Mobile devices used for medical applications: Insights won from a usability study with diabetes patients

André Calero Valdez, Martina Ziefle
Human Technology Centre
RWTH Aachen University
Aachen, Germany
calero-valdez@humtec.rwth-aachen.de

Andreas Horstmann, Daniel Herding, Ulrik Schroeder
Computer-supported learning research group, Department of Computer Science
RWTH Aachen University
Aachen, Germany

Abstract

Mobile small screen technology increasingly penetrates health care and medical applications. However, usability research regarding the ease of using these devices, as well as acceptance issues did mostly address aspects of IC-Technologies in younger and healthy users. This study investigates impacts of ageing, technological expertise and domain knowledge on user interaction using the example of diabetes. First, the software for the monitoring of diabetes had been developed and implemented on a PC. The simulated diabetes assistant was displayed on a small screen device as well as on a very large display, simulating an ambient assisted living environment. In a second step, the navigation behavior of younger and older diabetes patients as well as healthy users was assessed. Results show that age and display-size have a big impact on the device interaction, independently of domain knowledge. Furthermore, it was found that simple usage of a bigger screen could be helpful for tutoring patients in using a mobile electronic living assistant. In order to assess usability of the device acceptance was measured after the experiment and was found to be positivley correlated with effectiveness.

1. Introduction

Future health care systems will face enormous challenges within the next few years. Drastically demographic changes such as increased life expectancy, improved medical healthcare reduced fertility rates, will lead to a growing number of frail older people, who will need medical treatment and long term care provided by official health care systems[1]. One of the central challenges for political and health care systems in the 21st century is therefore to master the demands of an ageing society.

Mobile technologies in combination with ambient technologies offer enormous potential to improve patients' medical care and reduce the financial pressure on health care systems alongside progress in biomedical sciences or genetics.

Supporting older patients in keeping mobility and maintaining an independent life style at home will only be achievable by systems that monitor and control health-related information that are portable and communicable and fit into the ecology of existing mobile devices as well as ambient assistant living.

Usable interfaces and slick user experience will be critical factors for acceptance, sustainability and competitive capacity of any mobile technical system. This pressure for improvement becomes ever more obvious in regard of demographic changes, world wide increasing life expectancy and the resulting increase of older users.

Over 20 percent of Germany's population already passed the age of 65 in 2008 and 18 percent more will follow within 30 years, squeezing more than one third of the population into the "senior users" category [2]. And similar forecasts apply to many western European countries [1]. Voluntary usage of electronic devices will ever more become a rarity as work or everyday life requirements will increasingly encompass handling these devices [3]. This impact will be even stronger concerning medical appliances of mobile devices.

Age related illnesses like diabetes accompany both demographic change and sedentary lifestyle and will change medical care and age appropriate independent domestic care to only be economically realized if done so through technical solutions [4].

The successes of combining mobile technologies and medical home care technologies – promising as they are – still have to overcome some obstacles. Designing such solutions in a way that is both self-explanatory and usable for heterogeneous user groups has not yet been realized [5]. Device development is still dominantly technical-oriented and criteria of usability and learnability are, if at all mostly applied subordinately.

This is a direct consequence of the development being in the hand of computer scientists and engineers and a lack of integration and harmonization with psychological and ergonomic knowledge, especially in regard to necessities, capabilities and cognitive structures of the end users.

Controlling and monitoring health states either on a small screen or in an ambient electronic device is desired but apparently obstructed by deficient

acceptance and willingness to use especially by the target user group of elderly people [6].

In this study we want to address these potential obstacles and evaluate the potential of medical electronic devices for younger and older patients. For two reasons we chose diabetes as application context. One reason is that diabetes is one of the major civilization diseases, which not only affects older patients, but also has increasing incidence rates in younger adults. A second reason is that one of the authors himself is a diabetic, and is highly motivated to develop user-centered technology, which is able to truly support patients handling their disease.

In the following, first, the importance of diabetes as a main civilization disease is outlined, followed by the status of knowledge regarding the usability of small screen devices, in combination with the impact of the diverse user group, which is using these devices. Lastly the research questions addressed by a usability experiment are presented.

1.1. Diabetes and secondary disorders

Diabetes mellitus is a metabolism dysfunction, which affects about 8 million people (10% of the population) in Germany alone and for the year 2010 an increase of up to 10 million affected is expected. Diabetes and secondary disorder treatment already covers 20% of Germany's compulsory health insurance funds expenditure. Diabetes alone is expected to cause a hole of 40 billion Euros in Germany's health care budget in 2010[6].

Diabetes mellitus type 1 occurs mostly in younger adults between the age of 5 and 50, but can also occur later in life. The main symptom of type-1-diabetes is absolute insulin deficiency. Only 8-10% of all diabetes patients are type-1-diabetics[8].

