
HERMES - Cognitive Care and Guidance for Active Aging
FP7-ICT 216709
Specific Targeted Research or Innovation Project

Start date of project: January 1, 2008
Duration: 36 months



D.7.2 User Evaluation Report Trial 1

Sebastian Prost, Arjan Geven, Manfred Tscheligi (CURE)

Mari Feli Gonzalez, Ana Navarro (INGEMA)

Version: 1.3
Date: 26/01/2010
Dissemination level: (PU, PP, RE, CO): PU

Project Co-Funded by the European Commission within the 7th Framework Programme

Abstract

This deliverable covers the procedure and results of the first user evaluation trials of the HERMES system. In particular it describes the conditions of the trials in Austria and Spain, how users were recruited and the two evaluation sessions, in which users tested HERMES MyFuture, PDA, MyPast and Cognitive Game by performing several tasks.

Results of the evaluation concentrate on the following aspects:

- General user acceptance of the technology
- Quantitative and qualitative enquiry about system acceptance and problems while performing the tasks
- Qualitative usability evaluation through think-aloud protocols and observation

The results for each of the components of the HERMES system are summarised and several conclusions about the performance of the system are drawn. In general the system performed fairly well, especially the MyFuture application. While HERMES is generally accepted by the users, all components need further improvement. Based on the results of the evaluation, instructions and advices are given about the changes that need to be made during the next iteration.

Table of Contents

1. INTRODUCTION	4
1.1 BACKGROUND	4
1.2 SCOPE OF THIS DELIVERABLE	4
2. COMPONENTS TESTED.....	6
2.1 AUSTRIA.....	6
2.2 SPAIN	12
3. PROCEDURE OF USER TRIALS.....	14
3.1 ADAPTATION OF THE USER EVALUATION PLAN	14
3.2 USER RECRUITMENT.....	18
3.3 FIRST SESSION	19
3.3.1 <i>Pre-evaluation</i>	19
3.3.2 <i>Tasks and Questions for HERMES MyFuture</i>	19
3.3.3 <i>Technology Acceptance Questionnaire for HERMES MyFuture</i>	21
3.3.4 <i>Tasks and Questions for HERMES PDA</i>	22
3.4 TASKS TO BE DONE AT HOME	23
3.5 SECOND SESSION.....	24
3.5.1 <i>Questions about HERMES PDA</i>	24
3.5.2 <i>Tasks and Questions for HERMES MyPast</i>	24
3.5.3 <i>Technology Acceptance Questionnaire for HERMES MyPast</i>	26
3.5.4 <i>General Questions about HERMES Cognitive Game</i>	26
4. RESULTS.....	27
4.1 CAS QUESTIONNAIRE	28
4.2 TECHNOLOGY ACCEPTANCE.....	30
4.2.1 <i>Results</i>	30
4.2.2 <i>Interpretation</i>	35
4.3 USABILITY EVALUATION OF COMMON COMPONENTS OF MYFUTURE AND MYPAST	36
4.3.1 <i>Results</i>	36
4.3.2 <i>Summary</i>	38
4.4 MYFUTURE APPLICATION	38
4.4.1 <i>Quantitative Evaluation</i>	39
4.4.2 <i>Qualitative Evaluation</i>	40
4.4.3 <i>Usability Evaluation</i>	42
4.4.4 <i>Summary</i>	45
4.5 HERMES PDA	45
4.5.1 <i>Quantitative Evaluation</i>	45
4.5.2 <i>Qualitative Evaluation</i>	47
4.5.3 <i>Usability Evaluation</i>	50
4.5.4 <i>Summary</i>	54
4.6 MYPAST APPLICATION	55
4.6.1 <i>Quantitative Evaluation</i>	55
4.6.2 <i>Qualitative Evaluation</i>	57
4.6.3 <i>Usability Evaluation</i>	59
4.6.4 <i>Summary</i>	61
4.7 THE COGNITIVE GAMES	61
5. ETHICAL ISSUES.....	63
6. OVERALL CONCLUSIONS	65
6.1 CONCLUSIONS FOR MYFUTURE.....	65
6.2 CONCLUSIONS FOR PDA.....	65
6.3 CONCLUSIONS FOR MYPAST.....	66
6.4 CONCLUSIONS FOR THE INTEGRATED SYSTEM	67
7. REFERENCES	67

1. Introduction

1.1 Background

Usability and user acceptance play a crucial role for the successful use of any technology. It is even more so for a system that is designed to meet the needs of older people, who might have limited previous computer experience. Therefore it is essential to know how older users interact with such a system under development in order to evaluate its performance and acceptance. Furthermore, the social and cultural background, attitudes and values strongly influence how a technical artefact is perceived. Older users mostly come from a non-computerised background, which puts them in the – nowadays unique – position of lacking many of the concepts and ideas that underlie our – i.e. the computer literate users’ – interaction with a computer. New technologies can frighten them, so they do not dare to even try things out. Older users can also simply ignore new technologies, because they don’t see any value in them for their lives or because they think they are not skilful enough. Skilfulness doesn’t simply mean being able to move a mouse or to double-click. Skilfulness means for example understanding what is a button that can be clicked on, understanding the concept of pointing and clicking on an object describing an action to actually perform the action, and understanding you need to point and click somewhere in order to do something, because otherwise the system doesn’t know where you want to perform an action. Designing complex technology for users that didn’t grow up with it is a challenging task and it needs testing with real users, especially since here the designers of the system are definitely *not* the users.

The first user evaluation trials of the HERMES system were conducted in late 2009 in Austria at CURE and in Spain at INGEMA in order to get these insights. By performing an evaluation using both quantitative and qualitative questionnaires based on literature as well as an expert observation looking for usability problems, thorough and deep information about user needs, wishes and problems has been gained. At both premises a lab setting was provided to enable older users to evaluate MyFuture, MyPast and the PDA.

1.2 Scope of this Deliverable

Input for this deliverable was mainly derived from Deliverable D.7.1. In this User Evaluation Plan the main research questions to be answered during the trials as well as the planned procedure was described. This deliverable aims to answer the questions raised in D.7.1, which are the following:

- Technology acceptance of the HERMES system in general
- Interface complexity of the MyPast application
- Information visualisation of audiovisual contents
- Game experience of the cognitive games

Technology acceptance of the HERMES system was evaluated during the first user trials after the users performed various tasks with its components. In particular a Technology Acceptance Scale (TAQ) based on Venkatesh et al. (2003) was used to assess acceptance for HERMES MyFuture and MyPast.

Interface complexity was evaluated by different means. First, users were asked quantitative and qualitative questions about the tasks they performed with the application. Second, users were encouraged to perform a think-aloud like realisation of the tasks. This means that, while

performing a task, users expressed any doubt, problem or idea they had immediately to the researcher, who noted this down. Finally, the researcher observed the user while he or she was trying to accomplish the tasks. Actually seeing users struggling with different interface elements and talking about their difficulties afterwards gave valuable insights.

Evaluation of information visualisation of audiovisual contents mainly concentrated on the timeline of HERMES MyPast (with or without filters applied), which contained different levels of complexity. However, as described in the later sections, only certain parts of the audiovisual components could be evaluated due to technical problems.

Game experience of the cognitive games could not be evaluated with the real application due to issues with transportation of the multi-touch screen from AIT to CURE's and INGEMA's labs. However, a video of the first cognitive game that was developed was shown to users. In a semi-structured interview users' options were enquired.

The results and conclusions of this user trial need to be incorporated in the next iteration of the HERMES prototype. The improvements made during the following months will be evaluated in the second user trial in summer 2010.

2. Components Tested

User evaluations were conducted both in Austria at CURE's usability labs and in Spain at INGEMA's premises. While in Austria the focus lied on the general interaction, the evaluation with Spanish users also deals with speech recognition components, which are developed with a focus on the Spanish language and therefore could not be tested in Austria. In the end, specific parts thereof were evaluated as described in more detail in Section 3.1. Table 1 shows which components were tested where. The following subsections present more details about the components tested in each country. Screenshots of the relevant elements are given as well.

Application	Component	AT	ES
MyPast	browsing using sliding bar	×	×
	changing date with date buttons	×	×
	specifying time filter	× ¹	× ¹
	specifying people filter	× ¹	× ¹
	changing a filter	× ¹	× ¹
	clearing a filter	×	×
	searching with search box	× ¹	× ²
	accessing of recorded audio		× ³
	accessing of transcribed audio		× ³
	using MyPast search to solve a quiz		⁴
MyFuture	browsing/finding appointments	×	×
	introducing a new appointment	×	×
	moving an appointment (Drag&Drop)	×	×
PDA	introducing a new appointment	×	×
	modifying an appointment (change time)	×	×
	modifying an appointment (add location)	×	×
	introducing a new appointment in locate mode	×	×
	recording of speech conversation	×	×
Cognitive Game	video of the first game that was developed		×

Table 1: Overview of evaluated applications and components by country. In Austria (AT) some components were not evaluated due to there speech processing dependency. These components were only assessed in Spain (ES).

2.1 Austria

User Trials conducted in Austria covered the HERMES MyFuture, PDA, and MyPast applications. Tasks with HERMES MyFuture covered browsing through scheduled events (Figure 1), introducing a new event (Figure 2), and modifying an event (Figure 3). A detailed description of the tasks performed with MyFuture is given in Section 3.3.2.

¹ Only the user interface element of this component was evaluated. For more details see Section 3.1.

² Spanish users could write a word in the search box, which was known to appear in the audio transcription of the PDA, and could hear the related audio. For more details see Section 3.1.

³ The participants could not access the audio of last year's recordings, but only the audio recorded in the PDA during the first session of this trial. For more details see Section 3.1.

⁴ It was not possible the test this component, because the last year's recordings were not available. For more details see Section 3.1.



Figure 1: HERMES MyFuture: On the main screen of MyFuture the user can browse through the calendar by pressing the “Previous Month” and “Next Month” buttons. Tapping on a specific day shows the events of this day.



Figure 2: HERMES MyFuture: As soon as a day is selected, a new event can be added on that day. Event details can be entered by the on-screen keyboard.



Figure 3: HERMES MyFuture: An event can be moved by dragging and dropping it to another day.

HERMES PDA was tested by letting the users to introduce and modify an appointment, browse through appointments (Figure 4 and Figure 5), introduce an appointment in locate mode and receive location-based notifications (Figure 6 and Figure 7), and record a conversation (Figure 8). The complete list of tasks assessed is presented in Section 3.3.4.

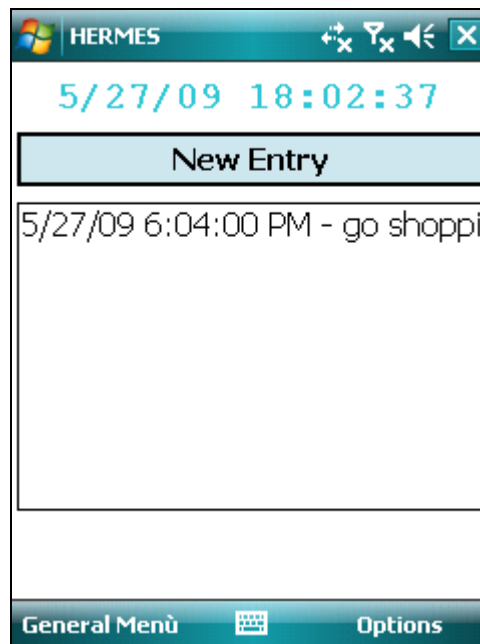


Figure 4: HERMES PDA: The main screen of the appointments application allows the user to add a new appointment and presents a list of appointments already recorded.

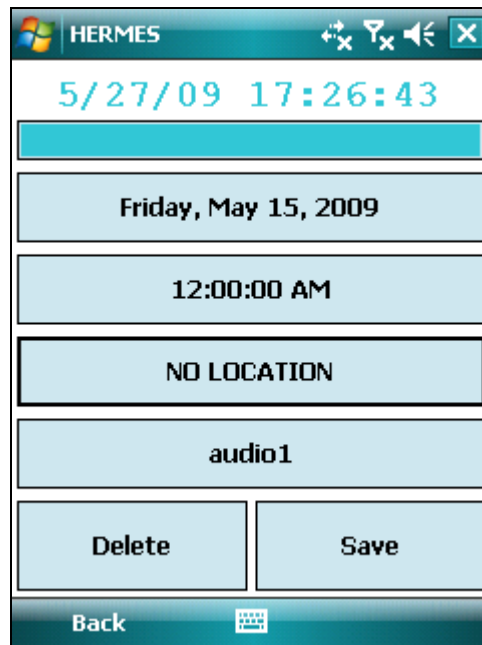


Figure 5: HERMES PDA: On the summary screen event details can be changed by tapping on the appropriate button.

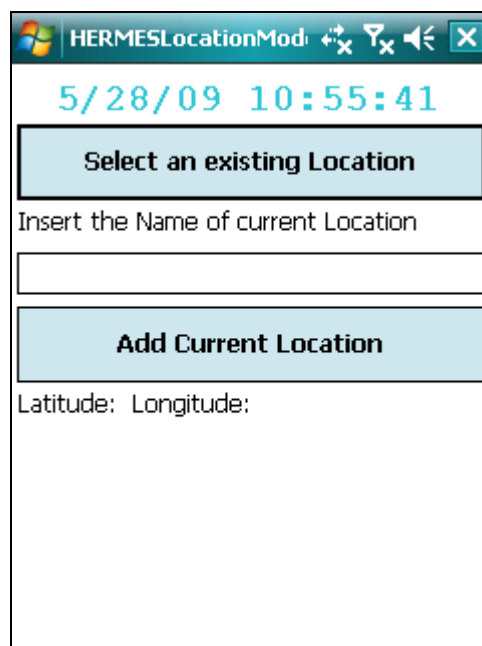


Figure 6: HERMES PDA: Location Mode offers the user the possibility to add the current location to the list of known locations. When the user approaches a location saved as a location-based reminder, he or she will be reminded.

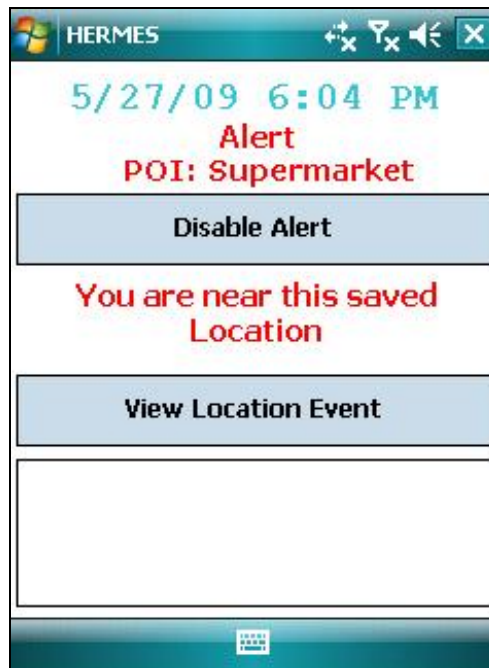


Figure 7: HERMES PDA: When a user reaches a location, whose GPS coordinates where previously recorded in an location-based reminder, the user receives a notification.

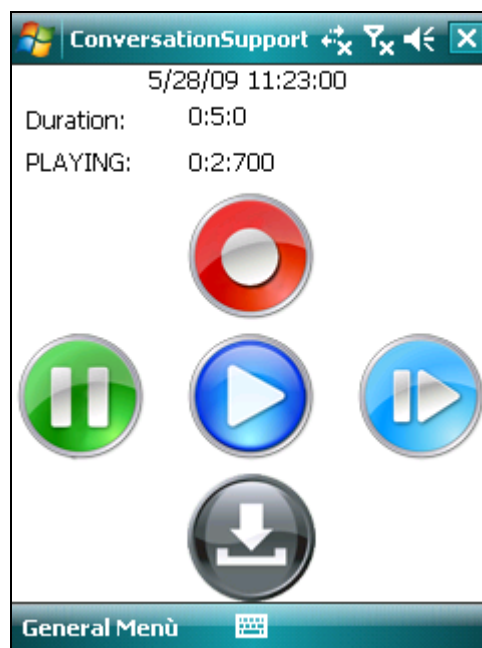


Figure 8: HERMES PDA: Conversation support allows the user to record and replay any type of speech via the on-screen record and playback buttons.

MyPast evaluation included browsing with the sliding bar and using the date buttons (Figure 9), specifying filters on time (Figure 10) and people (Figure 11), changing and clearing those filters and using the search box. While the interface components for adding filters and using the search box were evaluated, no actual functionality was present at the time of evaluation. In Spain however, the transcribed speech audio recorded with the PDA could be played when typing a word in the search box that appeared in the audio transcription. A full list of tasks performed with MyPast is given in Section 3.5.2, a more detailed description of the necessary adoptions of the evaluation plan in Section 3.1.

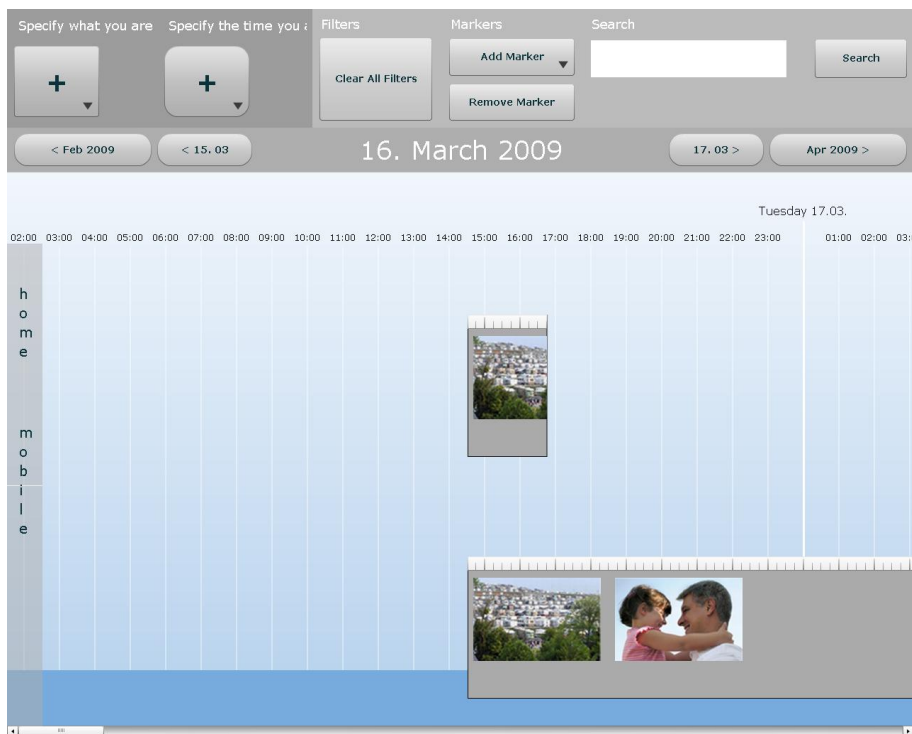


Figure 9: HERMES MyPast: Browsing through time is possible by using the sliding bar (dragging the finger to the left or right) or by using the date buttons to switch to next/previous day or month.



