Usability evaluation of a health care application based on IPTV

Vitor Simões Ribeiro\textsuperscript{a}\textsuperscript{*}, Ana Isabel Martins\textsuperscript{b}, Alexandra Queirós\textsuperscript{b, c}, Anabela G. Silva\textsuperscript{b, d}, Nelson Pacheco Rocha\textsuperscript{b, e}

\textsuperscript{a}PT Inovação, Rua Eng. José Ferreira Pinto Basto, 3810-106 Aveiro, Portugal
\textsuperscript{b}Institute of Electronics and Telematics Engineering of Aveiro (IEETA), Campus Universitário, 3810 Aveiro, Portugal
\textsuperscript{c}Health Sciences School, University of Aveiro (ESSUA), Campus Universitário, 3810 Aveiro, Portugal
\textsuperscript{d}Center for Health Technology and Services Research (CINTESIS), Faculty of Medicine, University of Porto, Porto, Portugal
\textsuperscript{e}Health Sciences Department, University of Aveiro, Campus Universitário, 3810 Aveiro, Portugal

Abstract

This paper presents the usability evaluation of the AAL@MEO, an application running in a commercial service of Internet Protocol TV that integrates technological solutions to support home care. The evaluation methodology comprises the application of usability assessment instruments (the Post-Study System Usability Questionnaire and the International Classification of Functioning based Usability Scale), performance tests and analysis of critical incidents. The results show that elderly users had a high degree of satisfaction towards the interaction with the TV set, but some of the interaction mechanisms based on a remote control with multiple functions need further developments.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Peer-review under responsibility of SciKA - Association for Promotion and Dissemination of Scientific Knowledge

Keywords: Usability Evaluation; Internet Protocol TV; Monitoring Devices; Home Care.

1. Introduction

Home care emerges as a potentially cost-effective solution to the challenges that health and social care systems need to face due to actual demographic trends. Home care allows the allocation of alternative resources to
hospitalization or institutionalization\(^{(1)}\) and can be defined as home-based health care and social care services provided by formal and informal caregivers, not solely focused on clinical purposes, but also on a range of essential activities for the maintenance of the individuals’ quality of life and that are part of the normal everyday life of every citizen, with the use of technology when appropriate\(^{(1-3)}\).

This paper introduces some results of the AAL4ALL project, as an approach to the provision of assistance and health care services (ie, AAL@MEO) using the commercial Internet Protocol TV (IPTV) deployment of the Portugal Telecom MEO service. First, we present the related work. Then, we introduce the methodological approach and we present the results of a prototype testing with real users. Finally, we present some concluding remarks.

2. Related Work

The recent developments related to smart components highly integrated and miniaturized\(^{(4)}\) together with the existing ubiquitous communication networks make transfer and access of health-related information such as measurement data or medical knowledge possible anywhere and at any time. Therefore, the advances on sensing and communication technologies make it possible to develop mobile and wearable sensors able to continuously monitor physiological parameters, activities and behaviours in out-hospital conditions. Monitoring physiological parameters, together with monitoring daily activities (ie, identifying consistency and completeness in these activities), to assess, in a naturalistic and continuous way, health and cognitive status\(^{(5,6)}\) might help to automate assistance and prevent accidents or disease exacerbations. Concerning emergency situations, monitoring patients and providing alerts to health care providers might facilitate a prompt intervention. In this respect, pervasive health applications might improve access to care, particularly when time is vital (eg, in stroke or acute trauma).

These systems are based, in most cases, on solutions designed to solve specific problems related to specific care process and specific patient situations\(^{(7-12)}\), which means that users need to acquire skills and abilities to interact with a broad range of systems deployed at home. In this respect the use of a TV set at home as the central interaction and communication system is advantageous\(^{(13-14)}\).

3. Methodology

The development methodology of the AAL@MEO, an interactive TV application for remote care services, was based on a series of procedures that go from the analysis of system requirements to the final evaluation of systems or services (prototype and pilot tests). The AAL@MEO application was the result of an in-depth requirements analysis and a conceptual validation involving several stakeholders\(^{(15)}\). The study reported in this paper is related to the usability evaluation of the AAL@MEO prototype.

