eWALL for Active Long Living

Project no.:	**610658**
Project full title:	**eWALL for Active Long Living**
Project Acronym:	**eWALL**
Deliverable no.:	**D6.4**
Title of the deliverable:	**Small scale studies report**

Contractual Date of Delivery to the CEC:	**31.10.2015**
Actual Date of Delivery to the CEC:	**31.10.2015**
Organisation name of lead contractor for this deliverable:	**Austrian Institute of Technology**
Author(s):	**Claudia Oppenauer-Meerskraut, Beatrix Zechmann, Cristian-Dan Bara, Harm op den Akker, Lex van Velsen, Clara Schaarup, Stine Veje Hangard, Mihail Mihaylov, Francesco Infarinato, Federica Rizza**
Participants(s):	**P01, P04, P05, P06,**
Work package contributing to the deliverable:	**WP6**
Nature:	**R**
Version:	**1.0**
Total number of pages:	**115**
Start date of project:	**01.11.2013**
Duration:	**36 months – 31.10.2016**

Abstract:

In this deliverable, the methodology followed for the small-scale lab trials and field trials is described and evaluation results from eWALL prototypes are summarized and discussed regarding identified problems and necessary implications for further eWALL prototype development.

Please note that the evaluation results of the M25 prototype will be described and submitted in an updated version of this deliverable.

Keyword list: evaluation, lab trials, prototype, usability, user experience,

Document History

Version	Date	Author (Unit)	Description
0.1	Jan 15, 2015	ATE	Proposal of ToC, Results 1st lab trials ATE
0.2	Jan 26, 2015	AAU (Clara Schaarup, Stine Veje Hangard, Mihail Mihaylov)	Input 3.2, 5.4 and Annex II
0.3	Mar 24, 2015	RRD (Cristian Bara)	Modified table input in chapter 3
0.4	Mar 26, 2015	RRD (Cristian Bara)	Updated 3.2
0.5	Mar 27	RRD (Cristian Bara)	Extended 4.1.1
0.6	June 1, 2015	ATE (Claudia Oppenauer-Meerskraut)	Modified table structure, Input COPD use case evaluations, Figure and Table numeration
0.7	Oct 12, 2015	ATE (Beatrix Zechmann)	5.5 Implications for the next eWALL prototype; Extended table structure for input of Small Scale Evaluations → chapter 6
0.7	Oct 13	RRD (Cristian Bara)	6.2 M19 Prototype Evaluation
0.7	Oct 14	RRD (Cristian Bara)	6.3.6 Participant description and results in NL
0.7	Oct 15	RRD (Cristian Bara)	Finalized 6.3.6
0.8	Oct 21	ATE (Claudia Oppenauer-	6.3 Results M22 Prototype Evaluation Circle 2

		Meerskraut)	6.3.7 Implications for M25 Restructuring the document according to Harm's recommendation
	Oct 28	ATE (Beatrix Zechmann)	Annex III: M22 Prototype Evaluations Protocol Fixed typos in Annex IV and some other parts Input for 5.1 and 7 General formatting (tables, figures etc.)
0.9	Oct 29	RRD (Harm op den Akker, Lex van Velsen)	6.3.7 M22 Prototype Evaluation with Patients with a Lung Disease
0.91	Oct 29	IRCCS (Federica Rizza, Francesco Infarinato)	M22 Prototype Evaluation with experts and MCI subjects
1.0	Oct 30	AAU (Sofoklis Kyriazakos)	Final review and approval

Table of Contents

1   Executive Summary

Preliminary validations of the working prototypes were conducted to find relevant input from the user and expert perspective for the optimization of the service delivery, in terms of technology, design, and clinical protocols. The M12 prototype evaluation and the Small Scale Evaluations (SSE) work plan as defined in Task 2.6 was carried out, leading to valuable input for the integration and service development tasks. In line with the methodology of DeChant the focus of these evaluations were on the usability of the platform, the patients' satisfaction with the services as well as experts' views to the services provided for people suffering from MCI, COPD and older people with age-related impairments.

After each evaluation, all four evaluation sites summarized their results in form of recommendations in a table to communicate them effectively to evaluation and design partners. Thus, any emerging need of changes and problems were addressed in the subsequent prototype optimization. We used weekly teleconference calls under the banner of the newly formed *Cross-Work-Package on User Experience and User Interaction* to discuss adaptations. The attendance of both technical and medical partners in these calls was a perfect basis to discuss the results on an interdisciplinary level. A final evaluation of the integrated eWALL system will be given in the field tests in WP8.

Please note that the evaluation of the M25 prototype, which is the last part of the "Small Scale Evaluations" will start by the mid of November due to on going development. The results will be subsequently described in an updated version of the D6.4.

2   Introduction

In this document we will demonstrate the eWALL prototype evaluations from M12 to M22. The different chapters explain evaluation goals, methodology and results of the different user studies. Based on users' feedback and user experience, the eWALL prototype has been redesigned and functionalities and services adapted in order to guarantee a well-accepted system. Final results of each evaluation process have been discussed together with all responsible partners and new prototype iterations are based on the conclusions made in the project consortium.

Within the prototype evaluation studies primary and secondary users as well as experts were included to investigate different aspects of the eWALL system and increase prospected benefits of eWALL.

The different chapters will report all the relevant results of the prototype studies and will illustrate the conclusions for the next prototype iteration process and finally for the last M25 prototype within the small scale validation process.

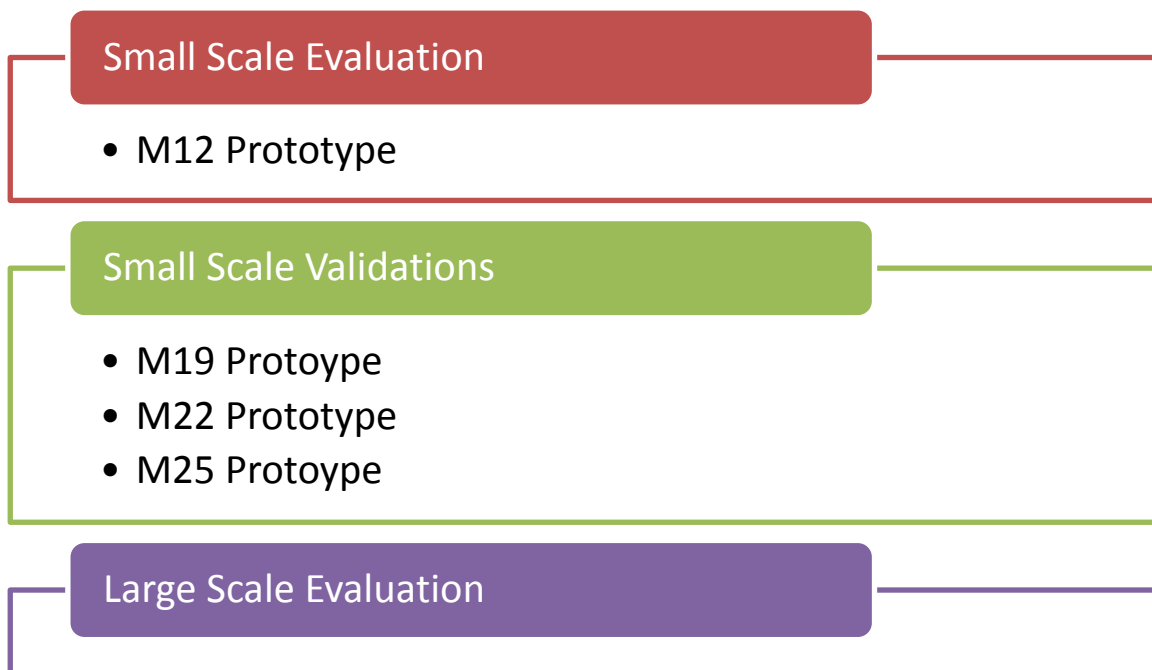Figure 1gives an overview of the evaluation studies presented in this document:

Small Scale Evaluation

- M12 Prototype

Small Scale Validations

- M19 Protoype
- M22 Prototype
- M25 Protoype

Large Scale Evaluation

Figure 1: Evaluation overview

3   Evaluation goals and methodology

During the writing of the eWALL proposal, DeChant's Staged approach to evolution of telemedicine is adapted as the framework for all the evaluation related tasks [1]. This framework suggests to tailor the type of assessment to the development cycle of the technology. Roughly, the framework differentiates between technology evaluation at application (stage 1–2) and global levels (stage 3–4). The framework suggested by DeChant et al., 1996 is generally accepted to be an appropriate framework for the evaluation of telemedicine- or related services. However this framework is published in 1996 and in the past years a lot of experience has been gained in its application for telemedicine evaluation. This has resulted recently in a specification of this framework [2]. This specification is shown in Figure 2.
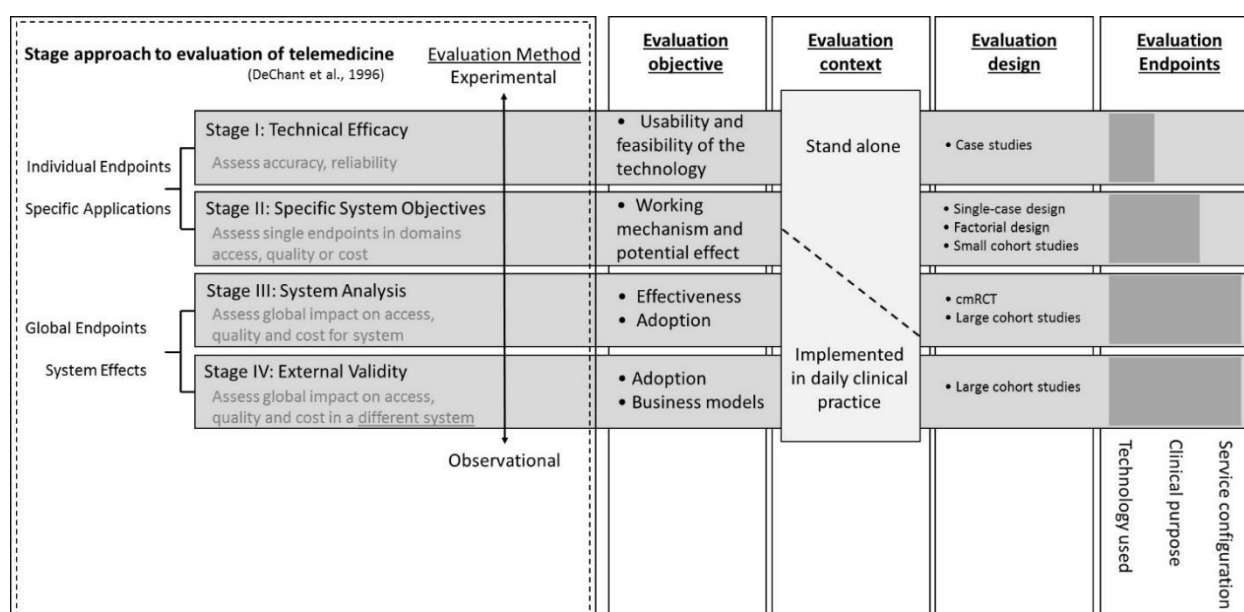
Figure 2: The specification of the stage approach to evaluation of telemedicine

This framework describes how evaluation can best be focused by addressing for each stage the evaluation objectives, evaluation context, evaluation designs and evaluation endpoints. These aspects are considered essential to prepare an adequate evaluation of telemedicine services:

Stage I: The first stage of telemedicine evaluation focuses on the feasibility and usability of the technology used in an experimental design with a small number of subjects or even case studies. This type of evaluation design allows researchers to gain detailed information which can be used for further improvement of the telemedicine service. The telemedicine service is evaluated as a stand-alone service and evaluation endpoints focus on feasibility and usability of the technology used.

Stage II: The technology used in the second stage is stable and evaluation is focused on gaining an initial idea about the potential added value for clinical practice and possible working mechanisms. For this, evaluation can be performed using the telemedicine service as a

stand-alone service. Designs that can be used focus on studying processes in often small group of subjects rather than on examining the effectiveness. Suitable designs are:

1. cohort studies with a small sample size (n<50);
2. single-case design (or N = 1 designs) that are shown to be suitable, useful and feasible to assess the natural history of relationships between treatment processes and outcomes within individuals [3];
3. factorial designs that enables a researcher to simultaneously evaluate two treatments and their interaction. Besides being efficient, factorial designs are an effective way to examine interaction effects [4, 5]. By using this design the potential effect of various modalities of the telemedicine service could be compared in one study. The evaluation endpoints in this stage should focus on the potential added value of the telemedicine service mapped on both the technology used and the clinical purpose that is supported.

Stage III: This stage starts when earlier studies indicate that the telemedicine service has potential and focuses on showing the effectiveness of the telemedicine service and/or adoption of the service by its end-users. In order to identify these aspects, it is important that the telemedicine services are evaluated in the way they will be implemented in daily clinical practice. Although, randomized controlled trials (RCTs) are considered the gold standard for evaluating the safety and effectiveness of medical interventions their characteristics do not fit well with the evaluation of telemedicine services [6]. An alternative for a conventional RCT might be the "cohort multiple randomized controlled trial" (cmRCT) being introduced by Relton et al., [7]. This design tackles some of the problems associated with pragmatic trial designs. The key features of this design are:
1. recruitment of a large observational cohort of patients with condition of interest;
2. regular measurement of outcomes for the whole cohort; and
3. capacity for multiple randomised controlled trials with new releases of telemedicine technology over time.

In addition, the cmRCT design aims to replicate routine healthcare in the real world. For each randomized controlled trial, information from the cohort is used to identify all eligible participants. Some eligible participants are randomly selected and offered the intervention. Data on participants who are not willing to participate provides information on the acceptability of the intervention. The evaluation endpoints at this stage should not only focus on a previously defined value expected for each technology used and the clinical purpose that is supported but also take into account the way the telemedicine service is being implemented in daily clinical practice (service configuration).

Stage IV: The fourth stage evaluation elaborates the adoption as addressed in stage III. To ensure further implementation, involvement of every stakeholder (healthcare professionals, patients, technology providers, insurance companies and policy makers on a local and national level) is important. This means that evaluation here should

focus on the business models and concrete business cases. Without information on the cost and effectiveness of telemedicine services, decision makers run the risk of introducing services that are not cost-effective for society [8]. This evaluation can only be performed in an adequate way when the service is implemented in daily clinical practice as only in this cases the true added value can be evaluated. The studies performed in this stage are large-scale cohort studies (n≥50) [9]. As addressed in stage III the evaluation endpoints in this stage should focus on the expected value of the telemedicine service depending on the application that is being used (technology used and clinical purpose) but also on the way it has been implemented in daily clinical practice (service configuration).

Evaluation of the eWALL platform
When we consider this evaluation framework, it becomes evident that the eWALL project focuses on stage one, focusing on the feasibility and usability of the eWALL technology as well as stage two, focusing on the working mechanisms and the potential effects (see Table 1). Given the timeframe, the limited resources and the immaturity of the technology, it is not feasible to enter phase three and implement the eWALL platform in daily clinical practice and carry out a clinical trial.

Stage one	- Evaluation objective focusing on usability and feasibility of the technology. - Not implemented in daily clinical practice but evaluated as stand-alone service. - Evaluation design: case studies (n=10). - Evaluation endpoints mainly focus on technology used.
Stage two	- Evaluation objective focus on working mechanism and potential effect. - Not implemented in daily clinical practice but evaluated as stand-alone service. - Evaluation design: - Single case studies - Factorial design - Small case study (n≤50) - Evaluation endpoints mainly focus on technology used and clinical purpose.

Table 1: Task description WP6 and WP8 related to the DeChant Framework

4   M12 Prototype Evaluation

4.1 Introduction

The M12 prototype was assembled in a lab environment involving various end-users as well as informal caregivers and Human-Computer–Interaction-Experts to evaluate the prototype and provide feedback on basic setups, apps, services, notification messages and dialogue windows. Table 2contains the number and type of participants who evaluated the M12 prototype.

	M12 prototype evaluation				
	Participants	**Primary end-users**			secondary end-users	Expert Reviewers
	In total	MCI	COPD	ARI	Relatives, nurses, doctors	HCI experts, eHealth experts
Austria	10			8		2
Italy	10				5 (2 carers; 1 Occupational Therapist; 2 Physiotherapists)	5 (1 Geriatrician; 1 Bioengineer; 2 Neurologists)
Netherlands	8					8
Denmark	7				2 nurses	5

Table 2: Number and type of participants per country for M12 prototype evaluation.

The evaluation process in eWALL aims to create through users' and experts' feedback and design iterations, a functional, useful and well-accepted product for the end users. In this 1st lab evaluation session, the eWALL interface was assessed with a task-based approach combined with user experience questionnaires in the post-interview phase. Participants were requested to complete a set of defined tasks in order to evaluate the overall usability of the eWALL system. Usability is crucial for proper user-system interaction, and therefore serves as a prerequisite for the evaluation of other user experience factors. Once usability is improved, based on the results of the first evaluation, we will include in the assessment the specific factors related to the personalization and user experience of the system (modularity, adaptability, unobtrusiveness) in the Small Scale Evaluation. Additional factors such as motivation and acceptance were assessed in the subsequent Small Scale Evaluations.

4.2. M12 Prototype Evaluation - Primary Users

4.2.1 Methodology

- Task based approach:
  - o Scenario 1 evaluates the eWALL first prototype appearance and interaction modality. The following questions were asked:

    - *This is eWALL in inactive mode. It is mounted on the wall and it will be in the living room of the patient. What do you think about the information on the screen?*
    - *Do you think this setup makes sense for your patients/relatives/customers compared for example with a tablet?*
    - *Please now, approach the device. What do you observe?*
    - *What is your first impression?*
    - *Do you see any benefits or disadvantages for this setup?*

  - o Scenario 2 aims to provide an insight into the usability of the main screen in its active mode. Participants are given time to explore the interface, click freely around and think aloud. The following questions were asked:

    - *What is your overall impression?*
    - *Do you want to suggest any application that would make sense for such a device?*
    - *How do you think your patients /relatives/ customers could benefit from such a device?*
    - *How could you benefit? In particular, what kind of features in such a device could be useful for your work with older patients?*

  - o Scenario 3 focuses on evaluating the usability of the sleep application. For expert evaluators, the method of thinking aloud is being used as well. Participants are being asked the set of questions presented below:

    - *Can you find your sleep summary from last night?*
    - *Can you please rate your sleep from last night?*
    - *What do you think about this application?*
    - *What do you think about the tags used to describe sleep? Is there a better way to describe sleep?*
    - *From your experience, would older patients be interested in using such an application?*
    - *Any ideas/comments about how to move on with the implementation of this app?*

  - o Scenario 4 focuses on evaluating the usability of the daily functioning monitoring application. The same thinking aloud technique is being used. Participants are allowed to

freely explore the interface of the application during the thinking aloud process and the question answering process. This set of questions is presented below:

- *Can you find and open the DFM application?*
- *Can you describe what it does?*
- *Would it be interesting for older adults to review their day in this way?*
- *On what aspects of the life of your patients/relatives/customers would you like to be informed? And how often - daily, weekly, monthly, only when there is an emergency?*
- *What do you think about this application? How can we improve it?*

o Scenario 5 evaluates the usability of the daily physical activity monitoring application. Participants are being asked a specific set of questions to which they need to provide a verbal answer. This technique is being coupled by the thinking aloud process described also in previous scenarios.

- *Can you find and open the DPAM application?*
- *Can you describe what it does?*
- *What do you think about the presentation of the physical activity in this app? How can we improve it?*
- *How can we motivate older adults with health problems to exercise more? What works well in terms of motivation from your experience? Do you have examples?*
- *Would it make sense to add specific medical exercises for the COPD patients? And which ones? (only if the participant has experience with this disease).*

- Questionnaires
  o The UMUX is a standardized questionnaire measuring perceived usability using a four items Likert scale. The UMUX was developed as a shorter version of the SUS [10].
  o The HED/UT is a two-dimensional consumer attitude scale, consisting of 17 items measuring hedonic and utilitarian values. The HED/UT was initially designed for evaluating the consumer attitude towards products and for benchmarking between similar products of different brands [11]. In eWALL, HED/UT questionnaire will assess the user experience.
- Pre/Post Interviews
- Expert Reviews

4.2.2 Results

In December 2014 ten older adults from age 65 to 73 were invited to the ATE laboratories. Two men did not attend the test session and we tested the first eWALL prototype finally with five women and three men (mean age = 67.1). All persons are retired and live independently at home. After a short introduction about the objectives of the eWALL project the participants signed the informed consent. The whole evaluation session was video recorded with the participants' agreement. Next the participants were asked to answer a short mobile technology questionnaire asking for use of common mobile and Internet technologies like smartphone, WLAN or Facebook.

For the task based test phase the ATE facilitator encouraged the older adults to think aloud and comment their actions. Five different eWALL scenarios were tested:

- eWALL appearance and interaction trigger
- eWALL main screen
- eWALL sleep
- eWALL Daily Functioning Monitoring
- eWALL Daily Activity Monitoring

After the task-based test phase we asked the participants several general closing interview questions to assess their subjective opinion about eWALL. In the end of the evaluation the older adults were asked to answer the HED/UT questionnaire and the UMUX:

- The HED/UT is a two-dimensional consumer attitude scale, consisting of 17 items measuring hedonic and utilitarian values. The HED/UT was initially designed for evaluating the consumer attitude towards products and for benchmarking between similar products of different brands [2]. In eWALL, the HED/UT questionnaire will assess the user experience.
- The UMUX is a standardized questionnaire measuring perceived usability using a four items Likert scale. The UMUX was developed as a shorter version of the SUS [1].

Before we started with the task based test phase, participants were asked to answer the Mobile Technology questionnaire. Figure 3 shows the answers of the eight participants. All participants regularly use E-Mail. Interviews have revealed that many older adults of our target population already check their new E-Mails in the morning. In general there seemed to be a correlation between smartphone use and interest in new technologies: the facilitator noticed that older adults using a smartphone were more curious about new technologies and about eWALL in particular. Agarwal and Prasad [12] have defined the construct of personal innovativeness as the degree to which an individual is willing to try out any new information technology. Personal innovativeness is expected to be a stable trait which means that it is stable across different situations. Early adoption of technology is strongly linked with positive beliefs about the concerned technology and finally personal innovativeness. All five smartphone users regularly use WLAN and three out of four GPS users are smartphone users. All regular Google Maps users are smartphone users.

Further smartphone users had no problems with using the Touch screen we used for eWALL testing while two of the remaining no-smartphone users had severe problems with clicking the buttons on the screen since touch pressure was too soft or the touch interactions were too short.
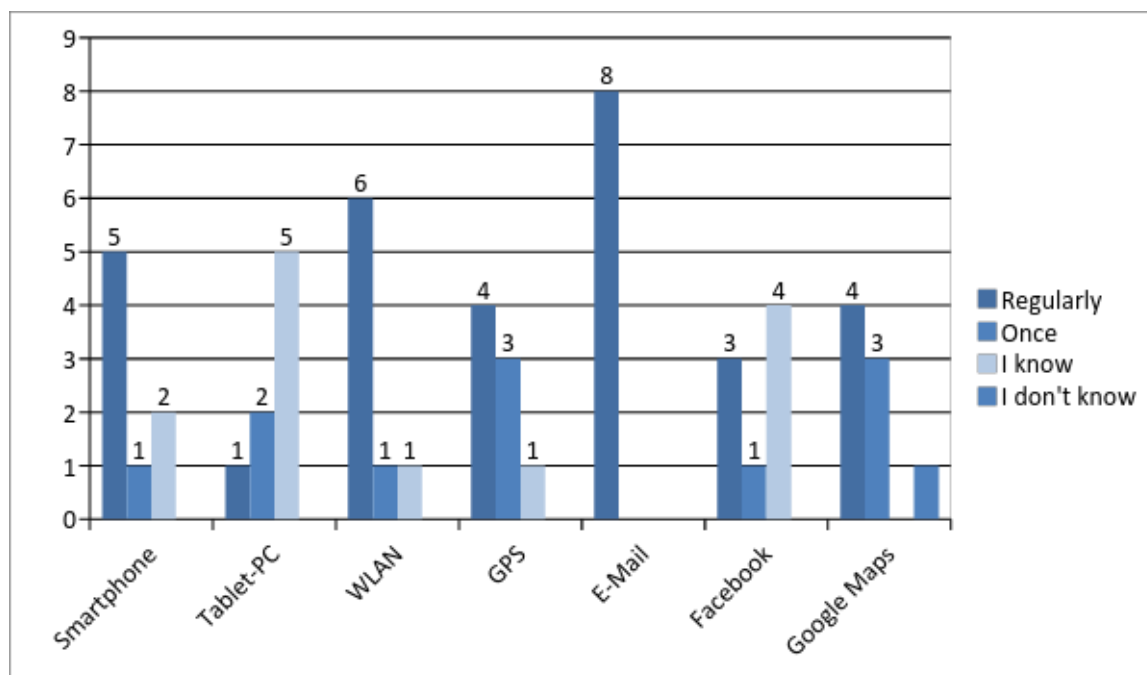
Figure 3: Results Mobile Technology Questionnaire.

4.2.2.1   *Results task based evaluation elderly users*

Scenario 1- eWALL appearance and interaction modality

In this scenario, our goal was to understand participants' first impressions of the setup, the information displayed in inactive mode and the transition to active mode.

First, all participants recognised the background image and reflected what the picture should show. They thought about a flower, an artichoke or a rose. Many of the test persons found the picture too pale and wished for more strong colours. Overall the background image seemed to be quite important for the participants.

The clock display was experienced positive, some wanted the date too. Temperature and location display were also recognized. Two persons would like to have some background music for eWALL in inactive mode.

Switching from inactive to active mode was for all participants comprehensible. They recognized the four menu buttons at the bottom of the screen. A few participants also observed the user profile box at the top left of the screen. At this early point of testing almost all users asked how eWALL would eventually be used: with a keyboard, a mouse, or a Tablet PC since all participants agreed that using eWALL with a large monitor on the wall would not be possible.

Older adults' first spontaneous ideas about using eWALL were quite similar. Since the idea of eWALL was connected with health, some of the test subjects wanted to be reminded for medicine

intake. Related health topics were nutrition and weight control (one person said that he would like to control every day his weight and adjust his meals accordingly). Nevertheless, the majority of participants associated eWALL with contacts and calendar: they wanted to manage all their contacts and calendar dates. Later in the eWALL evaluation process the test persons were told that they could directly contact their family members, friends and physicians and this made a positive impression to them. Table 3summarizes observations and following recommendations of Scenario 1.

Observations and Comments	Recommendations for Adaption
Importance of background picture	More strong colours for inactive modus
Clock display positive	Add display of date
Switch from inactive to active mode was noticed	No need for change
Need for management of calendar dates and contacts	No need for change

Table 3: Evaluation eWALL Appearance (Scenario 1).

Scenario 2 – eWALL main screen
Secondly, we evaluated meaning and understanding of the terms used in the eWALL main menu: Health, Contacts, House, and Games. Further, we went through the menu and asked for understanding of the applications listed in each main category.

After the menu term evaluation we asked the participants if they had spontaneous ideas for further applications. The majority of the tested users missed a calendar with all necessary appointments, reminders, to-do lists and shopping lists. One very intensive smartphone users wished for all his apps he was already using on his smart phone. Table 4summarizes observations and recommendations of Scenario 2.