Diabetes mellitus type 2 occurs mostly after the 40th year of one's life but increasingly often occurs in younger people today. Main causes for type-2-diabetes are obesity and lack of physical exercise. The main symptom of type-2-diabetes is body cell insulin resistance. This relative insulin deficiency turns into an absolute insulin deficiency when the pancreas' insulin production collapses. About 90% of all diabetes patients are type-2-diabetics [8]. A persisting maltreated type-2-diabetes can lead to blindness, renal failure, amputations and heart failure and may also cause cognitive deficiencies. Especially in patients older than 50 years executive functions and the neurocognitive processing speed are affected. Episodic memory, word flow and semantic memory though seem to be unaffected [9]. Fortunately type-2-diabetes can sometimes be treated by diet and physical exercise alone, since reduction of body weight can in some cases cause full remission [10].

Almost all diabetes therapies require or can at least be assisted with mobile small screen devices, since monitoring, persisting and analyzing of blood glucose levels, insulin dosage and caloric intake increases therapy success, since their correlation and

behavior can vary drastically between individuals [11].

1.2. The usability demands in small screen devices

It is a central claim that mobile devices are designed to be in line with users' specificity and diversity. However, the intelligent interface design of mobile devices, which meets the demands and abilities of especially older users, is an extremely sophisticated task. Aging itself represents a highly complex process. Not all users age in the same way, and the onset of aging processes as well as the consequences show considerable differences across human beings. Design approaches should therefore take the user-perspective seriously.

The miniaturization of small screen devices may also contribute to usability shortcomings. Beyond handling and visibility problems, the restricted screen space allows only little information to be displayed at a time. By this, memory load is increased. In addition, orientation in the menu is complicated, because users do not experience how the menu might be "spatially" structured and how the functions are arranged [12][12]. In hierarchically structured menus disorientation occurs when complexity is high with respect to the depth and breadth of menu levels [14]. With respect to effects of users' age, the profound changes in sensory, physical, psychomotor and cognitive functioning over the life span are well known [15]. These changes may account for older adults' lower performance when using technical devices. Furthermore, due to a different upbringing, the majority of older adults possess limited computer knowledge, which may also account for differences in computer-based performance [16][17]. However, it was found that age-related decreases could be compensated by expertise [18]. Thus, performance of older adults can be just as good as that of younger adults when they can rely on elaborated domain-specific knowledge.

1.3. Questions addressed and logic of experiment

The present experimental study addresses the impact of aging and domain knowledge of diabetes on task performance when using a small-screen diabetes living assistant in comparison to using an ambient assisted living screen. Can users compensate their performance disadvantages by facilitating better domain knowledge? This question is addressed by comparing younger and older diabetes patients to younger and older healthy users. In order to learn if participants would accept the monitoring of their diseases in an ambient environment, two conditions were realized. Participants had to use the electronic diabetes device on a simulated small-screen device (via a touch screen) and on a large display (simulating a wall), which could also be used by touch input. Here it was of interest how participants evaluated the

usefulness of the application and their acceptance to use the device in an ambient environment as well.

Although this study was primarily designed exploratory, the following outcomes were expected:

Younger users due to aging impacts on both cognitive and perceptual abilities outperform older users [5][9].

Users with higher technical expertise outperform users with lower technical expertise due to conceptional transfer of navigation user interfaces.

Users with higher domain knowledge outperform users with lower domain knowledge due to improved understanding of tasks and higher appreciation of purpose behind function of the user interface (e.g. a diabetes patient knows about bread unit calculation and it's importance) (see [16][18][19]).

2. Method

The objective of the study was to understand influence of aging, technical expertise and domain knowledge on task performance on a living assistant for diabetes patients and to gain knowledge of determining factors on navigation performance in small screen touch enabled devices as well as in ambient assisted living environments. Since the current study claims to extend earlier research, efforts were made to keep the method very similar to that used before. In this section the conceptual design and the procedure of the experiment are described.

2.1. Experimental variables

In our study we considered four independent and five dependent variables. The first independent variable we examined is user age in order to measure influence on both task effectiveness (i.e. the amount of tasks solved correctly) and efficiency (the amount of time required to solve a task).

Secondly the influence of an ambient assisted living screen vs. a mobile phone sized small screen on performance and learnability of the application was studied.

In order to relate effects on usage performance to prior experience with modern technology or experience with medical technology, participants were asked about their experience with different nonmedical (mobile phone, computer, GPS navigation, digital camera, microwave oven, alarm clock, gaming console) and medical devices (blood glucose meter, hearing aid, blood pressure meter, heart rate monitor, in-house emergency call). Since the study was performed using a small screen device user experience with mobile phones was surveyed thoroughly as well.

To measure impact of domain knowledge on effectiveness and efficiency, knowledge of four key health parameters (blood glucose, HbA_{1c}, blood pressure, body fat percentage) were surveyed and aggregated as an independent variable.

As dependent variables five performance criteria were measured: success rate, total steps carried out to solve the tasks, detour steps, total time, and time per step. Success rate (effectiveness) is measured as the percentage of successfully performed task steps of each task. Total steps are the amount of program interactions performed for a certain task. Pressing on the non-interactive background or missing a button is not included in total steps. Detour steps are all program interactions that do not lead into solving the task at hand (navigation failures, accidentally pressed buttons, unnecessary repeated input). Total time is the amount of time the users take to finish all tasks cumulated. Time per step is the average amount of time a user takes between two program interactions. A lower value represents a faster navigation pace but not necessarily a better navigation performance.