Figure 10: HERMES MyPast: A time filter (e.g. hour range) allows filtering those events that do not match the given criteria. This allows better orientation over the recorded content.



Figure 11: HERMES MyPast: Content-based filters like the people filter allow the users to hide all those recordings, where the person he or she is interested in does not appear.

2.2 Spain

Besides the components evaluated in Austria, in Spain an additional focus was on speech components. Since speech processing is available for Spanish speakers, this part of the system was not evaluated in Austria. Evaluation of speech components includes reading transcribed text, keyword based search, and practicing with the search function of MyPast. However, only searching in MyPast could be used with real functionality. When typing a word in the search box, audio was played back, if the word appeared in a transcribed audio recording from the PDA (see Section 3.1 for more details).

The attempt to transport a multi-touch screen resulted in a broken screen. In order to evaluate the games on location in Spain anyhow, a video of the game has been shown to the users instead of testing the cognitive game itself (Figure 12). While this solution is not optimal, it prevented any further delays. For the second user evaluation the options of building a multi-touch surface on location or buying commercial off-the-shelf products are currently investigated. The video of the first cognitive game developed by AIT was used in a semi-structured interview carried out at INGEMA in order to explore the first impressions of elderly people about this cognitive game.

Unlike the Austrian setup the Spanish hardware setup for the first lab evaluation did not include a touch screen. As it was expected to have the multi-touch surface available, no other touch-screen was present during the user trials. Users were required to imitate touch interaction with a standard computer mouse. As it will be shown later on, this caused some difficulties, while the touch screen itself carried some unique problems.

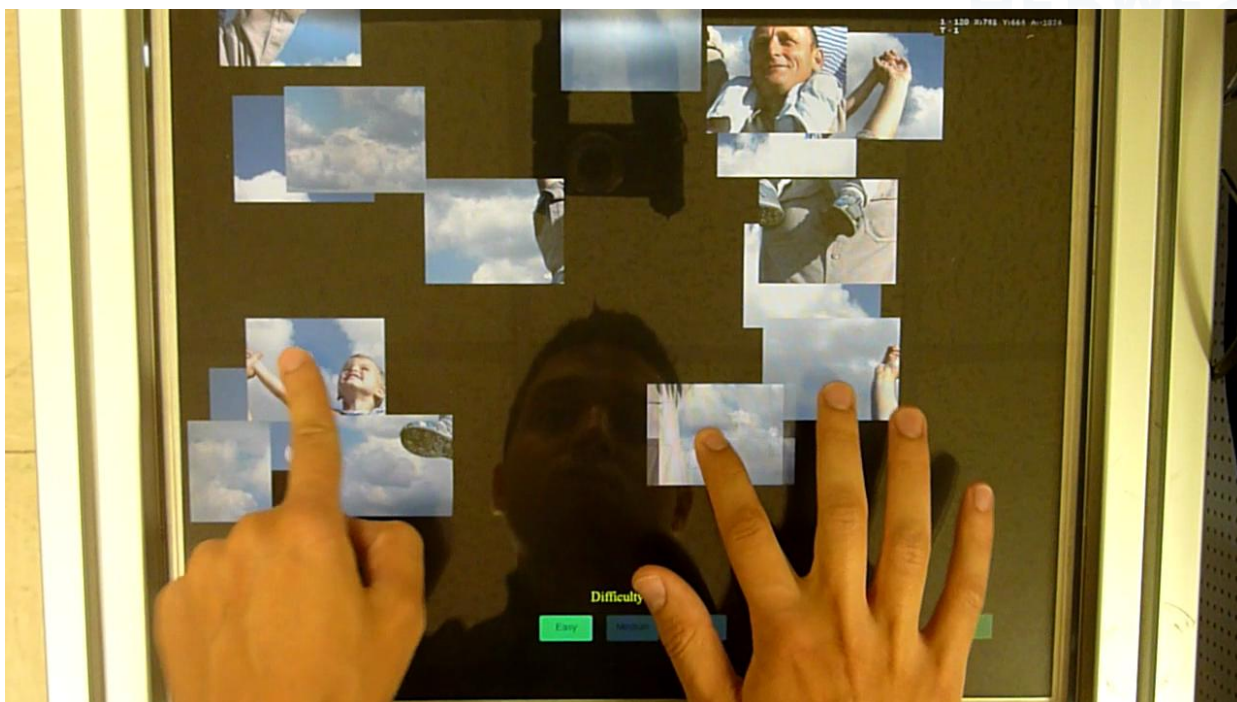


Figure 12: Image of the video of the Cognitive Game that was shown to the test participants at INGEMA. After seeing the video users were asked about their impressions in a semi-structured interview.

3. Procedure of User Trials

User evaluation followed the same procedure both in Austria and in Spain, besides the differences in the tested applications described in Section 2. The procedure, including user recruitment and trial schedules is described in the following sections. The first session included the evaluation of the HERMES MyFuture application and an introduction and first evaluation of the HERMES PDA application. After the first session a number of users – 5 users in Austria and 9 users in Spain – took a HERMES PDA home to test it during the following four to seven days. They returned it in the second session, where feedback of the HERMES PDA was collected as well as HERMES MyPast and in Spain HERMES Cognitive Game were evaluated. The users who did not take the PDA home performed the tasks with the device in the lab.

3.1 Adaptation of the User Evaluation Plan

The goal of the user evaluation was to involve 12 to 16 participants per country in the evaluation. The actual user tests consisted of 13 people in Austria and 14 in Spain. In both countries 16 people were invited, but in the end some couldn't come, because of conflicting appointments or other reasons. In Austria the number of users taking the PDA home was planned to be 6. In the end 5 participants took the PDA home, because 1 person decided not to take it home because it appeared to this user that the device would be too difficult to use. In Spain 9 people took the PDA, 5 did not want to take it. They did not feel confident and they perceived that to take the PDA is something stressful. The participants that did not take the device home (8 participants in Austria and 5 in Spain) performed the tasks with the PDA at the lab. For efficiency reasons, the number of days people kept the PDA at home was reduced from seven to four, in which all planned tasks could be performed. Due to scheduling issues the number of days people actually kept the PDA varied between 3 and 4. In Spain a second PDA was bought and sent to TXT to configure it to use the HERMES application. In Spain 2 people kept the PDA for seven days, 2 people for six days, 4 people for three days and 1 participant only kept it for two days.

Since no multi-touch screen was available for deployment during the first user trials, the Cognitive Game that was developed could only be evaluated by showing the users a video how to play the game. Afterwards the users' opinions about the game were asked in a semi-structured interview.

In Spain the following tasks were planned at the second session to assess accessing recorded and transcribed audio as well as using MyPast search to solve a quiz:

1. Accessing Transcribed and Indexed Audio

- Tutoring and getting feedback on audio playback and reading transcript
- Tutoring and getting feedback on keyword based search
 - Searching for a single word
 - Stemming (reducing inflected or derived words like “going” to “go”), accent tolerance
 - Searching for a combination of words (implicit AND)
 - Searching for a <word1> OR <word2> construction
 - Searching for an exact phrase, e.g. “drink coffee”
 - Searching with complex constructions (<phrase> OR <word1> <word2>)
 - Misspelling and “did you mean” interface

2. Practicing MyPast Search

- Refreshing user's memory about a conversation that the user was involved at the first session and that was recorded and processed by the HERMES system.
- Performing a quiz by referring to a question the user has been asked. The user then should use MyPast Search to find what he or she has answered.

The original intention was to use one year old audio/video recordings for practicing MyPast Search. These recordings have been made with the same older people in 2008 during the data collection task in the framework of the user requirements study. However, it proved that significant effort would have been needed to transfer the video material to the deployed system in a realistic timeframe. Due to the different data structure of the database adaptations to the video meta data would have been necessary. Moreover, this bore the risk of inconsistency in the database upon transferring the material. Finally it was decided not to transfer the material. In order to allow personalized data to be revealed to the users in HERMES MyPast, as well as to allow adaptation to the speaker to be performed, the first session was used to record user data, which was accessed through MyPast in the second session.

This also implicated that MyPast (which was scheduled for the first session) was evaluated in the second session, while MyFuture was moved from the second to the first session in order to balance the length of the sessions. Due to technical problems in processing the recorded videos, the material recorded for some users during the first session (when MyFuture was evaluated) was not available. However, while the thumbnails of this material were shown in the timeline, it was not possible to tap them and to play back the video. Besides that, accessing of transcribed audio that was recorded on the home system was not available. This implicated that using the search box to solve a quiz by typing a keyword used in the conversation could not be performed. Therefore this part of the evaluation plan was omitted. It was, however, possible to test playback of transcribed audio that was recorded with the PDA. When a word that appeared in the transcription of PDA recordings was typed in the search box, the corresponding audio recording was played back. When it will be possible to test the other speech related components, INGEMA will contact the participants of this first trial and assess these tasks.

Table 2 details description and status of the HERMES components. The status of the components reflects how they were experienced during the first user trials. "Not evaluated" means that this back-end component is working but was not assessed since it was not relevant for the user trials.

Technology Name	Description	Status	Main Partner(s) involved
Person / Face Identification	Identifies Actors entering HERMES space	Not evaluated. Only dummy persons are present in the MyPast user interface	AIT
People Counting (Visual)	Counts People within the room	Not evaluated	AIT
Speech Info Indexing & Retrieval	Search in spoken data	Working	IBM
Voice Activity Detection	Detects Human Speech	Scheduled for second prototype.	AIT

Visual Person Tracking	Tracks people movement within the HERMES space (in-door)	Working	AIT
Speech-to-text	Transcription of spoken conversations (Spanish only)	Working. Transcriptions of appointments and conversations recorded on the PDA available	IBM
MyPast Application	Presents the user interface functionalities in a GUI to the user and allows the user to retrieve and edit information that has been captured by the input processing components (as described in the deliverable D.5.1)	Working with some functionality limitations	CURE/UniBrad
MyFuture Application	Presents a user interface that allows the user to <ul style="list-style-type: none"> • set appointments • store notes as addition to appointments • edit appointments • synchronisation with HERMES PDA imports appointments from the PDA and exports appointments to PDA (within a defined time interval) 	Working Working Working Working in Spain Not working in Austria (deployment issue)	CURE/TXT
Fingertips tracking for the Multi-Touch Surface	Track users' fingertips to enable multi-touch surface interface	Not evaluated	AIT
Cognitive Games Application	Cognitive Training Games Logic (as Defined in D.6.1)	1 game developed, video of game evaluated	AIT & INGEMA
Mobile POI Manager	<ul style="list-style-type: none"> • GPS integration • Possibility to add POIs basing on the current location 	Working	TXT

Mobile device synchronisation	<p>Synchronisation with the home-based workstation with focus on:</p> <ul style="list-style-type: none"> • Available audio transcriptions related to audio notes attached to appointments • Audio recordings for Conversation Support. • New/Updated stored POIs (using the “locate mode”) • Appointments (to and from the PDA and the home-based workstation) • Audio notes attached to the appointments 	<p>Working in Spain</p> <p>Not evaluated</p> <p>Not evaluated</p> <p>Working in Spain</p> <p>Working in Spain</p>	TXT
Mobile Shopping List Manager	Allow for the user to manage shopping lists includes GPS integration (the user is alerted when close to the food store)	Scheduled for second prototype	TXT
Speech-to-text	Transcription of spoken appointments to make them searchable and to make to display them in textual form	Working. Transcriptions of appointments and conversations recorded on the PDA available	IBM
Speech info indexing	The speech transcripts are stored in indexed form for fast search	Working	IBM
Image/Video Tagging and Retrieval	Stores images/video with appropriate tags / Retrieves video segments based on queries	Scheduled for second prototype	AIT
Audio recorder	<ul style="list-style-type: none"> • The mobile application allows the user to record the audio note and attach it to a chosen event (Appointment Dictation) • The mobile application allows also to record a discussion (Conversation Support) 	Working	TXT

Hermes PDA Application	The mobile application allows the user to <ul style="list-style-type: none"> • insert/delete/edit/browse /notify appointments (including information like date, time, notes, POIs, etc.) • record discussions • attach audio notes to appointments • display post processed speech transcriptions • take pictures • browse HERMES-registered people profiles • manage shopping lists 	Working Working Working Scheduled for next prototype Scheduled for next prototype Scheduled for next prototype	TXT
Video summarisation	Summarisation of videos through frame analysis	Scheduled for next prototype	UniBrad

Table 2: Description and status of components of the first HERMES prototype. While all future and not yet developed components of HERMES are present in this table, their status is “Scheduled for next prototype”. Components that don’t reveal their functionality directly to the user have the status “Not evaluated”. Assessed components either have a “Working” of “Not working” status.

3.2 User Recruitment

In Austria the tests were carried out with 13 potential HERMES users (5 males and 8 females, with an age average of 71.92 years, from 62 to 86 years old) and in Spain 14 potential HERMES users (7 males and 7 females, with an age average of 69.57 years, from 63 to 76 years old) from the target group were involved.

To be included in the user trials the following inclusion criteria had to be met by potential HERMES end users:

- users are over 60 years of age,
- either without cognitive impairment or diagnosed with AAMI or MCI,
- do not suffer from any severe sensorial and/or motor problems, and
- are living independently in their own homes.
-

Additionally, the main target group of HERMES is people who have retired. However, it was not a specific inclusion or exclusion criterion.

Both CURE and INGEMA had a pre-existing pool of test participants that agreed at the beginning of the project to belong to the target group and to participate in the several trials with end users in the project when required. For this trial, all the participants in Spain were recruited from this pool. In Austria there were not enough people from the existing pool of HERMES test persons available, since a number of them didn’t want to participate (due to physical or mental decline) or could not be reached anymore. Therefore, additional 4 older people were recruited who met the criteria mentioned above. Since the recorded material from the data collection

during requirement analysis phase was not available in the integrated prototype (as described above), to recruit the same persons selected last year was not critical.

3.3 First Session

The goal of the first session was to evaluate HERMES MyFuture Application and to introduce the user to the HERMES PDA. Following a general introduction, the evaluated applications were presented to the user. The first session also included a pre-evaluation of the users' attitude towards computers and the presentation of the test material (questionnaires).

3.3.1 Pre-evaluation

The users were asked questions regarding their general attitudes towards computers as proposed in the Computer Attitude Scale (CAS) by Nickell et al. (1986). The rating of the given statements follows a 1-5 Likert scale, where 1 means "strongly disagree", 2 "disagree", 3 "undecided", 4 "agree", and 5 "strongly agree". The statements are listed in Table 3.

Item
1 The Computer will never replace human life.
2 The Computer makes me uncomfortable because I don't understand it.*
3 People are becoming slaves to the Computer.*
4 The Computer is responsible for many good things we enjoy.
5 Soon our lives will be controlled by the Computer.*
6 I feel intimidated by the Computer.*
7 There are unlimited possibilities of Computer applications that have not been thought of yet.
8 The overuse of the Computer may be harmful and damaging to humans.*
9 The Computer is dehumanizing to society.*
10 The Computer can eliminate a lot of tedious work.
11 The use of the Computer is enhancing our standard of living.
12 The Computer turns people into just another number.*
13 The Computer is lessening the importance of too many jobs done now by humans.*
14 The Computer is a fast and efficient means of gaining information.
15 The Computer's complexity intimidates me.*
16 The Computer will replace the working human.*
17 The Computer is bringing us into a bright new era.
18 Soon our worlds will be run by the Computer.*
19 Life will be easier and faster with the Computer.
20 The Computer is difficult to understand and frustrating to work with.*

Table 3: Computer Attitudes Scale (CAS) (1-5 Likert scale) based on Nickell et al. (1986). The purpose of the questionnaire is to assess the general attitude and anxiety of elderly people towards the computer. Note: * indicates reversed scale.

3.3.2 Tasks and Questions for HERMES MyFuture

HERMES MyFuture application was tested by giving users several tasks and evaluating users opinions afterwards with a mixed questionnaire (open and closed questions), which is listed below. Each task description was read to the user by the tester and was given in a written form. For each task questions from the test person and reported problems were noted down in a structured way. At the beginning of the trials it was agreed to time the tasks, but after some trials the time needed for completing the tasks was not evaluated for the following reasons:

- The users were not aware they were timed performing the task, so they started with the tasks, but eventually they stopped, talked to the interviewer about the difficulties they found or other ideas they had. The interviewers tried to focus the attention of the users on the tasks again, but some of them were talking at the same time they were performing the task.
- Since the user interface of the first HERMES prototype is in English, the users needed more time because they needed the interviewer to translate the words and messages on the screen in all the applications while they were doing the tasks.

For each task, there were several questions asked to the test person. The following text was read to the user:

By using the HERMES MyFuture application you can record several events and you can organise them. The buttons on the upper part help you to change the date of MyFuture application. By clicking on the date you can see the related events on this date. In the calendar you will see that on the top of some days there are lines. These lines indicate the event number on this day. If the line is longer, it means that you have more events in this day. Now we are kindly asking you to go through the following tasks with the system.

Afterwards, the user had to perform the following tasks:

1. **Standard browsing /finding events:** Please go to next month (e.g. September 2009) and try to find the day, on which you have the most events.
 - How many events did you find?
 - What are the next events in your calendar and when do they take place?
2. **New entry:** Please go to today in order to enter new events. Please add the following event to your list.
 - Doctor Appointment, 18:05, 1010 Vienna (or San Sebastian)Please save another appointment for the same day.
 - Meeting 20:00, 1060 Vienna (or San Sebastian)
3. **Move entry:** Please move the entry you made to another date (XX.XX.) by dragging and dropping it.

