Usability Evaluation Plan for Advanced Technology Services for Prevention and Management of Chronic Conditions for the Elderly

Angelina Kouroubali, Ilia Adami, Michalis Foukarakis,
Margherita Antona, and Constantine Stephanidis

Institute of Computer Science, Foundation for Research & Technology-Hellas
100 N. Plastira St, Heraklion, 70013, Crete, Greece
{kouroub,iadami,foukas,antona,cs}@ics.forth.gr

Abstract. This paper analyzes the methodology and preliminary results of the usability evaluation of large scale technology services for the prevention and management of chronic conditions of elderly people. REMOTE, a three years European project in the Ambient Assisted Living domain, aims to define and establish a multidisciplinary and integrated approach to research and development of advanced technology services for addressing real needs of elderly people. The target population included citizens at risk due to geographic and social isolation in combination with specific chronic conditions and the coexistence of lifestyle risk factors, such as obesity, blood pressure, poor eating/drinking habits, stress and others. Technology services provided by the project include wearables and sensors for detecting body temperature, heart rate, human posture, as well as sensors and actuators to be installed in premises for providing context information, such as air temperature, human location and motion. The article presents the usability evaluation plan that was developed for the purposes of the project, as well as some preliminary results from the usability evaluation.

Keywords: Usability evaluation, ambient assisted living, sensors & wearables.

1 Introduction

The elderly population is rising throughout Europe and it has been estimated that between 2010 and 2030 the number of people aged from 65 to 80 will rise by nearly 40%, posing enormous challenges to Europe's society and economy [1]. Policies which address the ageing of the population and the work force focus on enabling older workers to remain active and productive for a longer proportion of their life span [2]. Advanced information and communication technologies (ICT) have the capacity to assist in this issue [3]. For example, using telemonitoring for disease management of heart conditions reduces mortality rates by an estimated 20%. It has also been demonstrated that telemedicine influences the attitudes and behaviour of patients, resulting in better clinical outcomes. In addition, ICTs can assist elderly people, especially individuals with chronic conditions and/or at risk of exclusion, in

B. Godara and K.S. Nikita (Eds.): MobiHealth 2012, LNICST 61, pp. 445–454, 2013.

increasing their self-management capacity [4]. ICTs are particularly useful in establishing a feeling of security and command for aged people in the European countryside [5], strengthening people confidence in leading an independent life at home and delaying, or even fully avoiding, institutionalisation.

The REMOTE project, recognising that isolation of the elderly- both geographical and social- is often a common multiplying risk factor, developed and tested ICT applications and services for health care and management [6]. A user-centred design approach was adopted from planning to execution of the project, with continuous involvement of users to ensure that all relevant facets of subjective quality were successfully delivered, ranging from ethical aspects to accessibility, usability, safety, security, and "customer relationship". The REMOTE technological platform includes

Table 1. The REMOTE platform applications

Application	Explanation
Brain skills trainer	Activities and games for memory support and memory assessment.
Calendar	Allows the user to keep track of daily tasks, schedule appointments, and provides reminders. Medication list and schedule is also viewed in this application.
Environmental control	Allows for remote control of the environment such as lighting and heating.
Guardian Angel	Connects with the BioHarness belt sensor to monitor vital signs and detect falls. It provides alerts based on the thresholds for each user as they have been set in the Healthcare advisor application.
Nutritional Advisor	Provides the user with personalized nutritional content such as menus, nutritional tips, shopping assistance and cooking assistance.
Activity Advisor	Provides an activity plan which the user should follow, joint with advice and tips. The professional has access to questionnaires to understand user needs and design the appropriate activity plan.
Emergency Management	Visualise and manage the automatically driven and patient driven health emergency calls and alerts.
Healthcare Advisor	The physician sets the user's medication schedule. Establish thresholds for vital sign monitoring and review alerts based on these thresholds.
Medical Contact Center	Administrative tool to register new patients, assign services, connect them to health professionals and store personal data ensuring privacy and confidentiality.

mobile, desktop and web applications for remote monitoring of vital signs, and activity status, emergency alerts, organization of activities calendar, environmental control sensors and others[1]. The REMOTE platform offers integration of services to encompass all range of users including physicians, patients and carers. Table 1 presents the applications of REMOTE which were evaluated for usability.

In order to ensure the user-centred approach, the REMOTE project has developed and executed an extensive usability evaluation plan to address usability issues throughout application and system lifecycles. This paper describes the design and implementation of the usability evaluation plan and explains the advantages and disadvantages of this plan for large scale usability evaluations of ICTs. Some preliminary analysis of the evaluation results is also presented.