Menu Terms	Users' Associations	Recommendations for Adaption
Health	Internet information about diseases/health, connection with eCard, nutrition and physical exercises recommendations, monitoring of health state markers (e.g. blood pressure), medication reminder, doctor appointments reminder, physical health feedback, guidance for day structure (to dos, shopping list, repairs), calendar and reminder of appointments, doctor recommendations and information	No need for change
Sleep	Many participants report sleep problems – tips for better sleeping, music for falling asleep, sleep diary for neurologist, sleep monitoring like in a sleep lab, sleep information about REM phases etc., smart home connection: is enough oxygen in the room, is the temperature ok?	No need for change
Life	Healthy nutrition, weight monitoring, nutrition and exercise program, calendar program, to do lists, shopping lists, smart refrigerator, feedback and monitoring of mood, doctor appointments, medication reminder, reminding of drinking, 3/8	Term "Life" is too broad. Users mixed it with movement and health

	had no idea what "Life" could mean	
Movement	Motivation for physical exercises at home, different individual exercises, exercises could be audio or video guided, motivation for going for a walk, bicycling, swimming,	No need for change
Contacts	All possible contacts (family, friends, caregivers, doctors, food delivers, workmen) and addresses, view if new messages are in, Call list, who is online? SMS, Whats App, Skype; pictures of contacts, reminder of birthdays of contacts, emergency button would be useful	No need for change
Family	Family members, friends, video calls	Extra category for friends? Other contacts
Caregivers	Home caregivers, list of agreements with caregivers,	No need for change
Doctors	General practitioner, Individual specialists, Google maps for navigation to the doctor, doctor appointments	No need for change
House	Majority had no ideas. Household lists, craftsmen, cleaning, checklist if herd or TV are off, alarm system, smart home (lights, windows), where are my glasses? Notice if a device is out of order, recipes and appropriate shopping list,	Term "House" is too broad
Games	Card games, cognitive training, chess, TV transmission of football or tennis games, crossword puzzle, one participant plays with the Wii Golf and Bowling	No need for change

Table 4: Evaluation of eWALL menu terms (Scenario 2).

Scenario 3 – eWALL sleep

In the third scenario, the participants had two tasks: to find the sleep summary and rate their own sleep from last night. The majority immediately found the summary. Most subjects were surprised that eWALL could monitor their sleep and in a first reaction they did not understand the motivation of sleep monitoring. This was also the first moment the participants recognized that eWALL is able to monitor them with sensors, and they did not feel comfortable with this. One woman noticed that a sleep protocol would be very helpful for caregivers to complete the health status of the older person. A few test persons found the sleep protocol interesting since subjective opinion about sleep quantity and quality is mostly incorrect. Many participants had sleep problems and one man even has to take sleeping pills in order to fall asleep. Most of them have problems to sleep through the night: they have to go to the toilet, or awake too early. Nevertheless, the majority would only read the sleep protocol regularly if severe problems with sleep occur.

Further, the participants had to rate their own sleep. This could be quite easily done by all persons but some were irritated that the tags were not able to choose in one step but the user had to click three times the "Add tag" button in order to choose the next tag. One woman remarked that she missed an "Enter" button. Rating sleep with stars could be done without problems.
The tags used for sleep rating were experienced as useful for older adults. Most participants would use the rating system once a week or once a month. Since their sleep was almost every night the same, sleep rating would be only interesting at the very beginning of using eWALL in order to

obtain a sleeping pattern. Table 5summarizes the observations and recommendations from Scenario 3.

Observations and Comments	Recommendations for Adaption
Sleep rating is irritating	Choosing tags in one step; completing with enter button
Sleep tags useful	Ideas for further tags: Falling asleep: uneasily, pain in general Waking up: stiffly, back pain, annoyed, lovingly, depressed If user has a specific disease a text tag would be helpful in order to have more detailed information

Table 5: Sleep rating (Scenario 3).

Scenario 4 – eWALL Daily Functioning Monitoring
In scenario 4, participants were asked to open the Daily Functioning Monitoring (DFM) application. They should describe what they see on the screen and tell their opinion about it.

Participants had problems to distinguish between "Life" and "Movement" in the eWALL menu. The majority would have not found the DFM without help of the facilitator. After having a look at the DFM, all participants were surprised that eWALL was able to monitor all their activities at home. They could not understand the motivation behind and disliked it very much. One woman noticed that this would be helpful for caregivers to see if the older adult had left the bed or had done some exercises if recommended. Another person remarked that this application could be useful for people with dementia. Only one test person could interpret this eWALL application positively and would use it for personal self-controlling: to motivate herself to be more active, do more sports. Many older adults wished for the possibility to switch the sensors off since they did not want their family members or others to have a look at their DFM. Almost all test persons felt too much controlled with this eWALL application and feared loss of self-determination. Table 6summarizes the observations and recommendations of Scenario 4.

Observations and Comments	Recommendations for Adaption
Life Term is not understandable	"Daily Routine" instead of "Life"
Icon in the activity bar is not understandable "lamp shade with rain drop"	Change icon
Test persons feel too much controlled	Switch sensors on/off

Table 6: Daily Functioning Monitoring (Scenario 4).

Scenario 5 – eWALL Daily Phys. Activity Monitoring
In scenario 5, participants were asked to open the Daily Physical Activity Monitoring (DPAM) application. They should describe what they see on the screen and tell their opinion about it.

The majority of test persons like this eWALL application very much. They are interested in monitoring of steps, activities and calories and appreciate this form of activity motivation. One man would like to have concrete nutrition recommendations. In general almost all persons were enthusiastic about the activity application and would use it if they had such a system. Especially

individual motivation and recommendations for activities by the system were appreciated. Nevertheless one male test person remarked that there are already Apps for smartphones available such as "Runtastic" which focus exactly on the same issue and are widely accepted by users. Table 7 summarizes observations and recommendations of Scenario 5.

Observations and Comments	Recommendations for Adaption
Movement category understandable	No need for change
Colours in the activity graphic for achieved and unachieved daily aims	Instead of yellow curves and bars: red for unachieved activity aims
Indoor and outdoor activities could be distinguished	Different colours for indoor and outdoor activities

Table 7: Daily Physical Activity Monitoring (Scenario 5).

4.2.2.2 *Results Post Interview and Questionnaires*

In the end of the evaluation session participants were asked some final general questions about eWALL. Further they had to answer the questionnaires HED/UT and UMUX.

In spite of participants' criticism overall impression about eWALL was quite positive. The three most common appreciated aspects of eWALL were health and movement, contacts and games. Especially the option of video calls pleased many participants. The majority of the interviewed participants could imagine using such a system if they would need more assistance because of illness and agreed that eWALL would be a great support for care persons and medical staff. At the same time there was one comment if care persons who are not very familiar with new technologies could handle eWALL appropriately. On the other hand, monitoring of eWALL was not appreciated for our very healthy sample of older adults. Furthermore, some persons asked for switching on/off of eWALL's monitoring function. Many test persons worried if the system with all its sensors and smart home functions would be affordable and if there would be enough space for such a huge monitor which is necessary for eWALL.

A male participant who is very experienced in smart phone use recognised that eWALL would not fulfil his personal needs since he is using a lot of apps on his smartphone already. He also criticized that eWALL integrated a lot of functions which were already available on the App market and could not understand why these already existing Apps could not be simplified for eWALL.

Furthermore, we asked the test persons if they would like to add functions for eWALL. The majority of asked persons were satisfied with the variety of functions already offered. A list of mentioned additional functions is as follows:

- Google Earth
- Navigation System
- Calendar
- Music and Photo library
- Collection of recipes

Finally the participants answered the questionnaires HED/UT and UMUX.Figure 4 shows the average values for all 17 two-dimensional user experience values of the HED/UT. Test persons had to choose 1 or 5 for very, 2 or 4 for rather and 3 for neutral. Particularly high average values can be found for "useful", "practical", "functional", "helpful", "efficient", "effective" and "enjoyable". In spite of this very positive rating 7/8 test persons evaluated eWALL as neutral concerning unnecessary and necessary and one person rated it as rather unnecessary.
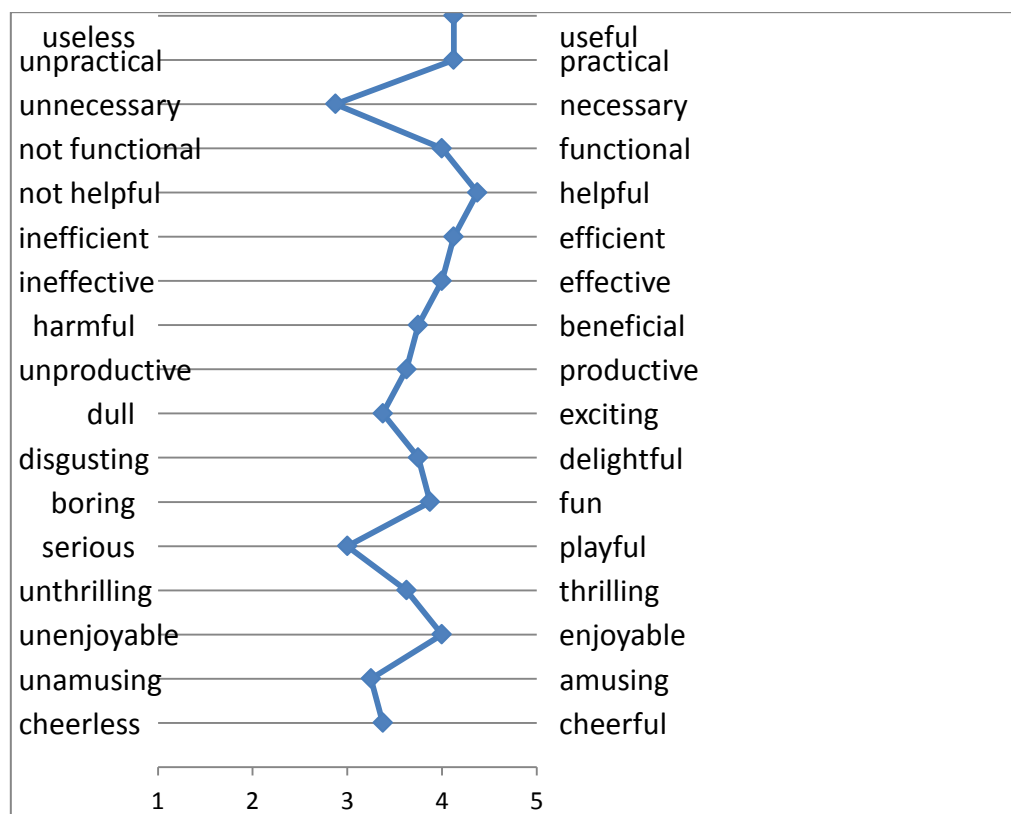

Figure 4: HED/UT Lab Testing eWALL M12 Prototype.

Regarding the results of the UMUX (Figure 5) none of the test persons was frustrated after using eWALL in the evaluation session. A total of 7 out of 8 participants agree or rather agree that eWALL is easy to use and only 2 out of 8 persons disagree or rather disagree that eWALL meets his/her requirements.
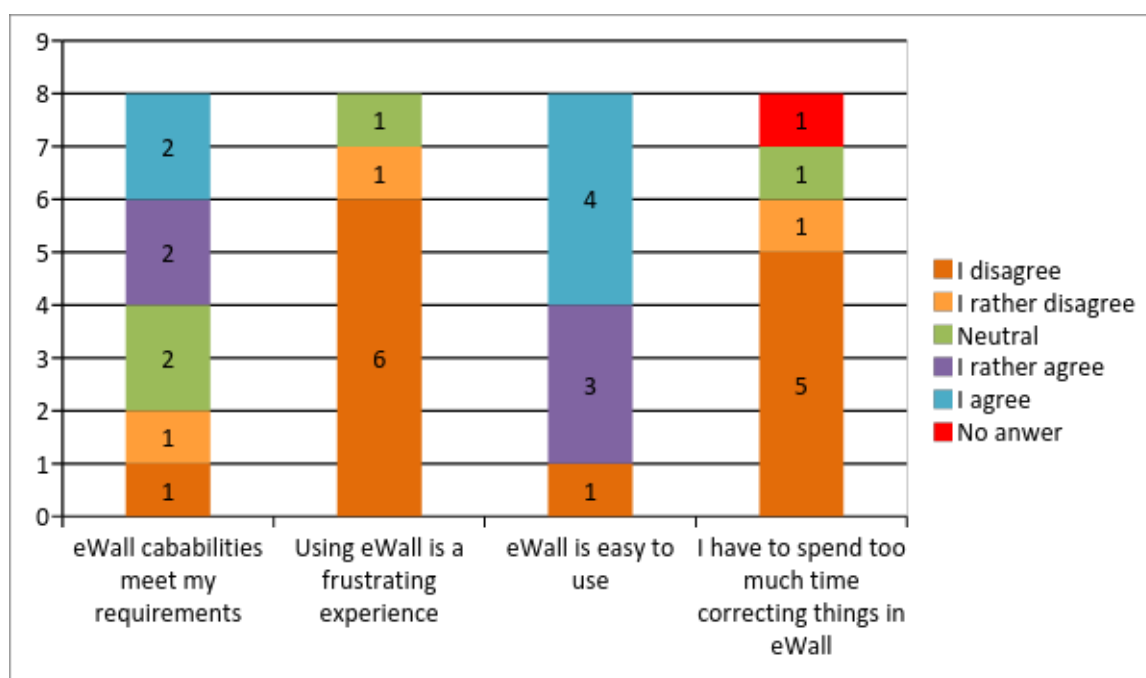
Figure 5: Results UMUX M12 Prototype Lab Evaluation.

4.2.2.3   *Discussion of primary user evaluation results*

Overall, older adults interviewed in this 1st lab evaluation study were very curious about the project and functionalities of eWALL. Particularly the health and movement section was something they were very interested in and the system's aim to motivate for physical training was well accepted by the participants. Further, video communication offered by eWALL was an idea the older adults liked very much. Nevertheless when the test person recognized that eWALL's functionalities were based on home monitoring they refused to have this system at home. Some mentioned the wish for switching eWALL on/off. Participants tended to forget the monitoring issue again and again during the evaluation sessions but were very clear in their attitude when monitoring came up again. This aspect would be interesting to evaluate in a future field study: Would older users forget about the monitoring aspect after some time using eWALL? Interviews of this study showed that users do not want to be monitored for 24 hours. Benefits of the system do not overcome the negative feeling of being monitored. Indeed, this is a main issue since eWALL is not possible to work without monitoring, and regardless usability or user experience of eWALL users will only use eWALL if they agree to be monitored. Following these results, it is necessary to discuss how to overcome this non acceptance of monitoring and find possible solutions.

From other AAL projects it is known that older adults tend to refuse technology solutions linked with care services if they do not need care or other assistance at the moment of interview. This is a problem of matching actual needs with a presented technology prototype. Since eWALL and also other systems of AAL projects tend to address many user groups (healthy users, MCI users and user with a chronic disease) these technologies also offer some kind of care services for healthy users who do not want to imagine themselves needing care. From a psychological point of view this provokes automatically resistance of the tested users. Further this diversity of user groups also includes another problem: active older adults tend to be quite interested in new technologies and

have no need for technologies especially designed for older adults whereas older adults with existing care needs mostly are not familiar with new technologies and have very specific needs concerning a technology they would use. In this evaluation session we interviewed a man who intensively used his smart phone with many apps and a very first reaction to eWALL was that he was already using a lot of apps similar to eWALL functionalities on his smartphone. Probably eWALL will not be attractive for older adults already using mobile technologies such as smartphone or tablet PCs since these persons do not need usability assistance in form of a specific designed interface such as in the eWALL project.

Another concern of the participants was if they could afford a system like eWALL. Besides the screen for eWALL, installation of sensors and linked smart home applications will require costs which are not imaginable for older adults and probably also not for other age groups since the idea of Smart Homes is still somehow futuristic. For future evaluation sessions it is necessary to provide some information about financial costs for the test users in order to achieve trust that the tested system is a realistic technology researchers in a test session are discussing with the participants. Otherwise, the tested technology stays unreachable and user can not sufficiently imagine the tested scenarios.

4.3   M12 Prototype Evaluation– Secondary Users

4.3.1 Methodology

Participatory heuristic evaluation (PHE) is a participatory inspection technique that serves as an extension to heuristic evaluation defined by Nielsen. In PHE, experts in software engineering, usability or human factors do the inspection as in traditional heuristic evaluation. However, work-domain professionals (WD-professionals) are added as a group of users doing the same inspection. The purpose of extending the heuristic evaluation with these WD-professionals is to complement the traditional inspectors' more abstract knowledge with very specific knowledge about the work-domain. In addition, the thinking aloud technique by Nielsen was added. The technique is characterized by the expert who has to describe the first expression of what he observes on the interface, what he likes and dislikes regarding the interface, what irritates him when he uses the system, etc. By using the thinking aloud- technique, we received the experts' 'real' opinion of the interface and thereby collected information to improve the interface.

In January 2015 five experts in Biomedical Engineering and usability – on going PhD fellows or associated professors – participated in PHE sessions of the eWALL interface in Denmark. The experts had different levels in heuristic evaluations. In general, the experts had collected or participated in PHEs before with more or less experience. The sample consisted of four men and one woman, (mean age = 31 year). Two healthcare professionals, e.g. nurses, participated as WD-professionals. Both of them had experience in doing in PHE. Before the PHEs began, the seven experts (5 usability experts and 2 WD-professionals) received information about the objective of the test, followed by an introduction to 15 heuristics by Muller et al., 1998 [13], a severity rating scale,

a log schema, thinking aloud technique, and finally information about some tasks scenarios (see also Annex II).

In the following chapters, the aggregated number of times each heuristic was used during the five PHEs will be presented. Afterwards, the results related to: eWALL Login; eWALL Main Screen; eWALL Sleep; eWALL Health 'Daily Functioning Monitoring'; eWALL Health 'Daily Activity Monitoring'; eWALL Contacts, will be presented. Finally, in the end of the last chapter, a recapitulation of the general findings will be listed.

In March 2015, eight experts in Human Movement Sciences, Biomedical Engineering, Psychology and Computer Science with expertise in usability of telecare and telemedicine technology participated in usability tests in the Netherlands. Following the usability protocol for task based evaluation, participants were explained the nature of the experiment, and the workflow of the experiment session. After the task based evaluation, each participant responded to a short interview. The method of thinking-out-loud was used to get participant input. Each session was sound recorded, after participants agreed to it and signed the informed consent form.

4.3.2 Results

Figure 4 illustrates an overview of the aggregated number of times each heuristic was used during the five PHEs. The usability experts used the two heuristics: *Aesthetic and Minimalist Design* and *Consistency and Standards* most, respectively 52 and 34 times. The WD-professionals used the three heuristics: *Consistency and Standards, Match Between System and The Real World,* and *Pleasurable and Respectful Interaction With The User* most, respectively, 21 and 2 x 20 times. The total use of heuristics (the usability experts and the WD-professionals together) were: *Aesthetic and Minimalist Design, Consistency and Standards,* and *Match Between System and The Real World* (Figure 6).
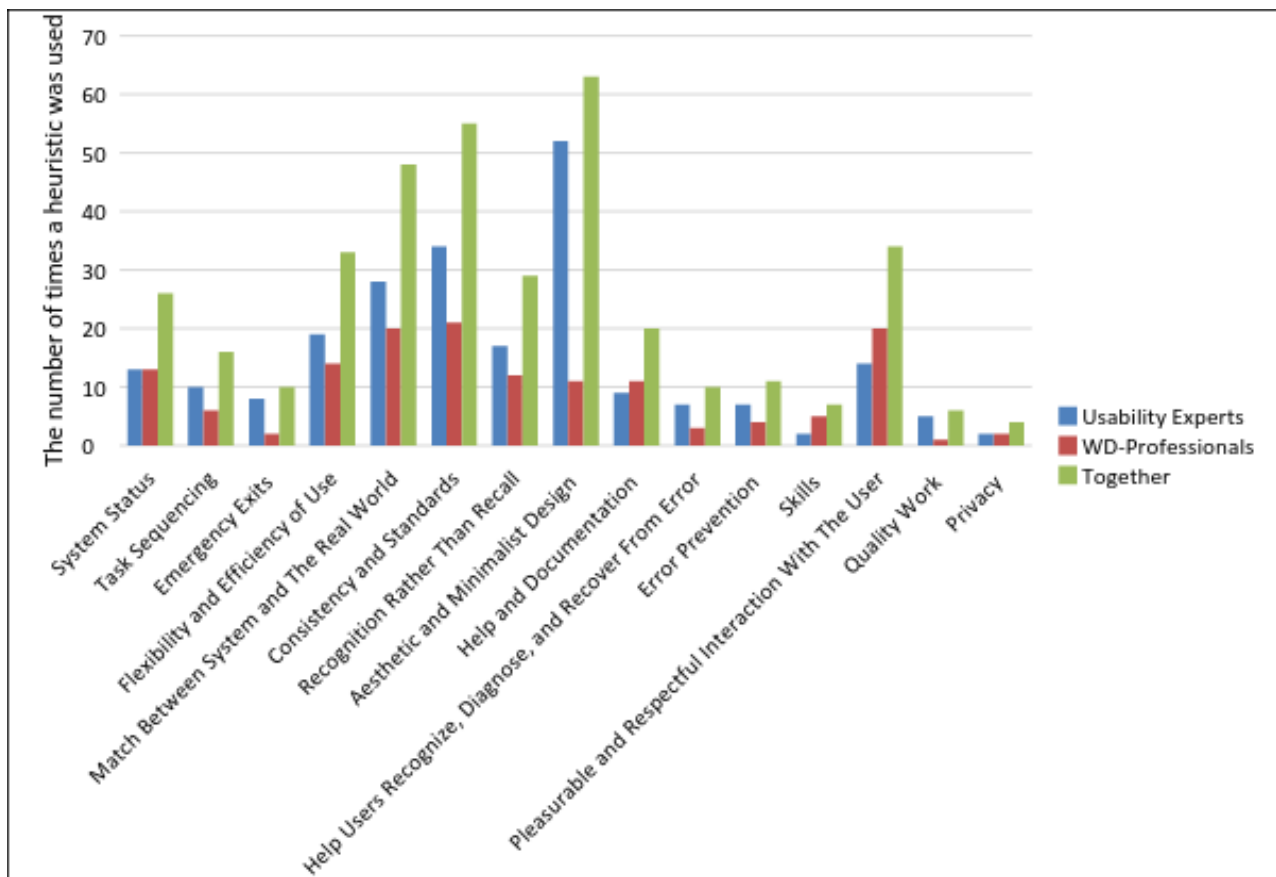
Figure 6: An overview of the aggregated number of times each heuristic was used during the five PHE.

From the severity rating scale, the usability experts used the severity 1-*Cosmetic* 49 times; (21.6%), the severity 2-*Minor*; 49 times (21.6%), the severity 3-*Major*; 67 times (29.5%) and the severity 4-*Catastrophe*; 62 times (27.3%), (Table 8).

Severity Rating	Usability Experts	%	Work-Domain Professionals	%
1 – Cosmetic	49	21.6 %	9	6.2 %
2 – Minor	49	21.6 %	33	22.8 %
3 – Major	67	29.5 %	68	46.9 %
4 – Catastrophe	62	27.3 %	35	24.1 %

Table 8: The table gives an overview of the usability experts and the work-domain professionals' use of severity during the participatory heuristic evaluations.

eWALL Login

Regarding eWALL Login (Figure 7), our goal was to understand the participants' first impression of the setup, the information displayed in inactive mode and the transition to active mode. The usability experts and the WD-professionals were asked to choose their native language, Danish, type their user name and come up with comments of the interface design.
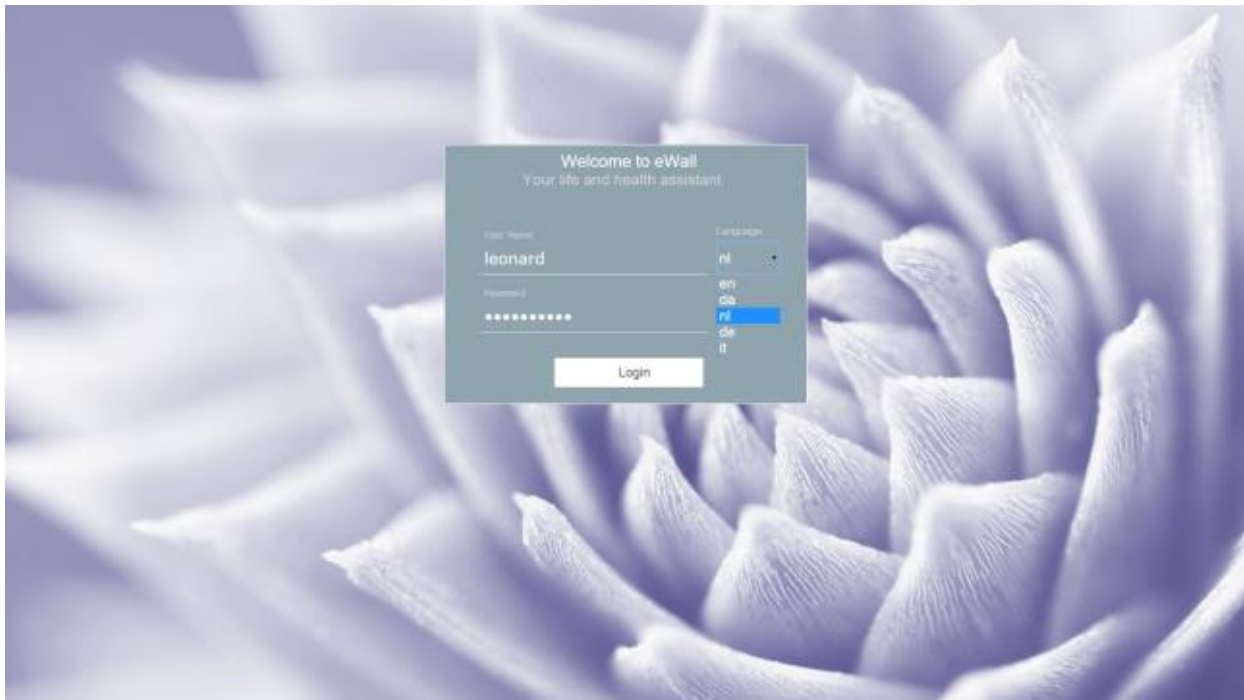
Figure 7: A screenshot of the login pace, as experienced by the test participants.

The experts and the WD-professionals found the following remarks:

- If you don't speak English or understand English, it is difficult to choose your native language because the command language is English.
- The size of the font is too small.
- The font is unclear.
- The colour contrast is too small.
- The written text does not change even though the experts and the WD-professionals pressed "da".
- The experts would rather like a small flag instead of the abbreviations: "da", "en", "nl", "de", or "it".
- The experts and the WD-professionals missed a keyboard to type in their user name.
- A required password is missing.
- It is not obvious how the users can change the language.
- The artichoke does not make sense. It is irritating.
- A keyboard is missing.

4.3.2.1 eWALL first impression and passive/active modes

In this section we describe how the tested participants perceived the entire eWALL metaphor, the passive screen (Figure 8) and active screen (Figure 9) modes and switching between them. Prior to this test, participants were explained the metaphor of the wall in the context of use. The proxemics trigger was experimentally set up at a distance of 1.5 meters from the screen, allowing the average human to touch the screen surface. The switch trailer was emulated by the experiment coordinator

using The Wizard of Oz technique, by pressing a key which would trigger the switch in the interface.

