Behaviour & Information Technology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tbit20

Effects of age, cognitive, and personal factors on PDA menu navigation performance

Katrin Arning [a] & Martina Ziefle [a]
[a] Human Technology Centre (HumTec), RWTH Aachen University , Aachen, Germany
Published online: 23 Apr 2009.

Behaviour & Information Technology
Vol. 28, No. 3, May–June 2009, 251–268

Effects of age, cognitive, and personal factors on PDA menu navigation performance

Katrin Arning* and Martina Ziefle

Human Technology Centre (HumTec), RWTH Aachen University, Aachen, Germany

(*Received October 2006; final version received August 2007*)

The present study examined the PDA menu navigation performance of younger and older adults. The research focus was directed to the understanding of the combination and interaction of user characteristics with PDA menu navigation performance. In order to detail individual factors that influence user's performance, users' age, spatial ability, verbal memory, the confidence to use technical devices and computer-expertise were studied and related to performance outcomes. Younger and older adults, experienced with the usage of different technical devices, but PDA novices, had to complete four common tasks in the digital diary of an emulated PDA and users' effectiveness and efficiency were surveyed. Even though the users of both age groups had a comparably high computer experience, participants had considerable difficulties to solve the PDA tasks successfully. Especially older adults were strongly disadvantaged when navigating through the PDA menu. Among the user characteristics which were revealed to be essential for performance, spatial abilities were the best predictor to explain PDA performance. In addition, an adequate mental representation of the PDA data structure was decisive for navigation performance, especially in the older adult group.

Keywords: user characteristics; age; personal digital assistant (PDA); menu navigation performance; spatial ability; mental model

1. Introduction

Currently, two major developments characterise modern societies. The first development refers to the profound demographic change with an increasingly aging population. In Germany, for example, due to low birth rates and increased life expectancies (Leonhardt 2006) the population will decrease from currently 82 million to 66 million in 2040 (Siebert 2001). According to population-statistical forecasts, the average age of the population will increase from 40 years (1999) to 46 years (2025) and 49 years in 2050. Overall, the proportion of people older than 65 years is assumed to rise from 25% to 56%. Similar forecasts also apply for many other Western countries (Siebert 2001). The second development is the ongoing proliferation of information and communication technology (ICT) in many parts of daily life. The usage of ICT is not longer limited to professional areas or to technology-prone user groups. Nowadays broader user groups have access to ICT and its effective use has become an essential requirement in today's working and private life. Parallel to the increasing diffusion of ICT, the technology itself has changed rapidly. In the 1980s, stationary PCs were predominantly used; the 1990s were characterised by the Internet and a worldwide information access. Nowadays, mobile communication technologies have become predominant, e.g. mobile- and smart-phones, communicators and electronic organisers, which show continuously increasing rates of growth each year (Shiffler *et al.* 2005). Apart from business applications, mobile technologies are assumed to specifically support older adults in their daily lives. In this context, mobile devices are applicable for medical monitoring, as wayfinding aids, as general memory aids or for conventional personal data management (e.g. a diary with a reminder for medical appointments or birthdays).

Both developments, the ageing trend and the proliferation of ICT, impose considerable demands on modern societies and affect private and work lives, training-on-the-job programs, educational and schooling programs, software and hardware concerns up to design of technical devices. It is a central claim that these mobile devices are designed to correspond to older users' specificity *and* diversity. However, referring to (1) the special demands of older users and (2) the menu design of applications in mobile devices like PDAs, there are many unsettled questions remaining.

First, as the user group of older adults receives more attention due to demographic changes, an interface design is required, which is easily accessible for older users. The utilisation and acceptance of

*Corresponding author. Email: arning@humtec.rwth-aachen.de

ISSN 0144-929X print/ISSN 1362-3001 online

DOI: 10.1080/01449290701679395
http://www.informaworld.com

mobile technologies by older users depends substantially on the extent to which these devices satisfy the specific needs of older users, i.e. their cognitive abilities and attitudes. However, the interface design of mobile devices, which suits the demands and abilities of older users, is a highly challenging task, because the process of aging itself is extremely differential and complex. Not all users age in the same way, as the onset of ageing processes and the affected cognitive abilities and functions show considerable inter-individual differences (e.g. Christensen *et al.* 1994, Birren *et al.* 2006).

Second, the structure and design of menus is a central issue of human computer interaction research, which is mainly focused on the analysis of users' menu navigation. The problem most often cited in menu navigation is disorientation and distraction from the correct navigational path (Conklin 1987). Users get lost in a menu system, without knowing where they are, where to go next, and how to get back to previous navigation routes or known parts in the menu. This especially applies for menus implemented in small screen devices, presumably because of the restricted screen space, where users only see parts of the menu they have to navigate through (e.g. Ziefle and Bay 2006). In hierarchically designed menus disorientation occurs when menu complexity is high with respect to the depth and breadth of menu levels (e.g. Parush and Yuviler-Gavish 2004, Ziefle and Bay 2005) and the naming and allocation of functions to categories in the menu is ambiguous (e.g. Schröder and Ziefle 2005, Pitman and Payne 2006). In network or hypertext structures, disorientation is even more likely to occur. In contrast to the exclusively hierarchically structured data in cell phones, PDA-applications provide both, network structures as well as hierarchical menu parts. Although navigational control in network systems should be higher – because users can directly select the nodes they want to visit – the nonlinear structure of hypertext and the number of cross-references is likely to provoke even more disorientation (e.g. Lin 2001, Ziefle *et al.* 2007). Up to now navigation behaviour in hierarchical menus of small screen devices (e.g. mobile phones) has been studied thoroughly (e.g. Omori *et al.* 2002, Bay and Ziefle 2003, Parush and Yuviler-Gavish 2004). However, in contrary to the profound knowledge about menu navigation in mobile phones and computer systems, only restricted knowledge is present regarding menu navigation in PDA. (Dorn *et al.* 2004, Goodman *et al.* 2004, Arning and Ziefle 2007a,b).

1.1. *Individual differences and their impact on navigation performance*

Considering the aforementioned navigation problems in complex menu structures, not all users face the same difficulties. Menu navigation performance is strongly influenced by a number of individual factors, which differentially benefit or hamper navigation (e.g. Pak 2001, Downing *et al.* 2005). Individual differences that affect computer interaction include cognitive abilities, emotional factors, personality, gender, age, and computer-expertise. In many studies, performance was mainly determined by spatial abilities, especially by *spatial visualisation abilities*, which can be measured by psychometric testing procedures (e.g. Kit of Factor-Referenced Cognitive Tests; Ekstrom *et al.* 1976). Users with a high spatial ability outperformed users with low spatial abilities in completing computer tasks (e.g. Westerman 1995), using hypertext (e.g. Lin 2001) or mobile phones (e.g. Ziefle and Bay 2006).

