Project FoSIBLE
Fostering Social Interactions for a Better Life of the Elderly

Deliverable

D6.2: Report on usability tests

Responsible
Fraunhofer IMS
UDE

Participants
Mauser

Version: Final
Date: 31/07/2013
Dissemination level: Public

Abstract

Our work on WP6 started with early mock-up testing. The focus of these testing was error handling, stability and usefulness especially for elderly users. The results have been taken as a basis for the prototype development. In regard to the prototype of the FoSIBLE widget, we established an iterative process using a bug tracking tool that allowed a close collaboration with the development partners.

Furthermore, we conducted a heuristic evaluation in the Fraunhofer inHaus to identify usability issues of the FoSIBLE platform. Participants of this evaluation were usability experts as well as persons of the target group.

As a last step, an evaluation with members of the target-group took place in the living environment that has been set up in the Fraunhofer inHaus. For this purpose, 15 members of the AlterAktiv e.V. computer club in Siegen were invited to have a “hands on” interaction session with the sensor devices developed by Mauser and installed at the Fraunhofer inHaus by IMS. This was followed by a focus group discussion in order to collect feelings and opinions about the sensor devices as well as suggestions for improvement. Furthermore, this was accompanied by an evaluation session of the re-designed version of the Gameinsam application developed by researchers of UDE.

Table of Contents

1. Introduction

1.1 Scope of This Deliverable

Objectives of the WP

The objective of this work package is to gather first feedback results with respect to the quality and user friendliness of the developed system. To this end, the inHaus Lab facilities will be exploited in order to

1. Test the developed material in a controlled use situation,
2. Boost the quality of the developed products by deeply involving test users from the target community into the development process.

The inHaus Labs will allow collecting valuable user feedback by allowing them to test our system within a home environment. This shall enable non-technical skilled users to gain hands on experiences with our system. Nevertheless, this work package is not intended to replace any real world field tests. Instead, it will be exploited to collect end user feedback information at an early stage within the development cycle. As a result, these tests will be executed on a regular basis whenever significant development or research results become available.

Description of work

Task 6.1: Early testing (mock-ups) (Responsible: FhG IMS and Uni DUE, Collaboration: Mauser Care): In the first evaluation phase mock-ups of the applications and components will be tested with end users in the inHaus Labs. These tests will be done applying a Wizard of Oz approach. The goal is to gain usability results about the system and its components in an early stage within the development phase. Mauser Care will contribute and install the corresponding sets of furniture at the inHaus lab home.

Task 6.2: Prototype Testing (Responsible: FhG IMS and Uni DUE, Collaboration: Mauser Care): In the next lab evaluation the application and prototypes will be evaluated in a monitored inHaus2 environment by professional testers and members of the target user groups. In this lab study, the focus will be on bugs and usability aspects. These tests will be carried out in order to (a) detect bugs and errors which will be corrected as soon as possible (before the field evaluations) and (b) to get better testing results with more information than the field evaluation can provide. Therefore, elements are tested, which are not implemented in the test households. Test tasks (in form of use cases) will be prepared. With given tasks every possible function of the software/hardware can be tested. Feedback on the tasks can help to find usability issues and fix them by interviewing the testers which had usability problems. (The two groups of testers will be used for different focuses of testing. The professional testers will have a deeper focus on bugs while the testers from the target user group will have a deeper focus on usability aspects. The focuses can be easily controlled by the tasks given to the testing persons.) After the first tests, the problems should be solved and the hard- and software should be redesigned and at this point be bug free. Mauser Care will contribute and install the corresponding sets of furniture at the inHaus lab home.

Task 6.3: Usability tests: CURE with support of AIT will test individual measurement components in their Experience Lab in Vienna, Austria, prior to integration in the overall system that is tested in inHaus. The facility of CURE includes a usability lab, so both functional and usability testing can be performed.

Deliverables of the WP: no., brief description and project month of delivery

D6.1: Report on inHaus lab evaluation results. Delivery date: M30

D6.2: Report on usability tests. Delivery date: M39.

D6.3: Report on usability tests in laboratory in Vienna. Delivery date: M30

1.2 Overview on WP6 Tests and Evaluations

In order to ensure the quality and usability of the developed FoSIBLE system in a holistic manner, the different partners set up several evaluation scenarios. One of the central aspects, besides usability, was to investigate the sociability and social interaction that can be created during the act of watching TV using presence and awareness tools.

For the entertainment applications like "Gameinsam", UDE evaluated the sociability and playful interaction of the elderly as well as a re-design of the interface of "Gameinsam", to point out adequate interface designs, focusing the specific needs of older people. In addition, UDE arranged an expert evaluation, which focused on specific usability heuristics for older persons to assure a high degree of usability for the FoSIBLE system. In a final step, a target group evaluation was conducted in order to investigate the suitability of several ambient applications for older people. Therefore, a sensor system working together with an ambient application was integrated in the Fraunhofer inHaus laboratory which has been used by members of the target group in a "hands on" session during the evaluation. These evaluations as well as their results are the major focus of this deliverable. Figure 1 gives an overview on the different evaluations conducted by UDE in cooperation with the project partners.

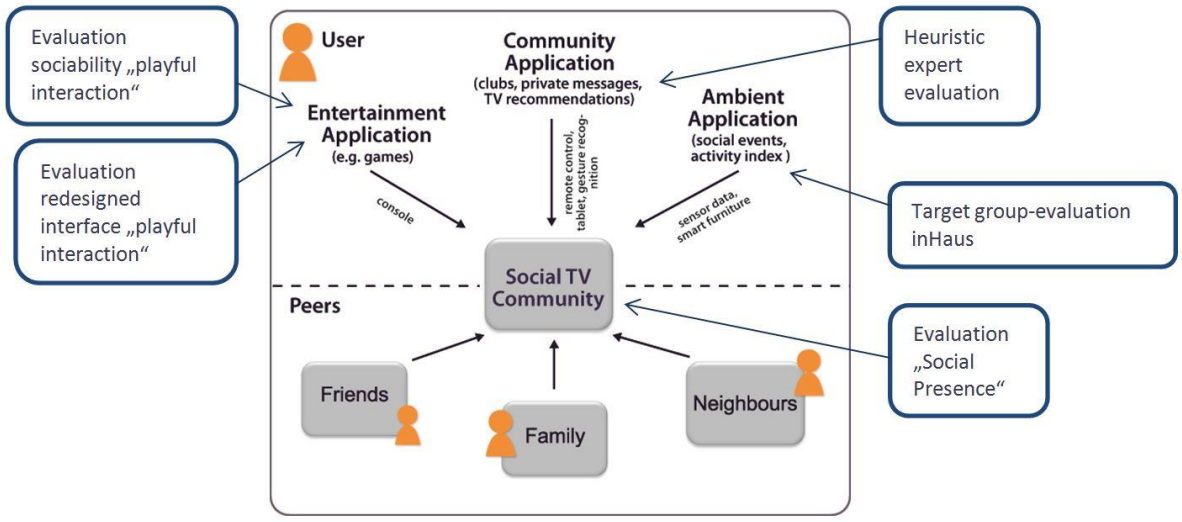

Figure 1 -Elements of the FoSIBLE system and their evaluation

2. Mock-Up and Prototype Testing

2.1 Early testing of mock-ups and concepts

Regarding work package 6, we started with our evaluation and test phase very early in the project in order to integrate the results in the widget development. As soon as the first version of the widget was presented by Kaasa (first version included a news section and first runnable applications on the real TV-setup) and University of Siegen (the social Widget which handled the chat-, buddy list-, and first network structure), we tested the available parts in particular with the focus on error handling, stability and usefulness regarding elderly users. Subsequently, constant feedback in form of e-mails, bug reports and written requirements to the responsible partners was provided.

We have strongly participated in the selection and development process of an appropriate implementation tool of the FoSIBLE system. For this purpose, different possibilities of social community platforms were evaluated and reviewed in respect to the goals of FoSIBLE. The results were integrated in Deliverable 2.1 and contain a detailed overview about costs, possibilities and restrictions of the reviewed social community building tools.

We coordinated and participated in developing new “demonstration scenarios” based on the developed preliminary scenarios, the developed persona and the widget functionalities. The demonstration scenarios (D2.2) provide a deeper understanding of the FoSIBLE functions and operations in the application context. In order to develop these scenarios, several coordination meetings via Skype and one meeting at the Fraunhofer inHaus in Duisburg were coordinated by UTT and UDE. The results were considered at the project review (28, 29, November 2011).

We also took a deeper look into the architecture of the developed system by reviewing the provided code of the different versions. During the whole development process of the FoSIBLE widget (delivered by Kaasa), detailed feedback in form of Excel sheets, emails and via the bug tracking tool Flyspray that has been used in the project (see figure 1) was provided.

2.2 Prototype Testing

Since the first version of the widget (or widget parts), we have strongly worked together with Kaasa and all other partners on improving the widget and identifying bugs. An iterative process of direct update loops has been established in order to fix bugs in the widget that were discovered in our Living Lab environment. Following the description of the bug-fixing procedure, a short overview about some of the produced versions in the project (figures 2-9) is given.

Especially in the later versions of the widget, the problems, bugs and feature requests got more complex so that it was decided to integrate a protocol for bug reporting or feature reporting of any kind via a bug-tracking tool. Therefore, a Flyspray system under the URL: http://Flyspray.kaasa.com/ has been launched together with Kaasa. Every project partner has access to this tool and can report bugs for any kind of problem (or even discuss features in general with all participating partners). The bug tracking tool was used to collect bugs for the TV widget as well as for the gesture control and the tablet application. The goal was to establish a process for all technical problems and to monitor the progress of the project better.

