ICT-2009-248730

Florence

Multi Purpose Mobile Robot for Ambient Assisted Living

STREP
Contract Nr: 248730

Deliverable: D1.4 Report on Wizard-of-Oz experiment

Due date of deliverable: (31-01-2011)
Actual submission date: (14-02-2011)

Start date of Project: 01 February 2010 Duration: 36 months

Responsible WP: WP1 FASS

Revision: final

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 Services)	
CO	Confidential, only for members of the consortium (excluding the Commission Services)	

0 DOCUMENT INFO

0.1 Author

Author	Company	E-mail
Victor Manuel Hernandez	FASS	victorm.hernandez@juntaandalucia.es
Melvin Isken	OFFIS	melvin.isken@offis.de
Dirk-Jan van Dijk	Novay	dirk-jan.vandijk@novay.nl
Henri ter Hofte	NOVAY	henri.terhofte@novay.nl

0.2 Documents history

Document version #	Date	Change
V0.1	2011/01/07	Starting version, template
V0.2	2011/01/10	Definition of ToC
V0.3	2011/01/20	First complete draft
V0.4	2011/01/26	Integrated version (send to WP members)
V0.5	2011/01/28	Updated version (send to project internal reviewers)
Sign off	2011/02/11	Signed off version (for approval to PMT members)
V1.0	2011/02/14	Approved version to be submitted to EU

0.3 Document data

Keywords	Wizard of Oz, human technology interaction
Editor Address data	Name: Pablo Quinones Delgado Partner: FASS Address: C/ Marie Curie, 6. Parque Tecnológico de Andalucía . 29590 Campanillas (Málaga) Phone: +34 946 430 850 E-mail: pablo.quinones@juntaandalucia.es
Delivery date	14-02-2011

0.4 Distribution list

Date	Issue	E-mailer
	Consortium members	al_florence_all@natlab.research.philips.com
	Project Officer	Luiz.Santos@ec.europa.eu
	EC Archive	INFSO-ICT-248730@ec.europa.eu

Table of Contents

0 DOCUMENT INFO 2
0.1 Author 2
0.2 Documents history 2
0.3 Document data 2
0.4 Distribution list 2
1 INTRODUCTION 4
2 WIZARD OF OZ CONCEPT AND ITS USE IN HUMAN INTERACTION RESEARCH 5
3 IMPLEMENTATION OF THE WIZARD OF OZ EXPERIMENT 7
3.1 User Interaction Aspects 7
3.2 Germany 7
3.3 The Netherlands 8
3.4 Spain 8
4 SOFTWARE AND HARDWARE REQUIREMENTS 11
4.1 Hardware 11
4.2 Software 11
4.3 Users Profile 12
4.4 Procedure 12
5 SUMMARY FROM THE WOOZ SESSIONS 13
5.1 Tests in Germany 13
5.2 Tests in Netherlands 15
5.3 Tests in Spain 18
6 CONCLUSION 25
7 REFERENCES 26
8 ANNEX 27
8.1 Information & Informed Consent - Netherlands 27
8.2 Information & Informed Consent - Spain 28
8.3 Information & Informed Consent - Germany 31
8.4 Study procedure - The Netherlands 33
8.5 Instruction Slides – Netherlands 36
8.6 Study procedure - Germany 40
8.7 Questionnaire 42

1 Introduction

Within the scenario re-definition targeted for D.1.3 the Wizard of Oz test was the final stage in order to deliver results in terms of the final robot based scenarios.[1]

Regarding the generation of user's requirements and services, user's opinions were not enough. There was an actual need to capture and analyse the interaction between the human and the system. Even having iterations in the project later on, the costs of building a proper prototype and test it in order to modify it iteratively was high, therefore the Florence Consortium saw the opportunity to use the Wizard of Oz method in order to anticipate some issues. In a Wizard of Oz test, the robot is not autonomous and but is remotely controlled by a technician (wizard) leading the interaction. Therefore, the development needed was minimal and the result was close to what a robot should be. This test set a meeting point for the technical and user organization partners since both groups have been involved in this process with the aim of polishing the final requirements and, at the same time, to define concrete interactions between technology and users.

[1] The final scenarios can be found in Deliverable D1.3

2 Wizard of Oz concept and its use in Human Interaction Research

This test is normally framed epistemologically in the field of human-computer interaction. In general, Wizard of Oz test can be described as a "technique that enables unimplemented technology to be evaluated by using a human to simulate the response of a system. The 'wizard' observes the user's actions, and simulates the system's responses in real-time. […] Often users are unaware (until after the experiment) that the system was not real."[2] This iterative design methodology simulates the behavior of a theoretical intelligent computer application. This allows testing of some difficult interface concepts before a system is fully working. The origin of this test dates back to 1980 at John Hopkins University by Jeff Kelley (Libin, 2004)

The participant in this case thinks he is actually interacting with the Robot interface (screen, video or voice) without the help of any "middle man" (Wizard), whereas in reality the main protagonist is the Wizard who controls the Robot in order to simulate the interaction with the user. In some cases the user noticed that the robot was being manipulated by someone. However, we estimate that this does not really affect the outcome.

The Wizard of Oz test adds a number of benefits to the Florence Project. This test encounters actual human responses concerning an artificial interaction. It also enables and effective interaction with a system before it is actually built, therefore this experiment can save money and effort.

In our case, we can see that we have been able to test a future technology without testing an expensive system. As shown in Table 5-1, some important aspects were revealed in an early stage of the project that otherwise would not have been revealed. Some challenges presented by these methods were the following:

- Wizard simulations require significant training so the wizard can respond in a way that is credible.
    - Involving and training a Wizard is an additional resource cost.
- It is difficult for wizards to provide consistent responses across sessions.
    - Thus, proper program code, or 'behaviour instruction' should be prepared and given to the wizard.
    - This 'Behaviour instruction' should not describe every single reaction, but try to control predictable and typical situation, and guide the session to answer the target questions
- If a research team member plays the role of Wizard, there is a risk that they will improvise beyond the programmed behaviour.
    - To avoid this risk, hire someone who can be instructed [programmed] with simple rules and play as a wizard.
- Computers respond differently than humans so the wizard needs to match how a computer might respond (for example, the Wizard should not make typing errors).

[2] http://amsterdamlivinglab.novay.nl/learn/techniques/wizard-oz

- Playing the wizard can be exhausting, meaning the wizard's reaction may change over time, mainly due to cognitive fatigue.
- It is difficult to evaluate systems with a large graphical interface element since scope for possible actions, impressions, etc is much larger.
- This approach does not uncover errors that arise as a result of system performance and recognition rates (unless these are specifically simulated), so it is more effective in revealing problems than predicting real world usability.

3 Implementation of the Wizard of Oz Experiment

The key aspects of the Florence Wizard of Oz experiment were decided after some coordination meetings. Work Package 1 partners decided that instead of focusing on scenarios use case functionalities as done for focus groups, more generic aspects would be analyzed. Nonetheless, those scenario use cases would be experienced and simulated. The interaction to be tested, consisted of common characteristics concerning usability, acceptance and functionality, therefore scenario use cases were the context in which interactions would be triggered and developed but at the end, what was being tested were the actual aspects involved in the interactions, such as, approaching, noticing, awareness, etc.

3.1 User Interaction Aspects

This section explains the aspects tested by this experiment. We give an account of each country separately, since every partner in charge of different use cases was able to point out the main aspects to test, regarding a simulated interaction.

3.2 Germany

3.2.1 Feasible aspects to test in WoOz experiments

It was impossible to test all scenarios therefore we decided to have a minimal set of open questions that should be answered with the experiments. They don't strictly relate to a single scenario, most aspects will provide information for more / all of the scenarios. These are the main functional blocks that we wanted to test. The robot's current battery runtime was also an inconvenience to test more aspects. We didn't want to connect it directly to scenarios. Things that will be used quite often and are easy to implement like video conferencing were tested as well. Videoconferencing was left to the end of the session because it did not require the robot moving any longer and would thus provide more running time for other activities.

Approaching the person is of fundamental interest for interaction with the robot. This is about how near the robot should come, which speed is used to approach the person and so on. Another question would be in case of a lying person (FALHAN) which side the robot should approach, either the head or some other area that might be not as threatening and frightening as the head region.

Following the person is another important behavior of the robot since such movements may disturb or annoy the user. We had two principles in mind, the first one was direct following and the second one was intermittent following. Direct following means the robot tries to keep a certain distance to the user. If the user moves one step, the robot directly reacts and also moves. This behavior could be annoying if the user just wants to move inside the room and the robot should wait in one corner. The intermittent following will try to solve this problem by keeping a variable distance to the user. So the robot will only begin to move if the user leaves a certain area or gets out of sight of the robot.

If the robot is accidentally in the way of the person and the robot has to move aside, different movement patterns can be considered to use. The robot could quickly move to a safe corner, just move one step backwards or try to predict the user's movement and go along that path. During the movement it is a possibility to play sounds to alert the user that the robot is standing in the way. Most of these actions will probably happen in corridors with not much space for the robot to escape.