The evaluation plan was applied to all applications and systems in all pilot sites of the REMOTE project. The usability evaluation methodology aimed at examining what elderly patients, nurses, doctors as well as healthcare professionals and informal carers expected from the proposed technology, as well as assessing the developed applications against the four main usability principles, as outlined by Booth [7], usefulness, ease of use, learnability/memorability, and likeability or satisfaction.

2 Usability Evaluation Plan

This section describes in detail the evaluation plan that was formed for the usability testing of the applications that were developed for the REMOTE project. The evaluation plan was based on existing expert and user-based evaluation techniques and used a combination of recognized qualitative and quantitative usability analysis tools to report the findings. This plan was followed by a total of 9 pilot sites in 6 countries, Germany, Israel, Greece, Spain, Romania, Italy, and Norway, and involved the testing of 9 main applications implemented for mobile and desktop use.

Evaluating large scale systems that involve a variety of applications with integrated services that can be accessed through a unified platform is not an easy endeavour. The evaluation plan needs to follow a solid methodology that can be carried out easily by all parties involved, as well as to produce findings that can be presented in a consistent and homogeneous manner. There are many recognized usability methods available to choose from and many usability metric tools respectively. Jakob Nielsen recommends using more than one method of evaluation, if possible, because one method may uncover issues that the other one missed [8]. Based on this premise, an evaluation plan was designed for the REMOTE project which consisted of a combination of expert-based, user-based, and home-based evaluation tools. Also, the evaluation plan included qualitative analysis components such as user personal comments and expert observations, as well as quantitative analysis components such as metric usability questionnaires (System User Satisfaction (SUS) questionnaire and user success rate) [9, 10]. The qualitative and quantitative components of the evaluation plan are explained in more detail further down the paper. The evaluation plan consisted of three phases: 1) Expert-based evaluation

[1] http://www.remote-project.eu/

phase, 2) User-based evaluation phase in a controlled environment, and 3) User-based evaluation phase at home.

Phase 1 Expert Based Evaluation: The main purpose of the first phase of the REMOTE evaluation plan was to identify and correct any major design flaws and problems before they reached production and real user testing. Expert-based evaluation is generally used to identify usability problems based on established human factors principles [11]. The experts conducting this type of evaluation can be human-computer interaction specialists, usability, and accessibility specialists, or even interface designers with experience in user-centric design principles. There are a few inspection techniques available [8], but the two that were used in the evaluation of the REMOTE applications were expert walkthroughs and heuristics analysis [11, 12]. During the expert walkthroughs, two to five evaluators performed a series of application specific user tasks on working or non-working prototypes, just like a real user would, and identified the areas that could potentially cause confusion or errors to the real users. At the same time, the experts were also asked to rate the application against Jacob Nielsen's Heuristics list of usability principles and guidelines [11]. At the end of the expert evaluation, each evaluator produced a report on the observations he or she made during the inspection and completed a Heuristic analysis questionnaire. These reports were then aggregated in a single report that included the results from all the inspections and was given to the development and design team of the evaluated application. Upon completion of the expert evaluation, the developers incorporated the most important comments into the applications/systems and released the working version of the software in order to proceed for testing with real users.

Phase 2 User Based Evaluation: Once the improved working versions of the applications were released from the developers, the actual user-based evaluation phase began. During this phase, real life usage scenarios and tasks were written for each application and actual users were invited to participate in the evaluations. The tasks had to be clear, precise, and relatively short to accomplish and the chosen evaluation participants had to fit the profile of real life users of the applications/systems as close as possible.

At the beginning of each evaluation session, the participant was introduced to the purpose and the process of the evaluation and was given a brief explanation of how the application worked by the evaluator. The participant was given a User Test Consent Form to sign that stated his or her agreement to participate. The fact that the evaluation was testing the performance of the application and not the performance of the user was also emphasized at that point. A series of representative user tasks was then given to the user to complete sequentially. The participant was requested to openly express his or her thoughts, observations, feelings, and comments to the evaluator during each task. This is known as the Think Aloud method [13], which enables the evaluator to capture the thinking process of the user. The evaluators were instructed to provide assistance only when absolutely needed and keep notes on what was happening and what was being said during each task.

The following three usability measurement tools were used during Phase 2 of the evaluation: 1) System Satisfaction Scale Questionnaire (SUS), 2) User Success Rate,

and 3) User Test Analysis. The first two tools were used for the quantitative analysis and the third tool was used for the qualitative analysis of the results.