While performing the tasks the users were observed by a usability expert tracking usability problems they were encountering. After each task the user was asked to rate the given statements on a 5-point Likert scale and to answer the following open questions about the completed task in HERMES MyFuture:

- 5-point Likert scale (1, strongly disagree – 5, strongly agree)
 1. The system supported me well in fulfilling this task.
 2. The system can be improved in supporting the fulfilling of the task.
 3. It is easy to understand the icons displayed on the screens.
 4. I was able to follow the changes on the screen easily after tapping it.
 5. I felt comfortable doing this task.
- Open questions following statement 2:
 6. What can be improved to better support you with this task?
 7. What specific difficulties have you identified by completing this task?
- Open question following statement 3:

8. If you had problems, what was difficult, apart from the English language?
 - Open question following statement 4:
9. If you had problems following the changes, what was difficult?
 - Open question following statement 5:
10. If you had and special feelings by completing this task, which ones?

3.3.3 Technology Acceptance Questionnaire for HERMES MyFuture

After performing the tasks with HERMES MyFuture and filling out the questionnaire, users were asked to rate several general statements about HERMES MyFuture. The 31 questionnaire items shown in Table 4 are adapted from the Unified Theory of Acceptance and Use of Technology (UTAUT) study of Venkatesh et al. (2003). For an in-depth description see Section 2.1.1 in the deliverable D.7.1. Other than wording modifications to fit the specific technology studied in this research, no changes were made to the user acceptance scale. All items were measured on a 7-point Likert scale (1, completely disagree – moderately disagree – somewhat disagree – neutral (neither disagree nor agree) – somewhat agree – moderately agree – 7, completely agree).

Scales / Items
Performance Expectancy (PE)
PE1: I find HERMES MyFuture useful in my life.
PE2: Using HERMES MyFuture enables me to accomplish tasks more quickly.
PE3: Using HERMES MyFuture increases my productivity.
PE4: Using HERMES MyFuture increases my chances of leading an active lifestyle.
Effort Expectancy (EE)
EE1: My interaction with HERMES MyFuture is clear and understandable.
EE2: It is easy for me to become skilful at using HERMES MyFuture.
EE3: I find HERMES MyFuture easy to use.
EE4: Learning to operate HERMES MyFuture is easy for me.
Attitude toward Using Technology (AT)
AT1: Using HERMES MyFuture is a good idea.
AT2: HERMES MyFuture makes life more interesting.
AT3: Living with HERMES MyFuture is fun.
AT4: I like living with HERMES MyFuture.
Facilitating Conditions (FC)
FC1: I have the resources necessary to use HERMES MyFuture.
FC2: I have the knowledge necessary to use HERMES MyFuture.
FC3: HERMES MyFuture is not compatible with other systems I use.*
FC4: A specific person (or group) is available for assistance with HERMES MyFuture difficulties.
Anxiety (AX)
AX1: I feel apprehensive about using HERMES MyFuture.
AX2: It scares me to think that I could lose a lot of information using HERMES MyFuture by hitting the wrong key.
AX3: I hesitate to use HERMES MyFuture for fear of making mistakes I cannot correct.
AX4: HERMES MyFuture is somewhat intimidating to me.

Behavioural Intention to Use the System (BI)
BI1: I intend to use HERMES MyFuture in the next semesters if I would have access to it.
BI2: I predict I would use HERMES MyFuture in the next semesters if I would have access to it.
BI3: I plan to use HERMES MyFuture in the next semesters if I would have access to it

Table 4: Technology acceptance questionnaire for HERMES MyFuture (1-7 Likert Scale). It's intention is to assess the users' acceptance of the functionality offered by MyFuture. Note: * indicates reversed scale.

3.3.4 Tasks and Questions for HERMES PDA

In order to prepare the user to use the PDA at home, a training session was held in the lab, which introduced the user to the options available and also included a questionnaire. The following text was read to the user:

The HERMES PDA that we have given to you can record reminders when you are away from home. For example, if you are on the street and you just remembered that you have to call your sister when you get home, you can take the PDA and record a message to remind you.

In the future, with the HERMES system fully developed (by the end of 2010), you will be able to download these reminders when you get home. "Download" means that the reminders that you have recorded away from home will be transferred automatically to your home-computer and it will remind you of events.

After this introduction a practice session with the HERMES PDA was held. This session included the following tasks:

1. **Introduce an appointment:** Please get into HERMES PDA, go to Appointments, and try to introduce a new entry and to specify its time. Select "No location" when the system asks you about the location of the appointment.
2. **Modify the appointment:** Please get into the appointment you have just introduced. Please edit "NO LOCATION" by selecting RIGHT HERE and then ADD CURRENT LOCATION. Remember that you have to wait some minutes until the GPS has coverage and this function is enabled.
3. **Introduce an appointment in Locate Mode:** Please get into HERMES PDA, go to Appointments and introduce a new entry. Now, select LOCATE MODE and record your appointment for this place.
4. **Browsing through the appointments:** Now you have recorded an appointment and you have some appointments recorded beforehand. Please go through this month and tell how many appointments you have for this month.
Can you go to the first appointment you have and tell the details of this appointment?
5. **Conversation.** Please get into HERMES PDA, go to Conversation and record a piece of speech after pressing the red button. Stop it by pressing the yellow button and listen to it pressing the dark blue one.

After each task the user was asked to answer the same quantitative and qualitative questions as described above for HERMES MyFuture (Section 3.3.2). Again users were asked to rate the given statements on a 5-point Likert scale about the completed task. While the users who didn't take the PDA home answered all the questions, the users who took the PDA home were only asked to rate the first statement ("The system supported me in fulfilling this task"). The other questions were assessed in the second session (Section 3.5).

After this practice, if the user felt confident, she or he took the PDA with her or him in order to use it at home. Besides the PDA, an instruction manual developed by INGEMA and CURE was given to the user. In the manual all the tasks that can be done with the PDA and the steps necessary to do them were specified. Also one section of the manual refers to the ethical issues about having a PDA being able to record other conversations and the fact that both CURE and INGEMA do not want the participant to record real appointments or conversations (these can be fictitious), and do not want any recordings of other people. Finally a very simple list of "Tasks to Be Done at Home" was also given to the users as it can be seen below.

3.4 Tasks to Be Done at Home

Depending on available time resources the users kept the PDA between two and seven days. Accordingly, a list with five tasks plus one or two free use options distributed over four or seven days was given together with a HERMES PDA to the voluntary participants. The tasks were the same for all the users, the difference was that for those users who kept the PDA for a week did one task each day and the ones who had the PDA for less than a week performed more than one task every day. During the time at home the user performed the tasks and filled out the form. When returning back the PDA the experiences during the days they had the PDA were discussed with the user and a second session with the PDA was performed, as described in Section 3.5.

	Task	Felt confident	Felt comfortable	Difficulties
Day 1	Introduce an appointment	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	
	Record a conversation	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	
Day 2	Edit an appointment (change the time)	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	
	Edit an appointment (introduce location)	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	
Day 3	Introduce an appointment with locate mode. Move away and go back to the former location, in order to try the function.	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	
Day 4	Optional free use	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	

Table 5: A task sheet containing this table with tasks for every day was given to the users that kept the PDA four or three days. The table originally filled a whole A4 sheet to give the user enough space to write comments.

	Task	Felt confident	Felt comfortable	Difficulties
Day 1	Free use	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	
Day 2	Introduce an appointment	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	
Day 3	Edit an appointment (change the time)	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	
Day 4	Edit an appointment (introduce location)	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	
Day 5	Introduce an appointment with locate mode. Move away and go back to the former location, in order to try the function.	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	
Day 6	Record a conversation	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	
Day 7	Free use	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	

Table 6: A task sheet containing this table with tasks for every day was given to the users that kept the PDA seven days. The table originally filled a whole A4 sheet to give the user enough space to write comments.

3.5 Second Session

When the user returned the PDA, a second evaluation session was held, beginning with questions about the use of the PDA at home. Afterwards several tasks were performed with the HERMES MyPast application, followed by an evaluation of general statements about HERMES MyPast. Tasks and questionnaire followed the same structure as the test with HERMES MyFuture during the first session (Section 3.3). In Spain the HERMES Cognitive Game evaluated was also, specifically the video of the Puzzle Game. This is a five-minute video in which the participants can see a person playing the game on a multi-touch screen. After the video was shown a semi-structured interview was carried out.

3.5.1 Questions about HERMES PDA

In the second session some questions about the usage of the PDA at home were asked, as well as the remaining questions from Section 3.3.4, because it was hoped the users will give more valid feedback after using the device for several days. The tasks the users were asked about where the same they tried to perform at home:

1. Introduce an appointment
2. Modify the appointment
3. Introduce an appointment in Locate Mode
4. Browsing through the appointments
5. Record a Conversation

3.5.2 Tasks and Questions for HERMES MyPast

Like HERMES MyFuture, the HERMES MyPast application was tested by giving users several tasks and evaluating their opinions afterwards. Each task description was read to the user by the tester and was given in a written form. Again, questions from the test person and reported problems were noted down in a structured way. Afterwards a mixed questionnaire, which is listed below, was used to evaluate user opinions. The following text was read to the user:

You are now on HERMES MyPast application. HERMES MyPast is a reminder application where you can record data including videos and photos and browse your recorded data by using different methods. HERMES MyPast helps you to search your photos, videos, voice recordings with your own selected criteria. Now we are kindly asking you to go through the following tasks with the system.

Afterwards, the user had to perform the following tasks:

1. **Standard browsing / sliding bar:** By using the sliding bar on the bottom of HERMES MyPast main window you can browse in HERMES MyPast application and you can go through the events randomly in a month, season or even a year without searching a particular event. By this way you can see a smaller picture or a sign for a recorded event. Please try to go between XX.XX and XX.XX⁵ by using this option.
 - How many events have you found?
 - How long is the longest event you recorded?
 - How many events did you find on the XX.XX?
 - Can you say how long Paul's visit on the XX.XX was?⁶
2. **Using buttons for changing date:** You can also use the buttons on the left and right sides of date on the upper part of HERMES MyPast application. To go faster in a period during your search or to go to a particular date you can use these buttons. Now please try to go to date XX.XX by using this option.
3. **Specifying time:** If you know the exact time you are searching for you can use "specifying time option" to narrow down the results by using the filters on the upper side. Please specify the following time period by using "specifying time" option. Please set the time filter for 9:00-16:00 and please specify the time period as spring.⁷
4. **Specifying people:** By using HERMES MyPast application you can "specify what you are looking for". For this you can use the button on the upper left of the screen. Please use this option to specify your search for searching people and please order the results according to their importance. And then choose "frank" to search for. MyPast will show your results in selected time related to frank. By using same filters you can also specify the place where you and Frank met, please open places filter and select "central Park" to narrow down your search about Frank.⁸
 - Can you say how many filters have you used to reach to the results?
 - What other possibilities do you think that you have to do the similar search?
5. **Clearing/changing filters:** If you do not reach the search results you can change or clear some of your filters. Please clear the place filter for your search. Please change the timeline from 9:00-16:00 to 9:00-19:00. You can also clear all the filters you created by using "clear all filters button". Please clear all filters.
6. **Using search box:** You can also conduct similar search by using the search button on the right top of the screen. Please use this search option to search for a specific word used in the previous conversation. You can also add a marker – a reminder that gives an idea about where to look at your search. Please add a health marker to your search.⁹

⁵ While the task description indicated to set a visible time frame of more the one day, in the user interface the time frame was fixed to 24 hours. It was not possible to scale or zoom the amount of time visible.

⁶ While the four questions were planned to be asked to the user, none of them actually made much sense. There was only one event available (the one from the first session) and users couldn't see how long the event was. Moreover, there was no event with Paul's visit.

⁷ As described in Table 2 this was not working. Applying a time filter did not have any effect.

⁸ As described in Table 2 this was not working. Applying a people or place filter did not have any effect.

⁹ As described in Table 2 this was only partially working. Conducting a search did only give results for audio recordings that were recorded on the PDA.

- Did you have difficulties when using the touch keyboard?
- Did you find the “Umlaut mark” easily?
- Were you able to switch to lowercase?

After each task the user was asked again to rate statements and to answer the open questions about the completed task in HERMES MyPast already mentioned when describing the procedure for MyFuture (Section 3.3.2).

3.5.3 Technology Acceptance Questionnaire for HERMES MyPast

As with HERMES MyFuture, after performing the tasks with HERMES MyPast, users were asked to rate several general statements about HERMES MyPast. The 31 questionnaire items are the same as already shown for HERMES MyFuture in Table 4. The only changes made was replacing the term “MyFuture” with “MyPast”. Again all items were measured on a 7-point Likert scale (1, completely disagree – moderately disagree – somewhat disagree – neutral (neither disagree nor agree) – somewhat agree – moderately agree – 7, completely agree).

3.5.4 General Questions about HERMES Cognitive Game

The cognitive game tested, HERMES Puzzle, uses pictures related to appointments already stored at the HERMES database. The game initially presents a photo album on the multi-touch screen. The user is able to browse among the photos using simple gestures to move, rotate, shrink or magnify the photos. The puzzle game is initiated when the user double-taps on the photo of his preference. The picture gets fragmented on several pieces that have to be merged in order to complete the puzzle. The way in which the user changes from an easy difficulty level to one level more difficult is shown. At the end of the game, the score obtained by the participant is shown on the screen.

HERMES Puzzle was evaluated through a semi-structured interview at INGEMA. The reason for using this methodology was to get a first impression of elderly people about the game. The following questions were asked to all of the participants:

- Do you think that the game looks interesting?
- Do you think that the game seems tricky?
- Do you think that the game could be boring?
- Do you think that this game would be effortful?
 - Do you think that it would be frustrating?
 - Do you think that it would be stimulating?
- Do you think you would be engaged with this game?
- Is it clear how to select the picture to play with?
- Is it clear how to change to another different picture?
- Are the levels of difficulty clear enough to be selected?
- Are the levels of difficulty complex enough?
- Is the score-screen clear enough?
- Is the score-screen motivating enough?
- Do you feel that this game would be helpful in your social relationships?
 - Would you play the game with your family or friends at home?
 - Would you play the game with your grandchildren (or against other children)?
 - Would you play the game on-line with other players?

4. Results

The user trials performed gave valuable insights in the aspects of the HERMES System that need further improvement and refinement. This section describes the results obtained from the user trials as detailed in Section 3. In particular it covers the:

- results of the CAS questionnaire
- results of the technology acceptance scale for HERMES MyFuture and MyPast
- for each HERMES component (MyFuture, MyPast and PDA):
 - quantitative evaluation of post-task questionnaire
 - qualitative evaluation of post-task questionnaire
 - usability evaluation
- evaluation of the first HERMES Cognitive Game
- evaluation of the cartoon for the avatar
- conclusions that can be drawn from each part of the evaluation

While quantitative and qualitative evaluation of the post-task questionnaires reflect solely the opinion of the user, the sections titled “Usability Evaluation” contain results from the observation of the user and the think-aloud-protocols obtained while they were performing the tests. The severeness of the usability problems found are rated on a 4-level scale: “serious”, “important” and “marginal” problems, supplemented by adjustments or add-ons that are “nice to have”.

1. Serious Problems

- are expected to prevent successful task complementation
- lead to abortion of user interaction
- an adjustment is essential

2. Important Problems

- successful task completion is difficult
- abortion of user interaction is probable
- an adjustment is highly recommended

3. Marginal Problems

- task completion is displeasing
- an adjustment is advisable

4. Nice to have

- an adjustment will cause more contentment

Important note: While from a pure usability point of view so-called “marginal” problems might seem less important, since they do not prevent the system of being “usable”, from a user experience point of view they usually have a great impact. A system that is displeasing to use will not be accepted by the users or will be the subject of constant complaint and discomfort. Therefore it is as important to solve “marginal” problems as it is to solve serious and important problems. This also applies, although to a smaller extent, to “nice to have” adjustments.

The usability results are categorised by the type of usability problem. The following list explains what kind of problem each category represents:

- **Interaction:** While all problems could be seen as interaction problems, because they hinder the interaction of the user with the system, some problems are specifically related

to way the interaction is structured. Interaction problems can be the way of interaction with the system in general (e.g. using a mouse or the on-screen keyboard) or the way the system determines certain tasks need to be done (e.g. the sequence of buttons you need to press in order to apply a filter in MyFuture).

- **Metaphor:** Metaphor problems are always serious problems, because they arise when certain metaphors that might even be commonly used in a computer context (like Drag & Drop) are not understood by older users.
- **Language:** This single problem arises when users have problems understanding the English language. Certain other usability problems might be related to or even caused by language problems.
- **Wording:** This category covers all problems that are related to bad phrasing or choice of words. While the problems can be quite serious they can usually be solved easily (e.g. changing the caption of a button).
- **Content Layout / Structure:** Often problems arise when the structure of information and interaction elements is inconsistent or illogical. Re-arranging, adding or removing interface elements is necessary.
- **Navigation:** This type of problem often goes together with structure, but navigation problems refer less to a static context but more to the flow of different screens.
- **Affordance:** Problems with affordance are often connected to metaphor problems. If users don't know certain interaction techniques like dragging an object across the screen, this object needs to tell the user somehow how he or she can interact with it.
- **Feedback:** Problems of this category arise when the system does not provide clear enough feedback of what it is doing or what it wants the user to do. This is especially important for elderly users as they lose track of their interaction easier.
- **Playability:** Playability means robustness, but not in a technical way. Users should be allowed to play around with the system without the risk of losing data (and crashing the system, of course).
- **Icons:** A common problem with older users is that they are unfamiliar with icons used in technological contexts.
- **Colours:** Colours can be confusing (by ignoring their signalling effect (e.g. red means "stop")) or can be too pale or simply disliked.

The usability problems presented in the next sections might give a negative impression of the system due to their quantity. However, it needs to be considered that the evaluated system is a first prototype. This means, that several features are still missing that affect the usability and acceptance of the system (see Sections 4.2 and 6 for more details). Additionally many problems listed separately in the tables do in fact come from the same source and can be solved with a single solution, as indicated in the "Proposed solution(s)" column on several occasions.

The following sections present the results collected both in Austria and Spain. In the cases where the results reflect insights gained only from Spain or Austria, it is noticed in the text. Where both INGEMA and CURE came to the same results, no further notice is given.