Figure 8: Screenshot of the passive screen as experienced by the test participants.
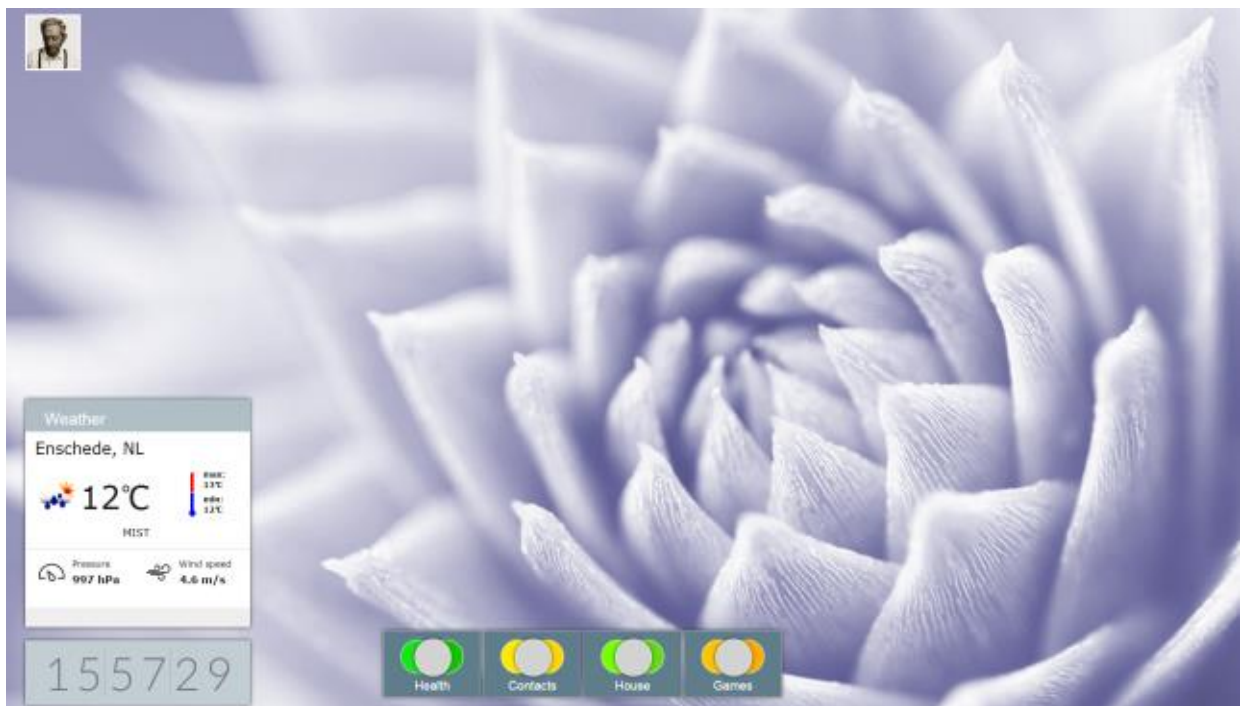

Figure 9: Screenshot of the active screen as experienced by the participants.

The experts in work domain professionals found the following remarks:

- The information being displayed in passive screen, namely the widgets and the background image are understandable and readable from distance.
- The passive screen is perceived as sleep mode.
- Persons with mild dementia would have trouble understanding the concept of interaction with proximity.
- The date is missing from the clock widget.
- A personalized picture should be displayed instead of the general background image.
- Generally speaking elderly should be educated in what passive and active screen mean and how they switch by approaching the screen.

When asked about the difference between the current screen and a tablet 50% of the participants were undecided about the added value of the current large screen, 25% were pro tablet and 25% saw the added value of the large screen. However, the following remarks were made:

- The visual change between states is not relevant enough, not noticeable enough.
- The main menu does not make a noticeable appearance in the active mode.
- Changes in the weather widget were not significant enough to be noticed.

When asked about their first impression in the Netherlands: 3 out of 8 participants considered this approach bad another 3 considered a good and the remaining 2 were undecided. The following remarks have been made:

- The screen could be overwhelming for elderly when experienced from close proximity.
- The screen is true too bright, especially for a small living room.
- It makes sense to enable the screen, when the user approaches it, in order to protect the user privacy.
- The transition between passive and active should be smoother and more meaningful.
- Only one person can use the system right now there could be more than two persons willing to use the system together.
- There is no possibility at the moment to adapt the height of the screen; elderly cannot sit when operating with the screen.
- Mild dementia users may forget that it won't screen changes when approached; the screen should invite user to interact with it.

4.3.2.2   eWALL Main Screen

Regarding the Main Screen, depicting in Figure 7, our goal was to discover how the experts and the WD-professionals explored the interface of the main menu, and comment the different elements and the usability of the side. The experts and the WD-professionals in found the following remarks:

- The watch is unclear to understand. Choose an analogue rather than a digital watch.
- It is difficult to interpret all the weather-information.

- Regarding the weather-application - there is information overload.
- The information about the weather is only necessary for users with COPD.
- The COPD-users are not interested in Celsius – but are interested in air humidity – that information is missing.
- It is unclear, whether it is possible to make the weather-application bigger or interact with it.
- Nothing happens when you press at the profile-button in the left corner.
- There is no logout button.
- The symbols (circles) in the button are not understandable.
- A date is missing.
- It is unclear what "house" covers.
- The main screen seems empty.
- The grey buttons look like they are moveable.
- Remove the seconds from the watch – they disturb.
- The clock is not directly recognizable as a clock.
- In the weather widget pressure is not necessary and not understandable.
- The main menu should be placed in the center of the screen; the menu buttons should be bigger.
- The screen light management has to fit the living-room ambient, it should be fully customizable.
- An accidental proximity switch could produce anxiety or confusion.
- The main menu labels are not representative for the button functionality; it is not clear that the main menu buttons are groups of buttons.
- The title in the launched applications are different than the labels on the buttons at launch them.
- It is nice to have the confirmation that the system recognizes the user; this feature should be emphasized.
- It makes little sense to group applications if they are so few 3 or 4; instead a brief overview of the outcome of using this application should be shown, when pressing the group buttons.
- The device is very visible in the living room due to its size.
- The buttons in the menu are too small and don't really stand out.
- When the screen switches to active mode one participant expected already some results regarding health state of the user.
- The screen is very clean and this makes the user have to go further into the system to discover functionality.

Get the above remarks are a combined list of outcomes from both Danish and Dutch usability tests. In the Dutch trials 90% of participants acknowledged the need for elderly users to be trained in understanding evil. A recurring remark was that the current set-up looks very much like a TV, therefore elderly will not perform a touch-based interaction with the device, if not instructed to do so.

4.3.2.3   eWALL Sleep

Regarding the eWALL Sleep, our goal was to collect information about how usable the experts and the WD-professionals found the application. They were asked to rate their sleep, choose adjectives for their sleep, and give comments of the interface in general. The interfaces are visible in Figure 10, Figure 11, Figure 12 and Figure 13.

Figure 10: Screenshot of the first view.

Figure 11: Screenshot of the sleep summary view.

Figure 12: Screenshot of the tag overlaying window.

Figure 13: Screenshot of sleep details page.

They identified the following remarks:

- The headline is too small.
- The font is too small.
- The headline is not placed in the middle of the side.
- It is unclear how to use the stars (rate your sleep).
- It is irritating, the way to choose information about your sleep.
- It is not possible to save entered information.
- A 'save'-button is missing.
- An 'end'-button is missing.
- The content is difficult to see through.
- Too much waste of space of the interface.
- The cats are annoying.
- It is unclear how the pictures correlate to the text.
- Information overload regarding your sleep.
- A help-service is missing.

- The consistency is missing regarding using number and letters (1 hour, two minutes).
- CAPS-log has been used regarding one of the subheadings.
- The information is irrelevant regarding the boxes.
- All the experts thought they had received all the information when they pressed the button further information, but it is possible to get further information than already.
- There is no correlation between the time and the line between the boxes with sleep information.
- The backward arrow is too small.
- The backward arrow is placed in the wrong corner – move it down to the button.
- Nothing happens when you press the cog wheel in the left corner.
- It is only possible to add one information regarding your sleep before the box closes down.
- It is only possible to add three information regarding your sleep.
- It is not possible to exit without adding one sleep information.
- The icons don't make sense.
- It is difficult to identify "what's in it" for the users.
- The sleep application can be easily found in the health application group.
- The button labels industry publication were mostly clear although the font is too small.
- To sleep summary understandable but it is too long, it is tedious read it should be shortened.
- The sleep summary was expected to be shown immediately after launching the application, not after pressing the SHOW SUMMARY button.
- The star rating is understandable, although not highlighted. Users will not use this feature unless directed to. It should be made more obvious.
- The tag/label rating is confusing. Participants expected to select all three labels at one time and expected to find a done button to save them.
- The tags / label panel was overlaying the summary. This should not happen.
- It is not obvious why one should use the tags / labels. This feature should be visually highlighted. Users should be asked for this input.
- The five star label can be somewhat restrictive; include the possibility of rating half stars.
- The sleep details are more clearly structured then the sleep summary. The sleep detail information should be extended and written in a more understandable way.

As a conclusion the sleep application will be redesigned to fit most of the participant's feedback. A first action will be to combine the textual information from the sleep summary with the good structure of the sleep details into one view. Button phones will be enlarged buttons with the redesigned to a more understandable shape and size. Subjective measurements of sleep will not be carried out through the labelling mechanism, but a questionnaire compliant with current subjective sleep evaluation practices will be integrated into the application. A reduced sleep summary will offer a one sentence conclusion over the last night sleep measurement.

4.3.2.4   eWALL Health 'Daily Functioning Monitoring'

Regarding the eWALL Health 'Daily Functioning Monitoring'-application (Figure 14), the experts and WD-professionals were asked to describe their experiences with the elements and the interface.
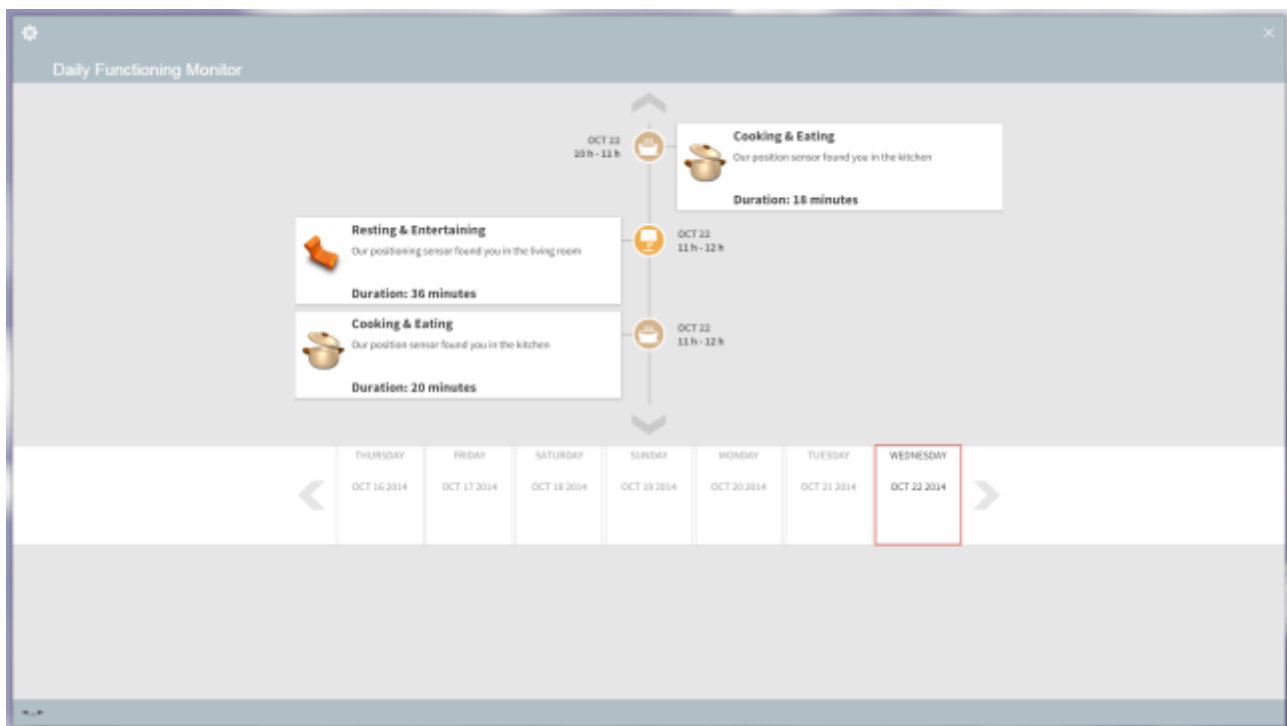
Figure 14: A screenshot of the Daily Functioning Monitoring Application, as experienced by the participants.

They identified the following problems:

- It takes too long time to open the application.
- You can't see if the system is working or if it has frozen.
- It is unclear how the users can interact and use the information to.
- You don't get a calendar when you press the button 'calendar'.
- Irrational that the dates are placed in the button of the screen.
- It is not clear which date is showed.
- The design of this application does not look like the others.
- In general, the MD-patients may feel like they are monitored.
- The label in the button life does not clearly indicate the purpose of this application. An alternative must be found.
- The title of the window daily functioning monitoring is too complex to be understood by general population. An alternative must be found.
- 50% of the Dutch test participants did not understood what the application that was at first glance, 38% could identify what the application basically does, 12% could not formulate an answer due to two applications failure to show data.
- The arrow buttons should be labelled with proper words such as: previous, or next; up, or down.
- The application has a thick bottom border which is currently not being used to display any kind of information this border should not be displayed after redesign.
- The vertical timeline with everyday events should be extended to show more than 3 events.
- All buttons are big enough to be touched constantly.

- The labels on the day timeline buttons should have a much bigger font.
- The labels describing the date and time in the vertical timeline for each daily functioning event should not shorten the temporal information, because it creates confusion. A full description of the date and time should be used example: 27th March 2015, 4:05 p.m.
- The icons in the daily life events are fairly intuitive but the text in the daily life events need to be written in a more human understandable way.

When asked to estimate frequency of use for chapped occasions by elderly users the Dutch participants could not formulate a decisive answer in 7 out of 8 participants, yet they all estimated a rare use of the application, as it is presented at the moment. They could, however, understand the usefulness of such information, primarily for caregivers. One participant expressed that the application will be useful for elderly and we'll create self-awareness. Two participants expressed interest in representing some form of physical activity with each daily functioning event displayed on a vertical timeline. One participant expressed a wish to integrate possible or scheduled daily activities for the future such as: scheduled visits, future house chores, medicine intake timings.

4.3.2.5   eWALL Health 'Daily Activity Monitoring'

Regarding the eWALL Health 'Daily Activity Monitoring' (Figure 15) the goal was to collect information about the availability of the interface. The experts and the WD-professionals were asked to check their daily steps, go back and see their results from the last month, and were asked to check their goal. In the Dutch trials, 7 out of 8 participants found the application with ease, with the general remarked that the application name should be more intuitive than the current "Moment". 6 out of 8 participants understood at first glance what the application does; one did not understand the functionality of this application and in one case the application did not load due to software failure.
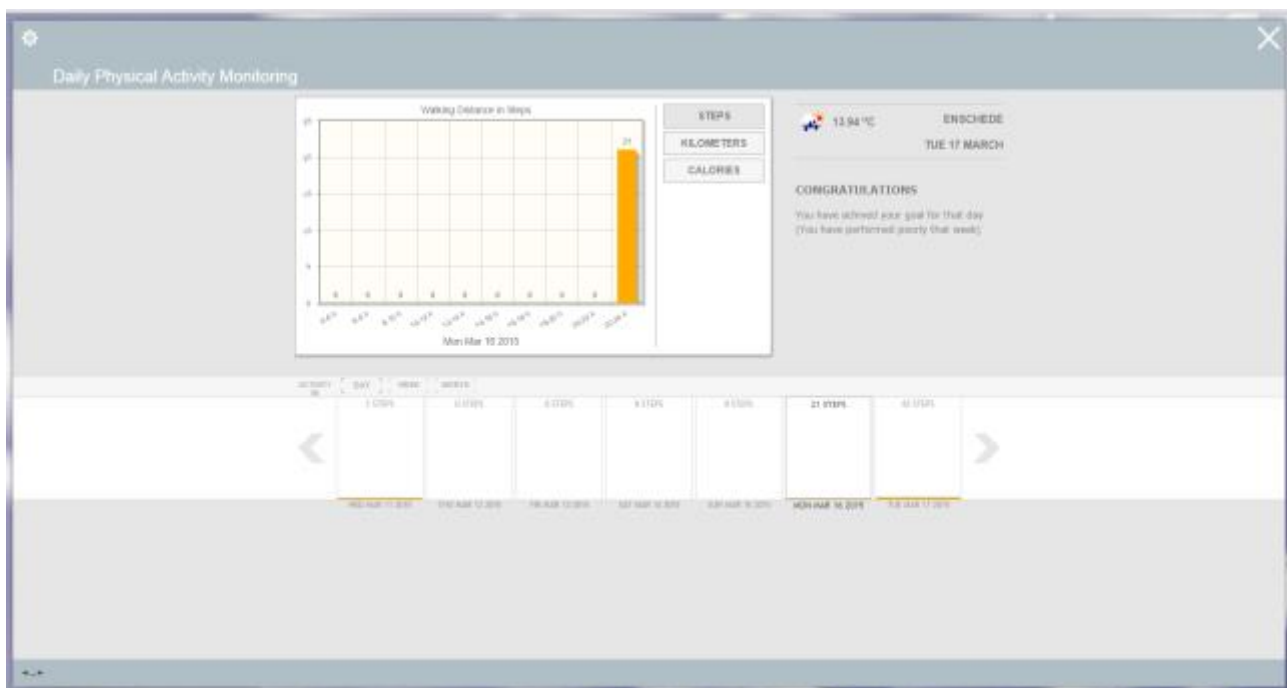
Figure 15: A screenshot of the Daily Physical Activity Monitoring Application, as experienced by the participants.

The experts and the WD-professionals identified the following remarks:

- It takes too long time to open the application.
- You can't see if the system is working or if it has frozen.
- It is unclear how the users can interact and use the information to.
- You don't get a calendar when you press the button 'calendar'.
- Why are the colours changing from grey to green, orange, and yellow?
- The graph is difficult to understand.
- The font size is too small.
- Who decides the goal and why the congratulation?
- The diagram misses terms (X,Y).
- The time interval is not understandable.
- The difference between the used colours is too small.
- The temperature should not be displayed with decimals. Make the temperature an integer number.
- It is not clear what the temperature role is in the current application interface: does it show the current weather (and why)? Does it show the weather in that day, when viewing a passed day?
- The chart is hard to understand because the bottom line doesn't have a linear distribution of values it starts first with a 6 hour interval and then the following intervals describe each two hours. The vertical axis is not labelled properly and the numbers vary for each day according to maximum and minimum activity detection.
- There is no goal representation of the activity. For elderly it will be not clear if the current chart illustrates good or bad behaviour.

- The colour coding of each data column in the chart is not intuitive. Elderly will need to be explained why green is good behaviour (physical activity quantity closer to the goal) and orange represent bad behaviour (not enough physical activity).
- The weather association with the quantity of movement is not is not clear as is displayed in the current prototype. Participants could not understand if the current weather is displayed or a weather recording of a specific day is being shown. If the purpose is to associate weather with amount of physical activity the weather should be recorded for each graph column.
- Measuring a physical activity in kilometres is not meaningful for elderly. It creates a feeling of not doing enough because it shows low the channel numbers. Doing zero point something does not create a feeling of accomplishment. This should be transformed in meters as it will provide a larger natural number.
- The calorie measurement is not intuitive enough. Generally elderly will not understand how much burn calories are good and how much is bad. What are the minimum calories that one should burn everyday? What are the maximum calories that one should burn everyday? Will this be calculated according to the personal energy expenditure level?
- The motivational message is not displayed clearly it should be much more emphasized.
- When reviewing a past date a short summary would be very useful. Was it a good day or bad, in terms of physical activity?
- It is not possible to compare between days unless one memorizes the first day and compares the mental model with the second day. This activity could be very difficult for elderly. In the horizontal timeline switching between days, weeks and months measurements is fairly easy and intuitive.
- Calories on talk time line items however have small fonts, the date label under them is shortened and how to understand. The date label should contain the full description of the day. Example: Friday the 27th of March 2015.
- In the bottom of the application there is considerable unused space which is not used. The chart should be made bigger and horizontal timeline should be moved at the bottom so as this space disappears and there is more place for larger interface items.

As a conclusion after using deprecation all Dutch participants to the trials have understood its purpose an added value as it shows elderly overview of their physical activity in the present day and all the past days since installing the system. The weather and the congratulatory message are not relevant for this application as they are presented the right now. They should be redesigned, according to the remarks above or removed from the application as they create confusion at the moment. The measuring units for physical activity should be made more intuitive. At the moment only steps are directly understandable measure by elderly, with previous experience of physical activity monitoring applications. Using this application will require a minimum amount of training for the primary users. Representing the goal of effective it is essential for this application each day should have a conclusion about the activity quality whether the user perform good or bad thing to the set daily goal. Loading the application took a considerable amount of time; in the Netherlands, on average, more than 5 minutes until the application rendered interface elements. After selecting another day or showing monitoring data the application would require similar amounts of long time to display the new requested information.

4.3.2.6   eWALLContacts

The goal of this scenario was to collect information about the game

- It is unclear what the difference is between doctors and caregivers.
- Use an iPad keyboard – the users know that.
- The keyboard does not contain the Danish 'æ, ø, å'.
- A keyboard is missing.
- Help-information is missing.

4.3.2.7 A recapitulation of the general problems

This recapitulation contains more general usability issues and problems, the usability experts and the WD-professionals identified on eWALL: Login, Main Screen, Sleep, Health 'Daily Functioning Monitoring', Health 'Daily Activity Monitoring', and Contacts. The remarks were:

- It is unclear what is interactive.
- Missing consistency.
- Irritating that everything is not translated into Danish.
- The headings are in general too small.
- Feedback is missing when the users press on the buttons.
- It is unclear how the users easily can exit.
- It can be too hard for the user to stand and do the exercises.
- The experts and the WD-professionals would like that the watch is visible all the time and not only on the main screen.
- A keyboard is missing.
- Help-information is missing.

4.4   M12 Use Cases Evaluation - COPD patients

4.4.1 Methodology

In April 2015 seven COPD patients from 49 to 77 years (M= 63,3) old participated in Vienna in order to evaluate five suitable use case scenarios. There was a completely new interface (Figure 16) designed by ATE - focusing only on physical training and related videos for the sections: breathing, mobility, strength and endurance.

Figure 16: Main menu page eWALL physical training.

After signing the informed consent forms main issues of the eWALL project were explained and participants answered two questionnaires:

- CAT (COPD Assessment Test): a helpful tool to measure COPD patients' burden and impact of the disease on every life. Issues are typical symptoms of COPD such as being out of breath, feeling of tightness in one's chest, coughing, and consequences on sleeping and every day activities. Score ranges from 0 to 40 (maximal burden).

- Mobile Technology Questionnaire: use and knowledge of technologies such as smartphone, WLAN or Facebook is asked.

In a next step specific COPD use cases were evaluated with a new designed physical training interface including diverse training videos for different training issues (breathing, flexibility, strength, endurance). For this reason task-based evaluation was chosen for two COPD specific use cases: "Video assisted training: supervised session" and "Personalized goals". Participants were motivated to think aloud and talk about their ideas how to navigate in the user interface.

Further eWALL notifications for following eWALL themes were evaluated using pictures from the Pexels free photo database. (www.pexels.com)

- Alarm detection (3rd COPD use case)

- Poor air quality (4th COPD use case)

- Weather and appointments

- Drinking
- Showering
- Sleeping
- Breakfast
- Eating
- Going out for a walk
- Riding the bike
- Physical activity
- Cognitive training
- Social activity

Finally the COPD patients were interviewed concerning their physical training habits at the moment and use of measurement technologies such as pulse oximeter.

4.4.2 Results

COPD patients CAT score ranged from 18 to 33 (40 is the highest score to achieve). Five of seven participants were already retired and two live in a care home. Figure 17 shows that majority of COPD patients is not used to smartphones or tablet PCs. Use of e-mail, WLAN or Google maps are more common.
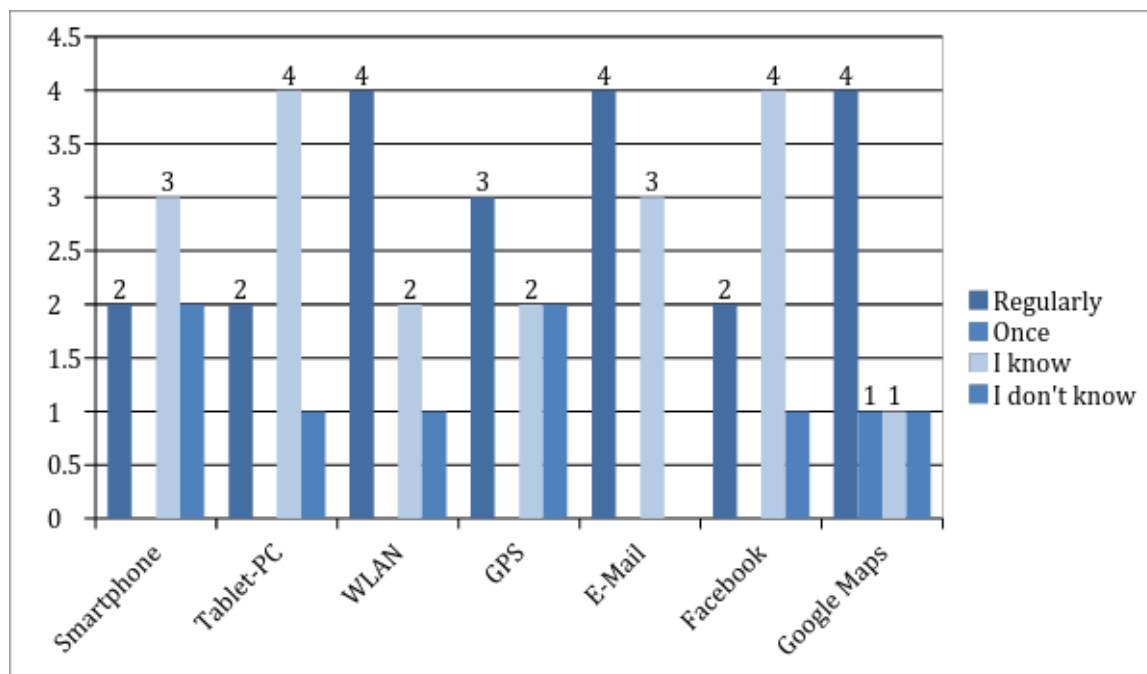

Figure 17: Mobile Technology Questionnaire Use Case Evaluations with COPD Patients.

4.4.2.1 Results task based evaluation

Use-case 1: Video Assisted Training: Supervised session

Task 1: Start eWALL individual training plan
Four of seven users find it easily to start the individual training plan. Nevertheless one user has problems to navigate through the further individual training plan menu and does not understand the concept of arrows to click further. Another user has no experience with computers and is too busy to deal with computer use in general.