In terms of the spatial metaphor (Kim and Hirtle 1995) the benefiting nature of spatial abilities is explained by having an (in)appropriate mental representation – or mental model – of the devices' information structure. The concept of mental models stems from cognitive psychology and is defined as users' internal representation of how systems work and how they are structured (Gentner and Stevens 1983). Mental models are assumed to be formed through increasing expertise with the system, combining knowledge of prior experience, cognitive schemata, and problem-solving strategies (Sasse 1991). Users with high spatial abilities construct a mental representation – a mental model – of the system's structure and thereby stay aware of their relative position while navigating through the system (e.g. Sein *et al.* 1993). Users who fail to develop an adequate mental representation of the menu structure are highly probable to experience navigational disorientation and show a poorer menu navigation performance (Edwards and Hardman 1989), whereas an appropriate model supports menu navigation performance (Gray 1990, Bay and Ziefle 2003, Bay and Ziefle 2008). The existence of an adequate system representation is especially essential for the usage of small screen devices. Due to the limited display size, presenting only few menu functions on a screen at one time, high cognitive demands arise, as users have to navigate through a menu, which is not transparent with respect to its structure, form, or extent.

Beyond the impact of an adequate mental model and spatial abilities, *verbal memory* is a further cognitive factor, which explains performance variability in the usage of small screen devices. Menu navigation requires the memorisation of names and the exact location of categories and functions. Furthermore, during navigation users have to remember visited nodes and the correct sequence of navigational steps required for the accomplishment of navigational tasks. Accordingly, individuals with higher verbal memory abilities showed a superior navigation performance in technical menus (Larson and Czerwinski 1998, Bay and Ziefle 2003, Oulasvirta 2004, Ziefle and Bay 2006).

Another influential factor of menu navigation performance is *computer-expertise* (Davies 1994, Ye and Salvendy 1994, Mayer 1997, Ziefle 2002, Downing *et al*. 2005, Mitchell *et al*. 2005). Computer experts reach a higher effectiveness and efficiency in menu navigation performance, because highly organised knowledge structures and a deeper system understanding lead to a detailed perception of problems, to flexible problem solutions, and to an efficient usage of short- and long-term memory functions (Chi *et al*. 1988).

Besides cognitive variables, personal characteristics like *self-efficacy* are able to explain performance variability. Computer self-efficacy refers to the individual confidence in one's capability to use technical devices. According to Bandura (1986), self-efficacy is a better predictor of performance than actual capability because self-percepts are highly instrumental in determining how individuals actually use their knowledge and skills. Studies have shown that high scores in computer self-efficacy are related to navigational performance and the reported ease of use (e.g. Brosnan 1998, Liu and Grandon 2003, Arning and Ziefle 2007b).

Moreover, the factor *gender* is widely discussed to play an important role in the explanation of computer attitude and performance. Women usually report lower levels of computer-related self-efficacy and a higher computer anxiety (e.g. Busch 1995), which, in turn, reduces the probability of active computer interaction and may lead to a generally lower computer-expertise level. Likewise, Rodger and Pendharkar (2004) reported noticeable differences in computer experience between males and females and suggest '... when both females and males are well trained in the use of computers the user's gender may not have an impact on the performance ...' (p. 539). According to that, gender effects might be explained by a reduced computer-expertise in female users.

1.2. *Age differences in computer interaction*

The relevant literature concordantly demonstrates profound changes in sensory, physical, psychomotor, and cognitive functioning over the life span (e.g. Fisk and Rogers 1997, Birren *et al*. 2006). Age-related decreases can be found in processing speed, i.e. the speed with which problems are solved and mental operations carried out (e.g. Czaja 1997, Park and Schwarz 1999). Another profound decline over the life span concerns spatial abilities (e.g. Vicente *et al*. 1987, Stanney and Salvendy 1995) as well as spatial memory (Cherry and Park 1993). Regarding visual abilities, the eyes' ability to accommodate and adapt to light changes declines and the sensitivity to glare and reflections increases (e.g. Craik and Salthouse 1992, Haegerstrom-Portnoy *et al*. 1999). A final consideration is concerned with the aging of psychomotor abilities. Even though older adults take longer to adapt to novel tasks and to learn new motor routines, practise and training can considerably reduce decrements in motor performance (e.g. Vercruyssen 1997). Furthermore, older adults are mostly less experienced in computer usage and technical understanding. As a result, the majority of older adults possess limited computer knowledge about how to operate technical devices. Those age-related changes may increase the difficulty of computer-related tasks and may therefore account for differences in computer-based performance (Czaja and Sharit 1993). However it was found that age-related decreases could be compensated by expertise (Morrow *et al*. 2004).

Contrary to current popular stereotypes, older users do take great interest in computer and mobile technology (e.g. Noyes and Sheard 2003, Arning and Ziefle 2006), but they are often confronted with the problem of a diminished usability, as was shown in recent studies (e.g. Czaja and Sharit 1993, Morrell *et al*. 2000, Arning and Ziefle 2007a). Therefore many older users reject the utilisation of ICT. However, the factor 'chronological age' cannot be used as an explaining variable itself, when it comes to usability research and design for older users. The underlying factors contributing to the aging disadvantages have to be adequately considered in research. Hence it would be more helpful to learn which cognitive processes and life-span changes account for the age-related differences in technology interaction.

1.3. *Research question*

The present experimental study addresses three basic topics.

(1) **Aging impact**. First, we aimed at a detailed examination of ageing effects in combination with the usage of current small screen devices. Therefore, users were surveyed with respect to their spatial ability, verbal memory, the self-confidence regarding the usage of technology and the extent of computer experience.
(2) **The specific information structure of PDA menus**: In contrast to the profound knowledge about menu complexity in hierarchically structured mobile phones and computer systems, there is only restricted knowledge about PDA

menu navigation. Therefore this study aimed for an analysis of navigation performance in a PDA menu, which contains hierarchical and network-structured menu parts.

(3) **Benchmark-testing procedure**. As technical devices are increasingly included in current working processes, we aimed at examining users, which are active members of the workforce. Thus, comparably young and well-trained older users (age range: 50–69 years) were examined. To mirror the ageing impact, the performance of a young adult group (students with an age range of 18–27 years) was contrasted. The underlying methodological approach is benchmark testing procedure, with both age groups representing a 'best case scenario'.

Participants worked on common PDA tasks and were surveyed with respect to effectiveness and efficiency. Furthermore, we assessed users' mental model of the PDA menu-structure in order to relate performance outcomes to the cognitive model of the menu structure. Although the present study was primarily of exploratory nature, the following outcomes were expected:

- Older users are outperformed by younger users with regard to cognitive abilities and PDA menu navigation performance.
- PDA menu navigation performance is explained by variability in individual cognitive variables.
- A higher navigational performance is associated with a having an adequate mental model of the PDA-menu-structure.