3.3 The Netherlands

The LIFIMP (Lifestyle Improvement) scenario formed the context for the Wizard of Oz tests in the Netherlands. Before the Wizard of Oz tests started, it was decided in the project to remove the DEVCOA (device coach) scenario from the project (see also Florence deliverable D1.3). Nevertheless, various aspects relevant for the DEVCOA scenario were still present in the Wizard of Oz experiment executed in the Netherlands.

In the LIFIMP (Lifestyle Improvement) scenario, the user may be standing (doing an exercise), sitting (at the breakfast table, looking at the health stats and/or receiving a suggestion to go for a walk), or even lying (performing a different exercise). We wanted to find out the usability while interacting from these different positions, in particular:

- Is the screen visible and readable from these different positions?
- Is operating a touch screen usable from all these positions (if at all necessary?)

In the LIFIMP scenario, the robot needs to get out of the way for the exercises, but still be viewable for the exercise videos. It is highly likely that the user needs to influence the robot’s orientation / help the robot orient itself properly.

Both in the DEVCOA scenario and in the LIFIMP scenario videos are played on the robot screen while the hands of the user may be occupied. We wanted to test to what extent voice control of video play out (“start”/”stop”) is required, or whether play out control via touch screen suffices.

Ultimately, in the Netherlands, we decided to structure our Wizard of Oz experiments as follows:

1. Using *voice only*, let a person *seated* at a kitchen table instruct the robot to *approach and orient* itself towards the person, such that the person can see a message on the screen and interact with the touch screen.
2. Same as previous, now *using gestures only* (note: to attract the attention of the robot and to face the camera to receive instructions, the person would still have to call the robot by its name).
3. Using either voice or gestures let a person instruct the robot to *follow* a person and *orient* itself properly such that the person *standing* would be able to see an exercise video played on the screen.
4. Using *voice and-or touch*, let a person *standing* control *play and pause* an exercise video on the robot´s screen.

So, we covered aspects from table 3-1. The aspects we did not test in the Netherlands were steer, tag and lie.

For the complete study procedure used in the Netherlands, we refer to section 8.4.

3.4 Spain

For the use cases tested in Spain, AGEREM (Agenda Reminder) and FALHAN (Fall Handling), we thought of the following aspects involving the actions carried out.

In both scenarios the robot has to approach the user taking into account the different reasons that triggered the approaches. The approach for an emergency such as a fall may be desired to be different to the approach for a mere reminder. Within this aspect we detected different sub categories such as speed, continuity, etc.

Secondly, we acknowledged the aspect of awareness. This is a bi-directional aspect which implies how to make the user aware that the robot is realizing an action or accomplishing a task when he does not expect it, and on the other hand how to make the user aware of the robots actions when he does expect it.

There were different sub-categories planned to be tested such as melodies, screen displays, alarms, etc.

Since situations like a fall imply to face unusual positions to interact with the robot, this category was very important from us. We wanted to know what the most common positions were for our use cases and at the same time to see the problematic in some unusual positions to interact with the robot. This interaction also refers to the distance between the user and the robot for an optimum performance.

Regarding the AGEREM use case, the display and an easy way to confirm the intake of the medicine are crucial for success.
Finally a crucial and transversal aspect was security. This aspect overlaps parts of other many actions and even other aspects such as movement, following, etc. We can say that it is a dimension in which every variable will be tested.

Table 3-1 showing summary of the aspects tested[3]

WoOz aspect categories	short explanation	Option	option	option
move: approach / follow / steer	test(s) about the movement of the robot approach = test(s) how the robot should approach you follow = test(s) how the robot should follow you (or avoid you) steer = test(s) how you can steer the robot directly (if possible at all)	approach + follow	follow + steer	approach + steer
orientation: gesture / tag / steer	test(s) how the robot should orient autonomous = no influence, the robot orients itself gesture = test(s) with orienting the robot by gesturing (with your own hand, arm, body) tag = test(s) with orienting the robot by holding some easily recognizable tag	autonomous + gesture	autonomous + tag	gesture + tag
Interaction: stand / sit / lie	stand = test(s) interacting with the robot while standing sit = test(s) interacting with the robot while sitting lie = test(s) interacting with the robot while lying	stand + sit	sit + lie	stand + lie
Interrupt	test(s) how the robot should interrupt you	Melody	Alarm	Voice

[3] This table was made by all the Partners implementing the WoOz in order to have a unified map of the aspects tested with the different alternatives. It was also a way to allocate all aspects so that we were able to cover them all.

4 Software and Hardware Requirements

4.1 Hardware

The Florence system hardware used in the WoOz experiment is realized to a level that would represent the final system. This includes basic interaction modalities (movement, size, appearance, audio input/output, video input/output, and touch input) .A general hardware setup was agreed on. This consisted of the Pekee base, a stand holding a tablet PC and a webcam to interact with the user. The actual setup slightly differed between the sites. The tablet PC was not chosen by the consortium at the time of conducting the test so every partner chose hardware that was available. An overview is given in Table 4-1.

	OFFIS	NOVAY	FASS
Robot	Pekee II	Pekee II	Pekee II
Power	Via cable	Batteries	Cable
Stand	Plastic tube, stabilized with strings	easel, stabilized with tensioned nylon cords	Plastic tube,
Tablet	Apple iPad	Acer Aspire Timeline 1825PTZ-412G25N	Paceblade EasyPad E12
Tablet weight	0.7 kg	1.7 kg	1.8
Screen size	24.6 cm	29.4 cm	28.5
Screen resolution	1024x768	1366x768	1336X788
Touch type	Multitouch	Capacitive	Multitouch
Webcam	2 (one for control, one for interaction)	2 (one for control, one for interaction)	2 (one for control, one for interaction)

Table 4-1 Hardware setup

4.2 Software

The Florence system software is mostly faked[4], but to a participant, it seems the system is working and reacting, but actually, there's a human “implementing” the system's behaviour, i.e.: observing the interaction between participant and hardware and operating the robot. The following list sums up the software needed for the implementation of the Wizard of Oz test.

1. Software to play Audio and VOIP software (Videoconferencing. e.g. Skype)
2. Software to control the movement of the robot and orientation of the camera
3. Software to control and see what's displayed on the touch screen (e.g. VNC)

[4] E.g. use of a Power Point presentation to simulate the main functional interface of the system.

4. PowerPoint to display mock-ups of the user interface for selection of the make and model of the problematic device
5. Software that allows the wizard to see where the user touched the display
6. Ability to record audio and video streams from the Florence robot. Ability to record audio and video stream from an overview camera and microphone in the room (to review the session for later analysis).

4.3 Users Profile

The user's profile agreed was people from 59 to 75 years old in an even proportion male-female (Germany 50% Male 50% Female, Spain 40% Male, 60% Female, Netherlands 50% Male 50% Female) with some degree of technology acceptance and some degree on computer use. Their physical state had to be relatively good, people with serious cognitive impairments or illnesses were not considered to be selected.

4.4 Procedure

The test was done from November to December of 2010 within the living labs at OFFIS, Novay or FASS. In the sites with no Living Lab like FASS a living room was simulated.

The stakeholders of this test were: a participant who always was an elderly person (one at the time), a researcher/observer who fills the evaluation form, a person who is leading the session, explaining the experiments, and a wizard consisting of a person preferably out of sight from the participant that observes the user's actions via one-way mirror and/or audio/video/screen feeding, and simulating the system's responses in real-time.

The duration was about 40-45 minutes per participant. robot was active for about 30 minutes per participant (time was limited by the current battery autonomy of the robot).The following table shows a sample of a detailed procedure accomplished by Offis.

It is important to remark that after welcoming the participants, an information and consent form was filled (see annex I, II and III) which allowed the researchers to use the obtained data for investigative purpose. It also let us to record the audio while the experiment was taking place. Information about the project was given during the recruitment mainly telephonically, but it was also included and read out as an introduction of the consent form.

Appendix 8.6 shows the procedure implemented in Germany in detail, it was written as it was a script for the actual test even though there was some scope of flexibility.

5 Summary from the WoOz Sessions

In the following section a summary of the three Wizard of Oz tests carried out in Germany, Netherlands and Spain, as well as comments from users will be documented.

5.1 Tests in Germany

The Wizard of Oz tests in Germany were carried out in the IDEAAL living lab at OFFIS. Five elderly participated in the tests, four of them having already participated in the Focus Group sessions. During the tests the following spaces were used: the living room, kitchen and corridor. The Wizard was sitting in the bedroom controlling the robot. The questionnaires and most of the following test were conducted in the living room area.

Before the sessions, a questionnaire was conducted (see annex, section 8.7) in which the users were generally more concerned about robot features such as “should be small enough to fit under table”. In general it was seen as highly technically advance device, which would probably be rather complex.

