Columbia International Publishing
Journal of Healthcare Technology and Management
(2014) Vol. 2 No. 1 pp. 1-20
doi:10.7726/jhtm.2014.1001
Research Article

Usability Problems Experienced by Elderly Users in Home Healthcare Systems

Anders Bruun[1]

Received 9 October 2013; Published online 29 March 2014

Abstract

Several research papers have presented crucial work on the needed functionality of home healthcare systems as well as on methods for evaluating the usability of such systems. There is also considerable work on usability problems experienced with systems that are targeted at medical staff. Yet, much less effort has been devoted to identification of usability problems in systems targeted at elderly patients. This paper contributes to the existing body of knowledge by presenting two empirically based usability evaluation studies of different home healthcare systems targeted at elderly users. The aim is to understand key usability problems that this type of target group experience while using such systems. Findings show that a group of elderly test subjects experience a high number of problems compared to a younger control group. In particular, the group of elderly subjects primarily experienced a high number of information related problems regarding e.g. technical information and ambiguous menu labeling. Furthermore, it was found that the group of elderly subjects was more sensitive to lack of overview of menu items. In practice, the studies presented in this paper can be applied to inform designers of home healthcare systems to be particularly aware of the type of information given in a user interface and how this is given to an elderly target group.

Keywords: Usability Evaluation; Usability Problem; Home Healthcare; Elderly

1. Introduction

There is growing interest in devices for home telemedicine. At world level, the life expectancy will increase from 2005 to 2050 to 67 years, and in developing countries to 76 years (United Nations, 2012). This has considerable consequences for healthcare budgets. Another key challenge is that the number of people with chronic illness is increasing and, due to frequent checkups at hospitals, these patients face reduced quality of life, as they have limited freedom to perform their daily activities. A number of research activities have studied home healthcare systems and frameworks that aid in reducing the societal and individual costs of chronically ill elderly.

*Corresponding e-mail: bruun@cs.aau.dk
1 Department of Computer Science, Aalborg University, Denmark

The aim of home telemedicine is to reduce health care costs and at the same time increase the quality of life for patients. Home telemedicine systems allow patients to conduct measurements from their own home (e.g. glucose measurements for diabetes patients) and send the results to the hospital. Other systems put even more emphasis on self-management by supporting patients to take care of their own treatment. If home telemedicine systems are successful, they will reduce the workload of medical staff at the hospitals and in the patients' home, and relieve the patients from visits to the hospital or even hospitalization (Kaufman et al., 2003), (Bruun and Stage, 2011).

1.1 Related Work

For home telemedicine systems to be successful, they must be safe and provide the required functionality. Many researchers have inquired into these aspects. The main focus here has been on the functionality that is required from such systems. Examples are technology for ubiquitous biological monitoring using mobile phones, wearable sensory devices, multi modal platforms, framework and architectural descriptions and literature reviews of observed medical effects (Eikerling, et al., 2009; Fensli and Boisen, 2008; Jaana and Paré, 2006; Pascual et al., 2008; Sasaki et al., 2009; Sashima et al., 2008; Souidene et al., 2009). The target user group of these systems is primarily elderly people.

Unfortunately, there are numerous examples of systems that fail despite having the right functionality, because the prospective users cannot use the system for its intended purpose. A problematic or incomprehensible user interface is a typical source of such problems (Rubin and Chisnell 2008). Usability is a measure of the extent to which prospective users are able to apply a system in their activities (Rubin & Chisnell, 2008). A low level of usability means that users cannot work out how to use a system, no matter how elaborate its functionality is (Nielsen, 1993). The potential of home telemedicine systems can only be realized if the systems have a high level of usability, i.e. a high level of usability is a prerequisite for achieving savings on the healthcare costs and a better quality of life for the patients through use of home telemedicine systems. A high level of usability is particularly important when the main user group is elderly people, who may be constrained by motor, perceptual, cognitive and general health limitations (Fisk and Rogers, 2002) and, in addition, may have a low level of computer literacy.

A considerable number of studies deal with health care systems where the target user group is professional medical staff. This includes evaluation of the usability of desktop, mobile and other healthcare systems with the aim of reducing medical errors introduced by technology. Examples are systems designed for supporting handheld prescription writing, decision support, ordering of lab tests, patient records, family history tracking etc. (Ginsburg, 2004; Johnson et al., 2004; Kushniruk et al., 1996; Kushniruk and Patel, 2004; Kushniruk et al., 2005; Linder et al., 2006; Peleg et al., 2009; Peute and Jaspers, 2007).

Few papers have dealt with the usability problems experienced by elderly users. Below we present four empirically based papers emphasizing usability aspects of systems designed for older adults. Bühler (1996) reports from an empirical study based on two usability evaluations of a robotic wheelchair arm. A total of 14 disabled wheelchair users participated in these evaluations. Findings from the study show that the users experienced problems using the patch board interface in general and that chair mobility was decreased due to the extended width posed by the robotic arm.

Additionally a survey was conducted with experienced wheelchair users in order to assess user needs (Bühler, 1996). Hubert (2006) presents a study focusing on uncovering the usability problems experienced with a home healthcare device. In that study 21 older adults participated in a survey showing that users experienced problems interacting with the physical buttons on the device caused by the small size, lack of texture, location etc. These buttons were examined in more detail in two follow-up studies based on the conduction of usability evaluations. The paper concludes with a set of recommendations for designing device buttons (Hubert, 2006). Kaufman et al. (2003) conducted a case study where a home telemedicine system for elderly diabetes patients was evaluated through interviews, cognitive walkthrough and field usability testing. The evaluated system featured video conferencing, transmission of glucose and blood pressure readings, email, online representation of clinical data and access to educational materials. The study focuses on a methodology for conducting usability evaluations but also provides a basic overview of barriers such as individual competencies, system usability issues and contextual variables. Two user examples of these barriers are provided (Kaufman et al., 2003). The study presented in Kurniawan and Zaphiris (2003) focus on the design of a web health information architecture for older users. Researchers applied the card sorting technique together with 49 seniors. Categories for the card sorting were derived from menu items of an existing website and were printed on index cards, which the seniors were asked to sort into logical groups. Findings show that the obtained information architecture differed from that of the existing website. The architecture on the existing website had four branches with two sub-categories and each sub-category contained two items. The architecture derived by the seniors was less structured with varying numbers of items and sub-categories in each branch. Additionally, seniors grouped items together based on their function or service provided, instead of factors such as geographic location (Kurniawan and Zaphiris, 2003).

