8.7

User localization error in number of micro areas			
0 micro area	**1 micro area**	**2 micro areas**	**3 micro areas**
7	7	8	5
~25,9%	~25,9%	~29,6%	~18,5%

In particular, the user was located in the same micro-area where he/she was actually standing, in the ~26% of the trials, positively contributing to the service usability and acceptability. In the ~26% of the trials the user was located within 1 micro-area respect to the current position, ~30% within 2 micro-areas and ~19% within 3 micro-areas Table 8.7.

Regarding the error in number of micro areas between the user was localized and where the robot stopped in the ~35% of the trials not error occurred, ~39% within 1 micro-areas and ~26% within 1 micro-areas Table 8.8.

Table 8.8

error in number of micro areas between the user was localized and where the robot stopped		
0 micro area	**1 micro area**	**2 micro area**
8	9	6
~34,8%	~39,1%	~26,1%

9 Indoor walking support service results

In this scenario the elderly user had to imagine he had some temporary motor problem so he used the domestic robot as a walking support. This service consisted in the steps shown in Figure9.1

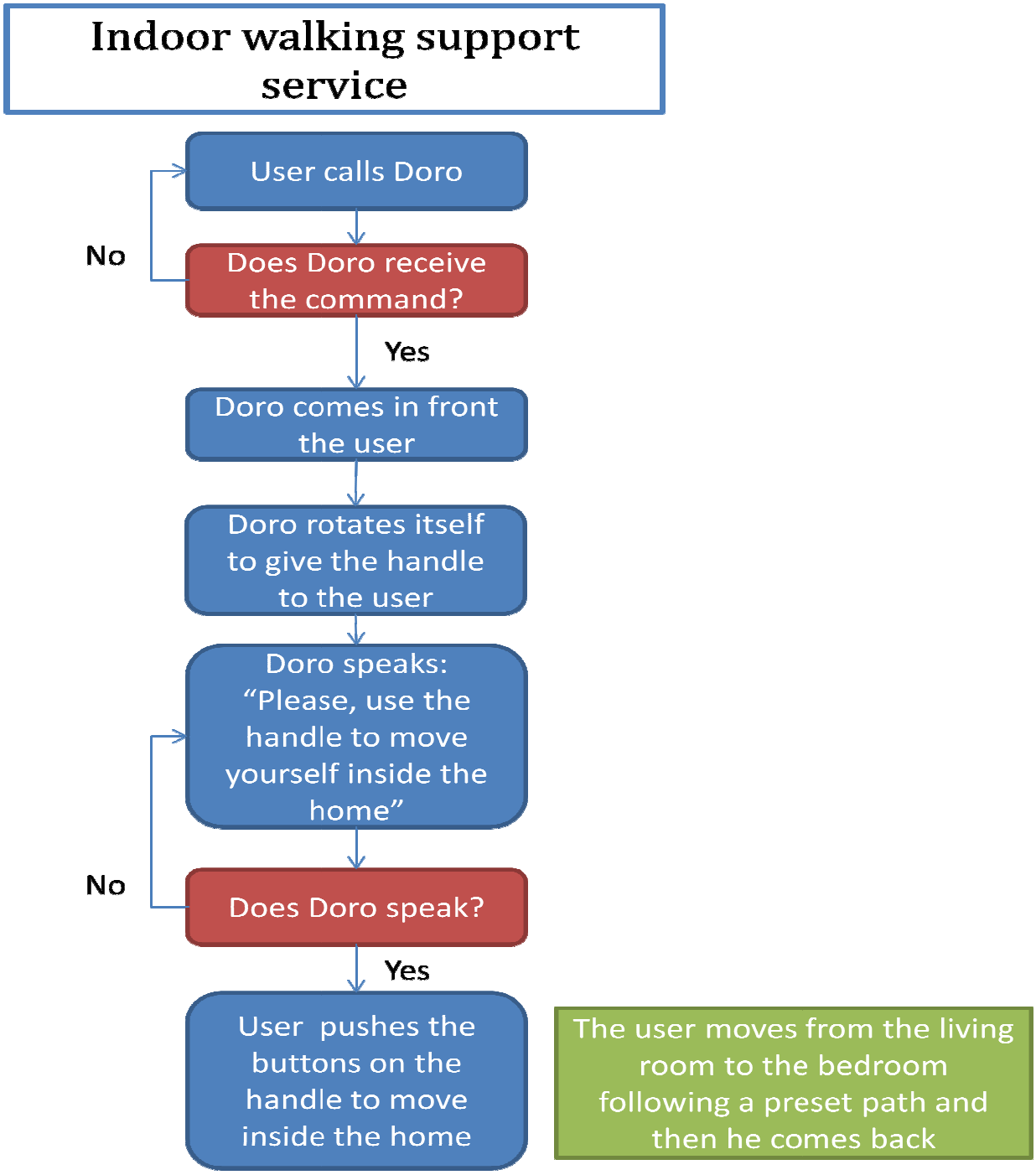


Figure 9.1 Indoor walking support service flow chart

In this scenario only DORO was involved working in the domestic environment.

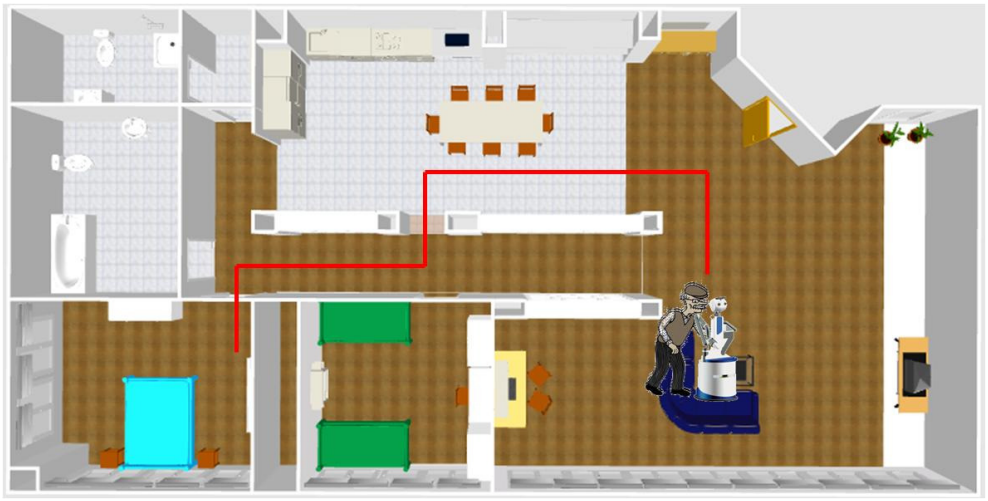


Figure 9.2 Domestic environment

Figure 9.3 User during experimentation

9.1 Usability

9.1.1 Usability SUS Scale

Immediately after the end of the "Indoor walking support service" , we proposed the system usability scale (SUS) referred to the service.

Table 9.1SUS Score: descriptive statistics

Indoor walking support	mean	std	#	%
Not Usable 0-64	-	-	0	0%
Usable 65-84	78,75	2,62	6	22%
Excellent 85-100	91,55	5,09	21	78%
Total	88,70	7,12	27	100%

The usability of the "Indoor walking support service" was evaluated as excellent (88,70±7,12) by elderly people. The actions performed by the domestic platform was perceived as well integrated (4,70±0,47) and easy to use (4,74±0,59), in fact nobody thought that he would need the support of a technical person to be able to use this service.

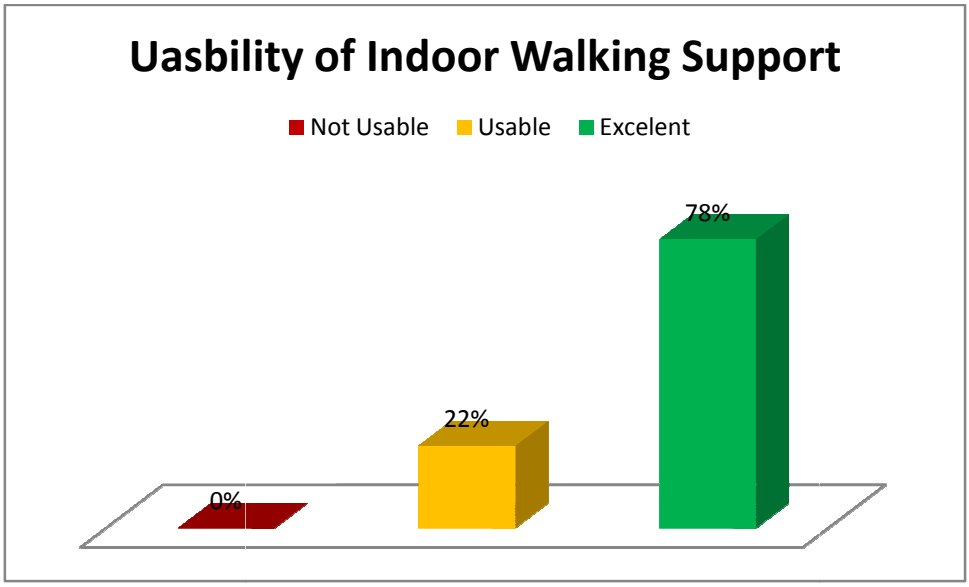


Figure 9.4 Usability result

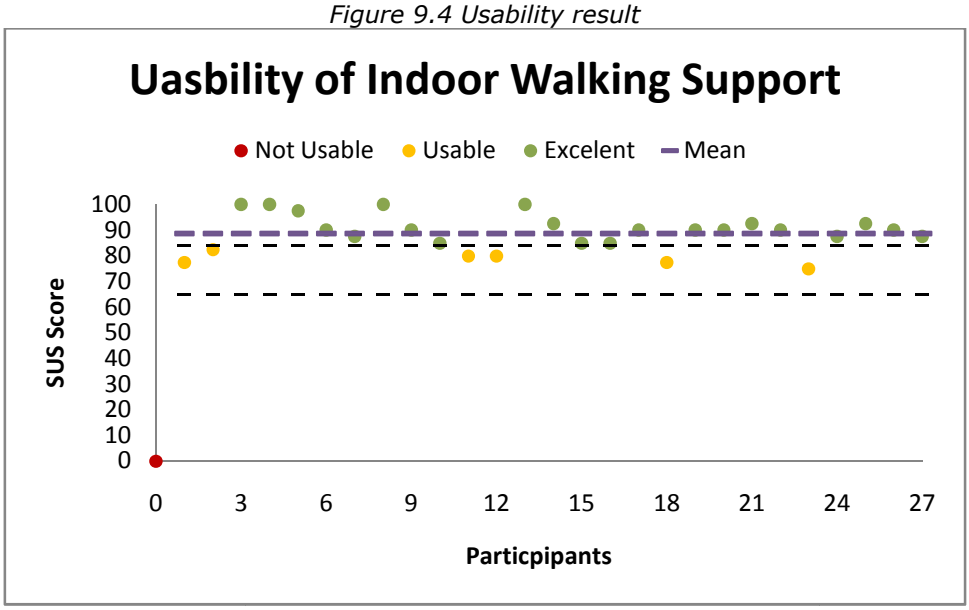


Figure 9.5 Usability result

9.1.2 Usability video observation

DORO does not calculate that the person next to it needs space to turn around (especially when there is a wall next to DORO). The test person does not know what to do, when she/he has not enough space between DORO and a wall to turn around. The person has to leave DORO or she/he is pushed against the wall. If an old person has problems to walk, the person has to trust that DORO will guide her/him safely to the bedroom

9.2 Acceptance

9.2.1 Attitude

The responses to attitude questions showed a high value (4,46±1,18) towards the "Indoor walking support service". We have to note that this positive results was due thinking a future use; (4,48±1,16) in case of need and (4,44±1,17) as help for family work. At the present the attitude was low (2,48±1,78) because the participants were autonomous for their daily activities and they saw this service as an assistance help.

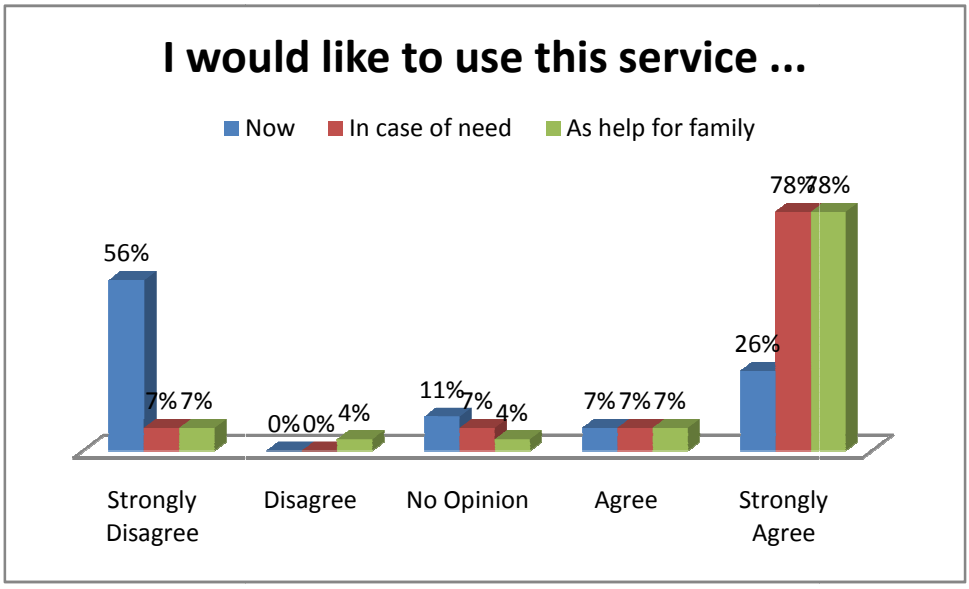


Figure 9.6

9.2.2 Acceptability

9.2.2.1 Acceptability by ad-hoc questionnaire

The questionnaire compiled by the old persons denoted a positive degree of acceptability (4,63±0,86) about the "Indoor walking support service". In particular using the domestic robot as a walking support didn't evoke anxiety in the participants who reported they enjoyed during the experimentation. However we have to consider that elderly persons performed the task as a "game" because they drove DORO from a room to another one.

9.2.2.2 Acceptability by video analysis

On the acceptability side, just two users seemed to be a little worried for the performance: in the first case, the user asked many further information on the task to be performed, looking around nervously, while in the second case, the user has stopped walking many times, stated of having the fear of damaging DORO. It has to be said, that the two users experienced more anxiety also during the other services performance, so it could be possible to hypothesize that the fear of not being able to use the robot for moving indoor is caused by a low attitude/personal experience with technology and not on the service itself. Finally, no one experienced fear of falling during the walking, giving some positive input on

the gait velocity ["It seems faster than ORO"], on the overall experience in walking with DORO ["This can be a nice service", "It's very easy to walk and to give command to DORO for walking!","How it moves fluently! - DORO", "The robot is nice!", "It's really nice!"] and most of all on the easiness to control DORO ["I'm really comfortable in moving with DORO!"]. On the enjoyment, it was noticed that the users were mainly concentrated on the task and on the physical interaction for this service, even if eight of them expressed a friendly predisposition to the robot, for example moving around DORO even at the end of the task or laughing during the experience.
The more critical aspect emerged from the observation regards the turning task. It was observed that mainly the female users have faced the same error in turning: when it was asked them to turn inside a smaller room, they usually have difficulty in organizing spatially the task, so, for example they have to move for doing the task. This problem can be related to the impact on the home environment issue, that is a well-known factor that sometimes can prejudice the adoption of the robot at home, if not approached properly ["You should have a bigger apartment than mine for moving with DORO!", "Sometimes is too small – the rooms - for turning"]. An alternative solution was given by one user, that is the implementation of a navigation function to go backward ["Can the robot go backward?"].
Finally, the execution of a double task for turning with DORO – press the button and move DORO – can be a little confounding, as observed in a small number of users.

9.2.3 HRI

9.2.3.1 HRI by questionnaire: Speech and GUI interaction

In this service the speech and GUI interaction were not evaluated because elderly users called the domestic without other interaction form.

9.2.3.2 HRI by video analysis

Concerning the HRI, the analysis was made mainly on the proxemics, communication/understanding and the difficulty in moving with the robot – partially reported also for acceptability -.
Regarding proxemics, it was observed that almost all the users have approached firmly to the robot and have taken the controller without hesitation. Just in two cases, the users were uncertain. This result can be quite important, thinking on the fact that the dexterity capability is a critical ability for the older people, that tends to worsening very soon, so having a good ergonomic for the controller is mandatory. In addition, some users reported that the controller should be customizable, taking into account the height of the person – for maintaining a comfortable posture -, and also the length of the step, to avoid hurts with the robot ["It's good, but the controller should be customizable, also because it influenced the step length not only the uncomfortable posture", "It has to fix the problem of the length of the step: if you touch DORO, he stops!", "It's too sensitive! I've just touch it slightly and he stops!", "The controller is very comfortable!"]. On the interaction/communication side, two users have expressed interest in the eyes of the robot: one user has recognized by himself the functionality of the eyes, at the beginning, while another one has touched the eyes after the testing session ["I really like his eyes!"]. This can suggest the importance of improving the communication modalities, most of all if the robot is the unique interface that

can be used for the service. Improvements should be made on the face of the robot, focusing on the eyes, for example, as they have captured the attention and gaze of almost all the sample, at the beginning of the service.

9.2.4 Quality of life

Regarding quality of life, elderly users reported neutral responses (3,41±1,55) to item *"I think my independence would be improved by the use of the robot for this service"* because, as said before, the participants were autonomous for their daily activities and they saw this service as an assistance help.