The AAL@MEO prototype was evaluated in terms of usability\(^{(16)}\) by an observational study. This study took place between July 22\(^{nd}\) and December 3\(^{rd}\), 2014. Each session consisted of four parts: i) Introduction – The investigator applied a social and demographic questionnaire and then delivered the session script, explaining orally all information contained therein; ii) Test – The participant performed the tasks described in the session script and the investigator collected data on the user interaction with the AAL@MEO; iii) Usability Evaluation Instruments – The investigator assisted the participant when filling in the usability questionnaires and scales; and iv) Summary – The investigator thanked the participant.

All adults able to fill in the instruments were eligible to participate if they gave written informed consent. The only exclusion criterion was the presence of limitation on the fine upper limb movements. All data collection was anonymized and the involved participants received all the information regarding the study and their participation before completing the informed consent. Additionally, the written informed consent was obtained prior to data collection.

A request of approval was directed to the administration of the *Santa Casa da Misericórdia de Aveiro* (SCMA), a church-sponsored charity institution.
3.1. Instruments

The instruments used to evaluate ALL@MEO were the Post-Study System Usability Questionnaire (PSSUQ) together with a general usability question, the ICF based Usability Scale I (ICF-US I) and the ICF based Usability Scale II (ICF-US II).

The PSSUQ is a usability evaluation questionnaire based on scenarios that was developed by IBM. It consists of 19 items aimed at addressing five usability characteristics of a system: rapid completion of the task, ease of learning, high quality documentation and online information, functional adequacy and rapid acquisition of productivity. The results and interpretation of the PSSUQ were performed within the defined parameters of the original scale. Lewis developed the rules to calculate the PSSUQ score and defined a global total score and partial scores for the subscales associated with PSSUQ: i) Global: average of the responses from items 1 to 19; ii) System Utility (SysUse): average of the responses from items 1 to 8; iii) Quality of information (InfoQual): average of the responses from items 9 to 15; and iv) Interface Quality (IntQual): average of the responses from items 16-18.

The final score of PSSUQ is calculated by the average of the scores of the items and the scores of the subscales. Higher scores are associated with better usability. If a participant fails to respond an item or classifies it as not applicable, then that item should be replaced by the average score of the remaining items.

Another evaluation tool that reflects the users’ opinion is the general usability question: “Overall how would you rate the application on a scale from 0 to 10?”. This question was based on the After-Scenario Questionnaire (ASQ) developed by Lewis and is accompanied by a numeric rating scale from 0 to 10, in order to facilitate the completion of the question, and 0 represents the worst possible score and 10 the best possible score.

In addition to these evaluation measures, two scales that reflect the opinion of the investigators were also used: the ICF-US I and ICF-US II. The ICF-US I allows a comprehensive usability evaluation, while the ICF-US II allows the evaluation of specific components of the application. The two scales allow the identification of barriers and facilitators (small, middle and big barrier or facilitator) and, consequently, the application weaknesses and strengths.

The ICF-US I scores all items from -3 to 3, which -3 is the less positive and 3 is the most positive response. If a participant fails to respond an item or classify it as not applicable, then it should be replaced by the average of the remaining items, rounded to the units. The final score of the ICF-US I is the sum of all items’ score.

In addition, the ICF-US II scale is composed of items that identify different application components (e.g., menu navigation, image or interaction with the remote control). This scale has three parts: i) the application components; ii) detailed usability; and iii) overall evaluation. Each item is classified as a barrier or a facilitator. Whenever a component is classified as a barrier, the user must identify the characteristic that is causing the component classification as a barrier (e.g., the characters can be considered as barrier due to their size, color or sharpness). Therefore, it is possible to identify which items require more work in order to improve the application.

Statistical analyses were performed using Microsoft Office Excel and SPSS - Statistical Package for Social Sciences (SPSS Inc, Chicago).