1.2 Objectives

The above mentioned research represents crucial work on the needed functionality of home telemedicine systems as well as on methods for evaluating the usability of such systems. There is also considerable work on usability problems experienced with systems that are targeted at medical staff. Yet, much less effort has been devoted to identification of usability problems in systems targeted at elderly patients. In this paper we present two studies where we evaluated the usability of two different home healthcare systems targeted at elderly people. The aim was to better understand key usability problems that such users experience using these types of systems. A better understanding of these problems is vitally important for future design of home healthcare systems with a high level of usability.

In the following we describe the applied research method and the healthcare systems. We then present and discuss our findings. Finally we present our conclusions and point out avenues of future work.

2. Method

We designed the research method to include two studies. With the initial study we aimed to broadly uncover the types of usability problems experienced by elderly users. A second follow-up study was set up to emphasize in depth analysis of the most prevalent type of usability problem and its implications. In the follow-up study we also included a control group consisting of younger test subjects. Both studies are based on classical and well established methods for conducting usability evaluations as per Rubin and Chisnell (2008).

The idea behind classical usability evaluations is to let representative end users apply a particular system while solving a set of realistic task scenarios. While solving these tasks, the users are asked to think aloud, which enables observations of usability problems. Such observations could, e.g. be discrepancies between what a user expects a particular element to do in a user interface and what it actually does. While solving these tasks a person from the evaluation team, the test monitor, sits next to the participants making sure that they think aloud. The job of the test monitor is also to introduce test subjects to the evaluation and to hand over tasks one by one. In an adjacent room, typically a control room, data loggers are also present in order to note down observations from each evaluation session. All data is typically also collected through audio/video recordings and analyzed afterwards, cf. Rubin and Chisnell (2008). The main end result of a formative usability evaluation is a list of usability problems where each problem is categorized according to critical, severe and cosmetic (Molich, 2000), depending on the impact it has on the user. A critical problem is typically a show stopper where the user is unable to continue on her/his own, while an issue that delays the user for several seconds will be categorized as severe etc. Thus, in a formative usability evaluation the main result is of a qualitative nature describing why users experienced problems.

In the following subsections we elaborate on the initial study and follow-up studies by describing the systems evaluated, the settings in which they were evaluated and the participants included in the studies. We then describe the procedures for conducting the evaluations and how we collected and analyzed data.

2.1 Systems Evaluated

The Home Healthcare System evaluated in the initial study is intended for home use by elderly people to monitor their health. It included a main device for data collection and transmission with a display, a speaker and four buttons for interaction, see Fig. **1**. As the manufacturer of the HHS wishes to remain anonymous we do not provide a reference to the system evaluated.

With secondary devices such as blood pressure meters, blood sugar meters and scales, users are able to conduct measurements at home and transfer these to the HHS via Bluetooth, an infrared link or a serial cable. At regular intervals, the device also asks the patients various pre-programmed questions regarding their health.

The HHS automatically transfers collected data to a health care center, where a nurse, doctor or other person is monitoring the health for a group of elderly patients. The system is sent to the patients in a package with a manual for installation.

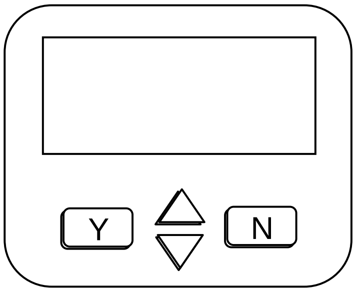

Fig. 1. Sketch of the main Home Healthcare System (HHS) for data collection and transmission.

In the follow-up study we used the SmartSenior healthcare portal as a case (SmartSenior 2013), see **Fig. 2**. The SmartSenior healthcare portal offers various features such as communication with the hospital and other seniors, advice for healthy diets and room climate, measurements of vital parameters etc. The system is aimed for use by elderlies in their homes and all user interaction is done through a web portal displayed on a TV using a remote control. As is the case for the HHS from the initial study, external devices are required to conduct measurements of blood pressure, weight etc.

Fig. 2. Screenshots of the SmartSenior healthcare portal (taken from video recordings), left: Main menu (level 1), right: Sub menu (level 2).

2.2 Settings

In the initial study we conducted the usability evaluation in a dedicated usability laboratory at the university, see Fig. **3**. In Subject room 1, a test subject was sitting at the table operating the system. The test monitor was sitting next to the participant, see Fig. **4**. Two data loggers were in the control room during all tests. Subject room 1 and the control room were separated by one-way mirrors such that test subjects could not see the observers in the control room, but the observers could see the test subjects. Fig. **4** illustrates placement of a test subjects and test monitor during an evaluation session (screenshot from the recorded video material).

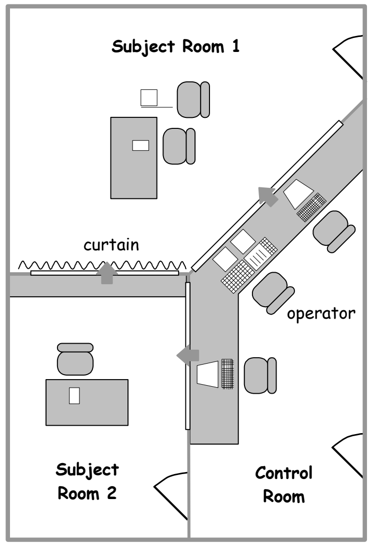

Fig. 3. The setting in the usability laboratory (initial study).

Fig. 4. A test participant and the test monitor. The picture is from the video recording. The small picture in the upper right hand corner shows the interaction with the HHS.