4.1 CAS questionnaire

The purpose of the Computer Attitude Scale was to evaluate the positive and negative attitudes towards computers the elderly, who tested HERMES, had about the computer in general (Nickell & Pinto, 1986; Venkatesh et al., 2003). The results for Austria and Spain are presented in Table 8. For each row, the cell, where the median is crossed, is highlighted either in green or orange colour. If the majority of the users are below median (i.e. disagree) the cell is coloured in orange, otherwise in green. If the median is crossed a middle cell, the colour is defined by which

extreme to the left or right has a higher percentage. If the distribution to the left and right is exactly even, the middle cell is coloured in grey. For statements with reversed scale (indicated by an asterisk) the colour scheme is reversed as well.

Looking at the individual statements, the majority of older users do not feel intimidated by the computer (81%). For some of them, the computer makes them feel uncomfortable because they do not understand it (33%). Interestingly, many of them do not agree to the statement that the computer will never replace human life (47%). Opinions are varying about the statement whether our lives will *never* be controlled by the computer or not. However, the differences are not significant (total agreement 59%). Users agree with the statement that overuse of the computer may be harmful and damaging to humans (85%). The majority of older Austrians and Spanish do not feel intimidated by the computer (81%) and agree there are possible applications of computers that have not been thought of yet (88%). They think the computer can eliminate a lot of tedious work (92%) and that it is a fast and efficient means of gaining information (100%). They also think that the computer will not replace the working human (73%). However, many of them think that the computer is lessening the importance of too many jobs done now by humans (67%).

The overall CAS questionnaire score can range from 20 to 100 with a neutral position at 60. The average score for the older users tested was 68.8 with a standard deviation of 7.74. This means that the participants have a slightly positive attitude towards technology.

When comparing the older Austrians and Spanish, we find that the average Austrian CAS score is 71.3 (SD=6.26) and the Spanish score is 66.9 (SD=8.44). There is no significant difference between these scores. For the individual statements Mann-Whitney Tests show the following significant differences between age, gender and country:

- Users aged over 70 years are more in agreement with the statements: “the computer can eliminate a lot of tedious work” ($p<0.01$) and “the computer is bringing us into a bright new era” ($p<0.01$) than users aged up to 69 ($M=17$ vs. 8.73 and 17.47 vs. 8.09, respectively).
- Older men are more in agreement with the statement: "The computer is responsible for many good things we enjoy" than older women ($p=0.01$). The means are 18.50 and 10.40, respectively.
- With respect to the differences between both countries, Spanish older users are more in agreement with the statements: "People are becoming slaves to the computer" ($p<0.01$) and “The overuse of the computer may be harmful and damaging to humans” ($p<0.01$) than Austrian users ($M=17.46$ vs. 10.26 and 17.89 vs. 9.81, respectively).

Item	Agreement in %														
	Austria					Spain					Total				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1 The Computer will never replace human life.	7.7	15.4	0	23.1	53.8	50	0	7.1	7.1	35.7	29.6	7.4	3.7	14.8	44.4
2 The Computer makes me uncomfortable because I don't understand it.*	23.1	61.5	7.7	7.7	0	35.7	14.3	7.1	14.3	35.7	29.6	37	7.4	14.8	11.1
3 People are becoming slaves to the Computer.*	7.7	30.8	46.2	15.4	0	7.1	7.1	21.4	35.7	28.6	7.4	18.5	33.3	25.9	14.8
4 The Computer is responsible for many good things we enjoy.	15.4	0	23.1	38.5	23.1	0	7.1	21.4	50	21.4	7.4	3.7	22.2	44.4	22.2
5 Soon our lives will be controlled by the Computer.*	7.7	38.5	30.8	23.1	0	21.4	42.9	7.1	14.3	14.3	11.1	25.9	18.5	33.3	11.1

6 I feel intimidated by the Computer.*	30.8	46.2	15.4	7.7	0	64.3	21.4	7.1	7.1	0	48.1	33.3	11.1	7.4	0
7 There are unlimited possibilities of Computer applications that have not been thought of yet.	0	0	8.3	50	41.7	0	7.1	7.1	21.4	64.3	0	3.8	7.7	34.6	53.8
8 The overuse of the Computer may be harmful and damaging to humans.*	0	7.7	23.1	53.8	15.4	0	0	0	35.7	64.3	0	3.7	11.1	44.4	40.7
9 The Computer is dehumanizing to society.*	15.4	15.4	46.2	23.1	0	14.3	7.1	14.3	42.9	21.4	14.8	11.1	29.6	33.3	11.1
10 The Computer can eliminate a lot of tedious work.	0	0	0	53.8	46.2	7.1	0	7.1	50	35.7	3.7	0	3.7	51.9	40.7
11 The use of the Computer is enhancing our standard of living.	7.7	0	15.4	38.5	38.5	7.1	0	7.1	42.9	42.9	7.4	0	11.1	40.7	40.7
12 The Computer turns people into just another number.*	23.1	38.5	15.4	23.1	0	21.4	7.1	28.6	35.7	7.1	22.2	22.2	22.2	29.6	3.7
13 The Computer is lessening the importance of too many jobs done now by humans.*	7.7	15.4	23.1	38.5	15.4	7.1	0	14.3	50	28.6	7.4	7.4	18.5	44.4	22.2
14 The Computer is a fast and efficient means of gaining information.	0	0	0	30.8	69.2	0	0	0	21.4	78.6	0	0	0	25.9	74.1
15 The Computer's complexity intimidates me.*	15.4	38.5	38.5	0	7.7	57.1	14.3	14.3	7.1	7.1	37	25.9	25.9	3.7	7.4
16 The Computer will replace the working human.*	7.7	61.5	23.1	0	7.7	35.7	21.4	7.1	28.6	7.1	22.2	40.7	14.8	14.8	7.4
17 The Computer is bringing us into a bright new era.	7.7	23.1	46.2	7.7	15.4	7.1	14.3	14.3	42.9	21.4	7.4	18.5	29.6	25.9	18.5
18 Soon our worlds will be run by the Computer.*	7.7	15.4	38.5	30.8	7.7	21.4	7.1	14.3	50	7.1	14.8	11.1	25.9	40.7	7.4
19 Life will be easier and faster with the Computer.	0	0	15.4	69.2	15.4	0	7.1	21.4	35.7	35.7	0	3.7	18.5	51.9	25.9
20 The Computer is difficult to understand and frustrating to work with.*	8.3	41.7	33.3	8.3	8.3	21.4	35.7	21.4	7.1	14.3	15.4	38.5	26.9	7.7	11.5

Table 7: Results of the CAS (1, strongly disagree – 5, strongly agree; Likert scale). Positive attitudes are highlighted in green, negative ones in orange. The overall score ranks at 68.8 with a standard deviation of 7.74. This indicates a slightly positive attitude towards the computer. Note: * indicates reversed scale.

4.2 Technology Acceptance

There were two Technology Acceptance Questionnaires (TAQ) handed out to the users, one after performing the tasks with HERMES MyFuture, and one after assessing HERMES MyPast. The results of both of them are discussed in this section.

4.2.1 Results

Table 8 lists the results of the TAQ for the MyFuture application, and Table 10 the results of the TAQ for MyPast. The colour codes used are the same described in Section 4.1 with the difference that positive and negative tendency is determined by the higher and lower three ratings respectively. The results indicate that technology acceptance in general is good or rather good, especially for MyFuture. MyPast is confronted with some clear rejection regarding the users' intention to use it. This issue is discussed in detail in Section 4.2.2 (Fehler! Verweisquelle konnte nicht gefunden werden.)

Scale/Item	Agreement in %																				
	Austria							Spain							Total						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Performance Expectancy (PE)																					
PE1: I find HERMES MyFuture useful in my life.																					
	15.4	7.7	0	0	30.8	23.1	23.1	0	7.1	0	7.1	28.6	21.4	35.7	7.4	7.4	0	3.7	29.6	22.2	29.6
PE2: Using HERMES MyFuture enables me to accomplish tasks more quickly.																					
	15.4	7.7	7.7	15.4	0	30.8	23.1	14.3	7.1	7.1	21.4	21.4	21.4	7.1	14.8	7.4	7.4	18.5	11.1	25.9	14.8
PE3: Using HERMES MyFuture increases my productivity.																					
	15.4	23.1	0	7.7	15.4	23.1	15.4	21.4	0	0	21.4	28.6	28.6	0	18.5	11.1	0	14.8	22.2	25.9	7.4
PE4: Using HERMES MyFuture increases my chances of leading an active lifestyle.																					
	30.8	7.7	23.1	0	23.1	7.7	30.8	21.4	7.1	0	35.7	14.3	21.4	0	25.9	7.4	0	18.5	18.5	14.8	14.8
Effort Expectancy (EE)																					
EE1: My interaction with HERMES MyFuture is clear and understandable.																					
	15.4	7.7	0	15.4	7.7	38.5	15.4	7.1	0	21.4	7.1	14.3	21.4	28.6	11.1	3.7	11.1	11.1	11.1	29.6	22.2
EE2: It is easy for me to become skilful at using HERMES MyFuture.																					
	7.7	7.7	0	7.7	0	61.5	15.4	7.1	7.1	7.1	7.1	7.1	21.4	42.9	7.4	7.4	3.7	7.4	3.7	40.7	29.6
EE3: I find HERMES MyFuture easy to use.																					
	0	15.4	0	7.7	23.1	38.5	15.4	7.1	0	14.3	0	21.4	14.3	42.9	3.7	7.4	7.4	3.7	22.2	25.9	29.6
EE4: Learning to operate HERMES MyFuture is easy for me.																					
	0	15.4	0	0	30.8	30.8	23.1	7.1	0	14.3	0	14.3	7.1	57.1	3.7	7.4	7.4	3.7	22.2	25.9	29.6
Attitude toward Using Technology (AT)																					
AT1: Using HERMES MyFuture is a good idea.																					
	7.7	15.4	0	7.7	7.7	15.4	46.2	0	7.1	7.1	7.1	21.4	7.1	50	3.7	11.1	3.7	7.4	14.8	11.1	48.1
AT2: HERMES MyFuture makes life more interesting.																					
	0	38.5	0	23.1	7.7	0	30.8	21.4	7.1	7.1	28.6	14.3	21.4	0	11.1	22.2	3.7	25.9	11.1	11.1	14.8
AT3: Living with HERMES MyFuture is fun.																					
	7.7	23.1	7.7	15.4	7.7	0	38.5	21.4	7.1	0	28.6	14.3	14.3	14.3	14.8	14.8	3.7	22.2	11.1	7.4	25.9
AT4: I like living with HERMES MyFuture.																					
	7.7	7.7	7.7	23.1	7.7	23.1	23.1	7.1	14.3	28.6	7.1	14.3	14.3	14.3	7.4	11.1	18.5	14.8	11.1	18.5	18.5

Facilitating Conditions (FC)																					
FC1: I have the resources necessary to use HERMES MyFuture.																					
	23.1	15.4	0	7.7	0	30.8	15.4	21.4	0	14.3	0	0	35.7	28.6	22.2	7.4	0	3.7	7.4	33.3	22.2
FC2: I have the knowledge necessary to use HERMES MyFuture.																					
	15.4	7.7	0	0	7.7	46.2	23.1	7.1	7.1	7.1	7.1	21.4	14.3	35.7	11.1	7.4	3.7	3.7	14.8	29.6	29.6
FC3: HERMES MyFuture is not compatible with other systems I use.*																					
	15.4	0	0	46.2	0	30.8	7.7	35.7	21.4	0	14.3	0	7.1	21.4	25.9	11.1	0	29.6	0	18.5	14.8
FC4: A specific person (or group) is available for assistance with HERMES MyFuture difficulties.																					
	38.5	23.1	7.7	0	0	15.4	7.7	0	7.1	0	0	28.6	21.4	28.6	18.5	14.8	3.7	0	14.8	18.5	18.5
Anxiety (AX)																					
AX1: I feel apprehensive about using HERMES MyFuture.*																					
	38.5	30.8	0	15.4	0	0	15.4	64.3	14.3	0	7.1	14.3	0	0	51.9	22.2	0	11.1	7.4	0	7.4
AX2: It scares me to think that I could lose a lot of information using HERMES MyFuture by hitting the wrong key.*																					
	53.8	23.1	0	0	15.4	0	7.7	35.7	14.3	0	7.1	42.9	0	0	44.4	18.5	0	3.7	29.6	0	3.7
AX3: I hesitate to use HERMES MyFuture for fear of making mistakes I cannot correct.*																					
	38.5	38.5	7.7	0	7.7	0	7.7	28.6	21.4	0	7.1	42.9	0	0	33.3	29.6	3.7	3.7	25.9	0	3.7
AX4: HERMES MyFuture is somewhat intimidating to me.*																					
	53.8	23.1	7.7	15.4	0	0	0	50	21.4	7.1	7.1	14.3	0	0	51.9	22.2	7.4	11.1	7.4	0	0
Behavioural Intention to Use the System (BI)																					
BI1: I intend to use HERMES MyFuture in the next semesters if I would have access to it.																					
	15.4	7.7	0	7.7	7.7	30.8	30.8	21.4	7.1	0	7.1	21.4	0	42.9	18.5	7.4	0	7.7	14.8	14.8	37
BI2: I predict I would use HERMES MyFuture in the next semesters if I would have access to it.																					
	23.1	7.7	0	7.7	0	30.8	30.8	14.3	14.3	0	7.1	14.3	7.1	42.9	18.5	11.1	0	7.4	7.4	18.5	37
BI3: I plan to use HERMES MyFuture in the next semesters if I would have access to it																					
	23.1	7.7	0	7.7	0	30.8	30.8	14.3	14.3	7.1	7.1	7.1	7.1	42.9	18.5	11.1	3.7	7.4	3.7	18.5	37

Table 8: Results of the Technology Acceptance Questionnaire for HERMES MyFuture (1, strongly disagree – 7, strongly agree; Likert scale). Technology acceptance is fairly good, as indicated by the many fields highlighted in green. Note: * indicates reversed scale.

Scale/Item	Agreement in %																				
	Austria							Spain ¹⁰							Total ¹⁰						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Performance Expectancy (PE)																					
PE1: I find HERMES MyPast useful in my life.																					
	8.3	8.3	16.7	0	16.7	25	25	25	25	0	12.5	25	0	0	15	15	10	5	10	25	15
PE2: Using HERMES MyPast enables me to accomplish tasks more quickly.																					
	0	16.7	8.3	8.3	25	26.7	25	25	37.5	0	12.5	0	0	12.5	10	25	5	10	15	10	20
PE3: Using HERMES MyPast increases my productivity.																					
	8.3	16.7	16.7	0	33.3	16.7	8.3	25	37.5	0	12.5	0	12.5	0	15	25	10	5	20	15	5
PE4: Using HERMES MyPast increases my chances of leading an active lifestyle.																					
	16.7	8.3	16.7	16.7	8.3	16.7	16.7	25	37.5	0	12.5	0	0	12.5	20	20	10	15	5	10	15
Effort Expectancy (EE)																					
EE1: My interaction with HERMES MyPast is clear and understandable.																					
	0	0	8.3	0	16.7	50	25	0	0	12.5	12.5	0	25	37.5	0	0	10	5	10	40	30
EE2: It is easy for me to become skilful at using HERMES MyPast.																					
	8.3	8.3	0	0	25	41.7	16.7	0	0	12.5	12.5	12.5	12.5	37.5	5	5	5	5	20	30	25
EE3: I find HERMES MyPast easy to use.																					
	0	16.7	0	8.3	25	25	25	0	0	12.5	12.5	0	37.5	25	0	10	5	10	15	30	25
EE4: Learning to operate HERMES MyPast is easy for me.																					
	8.3	0	8.3	0	8.3	41.7	33.3	0	0	12.5	12.5	0	12.5	50	5	0	10	5	5	30	40
Attitude toward Using Technology (AT)																					
AT1: Using HERMES MyPast is a good idea.																					
	0	16.7	0	8.3	16.7	33.3	25	12.5	0	25	12.5	12.5	12.5	12.5	5	10	10	10	15	25	20
AT2: HERMES MyPast makes life more interesting.																					
	0	25	8.3	8.3	25	25	8.3	12.5	12.5	37.5	0	12.5	0	12.5	5	20	20	5	20	15	10
AT3: Living with HERMES MyPast is fun.																					
	16.7	16.7	16.7	0	8.3	33.3	8.3	12.5	0	25	12.5	0	25	12.5	15	10	20	5	5	30	10

¹⁰ For Spain these percentages come from the 8 participants who tested the application and not from the 14 participants who did the trial. One of the 8 participants did not rate this scale because he or she was in a hurry. That is why the percentages do not sum up to 100%.

AT4: I like living with HERMES MyPast.																					
	25	0	16.7	8.3	16.7	25	8.3	25	0	37.5	12.5	0	12.5	0	25	0	25	10	10	20	5
Facilitating Conditions (FC)																					
FC1: I have the resources necessary to use HERMES MyPast.																					
	25	0	8.3	8.3	16.7	25	16.7	0	12.5	0	12.5	0	50	12.5	15	5	5	10	10	35	15
FC2: I have the knowledge necessary to use HERMES MyPast.																					
	16.7	8.3	0	0	0	58.3	16.7	0	12.5	0	12.5	12.5	37.5	12.5	10	10	0	5	5	50	15
FC3: HERMES MyPast is not compatible with other systems I use.*																					
	8.3	0	0	50	0	16.7	16.7	50	12.5	0	12.5	0	12.5	0	25	5	0	35	0	15	10
FC4: A specific person (or group) is available for assistance with HERMES MyPast difficulties.																					
	8.3	8.3	0	16.7	16.7	25	16.7	12.5	12.5	0	0	0	50	12.5	10	10	0	10	10	35	15
Anxiety (AX)																					
AX1: I feel apprehensive about using HERMES MyPast.*																					
	33.3	41.7	0	16.7	0	0	8.3	62.5	12.5	12.5	0	0	0	0	45	30	5	10	0	0	5
AX2: It scares me to think that I could lose a lot of information using HERMES MyPast by hitting the wrong key.*																					
	41.7	25	16.7	8.3	0	8.3	0	25	37.5	12.5	0	0	12.5	0	35	30	15	5	0	10	0
AX3: I hesitate to use HERMES MyPast for fear of making mistakes I cannot correct.*																					
	33.3	25	16.7	16.7	0	0	8.3	37.5	37.5	12.5	0	0	0	0	35	30	15	10	0	0	5
AX4: HERMES MyPast is somewhat intimidating to me.*																					
	33.3	25	16.7	8.3	8.3	0	8.3	50	37.5	0	0	0	0	0	40	30	10	5	5	0	5
Behavioural Intention to Use the System (BI)																					
BI1: I intend to use HERMES MyPast in the next semesters if I would have access to it.																					
	16.7	8.3	8.3	8.3	8.3	41.7	8.3	62.5	12.5	0	0	0	0	12.5	35	10	5	5	5	25	10
BI2: I predict I would use HERMES MyPast in the next semesters if I would have access to it.																					
	16.7	25	0	16.7	0	33.3	8.3	62.5	12.5	0	0	0	0	12.5	35	20	0	10	0	20	10
BI3: I plan to use HERMES MyPast in the next semesters if I would have access to it																					
	16.7	25	0	8.3	8.3	33.3	8.3	62.5	12.5	0	0	0	0	12.5	35	20	0	5	5	20	10

Table 9: Results of the Technology Acceptance Questionnaire for HERMES MyPast (1, strongly disagree – 7, strongly agree; Likert scale). Acceptance is not as good as for MyFuture, especially the behavioural intention to use the system. Note: * indicates reversed scale.