2. Method

2.1. Participants

A total of 32 participants, 16 young adults between 18 and 27 years ($M = 23.8$, SD $= 2.8$) and 16 older adults between 50 and 69 years ($M = 56.4$, SD $= 6.7$) took part in the study. Each group was balanced by gender, with eight females and eight males in each age group. Younger subjects were recruited at university and fulfilled a course requirement. Older subjects were reached through the social network of older participants and received a small gift after completing the experiment. Regarding the recruitment of participants a benchmark procedure was pursued: we aimed at the 'younger and healthy seniors'. All of them were active parts of the work force, mentally fit and not hampered by stronger age-related sensory and psychomotor limitations. All participants had a normal or corrected to normal visual acuity which was at least 14/14 (Snellen), as tested by the TITMUS Vision Tester®. In addition, participants reported not to have any history of eye-illnesses. In order to assure a proper input device usage, participants had to complete a few aiming tasks with the mouse (hitting targets on the screen surface with the mouse) prior to the task completion. Participants had no difficulties to use the mouse efficiently, which confirms participants' reported mouse experience. Furthermore, no psychomotor limitations were reported.

In the younger group, students of different academic fields (engineering, informatics, social sciences, psychology and communication science) took part. In order to achieve a representative profile of educational levels, the older adults came from different professional fields (engineers, administrative officers, secretaries, (high school) teachers, nurses, architects, physiotherapists, physicians, and psychiatrists). Ruling out further confounding variables, a careful screening of participants' abilities was undertaken. First, all participants were in good physical and mental health conditions. Visual acuity was normal or corrected to normal and no history of eye-illness was reported. Second, the computer experience for the older and younger group was determined. Participants indicated the length of time since they were using a computer, the frequency of computer (and mouse) usage and also the perceived ease of computer use. No significant differences were revealed between the age groups (Table 1).

2.2. Tasks and procedure

The experimental tasks referred to the digital diary – a standard software application implemented in each commercially available PDA. Participants worked on four prototypic tasks (Figure 1) and had to enter (task 1 and 3) and to postpone (task 2 and 4) appointments in the digital diary within a time limit of 3 minutes per task. An example of an entry-task is: '*You have just made an appointment at the coiffeur Salon Beauté on Monday, 04/18/04, 9:00–11:30 AM. Now you want to enter this appointment into your digital diary*'. A typical

Table 1. The surveyed facets of computer-experience.*

	Young	Old	p
Length of computer use [in years]	7.2 ($s = 2.2$)	7.2 ($s = 3.8$)	>0.05
Frequency of use [max = 4]	3.8 ($s = 0.5$)	3.3 ($s = 1.0$)	>0.05
Ease of use [max = 4]	3.3 ($s = 0.6$)	3.0 ($s = 0.7$)	>0.05

*Length of computer use (measured in years), frequency of use (rated on a 4 point scale: 1 = less than once a week, 2 = once per week; 3 = 2–3 times per week, 4 = daily) and ease of use (rated on a 4 point scale: 1 = very easy, 2 = quite easy, 3 = quite difficult and 4 = very difficult).

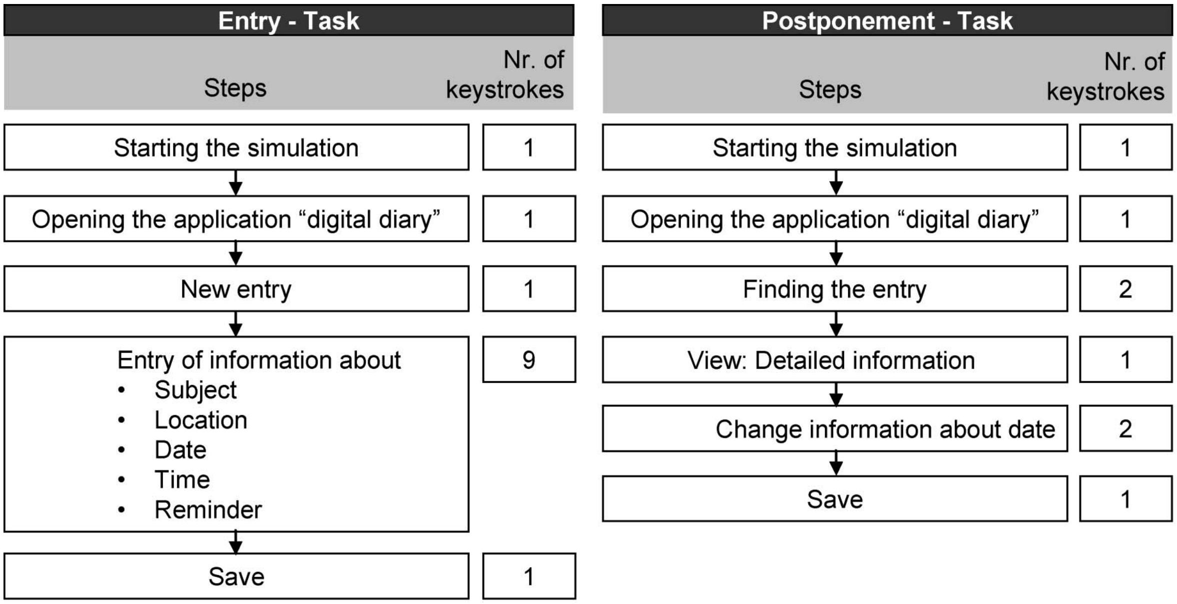

Figure 1. Flowchart with the type and number of steps required for a new-entry (left) and a postponement-task (right).

postponement-task was: '*You have an appointment for dinner with Tom and Mary on Sunday, 03/21/04 at 8 PM at the Restaurant Al Muretto. Tom just called you and asked to postpone the dinner appointment. Now you'll have an appointment on Sunday, 28.03.04 at the same restaurant at 8 PM. You want to enter these changes into your digital diary*'.

The instructions were presented in German, the native language of the participants. The minimum number of steps (i.e. clicks) to solve the tasks was determined. To enter a new appointment into the digital diary, a minimum of 15 steps was necessary. The postponement of an appointment required 8 steps. Figure 1 illustrates a flowchart with the type and number of steps required for a new-entry-task (task 1, 3) on the left side and for the postponing task type (task 2, 4) on the right side.