Table 9.2 Acceptance attributes: descriptive statistics

Indoor walking support		mean	std	Acceptance Score	std
ACCEPTANCE	Attitude	4,46	1,18	89,70	9,29
	Acceptability	4,63	0,86		
	QoL	3,41	1,55		

For reporting the final acceptance result, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the service was not acceptable, from 65-84 was acceptable and from 85-100 the result was excellent. At last the "Indoor walking support service" was acceptable (89,70±9,29) by elderly participants. We can note that 78% of the users the acceptance as excellent (94,18±4,76).

Table 9.3 Acceptance Score : descriptive statistics

Indoor walking support	**mean**	**std**	**#**	**%**
Not Acceptable 0-64	60	-	1	3,70%
Acceptable 65-84	79,20	5,02	5	18,52%
Excellent 85-100	93,62	4,41	21	77,78%
Total	89,70	9,29	27	100%

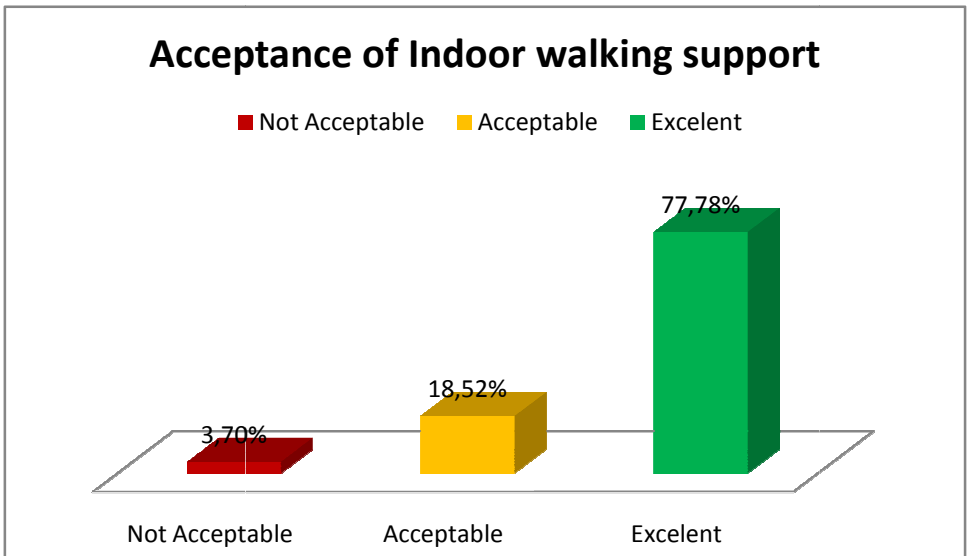


Figure 9.7Acceptance result

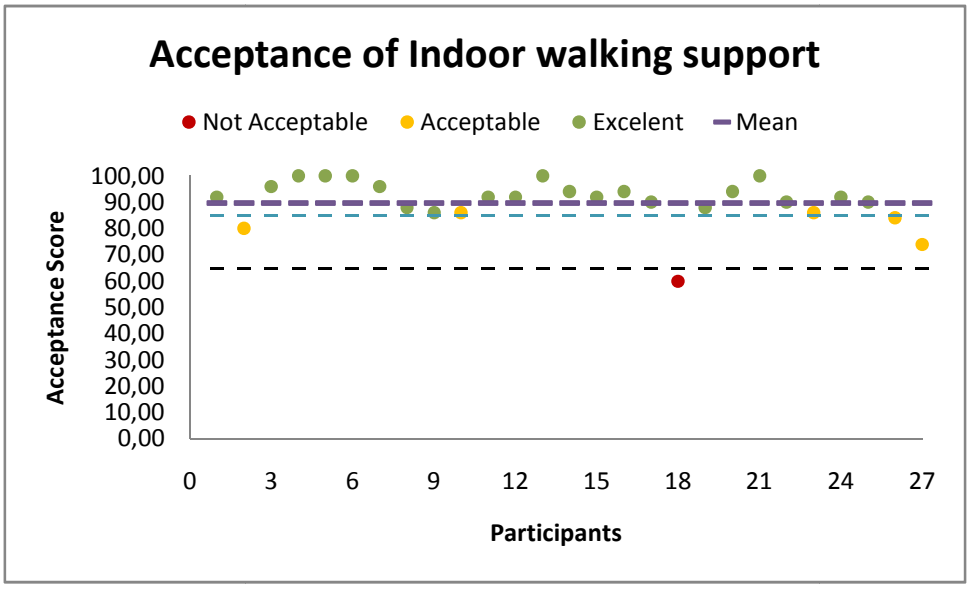


Figure 9.8 Acceptance result

9.3 Technical data

9.3.1 Speech recognition

With reference to Figure 9.1, we analyze the number of success about the speech interaction. As shown in the tables when the user called the domestic robot, it received the command at the first the time in the majority of cases. Also DORO gave a notification to the elderly person without problem.

Table 9.4 Speech interaction

User Calls DORO					
# Call				Recognized	
1	2	3	>3	Yes	No
22/27 (~81%)	4/27 (~15%)	-	1/27 (~4%)	27/27 (100%)	-

Did DORO talk? "Please Use the handle?"	
Yes	No
25/27 (~92%)	2/27 (~8%)

9.3.2 Time to perform the service

After the AmI detected the gas leak, the response time of the domestic robot was by 10sec (8±1) in ~78% of cases, by 20 sec (12±3) in ~36% of cases. Table9.5

Table 9.5

Time (sec) between the user calls DORO and DORO starts							
	0-9	10-19	20-29	30-39	40-49	50-60	>60
#	21/27 (~78%)	6/27 (~22%)	-	-	-	-	-
Mean	8±1	12±3	-	-	-	-	-

In the following tables9.6-9.7, the time to perform the task.

Table 9.6

Time (sec) for going from the living room to the bedroom							
	0-30	30-60	60-90	90-120	120-150	>150	*No evaluated*
#	-	1/25 (~4%)	12/25 (~48%)	5/25 (~20%)	6/25 (~24%)	1/25 (~4%)	2/27
mean	-	60	78±9	99±9	131±9	177	

Table 9.7

Time (sec) for going from the bedroom to the living room							
	0-30	30-60	60-90	90-120	120-150	>150	No evaluated
#	-	4/24 (~17%)	7/24 (~29%)	8/24 (~33%)	2/24 (~8%)	3/24 (~13%)	3/27
mean	-	54±7	76±9	104±11	136±6	240±108	

10 Outdoor walking support service results

In this scenario the user moves from the point A to point B following a preset path and then he comes back.
The user uses the joystick to drive the robot and he tries to open and close the robot bin, pushing the icon on the screen. In this scenario only ORO was involved working in the outdoor environment.

Figure 10.1 Outdoor environment

Figure 10.2 Users during experimentation

10.1 Usability

10.1.1 Usability by SUS Scale

Immediately after the end of the "Outdoor walking support service" , we proposed the system usability scale (SUS) referred to the service.

Table 10.1 SUS Score: descriptive statistics

Outdoor walking support	mean	std	#	%
Not Usable 0-64	57,50	3,54	2	8%
Usable 65-84	75,28	4,41	9	38%
Excellent 85-100	93,08	4,80	13	54%
Total	83,44	12,51	24	100%

The usability of the "Outdoor walking support service" was evaluated as usable (88,70±7,12) by elderly people. The actions performed by the domestic platform was perceived as well integrated (4,04±1,23) and easy to use (4,33±0,92without the support of a technical person to be able to use this service.

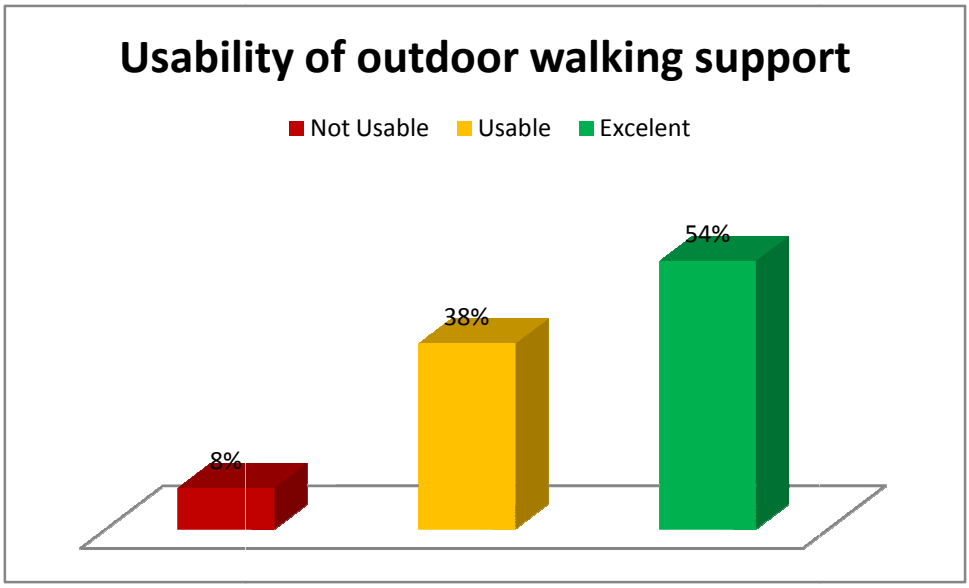


Figure 10.3 Usability result

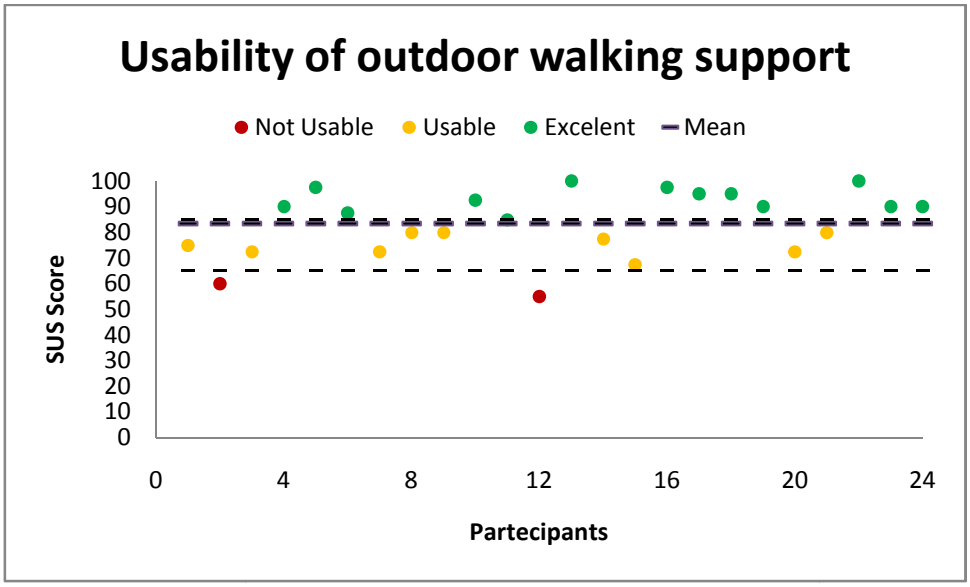


Figure 10.4 Usability result

10.1.2Usability by video analysis

The joystick might be too sensitive and does not give feedback in which position it is. The test person has to hold the joystick straight on, but instead she/he pushes it to the left or right and so it becomes difficult to navigate the robot straight on.

When the test person wants to talk and to navigate the robot at the same time, she/he needs to split her/his concentration at this moment and it becomes more difficult to navigate the robot. In addition the elderly might not have understood how to navigate via joystick because they are not used to it in daily life.

10.2 Acceptance

10.2.1Attitude

The responses to attitude questions showed a high value (4,33±1,36) towards the "Outdoor walking support service". We have to note that this positive results was due thinking a future use; (4,33±1,36) in case of need and (4,29±1,32) as help for family work. At the present the attitude was neutral (3,29±1,82) because the participants were autonomous for their daily activities and they saw this service as an assistance help.

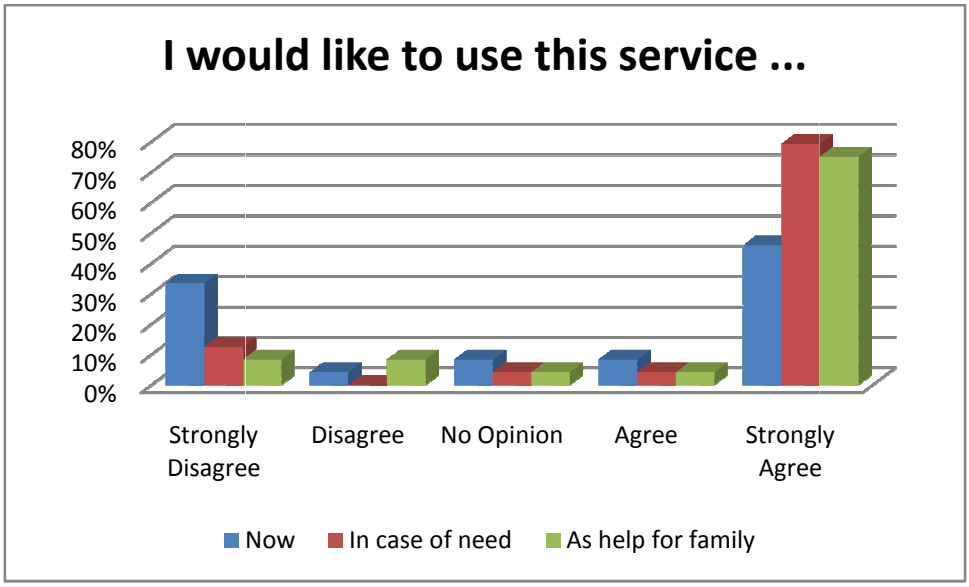


Figure 10.5

10.2.2Acceptability

10.2.2.1 Acceptability by ad-hoc questionnaire

The questionnaire compiled by the old persons denoted a positive degree of acceptability (4,44±1,05) about the "Outdoor walking support service". In particular using the domestic robot as a walking support didn't evoke anxiety in the participants who reported they enjoyed during the experimentation. However we have to consider that elderly persons performed the task as a "game" because they drove the outdoor robot from a point A to a point B.

10.2.2.2 Acceptability by video analysis

On the acceptability, of the 25 users whose videos were analyzed, slightly under half (12) appeared noticeably comfortable and confident during the service based on their posture and demeanor. It is interesting to note that several of these users appeared comfortable

despite having minor difficulties with control. Users also expressed enjoyment through smiling and laughing (11 users). It may not be appropriate to take all of these events as expressions of enjoyment, however. In a few cases, users smiled or laughed when they had difficulties controlling the robot, possibly to mitigate embarrassment. Only two users verbally expressed negative opinions about the service, one saying that it was tiring and the other saying, "I will never do to do shopping with the robot" (this user had significant and persistent difficulties controlling the robot).
A large number of the users (8) commented that the robot was simple or easy to drive. It seems that users appreciated the simplicity of the joystick interface even though it was not intuitive for all users. A majority of verbal statements were about control rather than overall impressions of the task. As this is covered by the usability analysis, they will not be discussed here.
In cases where there was significant difficulty driving the robot, this proved to be a source of anxiety for a minority of users. Four users (all female) made statements assigning blame to themselves for problems controlling the robot. There were only two cases (both female users) where the user also expressed nonverbal signs of anxiety through their gestures. In the case of the user whose negative experience was noted above, she was highly motivated to perform the task despite her difficulties, eventually refusing additional assistance from the experimenters.
In terms of their focus of attention, the majority (20) spent a significant amount of time during walking looking at the joystick and/or the experimenter (users were often in conversation with the experimenters during the task). Slightly fewer than half of users (11) spent a significant amount of time looking at their intended path or direction of travel without monitoring their joystick use.
The amount of time that users spent focused on the joystick while driving suggests that they found this controller difficult enough to use that they must frequently attend to it. The usability analysis supports this conclusion. However, users' statements about the ease of driving somewhat contradict this result. It may be that users find this interface acceptable because their self-perceived ease of use is high even in the presence of minor difficulties.

10.2.3HRI

10.2.3.1 HRI by questionnaire: Speech and GUI interaction

In this service the speech was not available and so it was not evaluated.
Regarding GUI interaction elderly persons reported very positive responses (4,86±0,51) but we have to consider that the GUI was very simple because it consisted in one icon to open and close the bin.

10.2.3.2 HRI by video analysis

On the HRI, slightly over half of the users (14) looked at the robot while walking, though they looked at the robot less frequently and for less time overall than at the joystick, experimenter, or path. The fact that not all users looked at the robot during the service could be due to their viewing the robot as a tool to use rather than as an agent or partner. Or it may simply be due to the fact that most or their attention was devoted to the task of driving. It is interesting to note that in several of the cases in which users looked at the

robot, they did so towards the latter half of the trial. It may be that as users gained confidence, they felt comfortable enough to shift their focus of attention to the robot. No users had body language that suggested that they were afraid of the robot. Users seemed confident and comfortable using the robot for physical support. One user went so far as to state, "it's as if robot gives me its hand."
Users who made note of the robot's appearance and behavior were also generally positive. Two users positively evaluated the robot's head motion. Only one user made a negative comment on the robot's appearance, calling it, "too big." Another user stated a preference for being able to drive the robot herself rather than having it move autonomously. This statement is in line with the recommendations collected prior to the implementation of the first experimental loop, confirming the benefit of the user-centered approach to the service design.
User's acceptance of this task is good overall, though difficulties with usability seem to have occupied much of the users' attention during evaluation. Further experiments should be conducted to refine the controller prior to the second experimental loop so that control does not distract from the task.
Some users verbally expressed that their confidence grew as they gained experience. For example, "after the beginning I feel better with the robot, " and, "now, in contrast to the beginning, I can drive the robot better." It may be that problems with control may disappear or be reduced with practice. Based on these types of statements and on the change in users' nonverbal behavior over time, it would be useful to include pre-experiment training in the second experimental loop. This would allow data about service acceptability and quality of interaction to be collected without interference from learning effects.