Users' device acceptance was assessed with original items from the Technology Acceptance Model of Davis (see [20]). The perceived ease of use (PEU) implies 'the extent to which a person believes that using a particular system would be free of effort', and secondly, the perceived usefulness (PU) which is defined as 'the extent to which a person believes that using a particular system would enhance his or her job performance'[20]. The validity and reliability of TAM items had been proven by several empirical studies, and also showed satisfactory values in this study. The five presented PEU items had to be judged on a six-point Likert-scale ranging from 1 (very easy) to 6 (very hard). PU items were rated on a six point Likert scale ranging from 1 (very useful) to 6 (not useful at all). Smaller values would reflect better acceptance of the device.

2.2 Participants

A total of twenty-two adults took part in this study. Among those, were eleven young adults (4 males, 7 females) with a mean age of 34.1 years (SD = 10.1; range: 25 – 55 years), eleven older adults (2 males, 9 females) with a mean age of 65.9 years (SD = 8.2; range: 58 – 87 years). The younger participants were mostly university students. Medium aged and older adults were reached by advertisement in local newspapers and through an exhibition on a local public diabetes convention and covered a broad range of professions and educational levels.

Twelve participants were non-diabetic adults, who were mostly recruited through their social networks (3 males, 9 females, mean age = 44.8; SD = 18.4; range: 25 – 71 years). The ten participants (3 males, 7 females; mean age = 56.7; SD = 16.8; range: 26 – 87 years) split up into a group of five participants diagnosed with Type-1-Diabetes (1 male, 4 females; mean age = 43.6; SD = 13.6; range: 26 – 64 years) and five participants suffering from Type-2-Diabetes (2 males, 3 females; mean age = 67.8 SD = 11.4; range: 59 – 87 years).

Regarding the recruitment of older participants a prototypical ideal participant "diabetic but otherwise healthy senior" was aimed at. All medium aged and older adults participating were either active parts of the work force or otherwise mentally fit and not

hampered by stronger age-related sensory and psychomotor limitations.

Technical expertise (TE) was surveyed by measuring the Perceived Ease of Use (PEU) and Usage Frequency (UF). Both PEU and UF were measured on a Six-Point Likert Scale. PEU was examined with questions like "How easy to use is for you..." (1 = very easy, 2 = easy, 3 = rather easy, 4 = rather hard, 5 = hard, 6 = very hard). UF was similarly examined with questions like "How often do you use a..." (1 = Daily, 2 = 2 - 3 times a week, 3 = once per week, 4 = 1 - 2 times a month, 5 = 1 - 2 times a year, 6 = never). Total expertise (mean = 4.1, SD = 1.1, range 2 – 6, N = 22) is calculated as the square root of the product of the mean of all PEU and the mean of all usage frequency (UF) in order to reflect a value that is also on a 6-point-Likert scale where 1 reflects a value with highest usage frequency and highest PEU and 6 represent the exact opposite. Intermediary values reflect both PEU and UF, with a tendency to rank equal values of PEU and UF higher than differing values. A person who uses a computer often, but finds it hard to use, scores lower (i.e. better), than a user that uses a computer not as often but ranks ease of use on a similar level. This effect is desired in order to account for misperception due to very extreme usage frequency (e.g. a person that only uses the computer to write letters once a year, but finds this task very easy). A person that finds a computer quite hard to use but uses it daily is expected to have a better computer expertise but a perception bias to find usage harder.

Significant differences only exist in expertise with technology (TE) and mobile phone usage ($p < 0.01$). Differences in experience with medical technology are not significant for the two age groups.

Domain Knowledge (DK) (mean = 4.0; SD = 1.7; range = 2 – 6, N = 22) was surveyed with a Five-Point-Likert scale ("How well do you know…"). Answers ranged from 1 = "very precise" to 5 = "I don't know", which considers the item at hand to be completely out of the knowledge of the person contrasting to 4, where the person has heard about the measurement before. t-Test-analysis of Domain Knowledge regarding to age shows that domain knowledge is not correlated with age (t = -0.243 p > 0.05). This is expected since in the case of diabetes domain knowledge should rather depend on the period of time being affected by the illness than on the numerical age. Differences in domain knowledge between age groups are not significant (see Table 1).

Table I. T-test for Equality of Means: technical expertise and domain knowledge regarding differences between the age groups

	t	dF	sig (2t)	MD	SE
TE	-4.06	20	<0.01	-1.39	0.34
DK	-0.243	20	>0.025	-0.18	0.75

2.3. Development of a small screen device for diabetes patients

Our goal is to develop a portable device that supports diabetes patients in their therapy and in their everyday lives, which could also be used in an AAL-Environment. Instead of creating a specialized application that is only useful for certain diabetes patients, we decided to include features that are required for the different types of the disease. On the first start, the user has to set up the application by entering his characteristic values (such as drugs that she/he has to use regularly for his therapy).