4.2.2 Interpretation

Regarding HERMES MyFuture, the above mentioned results indicate that the majority of Austrians and Spanish participants have a good *Performance Expectancy*, that is, they believe that using HERMES MyFuture will be useful in their life. They believe it is rather easy to use (*Effort Expectancy*), they have the resources to use it (*Facilitating Conditions*), and its use seems a good idea (*Attitude towards using technology*). Also, they do not feel fear or anxiety at the time of use (*Anxiety*) and would not mind using it if they had access to it (*Behavioural intention to use the system*).

Mann-Whitney Tests show significant differences in MyFuture between Spain and Austria in only one item: “A specific person (or group) is available for assistance with HERMES MyFuture difficulties” ($p=0.01$). Spanish older users are more in agreement with it than Austrian users ($M=17.75$ and 9.96 , respectively). Also, there were gender differences in only one item of MyFuture questionnaire: “It scares me to think that I could lose a lot of information using HERMES MyFuture by hitting the wrong key” ($p<0.01$). Older men are more in agreement with it than older women ($M=17.62$ and 11.10 , respectively). However, there were no significant age differences (between users up to 69 years and 70+ years) in HERMES MyFuture questionnaire.

The results of the HERMES MyPast TAQ indicate that participants have a rather neutral attitude towards the usefulness of the system. For example, they do not believe they can lead a more active life (*Performance Expectancy*). However, they believe they have the necessary resources to use the system (*Facilitating Conditions*) and they see it easy to use (*Effort Expectancy*). Many of them think it is a good idea to use it, they think it can be fun (*Attitude towards using technology*), and they show no fear or anxiety using it (*Anxiety*).

Considering the rather positive attitude towards the indicators of the TAQ mentioned above, it seems rather surprising that there is such a negative tendency regarding the users’ intentions to use HERMES MyPast, if they had access to it (*Behavioural intention to use the system*). Participants are somewhat reluctant to imagine using the system in the future. Considering the circumstances the prototype was evaluated in each country, it seems the status of the system is influencing its acceptance. Spanish users were exposed to the real speech search functionality which is underpinned by the automatic transcription of speech recorded with the PDA. It should be noted that the HERMES database was not (and could not be) filled up with numerous personal recordings but rather contained a small amount of short recordings made at the first session. The value provided by the speech search cannot be revealed under such circumstances. Apart of that, the users of course didn’t know about the query composition principles (e.g. use of words that likely were frequently said in the target conversation, use of multiple queries in a trial-and-error fashion). In the absence of huge volumes of personal user data the intention and main focus of the evaluators was the user interface rather than the search functionality itself. In order to demonstrate the interface the evaluators suggested certain query words taken by them on purpose from the ASR transcript while from the users perspective these words might not represent the essence of the conversation as they remember it or even represented the ASR errors. It explains why most of the Spanish users perceived this functionality useless. The lessons learned from this experience will be implemented in the final trial. First, we will strive to urge the users to collect as much recorded data as possible. Secondly, we will use the avatar to guide and advise the user during the interaction (e.g. how to increase the number of results by using more general keywords).

Comparing Austria with Spain, there are differences in the intention to use the system, but they are not statistically significant. In Austria, there was no real functionality at all exposed in the

user interface, partly due to the same technical issues that affected Spain, partly due to the fact that there is no audio transcription available in German. In order to test the system in Austria, the evaluator needed to communicate the idea of the system to the users, which they found interesting. These differences in acceptance, depending on whether a system provides useful data or not, need to be investigated further, as described in Section 6.3.

4.3 Usability Evaluation of Common Components of MyFuture and MyPast

The purpose of this section is to cover usability issues of those software components that are used both in MyFuture (Section 4.4.3) and MyPast (Section 4.6.3), which are mainly the on-screen keyboard and the clock.

4.3.1 Results

The following tables present usability problems found in the components mentioned above, grouped by categories serious, important and marginal usability problems. The rating of severeness follows the explanations in the introduction of Section 4. No nice-to-have adjustments were found.

Serious Usability Problems

Category / Problem	Proposed Solution(s)
Interaction / Gestures	
Users struggle with dragging the finger over or holding it on the screen. It is used in various parts of the application (setting time, moving an appointment, entering Umlaut/accent characters) and every use is causing problems. Since only in Austria a touch screen was used – which had a rather difficult surface that required the user to press quite firm and precisely – it remains unclear, how much the screen used evaluating the system affected the performance. Most of the time users were not thinking of continuously pressing the finger on the screen, but only tapped the element for a short period of time.	Possible gestures should be communicated somehow by the system. Maybe there could be a description telling e.g. “drag me to a new date” appearing after a draggable element was tapped.
Another effect that only became visible when using the touch screen was the fact that users are hiding interface elements with their own arm. This becomes especially evident when users don’t see the popup to enter Umlaut mark or other special characters. Also, users can’t read the time in the centre of the time selection dial, because they are hiding it with their own fingers.	A restructuring of the applications is necessary. All main interaction elements (like the navigation bar) should be put on the bottom of the screen. Show the keyboard pop-up above the button pressed. The pop-up also needs a clearer border to differentiate it more from the other keys.

Interaction / Buttons	
Users have difficulties understanding the concept of a “button”. Some don’t recognise a greyed-out button as non-clickable; some don’t understand why there are buttons they have to tap, to hold, rotate, or drag. The button itself does not communicate how to interact with it.	A clear differentiation between buttons (and other clickable elements) and non-clickable screen elements must be provided (by colour and/or border). Furthermore the buttons need to carry descriptive labels containing verbs and nouns that tell what will happen to which object if a user presses the button (e.g. “Create A New Appointment” instead of “New Appointment”).
Interaction / Text Entry	
Users don’t understand they have to tap into a text box to enter text (to make the keyboard appear).	An empty text box might carry a label saying “tap here to enter text”.
Interaction / Time Entry	
Users don’t know they could tap the buttons with the time. They are guessing various other locations (“Today”, Number-keypad, big up/down arrows in events list) or they want to enter the numbers with the keyboard, if it stayed visible from a former interaction.	Use a different approach to enter the time altogether. A digital clock with +/- buttons is suggested.
Users don’t know they have to hold the finger on the button to see the time dial	
Users don’t know they have to circle around the centre point to set the time.	
Users don’t know need to rotate counter-clockwise to get hours smaller than 12.	
In MyFuture, many users don’t know they have to set hour and minute in a different dial. Instead they tried to position the hour hand between two hours.	
Language	
Users experience difficulties with English text exposed by the GUI. Although terms are translated, they don’t remember their meanings. Consequently some usability issues might only be caused by language problems.	GUI Localisation

Table 10: Serious usability problems and suggested solutions in common components of MyPast and MyFuture.

Important Usability Problems

Category / Problem	Proposed Solution(s)
Interaction / Text Entry	
Confusing functionality of the green tick and the red cancel button on the keyboard.	Only one “Hide Keyboard” button, text entry is always saved
Users don’t know where to switch between upper and lowercase.	Might only be language related.

Users get irritated because they can't clear all text in a text box with one tap. (They press "Correct" several times or try the "Cancel" Button on the keyboard).	Provide a "Clear All" button next to or in the text box to clear all entered text with one tap.
Users don't see they can hold a key to get special characters with Umlaut or accents. Some recognise the small triangle pointing down, but get irritated, because it's shown also on letters where they wouldn't expect it (in German only letters A,O,U can have an diacritic mark)	Provide separate keys for common special characters of the selected language. Provide a button to switch to Umlaut/accent characters. Provide a textual description telling users they can hold a key to enter Umlaut and accent characters

Table 11: Important usability problems and suggested solutions in common components of MyPast and MyFuture.

Marginal Usability Problems

Category / Problem	Proposed Solution(s)
Interaction / Text Entry	
Users need time to realise they are already typing letters, because the text box, where the letters appear, is too far away from the keyboard.	Put keyboard and text box closer together.

Table 12: Marginal usability problems and suggested solutions in common components of MyPast and MyFuture.

4.3.2 Summary

The above mentioned results indicate that the following main changes to the common components of HERMES MyFuture und MyPast are necessary:

- localisation of the user interface in German and Spanish
- restructuring of the content (navigation bar to bottom)
- clearer highlighting of clickable elements by caption, colour, border and/or interactive feedback (pressed-down state)
- an easier and more reliable way of entering special characters
- a new way of setting the time (digital clock)

For further details the columns "Proposed Solution(s)" in each of the tables of Section 4.3.1 are referred.

4.4 MyFuture Application

All of the participants in Austria and in Spain tested HERMES MyFuture. They were asked to complete three tasks in this application:

1. browsing events
2. entering a new appointment
3. moving an entry to another day

The following sections present the evaluation of the quantitative and qualitative questions that were asked after each of the three tasks, as well as the usability evaluation based on the observation of the users performing the tasks.

4.4.1 Quantitative Evaluation

Table 13 summarises the results retrieved from the quantitative questions of HERMES MyFuture. It uses the same colour scheme as the tables before. The results indicate that the majority of Austrians and Spanish participants believe the system supports them in fulfilling all tasks. However, they agree the system can be improved to better support them. They think it is easy for them to understand the icons, many of them are able to follow the changes on the screen and they feel comfortable doing all tasks.

Mann-Whitney Tests show a significant difference in the MyFuture post-task questionnaire between Spain and Austria in only one item: “I felt comfortable doing move entry task” ($p=0.01$). Spanish older users are more in agreement with it than Austrian users ($M=17.57$ and 10.15 , respectively). There were no gender and age differences in the MyFuture post-task questionnaire.

Looking at the results, a general agreement tendency of the users towards any of the questions seems likely. This is supported by the fact that also second statement, which used a reverse scale, is generally agreed with. Also, these results coincide with the observations made by the evaluators while the participants were filling in the questionnaire. Therefore, conclusions drawn out of these results need to be carefully considered.

Task/ Question	Agreement in %														
	Austria					Spain					Total				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Browsing events															
The system supported me well in fulfilling this task.															
	0	23,1	7,7	23,1	46,2	0,0	14,3	14,3	34,3	7,1	9,0	18,5	11,1	44,4	25,0
The system can be improved in supporting the fulfilling of the task.*															
	7,7	23,1	7,7	53,8	7,7	0	0	14,3	71,4	14,3	3,7	11,1	11,1	63	11,1
It is easy to understand the icons on the screen.															
	0	0	23,1	46,2	30,8	7,1	21,4	14,3	42,9	14,3	3,7	11,1	18,5	44,4	22,2
I was able to follow the changes on the screen easily after tapping it.															
	0	7,7	7,7	53,8	30,8	0	7,1	0	85,7	7,1	0	7,4	3,7	70,4	18,5
I felt comfortable doing this task.															
	0	0	46,2	23,1	30,8	0	0	0	57,1	42,9	0	0	22,2	40,7	37
New Entry															
The system supported me well in fulfilling this task.															
	15,4	30,8	15,4	0	38,5	0	14,3	7,1	57,1	21,4	7,4	22,2	11,1	29,6	29,6
The system can be improved in supporting the fulfilling of the task.*															
	7,7	0	7,7	53,8	30,8	14,3	7,1	21,4	35,7	21,4	11,1	3,7	14,8	44,4	25,9
It is easy to understand the icons on the screen.															
	0	15,4	7,7	46,2	23,1	0	35,7	7,1	42,9	14,3	0	25,9	7,4	44,4	18,5
I was able to follow the changes on the screen easily after tapping it.															
	0	23,1	23,1	46,2	7,7	0	7,1	14,3	64,3	14,3	0	14,8	18,5	55,6	11,1
I felt comfortable doing this task.															
	0	23,1	7,7	46,2	23,1	0	0	0	71,4	28,6	0,0	11,1	3,7	59,3	25,9
Move entry															
The system supported me well in fulfilling this task.															
	15,4	30,8	7,7	30,8	15,4	0	14,3	14,3	28,6	42,9	7,4	22,2	11,1	29,6	29,6

The system can be improved in supporting the fulfilling of the task.*															
	0	15,4	15,4	53,8	7,7	7,1	14,3	28,6	42,9	7,1	3,7	14,8	22,2	48,1	11,1
It is easy to understand the icons on the screen.															
	0	23,1	15,4	53,8	7,7	0	21,4	14,3	57,1	7,1	0	22,2	14,8	55,6	7,4
I was able to follow the changes on the screen easily after tapping it.															
	0	23,1	30,8	30,8	15,4	0	0	7,1	85,7	7,1	0	11,1	18,5	59,3	11,1
I felt comfortable doing this task.															
	0	0	38,5	53,8	7,7	0	0	7,1	42,9	50	0	0	22,2	48,1	29,6

Table 13: Results of the quantitative evaluation of HERMES MyFuture. The results indicate a good performance of the system. However, a possible agreement tendency of the users needs to be considered. Note: * indicates reversed scale.

4.4.2 Qualitative Evaluation

The participants were asked five qualitative questions after each task:

1. What can be improved
2. Specific difficulties for completing the task
3. Problems (apart from the English language)
4. Difficulties following the changes on the screen (application flow)
5. Special feelings associated with completing the task

Due to the fact that the feedback of the majority of the participants did not target on a specific task but on the system in general, for presentation purposes the results are not grouped by task but by question instead. Also, questions 2, 3 and 4 are presented together, since the feedback was too indifferent to be separated.

Improvements

When asked about the improvements the users can think of, they gave the following feedback:

1. All users believed they needed some kind of written instructions manual, handbook, a help function within the system or more training in order to perform better.
2. The majority of the participants identified the line, which appears above the days in the calendar and which indicates that on this day there is an appointment, as one of the icons that can be improved. The problem with this line is that it is not very easy to notice. The solutions are divided into three groups:
 - To make a bigger line.
 - To change the colour of the line to a more visible colour. The participants who gave this suggestion proposed the use of a red colour.
 - To change the line to a circle, as they are used to make circles on their calendars at home on the days they have something to do. Three participants pointed out that circles are more intuitive than lines. One of the participants commented that it would be useful to have multiple concentric circles indicating the number of appointments on a given day.
3. Participants also felt that the interface itself could also be improved by means of:
 - Reducing complexity by fewer buttons.
 - Enlarging the font for the button labels.
 - Enlarging the calendar to full-screen. Some participants recommended the possibility of using a zoom button.
 - The colours of the application are not optimal for two reasons: (1) the colours do not motivate the elderly to play with the application; (2) for the elderly it is difficult to see the icons and the numbers in the application as they are shown at the moment. They demand more visible colours and more contrast between the colours on the screen. One

participant suggested the idea of distinguishing the buttons by colour (e.g. go to next month in yellow colour; go to previous month in blue colour).

4. In Austria two users requested an additional feature: They wanted to set the duration of the appointment.

Difficulties and Problems

General difficulties arose from the use of the English language. Many usability problems ascribed later on are related to the fact that users didn't find certain elements on the screen, because they didn't understand the words and also didn't ask the evaluator for translation. It remains unclear, however, which of the problems are only related to language difficulties.

Besides that, the users identified the following task-specific difficulties:

1. In the second task, only one participant stated that for her this task was difficult because a lot of steps were necessary to enter a new appointment and she was not able to memorize all of them.
2. Setting the time with the clock dial: As already mentioned in Section 4.3, Usability Evaluation of Common Components of MyFuture and MyPast, a general complaint was about the clock which appears when they have to set the time. Some users proposed changing the clock to a digital one.
3. Moving appointments by drag and drop: The third task has been the most difficult one. Some improvements proposed by the participants are:
 - In the case they have to move one appointment from one day to another, it is easier to directly point to the day they want to change instead of moving the appointment from one day to another with the mouse.
 - If this is not possible, they prefer that the square which appears when they are moving the appointment be transparent, because, in this way they can see the day under the square. Now it is not clear on which new day they are allocating the appointment.
4. In Spain the mouse usage was very difficult and slowed down the performance of the tasks.
5. When participants selected a day for viewing the details of their appointments, it was difficult to go back to the main screen of the application. For them to click on "close day" was not intuitive even when the evaluator told them the meaning of "close day" in Spanish or German respectively.
6. The buttons which contain words (e.g. previous month, next month and so on), should also contain icons. They commented that it would be easier for them to understand that they have to click on next month if an arrow marking to the right appears.
7. Test persons both in Austria and in Spain had a lot of problems with the use of the screen keyboard. In Spain this was partly related to the fact that the users were using the mouse to enter text. These users preferred to use the standard keyboard. They suggested however, that it would be possible for them to use the keyboard on the screen if it were a touch-screen.
8. When the users tried to write down the description of the new appointment they have to clear the previous text. It should not be any text in order the participants would avoid the erasing step.

Feelings

In Spain people do not report special feelings like anxiety or stress when they completed these tasks. In Austria however, feelings ranged from excitement (about the touch-screen) and good feelings to negative ascriptions: People described their feelings as insecure, inexperienced, overstrained, surprised, unfamiliar, not good, surprised (not as bad as expected, slower than with paper), sceptical and confused.