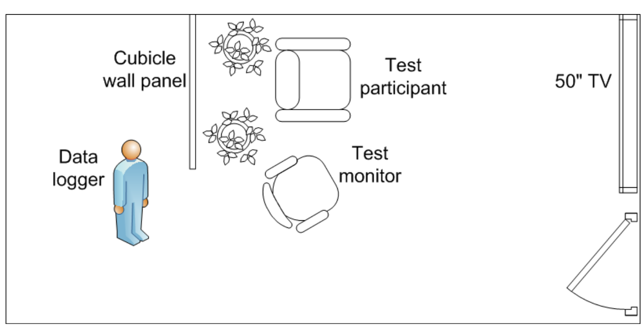

Fig. 5. Overview of settings (follow-up study).

In the follow-up study we conducted the usability evaluations in a room at the telemedicine center Charité, see Fig. **5**. Test subjects were sitting in a chair operating the healthcare portal using a

remote control. A test monitor was sitting next to the participant and a data logger was standing in the room behind the test participant and test monitor to observe all evaluation sessions and take notes.
From the above it can be seen that there are minor differences between the settings where the initial study was conducted in a dedicated usability laboratory and the follow-up study was conducted in a modified office room. However, previous studies have shown that differences in settings (remote versus local) do not impact results of usability evaluations (Adreassen et al., 2007).

2.3 Participants

The HHS device in the initial study was evaluated with 5 test subjects (3 males) and the SmartSenior system in the follow-up study was evaluated with 6 subjects (all female). All participated in a between-subjects design, i.e. all subjects participated in one study only. As both systems primarily are intended for use by elderly people, we selected test subjects with ranging from 61 to 78 years of age. Their mean age was 67.2 years (SD=3.9). None of them had previous experience with these or any similar systems. Their experience in using electronic equipment in general varied from being novices to being experienced. Educational levels varied from the lowest of 2 years to the highest of 4 years and educations also ranged from craftsmanship to academic levels.

In the follow-up study we also chose to include a control group of 6 younger test subjects with a mean age of 35.3 years (SD=5.6), all female. None of them had previous experience in using the evaluated systems, but had extensive experience in using electronic equipment in general. Educational levels varied from 3-5 years, all within academia.

The number of test subjects in each study may seem small, but it should be noted that in the area of Human Computer Interaction it is customary to conduct formative evaluations using 5 test participants as this, from a cost/benefit point of view, is the most feasible. This number is based on the classical studies conducted by Nielsen and Landauer (1993) showing that by using 5 test subjects, evaluators are able to identify 85 % of the total number of usability problems. Additionally, previous studies have shown that the number of test subjects in usability evaluations have limited effect on results compared to the types of task scenarios given, i.e. a high number of subjects does not affect the number of identified problems significantly (Lindgaard and Chattratichart, 2007; Bruun and Stage, 2012).

Three usability evaluators were involved in each of the studies. Two of them were experienced in conducting usability evaluations. In the evaluations, one of them served as test monitor and two served as data loggers.

2.4 Procedure

We applied the same procedure for the evaluations in the initial and follow-up studies. Before the evaluations started, the test subjects were asked to fill in a questionnaire with demographic information. The test monitor then introduced the system and evaluation procedure. This included an introduction to the think-aloud protocol. The tasks were given to the test subjects one at a time.

The test monitor's job was primarily to ensure that the test participants were thinking aloud and give them advice if they got completely stuck in a task. Subjects in the initial study were given the five tasks to solve using the HHS device, see Table 1.

Table 1 Overview of task descriptions for HHS device (initial study).

Task #	Task description
1	Connect and install the HHS and secondary devices.
2	Transfer the data from the blood sugar meter to the HHS. The blood sugar meter is connected using a cable.
3	Measure the weight and transfer the data from the scale to the HHS.
4	A new wireless blood sugar meter is used. Transfer the data from this to the HHS.
5	Clean the equipment.

In the follow-up study subjects were asked to solve the six tasks in Table 2 using the SmartSenior healthcare portal.

Table 2 Overview of task descriptions for the SmartSenior system (follow-up study).

Task #	Task description
1	Log into the healthcare portal.
2	Find the menu where you can change font size.
3	Find the menu from where you can get your heating fixed.
4	Find the log of your vital data.
5	Find the menu where you can create an audio/video connection to the telemedicine center.
6	Find the menu where you can log out of the system.

Tasks were developed based on their relevance for required and everyday usage. Required usage for the HHS device is for instance that the users themselves are asked to connect and install the device, this is a requirement posed by the vendor. Also, in the case of the SmartSenior system, the users are required to log in before usage. All other tasks were developed to reflect everyday use, e.g. data transfer of blood sugar readings in the HHS device and finding logs of vital data in the SmartSenior portal. Formulating tasks for usability evaluations needs careful considerations, not just for the above mentioned reasons, but also because these should not be leading. In other words, if tasks are formulated as users manuals, e.g. "choose this button, then select that menu", it would help the test participants rather than providing insights into how well they can interact with the system on their own. We applied the guidelines for task development provided in (Rubin & Chisnell, 2008).

2.6 Data Collection

In both studies we recorded all evaluation sessions using video cameras and a microphone. The videos showed the system screen in a picture-in-picture setup including the user's face, see Fig. **2** and Fig. **4**. We recorded a total of 9 hours of video. The evaluators who did not act as test monitor acted as loggers and made written log files during the tests.

2.7 Data Analysis

The data analysis was carried out by applying video based analysis as per Rubin and Chisnell (2008). The three evaluators analyzed the video material individually and made a list of identified usability problems. The severity of each problem was also categorized as either "critical", "serious" or "cosmetic" (Molich, 2000). The three individual lists of usability problems were discussed in the team and merged into one list of usability problems. When there was disagreement or doubt whether problems should be combined or split, or how they should be categorized, the video material was reviewed and discussed until agreement was reached. One evaluator will typically find one set of problems, while a second evaluator will find a set different of problems, in this case the overlap of identified problems, i.e. the intersection between the two sets should be determined. If there is a relatively high overlap, this means that evaluators have found many of the same problems, which is an indication of validity. Formally, this overlap is determined by the any-two agreement in usability evaluations, which expresses the extent to which all pairs of evaluators have overlap on observed problems (average agreement). The any-two agreement in the initial study was 40.2 %, which is well above the minimum of 6 % and close to the 42 % maximum found in other studies (Hertzum and Jacobsen, 2003). In the follow-up study the any-two agreement was 43.7 %.