The most important feature in everyday use is the so-called diabetes diary (Diabetes-Tagebuch). Every time the patient measures or influences his blood sugar concentration, she/he is supposed to insert the data into the diary, using a wizard-based input mechanism. For instance, when the patient has had a meal and measured blood sugar she/he creates a new entry. She/he enters the time, the measured blood sugar, and the bread units' equivalent of the ingested meal (1 bread unit (BU) = 12 g of carbohydrates).
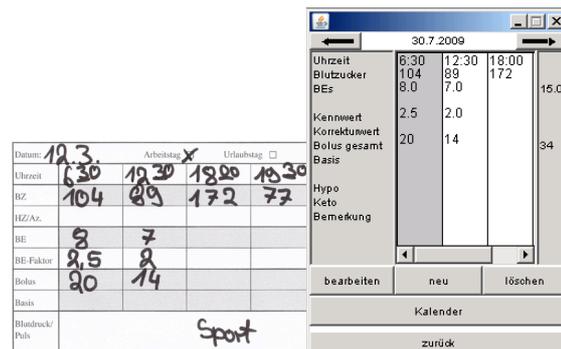

Figure 1. The paper version of the diary (left) and the diary function of our application (right).

In the wizard, she/he simply skips values that were not relevant for this particular entry. The entered data is then shown in a column of a table. The tabular representation is based on the layout of the paper-made diaries that are in common use in Germany (see Figure 1.), and that a large part of the target users are already familiar with. 61% of diabetes patients in Germany are using a diary to record their values; 91% of these are keeping their diaries on paper.

Another application feature that is inspired by a paper template is the health passport, or Gesundheits-Pass. After each quarterly examination, the doctor writes down the results into this booklet. Like the diabetes diary, the values (e.g. blood pressure, HbA_{1c} value, weight, etc.) are entered in a table. We only show the values of one quarter of a year, while the paper version has columns for four quarters, due to screen size constraints.

For a successful diabetes therapy, it is important to teach the patients a basic knowledge of the nutrient contents of groceries. Especially patients who inject insulin need to calculate their drug dosage on the basis of the food they consume. Most people weigh the ingredients and then look up the bread units (BU) per gram in a nutrition table to calculate the total bread units for their diabetes diary.

We included a feature that supports the user in looking up and calculating these values. She/he can choose a grocery from a predefined list, and then enter a weight or volume. The application then displays the bread units and kilocalories (see Figure 2.). This is repeated for all ingredients of the meal, and the resulting BU sum is copied into his diary. The user can store favorite dishes for later reuse.
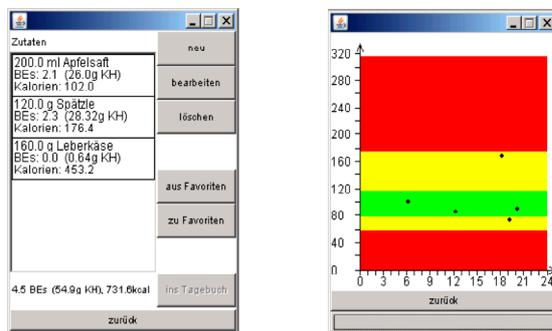

Figure 2. The bread unit calculator (left), the plotter showing one-day overview of the blood sugar concentration (right)

There are two features that help the patient to keep track of the progress of his therapy. The first one is the so-called plotter, which shows the course of measured values in a history diagram (see Figure 2.). The other feature is called screener. It displays the latest entry of characteristic values such as body weight and blood sugar concentration, and compares it to the previous entry. Colored arrows visualize the tendency.

For recording purposes the Jacareto-Framework was used.

2.4 Experimental procedure

In order to test the research model and to determine the effects of screen size, domain knowledge and age variables on performance and learnability, an experimental setting with a simulated small-screen-device and wall-screen device was conducted.

At the outset participants completed a paper-based questionnaire concerning demographical information (age, gender, educational achievement) and information about the familiarity with common technical devices (TE) as well as domain knowledge in diabetes (DK). The assessing of demographic data was performed paper-based.

After completing the survey participants were asked to perform a set of five tasks on both simulated devices. Each task regarded a different main function of the device an all medical values that were to be entered into the device were predefined on a paper based task description to create equal preconditions. Task-information was printed on hardcopy and was available throughout the experiment. The fastest user completed all five tasks in about six minutes and all participants finished under 30 minutes per trial. Users who started the experiment on the simulated small-screen-device then repeated the whole process for the large screen and vice versa.

After completion of the experimental tasks participants were asked to rate the perceived ease of use (PEU) and perceived usefulness (PU) of the used functions in the simulated device to assess the users acceptance of the software and the uses display technology.

2.5. Wall screen device simulation (or AAL-screen)

If new display technologies (such as OLED-Display foils) become available at reasonable pricing, whole walls, ceilings and floors could be used as displays. Our aim was to simulate an interface that can be projected onto the wall of the patients' living room, so that the interface blends in into the natural living environment of the future user. The wall acts as a touch-sensitive interface although only a smaller part of it is truly sensitive to touch input in this case.

Figure 3. AAL-Screen prototype and IR-pen.