The participants were asked to comment the graphic "New batch": only one user seems to understand the information transported with the graphic. Users try to interpret the meaning but the graphic is too unclear for them. One user suggests to show training process with a curve based diagram.

Task 1.1: Please open a training video of category "Mobility"
Five users are able to start the mobility training video easily. Two users needed minor support. None of the users complained about it, but it would be more user-friendly if the video would enlarge automatically full screen after selecting it.

Task 1.2: What does participant think about eWALL training program? Would they use it at home for training?
The majority of users like the idea of an individual home training very much. Only one user cannot imagine to use eWALL for this reason. Two users stressed the importance of measuring daily well-being before starting the training since physical well-being can change from day to day with COPD and the training plan should consider this. For two other participants it was very important that training videos were suitable for COPD patients and the female participant even wished that the trainer in the video should have COPD one his own to make the videos more authentic. In general users were quite concerned that physical training could be too exhausting to them since they used to have problems with physical exercises and regular training in the past.

Task 2: Finding the goal setting, let describe what they see
Since the interface for the physical training menu was in a mock up level designed the users had problems to find the goal setting in the menu. At least four users had the idea that changes of goal setting would be located in the "Hallo Claudia" button but could not start the goal setting.
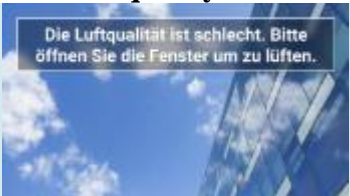
General comments on eWALL
In general participants liked the idea of a computer based physical training at home. One male participant who lives in a resident home recognized that the screen size of eWALL would be too big for this apartment and that he would prefer a very small screen which he could position like a picture frame since he would not be allowed to install the screen on a wall.

4.4.2.2 Results of notifications evaluation

After explaining the notification concept in eWALL to the participants the users were asked to comment on the pictures shown and rate them from 1 (=I like very much) to 5 (=I don't like at all). Results are listed in Table 9.

Notification issue	Summary of user comments	Average user rating	Need for change
Alarm detection level of Oxygen			
	Users like the picture, message is clear; one user misses head; only one user completely disliked the picture and felt motivated to get into the same yoga position which he felt unable to do	2,1	No
Poor air quality			
	Four users found the picture fitting to the message; the other three users were concerned about showing a big city where air quality uses to be bad; one user stated that the picture showed a building with air condition and windows cannot be opened;	2,6	Yes -> Opened window would be nice
Weather and appointments			
	Two users found the picture too childish; weather image is shown tree times – this was irritating for some users; one female user wished for more realistic information such as fine dust, air humidity or pollen information for which she already uses an app on her smartphone	3	Yes -> more realistic picture and more COPD relevant weather information
Drinking			
a)	Nearly all users liked the picture and agreed on importance of drinking; only one female participant disliked the pictured because she associated a commercial aspect with it	1,7	Yes -> delete question modus
b)	Three people preferred picture a) one female participant liked picture b) with the text of the first drinking picture	1,4	Drinking a) was favorited

Showering			
a)	For the majority the notification was clear. Two users found this kind of notification not important; feelings after showering are not addressed with this picture	2	Yes -> Picture addresses no emotions after showering
b)	Picture reminded participants more on holiday or swimming than showering; request completely unclear	4	Yes -> Picture not useful

Going asleep			
a)	Participants found notification not meaningful to them – can decide on their own	2,7	Yes -> Sleep notification only for MCI patients?
b)	Notification was for some very unclear and was associated with going outdoor for a walk or watch the stars	2,3	Yes -> New picture is needed for sleep notification

Breakfast			
	Notification was clear	2,1	No

Meal			

	Most participants associated the picture with a healthful meal and cooking;	1,8	No

Going out for a walk

a)	Participants found the request nice and clear	1,6	No
b)	Picture was not inviting enough for the participants. Picture a was preferred	3,3	Yes -> Picture is not suitable
c)	Picture and massage do not go together. Participants did not like this picture	-	Yes -> Picture is not suitable

Bicycling

	Users did not feel invited since bicycling is not possible for them or they find it too dangerous in the city;	3,2	Yes ->Bicycling maybe not suitable for our target groups

Physical Training

	Picture and text were accepted; maybe a more dynamic picture can be found for the final prototype	2,3	Yes -> more dynamic picture

Cognitive Training

	Picture not suitable; request not clear for many participants	3,5	Yes -> change picture and text
Social Activity			
	Request was clear for most participants; social activity is not necessarily linked with drinking coffee	2,5	Yes -> more focus on meeting others

Table 9: Well-being Advertisement overview and feedback.

In general users preferred large scale pictures instead of smaller picture with the eWALL main menu background since text can be read without glasses. Further, decision modus (YES/NO) was refused and majority liked more the motivation by citations or slogans. Most COPD patients would accept to be addressed on a first name basis.

4.4.2.3 Results Post interview COPD Patients

At the end of the evaluation session COPD patients were interviewed about their habits concerning physical activities. Since COPD is a very restricting chronical disease sports are very limited to this population. Nevertheless four participants did regularly some kind of sports: swimming, hiking, Kieser training or ergometer training at home.

Finally we asked the COPD patients if they were already using some kind of application or program in order to measure physical activities or health values like pulse or oxygen saturation (Figure 16). The most frequently used tool was blood pressure instrument. Other measuring possibilities or applications were hardly used by the COPD patients. The use of wearable technologies was quite aninteresting idea for some users.
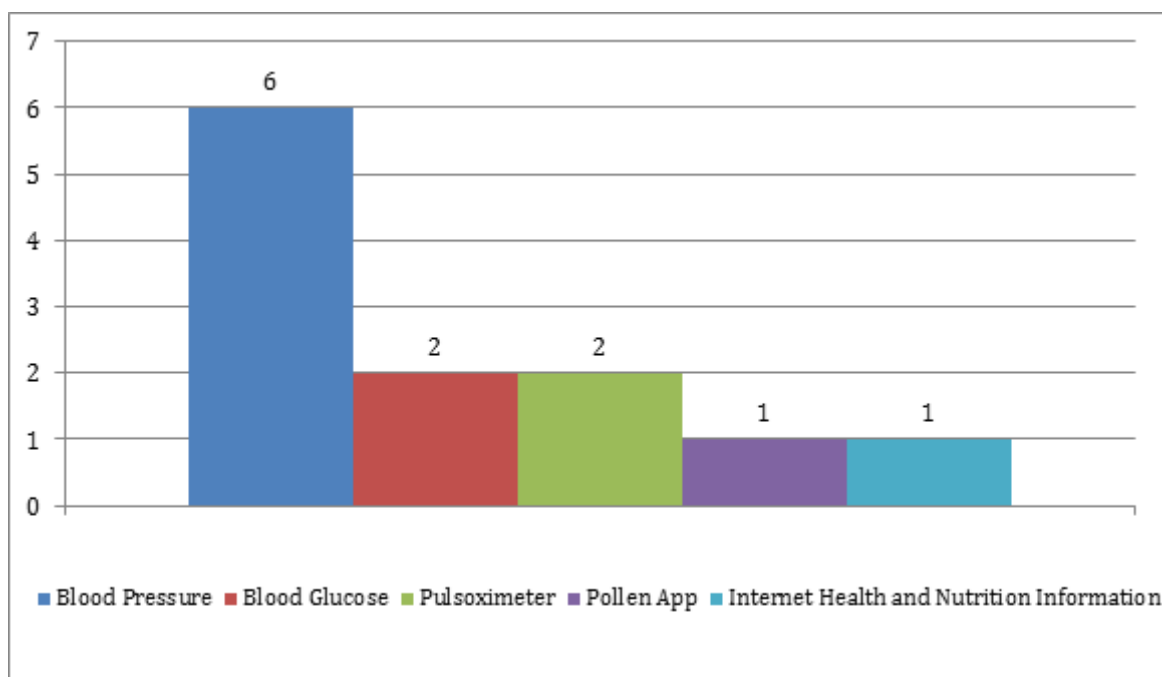
Figure 18: Physical Instruments used by COPD patients.

4.5   Implications for the next eWALL prototype

The feedback from user tests, conducted by evaluation partners, were summarized and provided to the technical and design partners in form of a table (see below). All inputs coming from users were represented as recommendations to have a clear picture about what to change or adapt in eWALL. We coordinated the prioritization of the recommendations with all evaluation partners by rating the priority from 5 (high priority) to 1 (low priority). This allowed to get an agreement of the importance of the recommendations within all evaluation partners.

The following two tables show the recommendations for technical (Table 10) and for the design adaptations (Table 11).

Rating Ø	Service	Recommendation
5	DFM	The daily information displayed should be reduced to only meaningful events. Not show if person is moving between

		kitchen and dining room with a high frequency.
5	Settings	Some notifications (showering, sleep, should only be used for MCI patients => other users feel uncomfortable with these notifications. The settings should give the user the decision what notifications he wants to have.
5	Settings	Test persons feel too much controlled --> possibility for user within settings to switch sensors on/off.
4,7	DFM	The typical/pattern comparison is not represented. How is the currently selected day different then a normal day? How can one compare two days? The user would have to get a complete mental image of the first day and compare it with the second one, at the moment. Add an overview/conclusion of the day and represent the lifestyle reasoner information.
4,5	Settings	The settings should contain a regulation of the screen's lightness (eventually day/night mode?)
4,5	Weather	The COPD-users are interested in air humidity – add that information to the weather application.
4,3	Settings	There should be a possibility for user to change the setting. Make a button in the main screen to get access to that functionality.
4	Weather App	Future weather prediction would be very useful.
3,8	Clock	Add display of date to clock display.
3,5	All services	Nothing happens when you press the cog wheel in the left corner. Remove the icon if it is not going to be activated.
3	Clock	Make the watch visible all the time and not only on the main screen. So when the users are in 'Contact' they can still see the watch.

Table 10: Recommendations for developers.

Rating Ø	Service	Recommendation

5	Main screen	Main screen requires redesign as a whole in terms of less sub-categories. "House" and "Life" is too broad. We may also delete Contacts? ("Daily routine" instead of "Life"?)
5	Sleep App	Redesign of the sleep app is required: Questionnaire instead of adding rates? More visual graph instead of text information? Cats or no cats? Other icons? There is too much noisy information now. The timeline should be horizontal. Last night's sleep should be visible in one page without scrolling. Make clear if the sleep was good or bad.
5	DPAM	The activity chart is understandable, but very scientific. - Remove the background grid and the numbers on top of the pillars because it would scare the users. - Make the graph bigger. - Add terms for X,Y - Use normal time intervals (not 0-6h, 6-8 h,...) - Difference between grey nuances is too small. Choose more different colours.
5	DPAM	Elderly want to see if they reached the goals, or if they have done enough for that day, and this information is not shown by the application.
4,8	Main screen	The main screen seems empty. No need to have so much free space in active mode. Make the buttons and widgets bigger for instance. The buttons are too small and they have to really stand out, they are too close together. Those 3 circles for the buttons are irritating; the grey buttons look like they are moveable. Icons would be better
4,5	All services	Be consistent with the font. Titles, content, message, accent text should be the same font and size across all UIs.
4,3	DFM	It's unclear how the user can interact with the applications. Change the design of buttons etc. to immediately see what is a button and what not.

4.3	DPAM	It's unclear how the user can interact with the applications. Change the design of buttons etc. to immediately see what is a button and what not.
3,8	Weather App	There is an information overload in the weather app. Change the look of it to make it clearer, otherwise choose another design.
3,6	Notifications	Pictures: - No childish pictures, more realistic pictures. - Show the pictures over the whole screen. - Persons shown with pictures should represent the target groups (e.g. physical activity shown should be realistic to do for an older person ->bicycling is not suitable for older adults)
3,5	DPAM	The distance in kilometres is not relevant. It should be displayed in meters, without decimals. It is easier to understand a natural number (how many? N many) as a quantity then a fraction (doing 0.something of activity). A feeling of quantity can make a user proud of.
3,3	DFM	Irrational that the dates are placed in the button of the screen. Move it up higher so users don't have to scroll and change to a horizontal timeline.
3,3	Notifications	Text: Motivation text can be direct or in slogan form but NO decision format and not too much text.
3	DFM	It is not clear which date is shown. Make that more clear for instance by using stronger colours.
2,8	DPAM	Why are the colours changing from yellow to green? Choose just nuances of one colour. (If it is possible to make it clear to the users it would be nice to have different colours for indoor and outdoor activities.)
2,3	Weather App	It is unclear, weather it is possible to make the weather-application bigger or interact with it. Make it look like an application if it is going to be an application.

2,3	DPAM	Calories are not easy to understand, there should be explained what it means to burn the displayed amount of calories.

Table 11: Recommendations for designer.

Those recommendations as well as on going new design or functionality concepts were discussed in the weekly xWP-d conference calls. The participation of technical and medical/evaluation partners made sure to discuss adaptations on an interdisciplinary level and allowed to get important input from both sides at a very early stage of development.

5 Small Scale Evaluations - M19 Prototype

5.1 Methodology

The M19 Prototype Evaluation together with the M22 and M25 prototype evaluation form the "Small Scale Evaluations" (SSE). These lab trials cover the first evaluation stage as presented in the DeChant framework (Figure 1). This stage aims at the evaluation of the usability of a *low* fidelity prototype of the eWALL system (Jansen-Kosterink, 2014). The tests were again conducted by eWALL partners in Austria, Denmark, Italy and the Netherlands. The main objective of the Small Scale Evaluation was to <u>evaluate the usability of this eWALL prototype,</u> based on the level of user satisfaction and their intention to use this prototype of eWALL.

The aim of the SSE was to test the eWALL prototype on different user groups and experts:
- Persons with age related frailty (ARI)
- Persons suffering from COPD
- Persons suffering from MCI
- Secondary Users (informal caregivers of the three groups above)
- COPD Experts
- MCI experts
- Usability experts

In April 2015 all evaluation partners collaborated for creating a table to plan the number of participants involved in the three circles of SSE (Figure 19).

	C1 (M19 / May)	C2 (M22 / Aug)	C3 (M25 / Nov)	Users involved
ATE	Participatory Heuristic Evaluation: **Usability + COPD Experts** (n=5)		Users: **Elderly** (n=10) **+ second. Users** (n=3 -> informal caregiver/ nurse/…)	**10 ARI** **3 Second. Users** **5 Experts**
RRD	Usability and UX evaluation: ARI elderly (n=5)	(perceived) Effectiveness ARI elderly (n=5) COPD elderly (n=5)	Potential for adoption ARI elderly (n=5) COPD elderly(n=5)	**15 ARI** **10 COPD**
AAU	Participatory Heuristic Evaluation: **Usability + COPD Experts** (n=5)		Participatory Heuristic Evaluation: **Usability + COPD Experts** (n=5). Given a successful PHE test on users: COPD-patients in lab (n=10).	**10 Experts** **10 COPD**
IRCCS		**Users (n=4) MCI** **Second. Users (n=5)** Physician/ informal caregiver/ nurse/ Telecare expert…)	**Users (n=4) MCI** **Second. Users (n=5)** Physician/ Neuropsychologist / nurse / …)	**8 MCI** **10 Second. Users**
Overall	5 ARI

10 Experts | 5 ARI 5 COPD 4 MCI 5 Second. Users | 15 ARI 15 COPD 4 MCI 8 Second. Users 5 Experts | **25 ARI** **20 COPD** **8 MCI** **13 Second. Users** **15 Experts** |

Figure 19: Small Scale Evaluation Participants Planning.

Due to delays in design adaptations after the M19 prototype evaluations, the two evaluation sites ATE and AAU decided to shift the evaluations till the new main screen design could be evaluated. Testing the "old" interface again would not give more insights from users. Thus, ATE and AAU conducted evaluations with the M22 prototype and just one site (RRD) evaluated the M19 prototype, which is described below.

5.2 M19 Prototype Evaluation - Primary Users

5.2.1 Goal

In the Netherlands, a better functioning version of the first year prototype has been tested with ARI elderly users. Our goal was to determine the user experience of eWALL as perceived by elderly in a laboratory setting.

The test will address questions regarding the usability and user experience which answers will determine design and development gaps in the interface and interaction concepts. The design and development team will have to address these gaps in the next prototype. The test was designed to address the following questions:

Usability
- Can people understand the eWALL concept?
- Can people understand where to find and how to use the sleep application?
- Can people understand where to find and how to use the DPAM application?
- Can people understand where to find and how to use the DFM application?

Experience
- Is it easy to get familiar with eWALL?
- Can people perform tasks in eWALL without unnecessary effort?
- Does the user feel in control of the interaction?
- Do users like/dislike eWALL?
- Is it exciting/motivating to use eWALL?
- Is eWALL innovative in the user's view?

5.2.2 Methodology

A total of 10 Elderly with age related impairments were tested in the Netherlands. The inclusion criteria were:
- the tests should be performed with elderly with age related impairments
- the elderly participants need to stand for the whole task based testing period (maximum 45 minutes)

The tests took place in a laboratory setup using the ELO 42" touch screen running the improved version of the first year prototype. The participants experienced the full user interface implementation. The test was designed around the participant thinking allowed while performing eWALL tasks and contains 3 main parts:

The **first part** introduces eWALL, it's concepts and Bob as a primary user persona. Participants are told to explore the eWALL interface freely while thinking allowed. The researcher performing the test said:

"This is Bob's eWALL. He is 65 years old and uses eWALL since a while now. eWALL is a technology that will help Bob live independently at home. It is still in the prototype phase and we are interested in your opinion on how to design it, in order for people to understand it and benefit from it.

Please approach the screen and feel free to press the buttons. Don't worry, you cannot break anything."

As the participant is using eWALL, the researcher aims to observe
- What is the user's first impression of eWALL: Screen size, distance to the screen, standing interaction, touch screen use, active/passive mode etc.
- How does the user interact with eWALL.

The **second part** consists of asking the participant to perform several tasks while thinking allowed, as listed below:
Task 1: Sleep Monitoring
- *Could you see when did Bob go to sleep last night?*
- *Explain what can you discern in the graphic of last night's sleep? What do you see?*
- *How many sleep interruptions does it show you for the last night?*
- *For how long did Bob sleep last night?*
- *Add a rating of your sleep in the last night.*

Task 2: Daily Physical Activity Monitoring
- *How much physical activity did Bob make last week?*
- *Did he meet his goal?*
- *Is physical activity important to you? Do you think this application could help you keep track of how much you exercise?*

Task 3: Daily Functioning Monitoring
- *Bob wants to keep track of his daily activities. Could you see what did Bob do today?*
- *Do you think this application could help you keep track of how you live?*

Task 4: General Comments
- *What do you think about this system?*
- *Would you use the system at home? Why (not)?*

The third part consists of assessing the user experience of the participant. They were asked to fill in the UEQ [14] (Figure 20) questionnaire based on the following question:

"Based on your experience today what do you think about eWALL?"

	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

Figure 20: The user experience questionnaire.

Figure 21 shows the procedure of the experiment:

Greetings and introduction 5 minutes	Informed Consent 5 minutes	Part1 10 minutes	Part 2 30 minutes	Part 3 5 minutes	Conclusion, Thanks, Good bye 5 minutes

Figure 21: Procedure of M19 prototype evaluations with primary users.

5.2.3 Results and Implications for the M22 Prototype

Part 1 outcomes:
Most participants understood Bob's condition and they associated it with eWALL functionalities.
Contrary to previous usability tests (performed with domain experts in the Netherlands), participants remained at the interaction distance from the screen. Some (3/10) occasionally approached the screen even closer, in order to better read text interfaces, when these were displaying small font sizes.

The touch gesture was comfortable, even though 8/10 users possess and use touch screen technologies, like tablets and smartphones. Swiping in order to scroll through the interfaces was less intuitive; 6/10 participants had trouble understanding when to perform this gesture.
Al most all participants had problems with the interaction flow due to inaccurate phrasing of button labels and the button clustering in the main menu.

Part 2 outcomes:
The results of these evaluations were reported as recommendations for developer and design partners.

Sleep Monitoring:
All participants were not monitoring their sleep behaviour prior to the tests. Reaching the sleep application was clear and direct for 8/10 participants, while the other 2 needed guidance. After launching the application, the information displaying Bob's sleep from last night was clear for 8 participants, although 2 out of these participants did not scroll through the interface (Figure 22 and Figure 23).
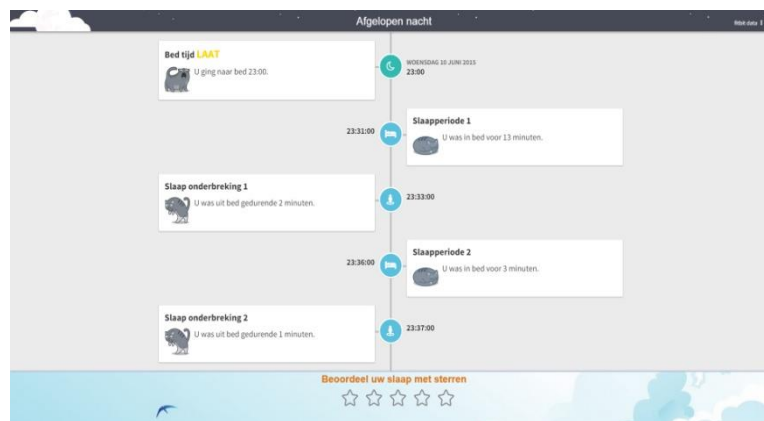
Figure 22: The sleep timeline as displayed when accessing the application.
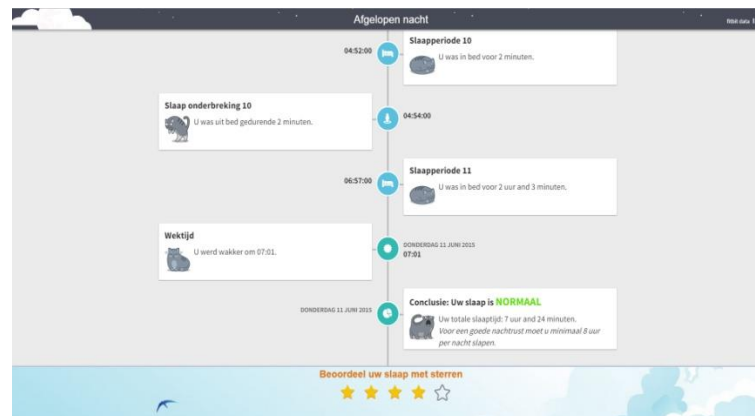
Figure 23: The sleep timeline after swiping until the bottom.

As participants had difficulties with scrolling, reaching information from the bottom part of the timeline was difficult. Therefore, only 3 participants could directly answer "How long did Bob sleep last night?", as this information is displayed in the last card and, therefore, required participants to scroll all the way down in the timeline (Figure 23).

We concluded that a better interface visualisation with more condensed data, displayed on one screen only and not needing additional interaction in order to discover more information is needed.

Daily Physical Activity Monitoring:
All participants claimed that physical activity is important for them and they see the value of this functionality as positive for them. Using the interface was difficult, as some participants (5/10) could not discern the interactive and non-interactive parts without explanations. The colour coding was not clearly presented and, therefore, 7/10 participants did not correlate the colour scheme presented by the application with the intended good/bad marking.

Only 5/10 participants could answer correctly, without guidance, to the question: *"Did he meet his goal?"*

We conclude that a more simplified information visualisation should be implemented. The colour scheme is not intuitive and will require an explanation. The concept of activity goal expressed in a number of steps/day was new to all participants.

Daily Functioning Monitoring:
Monitoring daily activities automatically was a new concept for all test participants.

All participants understood the data displayed on screen, but 6/10 did not scroll, or were not able to perform navigation interactions. The list of daily activities was vague for most participants; 3/10 tapped the activity boxes and expected to get more information about Bob's activity. When asked if they would use such application: 4/10 said they would not use such application because they would feel controlled, while 6/10 participants consider the application helpful and would sometimes use it, but if the information is clearly presented.

We conclude that this application should be available on user request only, as acceptance of the functionality proves to be debateable. The interface needs to be simplified in order to fulfil its primary purpose. Although some participants asked for more details when looking at Bob's daily activities, they requested less information when thinking of themselves as possible users, with the motivation that they already know what they've done.

General Comments:
This eWALL prototype was perceived as useful, but was not clear enough in terms of both information presentation and interaction modalities. Several redesign requirements were compiled (Annex IV) and promoted to the design and development partners. A total of 4 out of 10 participants said they would use the tested eWALL prototype, while the rest of 6 participants considered Bob as a much frailer person than themselves, therefore motivating their negative answer.

Part 3 outcomes:
This part explores the user experience of the participants using the UEQ questionnaire [14]. The resulted statistics show an imbalance between the experience and usability. This is interpreted as the effect the technological novelty. Therefore an above average attractiveness score contrasts the bad perspicuity score. This means eWALL is beautiful, but difficult to use and hard to intuitively learn. Most difficulties were reported in the main active screen, which hides the access to eWALL applications in a category menu. The application mase were not intuitive enough and the main

screen was generally perceived as empty. Some participants stated that the screen is too big and seems too heavy for the limited functionality it provides. The effectiveness parameter counts for each participant's perception of their own task accomplishment. This parameter is average. The Dependability score is low, does to the fact that Bob is a simulated user and some data provided during the tests did not make sense to the participants. For example: the DFM reported Bob resting for 1 minute. This induced the feeling of imperfect technology. eWALL is perceived as very stimulating, yet the novelty score is low. Most participants compared it to the tablet, or smart phone technology they possess and stated that more applications on eWALL would make it better than these. The size of the screen in contrast with the size of the interface elements contributed to the low scores in novelty (Figure 24).
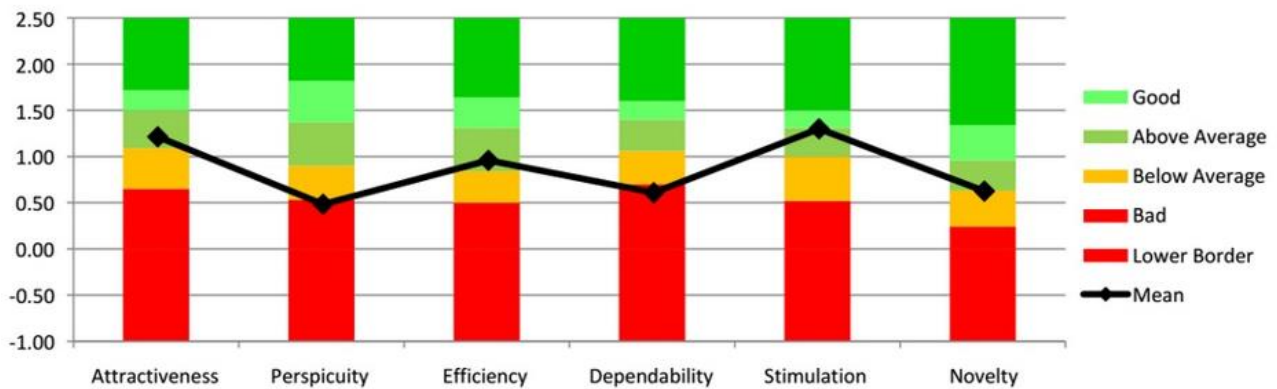
Figure 24: UEQ results M19 prototype evaluation.