10.2.4Quality of life

Regarding quality of life, elderly users reported neutral/positive responses (3,75±1,39) to item *"I think my independence would be improved by the use of the robot for this service"* because, as said before, the participants were autonomous for their daily activities and they saw this service as an assistance help.

Table 10.2 Acceptance attributes: descriptive statistics

Outdoor walking support		mean	std	Acceptance Score	std
ACCEPTANCE	Attitude	4,33	1,36	72,40	6,28
	Acceptability	4,44	1,05		
	GUI interaction	4,86	0,51		
	QoL	3,75	1,39		

For reporting the final acceptance result, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the service was not acceptable, from 65-84 was acceptable and from 85-100 the result was

excellent. At last the "Outdoor walking support service" was goodly acceptable (89,01±6,28) by elderly participants.

Table 10.3 Acceptance Score: descriptive statistics

Outdoor walking support	mean	std	#	%
Not Acceptable 0-64	-	-	0	0%
Acceptable 65-84	77,25	7,31	5	21%
Excellent 85-100	92,11	4,75	19	79%
Total	89,01	8,06	24	100%

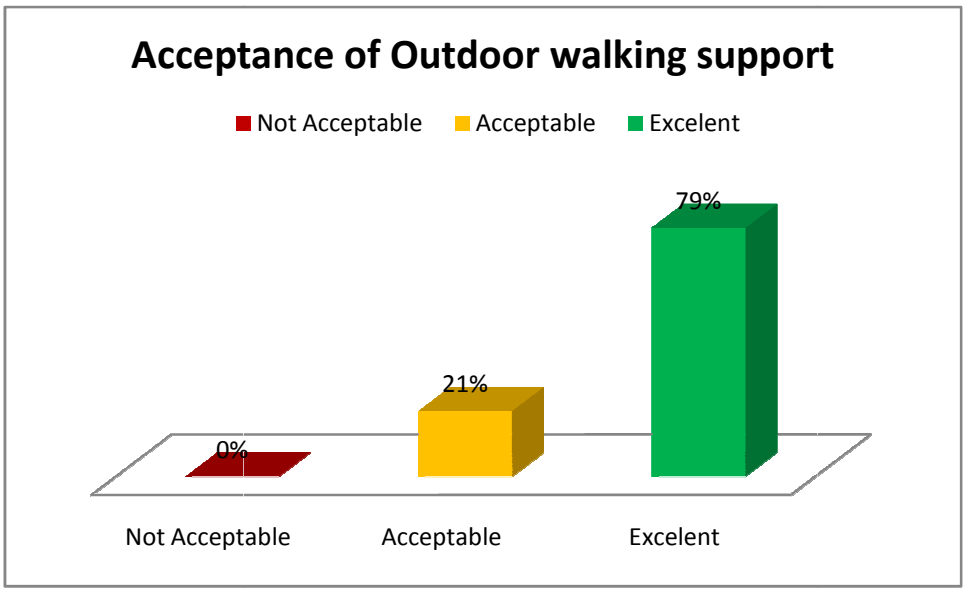


Figure 10.6 acceptance result

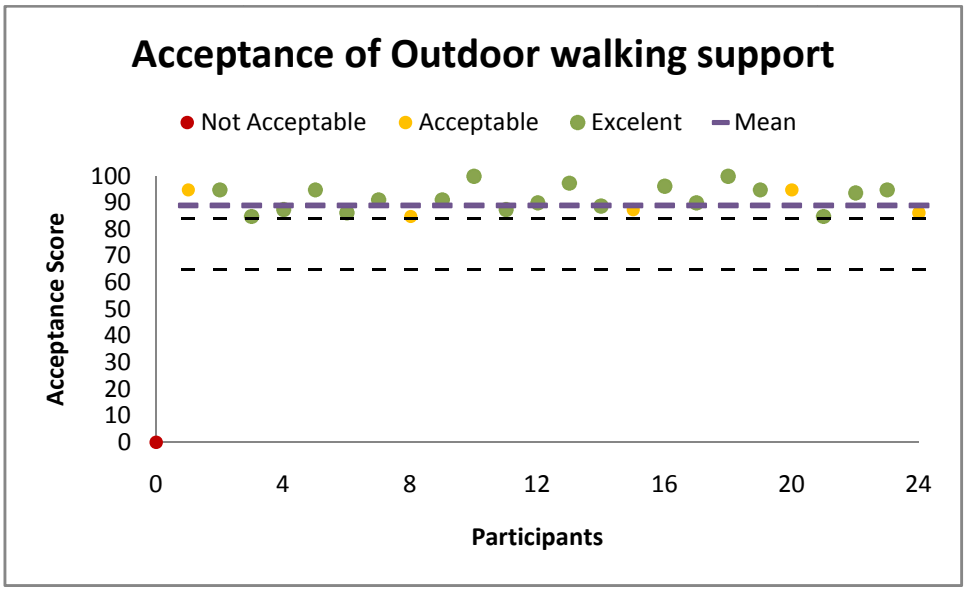


Figure 10.7 Acceptance result

10.3 Technical data

10.3.1 Time to perform the service

In the following tables10.3-10.4, the time to perform the task. Observing the table we can note an improvement between outward and backward demonstrating that elderly learned quickly to drive the robot.

Table 10.4

Time (sec) for going from point A to point B						
	0-1	**1-2**	**2-3**	**3-4**	**4-5**	**>5**
#	1/24 (~4%)	6/24 (~25%)	11/24 (~46%)	5/24 (~21%)	-	1/24 (~4%)
Mean	00:48	01:46±00:19	02:23±00:17	03:36±00:21	-	05:13

Table 10.5

Time (sec) for going from point B to point A						
	0-1	**1-2**	**2-3**	**3-4**	**4-5**	**>5**
#	1/24 (~4%)	10/24 (~42%)	9/24 (~38%)	3/24 (~13%)	1/24 (~4%)	-
Mean	00:51	01:43±00:15	02:14±00:08	03:47±00:11	04:37	-

11 UTAUT questionnaire results

11.1 UTAUT applied to the domestic platform

The UTAUT questionnaire was proposed at the end of the Shopping and Drug delivery, Communication and Garbage collection services experimentation in order to evaluate the acceptance of the domestic platform which was involved in the services. Table 11.1 gives an overview of the descriptive statistics concerning the several constructs.

Table 11.1

Construct	Mean	std
ANX	1,43	0,99
ATT	3,81	1,39
FC	3,97	1,13
ITU	3,99	1,34
PAD	4,49	0,84
PENJ	4,38	1,11
PEOU	3,90	1,49
PU	3,54	1,56
SI	4,20	1,23
Trust	4,89	0,36
SP	2,70	1,77
PS	3,56	1,54

ANX: Anxiety, ATT: Attitude, FC: Facilitating Conditions, ITU: Intention to Use, PAD: Perceived Adaptability, PENJ: Perceived Enjoyment, PEOU: Perceived Ease of Use, PS: Perceived Sociability, PU: Perceived Usefulness, SI: Social Influence, SP: Social Presence

ANX-The domestic platform didn't evoke anxious or emotional reactions (1,43±0,99) during its use to perform the Robot-Era services.

ATT-The attitude (3,81±1,39) towards DORO was positive, because the participants thought it was a good idea to use the robot.

FC-Regarding Facilitating Conditions (3,97±1,13) elderly people reported that they knew enough of the robot to make good use of it.

ITU-About Intention to Use (3,99±1,34) the results was very positive, but it should be explained in this way :" If I had DORO at home I would use it, but now I wouldn't purchase it because I do not need it".

PAD-A very positive results regards the Perceived Adaptability (4,49±0,84) because DORO was perceived as adaptable to user's needs.

PENJ-The overall result showed that the Perceived Enjoyment was high (4,38±1,11) because the participants enjoyed the domestic robot talking to them and performing task with the domestic robot.

PEOU-Regarding the Perceived Ease of Use (3,90±1,49) the users thought that DORO was easy to use and they knew enough about it after the experimental test.

PU-About the Perceived Usefulness (3,54±1,56) the participants' responses were neutral because they perceived the usefulness of the robot , but not the usefulness for them. We have to remember that the elderly volunteers were autonomous for their daily activities.

SI-Regarding the Social Influence (4,20±1,23) the elderly reported that if they should use the domestic robot it would give a good impression.

Trust-About Trust (4,89±0,36) the users would trust the domestic robot if it gave them advice and they would follow that advice.

SP-Regarding Social Presence (2,70±1,77) the responses were negative-neutral because the users perceived DORO always as a machine and not as living being.

PS-Regarding Perceived Sociability (3,56±1,54) the responses were neutral enough because elderly users preferred a person to socialize and we have to consider that speech capabilities of DORO were poor.

11.2 UTAUT applied to the outdoor platform

The UTAUT questionnaire was proposed at the end of the Outdoor walking support service experimentation in order to evaluate the acceptance of the outdoor platform. Table 12.1 gives an overview of the descriptive statistics concerning the several constructs.

Table 11.2

Construct	Mean	std
ANX	1,24	0,72
ATT	3,69	1,59
FC	4,48	0,87
ITU	2,89	1,86
PAD	4,15	1,32
PENJ	4,07	1,25
PEOU	3,83	1,59
PU	2,60	1,68
SI	4,04	1,44
Trust	4,75	0,70
SP	2,19	1,63
PS	2,75	1,61

ANX: Anxiety, ATT: Attitude, FC: Facilitating Conditions, ITU: Intention to Use, PAD: Perceived Adaptability, PENJ: Perceived Enjoyment, PEOU: Perceived Ease of Use, PS: Perceived Sociability, PU: Perceived Usefulness, SI: Social Influence, SP: Social Presence

ANX-The outdoor platform didn't evoke anxious or emotional reactions (1,24±0,72) during its use to perform the walking support service.

ATT-The attitude (3,69±1,59) towards ORO was positive enough, because the participants thought it was a good idea to use the robot.

FC-Regarding Facilitating Conditions (4,48±0,87) elderly people reported that they knew enough of the robot to make good use of it.

ITU-About Intention to Use (2,89±1,86) the results was negative enough because elderly people would use ORO only in case of need.

PAD-A very positive results regards the Perceived Adaptability (4,15±1,32) because ORO was perceived as adaptable to user's needs.

PENJ-The overall result showed that the Perceived Enjoyment was high (4,07±1,25) because the participants enjoyed the domestic robot talking to them and performing task with the domestic robot. We have to consider that speech capabilities of ORO were poor.

PEOU-Regarding the Perceived Ease of Use (3,83±1,59) the users thought that ORO was easy to use and they knew enough about it after the experimental test.

PU-About the Perceived Usefulness (2,60±1,68) the participants' responses were negative enough because they perceived the usefulness of the robot , but not the usefulness for them. We have to remember that the elderly volunteers were autonomous for their daily activities.

SI-Regarding the Social Influence (4,04±1,44) the elderly reported that if they should use the domestic robot it would give a good impression.

Trust-About Trust (4,75±0,70) the users would trust the domestic robot if it gave them advice and they would follow that advice.

SP-Regarding Social Presence (2,19±1,63) the responses were negative because the users perceived ORO always as a machine and not as living being.

PS-Regarding Perceived Sociability (2,75±1,61) the responses were negative-neutral because elderly users preferred a person to socialize and we have to consider that speech capabilities of ORO were poor.

Results of the first experimental loop in Angen

12 Participants

In Angen 12 elderly persons were recruited. They were 4 woman and 8 men over 65 years old and their mean age was 70,67±5,37.
About their skills regarding the everyday technological devices as expected, most of them knew very well : TV (~67%) and telephone (~92%) Table 12.1. An encourage data was that ~50% of them the use a PC and ~58% a tablet and ~45% a smartphone.

Table 12.1

Attitude: Use of technology							
	1	2	3	4	5	not applicable	don't know
Phone	0,00%	0,00%	0,00%	8,33%	91,67%	0,00%	0,00%
Mobile phone	0,00%	0,00%	18,18%	27,27%	54,55%	0,00%	0,00%
Smart phone	9,09%	9,09%	9,09%	18,18%	45,45%	9,09%	0,00%
TV/ Remote control	0,00%	0,00%	16,67%	16,67%	66,67%	0,00%	0,00%
Remote control for TV	0,00%	0,00%	25,00%	8,33%	66,67%	0,00%	0,00%
CD/DVD player	0,00%	8,33%	16,67%	33,33%	41,67%	0,00%	0,00%
PC	0,00%	0,00%	25,00%	25,00%	50,00%	0,00%	0,00%
Internet	18,18%	0,00%	18,18%	18,18%	27,27%	18,18%	0,00%
Tablet	0,00%	0,00%	16,67%	25,00%	58,33%	0,00%	0,00%
Bank ATM	0,00%	0,00%	16,67%	25,00%	41,67%	16,67%	0,00%
Devices to increase your safety	9,09%	0,00%	0,00%	27,27%	63,64%	0,00%	0,00%
Washing machine	11,11%	11,11%	11,11%	0,00%	11,11%	55,56%	0,00%
Telemedicine	10,00%	10,00%	10,00%	0,00%	50,00%	10,00%	10,00%
Alarm system	0,00%	0,00%	0,00%	8,33%	91,67%	0,00%	0,00%

13 Acceptance of Robot-Era Robot Appearance Results

In the Sweden experimental pilot site only two robotics platforms were involved: the domestic one, called DORO, and the Scitos G6 Transporter, called the RESIDENTIAL. The RESIDENTIAL platform was a commercial product by Metralab, so we decided to investigate only the reactions evoked in the participants by DORO which had the same appearance of the Italian one. In order to evaluate the acceptance of Robot-Era domestic platform appearance, the constructs were classified into six categories:

- Anxiety: Evoking anxious when the Robot-Era platforms were seen
- Aesthetics: Positive or negative judgements about Robot-Era platforms aesthetics

- Intention to interact: The outspoken intention to interact with the Robot-Era platforms when they were seen
- Trust: Inspired confidence when the Robot-Era platforms were seen
- Safety: The belief that the Robot-Era platforms perform reliability
- Facility Conditions: Factors of the Robot-Era platforms that could facilitate their use
- Perceived functions: The degree to which a person believes that Robot-Era platforms communicate their functions

Before starting the experimental test, the users observed DORO for some minutes and then they replied to the questionnaire items. The questionnaire was completed by 12 participants and in the following we present the results concerning the proposed constructs about the domestic platform Table 12.1.
The appearance of the domestic platform evoked little anxiety (2,71±1,24), the responses to aesthetics were neutral (3,06±1,04) and the participants had little confidence in the robot (2,92±1,24). Also elderly people were neutral-disposed (3,17±1,07) to interact with robot and the presence of an head on it was not considered an important factor to facilitate the interaction. After the first observation elderly people were neutral about the reliability of the domestic platform (3,36±0,80) and also the position of the touch-screen and the dimensions compared to a domestic environment got a neutral-positive feedbacks from users (3,49±1,08).
At last the appearance of DORO didn't communicate its possible functions explicitly, in fact elderly people expressed a negative judgment (2,61±1,32).

Table.13.1 Descriptive statistics

	DORO	
	mean	std
Anxiety	2,17	1,24
Aesthetics	3,06	1,04
Intention to interact	3,17	1,07
Trust	2,92	1,24
Safety	3,36	0,80
Facility Conditions	3,49	1,08
Perceived functions	2,61	1,32

For reporting the results, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the Robot-Era appearance platform was not acceptable, from 65-84 was acceptable and from 85-100 the result was excellent. At last the domestic platform was not acceptable (63,95±13,07 Table 12.2) by elderly participants (67%) and also we can note that over the appearance of DORO was acceptable only for 4.