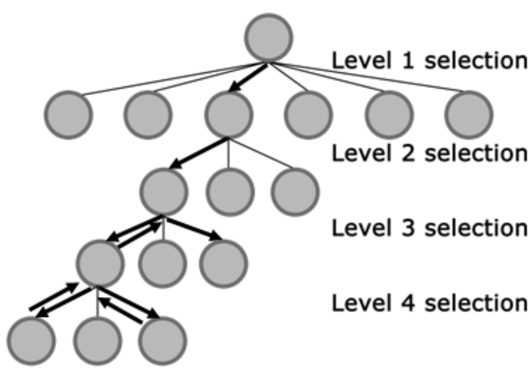

Fig. 6. Example of a path taken by one of the participants at different menu levels in the information architecture.

Based on the findings in the initial study (see results section below), we emphasized in-depth analysis of how test subjects interacted with the information architecture in the follow-up study. Thus, the written logs made by the data logger during the evaluation sessions in the follow-up study represented how subjects interacted with the menus in the form of the paths taken. Fig. **6** provides an example of a path in the information architecture where a participant made one selection at levels 1 and 2 after which she made multiple selections in the third and fourth menu levels. These notes were analyzed quantitatively by the data logger in terms of the number of menu selections made at different levels in the information architecture.

3. Results

This section describes the usability problems we identified in the evaluation of the HHS device in the initial study and the SmartSenior web portal in the follow-up study.

3.1 Types of Problems Experienced

Table 3 shows the number of critical, serious and cosmetic problems identified in the initial and follow-up studies. In the initial study of the HHS device we identified a total 51 problems of which 14 were critical, 15 serious and 20 cosmetic. In the follow-up we found that the elderly test subjects experienced a total of 42 problems where 6 were critical, 20 serious and 26 cosmetic. We also found that the control group of younger test subjects experienced considerably fewer problems. Here we identified 16 problems of which 2 were critical, 3 serious and 11 cosmetic.

Table 3 No. of identified usability problems in the HHS (initial study) and SmartSenior healthcare portal (follow-up study), n = number of participants in the evaluations.

	HHS Device	**SmartSenior Portal**	
	Elderly (n=5)	Elderly (n=6)	Young (n=6)
Critical	14	6	2
Serious	17	10	3
Cosmetic	20	26	11
Total	**51**	**42**	**16**

To get a better understanding of the different types of usability problems, we have categorized them in terms of 12 different usability themes. Below, we briefly explain the meaning of each theme based on Nielsen et al. (2006):

- **Affordance** relates to issues on the users perception versus the actual properties of an object or interface.
- **Cognitive load** regards the cognitive efforts necessary to use the system.
- **Consistency** concerns the consistency in labels, icons, layout, wording, commands etc. on the different screens.
- **Ergonomics** covers issues related to the physical properties of interaction.
- **Feedback** regards the manner in which the interface relays information back to the user on an action that has been done and notifications about system events.
- **Information** covers the understandability and amount of information presented by the interface at a given moment.
- **Interaction styles** concerns the design strategy and determines the structure of interactive resources in the interface.
- **Mapping** is about the way in which controls and displays correlate to natural mappings and should ideally mimic physical analogies and cultural standards.

- **Navigation** regards the way in which the users navigate from screen to screen in the interface.
- **Task flow** relates to the order of steps in which tasks ought to be conducted.
- **User's mental model** covers problems where the interactive model, developed by the user during system use, does not correlate with the actual model applied in the interface.
- **Visibility** regards the ease with which users are able to perceive the available interactive resources at a given time.

Table 4 Total number of identified problems distributed according to usability themes and severity, n = number of test subjects.

	HHS Device				**SmartSenior portal**							
	Elderly (n=5)				Elderly (n=6)				Young (n=6)			
	Critical	Serious	Cosmetic	**Total**	Critical	Serious	Cosmetic	**Total**	Critical	Serious	Cosmetic	**Total**
Affordance		2	2	**4**	1	2	1	**4**	1	1	1	**3**
Cognitive load												
Consistency		1		**1**			1	**1**				
Ergonomics			2	**2**			4	**4**			2	**2**
Feedback	1	3	3	**7**		3	4	**7**			1	**1**
Information	5	8	4	**17**	2	3	8	**13**	1	1	3	**5**
Interaction style					1			**1**				
Mapping	1			**1**			1	**1**				
Navigation	1			**1**			6	**6**			3	**3**
Task flow	1			**1**								
User's mental model	2	3	5	**10**	1			**1**				
Visibility	3		4	**7**	1	2	1	**4**		1	1	**2**
Total	**14**	**17**	**20**	**51**	**6**	**10**	**26**	**42**	**2**	**3**	**11**	**16**

Table **4** shows the total number of identified usability problems distributed on usability themes and severity ratings. This shows that test subjects across the two studies experienced problems from all categories with the exception of cognitive load.

In both systems we found that two of the most prevalent categories of problems are related to the information given and feedback. In case of the HHS the elderly test subjects experienced 17 information related problems and 7 feedback related, which sums up to 24 of the 51 problems (47 %). In case of the elderly test subjects in the SmartSenior evaluation we found 13 problems related to information and 7 regarding feedback, which amounts to 20 out of 42 problems (48 %). In case of the younger control group we found no clear differences between problem types. One of the problems related to the information theme concerned the user manual, which illustrates two possible ways of connecting the HHS to the phone line. In the manual layout, the illustrations were placed on opposite sides of an A5 brochure, which some of the elderly test subjects interpreted as a sequence of steps. This resulted in some subjects trying to connect the device like in the first picture and afterwards connecting the HHS as described by the second illustration. Also, the user interface of the device applied several technology oriented terms, which the elderly test subjects did not understand. An example of this is "Initializing connection" and "Detecting Phoneline", which they did not understand. In the SmartSenior portal we did not find problems related to technical terminology, but instead we found different forms of information related problems. In task 3 for instance (see Table 2) the test subjects were asked to find out how to get the heating fixed. In this case most of the elderly initially located the level 1 menu labeled "At Home" correctly, but they experienced problems in the underlying menus at level 2. One of the usability problems is that the elderly subjects expected to be able to contact the janitor through the menu labeled "Information and Advice", but this did not provide the correct information after which they returned to the "At Home" menu. The elderly test subjects would then traverse trough several other menus, which they believed could provide the information. Thus, the information related problems in the SmartSenior portal were not caused by technical terminology, but rather ambiguous naming of menus.