As a conclusion, the test outcomes prove that eWALL is a promising technology and elderly would adopt it if the information presented is clearer and better structured. eWALL is perceived as a personal device, but the metaphor is not well represented. This leads us to recommend a redesign of the eWALL metaphor and main screen.

6   **Small Scale Evaluations - M22 Prototype Evaluation**

6.1 Goal

In order to evaluate the M22 prototype from an expert point of view Participatory Heuristic Evaluation was chosen for the partners in Austria and Denmark. Overall 10 experts should evaluate the new eWALL interface.

6.2 Methodology

After signing the informed consent form, the eWALL project was particularly explained with the help of the three user scenarios Bob, Michael and Jane in order to demonstrate the three target groups of eWALL. For expert evaluation task-based approach and thinking aloud technique were used. The evaluation was divided into 11 tasks: main screen, personal data, daily functioning monitoring, daily activity monitoring, sleep monitoring, sleep rating, cognitive training, playing games, physical training, rating the training, general experience. The experts were asked to comment on the different interfaces of eWALL and related services and functions. Table 12 shows the main eWALL screens demonstrated for evaluation. The protocol of the M22 Prototype Evaluation can be found in the Annex II.

Task	eWALL Screen
Task1: Main Screen	

Task 2: Personal Data

Task 3: Daily Functioning Monitoring
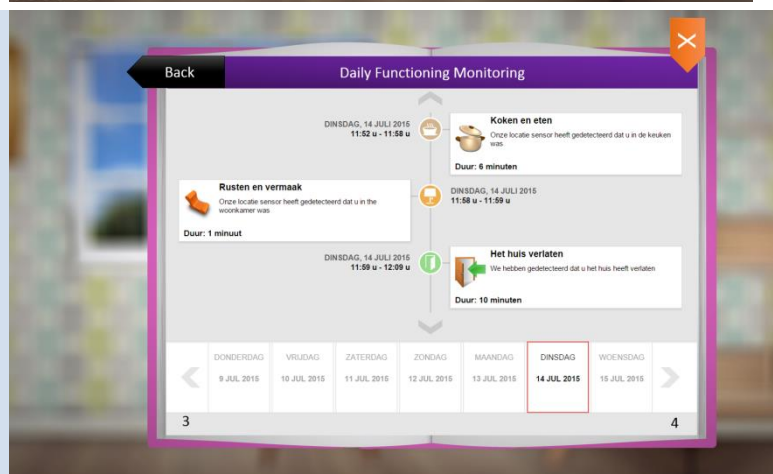
Task 4: Daily Activity Monitoring
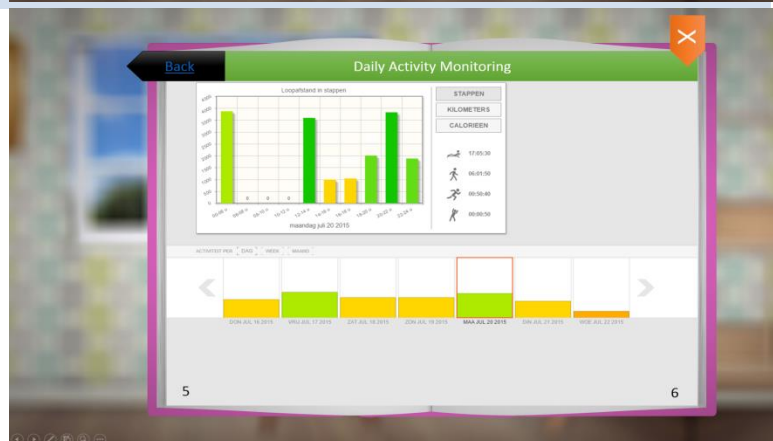
Task 5: Sleep Monitoring
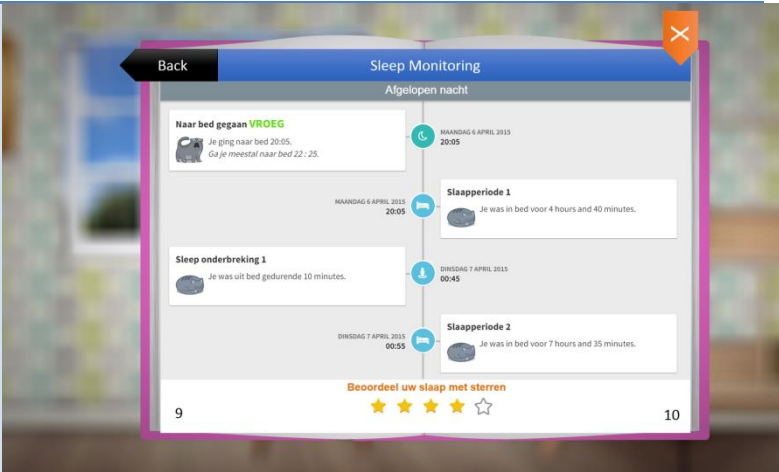
Task 6: Sleep Rating
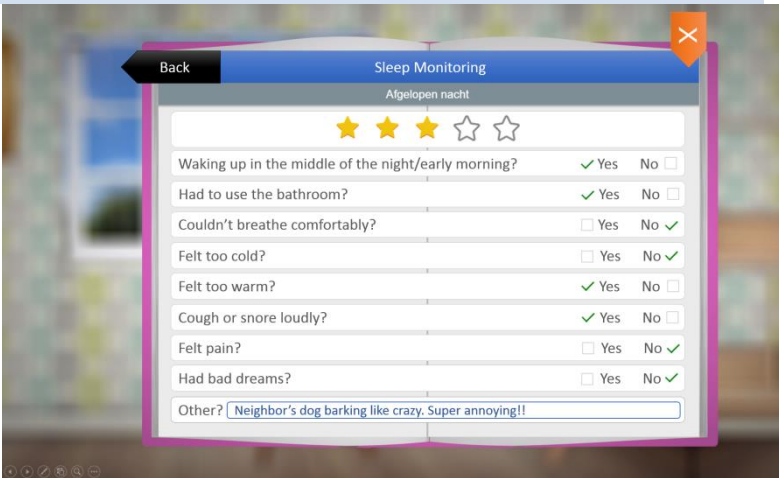
Task 7: Cognitive Training

Task 8: Game #1
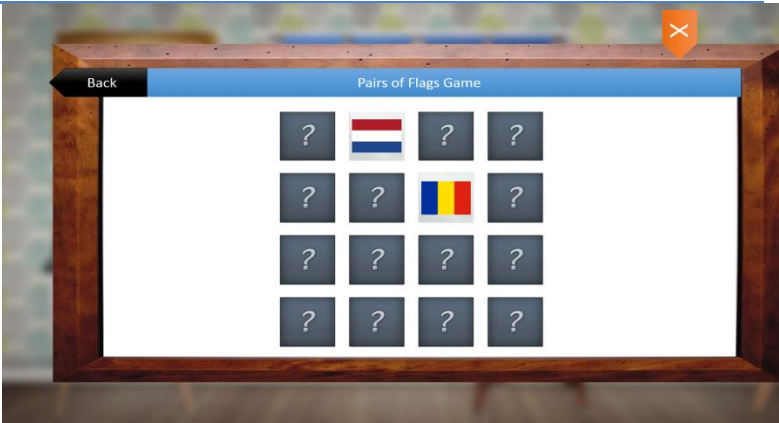
Task 8: Game #2
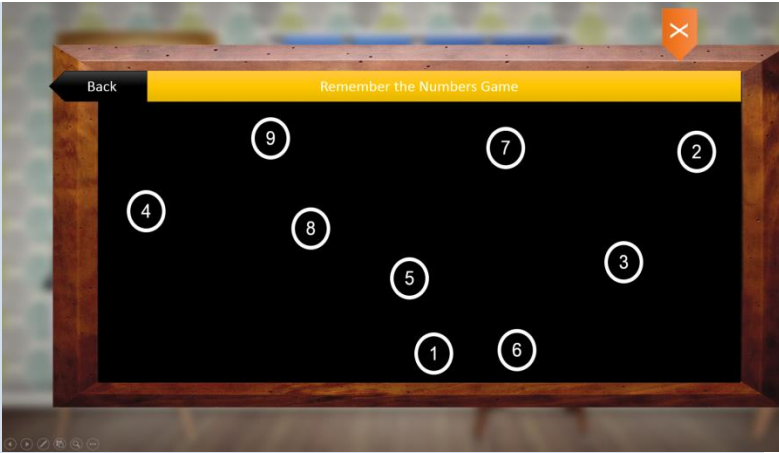
Task 8: Game #3
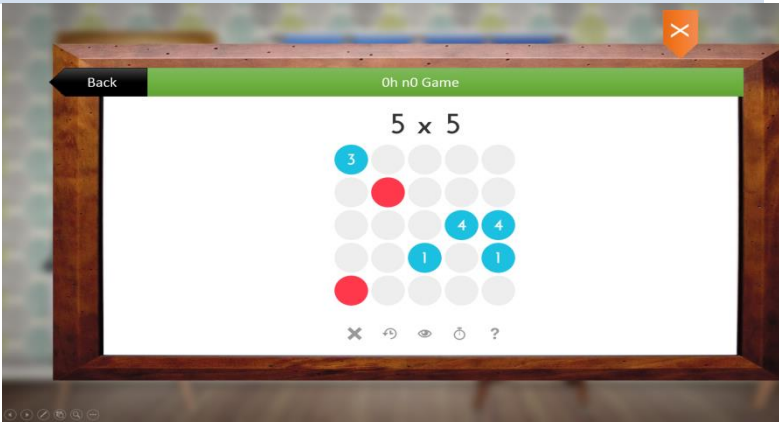
| **Task 9: Physical Training** | |
| **Task 10: Physical Training Rating** | |
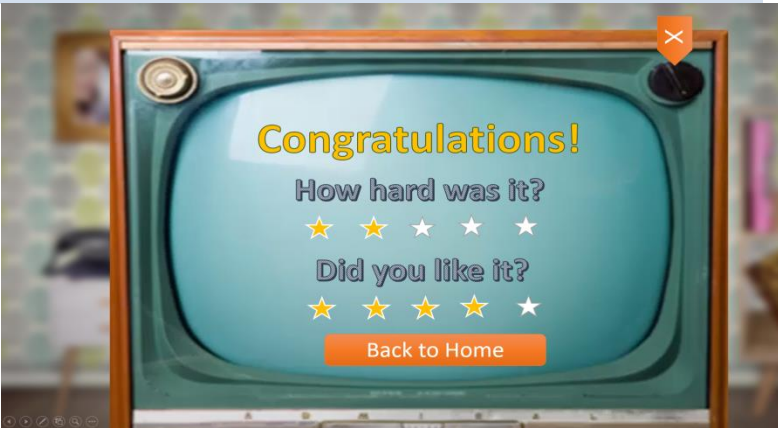
Table 12: Demonstration of screens in the SSE evaluations.

6.3 Participant description and results from Austria

6.3.1 Participants

Of the 6 participants recruited for the study, four (66%) were female. The small cohort of 6 participants consisted of 3 UX experts - all recruited from AIT Technology Experience department, one MCI experts/secondary user and 2 COPD experts (Table 13). The MCI expert is a clinical and health psychologist with a master's degree in psychology, the male COPD expert is leader of a self-help group in Vienna and qualified trainer for patient education and the female COPD expert is leader of a self-help group in Lower Austria. All target relevant experts were recruited by an email inviting them to participate in the evaluation.

#	Gender	Age	Role
1	F	37	UX expert
2	F	29	UX expert
3	M	30	UX expert
4	F	31	MCI expert/ second. User
5	F	41	COPD expert
6	M	63	COPD expert

Table 13: The demographics of the Austrian participants in the first Small Scale Evaluations.

6.3.2 Results

Task 1: Main Screen

The first impression of all experts was quite positive. The design of the interface demonstrated that it was not clear enough for all experts which objects were active and which not. Another aspect of the design was the subjective feeling about attractiveness - not all experts liked the design of the 60s and the usability experts recommended a customization of the interface design. Most objects were interpreted in the right way, nevertheless the television - "watching TV"- and the book - "reading a book" - were misinterpreted. Finally labeling of active objects and customization of the interface design were the main recommendations of this task.

Task 2: Personal data

The pink book as a symbol for personal data was not very convincing for the experts. Further the page numbers were found to be irritating since the different services were linked anyway with the headlines.

Task 3: Daily functioning monitoring

The amount of information presented in this service app was estimated to be too high in order to guarantee a structured overview and ease of use. Particularly the separation of the information into two columns was not interpreted logically. Some experts recommended to have one column for the time interval and a second column for the specific activity. Both COPD experts concluded that this service would be not very interesting for COPD patients, especially if the patients were not living a very active lifestyle because of their disease. The MCI expert found the service interesting with regards to alarms if activities were out of average such as leaving the house for longer than usually - but not necessarily for primary users but for secondary users in order to have an overview about the daily activities. Overall the experts wondered if the detailed amount of information would be attractive for the primary users and if this service would have an influence on fears about being monitored.

Task 4: Daily activity monitoring
Usability experts had no major recommendations for change and found the overview easy to understand. The COPD experts had concerns about the "running" category since COPD usually do not go running and probably the majority of older adults do not chose running as their favorite physical activity. Furthermore some experts worried if passive users could be motivated with this graphical overview and if more motivational aspects should be stressed.

Task 5: Sleep monitoring
The majority of experts missed the weekly and monthly overview like in the daily functioning and activity monitoring service. Moreover, some experts wondered how the system knew if the user was sleeping since it could also be possible that the user was laying in the bed and watching TV or reading a book. The COPD experts agreed that information about sleep quality and quantity is a very important factor in COPD patients concerning overall health and possible infects which highly constrain COPD patients. Finally, a major issue was that the sleep monitoring service should particularly show the user and overview about sleeping time with and without interruptions.

Task 6: Sleep rating
The experts quite concluded that the amount of questions rating sleep quality is too high. Besides there are some information asked which the eWALL system monitors anyway (e.g. sleep interruption). Once COPD expert suggested including a question about productive and non-productive coughing. Another COPD expert stressed the importance of specific pain symptoms which should be possible to describe as well as the need for oxygen during the night.
Final conclusion was to minimize the amount of rating questions in order to increase user motivation for rating the sleep.

Task 7: Cognitive Training
The experts liked the design of the cognitive training service, only one COPD expert had not recognized the chess board as an active object. Nevertheless the specific games object were hardly interpreted correctly which again stressed the importance for labeling the objects.

Task 8: Games
After going through the different games and trying out some experts doubted if the users would play the games in a standing position. One COPD expert remarked that the games would be an attractive opportunity to get familiar with the eWALL system and the touch use. In order to increase the motivation to play the games well-structured manuals and different levels of difficulty would be necessary.

Task 9: Physical Training
Especially the usability experts noticed that the training video should be shown full screen size. Further, the video should repeat as long as the user is doing the exercise in order to guarantee that the user is not doing the exercise alone. The majority of experts stressed the importance of an individualized training plan.

Task 10: Physical Training Rating

The experts agreed that rating each training video would annoy the users and disturb the training session. One COPD expert proposed to rate the videos only once. Further, some experts doubted if the user rating would be a valid instrument for generating an activating trainings plan: probably the users would rate demanding videos more negatively and consequently this would decrease their training outcome. One COPD expert asked how the system would know if the user had completed the exercise, which would be an important information as well.

Task 11: General questions
Overall the experts found the eWALL system very interesting and were especially enthusiastic about the big screen size of the system and the physical training service. One COPD expert asked if the system would be affordable for the user since affluent users would have anyway good possibilities for therapy. While the usability experts were quite positive about the learnability of the system for the users the COPD experts were more unsure and recommended a specific tutorial and written manual to increase use of the eWALL system and minimize fears and constraints of the users. One COPD expert even thought that MCI users would not be able to use the system actively and recommended to involve the informal caregivers since these have a major influence on the primary users for using the system.

6.4 Participant description and results from Denmark

6.4.1 Participants

Five experts in the field of usability and human computer interaction, and two work domain professionals participated in the participatory heuristic evaluation. Four out of the five usability experts were men, corresponding to 80% of the total number of usability experts. One out of the five usability experts was a woman, corresponding to 20% of the total number of usability experts. Both work domain professionals were women. The usability experts and work domain professionals were recruited by the Department of Health Science and Technology from Aalborg University. The same usability experts and work domain professionals had participated and evaluated the prior eWALL interface application, which gave them a reliable foundation of evaluating the redesigned eWALL interface application. All of the usability experts and work domain professionals had prior knowledge in the field of performing participatory heuristic evaluations of medical informatic systems. The usability experts all had a background as MSc in Biomedical Engineering and a PhD title. Furthermore, two of the usability experts also were Associated Professors. Both of the work domain professionals had a MSc in Clinical Science and Technology, and were on going PhD fellows.

From Table 14 the gender, age, and role is visualised over the usability experts and work domain professionals who participated in the Participatory Heuristic Evaluation.

#	Gender	Background	Role
1	M	MSc in Biomedical Engineering	Usability expert
2	M	MSc in Biomedical Engineering	Usability expert
3	M	MSc in Biomedical Engineering	Usability expert
4	M	MSc in Biomedical Engineering	Usability expert
5	F	MSc in Biomedical Engineering	Usability expert
6	F	MSc in Clinical Science and Technology	Work domain professional
7	F	MSc in Clinical Science and Technology	Work domain professional

Table 14: The demographics of Danish participants in the M22 prototype evaluation.

6.4.2 Results

The number of times each heuristic is used among the usability experts and the work domain professionals is visualised in Figure 25: Heuristics outcomes from first SSE in Denmark. From the first participatory heuristic evaluation performed in Denmark, the evaluation partners (usability experts and work domain professionals) identified 384 usability issues. In the second participatory heuristic evaluation, they identified 371 usability issues. 13 fewer usability issues identified in the second round compared with the first round. When looking at the distribution of the use of the different 15 heuristics among the evaluation partners, a small change has occurred. In the following section, this is showed.

Heuristic no. 1, *System Status*, was used 8 times (2,08%) in the first PHE and 21 times (5,66%) in the second PHE. Heuristic no. 2, *Task Sequencing*, was used 7 times (1, 82%) in the first PHE and 20 times (5,39%) in the second PHE. Heuristic no. 3, *Emergency Exits*, was used 7 (1,82%) in the first PHE and 10 times (2,7%) in the second PHE. Heuristic no. 4, *Flexibility and Efficiency of Use* was used 33 times (8,59%) in the first PHE and 28 times (0,78%) in the second PHE. Heuristic no. 5, *Match Between System and Real World*, was used 58 times (15,1%) in the first PHE, 49 times (13,21%) in the second PHE. Heuristic no. 6, *Consistency and Standards*, was used 50 times (13,02%)in the first PHE and 52 times (14,02%) in the second PHE. Heuristic no. 7, *Recognition rather than Recall*, was used 56 times (14,58%) in the first PHE and 32 times (8,63%) in the second PHE. Heuristic no. 8, *Aesthetic and Minimalist Design*, was used 72 times (18,75%) in the first PHE and 58 times (15,63%) in the second PHE. Heuristic no. 9, *Help and Documentation*, was used 27 times (7,03%) in the first PHE and 27 times (7,28%) in the second PHE. Heuristic no. 10, *Error Prevention*, was used 9 times (2,34%) in the first PHE and 11 times (2,96%) in the second PHE. Heuristic no. 11, *Skills*, was used 5 times (1,30%) in the first PHE and 8 times (2,16%) in the

second PHE. Heuristic no. 12, *Pleasurable and Respectful Interaction with the User*, was used 38 times (9,9%) in the first PHE and 33 times (8,89%) in the second PHE. Heuristic no. 13, *Quality Work*, was used 6 times (1,56%) in the first PHE and 9 times (2,43%) in the second PHE and finally heuristic no. 15, *Privacy*, was used 4 times (1,04%) in the first PHE and 3 times (0,81%) in the second PHE.

When looking at the distribution of the use of different heuristics among the evaluation partners from the first to the second evaluation round, it appears that the evaluation partners have increased their use of the heuristics: *System Status* and *Task Sequencing*, and has decreased their use of the heuristic: *Recognition rather than recall.*

The more use of *System Status and Task Sequencing* may corresponds to that the evaluation partners find that the redesign of the new main screen interface has a lack of easiness compared with the old main screen. The icons on the old main screen were 'homemade' and the layout was not close to finish. The redesign of the main screen contains recognisable icons and looks more like a real system. This positive development of the layout may be in correlation to the less use of the heuristic *Recognition rather than recall*. From Figure 25 it appears that the heuristic: *Aesthetic and Minimalist Design* still is the one that is used most often. This may indicate that the main screen interface still needs further rounds of iterations before we deliver the system to the COPD patients and MCI patients.
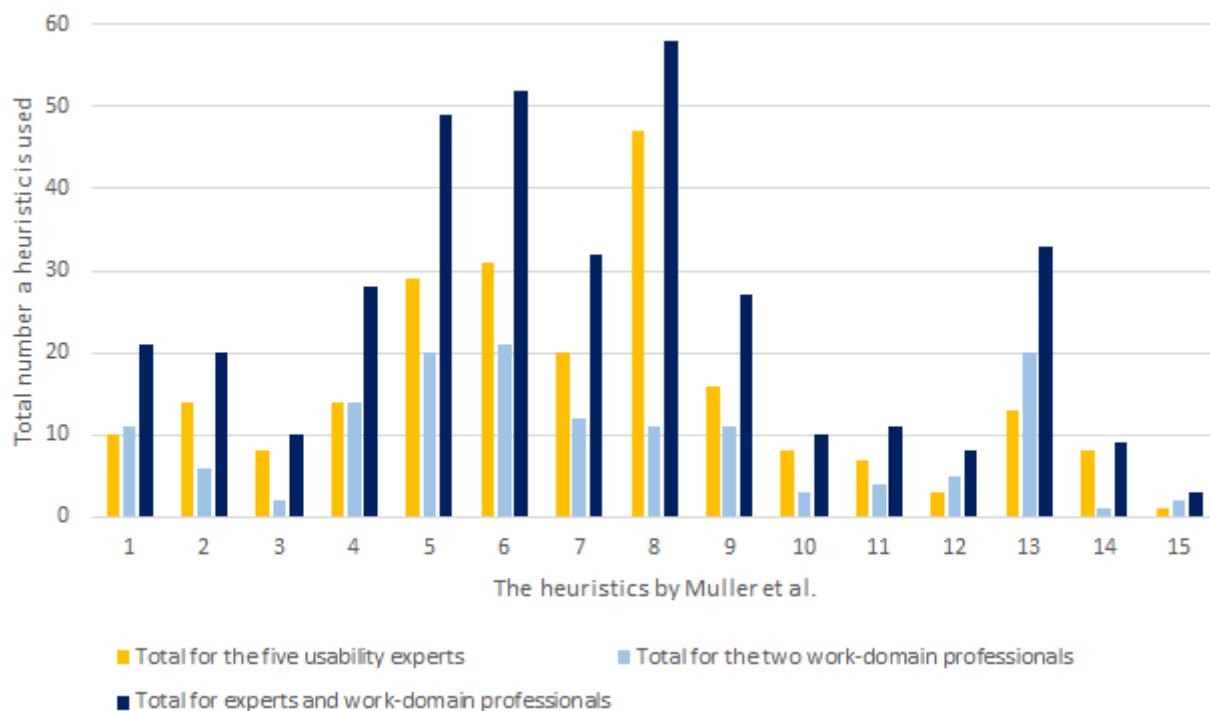
Figure 25: Heuristics outcomes from first SSE in Denmark.

The total number of times each severity grade (from 1-4) was used by the evaluation partners is visualised in Figure 26 below. In the first round, both evaluation partners used the severity gradeno. 3, *major usability problem,* most frequently. In the second round, the work domain professionals still used *major usability problem* most, but the usability experts used *Minor usability problem*

more often compared with the first round. The high use of the severity grade *major usability problem* indicates that the system is not yet ready to be delivered to the end-users such as COPD patients and MCI patients.
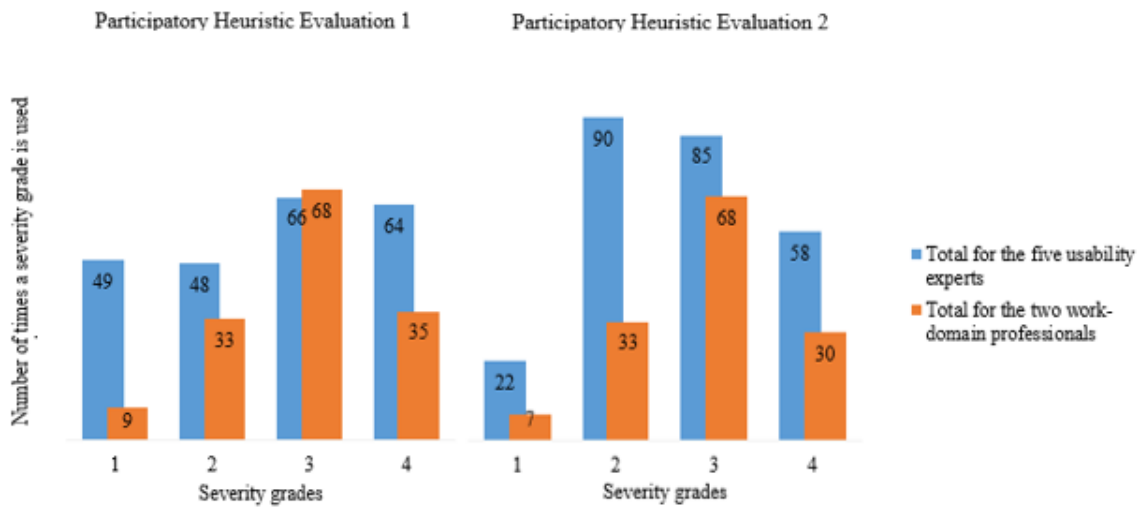
Figure 26: Severity grades of PHE.

Besides this second participatory heuristic evaluation performed in Aalborg University, we are ready to perform a user test with real users. But we will first perform that test when the system is frozen - meaning that the system we perform user test on, is the system we give the users in the large scale demonstration.

6.5 Participant description and results from Italy

6.5.1 Participants

The cohort consisted of 10 participants. Six MCI subjects and four expert recruited from the staff of I.R.C.C.S. San Raffaele Pisana (Table 15). In particular, the expert group consisted in one tele-rehabilitation expert, one Health Technology Assessment (HTA) expert, one Clinical Psychologist, and 1 Neuropsychologist. Three of four participants recruited for the study were females. The Psychologist and the Neuropsychologist were clinical and health experts with a master of sciences in clinical psychology and about 10 years of experience with patients with MCI. The Tele-rehabilitation expert was an engineer with ten years of experience in rehabilitation and ICT applications in the clinical field. The HTA expert was a bioengineer with five years' experience in clinical engineering. All experts received an email with invitation to participate at the evaluation. Gender, age and role of each participant are showed on table 15.

#	Gender	Age	Role
1	F	31	Clinical Psychologist
2	F	30	HTA Expert
3	M	38	Tele-Rehabilitation Expert
4	F	33	Neuropsychologist
5	M	78	MCI subject
6	M	66	MCI subject
7	M	74	MCI subject
8	F	72	MCI subject
9	F	65	MCI subject
10	F	77	MCI subject

Table 15: The demographics of the Italian participants in the M22 prototype evaluation.