In order to simulate this ambient assisted living environment, we decided to put up a back projection wall, which was considered a prototype for a wall-touch-screen interface (screen size 1600x1200mm, mounted at 600mm height). As projection surface we used PLEXIGLAS® 7D006, which is specifically designed for IR-based multi-touch back projection applications. Projection was performed with a Hitachi CP-A100W short distance projector with a resolution of 1024x768 pixels. In the living room scenario six of these are mounted in tiles to create a wall that in its

whole acts as a screen. Input was handled using a simple IR-LED-pen (VISHAY TSAL6400 LED) with a pressure sensitive tip. Upon lightly pressing the tip of the pen onto any surface (see Figure 3), IR light is emitted from the top of the pen. Touch tracking was performed using two Wii Remotes' IR-Cameras (both used 70% of their tracking resolution of 1024x768) and Wiimote-Whiteboard as a mouse driver. The Wii-Remote has a built-in hardware to track up to four IR-dots with its camera.

Our application was about five times larger on the AAL-Screen than on the simulated small screen (width = 38.2cm; height = 49.8cm) but used the same resolution of 245x319 pixels. The rest of the screen was filled with an abstract picture.

Participants were allowed to move freely in front of the wall but due to restrictions in room size could never move further away from the screen than 3 meters. Lighting conditions were kept the same by choosing a room with no exterior lighting and a fixed interior lighting system. If the participant required any corrective lenses, wearing those was obliged throughout the experiment.

2.6. Small screen device simulation

The diabetes living assistant was simulated as a software solution one a PC running Windows XP connected to an Iiyama AX3819UT touch screen (15" TFT-display, display resolution 1024x768 pixels). The simulated device spanned over 245x319 pixels (width = 7.27cm; height = 9.47cm) and was displayed in the center of the screen. The rest of the visible screen was covered with an opaque paper cutout to simulate a device like feel. All interaction was to be performed with a touch pen, suitable for this display in order to maintain comparability between trials.

Participants were seated on a height-adjustable chair in a comfortable seating position. In order to control viewing conditions, participants were not allowed to choose viewing angle, viewing distance or inclination of the TFT-Monitor.

2.7. Experimental tasks

Five Tasks were to be solved by the participants. In particular users were first asked to setup the "freshly unboxed" device and enter information about their current therapy (i.e. insulin type, dosage and schedule, etc.) with given fake values. Secondly users were asked to fill out their health passport. After completing the first two tasks participants should enter three blood glucose measurements along with bread units of a meal and insulin dosage for three times of a given day into the digital diary. Again all values were predefined. The fourth glucose measuring was preceded by a task in which the users had to calculate the bread units of a given meal using the BU-Calculator of the device. The last task required the user to simply view the daily blood glucose graph in the plotter of the device. All tasks

were described in natural language but data for all input forms was given numerically.

- Example for 'digital diary'-task: '*After finishing configuration of your device, daily blood glucose measurements can be stored in the devices digital diary. Please enter the following measurement into the digital diary. This morning at 9:20am. Blood glucose level 123; consumed 3 bread units, no correction of insulin dosage; no basal-insulin dosage; no hypo- or ketoacidosis was measured*'.

- Example for 'BE-Calculator-task': '*You are hungry and want to eat some fish sticks (200grams) and have a glass of apple juice (200ml). Please calculate the bread units for this meal using the BE-Calculator of the device.*'

3. Results

Results of this study were analyzed by multiple measures ANOVA, bivariate correlations and multivariate analysis of variance with a level of significance set at 5%.

The result section is designed as follows: first, we assess correlative relations and impact of individual factors (age, health status (HS), technical expertise (TE) domain knowledge (DK), etc.) on users' performance; second, a deeper analysis of aging effects on performance is conducted in regard to effects that occur between trials.

3.1 Effects of age, technical expertise and domain knowledge on performance

Table II. **Bivariate correlations between age and user characteristics and performance**

	Success Rate	Total steps	Detour Steps	Total Time	Time per step
Age	-0.664**	0.616**	0.472*	0.231	0.693**
HS	-0.179	0.342	0.181	-0.102	0.421
TE	-0.449*	0.330	0.244	0.476*	0.320
DK	-0.53	-0.167	0.097	0.314	-0.244

$**p < 0.01, *p < 0.05$

In order to reveal existing relationships between user characteristics, age, health status, technical expertise and domain knowledge were related in regard to performance criterias.

3.1.1. Relationship between Factors and Performance. To get a first insight into the data, correlations (Spearman rank) between individual variables and performance measures of the first trial (small and wall screen) were carried out.

Correlation analysis shows that only age and technical expertise show a significant correlation with performance measures (see Table 2). Younger age is highly correlated with better effectiveness (r = -

0.664) and efficiency. Younger users need less total steps (r = 0.616), make less navigation errors (r = 0.472) and have a faster navigation pace (r = 0.693).

Expertise with technology is mostly correlated with effectiveness (r = -0.449) such that users with better expertise are more effective than users that are more inexperienced. This correlation does only affect one efficiency measurement significantly (i.e. total time r = 0.476), which also shows that higher expertise is related with better performance.