3.2. Performance

During the test, an investigator recorded data on the execution of each task in a performance evaluation grid. Performance evaluation is a technique centered in the users and the tasks they perform, and involves the collection of quantitative data. For each task included in the script, the following data was recorded: the success or failure in carrying out the task, the execution time (in seconds) and the total number of errors.

Statistical analysis of the performance evaluation grid data were performed using Microsoft Office Excel.

3.3. Critical Incidents Record

The same investigator that recorded the performance also logged the critical incidents in order to systematically identify behaviors that contribute to the participants’ success or failure in specific situations. Several details were considered, such as easy/difficult interaction with the application or calmness/restlessness.
4. Used Technology

In Portugal, there are elderly people who feel uncomfortable when using PCs/laptops or Smartphones. The major reasons are the complexity and the small size of characters usually used in those terminals, namely in the small screens like the ones of the Smartphones. On the other hand, a TV set equipment is well known by the elderly population, so that the adaptation to a new application/service is usually more natural. For that reason, the AAL@MEO was developed over the commercial IPTV deployment of the Portugal Telecom MEO service. This allows LiveTV and VoD channels as well as a sort of interactive commercial services related with different IPTV contents, supported over a Mediaroom TV Platform. Furthermore, the development of new applications is supported by a special purpose Application Program Interface known as Presentation Framework (PF).

The Fig. 1 presents a screen capture of the AAL@MEO application, which was developed to:
- Present information, namely vital signs measurements (e.g., blood pressure), weight measurements, or information videos, starting with a general dashboard with a summary of all the information available to the end user.
- Receive pop-up in the TV screen with alerts or notifications.
- Allow interactions with the end-users for confirmation of the measurements or answering questionnaires related to their physical and psychological state.

The use of characters with big size and figures in the several TV screens of the PF application oriented to the elderly population, associated to an easy navigation menu with only five buttons (Up, Down, Left, Right and OK) of the remote control, were the major requirements for this developed TV interface in order to have a friendly service.

![Fig. 1. Screen capture of the application.](image)

5. Results

The sample was selected according to previously defined inclusion and exclusion criteria (consecutive recruitment). The sample consisted of 30 female collaborators of the SCMA that provide care to elderly people. The average age was 58 years old (SD = 3.3). In terms of education level most participants (n = 12) had the 4th grade, followed by the 7th grade (n = 9), 5th grade (n = 5) and finally the 12th grade (n = 2) and graduation (n = 2).

Among the participants, 28 reported to use a mobile phone regularly, 26 of which use a simple mobile phone and 2 use a smartphone. Regarding the computer utilization, only 11 participants have ever used a computer, but very sporadically and with help from a third person. Only 5 participants had already used the commercial MEO service before.

Most participants are very satisfied with their ability to perform day-to-day activities (n = 21). With regards to the quality of life, 16 participants considered it reasonable and 11 good. Most participants (n = 18) considered them as having good memory.
According to the experimental design, the following subsections discuss the results related to: i) PSSUQ and the ICF-US I and II scales; ii) the performance evaluation; and iii) the critical incident analysis.

5.1. Instruments

The average score of PSSUQ for all participants was 5.82 out of a maximum of 7 (SD = 1.01), which indicates a high degree of usability and satisfaction. The results for the subscales associated with PSSUQ were:

- System Utility (SysUse): average of the responses from items 1 to 8 = 5.52 (SD = 0.38).
- Quality of information (InfoQual): average of the responses from items 9 to 15 = 5.89 (SD = 0.55).
- Interface Quality (IntQual): average of the responses from items 16-18 = 6.34 (SD = 0.09).

In a more detailed analysis, the items that obtained higher scores were "The interface of this system was pleasant" (6,43 out of 7) and "This system has all the functions and capabilities that I expected it to have" (6,31 out of 7). Both items are related to the interface quality subscale. This is in agreement with the participants' expressed satisfaction about the visual part of the application.