6.5.2 Results

All clinical experts and MCI participants had a good overall impression of the platform. Experts highlighted the potential efficacy of the eWALL in terms of functional monitoring of the patient's daily activities. Although patients were accustomed to the use of software for cognitive training conducted at the IRCCS San Raffaele Pisana Day Hospital, all the MCI participants initially showed fear to use the technological tool (PC with a touch screen), saying that they don't know well how to use such a platform. However, after experimenter' instructions all patients had a good first impression about usability and the usefulness of the platform and five of them have offered to participate in the demonstration phase. Experts and MCI subjects have raised many issues, all these issues were gathered, categorized and summarized as follows:

Expert reviewers

Main Screen and Personal Data:
- It was not clear for the patients (especially elderly or MCI) if some images were buttons or static pictures (it could be explained, but if the system has not an intuitive UI it could be hard to remember. Interactive buttons could be pulsing or highlighted when tapping one time as we have no mouse);
- It would be practical to add a button with a sort of "personal assistant" to remember the aim of the platform and how to use the functionality in the page;
- The position of the TV set is too prominent, activities buttons should be more in the foreground (The risk is that participants will use more the video training functionalities and less the others);

- It could be useful if caregivers could remotely change family photo and window background;
- Something moving in the main screen could give the impression of an active system (for example a puppy or a fish bowl, and subjects could also feed or pet. Just thinking loud.);
- Books on the shelf are too thin. It is possible to enlarge them to allow greater usability;
- The icon seems more a bookmark for not expert people (if is possible, write " close " instead of X in the icon);
- Numbers may results confounding (it should be better to have an "open" with icon or text).

Daily Functioning Monitoring:
- If the screen is big (more than 32') fonts are ok;
- Have to use a more informal language. Better avoid sentences that can imply the control of the life: "we detected" or "we noticed" Seems something in "STASI style".
- It could be useful to link this information to the wellbeing of the participant using motivational messages.

Daily Activity Monitoring:
- Messages seems a bit confusing;
- Seems too scientific for the use with elderly people;
- A motivational message to summarize the screen could be useful;
- A small icon near the axes could ameliorate the comprehension of the graph (for examples a small watch and 2 feet);
- It seems something is missing on the right of the graph;
- Colours are too similar;
- More than kilometres is better to use meters (Elderly cannot walk for kilometres, to read numbers under the unit can be frustrating);
- The graph is not very clear for the patients. If possible use histograms.

Sleep application:
- Nice! It could be useful to link these information to the wellbeing of the participants using motivational messages (for example: "you should sleep 7 hours");
- It could be nice to add some tips (for example: "try to read a book before you go to sleep" "It is better not to watch too much TV late at night";
- The graph is not very clear for the patients. If possible use histograms.
- If possible, use icons or emoticons with emotions to have a quality index of the mood of the patient (for example: "I'm angry", "I'm sad", "I'm tired", "I'm happy" etc.).
- There have been difficulties in the selection of books because the icons do not correspond to the mouse pointer.
- We found difficulties in the selection of books because the icons do not correspond to the mouse pointer.

Rating Sleep:
- Questions are right for a good comprehension of sleep disorders.
- Snoring people often don't know they are snoring or they don't mind;

- "Had bad dreams?" The answer "don't remember" should be taken in consideration;
- Please insert a "finish" and a "skip" buttons;
- To write text on a big touch screen is something elderly are not comfortable.

Cognitive Training:
- Draughts, backgammon, domino, dice. None of these games is implemented.
- A button with a personal assistant could be really useful, his aim should change with respect of the functionality chosen (for example he could give tips related to the functionality or explain the rules of the games etc.);
- Clarify and repeat the instructions of the games especially for MCI.

Playing games:
- Elderly people will not find comfortable to play while standing;
- With no motivational message games will be a bit boring;
- Please insert an "instructions" button; at the opening of the icon, the instructions of some games are missing (such as for the game tests of memory and memory card). In addition, since the patient may forget the instructions you could add a small window/icon (top right), the patient can open by pausing the game to remember the instructions. Always remind MCI patient what to do.
- Write in Italian the name of the game. None of our patients knows English. (For example: Memory games etc.).
- A "pause" button could be useful during the gaming (less anxiety);
- Add a different level of difficulty (for the motivation).
- Memory test Game: This game allows you to make a workout numerical and visual-spatial working memory. The visual-spatial sequence of the game is long (a normal person has a span between 5-8) and often the patients have not the same span of a normal adult and can't remember the sequence of numbers associated with the location of the balls. When the patient makes a mistake, he cannot repeat the same sequence, because the game moves immediately to another different sequence. The patient may feel frustrated and give up this kind of game.

Physical Training:
- If we will use a 42" screen maybe the text will be giant and not easy readable;
- A "pause" button could be useful during the video training;
- Remove sounds (for attention deficits);
- The X button is not so nice; is more easily understood an "On-off" button.
- Picture in picture makes the design confounding.

Rating the training:
- The level of fatigue could be a useful index (more than the "like" rating).

E-Wall in general:
- Taken in consideration the possibility to have a "personal assistant" button. MCI people often need the instruction and the training to be repeated several time;

- Please avoid terms and phrases that could instil in participants the doubt or the feeling of being followed or monitored. (No STASI or KGB speaking, please);
- Use the Italian language for all activities (even in calendar etc.).

MCI subjects

MCI subjects performed a task-based evaluation, results are summarized as follows:

Task 1-Main Screen-:

- All patients would be happy to have a large screen at home. Some of the patients have asked if they could also watch TV with the screen or listen to music. 5 of 6 patients, prefer to be able to sit in front of the screen (not standing), the remaining one did not show preferences between sitting and standing.
- All patients had a good impression of the platform. Patients liked the possibility to customize the wallpaper; liked the window in the middle of the screen for a sense of freedom, and the family environment (photo frame);
- All participants understood the function behind clock, book, TV, game board and of the platform;
- Some participants asked if it was possible to change the image in the picture.

Task 2-Personal Data-:

- All participants understood the functionalities of the book and icons.

Task 3-Daily Functioning Monitoring:

- All participants had a positive general impression of the application, they can see it as a personal diary about daily activities, none of them was able to imagine the actual operation of the application when connected to wearable sensors.
- The patients imagine sharing this data with family;
- 2 of 6 patients say they don't see well what is written, and they asked if it is possible to enlarge text font.

Task 4-Daily Activity Monitoring-:

- All participants had a good first impression of the application;
- Initially participants showed difficulty interpreting the chart. One of them claimed he does not need such accounting functionality, because he not walks very well. Participants would share this data with family and doctors;
- The privacy was not a problem for DFM and DPAM applications.

Task 5- -Sleep Monitoring-:

- 4 of 6 patients had a good first impression of the application. 2 of them claim that they sleep well and not need it;
- Some participants say that the graph is not clear.
- All participants imagine sharing this data with the family and the doctor.
- Participants had a good first impression of the use of the "stars" for evaluation, but don't know if they should use the application every day.

Task 6- Sleep Rating-:
- Some patients say that the questionnaire was boring and that they should not be complete them every day;
- They prefer to rate with stars.

Task 7-Cognitive Training-:
- All participants had a good first impression of the games. Enjoy playing;
- Some participants already performed Cognitive Stimulation with "games" during their therapeutics program, so they knew what to do.
- 1 of 6 participants do not like to play;
- They like to "play" on the big screen.

Task 8-Playing Games:
- Some patients did not like this kind of games because they are too simple compared to games they know (on the phone, at home or in therapy);
- 2 of 6 participants failed to play (they did not remember the right number sequences), they proposed to change the rules and let them try with the same sequences until they can remember.
- All participants liked playing these games on a screen in their home.

Task 9 –Physical Training-:
- All patients liked this kind of app and said it is very useful;
- All participants consider this kind of training too simple and repetitive;
- They like to do "gymnastic" at home (as a physiotherapy);
- 2 of them smiled when they saw the lady (they said that they are not as old as her).

Task 10-Rating the Training-:
- The participants had a good first impression of the app. No problem during the evaluation of their experience;
- The rating of the training is a quick and easy task for all participants.

Task 11-General Experience-:
- Patients consider eWALL easy to learn and they would use eWALL at home. All participants think that they can learn to use the platform, five of them are curious to try if this can really ameliorate eventually the quality of their lives. Only one of them is not excited about this technology. Three of them are also happy to be monitored and remotely controlled by doctors.

6.6 Participant description and results from the Netherlands

In the Netherlands two evaluations in different points of time were conducted. The first evaluation is described in this chapter. The second evaluation will follow in chapter 6.3.7.

6.6.1 Participants (first round of M22 evaluation)

A total of 8 elderly users were recruited from the Enschede area in the Netherlands: 5 male and 3 female, aged between 66 and 88. Table 16 lists the users included with the first of M22 evaluations.

#	Gender	Age	Role
1	M	88	ARI
2	F	67	ARI
3	M	77	ARI
4	F	68	ARI
5	M	67	ARI
6	M	66	ARI
7	F	75	ARI
8	M	74	ARI

Table 16: The demographics of the Dutch participants in the M22 prototype evaluation.

6.6.2 Results (first round of M22 evaluation)

In the Netherlands, 8 ARI elderly have participated to the user walkthrough tests. They also filled in the UEQ questionnaire. Similarly to the 6.2 outcomes, the participants had a good first impression of the proposed technology. All participants understood Bob's story and the reason for choosing eWALL. Associating common object representation with a representative functionality proved to be a better solution. Participants could easily associate the book with the monitoring functionality when explaining that "eWALL writes Bob's like events in a book". Several redesign recommendations were made out of which the most relevant are:

- the pictures should show family (especially grandchildren).
- the phone is too old fashioned and should be removed.
- window should explicitly show the weather information.
- the clock could show upcoming appointments, medicine schedule, or meal times.

All participants understood that function of the clock, book, TV, game board, but when it came to the picture outside the window opinions varied between being able to set up favourite landscape to

viewing pictures from past vacations. No participant intuitively deduced it is an abstraction of the weather quality. We believe this is because the depicted landscape does not resemble with any natural landscape in the Netherlands.

Daily Functioning Monitoring (DFM):
Participants understood what the information in the application, yet 30% considered it to be inappropriate when asked if they would use it. Most participants could not understand the goal of this application.

Daily Physical Activity Monitoring (DPAM):
Participants focused on the chart part of this application and understood and agreed with the colour coding of activities. As we expected this minor learning curve was overcome immediately with the researcher's explanation. All participants considered this application adds value to the system, 2 of them claimed they do not need such accounting functionality, because they are sportive persons. Participants would share this data with family and doctors. The privacy did not present a concern as for the previous DFM application

Sleep Monitoring:
This application perceived positively by most participants (6/8). The major concerns about it were the privacy aspects (2 elderly do not feel comfortable with the idea of being monitored while sleeping, but the rest would share this information with close family and doctors) and daily use (a majority of participants would not use this application every day, only occasionally, claiming that they feel when they sleep bad, or good.). Interface improvements were suggested:
- make the sleep overview shorter and more concise (add a summary of last night awakenings and sleep periods instead of displaying them one by one)
- make the list centred and avoid the "zig-zag" display
- make the list viewable in one screen (avoid the need for information discovery through swiping)
- make the questionnaire as short as possible, because it will not be answered every day

The new functionalities presented in this prototype are the video exercise trainer and the cognitive games.

All participants consider this functionality useful stating that it is a useful functionality when they cannot go outside (this happens quite often in the Netherlands due to weather conditions). 4 participants would recommend this to other persons, 2 would use this every day.
Some redesign requirements:
- the video are should be bigger
- the video difficulty should be increased (2 participants claimed that the presented routine is too easy)
- the TV buttons on the left margin should be functional, or not displayed

Rating each exercise was considered and easy task and elderly would provide this information. They would trust eWALL to take these ratings into consideration when recommending future routines. 1 elderly suggested that it would be useful to share these ratings with the doctor.

Cognitive Games:

The cognitive games were not perceived as training cognition, but considered regular games. Most elderly considered the game quality low (some stating that they have better games on their phones/tables, others do not like to play such games in general). 2 elderly were enthusiastic to play games on such big screen.

eWALL in general:

Participants consider eWALL easy to learn, and they would use this version of eWALL. 80% are confided they can achieve a good command of the technology. Figure 27 below shows the participant experience rating.
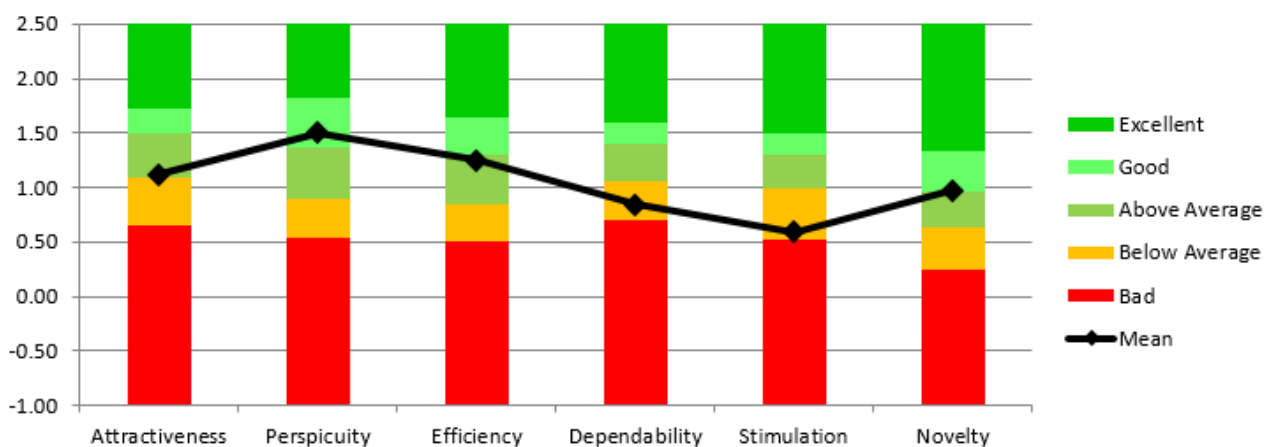
Figure 27: Participant experience rating.

We will present these outcomes in comparison with the previous user experience test. As the table shows, the new eWALL metaphor provides and above average score in attractiveness (similar to the previous prototype) and good perspicuity score (as opposed to the bad score from the previous prototype). This means that elderly consider the new metaphor more intuitive, motivating that the correlation between familiar objects and system functionality is clear and easy to understand. A much higher score in efficiency was achieved and participants did not express the feeling of emptiness (in previous tests the emptiness of the screen was correlated with the feeling of reduced eWALL functionality). As this test consisted of a user walkthrough, the dependability and stimulation scores are low, because participants did not operate with the system themselves. The novelty perception scored above average (higher than the average in the previous prototype).

6.7 Second round of M22 Prototype Evaluation in the Netherlands

6.7.1 Goal

The eWALL concept and prototype was tested with patients with a lung disease (mainly COPD) to assess its worth for patients and their context. As patients are likely to have different needs and

wishes than general, healthy older adults, this evaluation is an important addition to the evaluations conducted thus far. In particular, this evaluation served three goals:

1. Gauging the added value of the eWALL concept for patients with a lung disease
2. Gauging the understandability of the eWALL concept and metaphors for patients with a lung disease
3. Identifying strong points and weak points of the eWALL concept and prototype for patients with a lung disease.

Ultimately, this evaluation will lead to redesign advice for eWALL.

6.7.2 Method

We conducted focus groups with patients from a physical therapy practice in Enschede, the Netherlands (Figure 28). Following two group training sessions, interested patients could join a focus group. These focus groups used the following setup:

1. Introduction – setting out the goals of the focus group, explaining participant anonymity, and obtaining informed consent.
2. Getting to know each other – each participant introduced him/herself and explained how their lung disease is affecting their life.
3. First impression – eWALL was shortly introduced and then shown on the designated eWALL device (42" touch screen); participants were asked to give their first impressions.
4. COPD functionality – the eWALL COPD functionality was demonstrated and participants were asked to discuss their usefulness and points for improvement.
5. Best and worst aspects – the participants were given print-outs of the main interface and asked to stick red stickers (dots) on aspects they disliked and green stickers on aspects they liked.
6. Acceptance – The participants were asked to state whether they would like to have aneWALL in their own home, should the opportunity to buy or acquire one arise; they were also asked to give a rationale for their decision.
7. Wrap-up – We closed the focus group by summarizing the main findings, explaining our follow-up steps, and thanking the participants for their contribution.

The focus groups were audio-recorded. After the focus groups, these recordings were analysed on a per-question basis, using inductive thematic analysis [15]. For each predefined question that was posed, similar answers were grouped and we determined whether there was no agreement among the participants, or whether some, or half, or (almost) all participants gave the same answer.
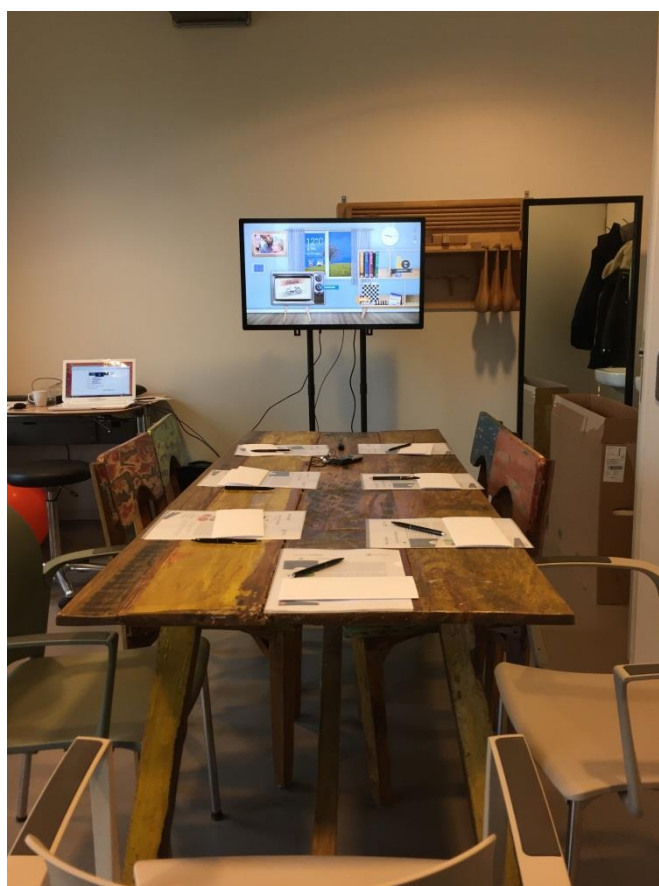
Figure 28: Setup of the location for the focus group with patients with a lung disease.

6.7.3 Participants

Our two focus groups resulted in a total group of 14 patients with a lung disease that gave their opinion on the eWALL concept and prototype. Table 17 and Table 18display their demographics. In the first focus group, a physical therapist attended as well. The participants hobbies included biking (on an electronic bike), walking, crafting, painting, gardening, reading, watching football, making furniture, working on cars, and making music. Their illness affected their hobbies in different ways: reduced stamina, and influences of weather and season (due to pollen in the air).

#	Gender	Age	Lung disease	Has lung disease since
1	M	77	COPD	30 years
2	F	26	Asthma	20 years

3	F	71	Asthma and COPD	Asthma since 65 years; could not recall when COPD began
4	M	70	Shortness of breath, multiple lung diseases	2 years
5	F	65	COPD	20 years
6	M	60	COPD	2 years
7	M	72	COPD, history of lung cancer	At least 25 years

Table 17: Patient demographics of focus group 1.

#	Gender	Age	Lung disease	Has lung disease since
1	M	58	COPD	2 months
2	F	49	Lung cancer	2 years
3	F	61	COPD	4 years
4	F	74	COPD	1 year
5	M	79	COPD	20 years
6	M	86	COPD	10 years
7	M	66	COPD	For a long time, could not recall

Table 18: Patient demographics of focus group 2.

6.7.4 Results

First impressions

When the participants first saw eWALL, they were drawn towards very different elements: the weather forecast, the thermostat, the step counter, the photo frame, the books, games (which some participants said they played themselves, while others did not), and the clock.

One participant stated that he would like to have an eWALL immediately. When asked why, he responded that he was drawn towards the books and the games. Those were things he could still do, given his lung disease, and could therefore give purpose to his days.

Initially, some participants thought that these books were meant for them to complete. In other words, they should be the persons writing down the number of steps they had taken during the day, instead of sensor technology automatically providing this input.

Multiple participants immediately asked whether or not this could also be provided via their own television or tablet, as they thought that eWALL was too big.

COPD functionality

Lava lamp / Step counter. After explaining the functionality and showing the Fitbit sensor, many participants commented that this was a useful function. Some participants were worried about this technology instructing them exactly what to do (e.g., do more steps), what they would not appreciate, or did not feel they needed a technology for telling them how active they were, or would feel demotivated after being confronted with the low amount of steps they were taking.

My activity book. After showing the "My Activity" book, participants asked whether or not goals would be personalized and geared towards their personal capacity. They would prefer a situation in which they could use the sensor for some time to set a personal baseline for the number of steps they can take daily. This should also take weather conditions into account, as these conditions influence the number of steps they can take on a day. Next, other disabilities also played a role (one participant explained that he had a worn hip and walks with a cane, and can therefore not walk that much), as well as the progress patients make while improving their stamina. At this point, the physical therapist stated that personal step goals could be best set by a physical therapist. Finally, the participants discussed that the number of steps they make each day can also be influenced by the day of the week (e.g., one participant worked and then made more steps; walking clubs that go out walking on a set day in the week).

My sleep book. The participants were shown the "My Sleep" book and were explained how the Beddit sensor works. The participants of the first focus group showed great interest in sleep monitoring. They were very interested in their sleep pattern over a longer period of time, as they suspected that their sleep was not optimal but had no way of finding this out for sure.

My health book. The participants were explained how health parameters (heartrate, oxygen saturation, and blood pressure) could be assessed, and incorporated and shown in eWALL. During the first focus group, participants showed interest, but also stated that one measurement a day would not tell them much. Instead, they were more interested in curves over the day. Furthermore, these measurements should be annotated so that they could also interpret what was happening (e.g., measuring heartrate after walking or while sitting on the couch). Other participants (and all participants in the second focus group) commented that they should like to know this information, but did not want to be occupied with measuring and viewing data all day as this would result in stress. After discussing this issue, most participants agreed that the measurements that were taken at the physical therapy office (two, three times a week) suffice.

Exercises. After explaining how exercises are shown and how an exercise schedule is created, participants asked whether or not the schedule could be programmed by hand. The participants receive exercises they should perform by their physical therapist and the eWALL exercise schedule

should comply with these exercises. The advantage they saw of this functionality is that it reminds them how they should execute exercises that were prescribed by their physical therapist. In all, the participants were positive about this functionality.

Best and worst aspects

After explaining all the functionality to the participants, we asked them to mark what they liked and disliked about eWALL by sticking red or green stickers (dots; green for positive, red for negative) on large sheets of paper on which the main interface was printed. See Figure 29 for an example. Figure 30shows a chart in which the times each eWALL functionality was positively and/or negatively valued are displayed. It shows that only the photo frame received more negative than positive values. Next, the exercise functionality (comprising of the Television and the 'gymnastics' button), the lava lamp/step counter, and the games received the most positive valuations.

Figure 29: Main eWALL interface with positive and negative evaluations (interface set to Dutch language).

Aspects of eWALL, valued positively and negatively
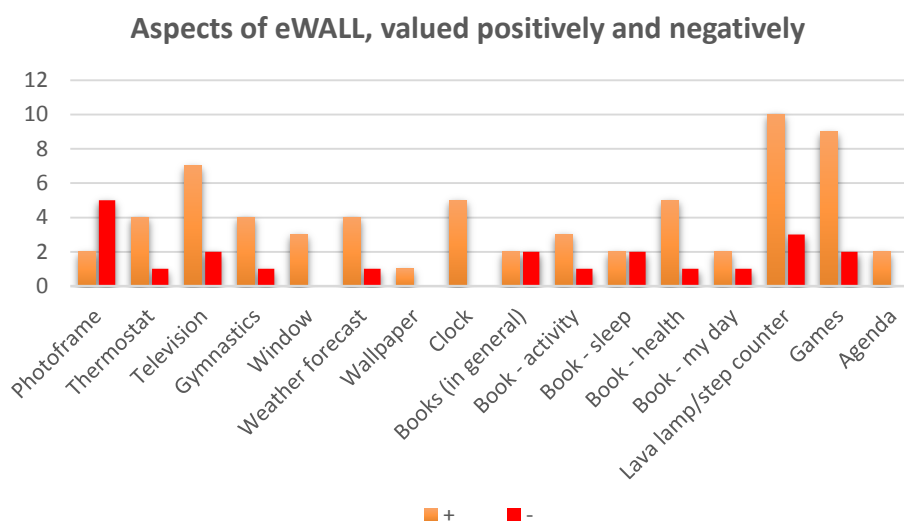
Figure 30: Times each eWALL functionality received a positive or negative valuation.

Acceptance

When asked whether or not the participants would like to have aneWALL, most participants said that they would not like to have it, as it would not fit into their homes. Instead, they expressed a very strong preference for a smaller version (on a Tablet PC or on their own television) as they thought this would be more practical, and they disliked the big screen as it would not fit with the style of their interior. Next, the participants preferred to have the option to turn eWALL off. They did not like having such a large device turned on in their homes all day.

Most participants agreed that they liked the concept and the functionalities. A few participants however, stated that they have enough motivation of their own to comply with their activity goals and do not need such a technology.

Finally, a discussion emerged on the sharing of data that is collected in eWALL. Most participants were fine with sharing this data with their healthcare professionals, but did state that their data were not to be used as a mechanism to check whether or not they complied with their goals.

6.7.5 Conclusions

In all, patients with COPD disliked the modality via which eWALL was provided (a big screen), while they did like the functionality that eWALL discloses. Therefore, the participants preferred to access eWALL via other means. There were several functionalities that were most valuable to patients with COPD. They are the daily activity monitoring, online exercising and games (the latter not from a 'health' point of view, but just because they like to play games).

In order to maximize the usefulness and acceptance of eWALL for patients with COPD, and based upon the two focus groups that we reported above, we have formulated the following redesign advice.

General redesign advice:
- The eWALL should be accessible via other modalities than the large screen alone, and should also run on a Tablet PC or on a WiFi TV.
- The eWALL should have the option to be turned off.
- The photo frame should allow users to upload a photo of their own and to disregard the connection to social media.

Redesign advice for Activity monitoring:
- The step goal for patients with COPD should be personalized on a daily basis and should take into account:
  - Their baseline condition;
  - the progress they make while improving their stamina;
  - disabilities besides COPD;
  - the daily weather conditions;
  - seasonal influences (e.g., pollen in the air);
  - a person's regular activities (e.g., work, participation in clubs (like walking clubs) with regular activities).

These daily step goals could be set by a physical therapist, or automatically.

Redesign advice for Monitoring health:
- The eWALL should be able to incorporate medical measurements (heartrate, blood pressure, oxygen saturation, weight) that are made at the physical therapy centre.
- The eWALL should allow users to measure their health parameters multiple times a day.
- The eWALL should allow users to annotate each measurement of a health parameter.

Redesign advice for Video Exercising:
- The eWALL should allow the physical therapist of a given user to create a personal exercise schedule, which is then disclosed to the user as a movie in the Exercise functionality.