Before participants worked with the PDA, they completed a computer-based questionnaire. In order to familiarise participants with the experimental setting and to control that older adults were – as they reported to do – able to handle the computer mouse, the questionnaire had to be filled in by using the mouse. Participants were instructed to use the mouse with their preferred hand (all were right-handed). The questionnaire assessed demographic variables (age, profession) and experience with computers by requesting the length of computer use, the frequency and the ease of use. The length of computer usage ('For how long have you been working with ... ?') was measured in years, the frequency ('How often do you use/work with ... ?') was rated on a 4 point scale (1 = less than once a week, 2 = once a week; 3 = 2–3 times a week, 4 = daily). Finally, the ease of use ('Using a mobile phone/computer is ... for me') was assessed on a scale with four answering modes (1 = very easy, 2 = quite easy, 3 = quite difficult and 4 = very difficult). Regarding the valid assessment of computer-expertise, different facets of computer experience (length of computer use, frequency and perceived ease of computer use) have to be taken into consideration. Therefore, a total measure was built by comprising all of the three facets. The ratings of length of computer use, frequency and reported ease of computer use were aggregated multiplicatively to build the variable "computer experience" (Computer experience = length of computer use $\times$ frequency of computer use $\times$ ease of computer use).

2.3. *Apparatus*

The PDA (iToshiba Pocket PC e740, system software Windows CE) was simulated as a software solution and run on a Dell Inspiron 8100 notebook PC that was connected to a TFT-screen (TFT-LCD Iiyama TXA 3841, TN, 15′, with a display resolution of 1024 $\times$ 768 pixels). The software prototype exactly corresponded to the real device, in size (chassis 80 $\times$ 125 mm), display size (3.5′), font size (9 pt for functions, and 11 pt (bold) for category headers), menu structure, and operational keys. In Figure 2 a screenshot of the PDA used in the experiment is provided. A specific logging tool was developed for experimental purposes, which guaranteed a precise and nonintrusive measurement of user navigation behaviour to derive parameters of effectiveness and efficiency. Based on these detailed

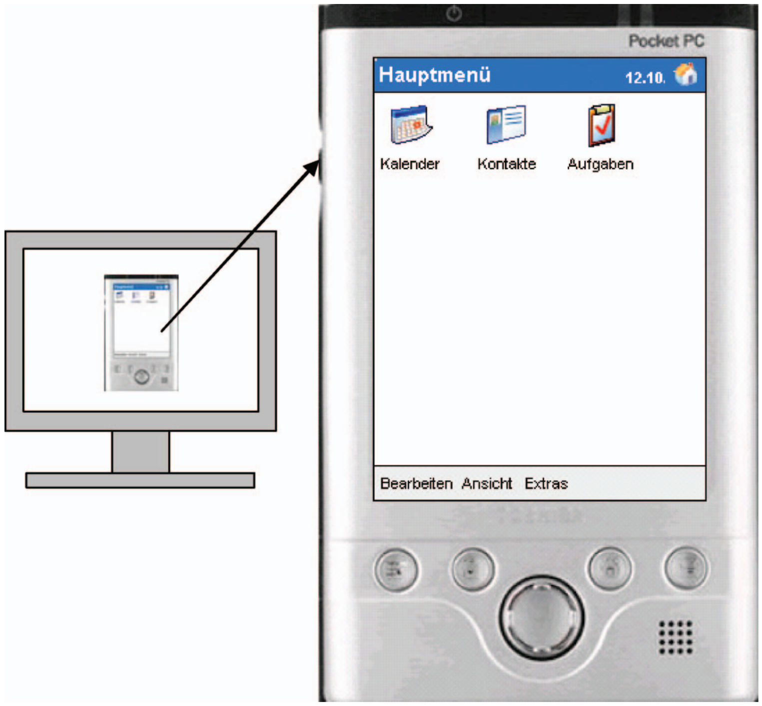

Figure 2. Screenshot of the PDA-emulation (left: PDA-emulation on the screen (scaled down view); right: view of the main PDA menu (scaled down view).

loggings it was possible to analyse individual navigational routes and to reconstruct the problem solving behaviour of participants, even when the tasks were not solved successfully in the given time limit.

Participants sat on a height-adjustable chair in a comfortable seating position. Controlling for inaccuracies due to stylus usage, participants completed the tasks with the computer mouse, as all participants were highly experienced in using the mouse. In order to optimise viewing conditions, participants were allowed to individually choose the viewing distance and the inclination of the TFT-monitor, where the PDA-emulation was displayed. Corrective lenses were worn throughout the experiment.

2.4. *Variables*

As independent variable, user age was examined, contrasting the performance of older and younger participants. As dependent measures task effectiveness and efficiency were analysed according to the standard for usability (EN ISO 9241-11, 1998). For task effectiveness, the percentage of successfully solved tasks was summed up. As efficiency measures, (1) the time needed to process the tasks, (2) the number of detour steps, and (3) the number of nodes revisited were collected. The number of detour steps describes the difference between the number of navigation steps actually done and the minimum number of steps necessary to solve the task. The number of revisited nodes indicates that users, in search of a specific function, selected nodes or functions more than once. Both, the number of detour steps and of nodes revisited, were shown to sensitively reflect users' disorientation behaviour in menus of technical devices (e.g. Tauscher and Greenberg 1997, Lin 2001, Ziefle and Bay 2005). Additionally, users' spatial ability, verbal memory, technical self-confidence (STC), the reported experience with computers and the users' mental representation of the PDA-structure were surveyed (see section 2.5).

2.5. *Materials*

2.5.1. *Spatial ability*

To measure spatial ability, participants completed a spatial visualisation test taken from the Kit of Factor-Referenced Cognitive Tests (Paperfolding test; Ekstrom *et al*. 1976), a widely used and validated ability test in the field of human-computer interaction (e.g. Evans and Simkin 1989, Dillon and Watson 1996, Hegarty *et al*. 2003). Each item shows successive drawings of two or three folds made in a square sheet of paper. The final drawing shows a hole punched in the folded paper. Participants had to select one of five drawings to show how the punched sheet would appear when fully opened (Figure 3). The maximum score to be reached was 20.

2.5.2. *Verbal memory*

To assess verbal memory abilities, a verbal memory test was conducted (Bay and Ziefle 2003). In order to diminish semantic processing of the stimuli and further elaborative processes, participants had to memorise Turkish nouns. These nouns are completely unfamiliar to native speakers of the German language. Fifteen Turkish nouns had to be memorised and were presented successively for 3 s on a computer screen. Immediately after the presentation (delay < 0.5 s), subjects had to recognise each target word among three phonologically and visually similar distractors. Here, the maximum score to be reached was 15 (Figure 4).

2.5.3. *Technical self-confidence*

The Subjective Technical Confidence (STC; Beier 1999) measures the subjective confidence of a person regarding the competency when using technology. The subjects were given the short version of the test containing eight items (e.g. 'Usually, I cope with technical problems successfully') which had to be rated on a five-point scale, ranging from 1 = totally disagree to 5 = totally agree. According to Beier's own results the reliability of the STC short version is high (Cronbach's $\alpha = 0.89$). In the present study the reliability of the STC-Scale was even higher (Cronbach's $\alpha = 0.91$).