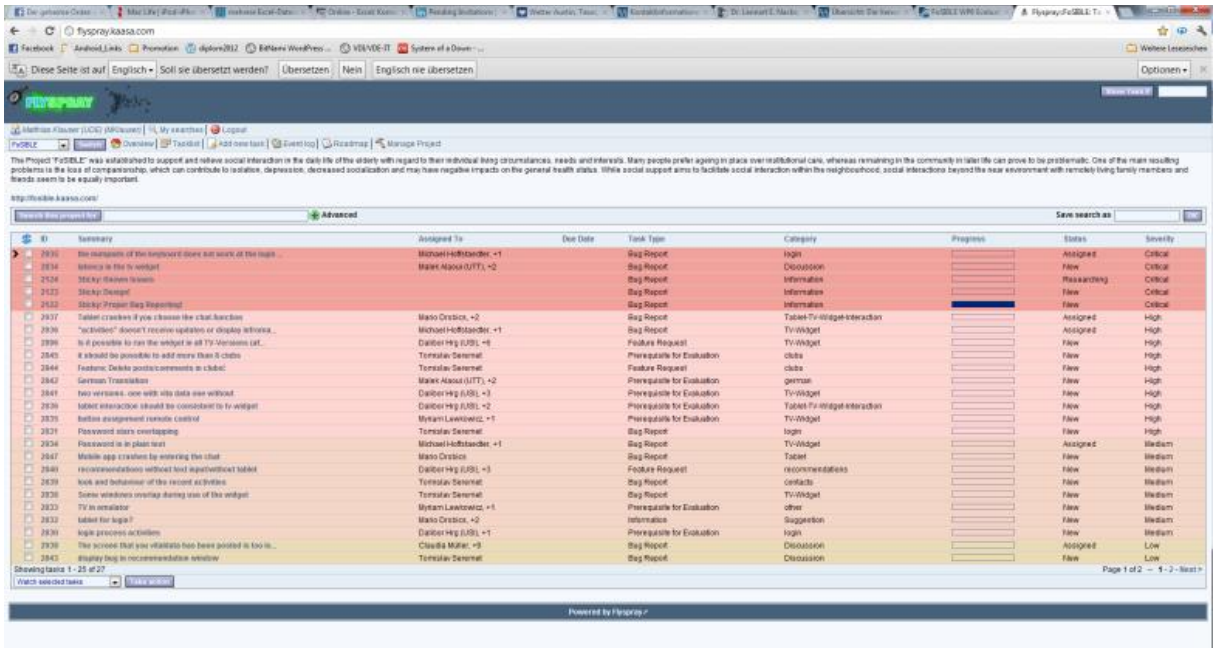

Figure 2 - Flyspray Bug Tracking System

For the first versions of the widget in May 2011, we mainly participated via code reviews and technology scouting. Therefore, the widget was evaluated by reviewing the presented functionality and the code behind these functions to improve the use of the Samsung API. By the time of the first integrated versions (integrated in the sense of combined version with tablet, social and basic functionality), tests of the functions and the application itself were started by providing detailed bug lists, feature lists and direct feedback, e.g. via Skype. Especially the versions launched in October/November 2011 were tested in the original inHaus-Lab environment. The identified bugs were provided in Excel sheets or communicated in live conferences with the developer team from Kaasa.

Figure 3 - FoSIBLE widget (platform & news) v1-26.05.2011

First version within the original Samsung platform. The version was a test for the capability and features of the platform and was able to display first reviews, and messages from different users.

Figure 4 - FoSible Widget (Social Widget) v1-26.05.2011

This is a second part of the first version produced by the university of Siegen. The focus here was to test the social functions of the widget and to create a first design draft.

Figure 5 - FoSIBLE widget v2-02.11.2011

This version was much more detailed and was able to handle different kinds of requests (i.e.: buddy list, chats, etc.). So it was the first version which integrated the social widget from the University of Siegen and the version produced by Kaasa.

Figure 6 - FoSIBLE widget v3-27.11.2011

Here the whole backbone architecture with TV recommendations and the interaction via the tablet has been added. Further interfaces were created to communicate with different services like the vita doc data, the inHaus sensors or high scores from games.

Figure 7 - FoSIBLE widget v4-19.04.2012

Following the review, the widget the widget was completely revised. The version 4 is the first stable version. It includes a whole new system architecture which facilitates the connection to other devices. Furthermore, the new architecture has a better performance. The most important update is the capability to run on different Samsung TV generations.

Figure 8 - FoSIBLE widget v5-24.04.2012

In this version further reported bugs from v4 have been fixed and the tablet can be used for the login.

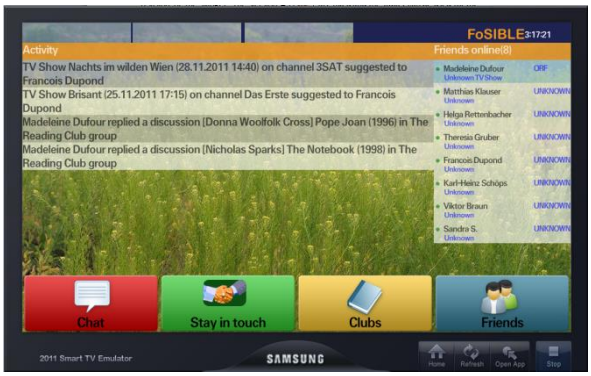

Figure 9 - FoSIBLE widget (Main Menu) v8-21.09.2012

In the latest version of the widget some functionality changes have been done. The buddy list now contains an indicator of the status (on-/offline). Moreover, the vital data element has been deleted. Instead, the new element “Stay in Touch” has been added.

Figure 10 - (Menu “Stay in Touch”) v8-21.09.2012

In the new section “Stay in Touch” users can for example search and add friends and send private messages.

The earlier tests by UDE were mostly conducted by using an emulator of the TV set. During the early testing period, only few tests by UDE were performed on the real device in the inHaus of Fraunhofer IMS in Duisburg. Since the end of 2012, all bug tests were performed on the real TV set in the Fraunhofer inHaus.

Further and more detailed information about the different versions and especially about the final sub-versions of the TV-widget and the tablet application can be found in the deliverables 4.1 and 4.2.

2.3 Evaluations

Besides the prototype testing of the components, different empirical evaluations as well as a heuristic evaluation described in chapter 4 and a focus group evaluation (chapter 5) were conducted.

A pre-study evaluation was conducted on the social community platform Facebook[1] via digital questionnaire with the focus on people of the age group of above 50. The questionnaire was designed to collect information about the habits regarding technical devices and social platforms. This first step of our evaluation reveals general information about the target group of FoSIBLE and provided the basis for our next evaluations and development of the system.

To focus further elements of the FoSIBLE interaction concept, an evaluation on playful interaction was conducted. The results show that playful interaction combined with co-located TV watching can foster sociability between family members and peers and how.

Simultaneously, another lab evaluation was conducted that focused on different forms of user representation in a mediated smart TV environment. In this study, it was investigated which effect these different representation forms have on the perceived social presence and awareness of the users and which role different TV genres play in this context. For further information on the pre-study, the evaluations of the playful interaction and social presence, and their results, see deliverable D6.1.

In addition, a heuristic evaluation was conducted in order to identify usability problems with the user interface design as well as with the interaction with the platform. For this purpose, a group of evaluators consisting of usability experts as well as persons of the target group

[1] Facebook - URL

examined the FoSIBLE platform in order to judge its compliance with existing usability principles.

In a final step, a focus group evaluation was conducted with focus on the social sensory environment installed in the Fraunhofer inHaus. Here, the different sensor functionalities were presented to members of the potential target group with a subsequent focus group discussion in which the participants were asked for their opinion regarding benefits, limitations and suggestions for improvement.

3. Expert Evaluation of Social TV Interfaces & Widgets

3.1 Heuristic Expert Evaluation

In December 2012, UDE organized a heuristic expert evaluation, which took place in the Fraunhofer inHaus. For this evaluation, the at that time current version of the FoSIBLE widget was used and run on the Samsung Smart TV version which is used in the field evaluation (cf. deliverable D7.3) as well. The aim was to identify major usability problems or bugs at this stage of the development, so that they could be eliminated before the start of the field trials.

The heuristic evaluation is a method to identify usability problems in the design of a user interface through a small group of usability experts on the basis of existing usability principles (“heuristics”). In this way, the evaluated product is supposed to become more user-friendly.

During the evaluation process, nine experts from UDE and IMS reviewed the user interface of the application (one person belonging to the target group). Two of the experts had the possibility to interact with the platform simulating a use situation and inspecting all elements of the application. In this context, the think-aloud-protocol-method (Lewis, 1982) was applied which means that the evaluating person has to say whatever he is looking at, thinking, doing or feeling while going through the platform. This helps identifying problems, misunderstandings etc.

Figure 11 - Expert evaluation in the inHaus

During this process, the other experts observed the interaction process and wrote down the usability issues they identified. Afterwards, all experts commonly discussed the existing

problems and compared them with the given heuristics. The whole evaluation process lasted two hours, followed by the interpretation and elaboration of suggestions for improvement.

The results were recorded (in a written form as well as with a video camera). The video helped to analyse and reflect the process afterwards and identify further, more general problems, e.g. with the navigation structure, while interacting with the widget.

The heuristics that were used during this evaluation session are described in the next section.