As an example of a feedback related problem, we found that when participants had answered all of the pre-programmed questions in the HHS device, the display showed the idle screen with the company logo and did not provide feedback to the users of whether they were finished or not. This resulted in some of the users looking for a way to finish and others thought they needed to answer more questions.

One the differences between the two types of systems is the number of problems related to the user's mental model, which is the second most prevalent problem in case of the HHS device (10 problems) but in case of the SmartSenior portal, we only found a single problem in this respect. A problem related to the user's mental model was identified during connection of the Bluetooth scale where participants were looking for a cable to connect this to the HCS, thereby exhibiting that they did not know how to connect these two devices.

Another difference between the systems is the relatively high number of navigation issues found in the SmartSenior portal where we identified 6 problems. In comparison, we identified a single navigation related issue in the evaluation of the HHS device. As an example of a navigational issue, we found that several test subjects in the SmartSenior portal could not figure out how to get back to the main menu after having entered sublevel menus.

3.2 Connection and Installation Issues

As can be seen in Table 1 the test subjects in the initial study were asked to connect and install the HCS device (task 1). We found that 32 of the 51 problems (63 %) were related to this task. To further support this finding.

Table **5** provides an overview of the time it took participants to complete each of the five tasks. A grey cell indicates that a particular participant was unable to solve this task on his/her own and therefore received significant help from the test monitor; this is referred to as task completion rate. All users spent considerable more time completing task 1 compared to any of the other tasks, and 4 out of the 5 users could not complete this task without help. Thus, the amount of problems identified and the task completion times indicate that connection and installation of the HHS device was particularly tedious and problematic in the initial study. Note that users of the SmartSenior portal were not asked to connect and install the devices required, as this is done by the vendor of the products, hence we did not evaluate this in the follow-up study.

Table 5 Task completion time and completion rate (HHS device).

	Task					
User	**1**	**2**	**3**	**4**	**5**	**Total**
1	33:25	10:10	08:47	07:37	01:15	1:01:14
2	33:44	09:34	04:30	04:54	01:00	53:42
3	28:25	02:26	02:45	05:08	01:24	40:08
4	18:43	02:43	04:24	04:07	01:19	31:16
5	26:05	01:06	03:31	04:45	00:41	36:08
Average	28:09	05:12	05:35	05:18	01:08	45:22

Table 6 shows the distribution of usability problems on themes, but only in relation to connection and installation of the HHS device. Almost all critical and serious problems were found during this task where 11 usability problems out of a total of 14 (79 %) are critical, 12 out of 15 (80 %) serious and 9 out of 20 (45 %) cosmetic.

Table 6 Identified problems related to task 1: connecting and installing the device.

HHS Device	Critical	Serious	Cosmetic	**Total**
Affordance		1	1	**2**
Cognitive load				
Consistency				
Ergonomics			1	**1**
Feedback	1	1	1	**3**
Information	3	8	4	**15**
Interaction style				
Mapping	1			**1**
Navigation	1			**1**
Task flow				
User's mental model	2	2	2	**6**
Visibility	3			**3**
Total	**11**	**12**	**9**	**32**

For this task, most problems also relate to information and user's mental model. The 15 problems with information contain all serious and cosmetic instances and 3 of the total of 5 critical problems were observed in task 1. When considering the 6 problems related to the user's mental model, we found that all critical (2) and almost all serious problems (2 of 3) were encountered in task 1.
The rest of the 11 problems identified in task 1 were distributed on affordance (2), ergonomics (1), feedback (3), mapping (1), navigation (1) and visibility (3). It is worth noting that all critical problems relating to these themes were observed in task 1.

As mentioned earlier, one example of an information related problem in task 1 was that most of the elderly test subjects did not understand the text "Detecting phone line" displayed during setup. This resulted in several participants lifting the nearby phone receiver and pressing various buttons on the HHS. Another example regarding information was the term "Line", which was represented on the back of the device and in the manual, but it made no sense to the participants.

One of the problems regarding the user's mental model was when the users, in order to install the HHS, had to connect a cable from the phone line in the wall to the correct port on the HHS device in order to communicate with the remote server. However, the participants did not know which cable to use and some mistakenly tried to connect the phone and the HHS. Another problem concerning the user's mental model was identified when the system asked the user to input a phone line prefix to bypass local in-building phone numbers. In our case, the users needed to press "0" as prefix in order to make the HCS able to communicate with a server outside the building. This prefix had to be selected on the HHS, but some participants pressed "0" on the phone with no result.

3.3 Menu Selections

In the above it is found that information related problems were the most prevalent type of usability problem in both studies, which indicates a need to further study the test subjects' interaction within an information architecture. The information architecture is a crucial element to consider as it determines the way in which information is categorized, labeled and presented, and, thus, whether users are able to locate relevant information efficiently (Gullikson et al. 1999). The SmartSenior healthcare portal provides richer information architecture than the HHS device, i.e. the menu structure consists of more branches and leaves at a greater depth. This gave us the opportunity to further explore information related problems and the strategies applied by the test subjects in order to locate relevant information.
Table **7** provides an overview of the mean number of menu selections made by the elderly and younger test subjects at the four menu levels in the information architecture of the SmartSenior portal. Note that task 1 (see Table 2) is not included as the log in screen was part of the front page and, hence, no menus were selected to complete this task.

Looking at the overall means, we found that the group of elderly subjects consistently made more menu selections than the younger control group and that this pertains to all four menu levels. At level 1 the elderly on average make 2.43 (SD=3.45) menu selections while the younger group made 1.73 (SD=0.98) selections. At level 2 this number is 2.99 (SD=2.99) for the group of elderly and 1.33 (SD=1.18) for the younger group. The mean number of menu selections at level 3 is 2.39 (SD=3.37) in case of the elderly and 1 (SD=0) for the younger group while the overall mean for the elderly at level 4 is 1.55 (SD=1.75) and 1 (SD=0) for the younger. A Welch's two sample t-test (unequal variances) reveals no significant differences between the number of menu selections made by the elderly and younger at level 1 (t=1.18, df=32.35, p=0.25). There are, however, significant differences between these groups at levels 2 (t=2.92, df=36.34, p=0.006), 3 (t=3.68, df=18, p=0.0017) and 4 (t=3.25, df=9, p=0.001). Thus, the group of elderly made significantly more menu selections at these levels than the younger control group. Fig. **7** presents a visualization of Table **7** where the line thickness corresponds to the mean number of selections, e.g. the level 1 line for the seniors is 243 pixels wide corresponding to the mean of 2.43 etc.