After working with the robot this opinion was completely turned around. The robot was perceived as easily controllable and had a quite nice appearance for a prototype system (see Figure 5-1). The results regarding the different interactivity sessions are presented in the following paragraphs.

Figure 5-1: Pekee Design for WoOz tests Germany

Results from the WoOz test helped in concluding that after the robot approached the user a comfort distance of 50-60 cm should be left between the user and the Florence robot. Regarding the direction of approach (see Figure 5-2), some users felt more comfortable with front-wise approaching, while others preferred side-wise approach. There was no clear conclusion. So it should be up to the user to command the robot to a place he or she likes and feels comfortable with. The Florence robot should be able to learn over time how to interact in this and other ways with the user, memorizing their preferred choices.

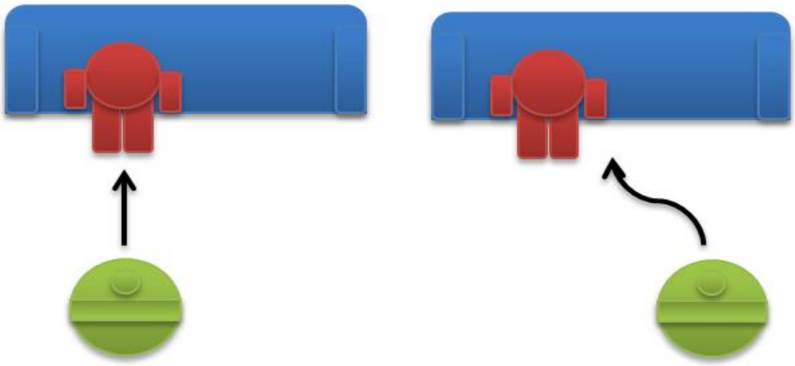


Figure 5-2: Modes of approach

Direct following and indirect following were also tested. Direct following meant that the robot continuously followed the user. Indirect following let the robot only follow if the user got out of sight. Users preferred indirect following; as they found direct (non-stop) following annoying. The users were impressed that the robot could find them behind a corner. The scenario itself was targeted at carrying things to the user.

Afterwards the user was ordered to go from the living room to the kitchen and the robot stood in the middle of the way. Two different tests were performed at this point: (i) the robot moved backwards (user is forcing robot to go back), (ii) the robot tried to get out of the way as fast as possible (turn robot a bit and go back). The users preferred the immediate “out of the way” strategy; the robot did not bother so much with this option. If the robot is just moving back in the direction of the user, the user has to walk really slowly.

For the next aspect to be tested the user forced the robot into a place where it couldn’t move anymore. Two strategies were tested: (i) the robot does not move anymore and stands still; (ii) the robot turns left and right to signal that it doesn’t know where to go. The users strongly preferred the second solution, because if the robot just stood still there was no way for them to recognize if the robot was working at all. So the robot should somehow show and give signs of operation by for example, trying to find a way to escape from that physical space even if there is no way out. This attitude helped the user understand and be aware of the situation.

During the different tests, at times, users tended to naturally try to stop the robot by just kicking the bumper gently. This seemed to be an intuitive way for an emergency stop that should be implemented in Florence.

The users liked the voice control features very much. They didn’t need much time to learn the commands and the robot reacted as expected. Among the commands used, left and right were included for controlled navigation. If the robot looked at them, it was sometimes confusing using left or right as a mirrored effect is introduced.

Comments regarding voice control should be carefully considered; the voice control of the robot was really good because no algorithms were behind the recognition and the human wizard is able to understand all kinds of command variations and does nearly always understand what the user wants. Variations on voices were also tested (man/woman, young/old etc.), opinions diverged considerably and therefore conclude this should remain optional, user should be able to choose.

Even if voice control was working extremely well, a touch screen would also be desired. The users seemed to still have more confidence in a system that is controlled by pressing buttons than by controlling it completely via voice.

The touch screen itself was very easily readable. On the touch screen there should also be some kind of robot steering functions, in order to control the robot's navigation if voice or remote control were to fail. The use of some kind of simple smiley to show emotions on the touch screen was considered to be good.

During the touch screen interactions the following was concluded:

- Messages and notifications should stay on the screen until they have been noticed
- The height (approx. 1.4 meters) was nearly perfect while sitting on the couch. If the user standing up, it could be a little higher: the stand should be height adjustable.
- The robot itself seemed to be a little big compared to the size of the room (the robot was very dominant). This may derive from the relatively small living room (about $10m^2$).

Users also commented that a tablet for carrying stuff would be nice to have. Some users indicated the display could be used as tablet whenever the touch screen is not being used (move tablet into a horizontal position). Again, a height adjustable stand would be useful in this case. If the stand is shaking like it was in WoOz tests, a bag would be better than a tablet.

If the robot is in its charging station, it should indicate that it is still turned on and working. Users indicated they would like to now in all circumstances whether the robot is inactive or inoperative.

Some users also suggested using the robot to clean the floor. Functionality that would allow the robot climbing stairs was also commented as most of the users have stairs at their homes.

In general, the users were impressed by the features of the robot (especially the voice control), before knowing that this functionality was faked. The overall (prototype) design was accepted. Regarding the movement, nearly all of our predicted behaviours were considered to be good.

5.2 Tests in Netherlands

In the Netherlands, 6 different aspects were tested with elderly users (aged 62-78) in the Wizard of Oz experiment with interviews before and after interacting with the robot. Users arrived in pairs (they were reluctant to come alone), and interactive tasks were done individually. All users had participated earlier in the focus groups. The interviews were carried out in order to obtain more information on robotic appearance preferences. This aspect was initially not to be considered in the wizard of oz tests, but was afterward included to comply with the European Commission request of special focus on the appearance which will consecutively increase the user acceptance. In the annex (section 8.4) we can see the graphic presentation that introduced the tasks to users, followed by the questionnaire in English about the appearance (section 8.7).

Before meeting the robot, the mobile robots without screen and pet-style look were the most preferred options[5]. After meeting and interacting with the robot, however, the mobile robot with screen and face (before meeting more disapproved of than approved), was the most favourite option. We believe that giving the robot a friendly face has a very high value. Two out of six users even gave the robot a compliment after having completed their first instruction: “well done”, with an intonation much like you would praise a dog or child.

[5] To go into more detail, results are shown in the next section

Figure 5-3: Florence robot as used in the Wizard of Oz test in the Netherlands

Two users commented that the looks of the robot should be a bit taller, like the sketch we've shown them, as indicated in the figure below.

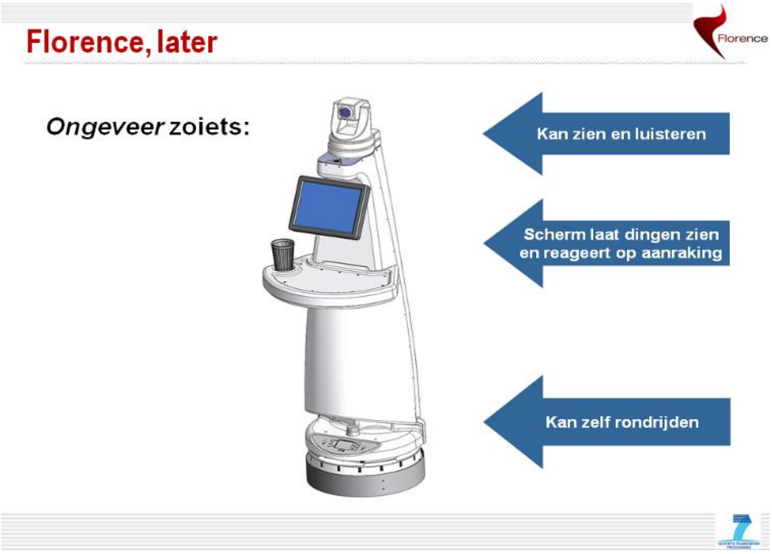


Figure 5-4 The sketch of the Florence we've also to the Dutch participants

With respect to movement, hidden wheels were the most preferred option; visible wheels were deemed acceptable, both before and after meeting the robot.

With respect to controlling movement, users mostly preferred giving the robot voice commands, both before meeting the robot, and shortly after interacting with the robot. When asked beforehand, on average users did not like the idea that the robot would follow them through the house. After having given a voice command or gesture however, no user mentioned he or she did not like the following.

With respect to controlling the robot in general, before meeting and interacting with the robot, a touch screen, buttons on the robot and voice control were all considered acceptable, whereas using a remote control was considered to be a bit unacceptable.

Our robot would not do anything until called by name "Florence", and then the robot would turn towards the speaker and make a short noise indicating it was ready to receive a command, which could be voice, gesture or both, depending on the mode that was set on the robot. After having tried voice control, gesture control and touch screen control, most users clearly preferred giving voice commands. Gestures also worked for users, and were reported to be easy to learn but voice was preferred by all except one user (who felt more “in control” by giving gesture commands). Most users would not object to having both gestures and voice as an option, but also most users said they would not mind if we were to remove the gesture option. However, when asked a few situations came up where gestures might be needed, e.g. `when you have a cold´. All these tasks can be found in a Power Point presentation included in annex, section 8.4.