The System Satisfaction Scale Questionnaire (SUS) [9] is a measurement tool that calculates the total system satisfaction. This questionnaire is a 5 point Likert scale based questionnaire that had to be completed by each test participant right after the test when everything was still fresh in the memory. The User Success Rate is a measurement tool that calculates the overall effectiveness of the system [10]. The evaluator marked in a form if the user completed each task with success (requiring no assistance from the evaluator), with partial success (requiring some assistance), or with failure (giving up on task). The User Test Analysis form was used to gather all the important comments and observations made by each test participant during the test.

It is important to note that the measurement tools are only indicative of the overall usability of the system. There are many factors that affect the scores such as user expertise, the complexity of the user tasks that were given to the participants to complete, the performance of the system and the network where the application was tested, etc. Therefore these scores are meant to be general indicators on how the applications performed.

Phase 3 – Home-Based Evaluation: During the third phase of the evaluation plan, the home-based evaluation, selected users were given a specific REMOTE application/system to use at home for a few days on their own. These users were given a demo of how the application/system worked beforehand. Each user was instructed to complete the home-based evaluation questionnaire at the end of the evaluation period.

3 Usability Evaluation Actions

Each pilot site explored the applications as part of an integrated scenario of daily living and decided upon the user tasks to test for each application. As a result, the applications were tested from different aspects focusing on the most important functions of the application/system. Table 2 provides an example of user tasks for the Guardian Angel application in three different pilot sites. User tasks were used to establish the user success rates. The evaluator introduced each task to the user, and were given to perform these tasks, the evaluator was noting the success, partial success or failure of the user to complete these tasks, in order to calculate the User Success Rate metric.

Participants were also asked to fill out the system user satisfaction questionnaire (SUS) marking each question with 1 to 5, where 1=strongly disagree and 5= strongly agree. The SUS score reflects the overall system usability as perceived by the particular user. Sample questions included *I would use this system frequently*, *The system is unnecessarily complex*, *I needed to learn a lot of things before using it*.

In addition to these tools, data about the users' profile were maintained taking into consideration, gender, age, health status, education and familiarity with technology. The average primary user profile was a person of 55-77 years of age, with chronic

Table 2. User tasks for Guardian Angel

Pilot Sites	User Tasks
GREECE Pilot A	T1: Once you login to the Guardian Angel application, follow the instructions displayed on the screen. Now press the "Start" button that's located on the belt and then the "Enter" button that is on the screen of the mobile phone.
	T2: Wait for the connectivity between the belt and the phone to establish and then read the vital signs measurements taken.
	T3: When you are done reading the various measurements, exit the application from the phone and take off the belt.
GREECE Pilot B	T1: You have been asked to enter the Guardian Angel application from the main menu. Navigate to the menu and check your breathing rate.
	T2: You have been asked to enter the Guardian Angel application from the main menu. Navigate to the menu and check your heart rate.
	T3: You have been asked to enter the Guardian Angel application from the main menu. Navigate to the menu and check your activity level.
	T4: You have been asked to enter the Guardian Angel application from the main menu. Navigate to the menu and check your posture.
SPAIN Pilot A	T1: Log in to REMOTE on mobile, choose health monitoring option and then Guardian Option
	T2: Place belt around chest, turn the system on; perform tasks for 10 minutes (sitting, sitting and reading; walking)
	T3: Turn off the recording on the mobile and turn off Guadian Angel; remove belt from chest
	T4: Using Health Advisor, check the data recorded

diseases and a slight familiarity with technologies such as personal computers and smart phones.

When dealing with the evaluation of large-scale systems with many sub-components and applications, as much emphasis must be given on how to aggregate and report the results of the evaluations, as on selecting the appropriate usability tools and methods to use. One of the main challenges in applying a common methodological evaluation approach to multiple sites is the accuracy and consistency

of reporting the results. To overcome this challenge in REMOTE, special templates were developed to report the results in an aggregate format and were sent to each pilot site along with clear instructions on how to use them. These templates helped towards a uniform reporting of results that allowed a common analysis despite the fact that each pilot site tested different applications. So, at the end of the evaluations, each pilot site was responsible for sending the appropriate filled-out templates and reports with the results, to one central person who then processed them to produce the final results per application accordingly.

The evaluation plan that was used for REMOTE's purposes proved to be efficient and effective in drawing a comprehensive picture of the usability of the entire platform. Its success was not only based on the fact that it incorporated expert-based and user-based evaluation techniques in addition to home-based user trials, but it was also based on the fact that it provided the pilot sites with the necessary reporting tools to aggregate and send-in their results in a homogeneous format saving a great amount of time and effort in their further analysis.