3.2 Existing Guidelines

There are a lot of existing guidelines or heuristics that can be considered for the heuristic evaluation of the FoSIBLE platform. In the following, some examples that give a good overview about important guidelines or heuristics that can be taken into consideration for the evaluation are listed (here, we focused on guidelines that are related to Social TVs):

- The "Sociability Heuristics for Interactive TV" of David Geerts (2009)
- The "Samsung TV Application SDK UX Guidelines" of Samsung Electronics
- The "Style guide for the design of interactive television services for elderly viewers" of Carmichael (1999)

The complete list of guidelines or rather the heuristics (in the following "heuristics") can be found in the Annex. For this evaluation, it was decided to use the "Sociability Heuristics for Interactive TV" (Geerts, 2009) as well as the "Style Guide for the design of interactive television services for elderly viewers" (Carmichael, 1999). Since the FoSIBLE platform aims at fostering social interaction between elderly people, the heuristics of Geerts were important in order to identify problems associated with the social use of our Social TV platform. Furthermore, these guidelines are well established. In addition, the style guides of Carmichael were used because of its direct focus on our target group – elderly people. In the next two subsections, we dwell on both heuristics.

3.2.1 Sociability Heuristics for Interactive TV

The guidelines of Geerts (2009, p. 173) present a set of heuristics especially designed for interactive television that are supposed to test the social uses of an interactive TV application.

They include the following 12 heuristics:

1. Offer different channels and levels for communicating freely
2. Use awareness tools for communicating availability

3. Allow both synchronous and asynchronous use
4. Support remote as well as co-located interaction
5. Exploit viewing behaviour for informing and engaging other viewers
6. Give the user appropriate control over actions and system settings
7. Guarantee both personal and group privacy
8. Minimize distraction from the television program
9. Notify the user of incoming events and situation changes
10. Adapt to appropriate television program genres
11. Let users share content flexibly
12. Encourage shared activities

According to Geerts, interactive TV systems should offer different channels to communicate allowing communication on different levels (quick responses to free communication) and integrate different awareness tools, which show, e.g., whether a user is available to interact, as well. Furthermore, the user should be allowed to use functionalities of the platform both synchronously and asynchronously; and it should be possible to interact co-located (multiple users at the same location) as well as at remote locations (simultaneous with different users at different locations). The heuristics further state that the information about the viewing behaviour of the users should also be used to create functionalities that foster social interaction. Besides that, the user should have sufficient control over his or her actions and it should be possible for the user to customize the system to his or her needs. Privacy is also an important aspect. In this context, Geerts mentions that the system should enable personal and as well group privacy. Distraction from the television program should also be minimized through special design features. Regarding the program genres, the developers should take into consideration that there should be different settings suitable for different TV genres. Another issue is that users should be notified about incoming events (e.g. via visible or auditory signals). In regard to possibilities to share content and activities, Geerts states that users should share the content flexibly meaning that users should be able to send content from different devices, and that the user should start and maintain shared activities easily (Geerts 2009, p. 154-172).

3.2.2 Style guide for the design of interactive television services for elderly people

The style guide of Carmichael (1999, p. 94-99) is a framework to design interactive TV systems for elderly people. These heuristics are very important for the project because they are catered to our target group and take the characteristics and needs of elderly users into consideration.

For a better structure of these guidelines, we arranged the heuristics in the following three sections:

- Visual elements of the interface
- Navigation through the interface
- Control devices

The first part of the heuristics concentrates on visual elements such as the text size, the on-screen presentation and its layout, meaningful icons, forms of highlighting and the design of interaction tasks.

In the second part the focus is on the navigation process. This includes that users should have the possibility to notice their mistakes and that these mistakes can be corrected in a suitable manner. The menu, keyword or other forms of search methods should also be suitable when the user is confronted with a large catalogue of items. In this context, the system should give a good overview presenting also the network the users are navigating through. To facilitate complex interaction tasks, interactive demonstrations should be offered to train novice users.

Furthermore, in the third part, Carmichael's guidelines deal with the control devices that should allow an easy usage by all users. It is also mentions that a variety of different control devices that are all compatible with the system should be taken into consideration.

3.3 Results

After transcribing the recorded video content and joining it with the taken notes, all feedback and the statements of the experts were put together and assigned to the different heuristics. The results regarding both heuristics can be found below.

3.3.1 Results regarding the Sociability Heuristics for Interactive TV

In the following, the results regarding the "Sociability Heuristics of Interactive TV" (Geerts, 2009) are described for each of the twelve heuristics containing in each case the summarized feedback of the experts. This includes positive and negative observation results as well as suggestions for improvements of the experts.

1. Communication modalities:
The FoSIBLE widget allows using different communication modalities, e.g. chatting, private messages or club entries. The problem that was identified in this context was that the users do not have the possibility to see the messages they have received or sent again. Furthermore, it was criticized that the chat overlays the TV program what is counterproductive regarding the concept that the TV program induces themes for

communication. Besides that, it was requested that users can only chat with the people in their buddy list instead of chatting with all people who are watching the same program as the user.

2. Presence and awareness:
The main awareness tools that were integrated in the platform are the buddy list that shows the online friends as well as what TV program and TV show they are watching currently, and the list of recent activities of the user. This was evaluated as good. The only negative aspect was the absence of opportunities to configure special settings, e.g. who can see what the user is watching.

3. Synchronous and asynchronous use:
The FoSIBLE widget allows communicating synchronous through a text chat and asynchronous through private messages or group entries. One point of criticism of the experts refers to the fact that the user cannot decide when reading private messages: the system creates automatically a popup window with the incoming message.

4. Remote vs. co-located interaction:
Since the project wants to foster social interaction between different, maybe lonely and isolated, elderly, the platform only supports the remote situation. Nevertheless, remote TV viewing is also possible.

5. Information about viewing behaviour:
A Buddy List gives information about the viewing behaviour of the registered friends including the TV program and channel they are currently watching. Besides, the information of the viewing behaviour should trigger interaction in the chat. This was rated well by the experts.

6. User control:
At the moment the user does not have the possibility to control system settings.

7. Personal and group privacy:
At the moment, the widget does not allow the user to configure any privacy settings. It was recommended that the user should have the possibility to adjust who can see what he is currently watching in the buddy list. However, if the user does not want to share any information, he can turn off the widget.

8. Distraction:
This heuristic was difficult to evaluate because during the test session it was not possible to receive a TV program. Nevertheless, the chat was considered as being critical in this context because it would overlay most of the program. If the users want to watch TV they have to turn off the chat. Furthermore, the popup windows could also disturb.

9. Notifications:
Different popup windows notify incoming events, messages or recommendations so that the users are always informed about system changes.

10. Adapt to appropriate TV genres:
At this state of the development, the users could only turn off the widget when they are disturbed while watching TV.

11. Sharing content:
The FoSIBLE platform allows the users to share their content flexibly, for example through TV recommendations, private messages or group entries. Furthermore, there is the possibility to add commentary to the group entries and the content can be sent from different devices like the remote control or the tablet.

12. Sharing activities:
The FoSIBLE users are encouraged watching TV (virtually) together by TV recommendations. These TV recommendations could also be starting points for communication.

3.3.2 Results regarding the Style Guide for the Design of iTV Services for Elderly Viewers

In the following, the results regarding the "Style Guide for the Design of iTV Services for Elderly Viewers" (Carmichael, 1999) are described for the three divided sections. Each section contains the summarized feedback of the experts including positive but also negative observation results as well as suggestions for improvements of the experts.

1. Visual elements of the interface:
The text size of texts of the widget was evaluated as being good, however referring to the presentation on the tablet it was mentioned that the text could be larger, especially because of some free space that could be used, too. Nevertheless, both single words and text presented for continuous reading are legible und clear and the contrasts were rated as being good. The on-screen presentations were not overfilled. Nevertheless, the experts criticized that the coloured buttons integrated in the widget take up too much space. In addition, the labelling of the buttons is also not meaningful enough. Another identified problem was the highlighting of an element when it is selected and that it does not become clear which element is actually selected. Concerning this issue, it was proposed that only the colour of the font could change from grey (not selected) to black (selected).

2. Navigation through the interface:
In regard to the navigation, it was mentioned that the navigation structure has to be more consistent. This belongs mainly to the labelling of the four coloured buttons that change their meanings irregularly. Furthermore, some buttons, like the red return button in the TV

guide, have no functions and induce confusion. Besides, it is not apparent how the user can switch to the program list in the TV guide. Another heuristic in this context is the opportunity to notice any mistake the user makes. A problem that occurred is that the widget will be closed when the user presses the return button by mistake: a dialogue like "Do you really want to close the widget" that allows to undo this mistake does not exist. Besides, a wrongly posted recommendation cannot be reversed. The experts also criticized that the widget should show where the user is in the navigation structure to relieve the burden on older people's memory. For the explanation of the interaction with the interface a guided tour could be helpful.

3. Control devices:
The users of the FoSIBLE platform can operate with different input devices like the remote control, a tablet PC or gesture recognition what is beneficial for the demands on different skills of elderly people. During the heuristic evaluation only remote control and tablet could be tested. The most important aspect that has been identified is that it is not evident which input device is supposed to be used and when. Furthermore, the TV does not provide feedback whether an input action (e.g. text input) was really executed.

3.4 Implications / Suggestions for improvement

The main reason why this Heuristic Expert Evaluation was conducted was that the FoSIBLE widget should run stable and free of bugs when tested in the real household Living Labs. Especially, for the identification of fundamental usability problems this method is well suited.

Based on the evaluation results, the collected data have been interpreted and suggestions for improvement have been elaborated. The results and suggestions for improvements were discussed with the developing partners as well as with the partners responsible for the field evaluation and thus having contact to the end-users.

The improvements have been rephrased as bugs or feature requests which have been added to the Flyspray bug-tracking tool as well as to a shared work document, so that they could be considered in the further development process.