Voice should certainly not be removed, e.g. in situations when a user has fallen down.

Participants learnt the short voice commands (“here”, “stop”, “follow”, and “go”, “play”, “pause”) very quickly and 4 out of 6 users mentioned in the interview afterwards that the robot listened very well. Three users mentioned that the short commands sounded a bit “dominant”, but that ultimately didn’t seem to be a problem. One user noticed that her voice was faint and she had to raise her voice a bit. Whether this is a problem in day to day usage remains to be seen. Voice control would have to be fine-tuned in two respects: the “stop” command seemed urgent enough for users to forget prefixing it with a name; therefore it should be available even without prefixing it with the name of the robot. Also, users did not always wait for the confirmation tone before proceeding with the command. Whether this is a problem in a real voice recognition situation remains to be seen.

The screen (11.6 inch) was considered large enough for images to be visible and text (a short sentence consisting of 4-8 words) to be readable. We encountered only one situation where low sunlight seemed to interfere with readability. We also observed one user having to bend town a bit to read what was on the screen. It would be good if the screen could tilt up a little to be better readable in those situations. Moreover, the users seemed to like the functions a screen can provide (show a face / facial expression, show messages, play video, confirm messages by a gentle touch), even to the extent that they rejected the pet-like look afterwards (‘How could you mount a screen in that? You would need a very large dog.”).

Using touch to confirm messages worked fine, but in the beginning one user tended to press too long. Another user with arthritis mentioned that she very much liked she did not have to press hard as her arthritis made operating some other devices like phones and ATMs painful, but not the Florence robot. Note that we used a capacitive, not a resistive touch screen. Nevertheless, while playing the exercise video, most users chose voice commands rather than touch, which save them walking up to the robot and back.

All in all, the users seemed very engaged in interacting with the robot. Users were amazed by how well the robot listened and executed tasks. Of course, being a Wizard-of-Oz test, we know that we have a challenge up ahead implementing the voice recognition. However, experience seems to indicate that a few short commands may go a long way.

Users were underwhelmed by the battery life of the robot (since we had to walk with an extension cord plugged into the robot most of the time), but we assured them that of course, in the final version the battery problem would be solved.

5.3 Tests in Spain

The Wizard of Oz carried out by FASS in Spain, took place in Seville where the FASS telecare centre is located. FASS tested the prototype with six users. Three of the users had previously taken part in the focus group held last September. All users interacted with the robot individually even though 2 of them came accompanied (son and friend).

Tests were scheduled every hour. The basic summary of each session was the following:

1. Welcoming
2. Explaining the project and signing the consent form
3. First Questionnaire (before showing the prototype)[6]
4. Showing the robot
5. First task
6. Second task
7. Second Questionnaire
8. Additions from the user

The speed of the prototype was considered to be acceptable for four users whereas other two pointed out that it might be too slow. Three users stated that the function of the robot following them may be disturbing whereas other three said they would not mind as long as they can activate or deactivate it.

There was a wide consensus about the rejection of an autonomous orientation of the robot, keeping the robot functions under user control was a constant request throughout the tests. Regarding the use of gestures or voice, all users were rather positive on both, with preference on voice command option.

After the users were shown the current prototype, they were invited to go to the couch and sit down while the robot would approach them. A greeting was displayed on the screen and the AGEREM scenario was simulated with a constant voice interaction between the user and the researchers in order to test different aspects concerning usability. New ideas were suggested about what to remind and what to display. Valuable contributions were made regarding the desire of two users to have the daily news displayed first thing in the morning along with the usual reminders.

Secondly, we proceeded with the simulation of the second scenario, FALHAN. The simulation as such was rather simple to accomplish hence the feedback generated by users was rich. The user-robot interaction from the floor in an emergency case was a sensitive matter. The average distance agreed from the robot to the user was about 1 meter since the user would feel safer and the communicative channels seemed to be clear enough. Finally the user filled the second part of the questionnaire leaving some time for suggestions, general impressions, etc.

5.4 Results

This section consists of the following table (Table 5-1) which conveys the main results commented by every partner. Results are classified by Partner and Aspect in order to give coherence to all the stages of this test.

[6] To see the actual questionnaire you can find it in ANNEX, section 8.7

Table 5-1: Main comments from participants

Key Aspect	Sub-aspect	Comment	Partner
Movement	Approach	The approach movement of the prototype has been seen by our users as convenient in terms of speed. Taking this speed, users see no problem in having a constant and steady movement. A remark was made thinking about emergency handling, where the robot may be faster	FASS
Movement	Approach	With respect to controlling movement, users mostly preferred giving the robot voice commands.	NOVAY
Movement	Approach	After approaching, the distance to the screen should be around 50-60cm for maximum comfort. Regarding the direction of approaching, some users felt more comfortable with front-wise approaching, others preferred side-wise approach.	OFFIS
Movement	Follow	This aspect should be up to the user depending on the functionality. Most of users saw this as annoying. The underlying affirmation is that users want to have the control of the robot all the time	FASS
Movement	Follow	When asked beforehand, on average users did not like the idea that the robot would follow them through the house. However, after having taken the initiative to give a `follow me` voice command or gesture, no user mentioned he or she did not like the following.	NOVAY
Movement	Follow	Direct following and indirect following was tested. The user preferred indirect following; it is not as annoying as direct (non-stop) following. The user was impressed that the robot can find them behind a corner. The scenario itself was predestined for the robot carrying things for the user.	OFFIS
Movement	General	Before meeting and interacting with the robot, a touch screen, buttons on the robot and voice control were all considered acceptable, whereas using	NOVAY

		a remote control was considered a bit unacceptable. After having tried voice control, gesture control and touch screen control, most users clearly preferred giving voice commands. However, when asked a few situations came up where gestures might be needed, e.g. `when you have a cold´. Voice should certainly not be removed, e.g. in situations when a user has fallen down. All voce and gesture commands always had to be preceded by calling the robot by her name: “Florence”. Participants learned the short voice commands (“here”, “stop”, “follow”, and “go”, “play”, “pause”) very quickly and 4 out of six users mentioned in the interview afterwards that the robot listened very well. Voice control would have to be fine-tuned in two respects: the “stop” command seemed urgent enough for users to forget prefixing it with a name; therefore it should be available even without prefixing it with the name of the robot. Also, users did not always wait for the confirmation tone before proceeding with the command.	
Movement	Steer	We did not test direct steering.	NOVAY
Movement	Avoiding	Two tests were performed, one with the robot just moving backwards in direction of the user (user is forcing robot to go back) the other one with the robot trying to get out of the way as fast as possible. The user preferred the immediate “out of the way” strategy, the robot does not bother so much	OFFIS
Movement	Avoiding	The user forces the robot into a place where it can’t move anymore. Two strategies were tested: 1. the robot does not move anymore and stands still. 2. The robot turns left and right to signal that it don’t know where to go The user strongly preferred the second solution, if the robot just stands still you can’t recognize if the robot is working at all.	OFFIS

Movement	General	Sometimes, user tend to stop the robot by just kicking the bumper gently	OFFIS
Orientation	Autonomous	Unless is necessary or it is justified by some reason (Domotic Checking or security Checks) users disliked this orientation option. Again the control affirmation becomes relevant.	FASS
Orientation	Autonomous	We did not test autonomous orientation; our robot would never move nor orient by itself, only by command. This even applied when notifying the user of a message: a sound would be provided, but the robot would only move after being called.	NOVAY
Orientation	Voice	The only way to orient our robot was to call its name, then it would orient towards the speaker. This worked fine.	NOVAY
Orientation	Gesture	All of them agreed that this is a very good idea. Controlling the robot by some gesticulation has been very popular, but even more predominant has been the voice option.	FASS
Orientation	Gesture	We did not test orientation (i.e. where it is facing) by gesture, our robot would only orient itself towards the user if the user would call it by name. For the rest, see movement.	NOVAY
Orientation	Tag	The most popular tag has been a bracelet even more popular than a pendant (broadly used in tele-assistance)	FASS
Orientation	Tag	We did not test this option	NOVAY
Interaction	Sitting	The main problem was the distance between the arm and the screen, some user suggested that the robot/screen should be oriented on the side of the chair/couch. The height of the screen must definitely be adaptable	FASS
Interaction	Sitting	The height of our screen (about 1 m) and distance was fine for interacting while sitting at a kitchen table on a kitchen chair; both for observing what's at the screen and for touching the screen. The screen (11.6 inch) was	NOVAY