Table.13.2 Acceptance of DORO Descriptive statistics

DORO	mean	std	#	%
Not Acceptable 0-64	57,03	6,65	8	66,7%
Acceptable 65-84	73,33	9,62	3	25,0%
Excellent 85-100	91,11	-	1	8,3%
Total	63,95	13,07	12	100%

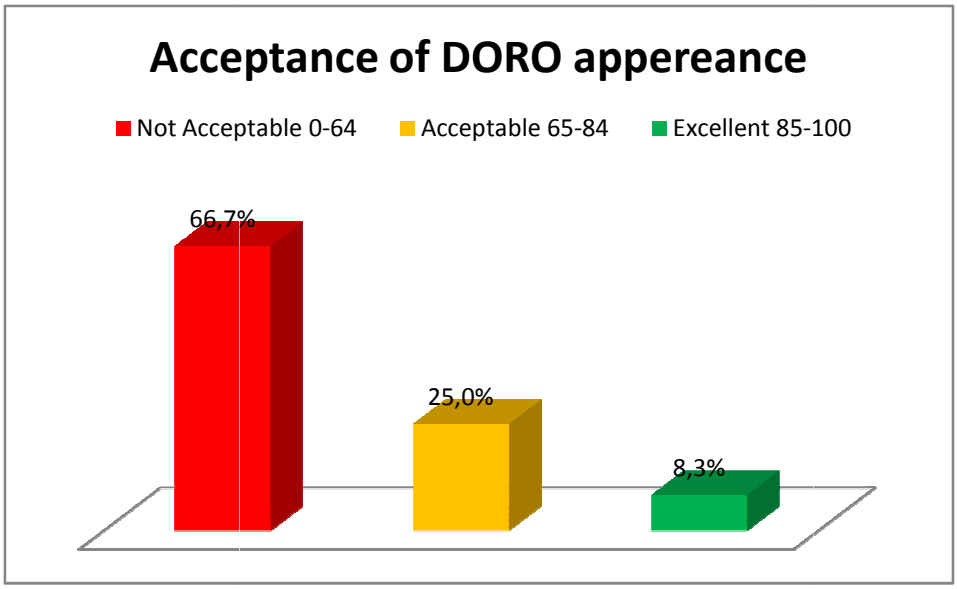


Figure 12.1 Acceptance of DORO Appearance result

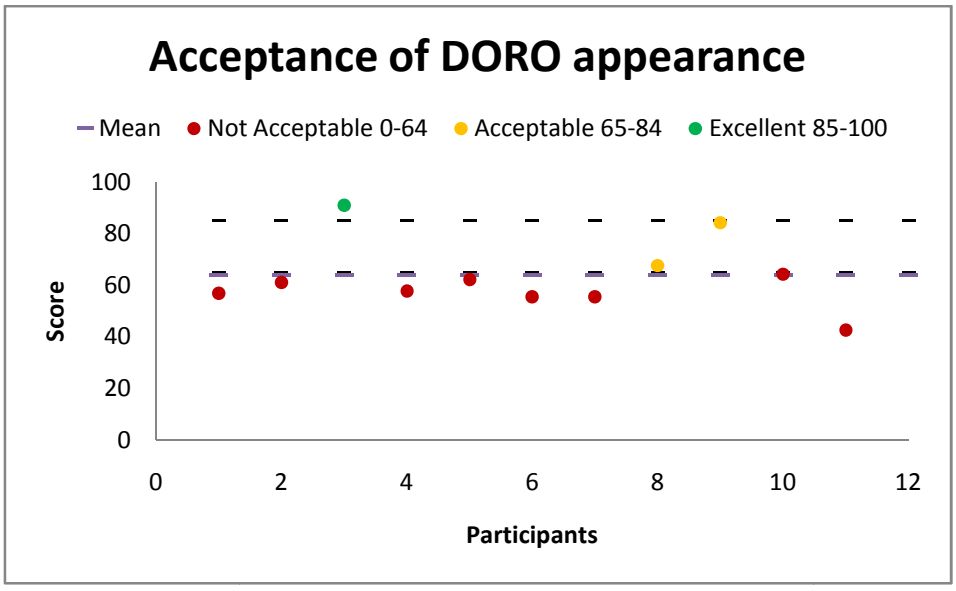
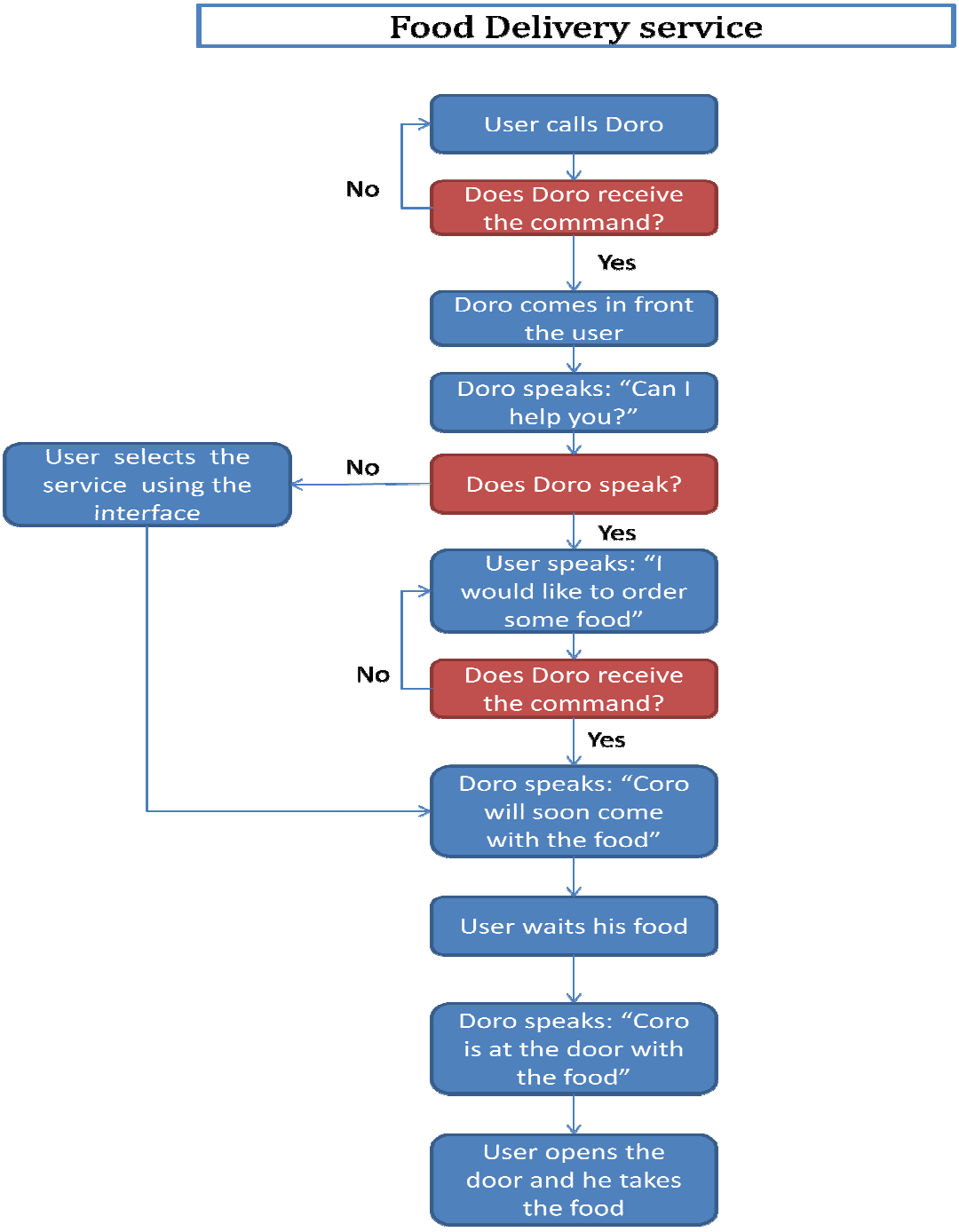


Figure 12.0.1 Acceptance of DORO Appearance

14 Food delivery service results

In this scenario the elderly user had to imagine he didn't want to cook his meal on his own, because he had little time. So he decided to order a meal The speech interaction and interface interaction were necessary to perform this service. This service was composed by several steps that are shown in Figure14.1.



14.14.1 Flow chart of the service

14.2 User during experimentation

14.1 Usability

14.1.1 Usability by SUS Scale

Immediately after the end of the "Food delivery service" , the system usability scale (SUS) referred to the service was proposed.

Table 14.1 SUS Score: descriptive statistics

Food delivery	mean	std	#	%
Not Usable 0-64	16,67	-	1	12,5%
Usable 65-84	70,83	3,4	4	50,0%
Excellent 85-100	88,89	2,41	3	37,5%
Total	70,83	23,78	8	100%

Given that data are few and they can't be significant however they can show the trend of the experimentation conducted in Sweden. The usability of the "Food delivery service" was evaluated as usable (70,83±23,78) by elderly people, but we have to note the high standard deviation. The service was easy enough to use (3,75±1,39) and nobody thought he needed the support of a technical person to be able to perform it.

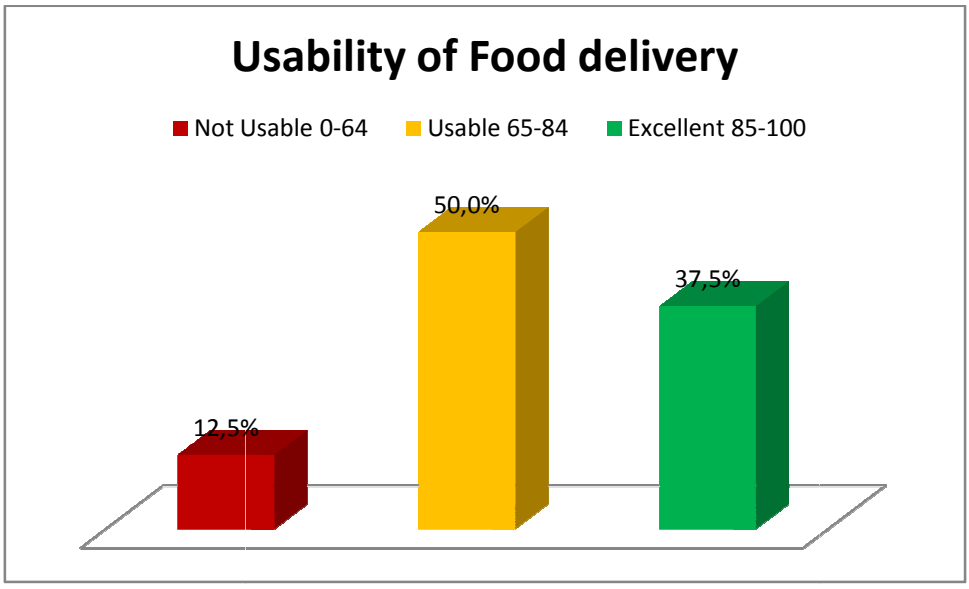


Figure 14.3 Usability of Food delivery result

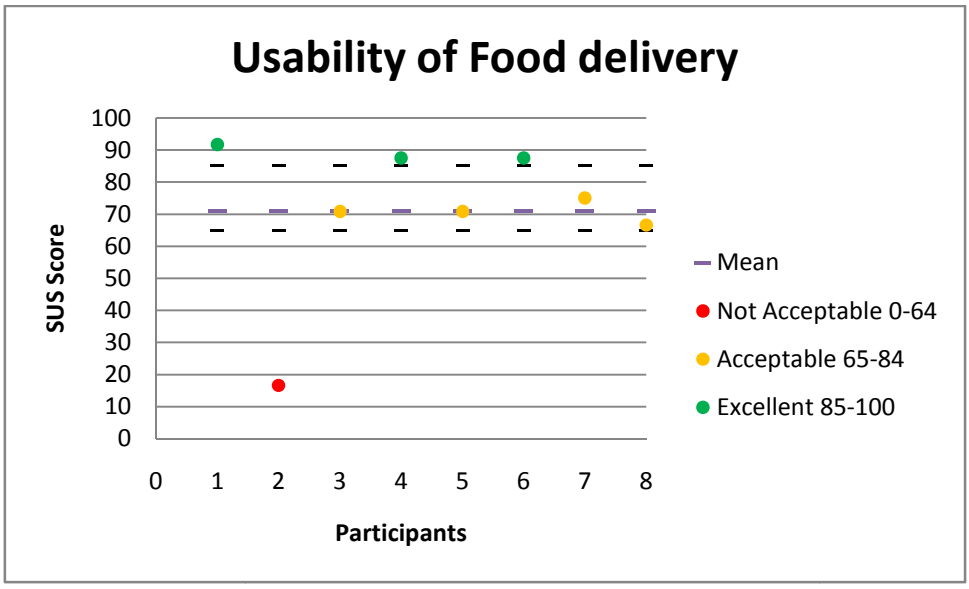


Figure 14.4 Usability of Food delivery result

14.1.2Usability by video analysis

Person fails to see the advice, maybe because the change between "No new notifications" and "Food order acknowledged" is not clearly visible for the user. The test person is unsure if the selection of one menu was accepted or not. The user fails to see the advice "Food order acknowledged" and expected some feedback as a confirmation of the order.

14.2 Acceptance

14.2.1Attitude

The elderly user demonstrated a high attitude (4,45±0,94) towards this service. However the users reported positive responses if they would use the Robot-Era platforms for this

service in case of need (4,18±1,25) or if the robots could help the family/caregiver's work (4,45±0,82).

14.2.2Acceptability

14.2.2.1 Acceptability by ad-hoc questionnaire

The questionnaire compiled by the old persons denoted a positive degree of acceptability (4,05±1,24) about the "Food delivery service". In particular using the Robot-Era system for the service performing didn't evoke anxiety in the participants who reported they enjoyed enough during the experimentation and they would trust in the robot ability to perform the service.

14.2.2.2 Acceptability by video analysis

The observation is based on eight videos. Speech interaction is used to call the robot, then the users used the tablet to command the service. The users have shown to have a good comprehension of the command to be given and they have demonstrated to understand the graphic interface, just one user has not understood the service and perform another task – shopping -. Some technical failures during the service were present. The users seem enjoyed when they found the tray with food outside the door, and some users say "Thanks!" to DORO and smiled at him, suggesting a friendly communication and satisfaction with the proposed service.

14.2.3HRI

HRI by questionnaire: Speech and GUI interaction

Regarding the speech interaction elderly users reported positive enough responses (3,77±1,51).
About the GUI interaction the old persons had a more positive approach (4,39±1,12) even if we have to note that the GUI was very simple.

14.2.4Quality of life

Regarding quality of life, elderly users reported neutral responses (3,64±1,36) to item *"I think my independence would be improved by the use of the robot for the service"* .

Table14.2 Acceptance attributes: descriptive statistics

	Food delivery	mean	std	Acceptance Score	std
ACCEPTANCE	Attitude	4,45	0,94	82,91	12,16
	Acceptability	4,05	1,24		
	Speech interaction	3,77	1,51		
	GUI interaction	4,39	1,12		
	QoL	3,64	1,36		

For reporting the final acceptance result, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the "Food delivery service" was not acceptable, from 65-84 was acceptable and from 85-100 the result was excellent. At last the "Food delivery service" was acceptable (82,91±12,16) by elderly participants. However we can note that 64% of the users evaluated the service as excellent.

Table114.3 Acceptance Score: Descriptive Statistics

Food Delivery	mean	std	#	%
Not Acceptable 0-64	52	-	1	9,1%
Acceptable 65-84	77,78	6,01	3	27,3%
Excellent 85-100	89,52	3,30	7	63,6%
Total	82,91	12,16	11	100%