The bugs and feature requests were categorized according to the different functionalities of the FoSIBLE widget: overall, clubs, chat, stay in touch, wall, TV guide. Furthermore, the priorities of the different bugs and requests were discussed by the partners and added to the shared work document. In this way, the development partners had a better overview of the problems and requests they had to fix before the final testing in the Living Labs.

In the annex, the bugs and feature requests for the six subdivisions with the particular prioritization are listed.

4. Evaluation of the Sensor Technologies in the Fraunhofer inHaus with members of the target group

A central part of work package 6 is the usability evaluation in the Fraunhofer inHaus with members of the target group. Since the FoSIBLE widget with its input devices (remote control and tablet) will be tested during the field evaluation tests in the Living Labs in Germany and France, this evaluation focused on the concept of sensor technology that was developed in the project. For cost reasons, these sensors could only be installed in the Fraunhofer inHaus and not in all Living Lab households, which is why it was decided to evaluate these concepts (see also D3.6) in a separate but complementary inHaus study.

For the preparation of the evaluation, UDE worked together with AIT and IMS in terms of the development of different social sensor use cases and the appropriate installation of the required sensors. The main aspect of the social sensor concepts is to trigger social interaction and social events by using an easy and intuitive input method consisting of sensors and actuators which have been seamlessly integrated in the furniture in order to avoid the feeling of being watched. These sensors can be used for the collection and provision of awareness and context information. D3.6 gives an overview on the sensors and the developed technologies.

In a number of projects, this way of sensor integrations is often used for the monitoring of the health status of elderly and not for social interaction concepts. Therefore, in the following we will first give a literature overview on the usage of the sensors in the context of social interactions for creating a theoretical background for the evaluation and for showing which different concepts can be realized with sensor technologies and Smart Homes. Subsequently, we report on the evaluation procedure and experimental set-up as well as on the participants invited for the study. At the end of this chapter, the results are presented and summed up.

4.1 Background

A main aspect that stresses the relevance of our evaluation of social sensor technology is the fact that only 19% of the Smart Home projects use sensor devices in order to increase interaction. Most of time, these projects use this technology for the monitoring of functional (71%), safety (67%), or physiological (47%) properties (Demiris & Hensel, 2008). For this reason, there are few studies that deal with the fostering of social interaction through integrated sensors and actuators what is therefore the main focus of the inHaus study.

Through research, it has been revealed that elderly often have a positive attitude toward sensor technologies. However, it is very important that the monitoring process is non-obtrusive, meaning that the users do not want to have the feeling of being watched and

observed (e.g. Demiris et al., 2008; Demiris et al., 2006). This feeling especially arises when a video sensor is being used what has been criticized by elderly users because of privacy concerns (Demiris et al., 2008). These results show on the one hand that it is necessary to consider the concerns of the target group for the development and the build-up of such sensor systems. On the other hand, our goal of integrating the sensors in a seamless way in the furniture of a household seems to be a good approach.

Furthermore, Zaad & Allouch (2008) found out that people who perceive more control over their well-being through the use of the sensor system also show the intention to use it more often. This indicates that elderly users, who experience the interaction with the developed interaction concepts through sensors as positive and easy, could also be interested in using the system further.

Based on this literature review, an inHaus study was planned and conducted, testing the concept of social sensor applications in order to trigger social events and interaction. The details of the evaluation are described in the following.

4.2 Evaluation

4.2.1 Research questions

As the conducted study was planned as an exploratory study in order to get a general impression about the opinion of elderly regarding social sensor applications and to identify the potentials and problems accompanied with its use, a number of rough research questions that should be answered through the evaluation were developed:

- What impression do the elderly have of the concept?
- Is the concept of social sensor applications useful and understandable?
- Do elderly people have problems with the interaction with the social sensor applications (what were problems) or was the interaction perceived as intuitive?
- According to the participants, what are benefits and potentials of social sensor applications for them?
- What are suggestions for improvement for the developed concept of social sensor applications? What further functionalities do the elderly want to have? Could the elderly imagine further use scenarios for their daily life?
- Could the elderly imagine using such a system in their daily life?

4.2.2 Participants

For the evaluation, 15 participants of the AlterAktiv e.V. computer club were invited. Six of the participants take part in the FoSIBLE Living Lab field tests and already have experiences with the interaction with the FoSIBLE platform. They use the system since four weeks on

average (Range: 2-12 weeks). Two of them use the platform a few times a day, one participant several times a week and two participants more seldom because of the fact that not all of their friends have been connected yet.

On average, the participants were 69.73 years old (SD = 4.758, Range: 62-77). Ten of the participants were female. In regard to the highest educational level, three participants have a graduation from boarding school, two have a secondary modern school qualification, five a secondary school certificate, one a higher education entrance qualification and four a university degree.

Furthermore, most participants watch television regularly. 11 participants watch TV once or a few times a day, three participants several times a week and only one participant more rarely than several times a week. The average time per week of watching TV among the participants is 18.29 hours (SD = 8.974).

It was also asked whether the elderly use social networks. Five of them affirmed. Two participants use social networks a few times a day, one participant once a day, one participant several times a week and one participant rarer.

Another important aspect for the evaluation was the willingness to use technology. This was measured with the scale of technology commitment of Neyer et al. (2012) that divides technology commitment in three subscales: technology acceptance, technology competence and technology control. We found that the participants have a high technology acceptance (M = 3.59, SD = 0.776) and a high technology control (M = 3.64, SD = 0.801). Technology competence can be located in middle range (M = 2.3, SD = 0.839). This results in a middle technology commitment (M = 3.17, SD = 0.542).

4.2.3 Method, User-Centred Evaluation Design and Ethical Issues

In order to assure that the evaluation was adequate for the older target group we tried to develop an experimental setting which is appropriate for the participants. Barrett & Kirk (2000) described several major issues and aid to support experimental settings with older persons. In order to foster group interaction and to get a better insight into the participants' experiences and opinions, we set up a "hands on" interaction scenario with the system in combination with a focus group discussion (Barrett & Kirk, 2000). In order to make the questionnaire and the interview as user-friendly as possible, we tried to keep the questions short, focusing only one issue at a time, and to use simple language. Furthermore we tried to keep the materials and letters as large as possible in order to support visually impaired participants. We asked for persons who have difficulties with hearing, but there was no participant in that sample. Considering minor needs of elderly, we planned large breaks, supplied meals and refreshments and offered constantly the opportunity to take a seat. We

also keep the duration of the evaluation session not longer than 1 hour. We also insured that there were at least two investigators present during the evaluation, to secure a high availability of support for the participants. In that case, we also had one well known person from USI in each room, to make sure that there is a contact person for the participants, they already know. Especially during the focus group, it was important that the participants felt comfortable in order to express the negative statements as well.

In regard of ethical issues, all aspects of the evaluation were in line with the declaration of Helsinki and the German ethical guidelines of the APA. In addition, the evaluation was approved by the ethnical commission of the University of Duisburg-Essen. Before the start of the evaluation sessions, the participants were informed that the experiment is fully voluntarily and that it could be broken up at any time without giving reasons. Furthermore, they were explained that their answers and logged interactions during the evaluation would be treated anonymously and that it could not be traced back to them. Knowing this, an informed consent was signed by all of the participants.

4.2.4 Experimental set-up and procedure

As interviews and focus groups are established instruments in this field and offer the opportunity to get insights in the fears, concerns, demands, and problems regarding such sensor-based interaction technologies, these evaluation methods were used for the inHaus evaluation.

For giving the participants a background for discussion, the platform with its integrated sensor functionalities was presented to the elderly in a first step of the evaluation phase. It was important to give them a general impression of the new components as they only knew the main functionalities of the platform through their participation in the Living Lab evaluation.

Subsequently, the elderly had time for one hour to interact with the platform. This interaction part was combined with different tasks the participants had to fulfil. Six tasks were communicated (see Figure 1 and Figure 2):

1. Get information about a specific book.
2. Recommend a book to the community group.
3. Invite a group of people to a game night.
4. Get an invitation for a game night or a walking tour.
5. Invite a group of people for coffee and cake.
6. Invite a group of people for a walking tour.

For the first scenario the participant had to choose a book and hold it in front of a sensor, which was installed in front of the TV. The application displayed the information as an

overlay to the normal screen on the TV. To recommend a book the participant had to hold the book in front of another sensor, which is on top of the cupboard and on the right side of the TV. When the sensor detects the book, the application opens a dialog, which asks the user to enter the recommendation and to confirm or abort it. The scenario for the invitation to a game night has been realized by a sensor in the shelf of the cupboard. The participant had to take out the game and a dialog was opened on the TV screen, which asks for conformation as well. When the participant got an invitation, he or she had to confirm or refuse it. To send an invitation for coffee and cake, the participant had to open and close the kitchen closet. In order to send an invitation to a go on a walk together, the participant had to take his or her walking stick or umbrella from the wardrobe. After those interactions the invitations had to be confirmed or refused as well. For the interaction with the smart TV the participants could choose between the normal remote control, three sensors in front of the TV and the gesture recognition, tracked by the UCOS-Sensor (see deliverable D3.6 and Figure 2).

During the interaction of the participants, one of the investigators triggered prepared actions via tablet input for simulating the interaction with a friend of the buddy list on the TV. In this way, several interaction patterns with the platform could be simulated, like answering an incoming invitation, sending invitations or recommendations, etc.

Figure 1: The scenarios "information about a book", "recommend a book", "invite to a game night" and "invite to coffee and cake".

Figure 2: The interaction with sensors to confirm an invitation and a screen cast showing the "sending an invitation for a game night" dialogue.