		considered large enough for images to be visible and text (a short sentence consisting of 4-8 words) to be readable. Using touch to confirm messages worked fine, but in the beginning one user tended to press too long. Using a capacitive touch screen seems to make touch operation light enough even for users with arthritis. We reckon that a short one-time on-screen tutorial about touching might be useful.	
Interaction	Standing	While playing the exercise video, most users chose voice commands rather than touch, which save them walking up to the robot and back. We also observed users may have to bend town a bit to read what's on the screen. It would be good if the screen could tilt up a little to be better readable in those situations. The screen (11.6 inch) was considered large enough for video to be visible from about 2 m distance.	NOVAY
Interaction	Lying	The main concern was to keep the distance in order to have a better video interaction. Camera movement has been proven to be essential in order to focus on the user. The stand must be adaptable. The main way to communicate from the lying position should be through voice.	FASS
Interaction	Voice control	The users liked the voice control features very much. They didn't need much time to learn the commands. The robot reacted as expected. If the robot looks at you, it is sometimes confusing using left or right because that is mirrored. Even voice control is working; a touch screen is desired as well. Regarding the relationship to the robot the type of voice is also important (man/woman etc.).	OFFIS
Interaction	Touch screen	The touch screen was good to read. On the touch screen there should also be some kind of robot steering functions, in order to control the robot if voice or remote control don't work. The use of some kind of smiley to show emotions	OFFIS

		was considered to be good.	
Remote Pointing			
Interrupt/ Awareness		Depending on the situation, for routinely interruptions they prefer a low profile melody whereas for something urgent they prefer an alarm. Note that certain melodies may confuse and not representing awareness.	FASS
Interrupt/ Awareness		Messages and notifications should stay on the screen until they have been noticed	OFFIS
Notification		We used a friendly sound to notify users that the robot recognized that it was called by name and open for further instructions. We used the default sound of MSN messenger to notify users of a notification. We received no complaints about this. One user even recognized the MSN sound.	NOVAY
Security		Unless the use is interacting with the robot via screen, the security distance should be of 1 meter or a similar where the screen is reachable by hand.	FASS same for OFFIS
Safety		Some questions regarding the safety of the robot itself were risen: What happens if the robot falls over? Is there some kind of protection? Can he get back up? Are there enough sensors for obstacle detection?	OFFIS
Appearance	Size	Users liked the fact that you don't have to look up to the robot while sitting; it is not frightening, but friendly.	NOVAY
Appearance	Size	The height (approx. 1.4 meters) was nearly perfect while sitting on the couch. If the user gets up, it could be a little higher -> the stand should be height adjustable. The robot itself seemed to be a little big compared to the size of the room (very dominant).	OFFIS
Appearance	Face	Before meeting the robot, the mobile robots without screen and pet-style look were the most preferred options. After	NOVAY

		meeting and interaction with the robot, however, the mobile robot with screen and face (before meeting more disapproved of than approved), was the most favourite option. We believe that giving the robot a friendly face has a very high face value. Two out of six users even gave the robot a compliment after having completed their first instruction: “well done”, with an intonation much like you would praise a dog or child.	
Appearance	Style	Two users commented that they didn’t like the cloth we draped around the robot and that the looks of the robot should be a bit tauter, like the sketch we’ve shown them.	NOVAY
Appearance	Wheels	With respect to movement, hidden wheels were the most preferred option; visible wheels were deemed acceptable, both before and after meeting the robot.	NOVAY
Appearance	Usage	A tablet would be nice to have. Some users indicated to use the display as tablet if it is not in use (move it in horizontal position) If the stand is shaking like it is now, a bag would be better than a tablet. If the robot is in charging station, it should indicate that it is still turned on and working. Some users also suggested using the robot for cleaning the floor. Climbing stairs would be another thing that users wish to have because most of them have stairs at their homes.	OFFIS

6 Conclusion

The Wizard of Oz test has proved a valuable technique to improve the requirements, since we have been able to clarify many questions that remained open after previous tests with focus groups.

This process represents the final stage for defining the user scenarios since real in-place interactions could be observed (see D1.3 Final Robot Based Service Scenarios for complete scenario and use case descriptions). There has been a redefinition of our view of the interaction techniques since we were able to notice the lack of proper sequence in real life interactions leading to consider more flexible environments.

Some challenges were presented while setting up the system regarding the technical part such as battery and WIFI problems along with the installation of VOIP software.

Comments from WoOz tests will also influence the technical WPs in which additional matters such as robot height, touch screen size and interaction means will be designed and developed.

This technique also presents challenges at the time of performing it, forcing “the wizard” to be trained in this performance. The forecast of different reactions taken by users and possible contingencies in a relatively uncontrolled environment is essential for the rapid reaction of the person performing the wizard. The cost of this test is low taking into account the feedback given in terms of usability and acceptance. From the contact with our users we have realized that robots are no longer a futuristic metal being, but a well understood and rather well accepted tool.

7 References

Bickmore, T., Caruso, L., Clough-Gorr, K. (2005); Acceptance and Usability of a Relational Agent Interface by Urban Older Adults, *CHI 2005*, Portland, Oregon, USA

Breazeal, C. (2003a); Emotion and sociable humanoid robots, *International Journal on Human-Computer Studies 59*, 119–155

Coenen T. (2006); “Knowledge Sharing over Social Networking Systems: Architecture, Usage Patterns and Their Application” *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*

Feil-Seifer, D., M. Mataric (2005); Defining Socially Assistive Robotics, In: *Proceedings of IEEE 9th*

Forlizzi, J., DiSalvo, C., Gemperle, F. (2004); Assistive Robotics and an Ecology of Elders Living Independently in Their Homes, *Human-Computer Interaction 19*, 25-59

Haigh, K., Kiff, L., Ho, G. (2005); The independent lifestyle assistant (I.L.S.A.): lessons learned, Technical Report ACS-P03-023, Honeywell Laboratories, USA

Libin, A. and Libin, E. (2004); Person-Robot interactions from the robopsychologists' point of view: The robotic psychology and robotherapy approach, In: *Proceedings of the IEEE Vol. 92 (11)*, 1789 – 1803

Lisetti, C., Nasoz, F., LeRouge, C., Ozyer, O., Alvarez, K. (2002); Developing Multimodal Intelligent Affective Interfaces for Tele-Home Health Care, *International Journal of Human-Computer Studies archive 59 (1-2)*, 245-255

Thrun, S., (2004); Toward a Framework for Human-Robot interaction, In: *Human Computer Interaction, Special issue: Human Robot Interaction.*

Wada, K., Shibata, T, Saito, T., Sakamoto, K., Tanie, K. (2005); Robot assisted activity at a health service facility for the aged for 17 months, *IEEE Workshop on Advanced Robotics and its Social Impacts*

8 Annex

8.1 Information & Informed Consent - Netherlands

GEÏNFORMEERDE TOESTEMMING

Ik, ………………………………………………………………… *(naam deelnemer)*

stem toe mee te doen aan een onderzoek dat uitgevoerd wordt door Novay.

Onderzoeker namens deze partij is: *Henri ter Hofte* (naam onderzoeker)

Ik ben me ervan bewust dat deelname aan dit onderzoek geheel vrijwillig is. Ik kan mijn medewerking op elk tijdstip stopzetten en de gegevens verkregen uit dit onderzoek terugkrijgen, laten verwijderen uit de database, of laten vernietigen.

De volgende punten zijn aan mij uitgelegd:

1. Het doel van dit onderzoek is de mening van eindgebruikers te peilen over het inzetten van een robot bij diverse onderdelen van het dagelijkse leven. Deze "Florence-robot" heeft tot doel het vergroten van het welzijn van ouderen en het verbeteren van de efficiëntie in de ouderenzorg.
2. Er zal mij gevraagd worden om een paar oefeningen met de robot te doen en daarover mijn mening te geven. Het hele onderzoek zal ongeveer 1 uur duren.
3. Er behoort geen stress of ongemak voort te vloeien uit deelname aan dit onderzoek.
4. De gegevens verkregen uit dit onderzoek zullen anoniem verwerkt worden en kunnen daarom niet bekend gemaakt worden op een individueel identificeerbare manier.
5. De onderzoeker zal alle verdere vragen over dit onderzoek beantwoorden, nu of gedurende het verdere verloop van het onderzoek.

Handtekening onderzoeker: …………………………… Datum: ……………..

Handtekening deelnemer: …………………………… Datum: ……………..

8.2 Information & Informed Consent - Spain

FLORENCE

GRUPO DE DISCUSIÓN

Como participante del test del prototipo Florence, hemos elaborado esta sencilla hoja informativa con la finalidad de introducirle en el proyecto así como en el test de usabilidad que le vamos a mostrar.

PROYECTO FLORENCE

El objetivo principal del proyecto Florence es mejorar el bienestar de las personas mayores y la de sus familiares /cuidadores, así como mejorar la eficiencia de la atención a través de nuevas tecnologias integradas en el entorno de la personas usuarias.

El sistema Florence, con su plataforma de multiuso robot móvil será pionera en el uso de robots así como en la prestación de nuevos tipos de servicios en favor de la autonomía de las personas mayores y de la tranquilidad de sus cuidadores. El objetivo principal es hacer que este concepto sea aceptable para los usuarios y rentable para la sociedad.