The items that obtained lower scores were "Whenever I made a mistake using the system, I could recover easily and quickly" (4,86 out of 7) and "I was able to complete the tasks and scenarios using this system" (4,90 out of 7). Both items are related to the interaction of participant-application. Another evaluation tool that reflects the participants' opinion is the general usability items whose average rating was 7,77 out of 10 (SD = 1.74), reinforcing participants' satisfaction regarding the application.

According to the ICF-US I, the application was considered a facilitator by 19 participants and a barrier by 11 participants. The mean score for all participants was 3,67 (SD = 14,20), out of a range from -30 to 30, indicating that in general the application was a small facilitator. The participant with the highest score had 24, and the participant with the lower score had -22.

Reviewing the items individually it is important to classify each one as a barrier or facilitator (see Fig. 2). The barriers, in order of severity, were: "The ease of use" (-0.70); "The similarity of the way it works on different tasks" (-0.43); "The ease of learning" (-0.40); "The achievement of expected results" (-0.27); and "Globally, I consider that the application was" (-0.10). Analyzing the items identified as barriers, most participants had difficulty in tasks related to interaction, learning, execution and repetition of actions.

In turn, the facilitators in order of importance, were: "The application responds to your actions" (1.77); "The understanding of the messages displayed" (1.63); "The degree of satisfaction with the utilization of the application" (1.40); and "The knowledge of what was happening in the application during its use" (0.77).

By analyzing the items identified as facilitators, it seems that most participants clearly understood the information provided by AAL@MEO. The users were aware of what was happening in the application and were satisfied with it. The item 6: "The ability to interact in various ways" was considered not applicable since there was only one form of interaction (via a remote control).

The ICF-US II was developed to evaluate the Application components (6 items), Detailed Usability (7 items), Overall Evaluation (1 item). Globally, the application of ICF-US II shows that the application was a facilitator for 20 participants, neutral for 1 and a barrier for 9 participants. The mean of the total scores for all participants was 8.90, in a range of -42 to 42 (SD = 18.31), indicating that in general the application was a small facilitator. The participant with the highest score had a score of 40, and the participant with the lower score had a score of -23.
The results of the subscales within the ICF-US II were analyzed according to the mean of the items (range -3 to 3):

- **Application Components**: average of the responses = 0.37 (SD = 2.40);
- **Detailed Usability**: average of the responses = 1.02 (SD = 2.41);
- **Overall Evaluation**: average of the responses = -0.47 (SD = 2.01).

The barriers, in order of severity, were: "The evolution graphs in the health menu" (-1.87); "Navigation on the horizontal menu" (-0.90); "Navigation with the remote control" (-0.67); "Measurements confirmation" (-0.60); "Navigation on the vertical menu" (-0.53); "The application functions" (-0.50); and "The session progress" (-0.47).

Analyzing the items identified as a barrier, it seems that most participants had difficulty in tasks related to the navigation in the vertical and horizontal menus, interaction with the remote control and the interpretation of the application behavior.

In the other hand, the facilitators were: "The contrast between the background and the information" (3.0); "The font" (2.97); "The menu colors" (2.97); "The information on the monitor menu" (1.97); "The information on the alerts menu" (1.97); "Authentication with the ID card" (1.50); "Icons/symbols/graphics" (0.13). Analyzing the items identified as facilitators, it seems that most participants found it easy to understand the information provided by the application interface and authenticating of the system.

5.2. Performance

For each task included in the session script, the data recorded was: the success or failure in carrying out the task, the execution time (in seconds) and the total number of errors. The data from the performance evaluation grid is shown in Table 1.

The most problematic tasks, with a higher failure and error rate were the tasks 8, 9 and 3. No participant was able to successfully complete Task 8 since it was not clear that swiping to the side it would display the graph of the monthly values and because it was not clear that to return to the previous menu the user should press the OK button in the remote control as Back was selected as default. Task 9, because in the vertical menu (on the left) the items are represented by icons and the participants were not familiar with any of those icons. For that reason they had great difficulty finding out for themselves how to logout. Task 3, because the participants tend to the press OK button to confirm, but in this case, they had to move to the confirm option first and then press OK button. Generally, these
were also the tasks that had higher execution times, except for the task 2 and 4, in which the execution time depends on the devices rather than in the difficulty performing the task.