In comparison with the PDA and with MyPast application, MyFuture was the easiest to use and the most useful application for the participants. The results described above are summarized in Table 14, where the major problems the users have found and the way to solve them proposed by the users are listed. The solutions to these problems recommended by the authors are presented in the next section.

Category/Problem	Solution proposed by users
General	
Need help in order to do the tasks on their own	Develop a manual
Small calendar and icons	Enlarge the calendar to full screen
Colours	To change the colours in order to get a more attractive application To use colours in order to achieve more contrast
Content Layout / Finding events	
The line which indicates there is an event it is not easy to see	Bigger line Change the colour of the line Replace the line with circles
Interaction / New appointment	
The clock to set the time it is difficult to use	Replace it with a digital clock
Interaction / Moving appointment	
It is difficult to drag the appointment with the mouse from one day to another	Just click on the new day they want to allocate the appointment

Table 14: Summary of Problems and solutions proposed by the users in HERMES MyFuture

4.4.3 Usability Evaluation

The usability evaluation of this section covers the problems observed during the use of MyFuture that were not already covered in Section 4.3.

Serious Usability Problems

Problem	Proposed Solution(s)
Wording	
“Close Day” Button: unclear description. Users don’t understand, what it means.	Users expect a “Back” or “Whole Month” button. Change the labelling and place the “Back”-Button in the navigation bar.
Content Layout / Structure	
Users don’t know what the big scroll buttons next to the event list mean. If they have more appointments than fit on the screen (i.e. 3), they do not see them.	Communicate the total number of appointments by writing it somewhere. Suggest that there are more appointments than fit on the screen by showing half of the next one. Tapping on the appointment that is only shown half needs to invoke scrolling (which again already shows half of the next appointment plus half of the first appointment that is disappearing on top of the list).

Affordance / Moving an appointment	
Most users don't know they have to press "Move" to move an appointment. Instead they try to directly move the entry from day view or from "next events".	Could be language related. Use a more instructive label like "Change date of appointment"
After given a hint to press "Move" users don't see the pulsing appointment shown in the lower left corner. Consequently some users try to move the calendar day to the new date instead of the appointment in the lower left corner.	The pulsing appointment needs to be more close to the centre of attention i.e. the centre of the screen.
Many users don't know they can drag and drop the appointment shown in lower left corner. Instead they just tap it. Sometimes this is connected with a total lack of understanding for the concept of "moving" an appointment. They would rather delete and recreate it as they do on paper.	Provide visual affordance (e.g. an already "floating" appointment supported by arrows pointing in all directions) and/or textual hints.

Table 15: Serious usability problems and suggested solutions in HERMES MyFuture

Important Usability Problems

Problem	Proposed Solution(s)
Wording / Feedback	
Users don't understand the meaning of "record". They tried to press the button in various situations (to enter time, to move an appointment, create a new event, or to save the new appointment).	Might be language related. Use "record voice" for more clarity.
Some users don't understand the meaning of "new event".	Use "Create New Event"
Users don't know what is actually deleted, when they press "Delete".	Button should be labelled with "Delete event" or "Delete whole event".
Icons	
The icon of the "Play" button is unclear. As discussed in Section 0 in general standard symbols for "Play", "Pause", "Stop", "Record" etc. are not recognised).	Use a textual description, group "record voice" and "play voice recording" together, use a different progress indicator.
Users don't see the blue event bar or think it belongs to the wrong day.	Use a stronger colour (red) and a clearer association with a day. Try the user idea of putting a circle around a day.
Users don't understand the audio playback controls. They don't know what "1/1", the time position indicator ("00:00"), and the playback button mean.	Simplify (i.e. reduce) the elements shown on the screen. Provide textual explanations.

Table 16: Important usability problems and suggested solutions in HERMES MyFuture

Marginal Usability Problems

Problem	Proposed Solution(s)
Layout / Structure	
Some users don't know what to expect by pressing "show event" in the "next events" section since for them the event is already shown there.	Use "Show Day" instead or make the whole line clickable and remove the button.
Users think "Next Events" shows the next (first) events of a given month. If a users switches forward to a month in the future, "next events" should show the first events of this month.	Don't adopt this behaviour. Instead communicate clearer which dates are shown in the "Next Events" section.
Users feel overstrained with the many buttons (day view, editing or moving an appointment)	Try to bring more structure to the many buttons in day view.
Users can't differentiate the individual appointment entries from the background.	Group each entry and the entire list more clearly.
Navigation / Feedback	
Users try to press on a greyed-out day of the next month shown already in the current month to go to next month.	If the user presses on one of these days the calendar should switch to the next month.
Feedback / Moving an Appointment	
Users don't know if they are finished after dropping the appointment over the new date	Provide some sort of feedback that the appointment has been moved.
Playability / New Entry / Deleting	
Users don't realise they have to press "Save" to fixate a new event. Some press "Close Day" and all data is lost. On the other hand, when editing an event, there is an auto-saving function.	Auto-save entered data immediately. If a user decides s/he doesn't want the event, s/he can press "cancel/delete" anyway. This would also be more consistent with editing behaviour, where data is saved automatically.
No undo for deleting an event.	Provide an "Undo" button.
Icons	
The icon of the "Close Day" button is unclear.	When using a different label ("Back" or "Whole Month") also change the symbol.
The icon of the "Move" button is unclear.	Replace the symbol together with the label. (See serious usability problems.)

Table 17: Marginal usability problems and suggested solutions in HERMES MyFuture

Nice-to-have

Problem	Proposed Solution(s)
The Calendar is too pale and small.	Make the calendar full screen and use stronger colours.
Users expect to see the number of events for a given day.	Instead of showing the blue bar a badge with the number could be shown.

Some users try to return to month view by pressing “This Month”	Make “This Month” active in day view.
Some users want to enter the duration of an appointment.	Add the possibility of setting the duration of an appointment.

Table 18: Nice-to-have adjustments for HERMES MyFuture

4.4.4 Summary

The results of quantitative and qualitative evaluation and the usability evaluation allow the following main conclusions to be drawn:

- Small changes can have big effects: Change the caption of some of the buttons to make them easier to understand.
- Improve the Drag&Drop operation necessary to move an appointment by better communicating affordance (“I am draggable!”).
- Make the system more robust by providing an auto-save function when creating a new event and by providing an undo button when deleting an event.
- Make the calendar larger (full screen) and use stronger colours.
- Allow alternative ways of navigation (tapping on greyed out days from next or previous month, using “This Month” button in day view).

For further details the columns “Proposed Solution(s)” in each of the tables of Section 4.4.3 are referred.

4.5 HERMES PDA

All of the test persons in Spain (9 in their homes and 5 in the lab) and 12 out of 13 in Austria evaluated the PDA application. The one user that didn’t evaluate it said he can’t read anything on the screen. The other users were asked to complete five tasks in this application:

1. introducing an appointment
2. modifying the appointment
3. introducing an appointment in locate mode
4. browsing appointments
5. recording a conversation

The following sections present the evaluation of the quantitative and qualitative questions that were asked after each of the five tasks as well as the usability evaluation based on the observation of the users performing the tasks.

4.5.1 Quantitative Evaluation

Table 19 presents the results retrieved from the quantitative questions of HERMES PDA. While for most statements the majority of Austrians and Spanish tend to agree, there is more disagreement than for MyFuture. In the task “Browsing through the appointments”, they were undecided whether or not the system can be improved (36.8%) or if they can follow the changes (26.3%). None of the Spanish and only few of the Austrian participants could answer the questionnaire for the task “Introduce an appointment in Locate Mode”. In the in-door lab it was not possible to perform this task due to the lack of GPS coverage. While it was tried to test the functionality outdoors, often the GPS signal could not be found by the device. Also, none of the users succeeded in using the locate mode at home. Therefore some table cells are empty or the numbers do not sum up to 100%.

Mann-Whitney Tests show significant differences in PDA post-task questionnaire between Spain and Austria in two tasks: “Introduce an appointment” and “Modify the appointment”. With respect to the first task, Spanish participants believe more the system has supported them ($p < 0.01$) than Austrian users ($M = 16.36$ and 8.73 , respectively). Also, Spanish users have felt more comfortable doing the task “modify the appointment” ($p = 0.02$) than Austrian users ($M = 15.82$ and 9.41).

There were no gender and age differences in PDA post-task questionnaire.

Task/ Question	Agreement in % ¹¹														
	Austria					Spain					Total				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Introduce an appointment															
The system supported me well in fulfilling this task.															
	18,2	36,4	18,2	27,3	0	0	14,3	7,1	57,1	7,1	8	24	12	44	4
The system can be improved in supporting the fulfilling of the task.*															
	0	9,1	9,1	36,4	4,5	0	14,3	35,7	14,3	0	0	12	24	24	28
It is easy to understand the icons on the screen.															
	9,1	9,1	36,4	27,3	18,2	0	35,7	14,3	21,4	7,1	4	24	24	24	12
I was able to follow the changes on the screen easily after tapping it.															
	9,1	27,3	27,3	27,3	9,1	14,3	35,7	7,1	21,4	0	12	32	16	24	14
I felt comfortable doing this task.															
	9,1	27,3	18,2	27,3	18,2	0	14,3	0	42,9	21,4	4	20	8	36	20
Modify the appointment															
The system supported me well in fulfilling this task.															
	9,1	18,2	9,1	36	9,1	7,1	0	7,1	42,9	28,6	8	8	8	40	20
The system can be improved in supporting the fulfilling of the task.*															
	0	9,1	0	54,5	18,2	7,1	7,1	42,9	14,3	7,1	8,0	16	16	16	28
It is easy to understand the icons on the screen.															
	9,1	9,1	27,3	18,2	27,3	7,1	21,4	7,1	14,3	28,6	4	8	24	32	12
I was able to follow the changes on the screen easily after tapping it.															
	0	18,2	45,5	9,1	18,2	7,1	14,3	14,3	28,6	14,3	4	16	28	20	16
I felt comfortable doing this task.															
	9,1	18,2	9,1	45,5	9,1	0	7,1	0	28,6	42,9	4	12	4	36	28
Introduce an appointment in Locate Mode															
The system supported me well in fulfilling this task.															
	9,1	36,4	0	9,1	18,2						8	20	0	8	8
The system can be improved in supporting the fulfilling of the task.*															
	9,1	9,1	0	27,3	27,3						4	4	4	12	16
It is easy to understand the icons on the screen.															
	0	9,1	27,3	27,3	9,1						0	16	12	16	4
I was able to follow the changes on the screen easily after tapping it.															
	0	9,1	18,2	36,4	9,1						4	8	8	24	4

¹¹Not all participants could answer all questions. That is why the percentages partly do not sum up to 100%.

I felt comfortable doing this task.															
	9,1	0	27,3	27,3	9,1						4	4	16	16	8
Browsing through the appointments															
The system supported me well in fulfilling this task.															
	20	0	0	0	20	0	0	7,1	35,7	42,9	5,3	5,3	5,3	26,3	42,1
The system can be improved in supporting the fulfilling of the task.*															
	0	20	0	20	40	7,1	14,3	50	7,1	0	5,3	15,8	36,8	10,5	10,5
It is easy to understand the icons on the screen.															
	0	0	20	40	20	0	7,1	35,7	35,7	0	0	5,3	31,6	36,8	5,3
I was able to follow the changes on the screen easily after tapping it.															
	0	0	20	40	20	7,1	0	35,7	42,9	0	5,3	0	26,3	42,1	5,3
I felt comfortable doing this task.															
	20	0	0	40	20	0	7,1	14,3	42,9	14,3	5,3	5,3	10,5	42,1	15,8
Conversation															
The system supported me well in fulfilling this task.															
	0	18,2	18,2	36,4	27,3	0	0	7,1	50	7,1	0	12	44	0	28
The system can be improved in supporting the fulfilling of the task.*															
	0	27,3	18,2	36,4	18,2	0	21,4	35,7	0	35,7	0	29,2	25	16,7	16,7
It is easy to understand the icons on the screen.															
	0	0	9,1	63,6	27,3	0	7,1	7,1	50	14,3	0	16	12	44	16
I was able to follow the changes on the screen easily after tapping it.															
	0	10	20	50	20	14,3	0	7,1	42,9	14,3	0	8	8	52	20
I felt comfortable doing this task.															
	0	20	10	40	30	0	14,3	0	21,4	42,9	0	12,5	8,3	33,3	33,3

Table 19: Results of the quantitative evaluation of HERMES PDA. Note: * indicates reversed scale.

4.5.2 Qualitative Evaluation

As in evaluating HERMES MyFuture the participants were asked five qualitative questions after each task of the PDA application:

1. What can be improved
2. Specific difficulties for completing the task
3. Problems (apart from the English language)
4. Difficulties following the changes on the screen (application flow)
5. Special feelings associated with completing the task

Due to the fact that the feedback of the majority of the participants did not target a specific task but the system in general, for presentation purposes the results are not grouped by task but by question. Also, questions 2, 3 and 4 are presented together, since the feedback was too indifferent to be separated. The results of the tasks for doing at home are included in this summary as well.

Improvements

1. It was suggested to reduce the steps necessary to create a new entry. Users came up with the idea of having a screen in which to select whether to enter an appointment by time or to enter an appointment by location. If the user chooses to enter an appointment selecting the time mode, the screen in which the user has to enter or select information regarding the

location will not be shown. And vice versa, if the user chooses to enter an appointment by location, the screen asking the user to enter the date and hour will not be shown. This solution will reduce the number of steps the user has to take in order to successfully record an appointment. The changes suggested in Section 4.5.4 will further reduce the number of steps.

- There are many icons and some of them are not very intuitive. For example, in the screen shown in Figure 13, the user has to press the button on the right if he wants to continue and enter the time of the appointment.

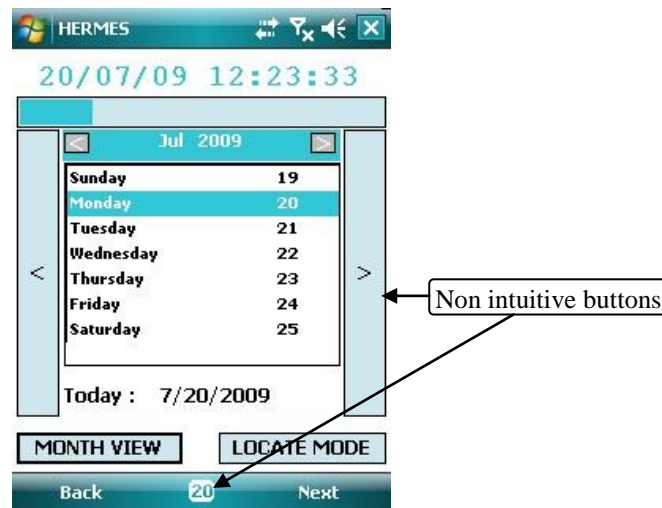


Figure 13: Non-intuitive buttons when creating a new entry on HERMES PDA: Users don't understand the ">" button. The numeric keypad that can be turned by pressing the "20" button is not recognised.

However, after selecting the day, the users are not sure about what to do next. They prefer to see an icon with the text "Continue". Another button, which is not very clear, is the "20" button that appears at the bottom of the PDA. The users do not associate that button with the keyboard. In general, users wanted clear descriptions (in German/Spanish) telling them what to do next. Readability of icons will also improve, if clear text labels are provided. This especially refers to the Conversation application. All users criticised the record and playback control icons, which were not intuitive for them.

- Participants say that at the end of the process they want a text message informing them if they have successfully recorded the appointment or if they have pressed the wrong button. They would like to see a message like "You have not finished recording your appointment, are you sure that you want to exit? Some of the participants were not able to record an appointment when they were at home because in the end they did not press the save button.
- Regarding the warnings of the PDA when they have an appointment, they know if the PDA is switched off they are not going to receive the warning until they switch on the PDA again. They would like the PDA to warn them even if it is off.
- As with the other applications, they would change the colours of the application and they would also increase the contrast between the colours and the icons.
- Some of them prefer to see the days of the following week rather than the list of the days on the screen. They also prefer the calendar screen and not the list of events as the main screen.
- When they enter an appointment on the PDA, the current date appears and not the date of the appointment, which confuses almost all the users. For example, in Figure 14 the users have already chosen one day and now they have to select the time for the appointment. At the top of the screen, however, they see the current day and this confuses them.

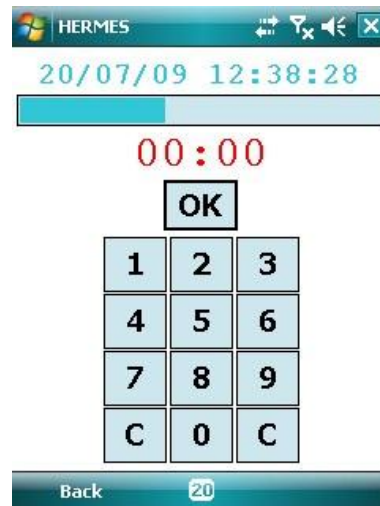


Figure 14: After having selected a day in the future to create an appointment, on the next screen the current date is confusing. Users think it is the day the new appointment will take place and think they selected the wrong date.

1. There is no need to place the “C” both sides of the button; this could be a source of distraction.
2. Some of the users mentioned the possibility of recording periodic events which happen always on the same date. For example, birthdays.
3. In the last task of the PDA (conversation), some of them suggested including something else to clarify what the different icons mean. For example, play, record, stop, etc. Also, if they want to listen to the recorded conversations, they can only listen to the last one and they do not know where the other conversations are. Some of them, would add a confirmation message after they have pressed the save button (e.g. “Your conversation has been saved”). Also they would like to know the date in which the conversation was recorded and to have a screen with all the recordings or some way of looking for them.

Difficulties and Problems

1. The use of the stylus is very difficult and slows down the performance of the tasks. For some tasks, they need to press with the stylus very accurately; older users with some kind of tremor would have more difficulties performing this action.
2. The users don’t know what to do next. The inconsistent navigation confuses them. They don’t know where to press to get to the next screen and they can’t follow what happened after they pressed a button.
3. The words are very small. The users commented that it would be easier for them to understand what they have to click on next month if an arrow marking to the right appears.
4. When they use the keyboard of the PDA, the letters are very small and it activates the dictionary, which makes the task harder. Besides that, the users had difficulties in removing the PDA keyboard by pressing the “20” button.
5. They required more support than that given by the application itself about what they should do on any screen or at any step, without requiring the evaluator’s help.
6. The buttons at the screen borders should be moved slightly inside the screen because of the lack of sensibility of the borders and corners on the touch screen.
7. Some participants have exited the HERMES application on the PDA and could use other programs of the PDA. Two participants in Spain entered the configuration tool (in which the user is selected, which should only be done by INGEMA or CURE), and they selected another different user. This has caused the loss of information for these people.