Subsequently, a focus group session was started with all participants and the experimenters. To allow all participants to join in the discussion intensively, we divided the participants in two groups (N_{Group1}= 8, N_{Group2} = 7). For the discussion phase, we prepared different questions investigating how the elderly experienced the interaction with the sensor applications and getting an impression of advantages, disadvantages and suggestions for improvement.

The following list gives a general overview of the questions asked during the focus group session:

- How did you experience the presented sensor technologies? What do you like, what do you not like?
- How did you experience the possibility to display information about a book by bringing the book near the sensor?
- How did you experience the possibility to recommend a book to your friends through putting it on the sensor?
- How did you experience the possibility to invite friends to specific events? Which procedures do you like more (automatic recognition or active proximity to a sensor)?
- How did you assess the possibility to ask others automatically for a common activity through the sensors (e.g. visit the senior club)?
- How did you experience the gesture control when using the platform?
- What is your opinion on the presentation if you are available, in community or not available in the buddy list? What would be your experience if your contacts would see this information?
- Did you have concerns regarding privacy? Which factors did these concerns depend on (e.g. availability display, contacts in the buddy list, kind of information)?

4.3 Results

The evaluation in the Fraunhofer inHaus sensor environment with the target group includes a behaviour observation while the elderly were performing the user tasks, as well as the subsequent focus group sessions. Both results are described in this chapter.

4.3.1 Results of the behavioural observation

In the following, the impressions of the participants during the user tasks will be presented. First the general handling of the participants with the system is described. After that, we will give an overview about the statements the users made during the interaction with the system.

General Handling:

Generally, all participants were curious to learn something about the unknown technology. The different tasks were more or less easy to understand for them. Most of the participants had no problems to get information about a specific book. Just two participants tried to get the information with the wrong sensor. To post a book in a specific group, most of the participants mixed up the sensors and tried to interact with the sensor, which gives information about a book. Furthermore, one participant tried to activate the sensor with her hand instead of a book. These mistakes occurred in the scenario to post a game as well. But there were also participants, which understood the system instantly. There occurred no problems during the handling with the invitations to specific activities as well as during the interaction with the kitchen sensors. It was also easy for them to refuse an invitation with the different input methods like remote control, gesture recognition and light sensors. Especially, the interface was comparable for them to other social applications they knew, which lead to a navigation through it free of problems. Just one participant tried to confirm her game invitation in the wrong dialogue.

Statement:
One of the participants suggested a progress bar for the detection of the book in order to get feedback of the system, whether a book is detected or not. Another participant misses a feedback when she finished recommending a book. In some cases the concept of feedback wasn't comprehensible for the participants, e.g. when they were supposed to refuse an invitation they had to confirm it after the rejection.
Most of the participants recognized their status in the application. Two mentioned that it would be helpful to see, whether another person is online or not like in the social community Facebook. In that case, it was not understandable that the widget is working asynchronously (see chapter 5.3.2) and that the system would run as a background application at home. Some participants tried to logout of the system and mentioned that

they are afraid that the application could be hacked and an abuse of presence could happen if one cannot log out.

One participant notified that the wording could be different. “Instead of ‘sharing’ an invitation it would be better to report or communicate an invitation”.

4.3.2 Results of the focus groups

In the following, the different results of the focus group evaluation will be presented, structured in general impression regarding the social sensor applications, the different integrated functionalities, gesture interaction, privacy and further ideas of the participants.

General impression
Most of the participants stated that social sensor applications are interesting in general. One participant mentioned the thought that the social sensor applications could be especially useful for the close environment, as appointments with friends for different activities are possible, while already existing social networks are more useful for staying in contact with distant living friends and the family.

Nevertheless, five participants say that they would not need the social sensor applications at the moment as they are still very mobile. They agreed that it would probably be more suitable for elderly who are bedridden or not mobile anymore. However, the participants were of the mind that it is necessary to deal with the software and sensors early in a stadium when elderly are still up to the mark as otherwise it would be difficult for them to use it later when it is needed. This often appearing lack of experience was emphasized by many of the participants: a lot of their friends are not interested in technologies or are frightened at new developments and therefore refuse to face up to technologies like smartphones, tablets or smart TVs.

Sending and receiving invitations
The majority of the participants rated the possibility of sending and receiving invitations through the sensors well, however with some restrictions:
It was the overall opinion that a confirmation is necessary. On the one hand it is important to know whether all friends received the invitation; on the other hand there should be a response dialogue about who will take part in the event. Furthermore, the sensor invitations were only called suitable for a fixed group, e.g. the hiking group or the walking group, as the group can be easier informed about dates and locations of meetings in this way. One participant mentioned in this context that he is member of a hiking group. The director of the group normally informs most of the members via email about the hiking route, but as some of them do not have an email account, he has to phone them in addition. Sensor invitations would facilitate this process.

Furthermore, it seems to depend on the kind of invitation. One participant perceived an invitation to her own home as being sensible and personal and therefore, she would want to do this via telephone. However, when a person wants to meet friends while going for a walk, the social sensor application is useful as it allows a fast and direct communication. Telephone would last too long here.

Another aspect that was mentioned was, that the responses given by the people which were invited for an activity that are provided by the system at the moment are not meaningful enough, as only a "yes" or "no" feedback appears. For the participants it is important to know why another person cannot comply with the invitation.

Further events which the participants could imagine using the sensor functionalities for are, e.g. going to the cinema or the theatre, organising a barbecue or also for finding or giving a ride.

Recommending books or getting information about books

Further tested usage scenarios were the possibilities to recommend a book to a friend or to gather information about a book by putting it on the adequate sensor. These functionalities were criticized as there are some limitations. First, this process would not be flexible enough as every book has to be equipped with a sensor chip. In addition, this would not be useful when a person wants to get information about a book he or she heard off and that he or she does not have at home. The same accounts for the case, that he or she only has the electronic version of the book. Then, a recommendation would not be possible.
Nevertheless, the idea behind this concept seems to be interesting. A participant proposes to integrate a group for the exchange of books, like an online swap meet. By putting the book on the sensor, the user can send a request into the group in which he or she can ask if someone else wants to read it.

Gesture interaction

It was also very important how the target group experienced gesture interaction for affirming or negating incoming messages or events. Only one participant answers that the concept was unfamiliar. In general, it was received favourably that different forms of interaction through remote control, tablet, gesture and sensors would be possible. According to the special situation or need, different input methods could be used. In this context, the limitations of other input devices like the remote control were addressed. The opportunity of choice seemed convincing. Further positive aspects in this context were the easier accessibility for elder or disabled persons and the intuitive concept behind the gestures as they are similar to gestures used in daily life.

Availability status created through person counting

The main idea of the availability status was that it would show if a person has visitors. It was the assumption that then he would not have the time to talk or to undertake something. However, it was shown that this assumption cannot be generalized. According to the participants the term “in community” is not meaningful as this would not directly indicate whether someone is not available. One participant, for example, explained that he - when he knows that his brother is at his parent’s home - wants to join them. The expression “not available” or “I do not have time” would be more explicit. These terms should be set manually and there should be different variants for availability messages.

During the discussions an interesting idea came up for using the sensors for the availability status: different cards could be used for different status displays and put on corresponding sensor surfaces when required. This idea was approved by the other participants.

Privacy

Another central focus of the conducted evaluation was privacy. The strong participation of the elderly at the discussions regarding this topic showed that it was very relevant for them. A general fear to be spied on out in various situations could be identified, e.g. when surfing in the internet, doing internet banking, sending important emails or paying with credit card.

Privacy issues concerning the social sensor applications referred only to the camera for person counting and the persons the invitations and recommendations were sent to. In the first place, it was important for the participants that they would be able configure who and how many people they invite and who can see the information posted on the social platform. Only one participant mentioned that he would not have a problem with strangers also having access to this information.

The main concerns regarding the camera existed for person counting as there is the fear to become spied on by others. In contrast to cameras of laptops, the UCOS sensors capture the whole room and (if also installed in other rooms) also other locations in the household. Therefore, a mechanism to turn off the camera (e.g. toggle switch) or hide it behind an occlusion was desired.

One participant also disliked the appearance of the availability status on the platform since criminals could maybe monitor the different status messages, derive patterns in daily life and abuse this information.

Interface

Although it was not the main topic of the evaluation, the participants also evaluated the interface of the platform. According to them the design was nice and clear, the navigation

was easy because of parallels to existing systems and the support through the different colours, and the use of the system was intuitive.

Further ideas of the participants

During the focus group sessions the participants also came up with own ideas for the improvement of the system and possible further functionalities. They mentioned that they would like to have the opportunities to share photos with others and to hear audio books together at remote locations. In addition, they would like to have the functionality to arrange meetings like it has been realized in doodle. At last, it should also be possible to create invitations through in the system by the use of the tablet in addition to the automatic process through the sensors.

4.4 Conclusion

As the developed concept of the social sensor applications could not be tested in the real household Living Labs, it was important to investigate their potentials in another user study. The conducted evaluation helps to identify benefits of the sensor functionalities, but also shows restrictions that have to be considered for the further development.

First of all, it has to be bore in mind that the participants from the AlterAktiv e.V. computer club in Siegen were very active and mobile seniors. Their statements in regard to that they would not use the sensor functionalities at the moment should be considered in this context. Nevertheless, they emphasized the necessity and potentials for bedridden and less mobile persons and could imagine using the system in a few years when their physical fitness has decreased. They especially recognized the benefits for appointments with people living in the direct, near environment, especially when they are spontaneous or when the invitations are for groups with a fixed set of persons.