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Success (nº of persons)</th>
<th>Average Time (seconds)</th>
<th>Number of errors (average)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - Authenticate in the system using your ID card and the card reader.</td>
<td>22/8</td>
<td>12,1</td>
<td>0.6</td>
</tr>
<tr>
<td>2 - After correct authentication, weigh yourself using the scale.</td>
<td>30/0</td>
<td>75,7</td>
<td>0.1</td>
</tr>
<tr>
<td>3 - Using the remote control, confirm the weight.</td>
<td>8/22</td>
<td>21,9</td>
<td>0.8</td>
</tr>
<tr>
<td>4 - Using the blood pressure monitor check your blood pressure and pulse rate.</td>
<td>30/0</td>
<td>101,7</td>
<td>0.0</td>
</tr>
<tr>
<td>5 - Using the remote control, confirm the blood pressure and pulse measurements.</td>
<td>12/18</td>
<td>24,6</td>
<td>0.8</td>
</tr>
<tr>
<td>6 - Enter the menu &quot;monitor&quot; and check for alerts (in red). If you have any unread alerts, access to this option.</td>
<td>24/6</td>
<td>14,9</td>
<td>0.2</td>
</tr>
<tr>
<td>7 - In the &quot;health&quot; menu check which was the last measurement executed.</td>
<td>17/13</td>
<td>24,0</td>
<td>0.7</td>
</tr>
<tr>
<td>8 - Access the blood pressure measurements and check the evolution of the monthly values and return to the previous menu.</td>
<td>0/30</td>
<td>90,6</td>
<td>2.5</td>
</tr>
<tr>
<td>9 - Logout of your session.</td>
<td>5/25</td>
<td>28,1</td>
<td>1.3</td>
</tr>
</tbody>
</table>

5.3. Critical Incident Analysis

In terms of critical incidents, about half of participants showed insecurity and lack of autonomy during the interaction as they constantly seek approval from the investigator in the interaction with the application. Some of the participants (7) had difficulty interacting with the remote control, especially the arrows buttons and must participants reported difficulties browsing the menus (17 participants had difficulties with the horizontal navigation while 13 participants had difficulties with the vertical navigation). These results are consistent with the results of the application of the ICF-US II scale.

6. Conclusion

The usability evaluation shows that elderly users have a high degree of satisfaction towards the interaction with the TV set. Despite the results being mostly positive, AAL@MEO presents some usability issues that occurred repeatedly, clearly indicating what features needed to be improved. Early on it was clear the lower computer literacy of the participants which was reflected in the execution of many of the tasks.

The application was developed to present information, namely on vital signs (eg, blood pressure), weight, or information videos, starting with a general dashboard with a summary of all information available to a specific end user. According to the proposed solution, the entire control interface is integrated in a TV set. Furthermore, considering the elderly population, characters with big size and figures were used in association with a navigation mechanism with five buttons (Up, Down, Left, Right and OK) of the remote control. However, it is relevant to mention the difficult of using the multiple functions button of the remote control due to the option of using the same type of remote control included in the available MEO service. As the arrows buttons are part of a "single" button it was difficult for participants to press the correct option, which had consequences in the navigation using the menus. This implies that the solution being used should be further analyzed. Results suggest that even the commercially available MEO service would benefit with a different type of remote control.

The services offered may be dynamically extended through the inclusion of third-party or new services. Having a common access interface based on IPTV allows elderly people to feel safer and have a perception of familiarity and confidence when interacting with new services.
Acknowledgements

This study was funded by Project QREN Nº 13852 I&D «AAL4ALL – Ambient Assisted Living for All», co-funded by COMPETE and FEDER.

References