Apparently domain knowledge (r = 0.53) and health status (r = -0.179) seem to have an unexpected adverse effect on effectiveness, but further correlation analysis shows that age is highly correlated with health status (r = 0.509) , and health status itself is highly correlated with domain knowledge (r = 0.799).

To examine how domain knowledge and health status predict performance, two analyses of covariance (ANCOVA) were conducted using 'domain knowledge' and 'health status' as a covariate. The ANCOVA revealed no significant main effect for domain knowledge (F = 1.817; p > 0.05) with 'domain knowledge as a covariate. Choosing 'health status' also reveals no significant main effect on effectiveness (F = 1.808, p > 0.05).

3.2. Effects of screen size and age on learnability

In order to reveal effects of different screens sizes, technical expertise and age on performance, two subsequent trials of the experiment were analyzed in regard to each other as well as in regard to age, technical expertise and screen size of the first trial.

3.2.1 Effects on effectiveness. Using repeated measures ANOVA analysis, interesting effects of screen size were found regarding effectiveness of the user. Effectiveness did not change significantly between trials (see Table 3) and did not differ significantly between users who started on different screen sizes (see Figure 4).
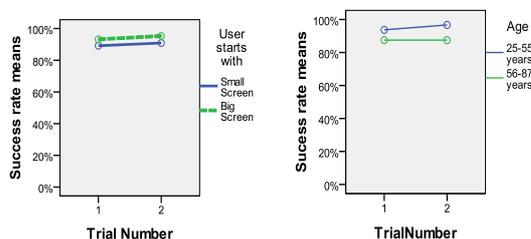

Figure 4. Success rate means between trials in regard to screen size and user age.

Age, technical expertise and domain knowledge similarly showed no significant effect on differences between trials but age showed a significant difference between groups (see Table 3).

Table III. **Wilks' Lambda Table for repeated measures analysis between trials and tests of between subjects effects regarding success rate**

	Wilks' Lamda Table between Trials			Test of between subjects effects	
	Value	F	Sig.	F	Sig.
SS[1]	0.91	1.50	> 0.05	2.39	> 0.05
Age	0.94	0.93	> 0.05	11.265	< 0.05
TE	0.92	1.30	> 0.05	0.948	> 0.05

3.2.2. Effects on efficiency. In contrast to effectiveness efficiency measurements were drastically different regarding the examined factors. Users who started the experiment on the big screen were significantly faster than users who started on the small screen. They were also able to maintain their low total time on the small screen device when working on it in the second trial. The other users were able to catch up in their second trial when using the bigger screen as well (see Figure 5).

Table IV. **Wilks' Lambda Table for repeated measures analysis between trials and tests of between subjects effects regarding detour steps**

	Wilks' Lamda Table between Trials			Test of between subjects effects	
	Value	F	Sig.	F	Sig.
SS	0.40	23.28	< 0.01	3.03	> 0.05
Age	0.316	32.54	< 0.01	22.6	< 0.01
TE	0.42	20.94	< 0.01	9.39	< 0.05

Similar effects can be found when looking into other efficiency measures. All users were able to perform the second trial with significantly less detour steps (see Figure 5). Users with both high and low expertise in technology cut the amounts of navigational mistakes in half. Additionally users who started on the big screen, who very technically experienced or who were young each completed both trials with significantly less detour steps than the respective other groups.

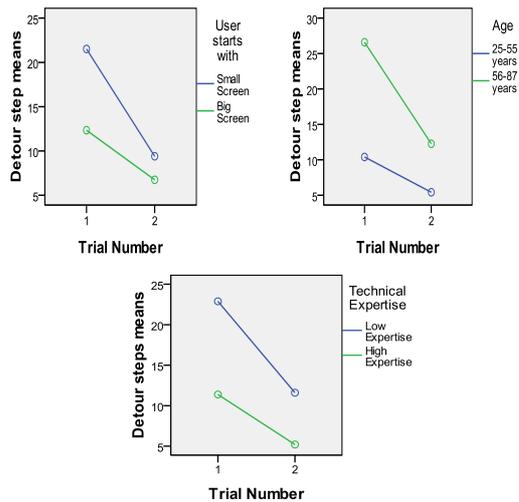[1] SS = Screen size, TE = Experience with technology

Figure 5. Detour step means between trials in regard to screen size and age (top row) and in regard to technical expertise (bottom).

The effect of screen size especially becomes obvious when examining the navigation pace (time per step). Users that started on the big screen took more time between each interaction when using the small screen device. In contrast users who started on the small screen significantly strode faster through the tasks when using the wall screen (see Figure 6). Younger users did not change their navigation pace significantly between trials, while older users improved their time between interactions. This could reflect a more careful approach to an initial interaction (see Figure 6). Interestingly users with high and low technical expertise performed almost equally fast in their second trial in contrast to an almost doubled demand for time during the first trial for inexperienced participants (see Figure 6).
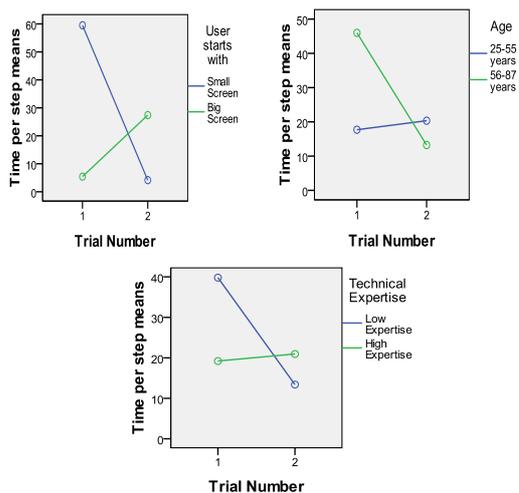

Figure 6. Time per step (seconds) means between trials in regard to screen size (left) and user age(right) and technical expertise (bottom).