El proyecto consiste en la integración de robots que presenten las siguientes características y utilidades, con la finalidad de permitir la comunicación de manera ágil y sencilla entre el usuario y su familia:

- Intercambio de de mensajes, fotografías, videos, ect.
- Comunicación a través de video-teléfono.
- El robot puede ser enviado a diferentes habitaciones para ver si todo marcha bien a través de la cámara, por lo que tiene una función de vigilancia (para dar una mayor seguridad a la persona usuaria).
- Intervención en momentos críticos. Por ejemplo, puede detectar una caída e inmediatamente poner en marcha todo el protocolo de emergencias.
- Asesoramiento en actividades de entrenamiento o bienestar.
- Sistema de video conferencia que puede dar un “diagnóstico” a partir del historial que posee en su memoria, que puede facilitar el trabajo a los profesionales sanitarios que actúen en caso de emergencia.

La prueba que se va a realizar consiste en simular diversas situaciones de uso del robot de una manera interactiva, con el fin de mejorar la utilidad y la usabilidad del modelo final.

El moderador no preguntará ni juzgará ninguna información directamente, sino que introducirá las situaciones para que ustedes puedan interactuar con el robot y dar sus impresiones.

De este intercambio de impresiones sobre los temas previamente expuestos por el moderador se detectarán aquellas cuestiones todavía no planteadas así como las posibles soluciones a temas todavía en debate dentro de nuestro proyecto. En definitiva queremos conocer la opinión de los participantes del test sobre las funciones diseñadas previamente por los diferentes socios del proyecto Florence en la dinámica de una discusión con actores que intervienen día a día en la teleasistencia.

Las identidades de cada participante se guardarán celosamente con arreglo a la jurisdicción española en privacidad (están protegidas conforme a la Ley 15/1999, de Protección e datos de carácter personal y demás normas conexas).

CONSENTIMIENTO INFORMADO.

PARTICIPACIÓN EN GRUPO DE DISCUSIÓN

He hablado con…………………………………………….............................sobre el proyecto.

Esto fue en
……………………………..

- He tenido la oportunidad de hablar sobre el proyecto y preguntar dudas
- Se suficiente sobre el proyecto ahora

- Entiendo que es mi decisión el formar parte de él o no
- Entiendo que si no quiero formar parte del mismo o decido dejarlo, esto no afectará a la ayuda que estoy recibiendo
- Entiendo que la entrevista puede que sea grabada. Puedo parar esto en cualquier momento. He sido informado que los datos personales gozarán de las medidas de seguridad previstal en la LO 15/1999, de protección de datos de carácter personal, y normas conexas.
- Estoy de acuerdo en tomar parte en el proyecto

Firmado…………………………………..................……….Fecha…………

Nombre(Mayusculas)……………………………….........................…..............

Firmado (Investigador)
……………………………......................…Fecha…………………..

Nombre
(Mayusculas)……………………………………………...............................……..

8.3 Information & Informed Consent - Germany

Informationsblatt zur Studie „Wizard of Oz Test Florence“

Liebe Studienteilnehmer!

Vielen Dank für die Bereitschaft, an dieser Studie teilzunehmen. Die Studie ist Teil des Florence Projekts, das von der Europäischen Kommission im Rahmen des 7. Rahmenforschungsprogramms finanziert wird.

Ziel der Studie ist es, die Einsatz- und Verbesserungsmöglichkeiten von Robotern im häuslichen Umfeld zu untersuchen. Hiermit soll die Entwicklung des geplanten Robotersystems unterstützt werden.

Es handelt sich bei dieser Studie um einen sogenannten Wizard of Oz Test. Es handelt sich hierbei um eine Simulation des später zu entwickelnden Gerätes im realen Umfeld. Das heißt, wir führen Ihnen den Roboter in unserer IDEAAL-Wohnung vor und befragen Sie zu Ihren Meinungen / Ideen / Anregungen.

Während der Studie werden Daten über den Studienverlauf erhoben. Dazu gehört die Audioaufzeichnung Ihrer Kommentare, Merkdaten zum Verlauf (bspw. Dauer des Gesprächs), sowie übliche (Alter und Geschlecht) und für diese Studie interessante (z.B. technischer Beruf) persönlichen Daten. Möglicherweise werden während der Studie ebenfalls Videoaufnahmen gespeichert. Die Audio- und etwaige Videoaufnahmen dienen der besseren Auswertbarkeit und werden nur zu diesem Zweck gespeichert. Sie werden Dritten nicht zugänglich gemacht.

Zusammenfassungen der Daten (gemittelt über die Teilnehmer) werden anonymisiert in Textform – bspw. in wissenschaftlichen Artikeln publiziert. Medien, wie Fotos oder Audio-/Videospuren, auf denen Sie als Personen zu erkennen sind, werden nur nach expliziter vorheriger Erlaubnis veröffentlicht.

Sie können sich jederzeit und ohne Nennung von Gründen aus der Studie zurückziehen. Sollten Sie Fragen haben, beantworten wir sie gerne. Vielen Dank, dass Sie mit ihrer Zeit unsere Arbeit und damit die Forschung im Bereich der medizinischen Gerätetechnik unterstützen!

1) Ich bin bereit an dieser Studie im Rahmen des Projekts teilzunehmen
2) Ich wurde über die Ziele der Studie aufgeklärt. Ich fühle mich ausreichend informiert.
3) Mir wurde erklärt, dass
   a. während der Studie schriftliche Aufzeichnungen gemacht und meine Kommentare auf Tonband und/oder Video aufgenommen werden,
   b. ich nicht dazu verpflichtet bin, die schriftlich und mündlich gestellten Fragen zu beantworten,
   c. alle persönlichen Informationen unter das Bundesdatenschutzgesetz[1] fallen, was bedeutet, dass meine Identität nicht ohne meine Einwilligung preisgegeben wird,
   d. alle gesammelten Daten ausschließlich und anonymisiert für wissenschaftliche Zwecke im Rahmen dieser Arbeit verwendet werden,
   e. ich jederzeit und ohne Begründung eine Aktivität oder die gesamte Teilnahme an der Studie abbrechen kann.
4) Ich kann die Wissenschaftlichen Mitarbeiter Melvin Isken [2] oder Hannah Baumgartner[3] kontaktieren, wenn ich Fragen zur Studie, dem Projekt oder meiner Teilnahme habe.

Teilnehmer____

Oldenburg, den____

Unterschrift____

8.4 Study procedure - The Netherlands

phase#	phase	task	duration total	item duration	description
			0:15		***Before the robot enters***
				0:02	Researcher welcomes participant to the room. The robot and wizard are out of sight. Researcher explains goal and structure of the session (also tells that the wizard is in the next room to monitor the robot and make sure nothing goes wrong. the researcher but does NOT reveal that the robot is actually operated by a human).
				0:03	Inform participant and ask to sign consent form.
				0:10	Pre-robot meeting interview by researcher, based on survey.
0	robot enters room	none	0:05		***Robot enters room***
				0:05	Researcher explains about the robot: looks of this prototype are not final (show picture how it might look), explains roughly what it can and can't do. Explain movement, camera (microphone), touch screen. If the robot is on the extension cord, explain that it will move without cord in a later version and ask user to imagine the cord is not there. *Briefly talk with participant about robot (what do like, what don't you like, what do you expect). In the end ask user to imagine it looks better (like the picture).*

phase#	phase	task	duration total	item duration	description
1	voice only	approach+orient	0:05		***1st interactive task: activity suggestion / voice only***
				0:02	researcher explains Imagine you've just had breakfast and the robot makes sound. You'd like to know about your health stats, which the robot can display on the screen. This time, the robot can only listen to voice commands. If you want the robot to listen, you always need to say "Florence" first. This will cause the robot to turn towards you. Then you give a command. Researcher shows and robot responds: "Florence … here". Let the participant try this command and robot responds. Tell participant there are three commands (note: in native language) "Florence: here" (nl: "hier") "Florence: stop" / (nl: "stop") "Florence: go" / (nl: "ga")
				0:03	interaction *Ask user to call the robot and read what's on the screen (advice to take a walk outside). Ask the participant to reflect briefly on how this went.*
2	gesture	approach+orient	0:05		***2nd interactive task: activity suggestion / gestures***
				0:01	researcher explains next task: same situation, but now the robot can only respond to gestures **(researcher presses some buttons on the base of the robot "to switch to a different mode").** But: to get the robot's attention, you still have to call it with "Florence". Now you can use three gestures (research shows the gestures and robot responds): "Florence" + here gesture "Florence" + stop gesture "Florence" + go gesture
				0:04	Ask user to call the robot and read what's on the screen (advice to take a walk outside). *Ask the participant to reflect briefly on how this went. Also ask the participant which of the two ways (voice only versus gestures)*