Sometimes it was hard for the participants to understand the context of the sensors and their functionalities, so they mixed up some interactions. This could have been caused by the novelty of the interaction with the sensors and the new application. On the other hand, a fast increase of the learning curve could be detected. Thus, considering the handling by the participants it could be discovered that the system is intuitive but needs some explanation, especially with the sensor interaction. Referring to this circumstance, it is required to add some kind of tutorials or help functions in the system to make it possible to learn the application during the interaction.

Another important aspect that has to be implemented in the system, is, on the one hand, a more detailed feedback in regard to the information whether everyone that was addressed has received the invitation and the possibility to post a response with further information, e.g. “I cannot take part because …” on the other hand.

The concept of getting and sending book recommendations has to be reconsidered in regard to books that are not physically in the households. It was asked to have a group in which books one only heard off can be described and recommended. As this is already implemented in the FoSIBLE platform but was not used in this evaluation because of the special focus on the sensor functionalities, this gives hints that a combination of both – using the sensor surfaces for book recommendations as well as the book club – is promising.

Regarding the availability status it has to be considered which other designations instead of "in community", "alone" etc. would be more expressive. Furthermore, several configuration possibilities for the status should exist, e.g. by using different "availability cards" that can be put on a corresponding sensor surface. In addition, it should also be ensured that the users can configure the setting of who can see the status in the buddy list. In this context, it would also be eminent to implement a dialogue that appears when more than one person is being recognized in the room that asks whether the user wants to change his availability status. Furthermore, this information about visitors could be used for collective events, e.g. by enabling a user who has two visitors to send this information to another friend, together with a message: "Come and join us, too". In consideration of the participants' privacy concerns regarding the camera, a cover will be developed for the UCOS sensors by Mauser that should provide a better feeling of security when using the system. At last, as the feedback for the gesture interaction for basic interactions with the platform was very positive, this approach should be pursued.

5. Empirical Interface Evaluation for Playful Interaction

In deliverable D6.1, we reported on the evaluation of the playful Social TV application "Gameinsam", which has been developed by researchers of UDE. This evaluation mainly focused on the creation of social presence and connectedness, user experience, and user acceptance with an intergenerational approach. In contrast, the foci of the evaluation described in the present deliverable are usability and design aspects considering the physical and cognitive restrictions of older people. Therefore, an appropriate interface design is significantly important. Thus, several usability heuristics were considered during the re-design process of the application, which will be discussed in the following.

5.1 The application

As described in D6.1, "Gameinsam" (Herrmanny et al., 2012) is a widget for the Samsung Smart TV, developed by the researchers of UDE. It offers playful remote interaction among family members and peers which refers to the current TV program, e.g. a quiz show, as part of the playful interaction. The application offers the opportunity of "shared shoutability" (shoutability = the need to e.g. answer questions (aloud) while watching quiz shows), allowing each participant to watch the program at home and share his or her guessed answer using the standard remote control. The four coloured buttons of the remote control represent the three possible choices and a question mark, which may be used for signalising the other player(s) that one doesn't know the correct answer. The information which family members do watch the same program and which answers they choose is displayed in a buddy list. Players can correct their answers all the time as long as the question is active. This allows them to react to the answers of the co-players as they play together. Family members commonly achieve joint high scores offering a collaborative playful interaction. When the solution of a question is given in the program, the question is set inactive, the correct answers are coloured green, the others red and the family score is updated. When there is a new question in the TV program, the interaction is set on active again, so that the users can make their input.

Figure 3 – Interface of Gameinsam, question active and solution

For the evaluation a bot has been integrated in the application. It simulates the co-player's answer inputs and answer switches in a realistic manner. Moreover, a videotaped game show was used in the prototype of the application during the evaluation runs instead of real TV program.

5.1.1 The Re-Design

In order to optimize the application for the special needs of seniors, a heuristic evaluation based on Nielsen's Usability Heuristics (Nielsen, 2005) and the Style guide for the design of interactive television services for elderly viewers (Carmichael, 1999) was performed. It identified some potential for improvement with respect to an older target group. Based on this, a re-design has been made by the researchers of UDE which included issues like the reduction of the cognitive load by reducing the number of elements, integrating a help menu, replacing some of the wordings with respect to the language of elderly, changing colour and contrast, optimizing the arrangement of the elements, integrating new feedback icons etc.

Figure 4 – Re-designed Interface of Gameinsam

5.2 Evaluation

5.2.1 Research issues

The aim of the study was to investigate if the re-designed interface of Gameinsam meets the demand of older persons, but also the acceptance of this kind of playful interaction.

In detail, it covered the following issues:

- The participant's general opinion regarding this playful kind of interaction

- Their preference regarding conceptual issues (number of co-players, collaborative or competitive gameplay)
- How comprehensive the game elements are for the target group
- Their aesthetic perception of the design in general
- The opinion towards arrangement of the interface elements
- If all elements are clearly recognizable
- Issues regarding the concentration

5.2.2 Method and Procedure

In order to investigate the topics presented above, in a first step the participants were instructed regarding the rules of the game. Afterwards, they used the Gameinsam application together as a group. This test lasted about 10 minutes and covered three quiz questions. During the use, the participants who were in possession of the remote control had to consequently type in the answer. The distance of the participants to the TV screen ranged from approximately 2 to 5 meters, depending on the different seating locations of the group members. This is supposed to be a distance usual in normal living rooms. The size of the smart TV screen was 55”.

After playing the game, the focus group method (see 4.2.3) was used to gather the opinions of the participants regarding the research issues described above.

The focus group was based on semi-structured guidelines, which have been established before the evaluation.

5.2.3 Ethical Issues

Analogous to the study described in section 4, ethical issues and specific demands of the elderly persons have been considered. All participants were volunteers and gave their informed consent. All aspects were in line with the declaration of Helsinki and the German ethical guidelines of the APA. Moreover, the study has been approved by the Ethics Comitee of UDE.

5.2.4 Participants

For the evaluation, 8 participants of the AlterAktiv e.V. computer club of Siegen were invited. Three of the participants took part in the FoSIBLE Living Lab field tests and were therefore already experienced in regard to the interaction with the FoSIBLE platform. They used the system since two weeks on average (Range: 2-3 weeks). One participant used the platform once a day, one several times a week and one participant more seldom.

On average, the participants were 68.75 years old (SD = 4.234, Range: 62-74). Four participants were female, four male. Regarding the educational level, two participants had a graduation at board school, one had a secondary modern school qualification, two a secondary school certificate, one a higher education entrance qualification and two a university degree.

Furthermore, the participants watched television regularly. All participants watched TV once or a few times a day. The average time per week of watching TV was 21.63 hours (SD = 7.425).

We also asked if the elderly would make use of social network communities. Three of them affirmed. One participant uses social networks a few times a day, one participant several times a week and one participant more seldom.

Another important aspect for the evaluation was the willingness to use technology. This was measured with the scale of technology commitment of Neyer et al. (2012) that divides technology commitment in three subscales: technology acceptance, technology competence and technology control. It was shown that the participants had a mediocre technology acceptance (M = 3.25, SD = 0.577), a high technology control (M = 3.57, SD = 0.450) and a low technology competence (M = 2.16, SD = 0.823). This results in a middle technology commitment (M = 2.94, SD = 0.438).

5.2.5 Results

The results show:

The participant's general opinion regarding this playful kind of interaction

All participants spontaneously found the playful interaction with Gameinsam to be entertaining. Nevertheless, they mentioned that they would prefer face-to-face games if possible. One participant even said:

> *When I'm watching TV, I'm watching TV. When I'm playing a game, I'm playing a game. And when I'm drinking my coffee, I'm drinking my coffee.*

Further comments revealed that for all participants the willingness to play Gameinsam mainly depends on the context: All of them think that the playful remote interaction is a good solution for interaction with family members and peers to bridge distance. Moreover it is seen as sense-making if a person is not able to leave the house anymore. One participant also found it positive for spontaneous interaction, when he sees that a peer is currently watching the same TV show.

The participant's preference regarding conceptual issues

The participants agreed on preferring a competitive gameplay (instead of cooperative) or at least the presentation of the personal score.

Regarding the optimal number of co-players one participant suggested four persons and the others agreed.

Comprehensibility of the game elements for the target group

We asked the participants about the comprehensibility of the different interface elements, such as visualisation of correct/incorrect answer, wordings and the calculation of the score. All these elements were stated to be clear and appropriate by the participants. Only one participant suggested "thumbs up" and "thumbs down" as an alternative for the feedback regarding correct/incorrect answers. The others preferred the implemented alternative.

Aesthetic perception of the design in general

The overall valuation of the design was very positive – without any exception. The participants claimed it as lovely and nice. They especially liked the choice of colours.

After the participants had formulated their first impression of the re-designed graphical user interface, they were shown two screenshots – one of the former and the other one of the re-designed Screen. We asked which they would prefer and why. Without hesitating, all of them found the green one (the re-designed alternative) more pleasant and more clearly arranged. One of them mentioned that the other interface was to dark and that the green one might be more eye-friendly.

Opinions towards arrangement of the interface elements

The arrangement of the elements in the re-designed interface was valued positive. Nevertheless, one participant mentioned that the right area (with the game information) would take up too much space. Another one suggested that - as there is enough space - the font size could be bigger using this blank area.

Recognisability of all elements

On a whole, the focus group as well as the observation showed that none of the participants had problems in reading the text and recognizing the other game elements. One participant mentioned that the info bar on the bottom of the screen could be a bit bigger as the text size is on the border to causing difficulties in reading.

An interesting aspect during the observation of the participants was, that no one seemed to have problems in recognising the screen elements, but some of them had to take there glasses when they wanted to use the remote control.