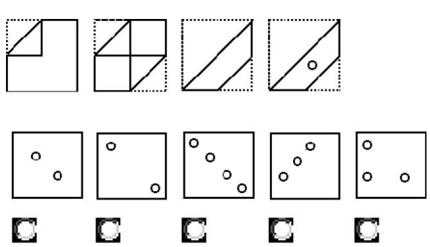

Figure 3. Item example of the Paperfolding Test (Ekstrom *et al.* 1976).

TATIL

☐ TAFIL ☐ TARAK ☐ TITAL ☐ TATIL

Figure 4. Item example from the verbal memory test (Bay and Ziefle 2003).

2.5.4. *Mental model*

After completing the four experimental tasks, the users' mental representation of the menu structure of the digital diary was assessed. It was a basic question if the younger and older adults had developed an idea of how the device might be structured, and if they have built up a representation of the mental room they were navigating through. Moreover it was of central interest to link this knowledge to performance outcomes in order to determine if an adequate representation of the data structure is essential for a proper navigation performance. We assumed that a correct representation of the system structure would "guide" the user while navigation, leading to a better navigational effectiveness and efficiency.

The assessment of mental models is difficult as users are not necessarily aware of having a specific representation and the representations are not directly observable. Different techniques can be applied to uncover mental representations, e.g. participants' verbalisation, drawings, or assessing the mental model with the card-sorting-technique (e.g. Sasse 1991, Evans *et al.* 2001, Ziefle and Bay 2004). As these methods are affected by restrictions in participants' verbalisation, drawing, and structuring abilities, we applied another procedure to assess the mental model. In order to support participants in the visualisation of the systems' structure, participants were offered drawings of five sample principles and asked to choose the most adequate structure principle (Figure 5). The models were developed in an earlier study, in which older and younger users were asked to draw the information structure of different technical devices (Bay and Ziefle 2003, Ziefle and Bay 2004). According to their proposals, the models used in this study were developed.

Participants were not given complicated details and form characteristics about the systems' structure in order to avoid biases. They were encouraged to look carefully at the different sample principles and to choose spontaneously the one, which corresponds most of all to their mental model of the PDA. If participants were not able to choose a model type (because they had not developed a representation at all or could not decide which model to choose), they were asked to state this. One could critically assume that the completion of this task requires a high level of abstraction abilities in order to understand the representations of the information architecture given in the PDA. On the other hand one could assume that the cognitive demand is lower, because participants had merely to reidentify the PDA structure according to their view. As observed, participants – the younger and the older – had no difficulties with the instruction,

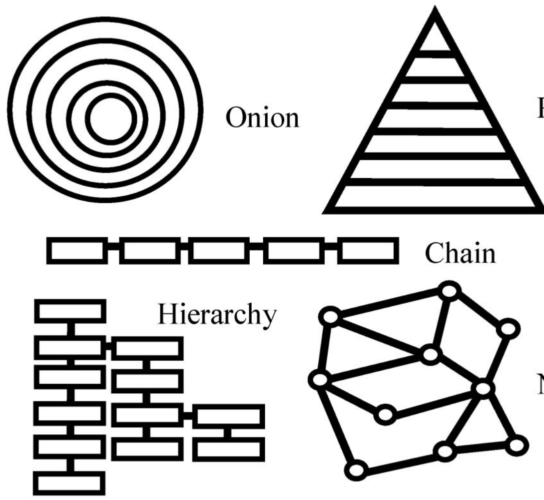

Figure 5. Structure principles of mental models.

and knew immediately what they were asked to do (even those which indicated to have no idea of how the model was structured).

Three sample principles were inadequate with respect to the real structure given in the PDA: an onion-shaped, a pyramid-shaped, and a chain-like structure. In contrast, network and hierarchy structures were considered as being adequate to reflect the PDA-structure, as both structure types existed in the digital diary. However, mental models reflect individual and subjective knowledge and are not necessarily congruent with the real structure of a device (Sasse 1991). Therefore one aim of the study was analyse the congruence between the subjective user model and the objective model of the information structure in relation to PDA menu navigation.

3. Results

Results were analysed by bivariate correlations, uni- and multivariate analyses of variance, and multivariate regression analyses with a level of significance set at 5%. The significance of the omnibus F-Tests in the MANOVA analyses was taken from Pillai values. In order to determine associations between performance outcomes and user characteristics, correlations were carried out (Pearson values were used for interval-scaled data, Spearman-Rho values for ordinal-scaled data and Eta values for nominal-scaled data). As all four experimental tasks stem from the diary application of a PDA and therefore share the same semantic context, the performance in the four tasks was comprised and the means for the user groups are reported.

First, outcomes regarding user characteristics and psychometric measurements are presented (section 3.1). Second, navigation performance outcomes are reported and related to user characteristics (section 3.2). Third, looking for the most powerful predictor of performance outcomes among different user characteristics, the findings from multiple regression analysis are described (section 3.3). Fourth, insights from the presence and adequacy of users' mental models are presented and related to navigation performance (section 3.4).

3.1. User characteristics, their outcomes and interrelation

Table 2 shows the correlation matrix for the psychometric test scores, revealing a very strong association between spatial ability and verbal memory (Kendall's tau $\beta = 0.43$, $p < 0.01$). Interestingly, high spatial ability was significantly interrelated with technology-related self-confidence (Kendall's tau $\beta = 0.43$, $p < 0.01$), but this was not found for verbal memory abilities. Apparently, high spatial abilities coact with a high technical self-confidence, which is not given in persons with high verbal memory abilities. Computer-expertise showed a strong association with technical self-confidence (Kendall's tau $\beta = 0.51$, $p < .01$), but no relations to other user characteristics.

Older subjects scored significantly lower in spatial ability and verbal memory than younger subjects (F $(2,27) = 26.5$; $p < 0.001$, Figure 6). In comparison to younger users the spatial ability of older participants was 45.5% lower ($M_{\text{Old}} = 8.4$, $M_{\text{Young}} = 15.4$), verbal memory abilities were reduced for about 37% ($M_{\text{Old}} = 7.6$; $M_{\text{Young}} = 12.0$). Male and female subjects did not differ, neither with respect to spatial nor verbal memory abilities ($F(2,27) = 0.3$, n.s.).

Table 2. Correlation matrix for user characteristics and cognitive ability tests ($n = 32$).

	Age	Sex	STC	Spatial ability	Verbal memory	Expertise
Age	–	−0.06	−0.42*	−0.80**	−0.67**	−0.04
Sex	–	–	−0.35	−0.03	−0.12	−0.23
STC	–	–	–	0.53**	0.30	0.51**
Spatial ability	–	–	–	–	0.75**	0.25
Verbal memory	–	–	–	–	–	0.17
Expertise	–	–	–	–	–	–

*$p < 0.05$, **$p < 0.001$.
STC, subjective technical confidence.