Table 7 Mean number of menu selections made by the participants groups at the four menu levels in the information architecture.

	Level 1		**Level 2**		**Level 3**		**Level 4**	
	Elderly	*Younger*	*Elderly*	*Younger*	*Elderly*	*Younger*	*Elderly*	*Younger*
Task 2	1.67	1	2.33	1	0.67			
Task 3	2.67	2	4.83	1	4.33	1	1.5	1
Task 4	1.33	1	5.33	1.33	5	1	2.33	
Task 5	1.5	2	1.5	2.33	2.17	1	0.83	
Task 6	5	2.67	1	1				
Overall mean	**2.43**	**1.73**	**2.99**	**1.33**	**2.39**	**1**	**1.55**	**1**

Table **7** also shows that the groups of elderly subjects made more selections at level 2 than at level 1 and that the team members make more selections at level 1 than any of the underlying menus. The grey cells indicate menu levels which were not reached by the test subjects in the respective tasks. Thus, in tasks 2, 4 and 5 the elderly subjects selected deeper rooted menu levels than the younger control group and they made between 0.67 and 2.33 selections in these cases. Thus, the strategy applied by the group of elderly tends to be to select a menu at level 1 after which they made most selections in the underlying menus, i.e. they are applying a “digging wide” strategy for relevant information in lower menu levels. The younger subjects used another strategy where they tend to make most selections at level 1 and tend to move up one level and try a different menu if they do not find the correct information, i.e. they are applying a “digging narrow” strategy in the underlying menus compared to the seniors. This is also illustrated in Fig. **7** where the lines for the group of elderly subjects are wider than those of the team members.

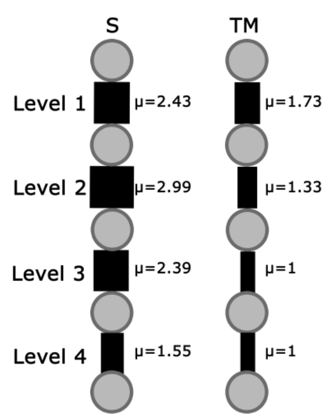

Fig. 7. Visualization of mean number of menu selections for the two participant groups at different menu levels. Line thickness corresponds to the mean number of selections at particular menu levels.

Table 8 and
Table **9** show an overview of the total number of correct and wrong menu selections at levels 1 and 2 made by the two groups of test subjects. We have chosen to focus on these two levels as this is where the elderly made most selections.

Table 8 Total number of correct (+) and incorrect (-) level 1 menu selections made by the two groups of participants.

	Level 1 selections											
	Task 2		**Task 3**		**Task 4**		**Task 5**		**Task 6**		**Total**	
	+	-	+	-	+	-	+	-	+	-	**+**	**-**
Elderly	6	1	10	2	10	0	8	0	6	11	**40**	**14**
Younger	6	0	6	2	6	0	7	1	6	6	**31**	**9**

Table 9 Total number of correct (+) and incorrect (-) level 2 menu selections made by the two groups of participants.

	Level 2 selections											
	Task 2		**Task 3**		**Task 4**		**Task 5**		**Task 6**		Total	
	+	-	+	-	+	-	+	-	+	-	+	-
Elderly	6	1	4	9	6	12	7	1	6	0	29	23
Younger	6	0	6	0	6	1	7	1	6	0	31	2

At level 1 (the main menu) the elderly made 40 correct and 14 incorrect selections. This gives a total of 54 selections with a 26 % error rate, i.e. 14 of 54 selections were incorrect. In comparison the younger group made 31 correct and 9 incorrect selections with a 23 % error rate. Looking at level 2 the elderly made 29 correct and 23 incorrect selections which reveals an error rate of 44 %. In this case the younger control group made 31 correct and 2 incorrect and an error rate of 6 %. Thus, considering level 1 selections the elderly had an error rate similar to that of the younger subjects while the elderly made considerably more errors at level 2 compared to the younger group.

4. Discussion

In this section findings are discussed with respect to related work.

4.1 Differences in Problem Types between Elderly, Younger and Medical Professionals

In this section we discuss our findings in relation to other studies with usability evaluation of health care systems. We have found relatively few papers describing usability evaluations of systems aimed at elderly users. One of these studies is similar to what we have done. This is Kaufman et al. (2003) who asked a group of elderly test subjects to solve the following tasks using a home healthcare unit: Measure glucose level, make blood pressure readings, access an educational website, send an email and change the calendar. These tasks differ partly from our initial study of the HHS device, but there is an overlap in the blood glucose reading task and the purpose of these two systems is similar. Although most tasks differed, we identified similarities in experienced problems. Kaufman et al (2003) identified problems related to unnecessarily complex tasks, which can be compared to the problems we found during connection and installation of the HHS, which most of our elderly test subjects could not complete on their own. Kaufman et al. (2003) also revealed problems concerning non-transparent screen transitions, which are comparable to our problems with missing feedback. Furthermore, we experienced system crashes and restarts that frustrated several test subjects, and they noted issues regarding system instability. Information-related problems were experienced both in their study, where the users did not understand the blood pressure expression "212/89" referring to the systolic and diastolic values, and in our study the users also experienced such problems, e.g. they did not understand the terms "initializing" and "detecting". Mapping problems were also found in both studies. Their users could not establish a correspondence between a set of numbers presented one way on the blood pressure meter and another way in the PC application. In our case the users could not establish a connection between illustrations in the manual and the actual layout of the physical system. Finally, both studies identified issues related to the user's mental model and visibility.