6.8 Implications for the M25 Prototype

During the consortium meeting in Budva from 1[st]to 2nd of October 2015 the main implications from the M22 prototype evaluations were discussed. The discussion was based on previous weekly teleconferences (X-WP_d telco) and summarization of results in a Google document. A detailed summary of important issues and recommendations is listed in the Appendix III*"Summary of Small Scale Evaluation Results"*.Table 19 demonstrates the major interface changes from M22 to M25 prototype.

M22 Prototype	M25 Prototype
Main Screen	
	Changes • Interactive Elements are labelled • Customization will be not realized in the project • Calendar central service for MCI patients • Contact service will be removed
Daily Functioning Monitoring	
	Changes • Left and Right positioning of information is renounced • Less information • Interface style improved (font size, contrast) • Additional information: medicine intake, visits • Combination of daily information and emotional rating • Feeling of being monitored has to be reduced • User should only see relevant information
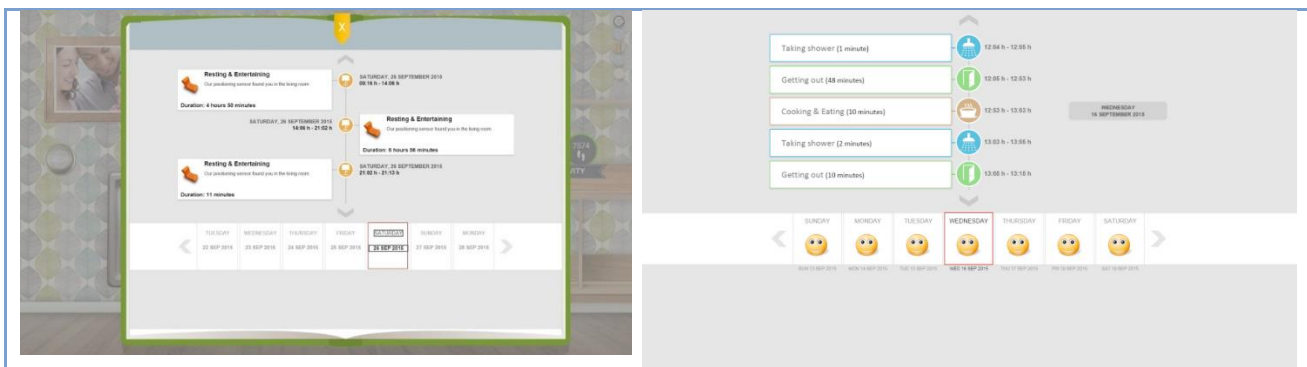
Daily Activity Monitoring

Changes
- Information overload will be reduced
- Only relevant information for user groups necessary to demonstrate (e.g. not running for COPD users)
- Exercises should be included in the bar diagrams
- Customization of activities
- Well-being adds will have to play a crucial role in motivation users to be more active

Sleep Monitoring

Changes
- Inconsistency of design compared to the DFM + DPAM
- Scrolling of days has to be included
- Left-Right Side reading will be replaced by a one column design

Sleep Rating
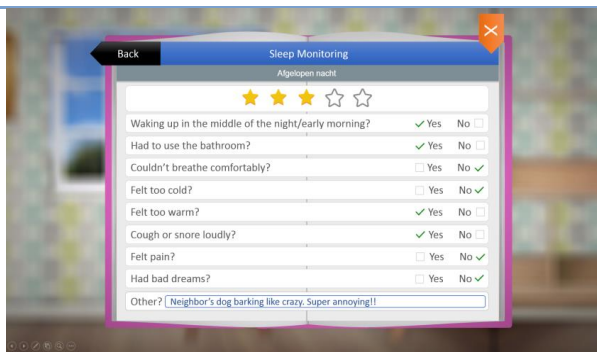
	Changes • Reduction of questions (some information is already monitored) • Personalization of questions will be possible • Stars will be replaced by another rating symbol • Touch screen pen for writing text on the touch screen

Games

	Changes • instructions of games → Personal assistant? • different levels of games • reward system for playing games and achieving points • inclusion of user's familiar games (Sudoku, playing cards, crossword puzzle) • COPD quiz • Explicit evaluation of standing position while playing games for the M25 prototype

Physical Training

Changes   • new design   • video will be enlarged   • inclusion of repeat/pause button   • display of time left for completing the exercise   • video will continue until user has finished the exercise

| **Physical Training Rating** ||

	Changes:
	• New design
	   • Feedback will be based on health measurements
	   • No rating after every exercise
	   • Symbol for rating will be replaced

Table 19: Overview of interface changes from M22 to M25 prototype.

7   Discussion and Conclusions

First evaluation studies results of the M12 prototype showed that participants of the target user groups were particularly interested in the health and activity monitoring functionalities of the eWALL system. Further the idea of video based exercise training was well accepted by the older adults involved. Nevertheless the majority of participants clearly refused to be monitored although they used to forget this aspect during the evaluation sessions. The large scale evaluation study will have to investigate whether this major aspect could be accepted by the target users and if benefits from the system would compensate this. Together with the experts' usability reviews the M12 prototype was mainly changed related to consistency issues, wording of the different services and headings, reducing information and general usability aspects.

Compared to the rather simple interface design of the M12 prototype, the next M19 prototype already focused more on attractiveness of the interface and the different eWALL applications were already more developed in terms of functionality and consistent design. In a first evaluation phase of the Circle 1+2 SSE older adults with age related impairments (ARI) participated in the Netherlands. These outcomes proved that eWALL is a promising technology and elderly would be motivated to use the system as a personal device if the information is presented more clearly and better structured.

Following the latest evaluations of the M22 prototype with experts and target users very detailed recommendations for prototype iteration were developed in the consortium. Despite from common design and usability changes some eWALL services will be adapted:
- More support for MCI patients
  - with a more assistive calendar service providing information about medicine intake, upcoming visits,
  - with the combination of calendar entries and emotions
- Reduction of relevant information in the Daily Functioning Monitoring
- Integration of a touch screen pen for writing information in the eWALL system
- Cognitive Training: better instructions and different levels for playing games

Due to some delay in the eWALL interface design two partners had to postpose the M19 prototype evaluation to the M22 prototype evaluation in order to get new feedback and not evaluate the "old" interface again. This did not interfere the aim of the Small Scale Evaluations.

The evaluation of the M25 prototype, which is the last part of the "Small Scale Evaluations" will start by the mid of November and results will be subsequently described in an updated version of the D6.4.

Bibliography

[1]   DeChant HK, Tohme WG, Mun SK, Hayes WS, Schulman KA. Health systems evaluation of telemedicine: a staged approach. Telemed J. 1996 Winter;2(4):303-12.

[2]   Jansen-Kosterink SM. The added value of telemedicine services for physical rehabilitation. Enschede2014.

[3]   Hobbs N, Dixon D, Johnston M, Howie K. Can the theory of planned behaviour predict the physical activity behaviour of individuals? Psychol Health. 2013;28(3):234-49.

[4]   Ryan M, Scott DA, Reeves C, Bate A, van Teijlingen ER, Russell EM, et al. Eliciting public preferences for healthcare: a systematic review of techniques. Health Technol Assess. 2001;5(5):1-186.

[5]   Ryan TP. Fractional Factorial Designs with Two Levels. Modern Experimental Design: John Wiley & Sons, Inc.; 2006. p. 169-247.

[6]   Kairy D, Lehoux P, Vincent C, Visintin M. A systematic review of clinical outcomes, clinical process, healthcare utilization and costs associated with telerehabilitation. DisabilRehabil. 2009;31(6):427-47.

[7]   Relton C, Torgerson D, O'Cathain A, Nicholl J. Rethinking pragmatic randomised controlled trials: introducing the "cohort multiple randomised controlled trial" design. BMJ. 2010;340:c1066.

[8]   Hailey D. The need for cost-effectiveness studies in telemedicine. J Telemed Telecare. 2005;11(8):379-83.

[9]   Zanaboni P, Wootton R. Adoption of telemedicine: from pilot stage to routine delivery. BMC Med Inform DecisMak. 2012;12:1.

[10]   Finstad, K. "The usability metric for user experience." Interacting with Computers 22.5(2010): 323-327.

[11]   Voss, K. E., Eric R. Spangenberg, and Bianca Grohmann. "Measuring the hedonic and utilitarian dimensions of consumer attitude." *Journal of marketing research* (2003): 310-320.

[12]   Agarwal, R, Prasad, J. A conceptual and operational definition of personal innovativeness in the domain of information technology, Information Systems Research 9 (2), 1998, pp. 204–21.

[13]   Muller et al., 1998

[14]   the User Experience Questionnaire, link: http://www.ueq-online.org/

[15]   Braun V, Clarke V: Using thematic analysis in psychology. Qualitative Research in Psychology 2006, 3(2):77-101.

Abbreviations

DFM	Daily Functioning Monitoring
DPAM	Daily Physical Activity Monitoring
HED/UT	Hedonic Utility Scale
SSE	Small Scale Evaluations
SUS	System Usability Scale
UMUX	Usability Metric for User Experience

Annex I: 1st lab evaluation protocol for primary and secondary end-users

The evaluation session can be carried out by one facilitator (minimum) who will be in charge of welcoming the participant, describing the scope of the project, asking questions, encouraging thinking-aloud and keeping notes of the participant's comments. It is however recommended - both for comfort but also for ensuring better coverage of participant's views - that 2 researchers conduct the study: 1 serving as the facilitator, 1 as the observer. In the room, except for the eWALL setup, there should be:

- Audio/video recorder placed behind participant's shoulder
- A laptop for entering information directly to a results form (not absolutely necessary)
- Pre-questionnaire, UMUX, HED/UT questionnaires
- Markers, pens in reach to the participant and blank paper, in case participants want to sketch ideas.

The checklist

Tasks	Check
1 WEEK BEFORE	
Explore the interface, try out all the scenarios proposed. It is necessary to have a first-hand experience with the system.	
Read and comment on the present evaluation protocol to: Angeliki.Angeletou@ait.ac.at.	
Book the evaluation facility.	
Recruit end-users for specific dates and times.	
Ask one of your colleagues to be available during the study as the observer.	
Prepare and send instructions for the participants how to find the facility.	
Print Informed Consent forms (consider some extra copies in case participants request to take it home).	
Print the questionnaires.	
Print this protocol.	
Arrange incentives for participants.	
3 HOURS BEFORE	
Check the prototype and the camera if they work.	
Make sure informed consent and questionnaires are in place.	
Rehearse the protocol.	

Facilitator tips

- ➢ Make small talk with the participant to warm her up.
- ➢ Explicitly encourage each participant to **think aloud** about their actions on the tasks given.
- ➢ Listen and do not make any judgments. Just **nod your head** to show that you understand her.
- ➢ If you think of more questions, go ahead and ask them. The more feedback, the better.
- ➢ Dig below top-of-the-mind answers. **Find out why and how.**

Timeline

Time (in min)	Activity
0-6	Introduction
6-9	Informed Consent
9-15	Pre-questionnaire (optional)
15-45	Discussions on Scenarios
45-48	UMUX (optional)
55-65	Post-interview
65-70	Conclusions, thanks, goodbye!

Introduction

The participants in most countries for this first evaluation are going to be secondary end-users ranging from carers, occupational therapists to bioengineers. Only in Austria, we will employ older adults for the evaluation. In the following scenarios, we propose a list of questions and small tasks to trigger discussion around features of eWALL. Obviously, the focus is different when the evaluation takes place with older users than with secondary users. Therefore all scenarios have 2 question lists marked as **a** and **b**.

With such a great variety of users, it doesn't make sense to apply questionnaires in this specific study. The goal is to gather qualitative input to update our design. However, in the timeline, we propose 2 questionnaires: the pre-questionnaire, which aims to measure the technology affinity of the participant and the UMUX, a standardized usability questionnaire.

The process takes about 1 hr and it starts with a participant entering the study room, where the touchscreen TV with the eWALL setup is prominently located. We (from now on, *we* refers to the facilitator and observer) welcome the participant and explain the process:

Welcome to our lab! We are participating in a large-scale European project called eWALL that aims to develop a device for use inside the home, for monitoring health of older adults, provide easy access to doctors and sensing your daily activities with the goal to inform relatives if emergency occurs. We are currently designing this application (we are in a preliminary stage) and we would like to have your opinion on it.

You will have some time to explore the application and tell us your first impressions on it. Then, I will ask you some questions and a couple of tasks to perform on the application. It will be very helpful if you express your thinking process during the scenario loud. I would like to inform you that our intention is not to judge you or your abilities, but only our designs and ideas. So feel free to ask any question or make any remark, because you are helping us understand what is wrong or right with our application. During the process and in the end, we will ask you to fill in some questionnaires. Thank you very much!

Informed Consent

The participant receives an Informed Consent (IC) that must be read and signed by her in order to participate. If needed, please go through it with the participant to gain her trust.

Scenario 1 – eWALL appearance and interaction trigger

In this scenario, our goal is to understand participants' first impressions on the setup (wall metaphor), the information displayed in inactive mode and the transition to active mode.

a. Older adult as participant
- *This is eWALL (inactive mode). What can you see on the screen from here?*
- *Please now, approach the device. What do you observe?*
- *Do you know why this (the change from inactive to active) happens?*
- *What is your first impression?*
- *Imagine eWALL on a wall of your living room. What would you like it to do for you?*

b. Secondary user as participant
- *This is eWALL in inactive mode. It is mounted on the wall and it will be in the living room of the patient. What do you think about the information on the screen?*
- *Do you think this setup makes sense for your patients/relatives/customers compared for example with a tablet?*
- *Please now, approach the device. What do you observe?*
- *What is your first impression?*
- *Do you see any benefits or disadvantages for this setup?*

Scenario 2 – eWALL main screen

a. Older adult as participant

Allow some time for the participant to explore the interface, click freely around and think aloud.
- *Can you describe each element you see in the screen?*
- *What do you expect to find under the category*
  - *Health?*
  - *Contacts?*
  - *House?*
  - *Games?*
- *Now, please click on each of these categories. What do you think each application listed under "Health" does? (same for all categories).*
- *Do you want to suggest any application that would make sense for such a device?*

b. Secondary user as participant

Allow some time for the participant to explore the interface, click freely around and think aloud.
- *What is your overall impression?*
- *Do you want to suggest any application that would make sense for such a device?*
- *How do you think your patients /relatives/ customers could benefit from such a device?*
- *How could you benefit? In particular, what kind of features in such a device could be useful for your work with older patients?*

Scenario 3 – eWALL sleep

a. Older adult as participant
- *Can you find your sleep summary from last night?*
- *Can you please rate your sleep from last night (objectively/honestly)?*
- *What do you think about the sleep summary? Would you be interested to read it?*
- *What do you think about the tags used to describe sleep? How (with which words) would you describe your sleep?*
- *Would you use such a rating system and how often?*

b. Secondary user as participant
- *Can you find your sleep summary from last night?*
- *Can you please rate your sleep from last night?*
- *What do you think about this application?*
- *What do you think about the tags used to describe sleep? Is there a better way to describe sleep?*
- *From your experience, would older patients be interested in using such an application?*
- *Any ideas/comments about how to move on with the implementation of this app?*

Scenario 4 – eWALL Daily Functioning Monitoring

a. Older adult as participant
- *Can you find and open the DFM application?*
- *Can you describe what it does?*
- *Can you describe all the elements in the screen?*
- *What do you think about this application?*

b. Secondary user as participant
- *Can you find and open the DFM application?*
- *Can you describe what it does?*
- *Would it be interesting for older adults to review their day in this way?*
- *On what aspects of the life of your patients/relatives/customers would you like to be informed? And how often - daily, weekly, monthly, only when there is an emergency?*
- *What do you think about this application? How can we improve it?*

Scenario 5 – eWALL Daily Phys. Activity Monitoring

a. Older adult as participant
- *Can you find and open the DPAM application?*
- *Can you describe what it does?*
- *Can you describe all the elements in the screen?*
- *What do you think about this application?*

b. Secondary user as participant
- *Can you find and open the DPAM application?*
- *Can you describe what it does?*
- *What do you think about the presentation of the physical activity in this app? How can we improve it?*

- *How can we motivate older adults with health problems to exercise more? What works well in terms of motivation from your experience? Do you have examples?*
- *Would it make sense to add specific medical exercises for the COPD patients? And which ones? (only if the participant has experience with this disease).*

Post Interview /Questionnaire

In case of older adult as a participant, we ask her to fill in a UMUX questionnaire.

a. Older adult as participant
- What is your overall impression of the system?
- Which were the best parts? Please give me 3 positive aspects.
- Which were the worst parts? Please give me 3 negative aspects.
- How would you redesign it? What would you change?
- What kind of applications would you add?

b. Secondary user as participant
- *What is your advice regarding the specific target groups we have and the system we are developing?*
- *Which needs your patients/relatives/customers have that we could potentially solve in the future with eWALL?*

Which needs are important and are not addressed b

Annex II: Protocol evaluation with expert reviewers

Participatory Heuristic Evaluation

Participatory heuristic evaluation (PHE) is a participatory inspection technique that serves as an extension to heuristic evaluation defined by Nielsen. In PHE, experts in software engineering, usability or human factors does the inspection as in traditional heuristic evaluation. However, work-domain experts are added as a group of users doing the same inspection. The purpose of extending the heuristic evaluation with these work-domain expert inspectors is to complement the traditional inspectors' more abstract knowledge with very specific knowledge about the work-domain. In this evaluation of the eWALL system, the work-domain experts would be healthcare professionals (EG nurses and/or doctors). We choose five experts in software engineering and usability and five work-domain experts, nurses, to participate in this PHE.

Before the PHE begins the users will receive information about the objective of the test, followed by an introduction to the 15 heuristics by Muller et al. (1998). While the experts and the work-domain experts are performing tasks, they are asked to think aloud. That includes: describing their first expression of what they observe on the system/ interface, describe what they like and dislike regarding the system, how they feel when using the system, what irritates them when they use the system, etc. By using the technique, thinking aloud, we will get the experts 'real' opinion of the system/interface and thereby improve the system/interface. Furthermore, the experts will be shown how to fill out a log schema over usability problems they will identify during the PHE. Hereby, they get information about a severity rating scale and finally they will receive six scenarios. This introduction is listed in bullets below here.

The PHE session

- Information about the objective
- Information about the heuristics
- Information about thinking aloud
- Information about the log schema
- Information about the severity rating
- Information about the scenarios
- Starting the PHE
    - Scenarios will be delivered to the experts and the work-domain experts
- Ending of the PHE, thanks, Goodbye!
- Administration of the results

The PHE session can be carried out by one facilitator (minimum) who will be in charge of welcoming the participant, describing the scope of the project, asking questions, and keeping notes of the participant's comments. It is however recommended - both for comfort but also for ensuring better coverage of participant's views - that 2 researchers conduct the study: 1 serving as the facilitator, 1 as the observer. In the room, except for the eWALL setup, there should be:

- An iPad to record what the participants say

- A laptop for entering information directly to a results form (not absolutely necessary)
- Markers, pens in reach to the participant and blank paper, in case participants want to sketch ideas.

The Objective

We are participating in a large-scale European project called eWALL that aims to develop a device for use inside the home, for monitoring health of older adults, provide easy access to doctors and sensing your daily activities with the goal to inform relatives if emergency occurs. We are currently designing this application (we are in a preliminary stage) and we would like to have your opinion on it.

The heuristics

SYSTEM STATUS		
1	SYSTEM STATUS	The system keeps users informed about what is going on through appropriate feedback within a reasonable time

USER CONTROL AND FREEDOM		
2	TASK SEQUENCING	Users can select and sequence tasks (when appropriate), rather than the system taking control of the users' actions. Wizards are available but are optional and under user control.
3	EMERGENCY EXITS	Users can: • Easily find "emergency exits" if they choose system functions by mistake (emergency exits allow the user to leave the unwanted state without having to go through an extended dialogue) • Make their own decisions (with clear information and feedback) regarding the costs of exiting current work • Access undo and redo operations
4	FLEXIBILITY AND EFFICIENCY OF USE	Accelerators are available to experts, but are unseen by the novice. Users are able to tailor frequent actions. Alternative means of access and operation are available for users who differ from the "average" user (e.g., in physical or cognitive ability, culture, language, etc.)

CONSISTENCY AND RELEVANCY		
5	MATCH BETWEEN SYSTEM AND THE REAL WORLD	The system speaks the users' language, with words, phrases, and concepts familiar to the user, rather than system-oriented terms. Messages are based on the users' real world, making information appear in a natural and logical order
6	CONSISTENCY AND STANDARDS	Each word, phrase, or image in the design is used consistently, with a single meaning. Each interface object or computer operation is always referred to using the same consistent word, phrase, or image. Follow

		the conventions of the delivery system or platform
7	RECOGNITION RATHER THAN RECALL	Objects, actions, and options are visible. The user does not have to remember information from one part of the dialogue to another. Instructions for use of the system are visible or easily retrievable whenever appropriate.
8	AESTHETIC AND MINIMALIST DESIGN	Dialogs do not contain information that is irrelevant or rarely needed (extra information in a dialog competes with the relevant units of information and diminishes their relative visibility).
9	HELP AND DOCUMENTATION	The system is intuitive and can be used for the most common tasks without documentation. Where needed, documentation is easy to search, supports a user task, lists concrete steps to be carried out, and is sized appropriately to the users' task. Large documents are supplemented with multiple means of finding their contents (tables of contents, indexes, searches, etc.).

ERROR RECOGNITION AND RECOVERY		
10	HELP USERS RECOGNIZE, DIAGNOSE, AND RECOVER FROM ERRORS	Error messages precisely indicate the problem and constructively suggest a solution. They are expressed in plain (users') language (no codes). Users are not blamed for the error.
11	ERROR PREVENTION	Even better than good error messages is a careful design that prevents a problem from occurring in the first place. Users' "errors" are anticipated, and the system treats the "error" as either a valid input or an ambiguous input to be clarified.

TASK AND WORK SUPPORT		
12	SKILLS	The system supports, extends, supplements, or enhances the user's skills, background knowledge, and expertise. The system does not replace them. Wizards support, extend, or execute decisions made by users.
13	PLEASURABLE AND RESPECTFUL INTERACTION WITH THE USER	The user's interactions with the system enhance the quality of her or his experience. The user is treated with respect. The design reflects the user's professional role, personal identity, or intention. The design is aesthetically pleasing—with an appropriate balance of artistic as well as functional value.
14	QUALITY WORK	The system supports the user in delivering quality work to her or his clients (if appropriate). Attributes of quality work include timeliness, accuracy, aesthetic appeal, and appropriate levels of completeness.
15	PRIVACY	The system helps the user to protect personal or private information— belonging to the user or to his or clients.

The thinking aloud technique

According to usability expert Jakob Nielsen, the technique thinking aloud is defined as:

> *"You ask test participants to use the system while continuously thinking out loud — that is, simply verbalizing their thoughts as they move through the user interface."*

The log schema

Heuristic no.	Location: Where is the problem?	Problem: Describe the problem.	Severity rating: (Category 1-4)	Solution – Suggestions for how to improve the eWALL

The severity rating

1 = Cosmetic problem only: need not to be fixed unless extra time is available on project
2 = Minor usability problem: fixing this should be given low priority
3 = Major usability problem: important to fix, so should be given high priority
4 = Usability catastrophe: imperative to fix this before product can be released

The scenarios

Scenario 1 – eWALL appearance and interaction trigger
In this scenario, our goal is to understand participants' first impressions on the setup, the information displayed in inactive mode and the transition to active mode.
It is the first time you have to log in to eWALL. Choose your language and type your user name.
The questions below can be used as support for the facilitator:

- *This is eWALL (inactive mode). What can you see on the screen from here?*
- *Please now, approach the device. What do you observe?*
- *What is your first impression?*
- *Imagine eWALL on a wall of your living room. What would you like it to do for you?*
- *This is eWALL in inactive mode. It is mounted on the wall and it will be in the living room of the patient. What do you think about the information on the screen?*
- *Do you think this setup makes sense for your patients/relatives/customers compared for example with a tablet?*
- *Do you see any benefits or disadvantages for this setup?*

Scenario 2 – eWALL main screen
Allow some time for the participant to explore the interface, click freely around and think aloud. Hereafter:

You want to know what time it is, how do you figure that out?
You want to edit your personal profile – how will you do that?

The questions below can be used as support for the facilitator:
- *What is your overall impression?*
- *Can you describe each element you see in the screen?*
- *What do you expect to find under the category*
  - o *Health?*
  - o *Contacts?*
  - o *House?*
  - o *Games?*

Scenario 3 – eWALL Health
You want to have a look at your Daily Physical Activity Monitoring. Where do you find more information about it?
The questions below can be used as support for the facilitator:
- *What do you think about this application?*
- *Can you explain the different elements on the screen?*
- *What do you think about the presentation of the physical activity in this app? How can we improve it?*

You were very tired yesterday so you went early to bed. Now:
1. You want to see how your sleep was during the night.
2. You want to see details about your night
3. Explain the symbols presented in your sleep monitoring
4. You want to rate your sleep and save your rating
5. You want to add descriptions of your sleep and save them
The questions below can be used as support for the facilitator:
- *What do you think about this application?*
- *Can you explain the different elements on the screen?*
- *What do you think about the presentation of sleep in this app? How can we improve it?*

Scenario 4 – Your contacts
You want to contact your wife, Maria, - please show me how to do that?
You want to contact your formal caregiver – how will you do that?

Post-questions

The post-questions can be used as a follow-up, where unsolved elements can be discussed. It is not necessarily relevant to go through all the questions, but they can be used as inspiration to cover all aspects of the users' experience of the interface.
- What is your overall impression of the system?
- Which were the best parts? Please give me 3 positive aspects.
- Which were the worst parts? Please give me 3 negative aspects.
- How would you redesign it? What would you change?

- What kind of applications would you add?
- What is your advice regarding the specific target groups we have and the system we are developing?
- Which needs your patients/relatives/customers have that we could potentially solve in the future with eWALL?
- Which needs are important and are not addressed by the system?

Two schemas the data was putted in.

Example 1:

Expert no.	Heuristic no.	Location	Problem	Severity Rating Category	Solution

Example 2

Heuristic no.	The heuristics	The usability experts. The total number of times each heuristic was used	The work-domain professionals. The total number of times each heuristic was used	Usability Experts +Work-Domain professionals

Annex III - SSE M22 Prototype Evaluation Protocol

The checklist

Tasks	Check
1 WEEK BEFORE	
Explore the interface, try out all the scenarios proposed. It is necessary to have a first-hand experience with the system.	
Book the evaluation facility.	
Recruit end-users for specific dates and times.	
Ask one of your colleagues to be available during the study as the observer.	
Prepare and send instructions for the participants how to find the facility.	
Print Informed Consent forms (consider some extra copies in case participants request to take it home).	
Print the questionnaires.	
Print this protocol.	
Arrange incentives for participants.	
3 HOURS BEFORE	
Check the prototype and the camera if they work.	
Make sure informed consent and questionnaires are in place.	
Rehearse the protocol.	