phase#	phase	task	duration total	item duration	description
					he/she liked better and why.
3	voice + gesture	follow+orient	0:05		***3rd interactive task: follow & orient / voice only + gestures***
				0:01	Researcher explains next task: robot suggests doing an exercise and can play a video of the exercise. You want to do that in the living room. You take the robot with you to the living room, position it so you can see the video. Now you have both voice commands and gestures available **(researcher presses some buttons on the base of the robot "to switch to a different mode").** In addition to come, stop and go away, you also have one new command and gesture: "follow me". "Florence: follow" (nl: "volg") / "Florence" + follow me gesture "Florence: here" / "Florence" + come gesture "Florence: stop"/ "Florence" stop gesture "Florence: go" / "Florence" + go gesture
				0:04	Ask participant to take to robot to the living room and position it so you can see the video and do the exercise. The wizard responds; in case of conflicting gesture and voice command: do the voice command. *Ask the participant to reflect on how it went and specifically ask why she chose voice and/or gestures.*

phase#	phase	task	duration total	item duration	description
4	voice + gesture + touch	play/pause video	0:05		***4th interactive task: play video using gestures or touch***
				0:01	Researcher explains next task: Play a video of the exercise. You can use two ways to control the playout: with voice or via touch screen. **(researcher presses some buttons on the base of the robot "to switch to a different mode").** *"Florence: play" (nl: "speel" / touch screen to play* *"Florence: pause video" (nl: "pauze") / touch screen to pause*
				0:04	Participant takes robot to the living room and positions it so she can see the video. *Ask the participant to reflect on how it went and specifically ask why she chose voice and/or touch screen.*
5		general	0:10		***After 4th task: post-robot meeting interview***
				0:10	Post-robot meeting interview by researcher, based on survey.
			0:05		***[5 min] Finalize***
				0:05	Researcher guides the participant to the exit. Note: we will tell participant that the robot was actually controlled by the wizard AFTER all participants did the test. We did not tell that beforehand, because it is very important for us that you to react as if you would respond to a robot, not to a human.

8.5 Instruction Slides – Netherlands

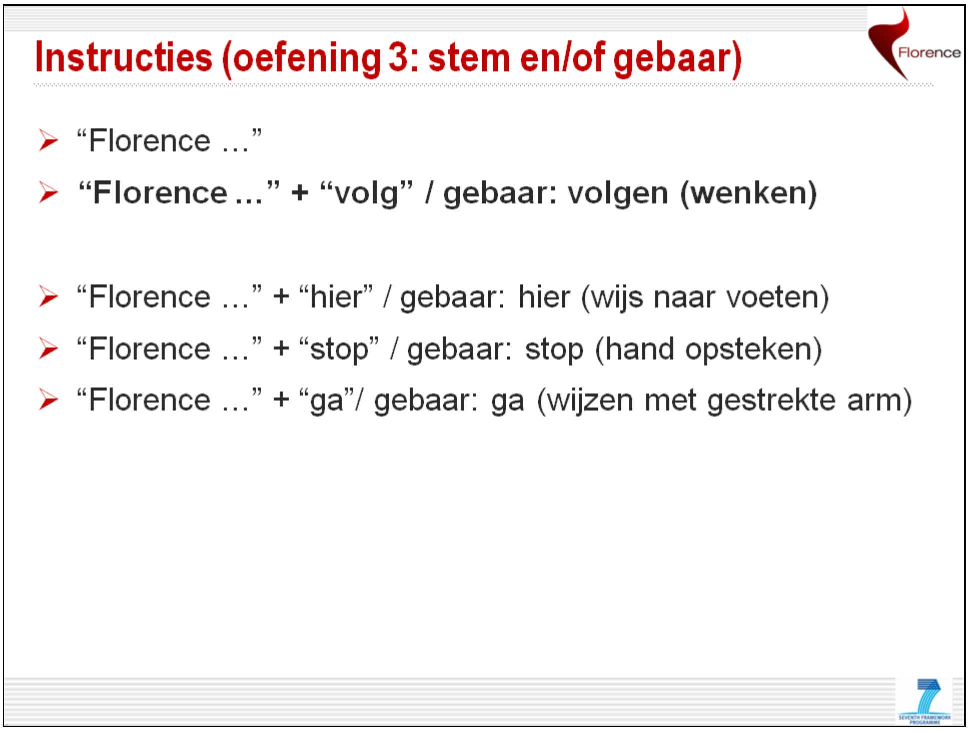
These slides have been used during the WoOz tests in the Netherlands.

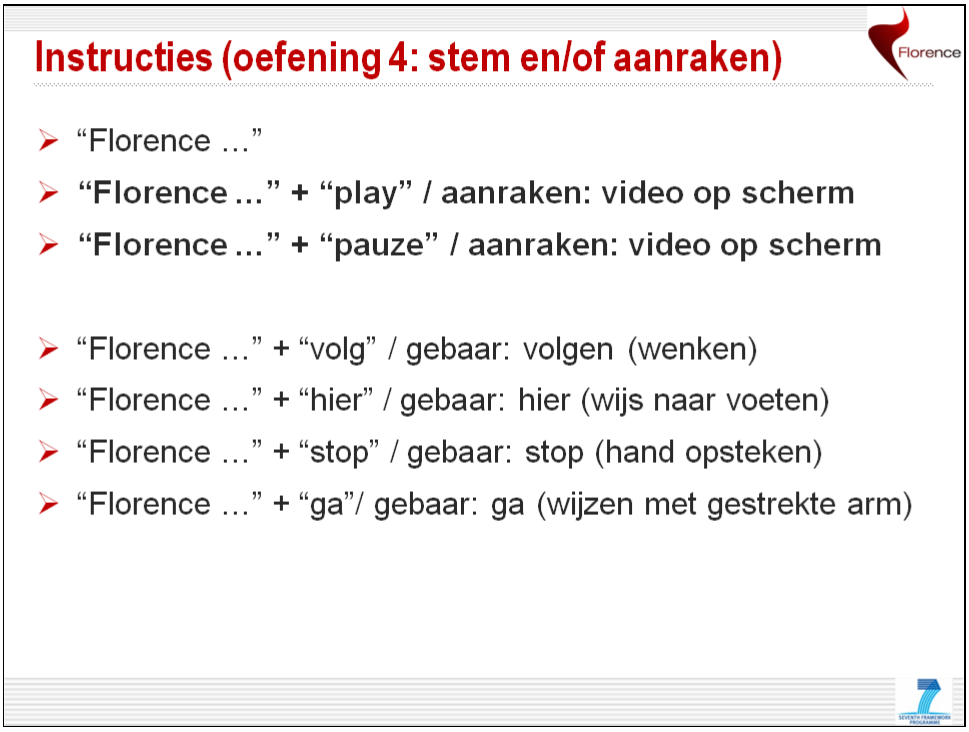
Instructies (oefening 1: alleen stem)

- “Florence …”
- “Florence … hier”
- “Florence … stop”
- “Florence … ga”

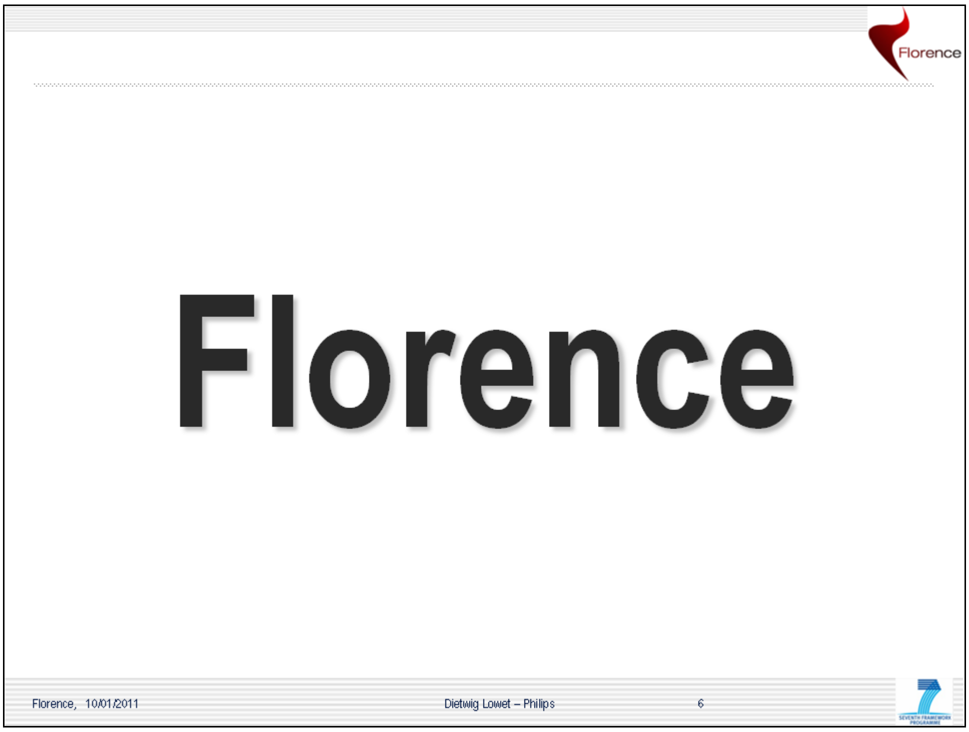
Instructies (oefening 2: gebaren)