8. The task of entering an appointment in locate mode could not be performed because it did not work correctly or users tried to use it when they did not have GPS signal.
9. Sometimes some warnings of the PDA (low battery, no GPS signal...), appeared and the users did not know what to do.

Feelings

1. When the users picked up the PDA at the lab they thought they would be able to record appointments and conversations by themselves without the help of the evaluator. However, when they gave us back the PDA no-one reported that he or she was actually able to do so.
2. Participants are afraid of losing information. They would like a method that allows them to retrieve the information.
3. Some of them reported feelings such as nervousness, anxiety or stress when they were completing these tasks. Some users even got annoyed and impatient because they were struggling so much with the PDA. While all users acknowledged that this application is only half-baked some said that they felt stupid. However, if it worked better, some users would prefer the PDA over the desktop MyFuture application due to the portability of the device.
4. The conversation application was easier than the appointments application.

The results described above are summarised in Table 20. It shows the major problems the users have found and the way to solve them proposed by the users. The solutions to these problems recommended by the authors are presented in the next section.

Category / Problem	Solution proposed by users
General / Interaction	
Use of stylus	None
Need of help in order to do the tasks by themselves	Develop a manual
Small letters and icons	Enlarge the letters and icons
Navigation / New Appointment	
Too many steps necessary in order to successfully record an appointment	Not having to press the save button so many times.
Having to write twice the audio description	
Modifying an Appointment	
Too many steps to modify only a date	Reduce steps
Feedback / Locate Mode	
The GPS was not always activated	None
Feedback / Conversation	
Listen only to the last recorded conversation	Indicate that the conversation is saved and where.

Table 20: Summary of problems and solutions proposed by the users in HERMES PDA

4.5.3 Usability Evaluation

The usability evaluation of this section covers the problems observed during the use of HERMES PDA.

Serious Usability Problems

Category / Problem	Proposed Solution(s)
General	
Users struggle with the inconsistent navigation. In general they are without any orientation and stranded. Buttons vary strongly in position, labelling and behaviour. Different buttons on one screen have confusingly similar labels. Screens have unclear description or no description at all. There is too much technical information present and too much technical knowledge necessary to use the device. The needs of the users are not met (e.g. it should not be necessary to assign every location with GPS coordinates).	A total re-design of the PDA “Appointments” application is recommended.
The users who took the PDA home, got lost in the PDA-system at some point, because they didn’t find the Hermes application anymore. They are currently required to navigate through the systems File Explorer to start the HERMES application.	Fix the error message at PDA start up, if possible deactivate PIN entry, auto-start the HERMES app (or at least put a short-cut on the main screen). Make it difficult to accidentally exit the application.
Interaction / Navigation	
Users don’t know what to do on each screen. They are unsecure because they do not know if the things they have entered on previous screens are saved or lost. This especially applies to the “audio recording” screen. The progress bar is not understood.	Provide a “task description” or “headline” on top of the screen. Provide “Step X of Y” Information.
Users don’t see or recognise the tall “>” button on the right side to proceed to the next screen. (Some press locate mode instead.)	Consistent position and labelling of “next” and “back” buttons
Users want to use the tall “>” (which is to proceed to next screen) to go to the next week or month.	See above.
Users need to enter a description after they press “Save”. Then they press “OK”, and then are needed to press “Save” again. The cumbersome navigation confuses them.	It is unclear why there is an option to enter a description, which is after all not visible in the screen showing an overview of the appointment. (It is only shown in the list in the main screen).
Interaction / Text Entry	
Users don’t know they have to tap the text field to get the keyboard and to enter text.	An empty text box might carry a label saying “tap here to enter text”.
Users don’t understand that the text field needs focus in order to be able to enter text. Users don’t know how to hide the keyboard afterwards	The keyboard should show and hide automatically, depending on if a text entry box has focus or not.
Interaction / Time Entry	
Users don’t know they have to enter a zero (0) first, if they want to enter an hour below 10.	Use a different approach to enter time (drop-down list).

Interaction / Description	
Users don't know how to proceed after entering a description, because the keyboard is hiding the "OK" button. They need to hide the keyboard first.	Don't let the keyboard hide the navigation buttons.
Interaction / Alert	
The Alert screen is confusing. Users don't know what to do there (not everybody saw it).	Re-design of this screen. Simplify by reducing options.
Content Layout / Audio message (Appointments)	
Users don't know what to do here. There is no description. They try to press "record" or "save" instead of "next" to continue.	Re-design of this screen.
Feedback / Location	
The terms "longitude" and "latitude" are too technical. Users don't know that they have to wait and what they have to wait for.	Put a spinning wheel, a progress bar, or a description "Determining Position. Please Wait..."
The Message Box "GPS turned off. Activate?" is too technical.	Auto-activate or (if not possible) use a phrase like "Location services need to be turned on in order to locate your position. Turn them on now?" – If "no" is pressed, return to previous screen.
The Message Box "GPS turned off. Activate?" is not recognised. Users try to press navigation buttons in bottom which of course don't work.	See above. Avoid message box altogether.
Users don't know how to proceed after selecting a location.	Re-design of this screen.
Icons / Colours / Conversation	
Users don't understand the symbols. They are not familiar with them even if they have a video or tape recorder.	Provide textual descriptions. Ideas: ear=play, mouth=record, microphone=record, stop-sign=stop.
The colours are misleading. (Red is for them not the colour to record, but to stop – like in traffic). Two shades of blue are confusing.	Recommended use of colour: red=stop, green=play.
Unsure if the colours are distinguishable by colour blind people.	Use colour blind-safe colours.

Table 21: Serious usability problems and suggested solutions in HERMES PDA

Important Usability Problems

Category / Problem	Proposed Solution(s)
Content Layout	
Users get irritated by the current date and time shown on the top as they think it's the time of the appointment they are about to create.	Remove it. It is unnecessary anyway.
At the Time Entry screen users don't know they have to enter the time.	Provide a label. Instead of typing the time, let them choose from a list (or spinning wheel like on iPhone). This also eliminates the strange possibility of entering invalid times.
Feedback / Conversation	
Users don't recognise that a recording is already running	Clearer feedback.
Interaction / Navigation	
Users don't know they have to press "OK" after entering the time.	Consistent Next/Back navigation.
General / Alert	
Disabling an alert also deletes the appointment from the list.	Don't automatically delete an entry.

Table 22: Important usability problems and suggested solutions in HERMES PDA

Marginal Usability Problems

Category / Problem	Proposed Solution(s)
General	
Users find the font size and buttons too small.	Provide larger interface elements.
Audio playback is not loud enough	If possible increase volume.
Content Layout	
Users don't know why a certain day (current day) has a black square around it and another day has blue background colour (currently selected day)	Provide clearer indications for current and selected day.
Some users get irritated, because some buttons have a thicker border (pre-selected buttons). Some think this is the button they have to press next.	Disable this feature (only necessary for hard-key navigation).
The event list contains technical information. It shows an acronym ("EID")	Remove this unnecessary information.
The wrong (and different) date formats are irritating.	Provide consistent and localised date formats.
Interaction	
"Right Here" and "No Location" are too indistinctive from the other locations	Make "Right Here" and "No Location" own buttons.
Delete can't be undone.	Undo button or "are you really sure?" box

The GPS doesn't work reliable	Clearer Feedback when the GPS has signal and when not. Time out the search for a signal after a given time period
Users don't find the button to switch to locate mode.	As suggested by users, offer a pre-selection to choose between time-based and location-based entries

Table 23: Marginal usability problems and suggested solutions in HERMES PDA

Nice-to-have

Category / Problem	Proposed Solution(s)
Navigation / Interaction	
There is no way to actually "browse" appointments like in the MyFuture desktop app. Only a cumbersome list with cryptic numbers and letters.	Offer a month-based and list view of events like in MyFuture
Users want to browse/view with the PDA events in the past. Actually they wanted it to use it as a dumbed-down version of MyPast	Allow to see past events

Table 24: Nice-to-have adjustments for HERMES PDA

4.5.4 Summary

The results of quantitative and qualitative evaluation and the usability evaluation allow the following main conclusions to be drawn:

- Some flaws in the design of the PDA application made it too difficult for the users to complete some tasks. However, it needs to be considered that, although users had a training session in the lab and a written manual, they were alone at home without the possibility of asking for help. They were invited to call in if they were at a loss, but few of them actually did so.
- In particular the usability issues reported during the trials indicate that the navigation should be more streamlined and consistent, the screen less cluttered and the text more readable, within the obvious limitations imposed by the PDA screen resolution and size.
- A redesign of the application is needed that has to focus on:
 - consistent next/back navigation,
 - clear and instructive labels,
 - clear and instructive icons,
 - fewer steps necessary to achieve task goals,
 - more robust and fail-save interaction, especially when dealing with external systems like GPS.

For further details the column "Proposed Solution(s)" in each of the tables of Section 4.5.3 is referred.

A new concept of the redesign of the PDA application is already in development. It addresses several issues by changing the basic content structure. One difference is the separation of time-based and location-based reminders. Since time-based reminders don't need a GPS-determined locations and location-based reminders don't need date and time, there is little need to mix them together. In a calendar view the user shall be able to view time-based appointments in a similar

manner as already now in MyFuture. This approach also increases the consistency across the whole HERMES system. The locations view offers a list of pre- and user-defined locations, on which the user will be reminded. Moreover the “assistant-type” of navigation (i.e. pressing “Next” on a predefined order of screens) shall be replaced by a “sheet-type” of creating a new appointment. This means, that the user gets to see an empty appointment immediately after pressing the “New Appointment” button. In whatever order the user then can add time, purpose and if needed reminder for this appointment. This new concept will be developed in close collaboration between CURE and TXT in the next months towards the second prototype.

4.6 MyPast Application

4.6.1 Quantitative Evaluation

Table 25 summarises the results retrieved from the quantitative questions of HERMES MyPast. The results indicate that the majority of Austrians and Spanish have few problems with performing the tasks with MyPast. Again, they agree that the system can be improved in supporting the task. Although to a lesser extent, again a general agreement tendency can be noticed.

Mann-Whitney Tests show the following significant differences in MyPast post-task questionnaire between age, gender and country:

- Regarding gender differences, older women are more in agreement with the statements: the system can be improved to better support them in two tasks: specifying time ($p < 0.01$) and clearing/changing filters ($p = 0.04$) than older men ($M = 14.19$ vs. 5.81 and 13.15 vs. 7.50 , respectively).
- With respect to the differences between both countries, Austrian older users are more in agreement with the item “The system supported me well in fulfilling this task” for the task “Using buttons for changing date” ($p = 0.04$) than Spanish users ($M = 13.15$ and 7.50 , respectively). Austrian participants believe the system has supported them at every moment. Also, Austrians agree more with the item “It is easy to understand the icons on the screen” ($p = 0.03$) than Spanish older ($M = 13.27$ and 7.31 , respectively).
- There were no age differences in the MyPast post-task questionnaire.

Task/ Question	Agreement in %														
	Austria					Spain					Total				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Standard browsing /sliding bar															
The system supported me well in fulfilling this task.															
	0	7,7	0,0	46,2	23	0	12,5	12,5	62,5	12,5	0	9,5	4,8	52,4	19
The system can be improved in supporting the fulfilling of the task.*															
	0	0	15,4	30,8	23,1	0	13	37,5	50	0	0	4,8	23,8	38,1	14,3
It is easy to understand the icons on the screen.															
	0	0	23,1	30,8	23,1	0	25	12,5	50	12,5	0	19	38,1	19	14,3
I was able to follow the changes on the screen easily after tapping it.															
	0	7,7	7,7	38,5	23,1	0	0	0	7,1	50	0	4,8	4,8	57,1	19
I felt comfortable doing this task.															
	0	7,7	15,4	30,8	23,1	0	0	7,1	42,9	7,1	0	4,8	14,3	47,6	19

Using buttons for changing date															
The system supported me well in fulfilling this task.															
	7,7	0,0	7,7	23	38,5	0	0	14,3	42,9	0	4,8	0	14,3	42,9	23,8
The system can be improved in supporting the fulfilling of the task.*															
	7,7	15,4	7,7	15,4	30,8	0	0	75	25	0	4,8	9,5	33,3	19	19
It is easy to understand the icons on the screen.															
	0,0	7,7	7,7	30,8	23,1	0	37,5	12,5	37,5	12,5	0	19	9,5	33,3	19
I was able to follow the changes on the screen easily after tapping it.															
	0	15,4	23,1	7,7	23,1	0	7,1	0	50	0	0	14,3	14,3	38,1	14,3
I felt comfortable doing this task.															
	0,0	7,7	7,7	23,1	30,8	0	0	13	75,0	12,5	0	4,8	9,5	42,9	23,8
Specifying time															
The system supported me well in fulfilling this task.															
	0	15,4	7,7	38,5	38,5	0	0	7,1	35,7	7,1	0	9,5	19	66,7	9,5
The system can be improved in supporting the fulfilling of the task.*															
	7,7	30,8	15,4	38,5	7,7	0	7,1	28,6	14,3	0	4,8	23,8	28,6	33,3	5
It is easy to understand the icons on the screen.															
	0	7,7	7,7	53,8	30,8	0	7,1	7,1	28,6	7,1	0	9,5	9,5	52,4	24
I was able to follow the changes on the screen easily after tapping it.															
	0	7,7	7,7	38,5	38,5	0	7,1	0	35,7	7,1	0	9,5	4,8	47,6	28,6
I felt comfortable doing this task.															
	0	7,7	23,1	38,5	30,8	0	0	0	14,3	35,7	0	4,8	14,3	47,6	28,6
Specifying people															
The system supported me well in fulfilling this task.															
	0	23,1	0	38,5	30,8	0	0	0,0	35,7	14,3	0	14,3	47,6	28,6	9,5
The system can be improved in supporting the fulfilling of the task.*															
	7,7	15,4	30,8	23,1	15	0	0	35,7	14,3	0	4,8	9,5	42,9	23,8	9,5
It is easy to understand the icons on the screen.															
	0	0	15,4	61,5	23,1	0	14,3	28,6	0	7,1	0	9,5	9,5	57,1	19,0
I was able to follow the changes on the screen easily after tapping it.															
	0	0	15,4	46,2	38,5	0	0	7,1	35,7	7,1	0	0	14,3	52,4	28,6
I felt comfortable doing this task.															
	0	0	23,1	30,8	38,5	0	0	12,5	75,0	0	0	0	19,0	47,6	23,8
Clearing/changing filters															
The system supported me well in fulfilling this task.															
	0	7,7	7,7	61,5	23,1	0	7,1	21,4	14,3	7,1	0	9,5	28,6	76,2	9,5
The system can be improved in supporting the fulfilling of the task.*															
	7,7	7,7	23,1	38,5	15,4	0	0	42,9	7,1	0	4,8	4,8	42,9	28,6	9,5
It is easy to understand the icons on the screen.															
	0	0	30,8	46,2	23,1	0	12,5	25	37,5	12,5	0	4,8	28,6	42,9	19
I was able to follow the changes on the screen easily after tapping it.															
	0	8	30,8	38,5	15,4	0	0	14,3	28,6	7,1	0	4,8	28,6	42,9	14,3
I felt comfortable doing this task.															
	0	0	30,8	53,8	15,4	0	0	12,5	75,0	0	0	0	23,8	61,9	9,5
Using search box															
The system supported me well in fulfilling this task.															
	7,7	23,1	7,7	46,2	15,4	0	0	12,5	62,5	12,5	4,8	14,3	9,5	52,4	14,3

The system can be improved in supporting the fulfilling of the task.*															
	7,7	7,7	15,4	46,2	15,4	0	12,5	25	50	0	4,8	9,5	19	47,6	9,5
It is easy to understand the icons on the screen.															
	0	7,7	30,8	53,8	7,7	0	25	50	0	12,5	0	14,3	19	52,4	10
I was able to follow the changes on the screen easily after tapping it.															
	0	0	38,5	38,5	15,4	0	0	12,5	50	25	0	0	28,6	42,9	19
I felt comfortable doing this task.															
	0	15,4	15,4	38,5	23,1	0	0	0	87,5	0	0	9,5	9,5	57,1	14,3

Table 25: Results of the quantitative evaluation of HERMES MyPast. The results indicate that the users generally succeeded in performing the given tasks. Note: * indicates reversed scale.

4.6.2 Qualitative Evaluation

As for HERMES MyFuture and the PDA application the participants were asked five qualitative questions after each task of HERMES MyPast:

1. What can be improved
2. Specific difficulties for completing the task
3. Problems (apart from the English language)
4. Difficulties following the changes on the screen (application flow)
5. Special feelings associated with completing the task

Due to the fact that the feedback of the majority of the participants did not target on a specific task but on the system in general, for presentation purposes the results are not grouped by task but by question instead. Also, questions 2, 3 and 4 are presented together, since the feedback was too indifferent to be separated.

Improvements

1. Reduce the number of steps to locate the date of the event you want to remember. They prefer that the system allows them to enter the date to search in a search box, so, they would spend less time.
2. Improve the touch interaction. The touch screen needs to be more precise and responsive.
3. Provide separate keys for diacritic symbols (umlauts, accents).
4. Provide a German/Spanish localised interface.
5. Again users had difficulties to understand the standard playback icons: Change the icon “play” (now it is represented by a triangle) with something more representative (e.g. an ear).
6. Change the "+" icons (add filter) for something else to clarify what the buttons mean.
7. Give greater visual contrast to the numbers representing the hours because they can not distinguish them very well.
8. The overview of the screen looks well although some of them would use a different background colour, e.g. something more cheerful.
9. Change the main screen. It should only include the buttons to perform an advanced search. Once the users have made the search, they want to see the events or audios saved, not before. They said that now the main screen is confusing because it does not show the current date.
10. Include the full date. Now, they do not know what day of the week is.
11. To use both keyboards (the computer and the application one) and not only the one of the application. For some people it is easier to use the computer keyboard.
12. Some of them have suggested alternative use to MyPast, e.g. as a personal diary in which they write their observations, feelings, etc. after the event has happening. Thus, they

eliminate the problem of recording a person. In this case, they would include a personal password.