Facilitator tips

➢ Make small talk with the participant to warm her up.
➢ Explicitly encourage each participant to **think aloud** about their actions on the tasks given.
➢ Listen and do not make any judgments. Just **nod your head** to show that you understand her.
➢ If you think of more questions, go ahead and ask them. The more feedback, the better.
➢ Dig below top-of-the-mind answers. **Find out why and how.**

Timeline

Time (in min)	Activity
5	Introduction
5	Informed Consent
50	Task based evaluation
5	Just last cycle: User Experience Questionnaire
5	Post-questionnaire (Demographic Data, Technology Experience)
5	Incentives/Sign reimbursement, thanks, goodbye!
70-75	**Time total**

Introduction

• Thank them for participating

- Explain briefly what will happen within the next hour.
- Underline that we are not aiming to judge their opinions but want to evaluate our ideas: "*There are no wrong answers. We are interested in your thoughts and opinions. Everything you say and do is helpful for the project. The more you contribute to this evaluation, the better it is.*"
- Encourage them for thinking aloud.
- Introduce shortly the project and the purpose of eWALL
-

Informed Consent

The participant receives an Informed Consent (IC) that must be read and signed by her/him in order to participate. If needed, please go through it with the participant to gain her/his trust.

Personas

Just for Expert evaluations → introduce Personas to demonstrate them a potential end user:

Persona:	Bob (65 years, male)
Domestic situation:	Lives alone
Characteristics:	COPD Stage 3, Ex-Smoker, Underweight, Hearing problems, Has trouble sleeping, No experience with modern technology

Persona:	Jane (74 years, female)
Domestic situation:	Lives alone
Characteristics:	Cognitive decline, Memory deficits, Sleeping problems, Anxiety, Avoid social contact, No experience with modern technology

Persona:	*Michael (67 years, male)*
Domestic situation:	Living at home with wife (Maria) in a big city
Characteristics:	Hypertension, Forgetfulness, Social Anxiety, Lack of Motivation, Social Isolation, Experience with modern technology

Task-based Evaluation:

Task 1: Main Screen
"Imagine you have the eWALL system at home and you are in your living room. You want to use eWALL so please walk towards the screen."
[When they arrive in front of the screen, close enough to touch it, the screen zooms out and shows the full screen.]

115
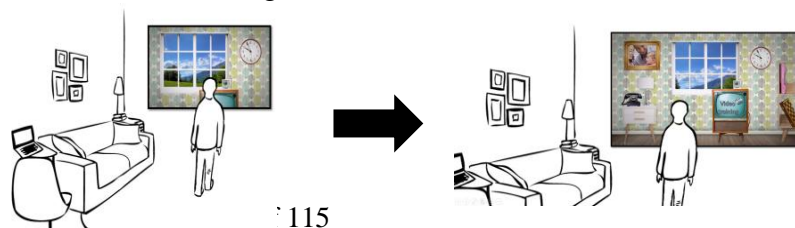
Questions:

Q1: What is your first impression?

Q2: Please describe how you perceive the interaction with the screen.

Q3: Can you imagine having a large screen like this in your house? Why (not)?

Q4: What would you expect behind each item on the screen?

Observation:

O1: How does the user interact with eWALL? → Screen size, distance to the screen, standing interaction, active/passive mode, Does the user find out that objects launch representative information by touching it?

Task 2: Personal data
"Please open the pink book in the shelf."

Questions:

Q5: Please explain what you can discern.

Q6: What do you think you can do here?

Q7: What do you expect by clicking on the buttons?

Task 3: Daily Functioning Monitoring
"Now please open the Daily Functioning Monitoring."

Questions:

Q5: What is your first impression?

Q6: Please explain what you did last Monday according to this overview.

Q7: Is it interesting for you to track your daily activities? Why (not)?

Q7a: Which of the information is useful?

Q7b: What would you additionally like to know about your daily activities?

Q8: How do you feel about seeing such an overview of your day?

Q9: Could you imagine sharing this data with your family? How about with your nurse or doctor?

Observation:

O2: How does the user interact with eWALL? →touch interaction, swipe functionality to see another day

Task 4: Daily Activity Monitoring
"Let's see your physical activity overview. Please open it and find out how active you've been the last 3 days."

Questions:

Q10: What is your first impression?

Q11a: Please explain what you discern in today's overview.

Q11b: How many steps did you do the last week?

Q12: Is it interesting for you to track your physical training activities? Why (not)?
Q13a: Which of the information is useful?
Q13b: What would you additionally like to know about your physical training activities?
Q14: How do you feel about seeing your physical activity overview?
Q15: Could you imagine sharing this data with your family? How about with your nurse or doctor?
Observation:
O3: Does the user find the way from the DFM to the DPAM?

Task 5: Sleep Monitoring
"Imagine you want to know how many hours you slept the last night. How would you proceed?"

Q16: What is your first impression?
Q17a: Explain what you can discern in the graphic of last night's sleep.
Q17b: How many sleep interruptions does it show you for the last night?
Q18: Is it interesting for you to track your sleep habits? Why (not)?
Q19a: Which of the information is useful?
Q19b: What would you additionally like to know about your sleep?
Q20: How do you feel about seeing such an overview of your sleep?
Q21: Could you imagine sharing this data with your family? How about with your nurse or doctor?

Task 6: Sleep rating:
"Add a rating of your last night's sleep."

Question:
Q22: What is your first impression?
Q23: Is it interesting for you to track your subjective perceived sleep quality?
Q24a: Which items are useful?
Q24b: What information would you like to additionally add?
Q25: How do you feel about answering these questions every day?
Observation:
O4: How does the user interact with eWALL? →ticking the boxes, reading it, rating with stars

Task 7: Cognitive Training
"Now we will look at some of the games which train the memory. Please go back to the main screen and open the chess set."
Q26: What is your first impression?
Q27: What do you expect to be able to do here?
Q28: What would you expect behind each item on the screen?

Task 8: Playing Games
"Please try out the games."
Q29: Do you like this kind of games?

Q30: Can you imagine playing these games on a screen like this at your home? Why (not)?
Q31: Are there any other games you'd like to see here?

Task 9: Physical Training
"Please open the video training application and start a video."

Q32: What is your first impression?
Q33: What do you think about this kind of training instruction?
Q34: If offered, would you perform such exercises yourself? If YES: How often?

Task 10: Rating the training
"When you are done with the exercise, you get to rate your experience of the training."

Q35: What is your first impression?
Q36: Do you think a technology can create a good personal training schedule for you, based on such questions*?
(* refers to the exercise ratings)
Q37: How do you feel about answering these questions after each video training?
Q38: What do you think about sharing this feedback with your family? How about with your nurse, physical therapist or doctor?

Task 11:General experience

Q39: What do you think about this system?
Q40: Would you use the system at home? Why (not)?
Q41: How easy or difficult will it be to learn to work with eWALL?
Q42: How easy or difficult will it be to use eWALL in your everyday life?
Q43: Do you expect to be able to make eWALL what you want it to do?
Q44: Do you like how eWALL looks?
Q45: Finally, if you could give some advice to the makers of eWALL, what would it be?

Post Interview /Questionnaire

o   Demographics

Annex IV - Summary of Small Scale Evaluation Results

Evaluation partner	User	Category	Observations and Recommendations/Suggestions
Task 1: eWALL Main Screen			
AAU, RRD, RRD	UX experts	design/aesthetics	The telephone does not exist today; probably the TV will also not be something that will exist in the future. Why does everything look old – the users are not old just because they have turned on 65 years – it might have a negative influence and demotivate the users or give them the feeling that they are old and "stupid".
IRCCS	Neuro-psychologist	design/aesthetics	Books on the shelf are too thin. It is possible to enlarge them?
ATE, AAU	UX experts	customization	It would be nice to individualize the interior - picture frame, TV, lamp, etc. (e.g. 70's, 80's style but also modern). | We do not have mountains in Denmark, so the background in the window has to be changed.
AAU, IRCCS, RRD	UX experts	motivation, reminder	It is unclear how the screen can be activated from passive mode. | Something moving in the main screen could give the impression of an active system. A puppy or a fish bowl, and subjects could also feed or pet… just thinking loud! | It is possible to add a button with a reminder of the use of the platform? I wonder if people who forget things will use the app | add a start button.
ATE	UX experts	human factors	The screen size is ok but if the user wants to see the complete screen he/she has to get some steps back and for clicking again forward many older adults already have vision problems and for a better orientation it may be better to have a smaller screen size - would also better fit to average room sizes of older adults.
ATE	UX experts	icons	Game objects are too unclear - expectations did not go along with the games.
AAU	UX experts	icons	It is not obvious what the different elements symbolize.
RRD	Elderly	icons	Elderly did not make distinction between widgets and app launchers.
ATE, IRCCS		layout	The size of window and television are in relation to the other objects more present. The position of the TV set is too prominent, activities buttons should be more in the foreground.
ATE, AAU, IRCCS	UX experts	user control	It is not clear which objects are buttons | labelling of all interactive elements is necessary
AAU	UX experts	user control	Missing feedback when clicking on an icon.
RRD	Elderly	new function	Clock should show time for meals, medication.
ATE, RRD, IRCCS, AAU	UX experts	new function	Regarding the picture frame: would be a nice function if family members could send pictures to the eWALL system. the picture frame to be used for FaceTime with family. It could be used also as a reminder: a cake for a birthday, depending on the creativity of carers.

ATE	MCI/COPD Experts	new function	Screen should be possible to be used as a TV.
ATE	MCI/COPD Experts	new function	It would be nice if system was talking with the user since a lot of older adults live very isolated.
IRCCS	Neuro-psychologist	new function	The possibility to have a "personal assistant" button should be taken in consideration. MCI people often need the instruction and the training to be repeated several times. A button with a personal assistant could be really useful, his aim should change with respect of the functionality chosen. For example he could give tips related to the functionality or explain the rules of the games etc…
AAU		new function	If you make the phone as an interactive tool – it could work as a phone book
ATE, RRD	MCI/COPD Experts	data sharing	Informal caregivers have to be included for using this system - without this support, the primary users will not use all functions of the system. I would like to share this info with my doctor, if it helps to make my sleep better. The partner and doctor can easily monitor the situation, but it needs to be secure (only trusted persons should have access to it). I would share this app with family (partner, kids) and doctor, if the data will be always neutral (DFM). It would be useful to share the answers with the doctor in video training and rating. It might be interesting to compare my data with other elderly | access to this information should be limited to family and friends and doctors (DPAM). I would use this every day, if doctors says so (video training).
RRD, ATE, AAU	Elderly	data sharing	Some elderly would have problems with privacy. | Users may feel irritated because they feel monitored. | I am afraid it is going to evolve into something forcing elderly to behave well. | The users should decide on their own, if the information is allowed to be seen by secondary users.
AAU	UI Experts	information relevance	Why is it relevant to know how long time you are cooking dinner?
AAU	UI Experts	customization	It should be possible to adjust the level of complexity in the games and in exercise videos.
AAU, ATE, IRCCS	UI Experts	human factors	The games are designed so the users have to stand up and play - but it might be a problem regarding the complexity and status of the users' disease. | Some of the COPD-patients do not have the strength to stand up for a longer period. | Playing in a standing position will not be accepted of older adults
AAU, ATE	UI Experts	human factors	It is a pity that the screen display isn't fully utilized, for example in the video training.
ATE	UX Experts	human factors	Older adults have small rooms - screen size is too big.
IRCCS	Neuropsychologist	human factors	In sleep rating, writing text on a big touch screen is not something elderly are comfortable with.

AAU, IRCCS	UI Experts	icon X	The cross in the right corner of the book is placed the same place everywhere - anyhow, it fits when you read/open the book with the DFAM but it does not fit when you have the games open. | The cross turns upside down. | If is possible, write "close" instead of X in the icon.
AAU	UI Experts	icon stars	The relevance of using stars as the way to score is not suitable. | "What do you think of this question" – is a very open-ended question and cannot be answered with stars.
AAU	UI Experts	layout	Several of the buttons are unclear.
AAU	UI Experts	colours	The meaning of the colours is not fully described in healthcare monitor | the pink colour is very girly | too many colours in the pink book when the users have walked too little, the difference in the colours are not that distinct to the average DPAM. Choosing green/red is a strange colour choice - if the users are colour-blind they cannot read the text. In the **sleep app**: why is the comment written in blue In **DFM:** contrast too less
AAU, ATE		font colour, contrast, type and size	Fonts are small; in DPAM, fonts in numbers are small, bad contrast. The contrast between the TV and the background is too small | Stripped font is bad. | Typography is not the same | the colours are confusing and the font type and size are confusing too. | In **video/physical training**, the question: 'are you ready to train'? Contains too much different font types. | In rating the training, stripped font and neon yellow are awful. In the **healthcare monit**or, the size of the font is not consistent. | There is no balance in the typography. In t**he main sc**reen, the colour of the temperature and the humidity is written with white. Little contrast, small font.
ATE, AAU		consistency	There is bad consistency between the activity functions: why does sleep app look different to daily activity monitoring? Scrolling of days at the bottom of the screen is missing in **DFM**: - the time intervals have to be equal, the time line appears asymmetric. in **Games**: the buttons in the bottom is not consistent in **Physical Training**: the buttons are not the same size in the **Healthcare Monitor**: the intervals are not equal | there is only 6 days shown in a week | it's strange that the calendar begins with a Friday.
AAU, ATE, IRCCS	UX experts	function help	A description of what the system is and the goals of the system is missing. Introduction of its elements is also necessary. Also: In **cognitive training and game**s: - instructions of games are missing | Clarify and repeat the instructions of the games especially for MCI. In **physical trainin**g: - an instruction how to use the Fitbit is missing

IRCCS, ATE, AAU	Tele-rehabexperts, UX Experts	wording	Please avoid terms and phrases that could instil in participants the doubt or the feeling of being followed or monitored. (No STASI or KGB speaking, please). "We detected" or "We noticed" seems something in "STASI style" (DFM) | "We recognized you in the living room!" --> Who is "we" --> this sentence can be deleted as it is not relevant to the user at **DPAM:**'monitoring' is difficult to understand in DFM and sleep app
ATE, RRD	UX Experts, Elderly	validity & coherence	In **sleep app**: - system does only know if I have been in my bed but what if I have been watching TV? In **games**: - these are just games, don't know what they train. Should we describe aim of the games? In **video training**: - users don't know if they are doing the exercise correctly | how does the system know if exercise has been done till the end?
		Task 2: Personal Data (pink book)	
ATE	UX Experts	information overload	Numbers of pages are irritating, no need for it as the overviews are linked by touching the button (would make sense if the user can browse through the book). | The numbers in the book are incomprehensible.
AAU	UX Experts	information overload	information overload on page 5
AAU	UX Experts	icons	It is unclear what the book contains. | The book looks like a romance. The book as a frame does not make sense.
AAU	UX Experts	icon arrow	The arrow heads do not fit to the book.
AAU	UX Experts	user control	It is problematic that the user is available to go back but not forward.
AAU	UX Experts	wording	How should the users know that NY stands for DPAM etc. | "monitoring" is not layman's language. | It is unclear what the different DFAM etc. covers
		Task 3: Daily Functioning Monitoring	
ATE	MCI/COPD Experts	meaning /aim / impression	First impression makes nervous, so much information on the screen - longitudinal effects are unclear, but maybe a prioritization of elements concerning size and colours would help.
AAU, RRD	UX Experts	meaning / aim	It is very difficult to see the meaning of having this monitoring. | I don't understand what the goal of the app is. | It is unclear what this functionality is for. | Using the page is not intuitive.
RRD	Elderly	meaning / aim	The app is clear and logical, can't see how to make it more fun.
ATE	UX Experts	information overload	Showing seconds is not necessary.
ATE	MCI/COPD Experts	information relevance	The information is only important for secondary users - what was my relative doing all the day, is everything ok - alarm if unusual activities are noticed.

ATE	MCI/COPD Experts	COPD information relevance	For COPD patients not important information - especially if activities cannot be commented.
ATE, AAU	MCI/COPD Experts	information relevance	Active COPD patients will profit from this information, very inactive users will not need this information since they are not motivated at all to change their activity. | A lot of information shown - will be overwhelming for untrained users
ATE	MCI/COPD Experts	information relevance	Information on how much has the user has drunken or eaten is important.
AAU	UX Experts	information relevance	The functionality is not relevant for COPD-patients.
ATE	MCI/COPD Experts	information relevance	Combination of daily information and emotional rating would be interesting.
ATE	UX Experts	information relevance	Additional information here: medicine intake, air quality inside the room, visits.
ATE	MCI/COPD Experts	motivation	Information that user has been very inactive can also be very frustrating.
IRCCS	Neuropsycho logist	motivation	It could be useful to link this information to the wellbeing of the participant using motivational messages.
AAU	UX Experts	icons	The icons are difficult to understand.
ATE	UX Experts	wording	How does system know that user is resting? --> better to write "sitting/lying on the couch" etc.
ATE	MCI/COPD Experts	wording	Term "Aufstehen" is not suitable; it should say "Hausverlassen".
AAU	UX Experts	layout	The buttons are unclear.
ATE	UX Experts	layout	The bar for choosing the day should be on top of the screen.
ATE,AAU	MCI/COPD/UX Experts	layout	Separation of activities is misleading; active phases should be one side and passive on the other. Points of time on the left side, icons with activity on the right side
AAU	UX Experts	layout	The vertical time line is not immediate clear how to read and understand. A title of the vertical time line is missing.
Task 4: Daily Activity Monitoring			
ATE, IRCCS	MCI/COPD Experts, Tele-rehab expert	meaning	For unexperienced users, graphical style will be difficult to interpret at the beginning. | It seems too scientific for elderly people.
ATE	MCI/COPD Experts	information relevance	For COPDpatients, "running" will be not relevant, since this form of activity is very unusual.
ATE	MCI/COPD Experts	information relevance	For very lonesome and inactive persons this overview will have no meaning.
AAU	UX Experts	information relevance	It is not relevant for all the users to have this functionality – e.g. if the users do not like to do exercise.
ATE	UX Experts	motivation	At the moment only colour is a motivational factor for being more active - think about other motivational aspects: for example distance from home to Eiffeltower=xx km, or I have achieved already xxkm, or have burnt xx calories = 50 burgers
IRCCS	Neuropsycho logist	motivation	A motivational message to summarize the screen could be useful.

ATE	MCI/COPD Experts	customization of goals	Defining personal goals is very important for COPD patients - medical assistance for thresholds of biological values is necessary.
AAU	UX Experts	icon	The average 10.000 is not easy to understand.
ATE	UX Experts	user control	Information about where I did the many steps would be useful --> different colours for inside and outside steps | would be good to see when I did the training, when I walked, when I was running.
AAU	UX Experts	user control	The average for a week is missing.
AAU	UX Experts	user control	There is a lack of information about the graph. | Units on the graph are missing.
IRCCS	Neuropsychologist	measuring units	More than kilometres is better to use meters.
IRCCS	Tele-rehab experts	graph issues	A small icon near the axes could ameliorate the comprehension of the graph. | It seems something is missing on the right of the graph: a small watch and 2 feet for example.
IRCCS	Neuropsychologist	wording	Messages seem a bit confusing.
ATE	UX Experts	new function	Graphs show only steps, but not training session or resting; no outdoor activities.
Task 5: Sleep Application			
AAU	UX Experts	information overload	Too much text.
RRD	Elderly	information relevance	I wonder if it's useful to know these things. When I sleep badly, I feel it.
AAU	UX Experts	information relevance	it is unclear what COPD-patients are going to use this for. | Sleep periods do not make sense for a COPD-patient. it is relevant - maybe - for the MD-patients.
ATE	MCI/COPD Experts	COPD info relevance	For COPD patients, time of coughing or dyspnoea would be relevant to interpret if an infection is coming.
ATE	UX Experts	ARI info relevance	Benefit of this service for ARI is questionable.
AAU	UX Experts	design	The cats make the screen look like a kindergarten.
IRCCS	Neuropsychologist	motivation	Nice! It could be useful to link this information to the wellbeing of the participant using motivational messages. | It could be nice to add some tips, such as "try to read a book before you go to sleep" "It is better not to watch too much TV late at night".
ATE	UX Experts	consistency	Bad consistency between the activity functions: why does sleep application look different to daily activity monitoring? Scrolling of days at the bottom of the screen is missing.
AAU	UX Experts	icons	The icons are difficult to understand.| the icon where a man is in a circle - it does not make sense.
ATE	UX Experts	user control	Bar to choose the day (as it is in the DFM and DPAM) is missing. | It would be good to have an overview of the week.
AAU, RRD	UX Experts	user control	From-to moments are missing | wake-up part is not clear - system has to say when I got out of bed.

ATE	MCI/COPD Experts	user control	Overview of sleeping time overall and interruptions overall would be fine to have.
ATE, RRD, AAU	UX Experts	layout	Again two columns - reading is difficult. | It should display 1 list, not left-right cards.
AAU	UX Experts	colours	The contrast in the colours is too small.
ATE	UX Experts	wording	"Early" or "late" are too interpretative.
AAU	UX Experts	wording	*EARLY* is confusing - why is the term highlighted? it is unnecessary information.
AAU	UX Experts	wording	There are 2 full stops after each other.
AAU	UX Experts	wording	The phrasing 'judge your own sleep' is too complicated.
Task 6: Rating Sleep			
ATE, AAU	UX Experts	information relevance	The aim of the rating is unclear. | What is the goal of asking these questions? | The questions ought to be individualised. | Several of the questions are irrelevant. | How are we using the sleeping data? That has to be discussed and described.
ATE	UX Experts	information relevance	Some questions are irrelevant since system is monitoring anyway: wake up, cold/hot also not informative. Delete unnecessary questions: did you wake up? Were you at the toilette?
ATE, RRD	MCI/COPD Experts	information overload	Too many questions, max 4, users will not want to rate everyday their sleep. | Answering the questions is fine for a week, for longer it will be annoying.
IRCCS	Neuropsycho logist	information relevance	Questions are right for a good comprehension of sleep disorders.
AAU	UI Experts	MCI info relevance	Given the fact that the people who suffer from mild dementia can have some problems remembering how they slept last night (for instance it can be difficult for them to remember how their night was), it can be difficult for them to rate the sleep of their previous night. So there is a risk that they will give a wrong answer.
ATE	MCI/COPD Experts	COPD info relevance	Differentiation between productive/unproductive coughing for COPD patients is relevant - please adapt.
ATE	UX Experts	icons	Pictograms instead of questions for a quicker rating.
AAU	UX Experts	user control	It is not clear how the users can 'save' and complete the session.
IRCCS	Tele-rehab experts	user control	Please insert a "finish" and a "skip" button.
ATE	UX Experts	user control	"Rating finished" button is missing.
IRCCS	Neuropsycho logist	user control	"Had bad dreams?" The answer "don't remember" should be taken in consideration Often, snoring people don't know they are snoring or they don't mind.
AAU	UI Experts	user control	There is missing a box/frame around the 'yes' and 'no' answers.
Task 7: Cognitive training			
RRD, ATE | RRD	Elderly	game relevance	Mixed opinions about game quality. | Elderly who play games on tablet, or phone were enthusiastic, those who do not play video games disliked all of them and considered the games unnecessary. | Games like Sudoku, playing cards, crossword puzzle are missing and are played by the target groups.

ATE, IRCCS, AAU	UX Experts	customization	Games need different levels of difficulty.
ATE	UX Experts	motivation	Reward system for achieving points is missing; there is no motivation for playing these games.
ATE	MCI/COPD Experts	icons	Interactive symbols are not clear enough - chess play was interpreted as deco.
RRD, ATE, AAU	Elderly	icons	I expected card games for card icon, domino, for domino icon, no clue for the spinning toy.	Icons do not represent the games behind.	The experts thought they were going to play backgammon when they saw the game - but they could not. -It is unclear what the icons symbolize.
IRCCS	Tele-rehab experts	new function	Draughts, backgammon, domino, dice; none of these games are implemented.
Task 8: Playing Games			
AAU, IRCCS	UI Experts	motivation	The game looks boring because of the way it is presented - With no motivational message, games will be a bit boring.
AAU	UI Experts	icons	Use text rather than icons.
AAU	UI Experts	icon	The flag confuses - it is not indicative. The users may think that the game is about speaking/learning language.	Why looking at flags - and not family members instead?
AAU	UI Experts	user control	A 'start-stop'-button is missing.	A backwards-button is missing (back to the main menu).	Regarding all the games it is unclear how to start the game -add a start button.
IRCCS	Neuropsychologist	user control	A "pause" button could be useful during the gaming (less anxiety).
AAU	UI Experts	user control	It is unclear where it is possible to click.
AAU	UI Experts	colours	The contrast in the colours is missing.
AAU	UI Experts	colours	The yellow tool bar does not fit in the design, not reader friendly.
Task 9: Physical Training			
RRD	Elderly	information relevance	These things are also on TV. How does eWALL differentiate?
ATE, RRD	MCI/COPD Experts	information relevance	A mix of easy and difficult exercises is important.	Elderly trust the program to create a personal training schedule based on the questions. / The exercises are too easy for active elderly, they think these are 80+ gym classes.
AAU	UI Experts	information overload	The screen "hello Michael" is unnecessary	It also misses a back button.
AAU	UI Experts	information overload	The text in the box is too much - remove! - write 'START* instead.
ATE	UX Experts	activity validation	Users don't know if they are doing the exercise correctly.	How does the system know if exercise has been done till the end?
AAU	UI Experts	user control	The opportunity to press "repeat" is not possible. It is unclear why there is a 'done'-button and a 'break'-button. A "pause" button could be useful during the video training. A watch is missing - it could be nice to know how much time you have left during the exercise.
ATE	MCI/COPD	user control	The video stops too soon, it should wait until the person is

	Experts		done.
AAU	UI Experts	layout	It is slipshod that the numbers and the text are not aligned.
AAU	UI Experts	layout	The slide 'find a chair' is messy.
IRCCS	Tele-rehab experts	use of space | layout	Picture in picture in picture makes the design confounding (slide n. 17)
AAU	UI Experts	wording	Change the green button "are you ready – let's start" to START.
AAU, IRCCS	UI Experts	sound problem	The drum roll is stupid. | Remove sound for attention deficits. | It is not possible to hear what the lady says because of the drum roll. | The sound of stars sounds stupid.
		Task 10: Rating the training	
AAU, ATE	UI Experts	information relevance	It is not necessary to rate "how do you like this exercise?". | Rating should be only after first time of the exercise, not every time. | It is not necessary to rate if the users stop. | Subjective rate for exercise selection could be misleading.| The level of fatigue could be a useful index (more than the "like" rating).
ATE	UX Experts	wording	Too "bad" in what? "Bad" should not be used as a feedback.
AAU	UI Experts	sound	The sound of "congratulations" is condescending. | "congratulations" with what?
		Task 12: The Healthcare Monitor	
AAU	UI Experts	icons	The icons are not meaningful.
AAU	UI Experts	icon arrow	The arrow head down in the left corner is incomprehensible.
AAU	UI Experts	user control	A 'back'-button is missing.
AAU	UI Experts	user control	An indication of what are outliers is missing.
AAU	UI Experts	user control	It is not possible to change week. | You don't get further information when you press month rather than week. | It is not clear enough the indication that you are 'standing' on week compared with 'day' or 'month'. When you click on 'month' the users see one picture which shows the development over the total month. That is difficult.
AAU	UI Experts	user control	A frame is missing around the graph values. | It is unclear whether it is the average value that is shown on the graph. | A frame is missing around the terms 'history' and 'measurements'.
AAU	UI Experts	user control	The graph is difficult to read. | An overview of the values in the graph is missing. | There are no units on the graph.