3.3 Effects of navigational performance on acceptance

In order to understand how different factors (age, health status, expertise with technology, success rate in experiment) influence user acceptance of our simulated device non-parametrical analysis of correlations (Spearman's rho) was conducted (see Table 5). Both age and success rate show a significant correlation with acceptance of the device. Users that are more successful show higher acceptance of the device than unsuccessful users (r = -0.507). Increasing age also seems to lead to higher acceptance of the device (r = -0.460). Interestingly health status seems not to have any effect (r = 0.027) on acceptance at all (p>0.05).

Table V.   **Bivariate correlation between user characteristics, performance and acceptance (*p<0.05)**

	Age	DK	HS	TE	SR
Acceptance	-0.460*	0.200	0.027	0.276	-0.507*

To examine how age and success rate affect acceptance, two analyses of covariance (ANCOVA) were conducted using 'age' and 'success rate' as a covariate. The ANCOVA analysis revealed no significant main effect for user age (F = 3.502; p > 0.05) with 'user age' as a covariate. Choosing 'success rate' as a covariate also reveals no significant main effect on acceptance (F = 3,378, p > 0.05).

Linear regression though contradicts this finding: both age and success rate explain 65.5% of the variance of acceptance of the simulated device, but success rate is a stronger predictor for acceptance (β = -0.486, p < 0.05) than user age (β = 0.241, p > 0.05). This suggests, that high performance in initial usage of a device might have a high impact on acceptance of a new device.

4. Discussion and Conclusion

The present experimental study was conducted to provide deeper understanding of small-screen-device menu navigation performance in respect to age, technical expertise and domain knowledge in a medical context and its' possible transferability to an ambient assisted living environment. A total of twenty-two participants accomplished five tasks designed for a diabetes living assistant in two different contexts. In order to analyze individual factors that may differentially affect user's performance, domain knowledge and their health status were surveyed and related to performance outcomes in two trials.

4.1. Impact of User Characteristics on Performance

The study confirmed the large impact of user characteristics on small-screen-device menu navigation performance as well as on AAL-screen navigation. The first influential factor found in the analyses was the user's age. Aged users tend to make more errors, require more time between interactions and are less effective in solving the tasks at hand.

The effect of technical expertise on performance was also confirmed in this study, although a big part of it's effect was lost during the second trial of the experiment. A higher technological expertise seems to make an intial run of our device significantly easier.

Domain knowledge and health status show no significant influence onto the measured performance criteria, but comparisons of means denote a correlation might exist. Thus, we can assume that the navigational performance is indeed facilitated if users show a high expertise in disease-related knowledge. The fact that we could not statistically confirm this on the significance level set is presumably due to the comparably small sample size. Future studies will therefore examine the relationship between the computer and disease-knowledge by enlarging the participant group.

4.2. Impact of screen size on learnability

Using a large screen has shown to improve navigation efficiency drastically. Especially the amount of time required to solve tasks in a first trial varied tremendously between different screen sizes. Navigational errors and detour steps were both significantly smaller if users started the experiment on a larger screen. This also reflects in a higher navigation pace on the larger screen in comparison to the smaller screen. Users who switched from small to large screen improved their efficiency measures to a large extent. This could especially be used in tutorial settings to improve acceptance of small screen devices by conducting training on large screens either with a health care professional or at home. A simple perceptual explanation fails to justify the effect, since resolutions were equal on both screens. Although taking a closer look at the screen, leads to bigger images on the retina, information density is not changed. These results could be explained as a cognitive effect because a large screen allows users to move freely in front of the screen, to choose either an overview perspective or a detail perspective when closing in on the wall for interaction which could allow the user to incorporate real space navigation (i.e. moving in the room) into the interface. Observations of the experimenter during the test sessions confirmed that participants indeed took advantage of the possibility to use a greater distance to the "wall", in order to improve overview and navigational control.

4.3. Qualitative Insights

Participants reported after the experiment that they had enjoyed trying out our software, and especially our hardware. All users particularly relished working with the AAL-Screen and agreed that they would like to have a product like this installed at home, if it were available for a reasonable price. Especially interacting with a large screen was perceived much easier by older adults, although no increase in resolution of the application was present.

4.4. Potential Applications and Limitations of this Study

The findings underline earlier research regarding usability and aging. Further research is required to prove or increase understanding of influence of domain knowledge or diabetic status on user interaction. Further analysis of task related problems and identification of required neuropsychological characteristics for different tasks might lead into input for further research. In this context, the comprehensibility of UI component labeling is of interest, as well as the investigation of the underlying mental model of device usage, which also could have impacted performance. Finally, individuals' disease-related coping styles should be incorporated into research scope.