Issues regarding the concentration

Being asked about their focus of attention during playing, the participants said that it was the TV program rather than the game. Considering this, it is surprising that they typed in their answers for each question in time and said that they realized when the co-player changed his answer. In contrast, they did not pay attention on the score, except one participant, but only at the end.

It could be observed that the participants, once having typed in an answer, did not change it anymore.

5.2.6 Discussion of the results

The participants agreed in most of the topics in the focus group discussion. The results of the focus group and the observation show that the Graphical User Interface of “Gameinsam” meets the demand of seniors. This concerns physical and cognitive issues, namely the aesthetic design, the arrangement of the screen elements, the recognisability and comprehensibility of the screen elements as well as the cognitive load and concentration. Some potential for improvement would be the enlargement of the info bar on the button and the reduction of the width of the green “gaming area”. However, the last aspect technically depends on the aspect ratio of the TV show.

Regarding the acceptance or the perceived usefulness of the application, the participants told that this depends on the context of use. If possible, they would prefer face-to-face games. If playing Gameinsam, they would like to have a personal score.

6. ANNEX

6.1 List of Guidelines Used in the Heuristic Evaluation

6.1.1 Sociability Heuristics for Interactive TV of David Geerts

Communication modalities:

- Offer different channels and levels for communicating freely
  - Enable voice chat as well as text chat if possible, otherwise use voice chat for a broad user audience including people with little chat experience or text chat for a specific audience including people with chat experience.
  - Allow communication on different levels, from low-activity quick responses such as emoticons, gestures or automatic replies, to free-form communication
  - Make sure the communication process is optimally supported so it can go smoothly

Presence and awareness:

- Use awareness tools for communicating availability
  - Give information about the current behavior of other users
  - There can be several levels of presence and awareness indication, from a simple "someone is watching television" to a list of buddies that each have a status and indicate the channel and program names they are watching
  - These tools should indicate if a user is available to chat or otherwise interact, or if there are special circumstances (such as watching in group) other users should take into account

Synchronous versus asynchronous use:

- Allow both synchronous and asynchronous use
  - Provide different functionalities for interacting and communicating synchronously as well as asynchronously, so users do not always have to be using the system at the same time, but can also benefit from it when they are using it at different times.

Remote versus co-located interaction:

- Support remote as well as co-located interaction

  - As social interactive television systems can be used with multiple users at the same location (co-located) simultaneous with different users at other locations (remote), it should make sure its functions are appropriate for both situations, and not disturb the interactions between co-located or remote viewers

Information about viewing behavior:

- Exploit viewing behavior for informing and engaging other viewers
  - Use information from a users' viewing behavior not only for showing currently viewed programs, but also for creating functions that aid communication, social interaction or recommendations for other users.

User control:

- Give the user appropriate control over actions and system settings
  - Users should have sufficient control over their actions and system settings, so they can adapt the system in general or specific features to their needs or to the current situation.

Personal and group privacy:

- Guarantee both personal privacy and group privacy
  - Make sure the system enables users to ensure their own personal privacy, by choosing what (not) to disclose, as well as group privacy, by taking into account the presence of multiple viewers in a co-located viewing situation

Distraction:

- Minimize distraction from the television program
  - Design system features so there is not too much distraction from watching the television program. Specific tools can help users control if they want to be distracted or not. Distraction of other co-located viewers should also be taken into account and minimized.

Notifications:

- Notify the use of incoming events and situation changes
  - When users are requested to respond to an action of another user, notify them visibly or audibly of these incoming events. When there are changes in another users' situation, e.g. when switching channels or when going from watching alone to watching in group, users that are actively interacting with this other user should also be notified of this change in situation.

Program genres:

- Adapt to appropriate television program genres
  - Take into account the properties of television genres, and offer features or settings that are appropriate for these genres. The plot structure and social uses of certain television genres are the important qualifying factors for this. The system should be tailored to television genres that are more suited for the synchronicity of the interaction or for specific platforms.

Sharing Content:

- Let users share content flexibly
  - Make sure users can send content to and from different devices to share this with other users
  - They should be able to edit the content as they wish, and add some form of commentary to it
  - As much as possible, integrate sharing content with other features of the social television system

Sharing activities:

- Encourage shared activities
  - Allow users to easily start and maintain shared activities around the television content, such as communicating, watching together, choosing programs or controlling the content. Make sure sufficient information is provided so efficient communication and interaction is enabled.

6.1.2 Style Guide for the Design of Interactive Television Services for Elderly Viewers

- Text should be presented as large as is reasonably possible.
- Text presented as single words generally only needs to be satisfactorily legible for older viewers. Whereas text presented for continuous reading needs to be relatively clearer than simply legible to ensure adequate understanding of the content and its inferences.
- On-screen presentations should not be overfilled with information or otherwise "busy". Ideally, a single screen should contain a single message or a single activity.
- The layout of a screen presentation should be designed to make what it has to offer easily understandable to the user. This may also involve the use of explicit instructions.
- The meaning of any explicit instructions used should be checked with naïve users.
- Icons that are meaningful are more beneficial than abstract or arbitrary ones (although the meaningfulness should be previously established with users)
- Designers of screen layout and their elements, should consider using a simulated reduction of visual acuity to check the clarity of their design.
- Various forms of highlighting can be useful for drawing users' attention to important areas of the screen. But care is needed to ensure that the highlight is suitable, given the context.
- Highlighting and (lowlighting) can be useful for guiding users through a sequence of operations (locations) on a screen.
- Some interaction tasks that can fit onto one screen may be easier for older people to deal with as a succession of screens containing one operation (and possibly associated instructions) on each
- A variety of techniques can be used to constrain progress through an interaction task, which can also guide users and will generally minimize errors.
- It is vital that users are given the opportunity to notice any mistakes they do make and are given the ability to make appropriate corrections or alterations in as efficient a manner as possible.
- If an on-screen pointer is used it must be clearly visible to older users and easy to use accurately.
- Visual highlights and other events can often be usefully augmented by sound, which ought to be rich (ie not pure tones) and preferably meaningful.
- Consider allowing a certain amount of personal customization of some presentation characteristics.
- When providing access to large catalogues of items, careful consideration should be given to the overall suitability of menu, keyword or others forms of search method.
- Older users will find it helpful if they are given an (suitable) overview of any large body of information (including the network they are navigating through)
- Whenever possible relieve the burden on older people's memory by providing equivalent information on-screen.

- For inherently involved or complex interaction tasks, consider providing an interactive demonstration to train novice users and prepare them for the real thing. (However, every effort should have already been made to ensure that the interactive service is effectively 'walk-up-and-use' regardless of the level of ability or knowledge of the user.)
- All effort should be made to ensure that the presentation of interactive services and the operations involved in using them throughout a particular system, are as consistent as possible from the user's point of view.
- Give very careful consideration to the control device intended for use with the system, on the basis of easy use by all users (given the operations required of it).
- Consider the benefits of providing a range of different control devices which are all equally compatible with the system.

6.1.3 Samsung TV Application SDK UX Guideline: Design Principles for Creating Samsung Apps Content

Simple:

- Application is not always good just because it has a lot of features and information. TV screen should not contain too much information. The screen layout should be easily accessed by user friendly features, clear nd conveniently arranged to use.
- Operations and time for entering each level need to be minimized and available to control.

Clarify:

- Accurate navigation for user operation is the most crucial among various factors in TV application
- If navigation is ambiguous, users always feel insecure
- Especially, application should be designed for users to figure out where they are in the application.
- Actions of navigation such as Move, Return, Enter should be clear

User Control:

- Application should provide operational methods and corresponding intuitive structures appropriate for Input Device. The structural design of application needs to be optimized for remote control
- Actions on TV OSD (ON Screen Display) followed by input of remote control are needed to come up to users' expectations

Consistency:

- Of Button Operations: in case users intend to allocate features to each button such as color button or simple menu on remote control, if the suggested buttons (for

example Yellow: Check/Release) mentioned in the guidelines are allocated, application users are able to learn how to operate application simply
- Of Screen Layout: Screen should be composed to provide information on application effectively, but each one line area indicating title on the top and Navigation Help at the bottom should be secured.
- Of Interactions: the identical way of interaction should be maintained in basic functions or screen factors like consistency of display location and style suche as message popup, option popup, action window, that of directional navigation in contents list.