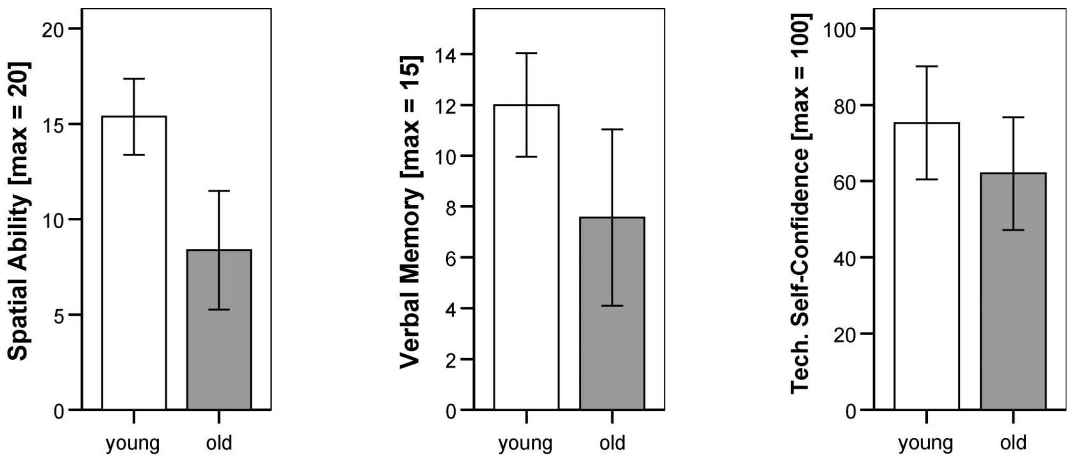

Figure 6. Age differences in psychometric measures and subjective technical confidence (STC). (left: spatial ability; centre: verbal memory; right: STC).

Interestingly, a significant ($F(1,31) = 5.2$, $p < 0.05$) disordinal interaction of gender and age was observed for technical self-confidence, with younger males reporting the highest technology-related self-confidence in contrast to female and the older users (Figure 7).

Two specific findings in this context are worth mentioning. First, younger users reached, on average, 'only' 75 out of 100 points – contradicting the popular stereotype that the young generation is highly confident and enthusiastic with respect to the usage of technology (older users reached on average 62 points). Second, although there were no differences in cognitive abilities and in device interaction performance between male and female users, there were significant gender differences in technology-related self-confidence.

Summarising the results so far, age-related differences are present with respect to the degree of spatial ability and verbal memory. Persons with high spatial competencies also show a high verbal memory and a high technical self-confidence. Contrary to cognitive abilities, where no gender differences were found, the younger males' technical self-confidence was highly pronounced. Interestingly, the two age groups showed no differences in self reported computer experience ($t(30) = 0.4$, n.s.).

3.2. *Task performance and user characteristics*

In this part, the effectiveness and efficiency (comprising all four tasks) with regard to user characteristics are reported. Regarding the analysis of task efficiency, two analysing strategies can be pursued: One strategy only includes users that successfully accomplished the tasks; the other one includes all participants, independent from their task performance. Methodologically, both procedures can be applied: On a first sight, the selection of successful task performers for further analysis seems to be advantageous, since their results can be directly related to effectiveness outcomes. However, if only a rather small proportion of participants was able to solve a certain task, only small and unequal samples would have entered statistical analyses. Moreover, from an ergonomic point of view, it is more insightful to learn about usability barriers in the PDA menu and to consider all user actions – even if users failed in the end. An analysis showed that result patterns for "task solvers" and "nonsolvers" were similar. Therefore, task efficiency was statistically analysed comprising the total group.

The ANOVA results showed highly significant age differences in all performance measures. Regarding the effectiveness in the four tasks (Figure 8), older subjects solved, on average, 1.25 out of the four tasks successfully, while the younger group was distinctly more successful ($M = 3.7$; $F(1,31) = 41.4$, $p < 0.01$). Age-differences were also revealed in efficiency measures. Younger subjects needed, on average, only 4.3 min. to process the four tasks whereas older subjects needed more than twice as much time ($M = 10$ min; $F(1,31) = 56.5$, $p < 0.01$). Looking at navigation behaviour, the same pattern was found. Older users ($M = 78$) made 58% more detour steps than younger adults ($M = 32.8$), yielding a statistical significant difference ($F(1,31) = 11.7$, $p < 0.05$). Given that overall 23 key strokes or menu steps (8 steps in the postponement tasks and 15 steps in the entry tasks) had to accomplished to solve the tasks on the most efficient way, the number of additional steps (detour steps) is considerably high in both groups. As expected, older adults made significantly more detour steps than younger adults. However, it is remarkable that even the younger adults made on average 32 detour steps, indicating that the PDA tasks were not easy to solve. The same pattern of results was revealed for the number of revisited nodes. Again, older users ($M = 37.4$) opened the nodes more frequently (38%) than younger users ($M = 23.1$; $F(1,31) = 7.6$, $p < 0.01$). Neither gender differences nor interactions between age and gender were found.

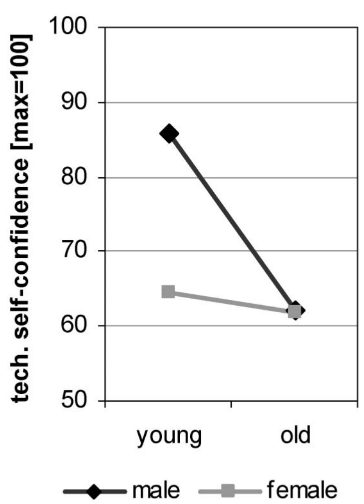

Figure 7. Interaction of age and gender regarding the subjective technical confidence.

To analyse the differential effects of user characteristics on performance, spatial ability, verbal memory and technical self-confidence, correlations between user characteristics and measures of performance were calculated (Table 3). Apparently, spatial ability and – to a slightly lesser extent – verbal memory were strongly related to performance outcomes, showing significant correlations up to $r = 0.8$. The association between technical self-confidence and performance measures was less pronounced in comparison to cognitive abilities, though reaching significant correlation coefficients of $r = 0.3$ to $r = 0.6$. Computer-expertise was not interrelated to performance, neither to effectiveness nor to efficiency.