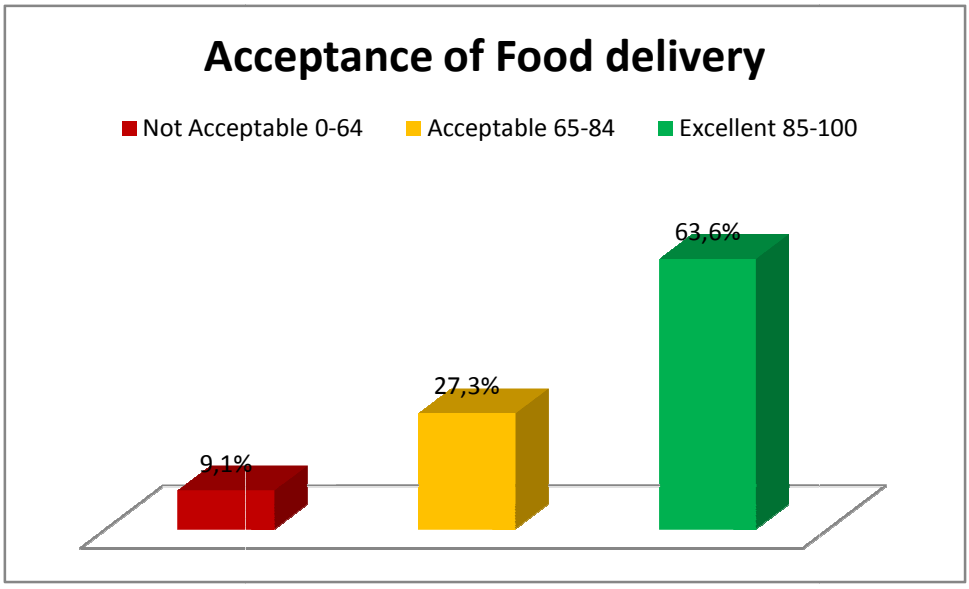


Figure 14.5 Acceptance of Food delivery result

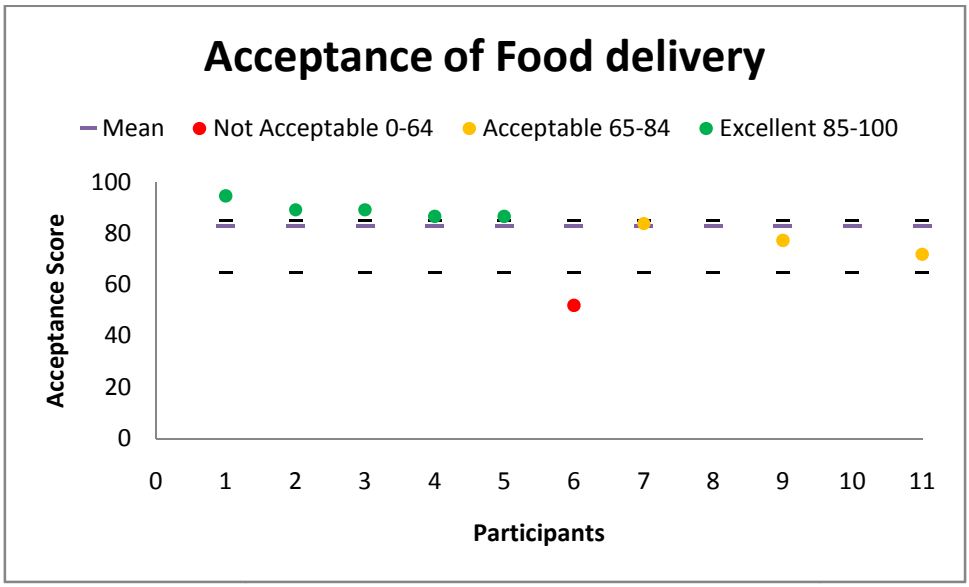
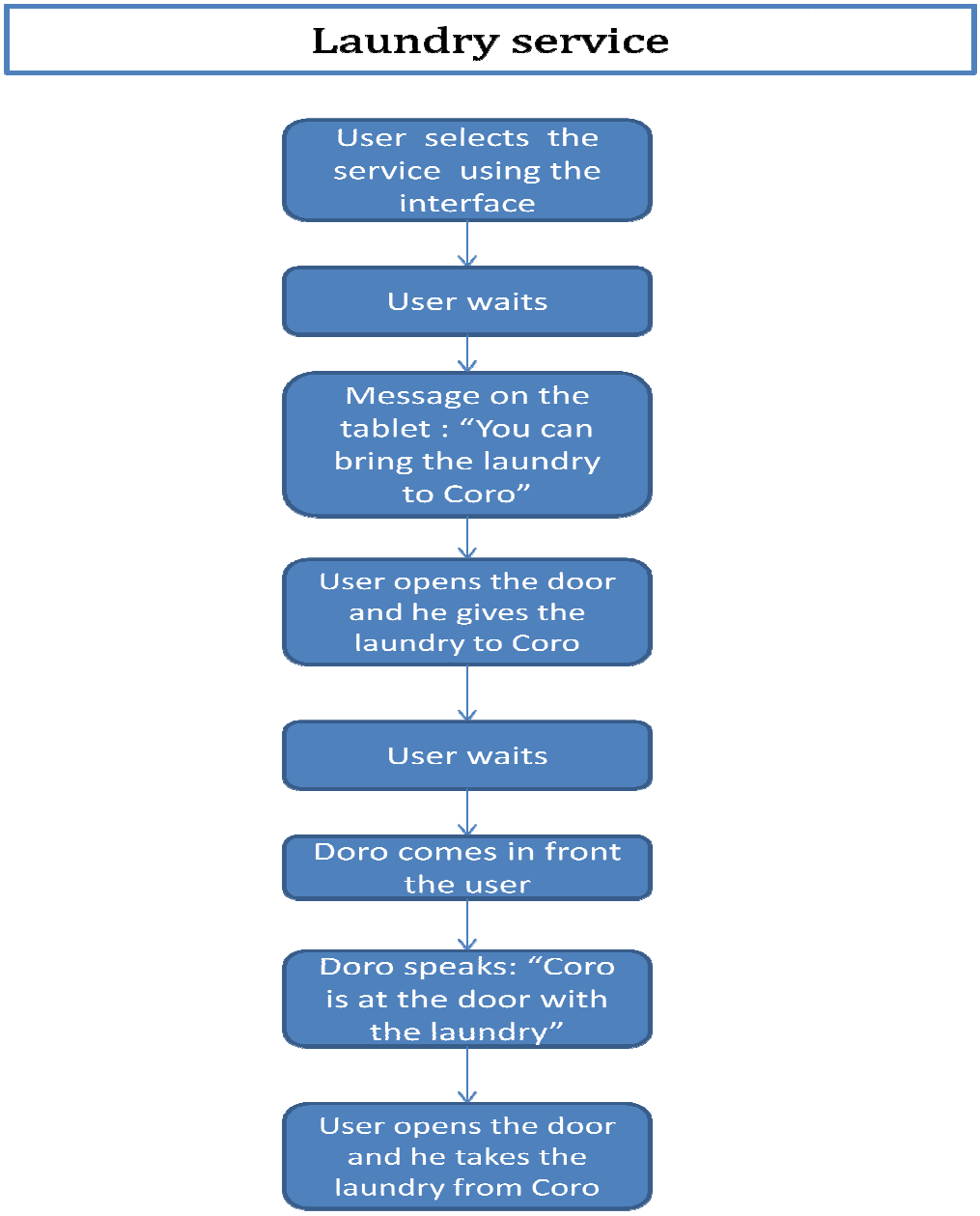


Figure 14.14.6 Acceptance of Food delivery result

15 Laundry delivery service results

In this scenario the elderly user he had to imagine he could not leave his home because he was sick so bearing in mind this presupposition the participant had to call the domestic robot in order to select and active the "Laundry service". The speech interaction and interface interaction were necessary to perform this service. This service was composed by several steps that are shown in Figure15.1.



15.1 Flow chart of the service

15.2 User during experimentation

15.1 Usability

15.1.1 Usability by SUS Scale

Immediately after the end of the "Laundry delivery service" , the system usability scale (SUS) referred to the service was proposed.

Table15.1 SUS Score: descriptive statistics

Laundry delivery	mean	std	#	%
Not Usable 0-64	56,94	9,62	3	33,3%
Usable 65-84	77,78	9,62	3	33,3%
Excellent 85-100	67,36	14,73	3	33,3%
Total	75,93	17,4	9	100%

Given that data are few and they can't be significant however they can show the trend of the experimentation conducted in Sweden. The usability of the "Laundry delivery service" was evaluated as usable (75,93±17,40) by elderly people, but the sample was divided equally between Not usable (33%), Usable (33%) and Excellent (33%).

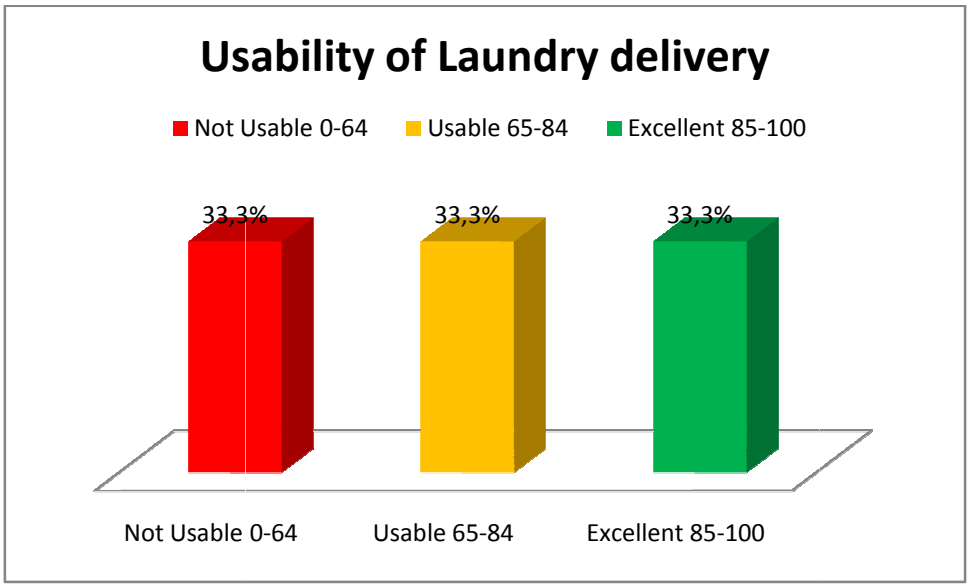


Figure 15.3 Usability of Food delivery result

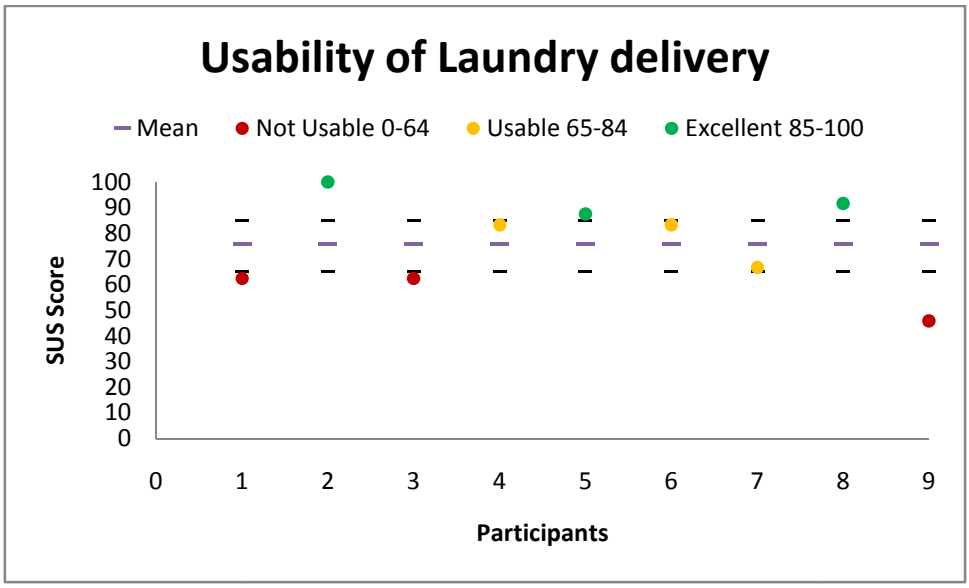


Figure 15.4 Usability of Food delivery result

15.1.2Usability by video analysis

The difference between the "Assist"- and "Collect"-button is not clear. The test person does not understand the wording and/or the icons of the "Assist"- and "Collect"-button. They do not seem to be intuitive enough.

15.2 Acceptance

15.2.1Attitude

The elderly user demonstrated a high attitude (4,50±1,33) towards this service. In fact the users reported positive responses if they would use the Robot-Era platforms for this service in case of need (4,70±0,67) or if the robots could help the family/caregiver's work (4,60±0,63).

15.2.2 Acceptability

15.2.2.1 Acceptability by ad-hoc questionnaire

The questionnaire compiled by the old persons denoted a positive degree of acceptability (4,10±1,63) about the "Laundry delivery service". In particular using the Robot-Era system for this service didn't evoke anxiety in the participants who reported they enjoyed during the experimentation and they would trust in the robot ability to perform the laundry delivery.

15.2.2.2 Acceptability by video analysis

The overall impression from the observation is that the users have conducted the task in a very formal way. It could be said that the users did not seem really enjoyed during the service and they seem to have some problem to understand the graphic interface to select the correct service. The support of the interviewer was necessary many times, during the task execution, suggesting an overall low level of comfort with the interface. All the subjects were focused mainly on the task more than on DORO, basically because the principal way of the interaction was constituted by the tablet, so it is difficult to evaluate the HRI and the overall satisfaction with this service. Nobody seem to feel anxious during the performance.

15.2.3HRI

HRI by questionnaire: Speech and GUI interaction

Regarding the speech interaction elderly users reported positive enough responses (3,90±1,12).
About the GUI interaction the old persons had a more positive approach (4,47±0,82) even if we have to note that the GUI was very simple.

15.2.4Quality of life

Regarding quality of life, elderly users reported neutral responses (3,50±1,43) to item *"I think my independence would be improved by the use of the robot for the service"* .

Table 15.2 Acceptance attributes: descriptive statistics

Food delivery		mean	std	Acceptance Score	std
ACCEPTANCE	Attitude	4,50	1,33	84,53	13,53
	Acceptability	4,10	1,63		
	Speech interaction	3,90	1,12		
	GUI interaction	4,47	0,82		
	QoL	3,50	1,43		

For reporting the final acceptance result, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the "Food delivery service" was not acceptable, from 65-84 was acceptable and from 85-100 the result was excellent. At last the "Laundry delivery service" was acceptable (84,53±13,53) by elderly participants. However we can note that 60% of the users evaluated the service as excellent.

Table 15.3 Acceptance Score: Descriptive Statistics

Food Delivery	mean	std	#	%
Not Acceptable 0-64	58,67	-	1	10%
Acceptable 65-84	74,67	5,81	3	30%
Excellent 85-100	93,78	4,59	6	60%
Total	84,53	13,53	10	100%

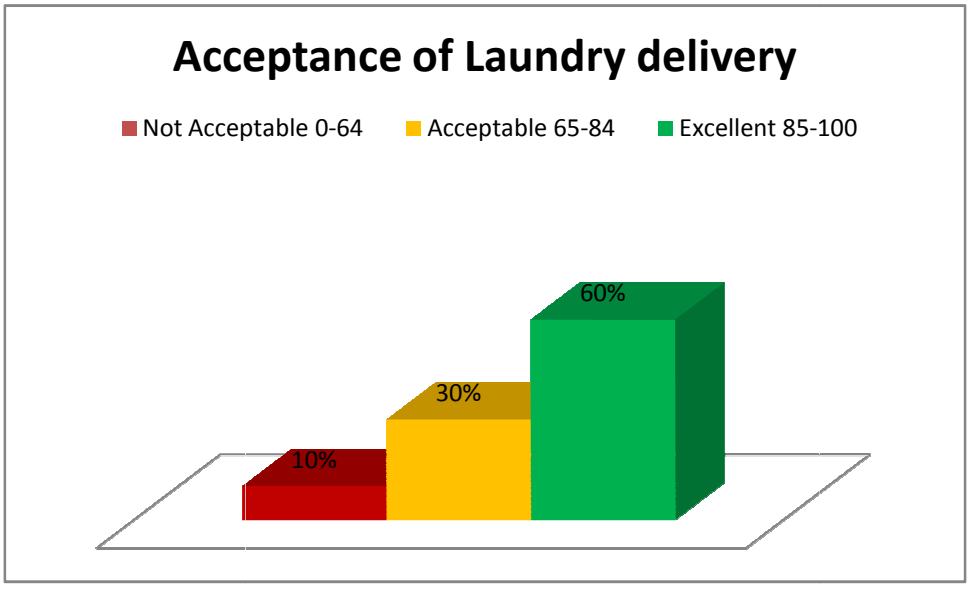


Figure 15.5 Acceptance of Food delivery result

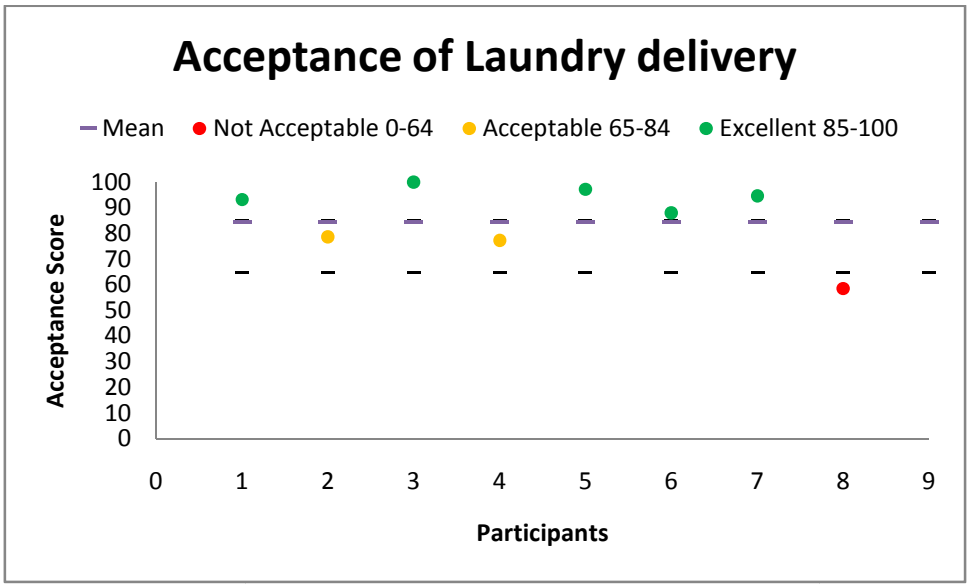


Figure 15.6 Acceptance of Food delivery result

16 Communication service results

In this scenario the users you wanted to use to robot to call a family member via Skype. This service could be performed by speech interaction and GUI use. The service was composed by several steps that are shown in Figure16.1 and Figure16.2