- “Florence …”
- “Florence …” + gebaar: hier (wijs naar voeten)
- “Florence …” + gebaar: stop (hand opsteken)
- “Florence …” + gebaar: ga (wijzen met gestrekte arm)





Also, on the floor, at the spots where Florence would go when given the "Go" command, we placed an A4 paper with the word 'Florence' on it:


Florence
Florence
Florence, 10/01/2011
Dietwig Lowet – Philips
6

8.6 Study procedure - Germany

phase#	phase	task	duration total	item duration	description
			00:15		*Before the robot enters*
				00:02	Researcher welcomes participant to the room. The robot and wizard are out of sight. Researcher explains goal and structure of the session (also tells that the wizard is in the next room to monitor the robot and make sure nothing goes wrong. the researcher but does NOT reveal that the robot is actually operated by a human).
				00:03	Inform participant and ask to sign consent form.
				00:10	Pre-robot meeting interview by researcher, based on survey.
0	robot enters room	none	00:05	00:05	*Robot enters room*
				00:05	Researcher explains about the robot: looks of this prototype are not final (show picture how it might look), explains roughly what it can and can't do. Explain movement, camera (microphone), touch screen. If the robot is on the extension cord, explain that it will move without cord in a later version and ask user to imagine the cord is not there.
					Briefly talk with participant about robot (what do like, what don't you like, what do you expect). In the end ask user to imagine it looks better (like the picture).
1	voice only	commands	00:05		*1 st interactive task: activity suggestion / voice only*
				00:02	researcher explains
					We've implemented some voice commands. If you want the robot to listen, you always need to say "Florence" first. This will cause the robot to turn towards you. Then you give a command.
					Let the participant try this command and robot responds by movement
					Tell participant there are commands like (note: in native language)
					"Florence: here"
					"Florence: stop"
					"Florence: go"
					"Florence: left / right"
				00:03	interaction
					Ask user to call the robot and read what's on the screen (weather report). Ask the participant to reflect briefly on how this went.
2	voice + movement	approaching	00:05		*2 nd interactive task: approaching the user*
				00:01	researcher explains next task: now try to get the robot in front of you, we will try different positions, to test which fits best **(researcher presses some buttons on the base of the robot "to switch to a different mode").**
				00:04	Ask user to call the robot and read what's on the screen (weather report). Try to touch screen, which angle is good etc.
					Ask the participant to reflect briefly on how this went.

3	voice + movem ent	following	00:10		*3rd interactive task: following person*
				00:01	Researcher explains next task: user should go from kitchen through living room to corridor. We have different ways the robot can follow **(researcher presses some buttons on the base of the robot "to switch to a different mode").** In addition to the other voice commands, you also have one new command and gesture: "follow me".
				00:04	Direct Following: the robot always tries to get into a certain distance to the user
				00:05	Indirect Following: the robot only moves if the user is out of sight. If the user is in sight, the robot only turns towards him/her
					Ask the participant to reflect on how it went
4	voice + movem ent	avoiding	00:10		*4th interactive task: avoiding the user*
				00:01	Researcher explains next task: We have two tests: First - the robot is in the way and you want to get through. Second - try to get the robot into a corner until the robot can't move anymore. We implemented different strategies how the robot could react. **(researcher presses some buttons on the base of the robot "to switch to a different mode").**
				00:03	First trial: the robot will just keep the distance to the user, if the user makes one step to the robot, the robot just moves back (still standing in the way)
				00:03	Second trial: the robot ties to get out of the way: turn around and then go back so that the way of movement is free
				00:03	Third trial: force the robot into a corner. If the robot can't move anymore, the robot just stops
				00:03	Fourth trial: force the robot into a corner. If the robot can't move anymore, it will turn forth and back to show that it doesn't know where to go
5		general			*Ask the participant to reflect on how it went and specifically ask why she chose voice and/or touch screen.*
			00:10		*After 4th task: post-robot meeting interview*
				00:10	Post-robot meeting interview by researcher, based on survey.
			00:05		*[5 min] Finalize*
				00:05	Researcher guides the participant to the exit.
					Note: we will tell participant that the robot was actually controlled by the wizard AFTER all participants did the test. We did not tell that beforehand, because it is very important for us that you to react as if you would respond to a robot, not to a human.

8.7 Questionnaire

The following questionnaire was filled by the participants before and after the WoOz tests.

Questionnaire before WoOz:

Summary of the project:
The project consists on building a Robot for helping users during their Activities of Daily Living (ADL), covering the emergencies, connectivity, security, monitoring and entertainment. Tasks where the robot could be used include:

- Emergency assessment, like falls or fainting
- Remember notices, like taking drugs or visits to the doctor
- User's physical activity measurements to advise healthy habits
- Promote the connection with other users and relatives with collaborative activities like making a puzzle

Part 1. Robot appearance

1 = COMPLETELY DESAGREE
4 = COMPLETELY AGREE

1. Taking into account the characteristics of the project, you would like that the robot has appearance of… (see ANNEX I)	☹ **1**	**2**	**3**	☺ **4**
a. Realistic humanoide				
b. Other kind of humanoide				
c. Mobile robot with screen				
d. Mobile robot with arms				
e. Mobile robot with arms and screen				
f. Mobile robot with arms, screen and face				
2. If the robot has to move, you would prefer that…	☹ **1**	**2**	**3**	☺ **4**
a. It used legs (see a. and b. in ANNEX I)				
b. It used wheels (see d. and f. in ANNEX I)				
c. It used **hidden** wheels (see c. and e. in ANNEX I)				

Part 2. Interaction

1 = COMPLETELY DESAGREE
4 = COMPLETELY AGREE

3. If the robot has to move, you would prefer that…	☹ **1**	**2**	**3**	☺ **4**
a. I move it using a remote control				
b. It follows me when I am moving				
c. It moves when I call it				
d. It moves by himself and I do not tell it when and where to move				

1 = COMPLETELY DESAGREE
4 = COMPLETELY AGREE

4. If the robot has to give you some information, you would like that it does so…	☹ **1**	**2**	**3**	☺ **4**
a. Showing some text through a screen				
b. Showing images through a screen				
c. Using voice				
5. If you have to order something to the robot, you would like to do so…	☹ **1**	**2**	**3**	☺ **4**
a. Choosing one option using a touch screen				
b. Using a remote control				
c. Using buttons located in the body of the robot				
d. Using voice				
6. When the robot gives you the information, you would like to confirm that you receive it…	☹ **1**	**2**	**3**	☺ **4**
a. Using gestures				
b. Using a remote control				
c. Using buttons located in the body of the robot				
d. Using your voice				
7. I would like to maintain a talk with the robot (Ask him some information using voice, and the robot answers me using voice also)				
8. I would like that the robot has a face and expresses himself using gestures (sad, angry, happy, …)				

Questionnaire after WoOz:

Part 1. Robot appearance

1 = COMPLETELY DESAGREE
4 = COMPLETELY AGREE

1. Taking into account the characteristics of the robot you tested, you would like that the robot will have the appearance of… (see ANNEX I)	☹ 1	2	3	☺ 4
a. Realistic humanoide				
b. Other kind of humanoide				
c. Mobile robot with screen				
d. Mobile robot with arms				
e. Mobile robot with arms and screen				
f. Mobile robot with arms, screen and face				
g. The appearance it has at the moment				
2. Because the robot has to move you think that..	☹ 1	2	3	☺ 4
a. It should use legs (see a. and b. in ANNEX I)				
b. It should use wheels (see d. and f. in ANNEX I)				
c. It is O.K. at the moment				
	☹ 1	2	3	☺ 4
3. I would like that the robot should have a basket or a bag to save things				
4. I would like that the robot should have a clothes hook to hang my clothes				
5. I would like that the robot should have a table to put there my mobile phone, glasses, … (to find this things easily)				

6. Will you add something more to the robot? What?

Part 2. Interaction

1 = COMPLETELY DESAGREE
4 = COMPLETELY AGREE

	☹ 1	2	3	☺ 4
7. The robot moved properly				
8. That the information was shown on the screen is proper				
9. The information was shown on the screen can be seen / read properly				
10. That the robot communicated with me using voice is proper				
11. When the robot communicated with me using voice I understood it properly				
12. The way I used to confirm the robot that I received the information it gave me is correct				

13. What did you like most from the robot you tested?

14. What did you like less from the robot you tested?

15. What will you change from the robot you tested?

The following page is ANNEX I of the questionaire (not of this deliverable document). This page is used by both questionnaires, before and after the tests.

ANNEX I

a. Realistic humanoide

b. Other kind of humanoide

c. Mobile robot with screen

d. Mobile robot with arms

e. Mobile robot with arms and screen

f. Mobile robot with arms, screen and face