Difficulties and Problems

1. The users don't remember the criterion by which they can look for if they have many events recorded.
2. The swiping interaction to move the sliding bar was found laborious.
3. Some users got irritated by the button to switch to the next day. When they needed to press it several times, they stopped one day before they should stop, because the button already showed the date of the next day.
4. In Spain the users preferred to search using the buttons and not with the sliding bar due to their difficulty in using the mouse. Some of them would prefer a wider sliding bar.
5. The date shown on the main screen can confuse because it is not the current date.

Feelings

1. Users are curious and interested in the possibilities of the application.
2. However, they do not believe they are going to use this application in the future, because they do not think it is very useful.
3. They are a little bit afraid of the ethical issues involving videotaping or audio of other people. They are afraid of a violation of privacy and asked about the possibility of recording only certain people.
4. When they didn't succeed in completing a task they felt confused, displeased and insecure.

The results described above are summarised in Table 26. It lists the major problems the users have found and the solution proposed by the users. The solutions to these problems recommended by the authors are presented in the next section (4.6.3)

Usability Evaluation).

Category/Problem	Solution proposed by users
Interaction	
Mouse usage difficult (Spain)	None
Need help in order to do the tasks of their own	Develop a manual
Colours	
Displeasing colours	Change the colours in order to get a more attractive application Use colours in order to achieve more contrast
Wording	
Filtering buttons	Change the filtering buttons to get a greater clarity. For example, change “+” icons

Table 26: Summary of problems and solutions proposed by the users in HERMES MyPast

4.6.3 Usability Evaluation

Serious Usability Problems

Category / Problem	Proposed Solution(s)
Interaction / Metaphor / Filters	
Users somehow lack understanding for the concept of filters. Some users want to use the timeline to set an hour range instead of using a filter. Many users don't know they have to press the "+" button to add a filter.	Possible merge search and filtering. Avoid the name "filter". Merge all filters into one panel (with only one "Start Search" button in the beginning). The free text search is one of the "filters".
The different position and grouping of the "+" buttons and the remove filters button is irritating to the user. They don't see they are connected.	
To set a second filter users use the same filter panel already used to set the first filter. Accordingly they overwrite the first filter	Provide textual descriptions like "add another search criterion". Maybe deactivate other filters, once one is applied.
When changing filters, users don't realise they have to press the specific filter again, once they opened the panel. They are looking for a "change" button.	Maybe: pre-open the filter as soon as the filter panel is open.
For removing a filter, users are looking for a "-" (minus) button. They don't see the "none" button. Some users tap the filter again and then press cancel.	Provide a "-" or a "Remove/Clear Search Criterion" button below the filter in the main bar or in the panel.
Users don't see or understand the "OK" button when adding a filter.	Use an "Apply Search Criterion" button where necessary. Otherwise avoid button altogether. Tapping a filter should apply it. Multiple selections are not understood by the users.
Interaction / Search	
Users try to press the search button to get the keyboard, or to "start the process"	This problem is avoided with a merged filter/search approach. Provide a text box with "press here to enter text" as a predefined label.

Table 27: Serious usability problems and suggested solutions in HERMES MyPast

Important Usability Problems

Problem	Proposed Solution(s)
Interaction / Search	
Users feel attracted by the "Search" button and want to press it on various occasions (after setting a filter, navigate to a day).	This problem is avoided with a merged filter/search approach.

Metaphor / Filters	
Users don't understand the concept of ordering things in different manners (name, importance, last visit). They don't know what to do with this.	Possibly remove this feature.
Users don't know the different meaning of the two "+" buttons.	Change labels. Maybe "Add time filter" and "Add Content Filter"
Users don't realise the effect of clearing all filters.	Provide clearer feedback, like animating the removal of filters (basic but ugly solution: message box "All Filters cleared")
Users don't realise the effect of having set a filter after pressing "OK". They are surprised that "everything vanished". Later on most of them realise the filter name instead of the "+".	
Navigation	
Users wanted to use the + buttons (add time filter) to change the date	This problem is avoided with different labelling.
Users think the button that shows the "next day" or the "previous day" is already the current day. Same applies to month buttons (especially, because it says e.g. "March" instead of only "3", which attracts more attention.)	For next day: write "Next day: 2. Feb. 2010"
Affordance	
Without explanation users don't know they can swipe to move the timeline. They don't get the concept of a "time band" you can move or scroll	Show textual description / use avatar to explain.
Users don't know how to initiate the search after typing the search word. They don't know if they need to press "search" or the green tick on the keyboard?	Simplify by offering a "start search" button and/or a dynamic button on the keyboard that takes the label of current context.

Table 28: Important usability problems in HERMES MyPast

Marginal Usability Problems

Content Layout	
Users don't see the weekday immediately. It is shown in the time line, but not in the large header.	Provide full date in header.
Navigation / Sliding Bar	
Users find the swipe gesture laborious, since they were only able to switch day by day.	Allow different scales (not only 24 hours), possibly add swiping with momentum.
Feature / Filters	
Users find it limiting being able to set only 4 different filters.	A merged search/filter approach might also offer more visual space for additional filters.
Users wanted to tap 3 months in the month filter instead of using the season filter.	Might be language related.

Table 29: Marginal usability problems in HERMES MyPast

4.6.4 Summary

The results of quantitative and qualitative evaluation and the usability evaluation allow the following main conclusions to be drawn:

- The main area of difficulties is filters. Not only don't the users understand the icons, but the whole concept and mechanism of filtering data is new to them.
- What they do understand is the concept of search. Therefore a redesign of the filter and search mechanism is recommended. Merged into one interface, it allows both standard keyword searching and filtering, i.e. searching for predefined content, such as people, places, as well as time frames.
- Small refinements in the presentation of data, especially the date headline and the navigation buttons are recommended.

For further details the columns "Proposed Solution(s)" in each of the tables of Section 4.4.3 are referred.

4.7 The Cognitive Games

The cognitive games were evaluated at INGEMA in Spain. All of the Spanish participants saw the video of the Cognitive Game, specifically the video of the Puzzle Game. The evaluator explained how the game works in general while the test person was watching the video.

First, the user of the game has to select one picture from the photo album. This photo then appears in the puzzle game. It is possible to change the picture just pressing to the bottom "Back" and selecting a new one. Furthermore, the user can choose between different levels of the game. So, if the player chooses a hard level, the picture would be divided in a large number of small pieces. Otherwise, if the player chooses an easy level, the picture would be divided in fewer and bigger pieces. When the puzzle is almost finished, the player has to press "Quit" and the score will appeared in the screen.

After seeing the video, a semi-structured interview was administered to the user. The percentages of agreement and disagreement with the questions are shown Table 30.

Agreement in % (Spain)			
Question	yes	no	don't know
Do you think that the game looks interesting?	92.9	7.1	0
Do you think that the game seems tricky?	7.1	85.7	7.1
Do you think that the game could be boring?	21.4	78.6	0
Do you think that this game would be effortful?	71.4	51.4	7.1
Do you think that it would be frustrating?	14.3	71.4	14.3
Do you think that it would be stimulating?	78.6	14.3	7.1
Do you think you would be engaged with this game?	71.4	21.4	7.1
Is clear how to select the picture to play with?	100	0	0
Is clear how to change to another different picture?	85.7	14.3	0
Are the levels of difficulty clear enough to be selected?	71.4	28.6	0
Are the levels of difficulty complex enough?	71.4	0	21.4
Is the score-screen clear enough?	92.9	7.1	0
Is the score-screen motivating enough?	64.3	21.4	14.3

Do you feel that this game would be helpful in your social relationships?	28.6	57.1	14.3
Would you play with it against your family or friends at home?	57.1	28.6	14.3
Would you play with it against your grandchildren?	71.4	14.3	14.3
Would you play with it on-line against other players?	28.6	57.1	14.3

Table 30: The results of the semi-structured interview performed in Spain about HERMES Cognitive Game indicate that users are interested in playing these games.

The following comments were collected in a semi-structured interview:

- 92.9% of the participants commented that the game was interesting, though 1 out of 14 participants said he prefers to play traditional games and he would not play with it.
- They also mentioned that they would feel more motivated to play if they could play with family or friends.
- The presentation of the score on the screen could be quite frustrating, so maybe it would be better to modify it to a non-numeric symbol or score (e.g. a message saying: “Well done, you have correctly completed this puzzle!”).
- They prefer to increase the number and variety of stimuli.
- For developing such games it must be important to try to prevent the frustration caused by age and physical limitations when elderly people play games.
- A training game before starting to play the game itself should be included.
- Some of the participants suggested increasing the difficulty of the puzzle giving the pieces different forms (not only squares) and also by adding the possibility of both moving and rotating pieces, if the player decides to play with this increased difficulty.
- Users would like to play with other people in person, especially with their grandchildren. However, they did not appreciate the idea of playing online against others.

Summarising the evaluation, it can be stated, that the participants showed interest in playing the game, especially if the grandchildren ask for it. One observation worth mentioning, that was made during the evaluation concerns memorisation of how to proceed to complete a task: The participants were asked questions about how well they understood the steps necessary to choose the level of difficulty: “Is it clear how to select the picture to play with?” and “Is clear how to change to another different picture?” When the participants who answered affirmatively were asked “How it was done?”, none of them could answer correctly (e.g. “By double clicking the image” or “By pressing the green buttons on the lower side of the screen”).

5. Ethical Issues

As stated in deliverable D.8.3, in the HERMES evaluations very personal information is used. The personal data collected will always be limited to and never exceed the scope and purposes of the project.

The process of the user trials from the ethical point of view was as follows:

1. The people who belong to the Hermes pool of users were contacted by telephone. They were given the explanations about the project and the first user trial was scheduled.
2. When the user arrived at the lab, the first step was to give more details about this phase of the study, mainly objectives, methodology and the collaboration we are asking for.
3. After answering their questions if any, the informed consent was read either by the user or by one of the professionals, and the doubts were answered.
4. All of the participants agreed to participate in this user trial and signed the informed consent (see D.8.3 for more details about it). Also one copy of the informed consent signed by the researcher was given to the users.
5. Once the informed consent was signed, the user trial started, and the cameras and the microphones started recording.
6. At the same time, all the evaluation protocols and so on were filled in with the user's code (the same the user received last year).
7. During the test, people were not obliged to give details about their own lives. For example, if we asked them to record an appointment, we insisted that they should invent the content of the appointment as we do not need to have real user appointments.

It was decided to record or take photographs of some people doing the user trial for dissemination purposes (e.g. to place them on the HERMES web site etc.). These people were informed about this action before video recording or taking the photos. Also they signed a copyright release form in order to allow their faces to be shown in the video and photographs.

The database which contains the correspondence between the codes and the name of the users is stored in only one computer and protected with a password. All of the computers at the lab which contain the users' videos, audios and databases have password protection and are only accessible for people conducting the trials. At the end of the trials this information has been saved on a hard disk protected by a password.

The informed consents have been safeguarded separately from the assessment protocols in order to prevent possible identifications.

While the evaluation of MyFuture, MyPast and the Cognitive Game applications did not have ethical implications, the PDA evaluation implied privacy risks for two reasons:

1. The participants can record other people's voices with the PDA. We warned the users about this. We asked them not to record other people's conversations both verbally before giving them the PDA, and also in the PDA manual. All of the participants have followed our warning and they have not recorded other people.
2. The participants' recordings are saved in the external memory card of the PDA. At the beginning of the trial we thought of transferring these recordings from the PDA to the PC in the lab and then delete the information in the PDA before giving it to a new participant. The problem is that this action implies some technical problems so it was not possible to do it.

And we have given the PDA to one participant with the other participants' recordings. But this has not been a very serious problem due to the fact that the recordings were not real recordings and it would be very difficult to find where the recordings are.

The way in which the information (video and audio recordings) will be interchanged between the partners of the consortium will be: in the following consortium meeting in Bradford both partners involved in the user trials (INGEMA and CURE) will take the information (audio and video recordings) on a hard disk with a password for the consortium members that need these videos in order to improve and test their algorithms. These partners copy the data to a hard disk protected by a password. At this moment, this partner will be responsible for the information in the same way as the organisation that originally recorded the data.

6. Overall Conclusions

With reference to Sections 4.3.2, 4.4.4, 4.5.4, and 4.6.4 this section summarises the conclusions drawn for the common components of MyPast and MyFuture, HERMES MyFuture, HERMES PDA and HERMES MyPast respectively. It needs to be said, that while a lot of problems were listed in the previous sections, many of them can be solved with quite simple means, e.g. changing the labelling of buttons. Moreover, often one solution covers several problems. However, there are a number of serious issues with all the applications that need to be addressed. The first prototype tried out several new concepts, like the analogue clock to set the time. Without performing a user evaluation no reliable feedback could have been collected about aspects of the system that perform well, and others that need to be changed or replaced. Overall it can be summarised that, while users had difficulties with the applications, they were interested in using them and saw the value such a system might have for their own life. Interestingly, which application in particular draw their attention varied greatly. While for some the calendar in MyFuture would be of great use, others valued the mobility of the PDA. Those users, who are already suffering of slight decrease of memory capacity or have relatives or friends suffering of such impairments, found a lot of usefulness in HERMES MyPast.

6.1 Conclusions for MyFuture

In comparison to the other applications HERMES MyFuture was evaluated best. Partly this is related to the familiar interface of a calendar, everybody has already used in the paper variant, but also thanks to a very clean and simplified user interface. Although few users asked for more complex features like setting the duration of an appointment or the possibility of categorising appointments, the generally good results of MyFuture lie in reducing complexity. Unnecessary features stripped off, it is still a functional calendar many users were imagining to use.

Parts of HERMES MyFuture that need revision or redesign include:

- Localised interface (German and Spanish)
- Improved symbols and text labels
- Better keyboard and clock
- Content repositioning in order to prevent objects from being hidden by the interacting hand
- Full screen calendar, stronger structuring elements and stronger colours to better differentiate user interface elements
- Better affordance for draggable elements (moving an appointment)
- More robust system through auto-save and undo functions

Overall the application performed fairly well in the tasks of navigating through the appointments and creating new events. Despite the problems with the keyboard and the clock mentioned above users generally succeeded in completing these tasks. While some users said they still prefer their old paper calendar, many users saw the value of an electronic calendar, when they look at the many corrections, cancellations and arrows pointing to other dates they have in their calendars.

6.2 Conclusions for PDA

One of the advantages of the PDA compared to MyFuture is the fact that many users saw in it a kind of mobile phone (which it was, in fact, too). Though only very few of the study participants

have already used a computer, almost everybody is using a mobile phone. Consequently they had comparatively less dread of using the PDA than using the desktop applications. Nevertheless, the HERMES PDA application was the cause of many difficulties and a lot of frustration among the older users. The following areas show the variety of problems the users struggled with:

- Localised interface (German and Spanish)
- Small screen elements (button size, font size)
- Inconsistent and unclear labelling of buttons and screens
- Inconsistent positioning of buttons providing next/back functionality
- Too many (and unclear) steps necessary to create a new appointment
- Error-prone interaction is causing different message or error boxes to appear or allows to user to accidentally exit the application.

Due to the fact, that most of the problems mentioned above are rooted in the way the user interacts with the application and when the application presents how much and which information to the user, a complete redesign is needed. The redesign has to focus on the usability problems mentioned above and needs to provide a pleasing and supportive user experience. The new concept indicated in Section 4.5.4 avoids the numerous steps when creating a new appointment in the current prototype and eliminates confusion due to mixing time-based and location-based reminders. After all, the mobility of the PDA application is highly appreciated among the users. Users want to use the device, but they need an application considering the task and user demands.

6.3 Conclusions for MyPast

For most users, the features HERMES MyPast offered or at least promised went far beyond their imagination. They never thought anything like that was possible and were keen on trying out the system. Only few of them expressed concerns about ethics and privacy. However, these concerns need to be taken seriously and must be addressed in the next prototype. The following list shows the main areas the application needs improvements:

- Ethical and privacy concerns need to be addressed by assuring absolute confidentiality of the collected data. Furthermore it needs to be possible to exclude certain people from being recorded (blacklist) or only allow the system to record certain people (whitelist).
- Due to the complexity of the matter, the system offers a variety of concepts that might be unknown to (older) users with no prior computer experience.
 - The timeline needs to show its affordance of being dragged
 - The concept of “filtering” data needs to be subsumed under “searching” as a search with pre-defined search parameters (people, places, time etc.)

Besides that, some issues also affecting HERMES MyFuture (and partly the PDA) also decrease usability of MyPast. In particular the following changes are necessary:

- Localised interface (German and Spanish)
- Improved symbols and text labels
- Better keyboard and clock
- Content repositioning in order to prevent objects from being hidden by the interacting hand

While some users said they would not use such a system, others saw the value of having an external memory aid. While all of the users can still live on their own, they are very well aware of their limitations or of impairments of other people they know. Considering this, they value in using MyPast. However, looking at the results of the TAQ and the conclusions drawn in Section 4.2.2, further research on this topic is necessary. A study is planned to be carried out to compare differing user acceptance of HERMES MyPast depending on whether the system has no real data the user can work with or offers a rich set of recorded events so the user can take full advantage of the possibilities of MyPast. It is expected that there is a difference in user acceptance, if a user is for example applying a certain filter and actually seeing changing results than just having to imagine these changes.

6.4 Conclusions for the Integrated System

Integrating a system consisting of HERMES MyFuture, MyPast, Cognitive Games and the PDA application from a user point of view is concentrated on two main areas:

1. How well does it work to launch the different application on the desktop device?
2. How well does it work synchronising data between the PDA and the desktop applications?

The second question can be answered positively, since in Spain synchronisation worked in most cases. The problems in Austria seem to be a local deployment issue that will hopefully be fixed for the next prototype. The first question however, was not addressed during this user trial. For the evaluation MyFuture and MyPast were launched by the interviewer from a standard Windows desktop. Cognitive Games is not available as a real application, due to the lack of a multi-touch screen. For the next prototype both issues need to be addressed in order to allow a successful integration of the system.

7. References

- Nickell, G.S., & Pinto, J.N. 1986. The Computer Attitude Scale. *Computers in Human Behavior*, 2, 301-306.
- Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D. 2003. User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly Vol. 27 No. 3*, 425-478