There is a higher number of studies focusing on usability evaluation of systems designed for use by medical staff, e.g. (Kushniruk et al., 1996; Kushniruk and Patel, 2004; Kushniruk et al., 2005). In these studies, the usability problems they have identified are distributed over a set of usability themes with various levels of abstraction ranging from information content, procedure (task flow), comprehension of graphics and text (user's mental model) and overall system understandability to data entry and printing. The results from these studies show that information-related problems were observed in all of the cases. However, the percentage of information-related problems experienced by medical professionals in these studies is considerably lower than what is experienced by the elderly subjects in our studies. In Kushniruk et al. (2005) 16% of the identified problems are information related, while the number is 7% in Kushniruk and Patel (2004) and Kushniruk et al. (1996) (the two latter are based on the same experiment). Our findings show that 33% of the identified problems in the HHS device (initial-study) were information related and 31 % of all problems in the SmartSenior portal (follow-up study with elderly subjects) were information related. The group of younger subjects in our follow-up study also experienced information related problems, and in this case the percentage is also 31 %, which is higher than that experienced by medical professionals in other studies as mentioned above. This is similar to the percentages experienced by the elderly subjects in our studies, yet, the group of younger test subjects experienced less than half of the problems experienced by the elderly subjects.

For problems regarding the user's mental model Kushniruk et al. (1996) observed 5% in this category. In our initial study of the HHS device 20 % of the problems experienced by the elderly subjects were related to this type of problem. This may be explained by a lower level of computer literacy, which is typical for elderly users (Lober et al., 2006). However, in case of the SmartSenior portal, we found that only 2 % of the problems experienced by the elderly subjects were related to the users' mental model and none of the younger participants experienced this. This considerable difference between the HHS device and the SmartSenior portal can be explained by the fact, that the subjects interacted with the SmartSenior portal through a TV set using a remote control, which both are well established technologies compared to the HHS device. Furthermore, most of the problems regarding user's mental were experienced during connection and installation of the HHS device, a task which was not evaluated with the SmartSenior portal.

In sum, when we compare the types of usability problems experienced by medical professionals in related work to the problems experienced by elderly subjects in our studies, we find that the elderly experience a relatively high number of information related problems. Furthermore, the group of younger subjects in our studies experienced considerably fewer problems than the elderly.

4.2 Elderlies are Sensitive to Lack of Overview

Due to the high number of information related problems experienced by elderly test subjects we wanted to further explore how such subjects interact with a healthcare system emphasizing the information architecture. Considering related work, the study in Kurniawan and Zaphiris (2003) focus on the design of a web health information architecture for older users. In that study the card sorting technique was applied where initial card categories were derived from menu items of an existing website. Seniors were asked to sort these cards into logical groups and findings show that the information architecture suggested by the seniors differed from that of the existing website. Our findings support this in the sense that the initial information architecture did not support the group of elderly subjects to the same extent as it supported the younger control group. We found that the elderly made significantly more menu selections at levels 2, 3 and 4 compared to the younger group and that the elderly had a considerably higher error rate at level 2. An interesting finding is that the elderly had a considerably lower error rate when interacting with the level 1 menu, i.e. they made fewer errors in this case. An apparent explanation of this difference could have been a variation in the number of menu items, e.g. if the level 1 menu had fewer items than level 2 menus. This, however, is not the case in the SmartSenior portal as the level 1 menu had six menu items while level 2 menus had in the range of five to eight items. A more likely explanation is the design of the menus which differed considerably. As shown in Fig. **2** the level 1 menu was designed using text and icons with all items visible on the screen at the same time. Deeper rooted menus were text only of which three to four items were displayed at once. These menu designs provided subjects with an overview of all items at level 1 but not at deeper levels. Thus, since the elderly made significantly more selections than the younger group at levels 2 and below as well as making considerably more errors at level 2, our findings suggest that elderly tend to be more sensitive to lack of overview in the information architecture.

Taken together, the lack of overview of items within deeper rooted menu levels and the above mentioned usability problems caused the group of elderly test subjects to apply an information retrieval strategy which can be described as “digging wide” while the younger control group were “digging narrow”.

5. Conclusions

This paper has presented two empirically based usability evaluation studies of different home healthcare systems targeted at elderly users. The aim was to understand key usability problems that this type of target group experience while using such systems. Findings show that the group of elderly test subjects experience a high number of problems compared to a younger control group. In particular, the group of elderly subjects primarily experienced a high number of information related problems regarding e.g. technical information and ambiguous menu labeling. This finding applied to both of the evaluated systems. Furthermore, it was found that the group of elderly subjects was more sensitive to lack of overview of menu items, i.e. this group made more menu selections and more selection errors than the younger control group before finding what they were looking for. Finally, the information related usability problems caused the group of elderly test subjects to apply an information retrieval strategy which can be described as “digging wide” while the younger control group were “digging narrow”. In practice, the studies presented in this paper can be applied to inform designers of home healthcare systems to be particularly aware of the type of information given in a user interface and how this is given to an elderly target group.

In relation to future work, a recommendation would be to conduct further usability studies with other types of home healthcare systems in order to increase generalizability of findings. This may additionally reveal other types of usability problems as in our study, where we found problems related to information and the users mental model to be the dominant types. Also, due to the low number of usability studies including elderly test subjects, it would be interesting to conduct studies with a higher number of subjects.

Acknowledgements

A special thanks goes to Professor Dr. Jan Stage, Dr. Claudia Hahn, Benjamin Voigt and Dr. Martin Schultz for their participation in these studies.

References

Andreasen, M. S., Nielsen, H. V., Schrøder, S. O. and Stage, J., 2007. What happened to remote usability testing? An empirical study of three methods. *Proc. Conf. on Human Factors in Computing Systems (CHI)*. ACM, New York.

Bruun, A. and Stage, J. 2011. Evaluating the Usability of Home Healthcare Applications. In Carsten Röcker and Martina Ziefle (Eds.), *Human Centered Design of E-Health Technologies*, IGI Global, Hershey, PA, USA.

Bruun, A and Stage, J. 2012. The effect of task assignments and instruction types on remote asynchronous usability testing. *Proc. Conf. on Human Factors in Computing Systems (CHI)*. ACM, New York.