3.1 Analysis and Preliminary Results

The analysis of all user responses across pilot sites per application based on the SUS scores and the User Success Rates scores indicates that the mobile applications received higher SUS scores and user success rates than the desktop ones, as it can be seen in Figure 1 and 2. However, it was concluded that this difference does not mean that one platform was more usable than the other because there are many factors that could have contributed to this difference. For example, the desktop applications offered richer functionality than their mobile counterparts and therefore the tasks used during the evaluation of the desktop applications were more complex in nature than the ones used in the evaluation of the mobile ones. It is very possible that the higher complexity of tasks used in the desktop applications could have caused the lower user success rates and SUS scores. Regardless of this differentiation, it is safe to conclude that both platforms performed above average in the scores, which means that the users perceived them to be easy to use, easy to learn, and useful.

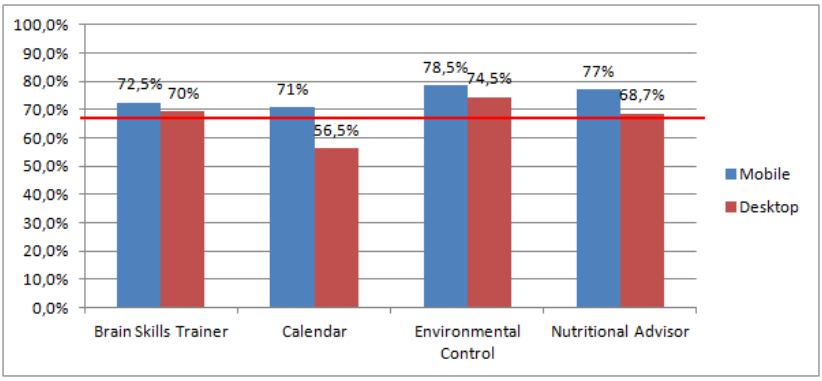

Fig. 1. Total SUS Scores Mobile vs Desktop Applications

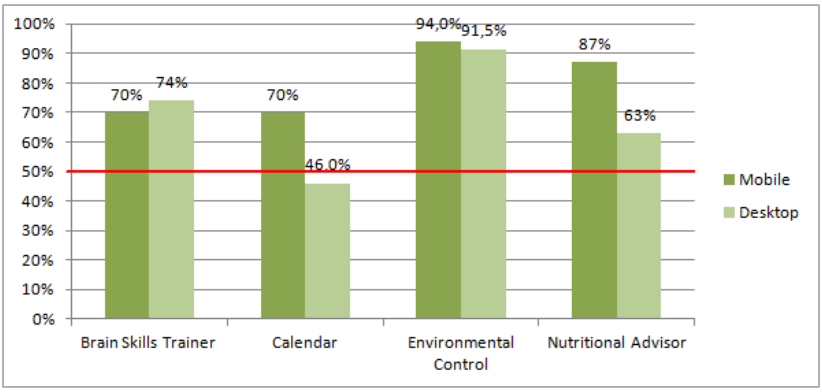

Fig. 2. Total User Success Rates Mobile vs Desktop

Looking at the user comments as they have been captured during the Thinking Aloud process, specific usability issues were identified. Overall, the users were satisfied with the concept of the interoperability of REMOTE's applications. However, difficulties were observed with using some of the application mainly due to the users' lack of experience and familiarity with the technologies involved. The touch screen functionality of the Smart Phones in combination with the small screen of these phones caused challenges in the users with chronic conditions such as Parkinson's, while the use of the keyboard and the mouse in the desktop application was also challenging for a lot of users, who had never used PCs before. Despite these barriers, the users were able to use the applications and complete the majority of the tasks given to them during the evaluations.

The home-based trials were carried out in 4 pilot sites and involved a total of 20 users who were given the Zephyr bio harness and the Nokia Smartphone to use along with the Guardian Angel and the health monitoring applications. The results aggregated from these trials further reinforced the conclusion that the users found the REMOTE platform a useful and usable tool for elderly people that can help in the monitoring of their health on an everyday basis. The platform received high marks for its ability to communicate the users' vital signs measurements directly to their assigned physicians through the mobile phone. The users appreciated the concept and thought that such health-monitoring tools can be very valuable for elderly people to have in their homes.

4 Discussion

Information and communication technologies have the potential to change health care as well as life style for people with chronic conditions [14]. These changes are occurring concurrently and need to be considered for appropriate design of sensor-enhanced health information systems. Basic functionalities of technology services include emergency detection and alarm, disease management, health status feedback and advice [14] as well as social and psychological support. These functionalities are

combined differently based on individual conditions as well as individual needs of chronic patients and their informal carers.