Up to this point, results have clearly shown that age as well as user characteristics showed strong associations with navigation performance. However, cognitive abilities decrease due to aging processes. Thus it still remains unclear whether age or cognitive abilities mainly contribute to performance variability. To disentangle the interactive effects of age, cognitive ability and self-confidence, performance outcomes were reanalysed using the user characteristics as a

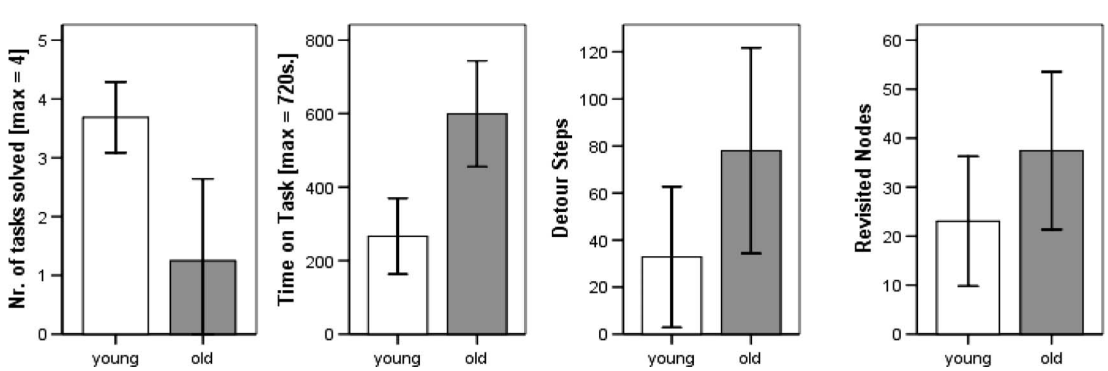

Figure 8. Age differences in effectiveness and efficiency. (left: number of tasks solved; centre left: time on task; centre right: number of detour steps, right: number of nodes revisited).

covariate. Based on the assumption that task completion mainly requires the ability to process spatial information, it follows that the lower spatial ability in the older group explains the group's lower performance. To examine how the measure of spatial ability predicts performance, an analysis of covariance (ANCOVA) was conducted using 'spatial ability' as a covariate. The ANCOVA revealed a significant main effect for 'spatial ability' ($F(4,26) = 3.3, p < 0.05$), but the main effect of age was no longer significant after the effects of spatial ability were controlled ($F(2,26) = 2.4, p < 0.1$). Multiple regression analyses confirmed this finding: spatial abilities were a stronger predictor of performance ($\beta = 0.5, p < 0.01$) than chronological age ($\beta = -0.4, p < 0.05$). This suggests that performance differences are predominately caused by differences in spatial ability and not by age per se. For further analyses two groups with low and high spatial abilities were built by median-splitting the spatial-ability scores of the two age groups, respectively. The cut-off points between low- and the high-spatial abilities group were 15 out of 20 points for the younger adults and 9 points for older adults. Figure 9 visualises the outcomes for users with high and low spatial abilities (for the two most sensitive measures, task effectiveness and time on task). With respect to verbal memory the ANCOVA revealed a significant effect for both, age and verbal memory ($F(4,26) = 7.3, p < 0.01$) and $F(4,26) = 4.7, p < 0.01$). Apparently both factors contribute to performance to a comparable extent. Finally, regarding technical self-confidence, age as a main source remained significant ($F(4,26) = 9.1, p < 0.01$), suggesting that differences in performance were caused rather by age and not by different levels of technical self-confidence.

Table 3. Correlations between user characteristics and performance ($n = 32$).

	Tasks solved	Time on task	Detour steps	Revisited nodes
Spatial ability	0.83**	−0.81**	−0.52**	−0.43*
Verbal memory	0.71**	−0.64**	−0.53**	−0.31*
STC	0.51**	−0.55**	−0.31*	−0.26
Expertise	0.26	−0.32	−0.24	−0.13

*$p < 0.05$, **$p < 0.01$.
STC, subjective technical confidence.

3.3. *Which of the user characteristics has the largest predictive power?*

As was shown, performance of both older and younger users was distinctly affected by cognitive abilities. The next question deals with the relative importance of spatial and verbal memory, respectively. In order to obtain insight into the specific relation and contribution of both abilities to navigation it was analysed, if performance is mainly predictable by spatial ability, verbal memory or by age itself. In multivariate blockwise regression analyses, the predictors age, spatial ability, and verbal memory were included and effects were observed for effectiveness (tasks solved) and processing time, respectively. In order to rule out validity and generalisation problems due to multicollinearity of predictors, a collinearity diagnosis was carried out for both regression models. In both cases the VIF (Variance Inflation Factor, values > 10 show multicollinearity) indicated, that multicollinearity could be ruled out and both regression models were acceptable (Table 4). It was found that spatial ability and age explained 73.2% of the variance of

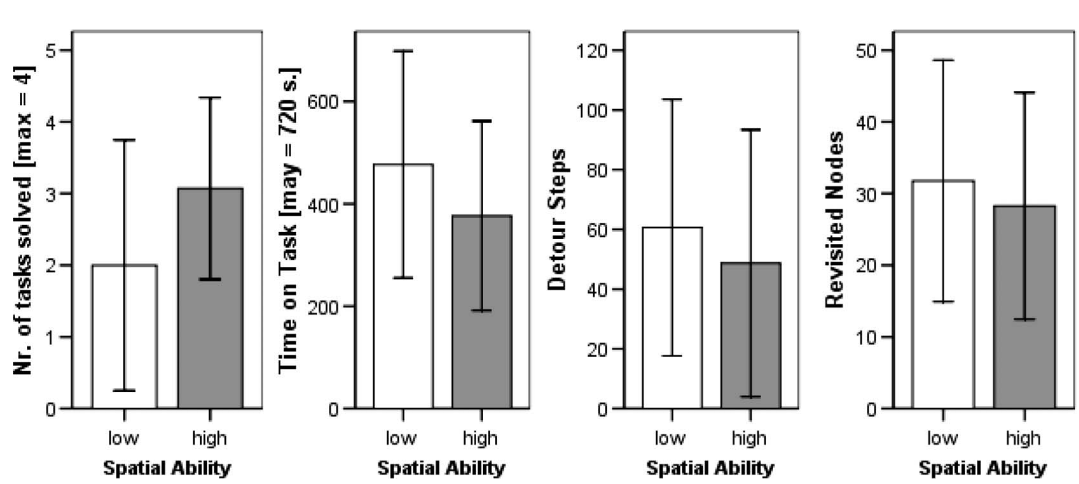

Figure 9. Performance outcomes for users with high and low spatial ability. (left: number of tasks solved; centre left: time on task; centre right: number of detour steps, right: number of nodes revisited).

effectiveness and 74.4% of the variance for processing time. Verbal memory proved to be comparably unimportant and was excluded as predictor in both stepwise regression analyses. Thus, high spatial ability – in contrast to verbal memory – accompanied successful PDA navigation performance.