Bühler, C., 1996. Approach to the analysis of user requirements in assistive technology. *Int. J. of Industrial Ergonomics 17*(2), 187-192.

Eikerling, H-J., et al., 2009. Ambient Healthcare Systems - Using the Hydra Embedded Middleware for Implementing an Ambient Disease Management System. *Proc. Int. Conf. on Health Informatics (HEALTHINF)*. INSTICC Press.

Fensli, R., Boisen, E., 2008. Human Factors Affecting the Patient's Acceptance of Wireless Biomedical Sensors. *Proc. Biomedical Engineering Systems and Technologies*. INSTICC.

Fisk AD, Rogers WA, 2002. Health Care of Older Adults: The Promise of Human Factors Research. In Wendy Rogers and Arthur Fisk (Eds.) *Human factors interventions for the healthcare of older adults*. Lawrence Erlbaum.

Ginsburg, G., 2004. Human Factors Engineering: A Tool for Medical Device Evaluation on Hospital Procurement Decision-Making. *J. of Biomedical Informatics 38*(3), 213-219. Elsevier.

Gullikson, S., Blades, R., Bragdon, M., McKibbon, S., Sparling, M. and Toms, E.G., 1999. The impact of information architecture on academic web site usability. *The Electronic Library 17*(5), 293-303. Emerald.

Hertzum, M. and Jacobsen, N. E., 2003. The evaluator effect: A chilling fact about usability evaluation methods. *Int. J. of Human-Computer Interaction 15*(1), 183-204.

Hubert, R., 2006. Usability Field Study of Home Health Monitoring Devices Used by Older Adults. *Proc. Student/Faculty Research Day*. Pace University.

Jaana M., Paré, G., 2006. Home Telemonitoring of Patients with Diabetes: A Systematic Assesment of Observed Effects. *J. of Evaluation in Clinical Practice 13*(2), 242-253. Blackwell Publishing.

Johnson, C. M., et al., 2004. A User-Centered Framework for Redesigning Health Care Interfaces. *J. of Biomedical Informatics 38*(1), 75-87. Elsevier.

Kaufman, D.R., Patel, V.L., Hilliman, C., Morin, P.C., Pevzner, J., Weinstock, R.S., Goland, R., Shea, S. and Starren, J., 2003. Usability in the real world: assessing medical information technologies in patients' homes. *J. of Biomedical Informatics 36*(1-2), 45-60. Elsevier.

Kurniawan, S.H. and Zaphiris, P., 2003. Web Health Information Architecture for Older Users. *IT & SOCIETY 1*(3), 42-63.

Kushniruk, A. W., et al., 1996. Cognitive Evaluation of the User Interface and Vocabulary of an Outpatient Information System. *Proc. AMIA Annual Fall Symposium*.

Kushniruk, A.W., et al., 2005. Technology Induced Error and Usability: The Relationship Between Usability Problems and Prescription Errors When Using a Handheld Application. *Int. J. of Medical Informatics 74*(7-8), 519-526. Elsevier.

Kushniruk, A.W., Patel, V. L., 2004. Cognitive and Usability Engineering Methods for the Evaluation of Clinical Information Systems. *Int. J. of Medical Informatics 37*(1), 56-76. Elsevier.

Linder, J. A., et al., 2006. Decision support for acute problems: The role of the standardized patient in usability testing. *J. of Biomedical Informatics 39*(6), 648-655. Elsevier.

Lindgaard, G. and Chattratichart, J., 2007. Usability Testing: What Have We Overlooked? *Proc. Conf. on Human Factors in Computing Systems (CHI)*. ACM, New York.

Lober, W., B. Zierler, A. Herbaugh, S. Shinstrom, A. Stolyar, E. Kim, and Y. Kim. 2006. Barriers to the Use of a Personal Health Record by an Elderly Population. *Proc. AMIA Annual Symposium Proceedings*.

Molich, R., 2000. *User-Friendly Web Design*. Ingeniøren Books.

Nielsen, J., 1993. *Usability Engineering*. Morgan Kaufmann. San Diego.

Nielsen, J. and Landauer, TK., 1993. A mathematical model of the finding of usability problems. *Proc. Conf. on Human Factors in Computing Systems (CHI)*. ACM, New York.

Nielsen, C.M., Overgaard, M., Pedersen, M.B., Stage, J. and Stenild, S., 2006. It's Worth the Hassle! The Added Value of Evaluating the Usability of Mobile Systems in the Field. Nordic Conference on Human-Computer Interaction (NordiCHI). ACM, New York.

Pascual, J., et al., 2008. Intelligent System for Assisting Elderly People at Home. Proc. Int. Conf. on Health Informatics (HEALTHINF). INSTICC press.

Peleg, M., et al., 2009. Using Multi-Perspective Methodologies to Study Users' Interactions with the Prototype Front End of a Guideline-Based Decision Support System for Diabetic Foot Care. *Int. J. of Medical Informatics 78*(7), 482-493. Elsevier.

Peute, L.W., Jaspers, M.W., 2007. The significance of a usability evaluation of an emerging laboratory order entry system. *Int. J. of Medical Informatics 76*(2-3), 157-168. Elsevier.

Rubin, J., Chisnell, D., 2008. *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests, 2nd. Edition*. John Wiley & Sons, Inc.

Sasaki, J., et al., 2009. Experiments of Life Monitoring Systems for Elderly People Living in Rural Areas. *Proc. Int. Conf. on Health Informatics (HEALTHINF)*. INSTICC Press.

Sashima, A., et al., 2008. Toward Mobile Healthcare Services by Using Everyday Mobile Phones. *Proc. Int. Conf. on Health Informatics (HEALTHINF)*. INSTICC Press.

SmartSenior, 2013. http://www1.smart-senior.de/enEN/ (downloaded October 4th 2013)

Souidene W., et al., 2009. Multi-Modal Platform for In-Home Healthcare Monitoring (EMUTEM). *Proc. Int. Conf. on Health Informatics (HEALTHINF)*. INSTICC Press.

United Nations, 2012. World Population Prospects. Dept of Economic and Social Affairs. Obtained at: http://esa.un.org/wpp/Documentation/pdf/WPP2012_Press_Release.pdf (downloaded October 4th, 2013).