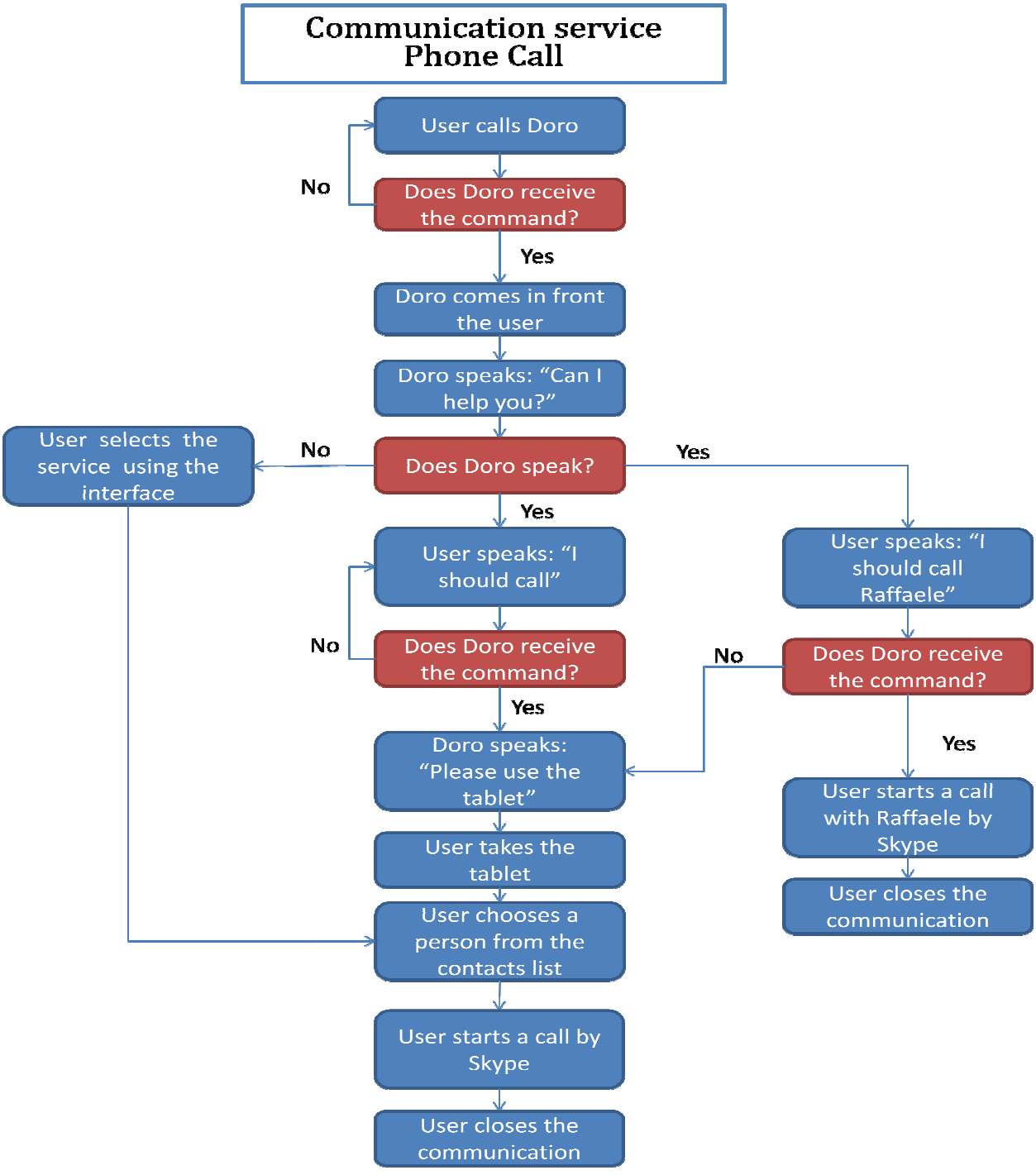
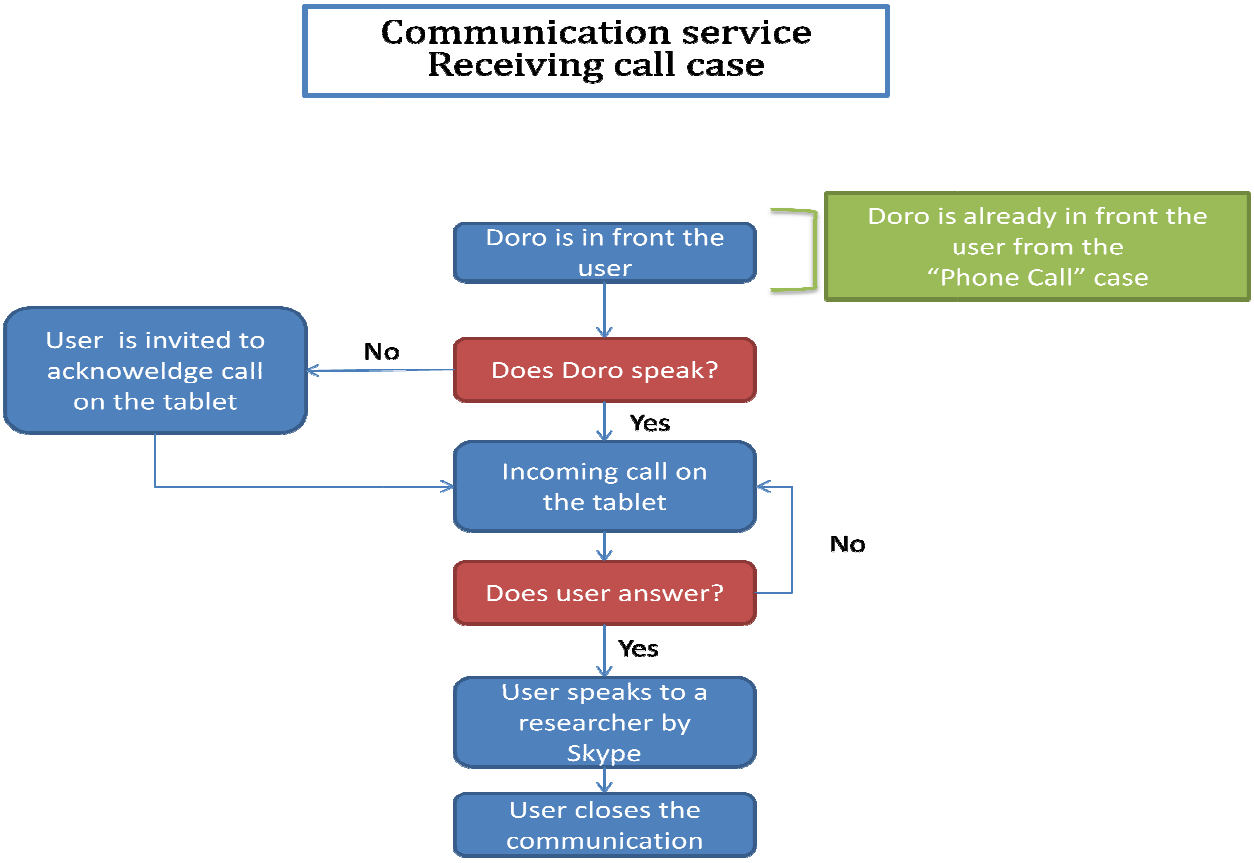


Figure 16.1 Phone call case flow chart



16.2 Flow chart of the service

16.3 User during experimentation

16.1 Usability

16.1.1Usability by SUS Scale

In order to evaluate the usability of "Communication service" we proposed the system usability scale (SUS), a simple ten item Likert scale that gives a global view of subjective assessments of usability.

Table 16.1Score: Descriptive Statistics

Communication	mean	std	#	%
Not Usable 0-64	42,71	17,14	4	33,3%
Usable 65-84	69,17	3,73	5	50,0%
Excellent 85-100	91,67	8,33	3	16,7%
Total	65,97	21,89	12	100%

The "Communication service" was perceived as barely usable (65,97±21,89) by elderly people and we can note that 33% of participants reported a totally negative evaluation. In particular the easy use of the service was negatively estimated (2,50±1,23).

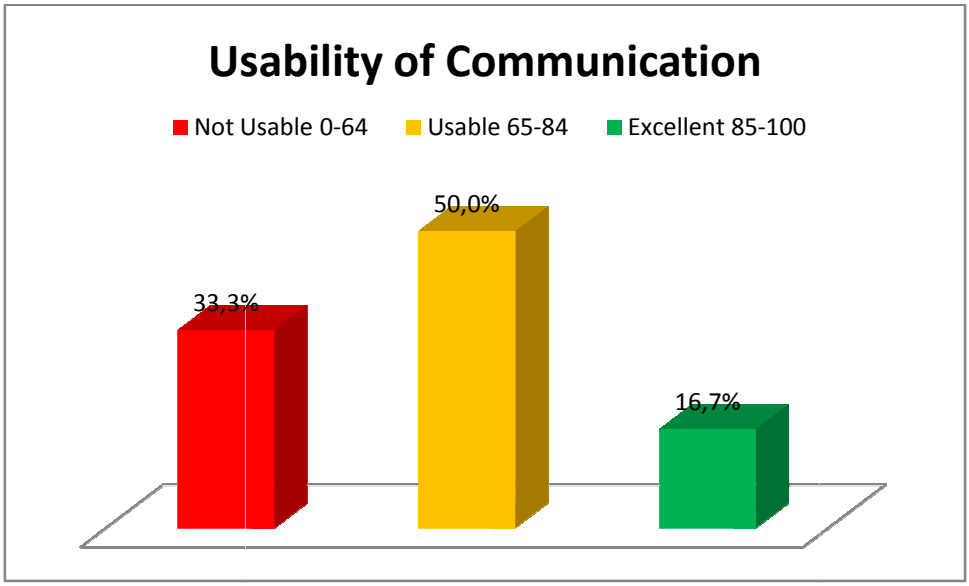


Figure 16.4 Usability result

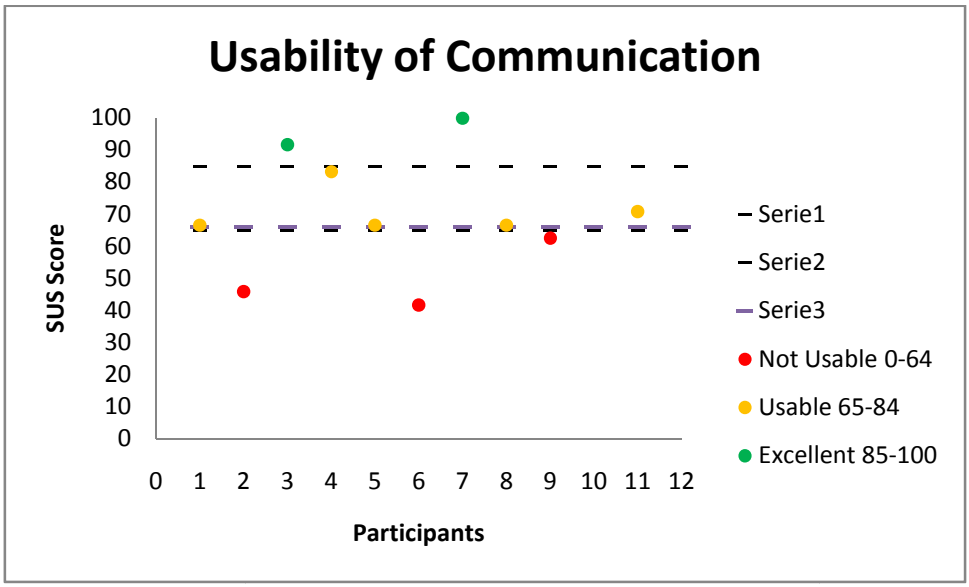


Figure 16.5 Usability result

16.1.2Usability by video analysis

The results do not differ from those reported in section 6.1.2

16.2 Acceptance

16.2.1Attitude

The elderly user demonstrated a medium-high attitude (3,81±0,82) towards this service.

16.2.1.1 Acceptability by ad-hoc questionnaire

The questionnaire compiled by the old persons denoted a medium-positive degree of acceptability (3,85±0,95) about the "Communication service". In particular using the Robot-Era system for the communication evoked a little anxiety in the participants.

16.2.1.2 Acceptability by video analysis

The overall impression is that the users seems highly enjoyed with DORO and they have understood perfectly the communication received by the robot without any difficulty. Nobody seem to be worried for the performance. The interaction with DORO was based both on the speech command and on the tablet. All the users are focused and concentrated on the task and the majority of the sample have shown to have a good predisposition toward the tablet: they did not need support, just some preliminary instruction from the experimenter. Also for making a call with Skype, they have shown to be highly competent by themselves. Only four users need to be supported by the interviewer during the task, suggesting an overall high level of comfort with DORO. No one has experienced anxiety during the task performance and there is not technical failures of the system. All the subjects are focused mainly on the robot, with the head, eyes and body orientation suggesting a good predisposition with DORO and the willingness to be more in contact with

it. The subjects are in open position, they smile at him and they look at DORO in-depth, supporting the idea of a good human robot interaction.

16.2.2HRI

HRI by questionnaire: Speech and GUI interaction

Regarding the speech interaction elderly users reported neutral responses (3,38±0,82). Also about the GUI interaction the old persons had a neutral approach (3,11±1,09).

16.2.3Quality of life

Regarding quality of life, elderly users reported negative responses (2,67±1,07) to item *"I think my independence would be improved by the use of the robot for this service"*

Table 16.2 Acceptance attributes: descriptive statistics

	Communication	mean	std	Acceptance Score	std
ACCEPTANCE	Attitude	3,81	0,82		
	Acceptability	3,85	0,95		
	Speech interaction	3,38	0,82	69,56	7,71
	GUI interaction	3,11	1,09		
	QoL	2,67	1,07		

For reporting the final acceptance result, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the service was not acceptable, from 65-84 was acceptable and from 85-100 the result was excellent. At last the "Communication service" was barely acceptable (69,56±7,71) by elderly participants.

Table 16.3 Acceptance score: descriptive statistics

Communication	**mean**	**std**	**#**	**%**
Not Acceptable 0-64	55,33	2,83	2	17%
Acceptable 65-84	72,40	4,22	10	83%
Excellent 85-100	-	-	0	0%
Total	69,56	7,71	12	100%

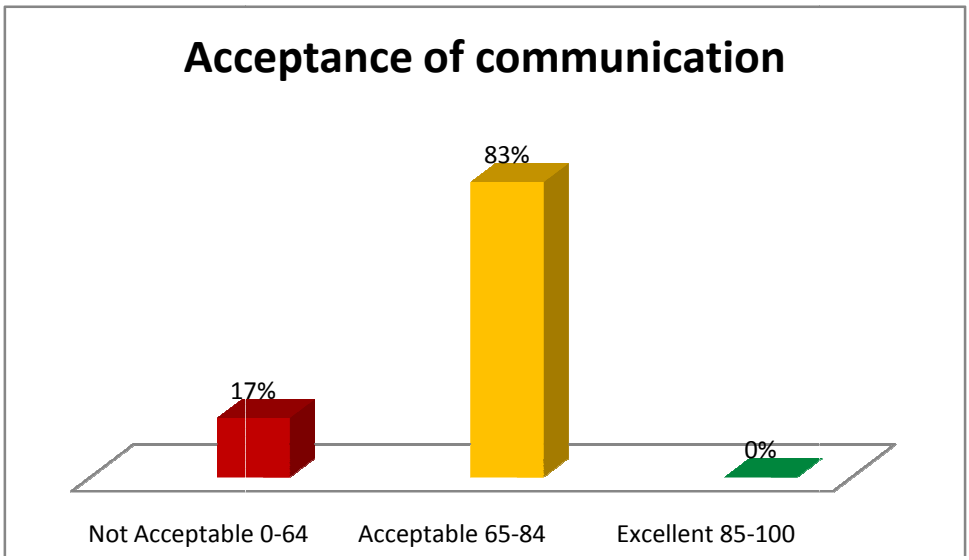


Figure 16.6 Acceptance result

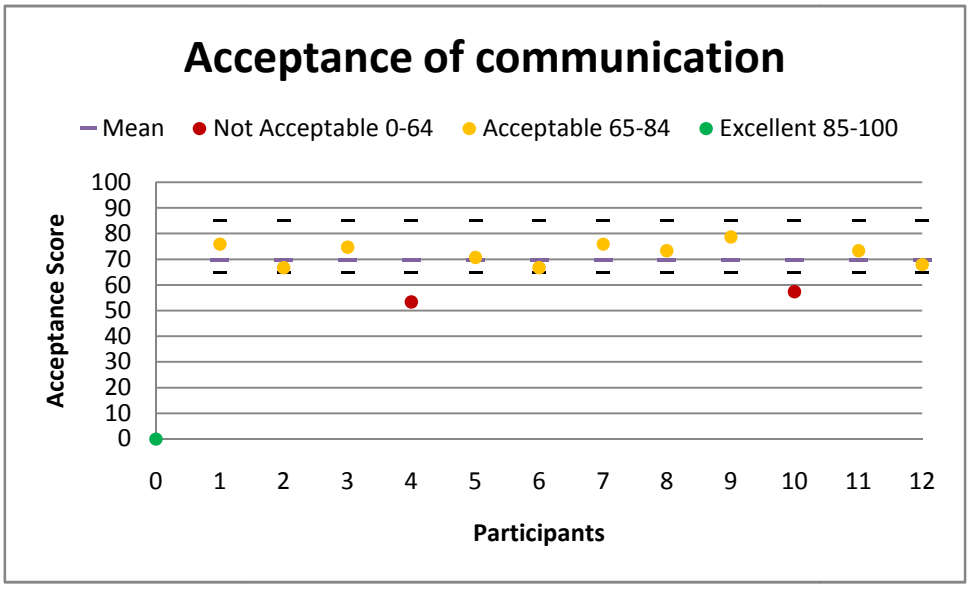


Figure 16.7 Acceptance result

17 Reminding service results

In this scenario the elderly user wanted to set a date on Robot-Era agenda. The user called the domestic robot to perform the task, then he moved to another room where he localized by AmI and the robot got through to him to remember the date. The speech interaction and interface interaction were necessary to perform this service. This service was composed by several steps that are shown in Figure17.1