However, the findings as promising as they are, also have to be looked at critically, especially as the participants here represented a kind of best-case scenario, which may not represent the whole group of ill and disease-limited patients.

A best-case homogeneous user group might have led to skewed findings compared to different populations. Older users were all mentally fit, of relatively high education and mostly all of them had experience with computers.

All diabetic participants were highly interested in contributing to advancement of usability of diabetes small-screen-devices and thus highly motivated. Such perfect preconditions cannot normally be assumed and devices must work perfectly even when the user is distressed, afraid or even in a case of emergency.

Simulating a small-screen-device on a 15" display is a simplification of the situational context, since holding and handling a real small-screen device requires more cognitive and motor load (coordination of both hands). Therefore all performance measures are probably an overestimation of real life performance especially in regard to using fake values and not real user data. Users might be more concerned about using the device correctly and thus be more disturbed by unexpected behavior in a medical device. This sandbox-like operation might have led to a more carefree approach.

5. References

[1] Siebert, H. Economic Policy for Aging Societies, Berlin: Springer (2002)

[2] Federal Statistical Office Germany: Germanys population until 2050 – 11th coordinated population projection, Wiesbaden 2006, http://www.destatis.de/ (Access Date: December 2009)

[3] Arning, K. & Ziefle, M. Understanding age differences in PDA acceptance and performance, in Computers in Human Behaviour, 23, pp. 2904-2927 (2007)

[4] Leonhardt, S. „Personal Healthcare Devices", in Satyen Mukherjee et al. (eds.), „AmIware : Hardware Technology Drivers of Ambient Intelligence", Chapter 6.1, pp. 349-370, Springer, Dordrecht, NL (2006)

[5] Ziefle, M. & Bay, S., Transgenerational Designs in Mobile Technology, In Handbook of Research on User Interface Design and Evaluation for Mobile Technology, Vol. 1, pp. 122-141 (2008)

[6] Röcker, C.; Janse, M. Portolan, N. & Streitz, N. User requirments for intelligent home environments. a scenario-driven approach and empirical cross- cultural study. ACM International Conference Proceedings Series, 121, pp. 111-116. (2005)

[7] German Diabetes Union: Report on Public Health - Diabetes (2007), http://www.deutsche-diabetes-gesellschaft.de/ (Access Date: December 2009)

[8] Tililil et al., Guideline of the German Diabetes Association for treatment of DM type 2, p.177 (1998)

[9] Yeung, Sophie E.; Fisher, Ashley L.; Dixon, Roger A., Exploring effects of type 2 diabetes on cognitive functioning in older adults, Neuropsychology. Vol 23(1), Jan 2009, 1-9 (2009)

[10] Dixon et. al. Adjustable Gastric Banding and Conventional Therapy for Type 2 Diabetes, The journal of the American Medical Association 2008, Vol. 299(3): pp. 316-323. (2008)

[11] Tatara, N. et. al.: A Review of Mobile Terminal-Based Applications for Self-Management of Patients with Diabetes, eHealth, Telemedicine, and Social Medicine, 2009. eTELEMED '09. Feb. 2009 Page(s):166 – 175

[12] Lin, D.M. Age differences in the performance of hypertext perusal. In Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting. Santa Monica: Human Factors Society, pp. 211-215. (2001)

[13] Ziefle, M., Aging and Mobile Displays: Challenges and requirements for age-sensitive electronic information designs. Proceedings of the 9th International Conference on Work With Computer Systems, WWCS 2009, Beijing, China. (2009)

[14] Parush, A. & Yuviler-Gavish, T., Web navigation structures in cellular phones. The depth-breadth trade-off issue. International Journal of Human-Computer Studies, 60, 753-770. (2004)

[15] Craik, F.I. and Salthouse, T.A., Handbook of Aging and Cognition. Hillsdale, N.J., Lawrence Erlbaum (1992)

[16] Downing, R.W.; Moore, J.L. & Brown, S.W. The effects and interaction of spatial visualization and domain expertise on information seeking. Computers in Human Behaviour, 21, 195-209 (2005)

[17] Czaja, S. J. & Sharit, J. Age differences in the Performance of Computer-based work. Psychology and Aging, 8, 59-67 (1993)

[18] Morrow, D.; Miller, L.S.; Ridolfo, H.; Kokayeff, N.; Chang, D.; Fischer, U. & Stine-Morrow, E. Expertise and aging in a pilot decision making task. In Proceedings of the Human Factors and Ergonomics Society 48th Annual meeting, pp. 228-232 (2004)

[19] Calero-Valdez, A.; Ziefle, M.; Horstmann, A.; Herding, D. & Schroeder, U. Effects of aging and domain knowledge on usability in small screen devices for diabetes patients. LNCS Vol. 5889 HCI and Usability for e-Inclusion, pp. 366-386, USAB (2009).

[20] Davis F. D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13, 319--340 (1989)

6. Acknowledgements