A unified usability methodology over pilot sites distributed across Europe is an essential tool for identifying application/system design and functionality issues that go beyond cultural or user profile differences. However, it is important to offer individual evaluation moderators instructions and training not only for conducting the evaluations, but also for reporting of the results. A unified methodological approach can only express its fullest potential if data capture and data reporting is also unified. This aspect can be a great challenge for large implementation projects involving different partners from different countries.

Despite the differentiation of data reporting across sites, the data that was reported offers a valuable basis for improving the design and implementation of the REMOTE applications. Preliminary analysis of the results has shown that elderly users are open to using mobile and desktop applications that may improve their daily life and support their activities. Participants in general welcomed the technology services offered, and thought it would improve their quality of life, resulting in fewer complications. These findings are in accordance with those of other studies about technology management of chronic conditions [15, 16]. The introduction of new technologies for supporting independent living would need to involve appropriate education and adjustment periods to ensure the motivational and psychological support of users. These practices are fundamental in enhancing use of technology services.

5 Conclusions

This paper has presented the methodology and preliminary results of large scale usability evaluation conducted in pilot sites in different European countries.

Participants followed an extensive usability methodology involving heuristic analysis, user-based evaluations, and home trials. It made use of a variety of qualitative and quantitative tools in order to examine the overall usability of the system. In addition, the methodology paid extra attention in the gathering of the results and provided the pilot sites involved with clear instructions and templates to use for reporting purposes. Preliminary results indicate that this is a suitable evaluation plan for large-scale systems and provide the necessary structure to gather data for applications/systems implemented in different settings. However, the interpretation of the results should be handled with care as there are several factors at play which cannot be measured such as cultural issues, background knowledge, etc.

Further work is needed to investigate whether implementation of changes to application and systems based on user comments will improve the opinion of users and their user success rates and user satisfaction scores.

Acknowledgment. This work was performed in the context of the project “ICT based solutions for Prevention and Management of Chronic Conditions of Elderly People” (REMOTE, AAL-2008-1-147) funded but the Ambient Assisted Living (AAL) Joint Programme.

References

1. Giannakouris, K.: Ageing characterises the demographic perspectives of the European societies. In: Population and Social Conditions, Eurostat Statistics in Focus (2008)
2. Eurostat, Demography report 2012: Older, more numberous and diverse Europeans, European Commission (2011)
3. Camarinha-Matos, L.M., Afsarmanesh, H.: Active Ageing Roadmap – A Collaborative Networks Contribution to Demographic Sustainability. In: Camarinha-Matos, L.M., Boucher, X., Afsarmanesh, H. (eds.) PRO-VE 2010. IFIP AICT, vol. 336, pp. 46–59. Springer, Heidelberg (2010)
4. MacAdam, M.: Frameworks of integrated care for the elderly: a systematic review. Canadian Policy Research Networks= Réseaux Canadiens de Recherche en Politiques Publiques (2008)
5. Heenan, D.: How local interventions can build capacity to address social isolation in dispersed rural communities: a case study from Northern Ireland. Ageing International 36(4), 475–491 (2011)
6. Bekiaris, A., Mourouzis, A., Maglaveras, N.: The REMOTE AAL Project: Remote Health and Social Care for Independent Living of Isolated Elderly with Chronic Conditions. In: Stephanidis, C. (ed.) Universal Access in HCI, Part III, HCII 2011. LNCS, vol. 6767, pp. 131–140. Springer, Heidelberg (2011)
7. Booth, P.: An introduction to human-computer interaction. Psychology Press (1989)
8. Nielsen, J.: Heuristic Evaluation. In: Nielsen, J., Mack, R.L. (eds.) Usability Inspection Methods, pp. 25–62. John Wiley and Sons, Inc., USA (1994)
9. Brooke, J.: SUS-A quick and dirty usability scale. In: Usability Evaluation in Industry, vol. 189, p. 194 (1996)
10. Nielsen, J.: Success rate: the simplest usability metric. Jakob Nielsen's Alertbox (2001)
11. Wharton, C., et al.: The cognitive walkthrough method: A practitioner's guide. In: Usability Inspection Methods, pp. 105–140 (1994)
12. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Empowering People. ACM (1990)
13. Lewis, C., Rieman, J.: Task-centered user interface design. A Practical Introductio (1993)
14. Haux, R., et al.: Health-enabling technologies for pervasive health care: on services and ICT architecture paradigms. Informatics for Health and Social Care 33(2), 77–89 (2008)
15. Bostock, Y., et al.: The acceptability to patients and professionals of remote blood pressure monitoring using mobile phones. Primary Health Care Research & Development 10(4), 299 (2009)
16. Verhoeven, F., et al.: The contribution of teleconsultation and videoconferencing to diabetes care: a systematic literature review. Journal of Medical Internet Research 9(5) (2007)