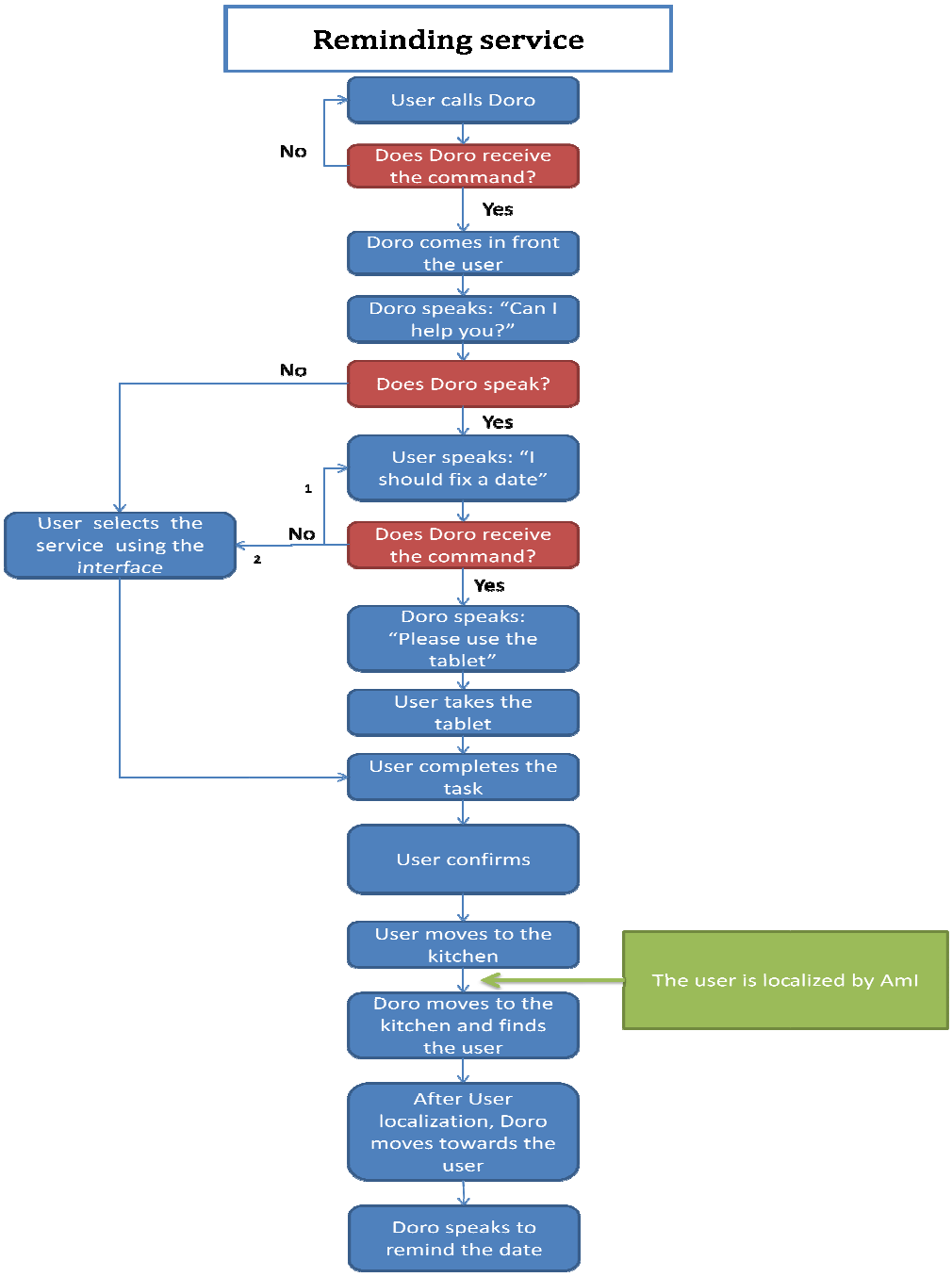
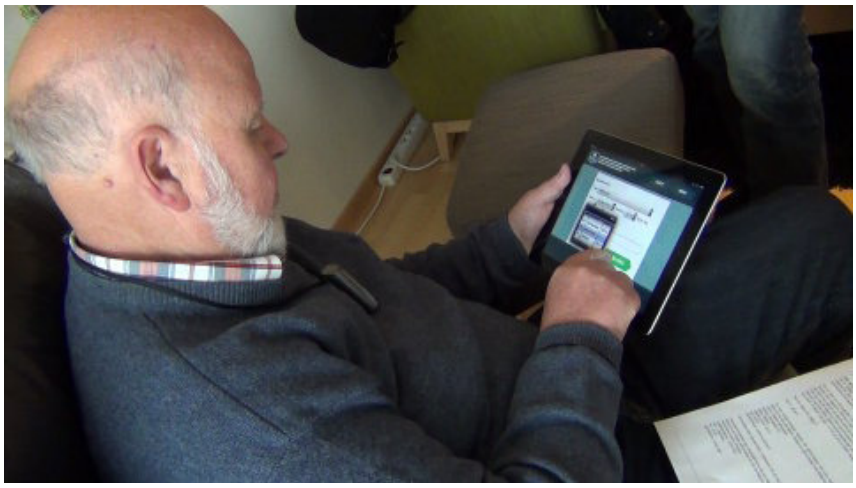


Figure 17.1 Reminding service flow chart



17.2 User during experimentation

17.1 Usability

17.1.1 Usability by SUS Scale

In order to evaluate the usability of "Reminding service" we proposed the system usability scale (SUS), a simple ten item Likert scale that gives a global view of subjective assessments of usability.

Table 17.1 SUS Score: descriptive statistics

Reminding	mean	std	#	%
Not Usable 0-64	32,50	21,13	8	88,9%
Usable 65-84	82,50	-	1	11,1%
Excellent 85-100	-	-	-	0,0%
Total	38,06	25,85	9	100%

The "Reminding service" was perceived as not usable (38,06±25,85) by elderly people. Because the participants reported that the easy use of the service was low (1,33±0,71).

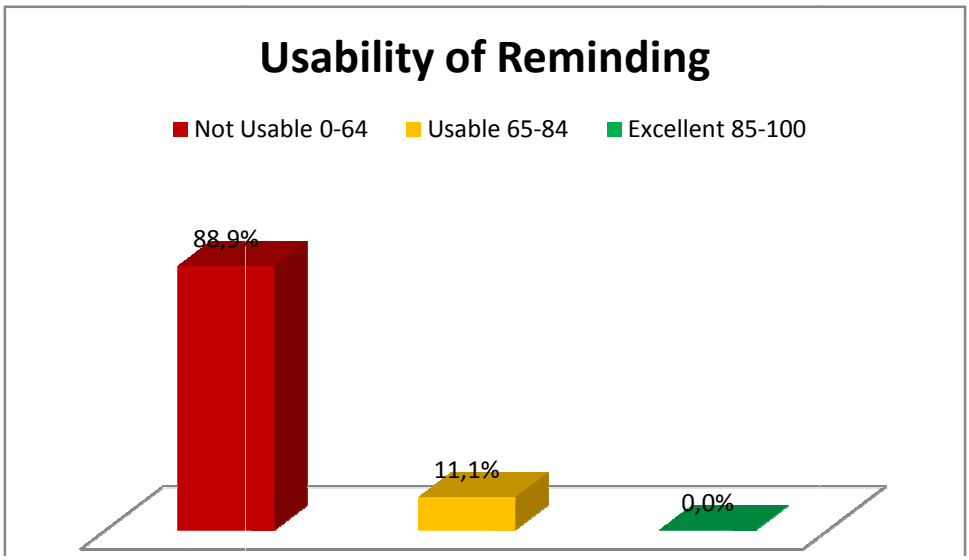


Figure 17.3 Usability result

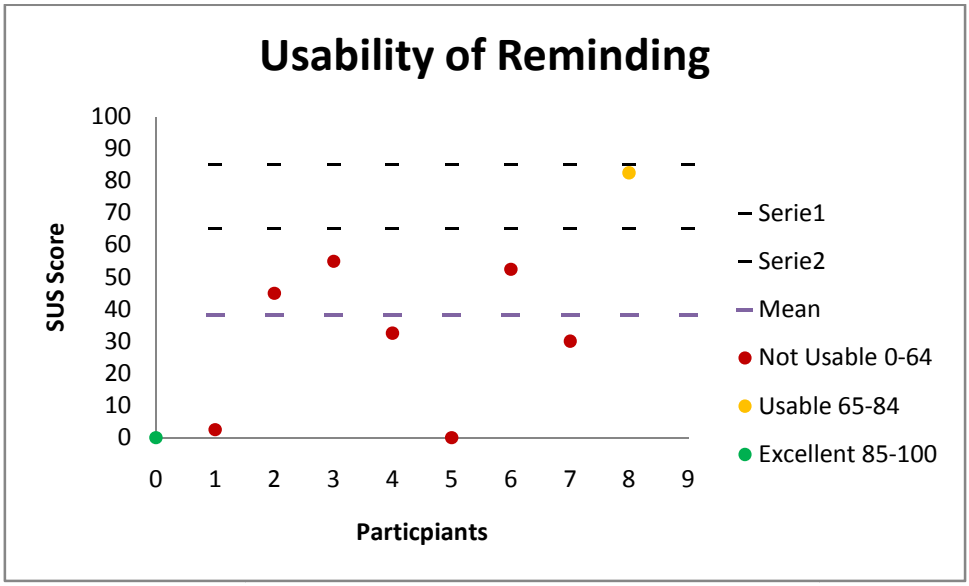


Figure 17.4 Usability result

17.1.2Usability by video analysis

The results do not differ from those reported in section 8.1.2

17.2 Acceptance

17.2.1Attitude

The elderly user demonstrated a high attitude (4,26±1,20) towards this service because they perceived it very useful even if the usability of the system was low.

17.2.2Acceptability

17.2.2.1 Acceptability by ad-hoc questionnaire

The questionnaire compiled by the old persons denoted a positive enough degree of acceptability (3,81±1,60) about the "Reminding". In particular using the domestic robot for this service the participants reported that it was not too much invasive for their lifestyle.

17.2.2.2 Acceptability by video analysis

For this service, the principal way of the interaction with the robot was constituted by the tablet. Seven videos were analyzed and from the observation it was possible to find that just three users are able to perform autonomously the task and using appropriately the interface. For the other subjects, it seems that the interface is felt as complicated and they need to receive support from the experimenter many times. Anyway, just one user has received a step by step support from the interviewer.

17.2.3HRI

17.2.3.1 HRI by questionnaire: Speech and GUI interaction

The HRI was not evaluated for this service

17.2.3.2 HRI by video analysis

On the interaction/communication side, it is very difficult to evaluate the HRI and the overall satisfaction with this service, mainly because the subjects were more concentrated on the tablet, and they have just called DORO at the beginning. Anyway, the vocal command was performed perfectly by all the users, even if it seems that the users were more concentrated on understanding what to do, instead of communicating more with the robot or just looking at it or feeling enjoyed.

17.2.4Quality of life

Regarding quality of life, elderly users reported medium-positive responses (3,67±1,41) to item *"I think my independence would be improved by the use of the robot for this service"*

Table 17.2 Acceptance attribute: descriptive statistics

Shop and Drug delivery		mean	std	Acceptance Score	std
ACCEPTANCE	Attitude	4,26	1,20		
	Acceptability	3,81	1,60		
	Speech interaction	-	-	79,17	16,68
	GUI interaction	-	-		
	QoL	3,67	1,41		

For reporting the final acceptance result, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the service was not acceptable, from 65-84 was acceptable and from 85-100 the result was excellent. At last the "Reminding service" was acceptable (79,17±16,68) by elderly participants. However we have to note that this data was not particularly significant because the HRI items were not evaluated.

Table 17.3 Acceptance Score : descriptive statistics

Reminding	mean	std	#	%
Not Acceptable 0-64	57,50	3,54	2	22%
Acceptable 65-84	67,50	3,54	2	22%
Excellent 85-100	91,50	3,35	5	56%
Total	78,61	15,96	9	100%

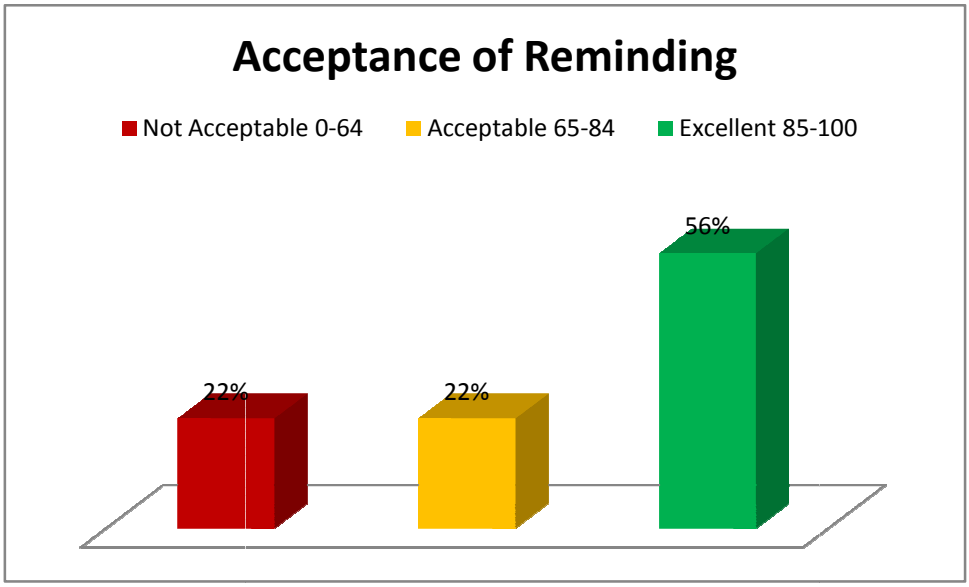


Figure 17.5 Acceptance result

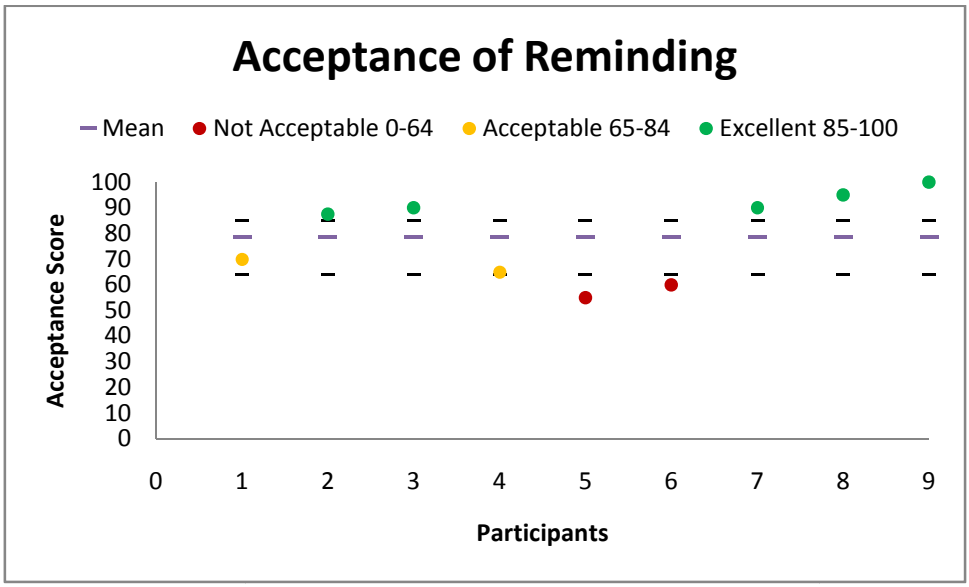


Figure 17.6 Acceptance result

18 Escort at night service results

In this scenario the elderly user had to imagine he had some temporary motor problem so he used the domestic robot as a walking support. This service consisted in the steps shown in Figure18.1

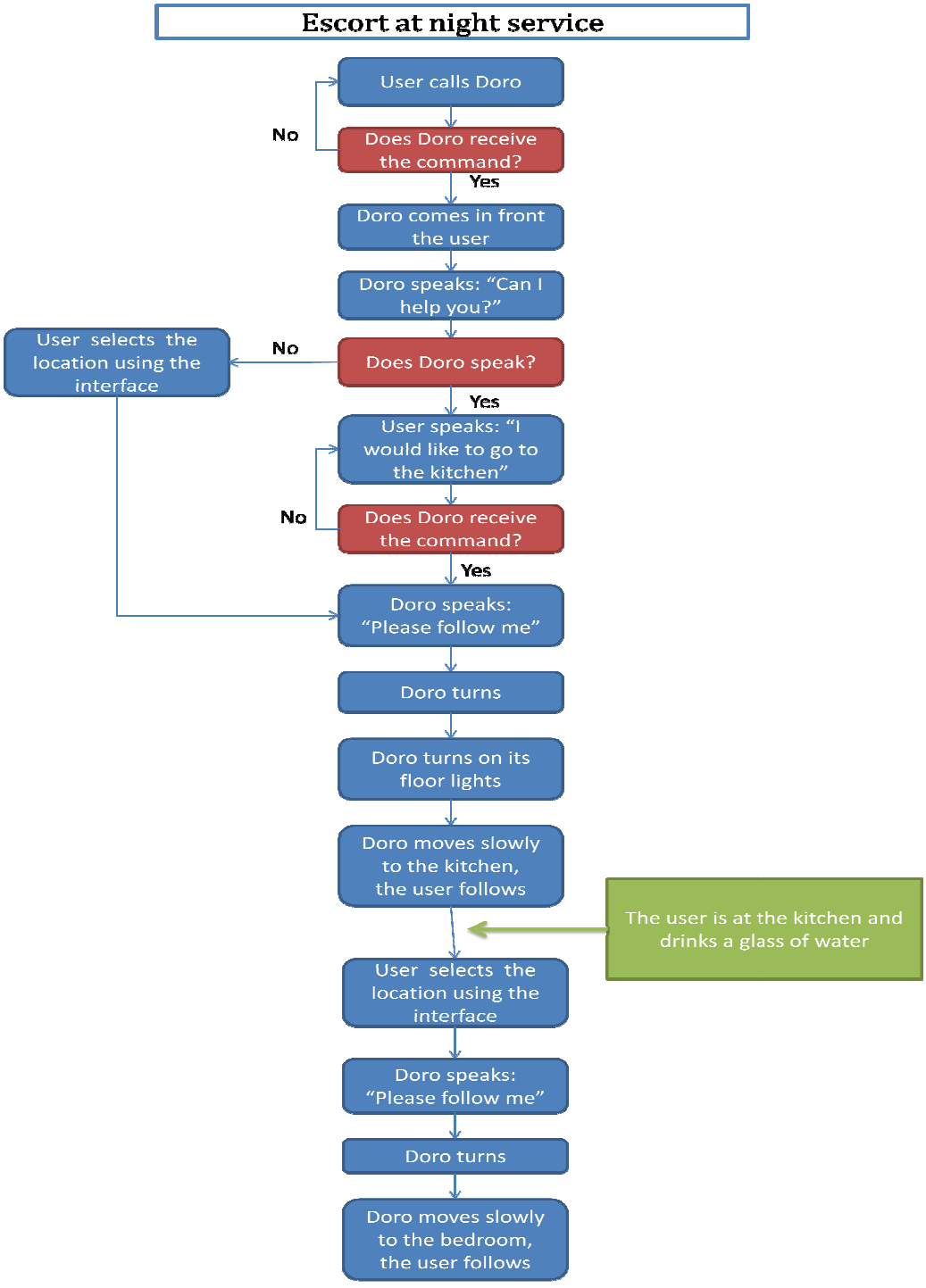


Figure 18.1 Escort at night service flow chart

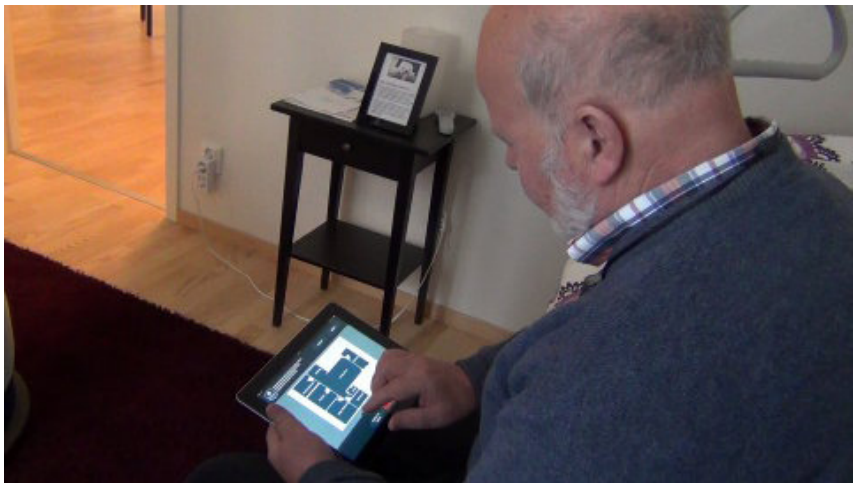


Figure 18.2 User during experimentation

18.1 Usability

Usability SUS Scale

Immediately after the end of the "Escort at night service" , we proposed the system usability scale (SUS) referred to the service.