Feedback:
- When displaying an item which can be focused, focused status and selected status by operations from remote control are displayed separately
- When entering service, or the time to bring data in accordance with user input exceeds the specific criteria while using application. Put a loading animation on the screen

Aesthetic:
- Aesthetic design offers more convenience to users than normal ones

6.1.4 TenUsabilityHeuristics (Jakob Nielsen)

1. Visibility of system status:

The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

2. Match between system and the real world:

The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

3. User control and freedom:

Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

4. Consistency and standards:

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

5. Error prevention:

Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

6. Recognition rather than recall:

Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

7. Flexibility and efficiency of use:

Accelerators - unseen by the novice user- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8. Aesthetic and minimalist design:

Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

9. Help users recognize, diagnose, and recover from errors:

Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

10. Help and documentation:

Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

6.2 Feature Requests and Bugs identified through the Heuristic Evaluation

6.2.1 Feature Requests and Bugs regarding the overall system

	Feature Requests	Prio	Bugs	Prio
Overall	The Yellow login button can be misinterpreted, because of the yellow key of the remote control. Especially the Login button doesn't have a function. Please delete those buttons and replace them through a clear instruction, how to enter.	2	There is a long delay after you delete a friend - instead there should be immediate visual (and optional auditive) feedback to confirm the action.	2
	All yellow buttons which aren't controled by the yellow key on the remote control: Change the color to a to a color not used in other contexts of the widget.	2	The time delay when pressing a button is too large. There should either not be more than 0.7 seconds delay, or an indication that the system is loading	2
	Add breadcrumb navigation to the whole widget (like the navigation in clubs)	1	The red arrows point to the wrong direction (the left direction is correct)	1
	Add an option menu to manage the privacy of the account. E.g.: Online State on/off Program tracking on/off Posting Activities on/off Receiving Messages on/off	1-3	At most points, it is not clear, that the cursor is within the textbox (the user does not know whether he may type or not). The widget should provide a clear feedback, if a text box is activated (cursor animation, text box highlighting, glow effect,...).	3
	Provide a clear difference between a selected and a non-selected button. Changing the color is not sufficient; although you can divide the two color, it doesn't get clear which is the selected one (e.g. if someone shall confirm a deletion or cancellation process via button the color red could either mean to confirm or cancel the action). Therefore another highlighting should be present as well (arrow, accentuation, glow effect...)	2	The Menu Bar isn't consistent in all screens (functionality and display of the button descriptions	1
	Add an update function, that you can see the user interaction with the tablet instantly (chat + message + clubs).		There is an error in displaying vowel mutations. Everything after the correspondent letter is not displayed.	1
	Increase the height of the textboxes.	3		

	The content overlays the TV screen. This is not very helpful, when wanting to watch TV!	5		
	Activity status messages should provide further information, when selected => providing a link to the activity (direct link to a TV show, a club etc.)	3		
	Implement a function which prevents spaming (messages + friendship invitations).	3		
	There is no tablet feedback (it should indicate, whether a message or text or whatever has been sent)(see. Tablet also). Provide also a clear, visible feedback for the connection with the tablet on both devices to give the user a better idea on which device the input takes place and where the input is visualized. There should be a message displayed on the tv screen to inform the user to use the tablet as input device as well as the other way round.	2		
	Implement a function to commit friendship invitations.	1		
	There should be one login for the whole session - no matter if you are using tablet or the smart tv widget. At the moment several logins on both devices are necessary within one session.	2		
	The function for adding contacts can be found in the "Stay in Touch" menu although the first best guess of users would normally be the "Contact" menu. Please make a suggestion to provide a clear separation of both menus and their contents.	3		
	Passive buttons, which are not accessible in specific screens or states are visualized with a light grey shade. The difference between the shades is barely visible - please use a stronger grey shade!	3		

	There are several inconsistencies between the style and the symbols of the used icons. Please use the same icons for the same actions - not only within the widget but also within the Android tablet app.	1		

6.2.2 Feature Requests and Bugs regarding "Clubs"

	Feature Requests	Prio	Bugs	Prio
Clubs	It is not possible to create a new club and/or invite others - please provide this functionality.	2	Inconsitence in wording (german version: clubs or klubs). In the activity posts club it's called group.	1
	The Club Articles should be sorted (newest first)	3	You can't add Themes or discussions with the tablet. The user input isn't displayed.	2

6.2.3 Feature Requests and Bugs regarding "Chat"

	Feature Requests	Prio	Bugs	Prio
Chat	Please add a selection (filter) to manage subgroups like friends.	4	Inconsistencies in wording (german version: close or schließen)	1
	There should be information on how to send a post in the chat to provide a clear affordance for users. (add it to the menu bar, eventually also near the text box)	2	Spelling mistake: "Schließen" is correct instead of "Schliessen"	2
			There is no update of the messages, when you type in the chat with the tablet application	1
			The chat window conceals the Close button (furthermore the button isn't necessary due to the MenuBar and the consistence in the whole widget). => Delete this Button	2
			The menubar is undefined	1
			If you type in a string, which is too long, the string overlays itself.	1

6.2.4 Feature Requests and Bugs regarding "Stay in Touch"

	Feature Requests	Prio	Bugs	Prio
Stay in Touch	Please add a feedback for successfully adding a friend. Make it more clear and make the "add button" invisible, when the friend is already added.	1	If you navigate beyond the last message (message inbox feature), you'll get a deadlock.	1
	To add a friend, you have to type in a name directly. There is no indication that you can browse through the list of people on the platform. Re-structure this information to make it clear.	1	After adding a friend the window is a deadlock.	1
	Also, don't list ALL of the people on the platform, but find a way to structure them (e.g. people in your city, people with same interest, friends of friends)	5	After sending a message the window is a deadlock.	1
	Integrate arrows or something like this, to indicate that the list does not end	2	When searching for a friend and you put e.g. an "a" in the text field, a randomised selection of friends is shown. => all of the friends with a name starting with "a" should be shown	2
	Indicate, that someone received a message, but don't open it instantly.	3		
	If you choose a message the reply field isn't necessary because there isn't a possibility to type something in.	3		
	Friends should be sorted in some way (e.g. alphabetically)	1		
	You should be able to send messages to friends that are offline (asynchronous communication)	2		
	Provide a name of the person who sent you a message. By the selection of a name you'll be offered a conversation window.(Message Inbox)	2		

6.2.5 Feature Requests and Bugs regarding "Wall"

	Feature Requests	Prio	Bugs	Prio
Wall	Timestamps for the Activities. (Display them this way: Fr. 14.12., 10:20)	1	Inconsitence in wording (german version: Activity or Aktivitäten => Aktivitäten)	1
	Activities must be structured in some way - the way they are presented is too confusing and does not provide any benefit	4	If you choose a friend you can't go back. The keys for navigation aren't working anymore.	1
	Add the possibilility to see offline friends. Distinguish them clearly from the online friends.	2	At the moment the friends online are simulated. The working functionality has to be added.	1
	Please add a feedback for the connection with the tablet.	1	You can recommend channels when your on your own name in the user list, but the recommendation will be send to the last friend you chose / interacted with --> It would be better to have your own recommendation (as a reminder)	2
	Visualize the channels each friend is watching.	1	There is no update of the activity display	1
	It should be possible to delete or manipulate own actitivies to offer users the full control over their status messages.	4	The recommendations aren't displayed.	1
			Private message: the button is mirrored/doubled	3

6.2.6 Feature Requests regarding the "TV Guide"

	Feature Requests	Prio	Bugs	Prio
TV Guide	The option to change channels is not intuitive because the status information on top has no affordance character and therefore is not explored by users. Furthermore the arrows on the left and right side can be misunderstood as providing this excact feature.	1	If you recommendate a channel to a friend, there's no "back arrow" and you can't go back either.	1
	There should be a list or some kind of indication, that a TV show was recommended by someone else or by yourself - the user should have an overview, what TV show might be of interest	4	The programm recommendation is still a dummy please change that.	1
	The whole navigation within the TV channel and programm selection menu is contra-intuitive. Please check other TV channel navigation systems (especially the one of the used Smart TV) for standard interaction input and behavior. Users with the tv set will get used to the standard navigation of their tv set and it makes no sense to force them in a completely different interaction scheme within the widget.	1		

7. Literature

Barrett, J., & Kirk, S. (2000). Running focus groups with elderly and disabled elderly participants. Applied Ergonomics, 31(6), 621-629.

Carmichael, A. R. (1999) Style guide for the design of interactive television services for elderly viewers. Winchester, UK: independent Television Commission.

Demiris, G., & Hensel, B. (2009). “Smart Homes” for Patients at the End of Life. Journal of Housing for the Elderly, 23(1-2), 106-115.

Demiris, G., Hensel, B., Skubic, M., & Rantz, M. (2008). Senior residents' perceived need of and preferences for``smart home''sensor technologies. International journal of technology assessment in health care, 24(1), 120.

Demiris, G., Skubic, M., Rantz, M., Keller, J., Aud, M., Hensel, B., & He, Z. (2006). Smart home sensors for the elderly: a model for participatory formative evaluation. human-computer interaction, 6, 7.

Deutsche Akkreditierungsstelle GmbH (2012) Leitfaden Usability. Retrieved on 13.03.2012 from http://www.dakks.de/sites/default/files/71-SD-2007_Leitfaden%20Usability%201.3.pdf

Geerts, D.; De Groof, D. (2009) Sociability Heuristics for interactive TV: Supporting the Social Uses of Television. ACM CHI 2009, 595-604.

Herrmanny, K., Budweg, S., Klauser, M. & Kötteritzsch, A. (2012) Gameinsam – A playful application fostering distributed family interaction on TV. In Adjunct Proceedings of 10th European Interactive TV Conference: Bridging People, Places & Platforms EuroITV 2012, July 4-6, 2012, Berlin, 21-22.

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

Neyer, F. J., Felber, J. &Gebhardt, C. (2012). Entwicklung und Validierung einer Kurzskala zur Erfassung von Technikbereitschaft. In Diagnostica, 58(2).

Nielsen, J. (1994) How to Conduct a Heuristic Evaluation. In Nielsen, J., and Mack, R. L. (Eds.), Usability Inspection Methods, John Wiley & Sons, New York, NY.

Nielsen, J.: Ten Usability Heuristics. Retrieved on 13.03.2012 from http://zonecours.hec.ca/documents/H2010-1-2357287.portionOK.pdf

Samsung Electronics: Samsung TV Application SDK UX Guideline. Retrieved on 08.03.2012 from http://freethetv2011.s3.amazonaws.com/Samsung%20TV%20Application%20SDK%20UX%20Guideline%5BV1.01%5D.pdf

Zaad, L., & Allouch, S. B. (2008). The influence of control on the acceptance of Ambient Intelligence by elderly people: an explorative study. In Ambient Intelligence (pp. 58-74). Springer Berlin Heidelberg.