Table 18.1SUS Score: descriptive statistics

Escort at night	mean	std	#	%
Not Usable 0-64	40,00	17,91	4	57%
Usable 65-84	82,50	-	1	14%
Excellent 85-100	97,50	3,54	2	29%
Total	62,50	31,22	7	100%

The usability of the "Escort at night service" was evaluated as not usable (62,50±31,22) by elderly people. The actions performed by the domestic platform was not perceived as well integrated and easy to use.

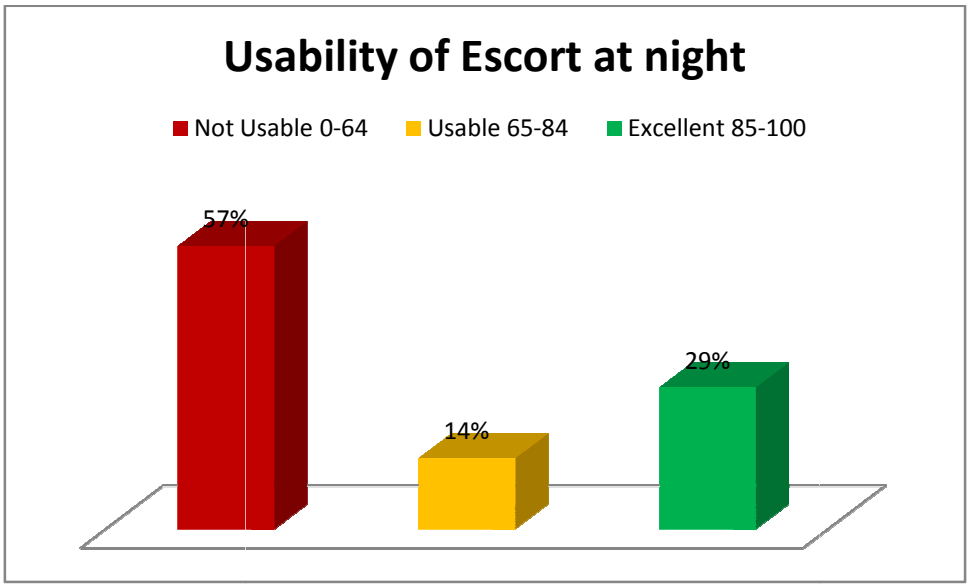


Figure 18.3 Usability result

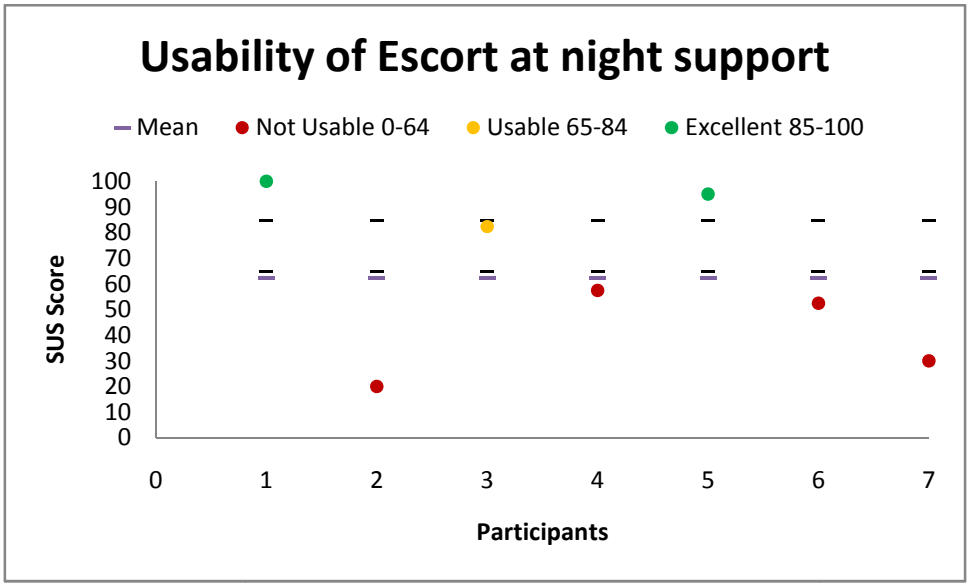


Figure 18.4 Usability result

18.2 Acceptance

18.2.1Attitude

The responses to attitude questions showed a high value (4,39±1,14) towards the "Escort at night service". We have to note that this positive results was due thinking a future use both in case of need and as help for family work.

18.2.2Acceptability

18.2.2.1 Acceptability by ad-hoc questionnaire

The questionnaire compiled by the old persons denoted a positive degree of acceptability (4,13±1,26) about the "Escort at night service". In particular using the domestic robot as a walking support didn't evoke anxiety in the participants.

18.2.2.2 Acceptability by video analysis

Differently from the experimental loop conducted in Peccioli, in Angen there was not a physical contact among robot and human. In fact, all the users walk behind DORO, without using the controller. Just in one case, the user walks beside the robot, touching often DORO. Speech interaction was not used to give command to the robot, the users have controlled DORO just using the tablet, for selecting the room to go to.
The overall impression from the observation is that the users have conducted the task in a very formal way. They have shown to have a good comprehension of the command to be given, but analyzing their facial expression, it could be said that the users did not seem really enjoyed during the service even if they expressed a friendly communication and seem to have a positive attitude to the robot, feeling comfortable with DORO. Just in one case, a user expressed enjoyment through smiling and laughing during the experimental session. Nobody seem to be worried for the performance and no one of the user experienced fear of falling or anxiety during the walking. In general, the users are able to use the tablet, they show to have a high competence with this technology.

18.2.3HRI

HRI by questionnaire: Speech and GUI interaction
The HRI was not evaluated for this service

18.2.4Quality of life

Regarding quality of life, elderly users reported neutral responses (3,67±1,63) to item *"I think my independence would be improved by the use of the robot for this service".*

Table 18.2 Acceptance attributes: descriptive statistics

Escort at night		mean	std	Acceptance Score	std
ACCEPTANCE	Attitude	4,39	1,14	83,33	17,87
	Acceptability	4,13	1,26		
	QoL	3,67	1,63		

For reporting the final acceptance result, we multiplied the sum of each construct score by a factor in order to get an acceptance score that could range from 0-100. From 0-64 points the service was not acceptable, from 65-84 was acceptable and from 85-100 the result was

excellent. At last the "Escort at night service" was acceptable (83,33±17,87) by elderly participants. We can note that 78% of the users the acceptance as excellent (94,18±4,76). However we have to note that this data was not particularly significant because the HRI items were not evaluated.

Table 18.3 Acceptance Score : descriptive statistics

Escort at night	mean	std	#	%
Not Acceptable 0-64	55,00	-	1	16,67%
Acceptable 65-84	67,50	-	1	16,67%
Excellent 85-100	94,38	4,27	4	66,67%
Total	83,33	17,87	6	100%

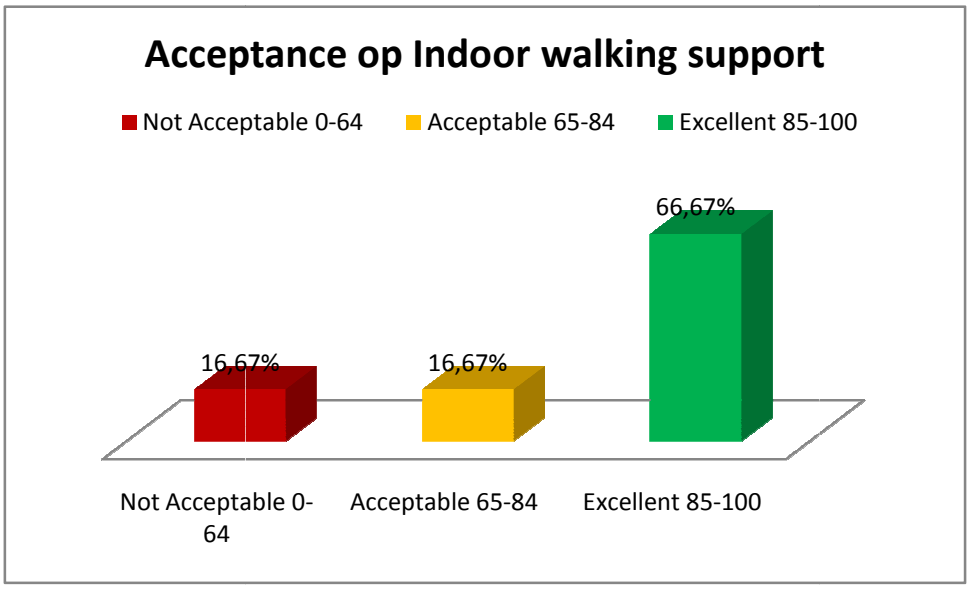


Figure 18.5 Acceptance result

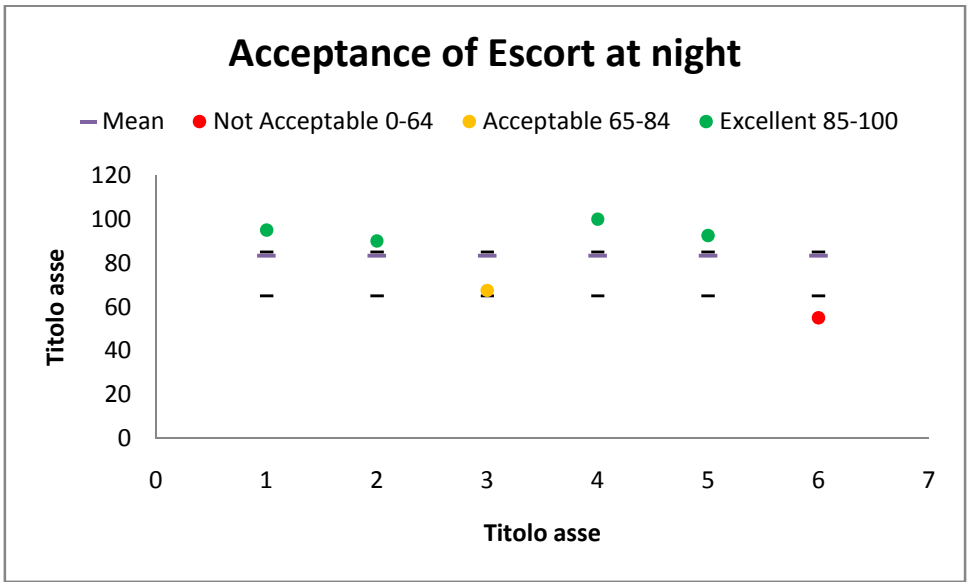


Figure 18.6 Acceptance result

19 Conclusions

19.1 Lessons learnt

Even if the first experimental loop is not yet finished , many factors came to light in order to improve for the second experimental loop. Regarding the HRI during the experimentation the majority of the participants reported to prefer the speech interaction than the GUI one, because according to them the first was more simple to use. Bearing mind this request we should to improve the dialog manager in order to get a natural language because elderly people expected to have a discussion with the robot. However the GUI has always be implemented as option in case of speech interaction interruption. Observing the conducted tests we can assert that elderly persons learned using the tablet interfaces quickly and currently few changes are necessary as adding a "back/return" button, bigger buttons and arranged more separately from each other in order to improve the usability.

With a view to the second experimental loop the user should have more awareness of the services provided by Robot-Era system perceiving the homogeneity and the integration of the several Robot-Era functions. In order to get this more feedbacks and monitoring tools should be had at user's disposal: e.g elderly persons should know the status of each robotic platforms involved in the project as the position, if they are ready for a new task and if not how soon they will be available again. Also after a service selection the user should can monitor the advancing of it and how much time is required to complete it.

At last in order to get a good degree of Robot-Era system acceptance, the robustness of the system should be improved aiming at a good integration and modularity of all components.

19.2 Summary and next steps

Table 19.1: Usability and Acceptance summary

	Users		**Usability**		**Acceptance**	
	Peccioli	Angen	Peccioli	Angen	Peccioli	Angen
Shopping and drug delivery	35		80,86 ±15,81		87,69 ± 8,23	
Communication	35	12	85,57 ± 14,09	65,97 ± 21,89	88,54± 7,94	69,56 ± 7,71
Garbage collection	35		90,14 ± 14,76		89,97 ±8,33	
Reminding	27	9	75,28 ± 23,12	38,06 ± 25,85	87,26± 12,32	78,61± 15,96
Indoor walking support	27		89,70 ± 9,29		88,70 ± 7,12	
Escort at night		6		62,50 ± 31,22		83,33 ± 17,87
Outdoor walking support	24		83,44 ± 12,51		89,01 ± 8,06	
Food delivery		11		70,83 ± 23,78		82,91± 12,16
Laundry delivery		10		75,93± 17,4		84,53± 13,53

The shopping and drug delivery service got a good degree of usability and acceptance even if at the present elderly persons didn't perceived the utility because they had a high level of autonomy. However for the second experimental loop this service could be more temping if the shopping would delivered inside the home as many elderly expected.

The communication service was performed both in Italy and Sweden but the results are discordant because Italian users reported high judgments while the service was barely usable and acceptable for Sweden ones. In the Peccioli pilot site the communication service consisted in a warning and phone call case and the first one was evaluated more useful because both Italian and Sweden elderly persons thought that currently there are many devices to communicate and a robot was only another one. For these reasons in the second experimentation the robot should not be seen as a way to communicate, but as a device integrated in a smart environment which is able to localize the users inside home and to give him an alert message in case of dangerous situation.

The garbage collection service was evaluated very useful by elderly persons who considered this task very boring and they would be glad if the robot performed it in place of them. With a view to the second experimental loop this service could be improved if all three Robot-Era platforms were involved. The garbage bin could be sensored and, after user's confirmation, the domestic robot could take it in order to deliver to the condominium one when a weigh threshold is over. Another option could be that the Robot-Era platforms perform the separate garbage collection on the base of day in an automatic way.

The reminding service was more difficult to use than other Robot-Era ones, in fact the usability score is sufficient in Italy and insufficient in Sweden. However the service was useful according to elderly and acceptance score, but it would be more usable with an improvement of speech interaction.

Initially the Indoor walking support and escort at night services were the same, but then they were tested in a different way in the two experimental pilot sites. In Italy DORO was used by elderly as a support to move from a room to another one and both the usability and acceptance score are high. This service was evaluated as useful in case of motor disease even only momentary. In Sweden DORO was equipped with floor lights and it moved from a room to another one and the user followed it. The usability score is very low while the acceptance one is high.

The outdoor walking support was both usable and acceptable according to elderly persons responses, especially in case of need. For the second experimental loop the service could be improved if the outdoor robot advised the user about obstacles and it was equipped with entertainment functions.

The food delivery and laundry services were barely useful and well acceptable. With a view to the second experimentation the food should be delivered inside home in automatic way and also for the laundry service the process would be automatic.

Addendum

In order to finish the first experimental loop both in Peccioli and Angen we undertake to test all provided Robot-Era services with 35 elderly persons as soon as possible.

References

[1] Lewis, J. R., & Sauro, J. (2009). The factor structure of the system usability scale. In Human Centered Design (pp. 94-103). Springer Berlin Heidelberg.

[2] McLellan, S., Muddimer, A., & Peres, S. C. (2012). The effect of experience on System Usability Scale ratings. Journal of Usability Studies, 7(2), 56-67.

[3] Knoblauch, H., Schnettler, B., Raab, J., & Soeffner, H. G. (2006). Video analysis: methodology and methods. *Qualitative Audiovisual Data Analysis in Sociology, Frankfurt am Main: Lang*.

[4] Mayring, P. (2000, June). Qualitative content analysis. In Forum: Qualitative social research (Vol. 1, No. 2).

[5] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 27(3).

[6] Bartneck, C., & Forlizzi, J. (2004, September). A design-centred framework for social human-robot interaction. In Proceedings of Ro-Man (pp. 591-594).