3.4. *Users' mental model and performance*

Finally, the question if user's mental model of the PDA information structure had an impact on performance outcomes was examined. After participants completed the tasks on the PDA, they were asked how the menu of the digital diary was basically structured. In order to support participants understanding of the question, they were offered five sample 'models' (Figure 3) and requested to choose the most adequate model. First, the frequencies of the chosen models will be reported for each age group (see Table 5), second the chosen type of mental model is related to navigational performance.

More than one third of the older users (37.5%) selected an incorrect mental model (Pyramid, Chain) and 25% did not build a mental representation at all (having no idea). Even if the younger group's representation was much better, it is remarkable; that 18.8% of the younger group chose an inappropriate model and even 6.3% did not chose a representation at all.

As the menu of the PDA's digital diary had a relational structure, i.e. a mixture of hierarchy and network, the models 'Hierarchy,' 'Network,' and 'Hierarchy/Network mixture' were categorised as adequate models, the models 'Onion,' 'Pyramid,' and 'Chain' were classified as inadequate. Nonparametric tests for unrelated samples (χ^2-test) showed that young subjects (75%) rather than older subjects (37.5%) were able to name an adequate mental model ($\chi^2 = 4.6$, $p < 0.05$). No relationship was found between computer-expertise and the existence of a mental model.

In order to find out whether the appropriateness of the mental representation is correlated with the menu performance on the one hand, and to analyse whether the appropriateness of the mental representation is related to users' age and the level of spatial abilities on the other hand, a MANOVA was run. As independent variables, age (young vs. old) and type of mental model ((1) inadequate model, (2) hierarchy, (3) network, (4) mix of hierarchy and network) were included and their effects on performance variables (tasks solved, time, detour steps, revisited nodes) was determined. The results revealed that both, age and the mental representation had an impact on navigation performance: A significant effect of age ($F(4,14) = 9.1$, $p < 0.05$) was found as well as a significant interaction of age and the type of users' mental model ($F(16,7) = 2.0$, $p < 0.05$). As can be seen in Figure 10 the different models yielded a different impact on performance.

From an ergonomic point of view, the interaction is of special importance, as it demonstrates that older users were able to reach the performance level of younger users when they had developed an appropriate mental model of the PDA structure. While younger users performed quite well independent of the type of their mental representation, the older group especially benefited from having a network model of the PDA-menu-structure. In contrast, older users where extremely disadvantaged when they reported an inadequate model (pyramid or chain model). Unexpectedly, the hierarchy type distinctly hampered performance, even to a larger extent than having an inadequate model.

Table 4. Results of the multiple blockwise regression analyses.*

Dependent variable	Predictor	Adj. R^2	Stand. β	T	p	VIF	ANOVA
Tasks solved (TS)	Spatial ability	73.2%	0.4	2.4	$p < 0.05$	2.8	$F(3,28) = 29.3$, $p < 0.001$
	Age		−0.3	−2.2	$p < 0.05$		
Time on Task	Spatial ability	74.4%	−0.4	−2.2	$p < 0.05$	2.7	$F(3,28) = 29.7$, $p < 0.001$
	Age		0.5	3.4	$p < 0.01$		

*Adj. R^2 gives the amount of explained variance (TS, tasks solved); VIF (variance inflation factor) is a parameter which indicates multicollinearity, and the ANOVA of the regression model.

Table 5. Relative frequencies of users' mental models of the PDA menu structure.*

	No idea	Onion	Pyramid	Chain	Hierarchy	Network	Hierarchy/network
Young	6.3	0	6.3	12.5	37.5	18.8	18.8
Old	25	0	12.5	25	25	6.3	6.3

*Given are percentage values.

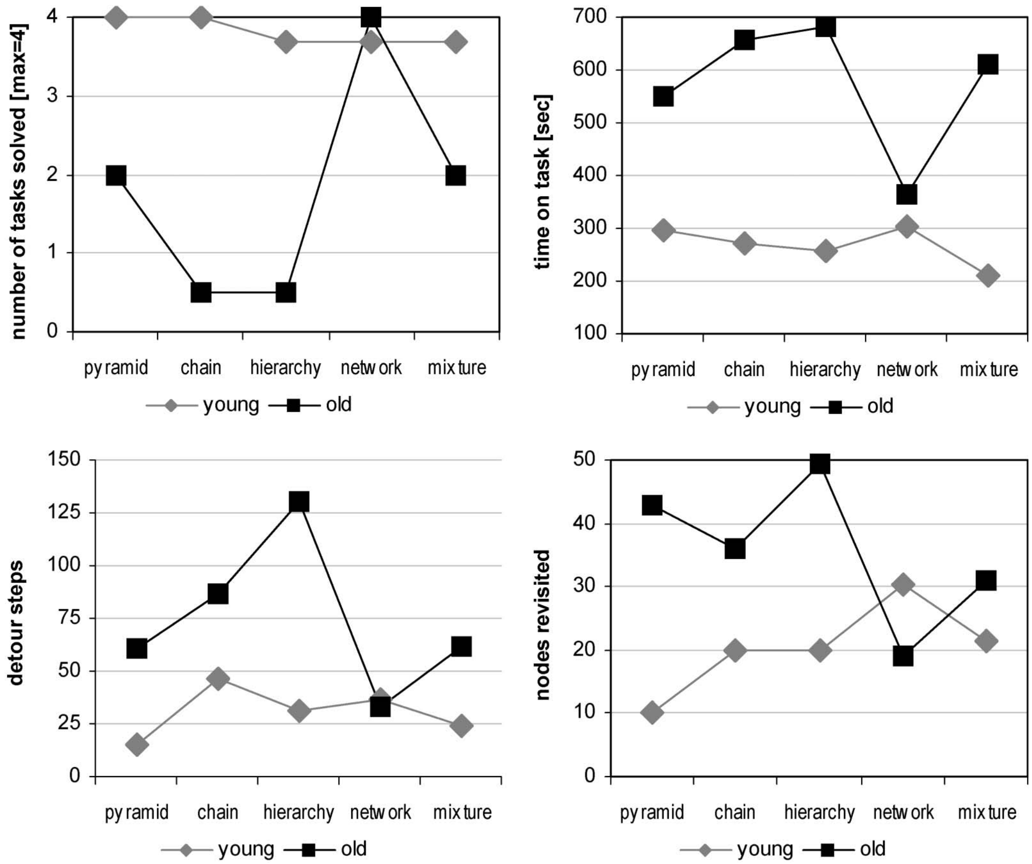

Figure 10. Type of mental model and navigation performance. (top left: number of tasks solved; top right: time on task; bottom left: number of detour steps; bottom right: number of nodes revisited).

To find out if spatial abilities affected the development of mental models, spatial ability scores and the adequacy of users' mental model were correlated. High spatial ability and the presence of an adequate mental model were associated ($\eta = 0.44$). This corroborates the finding that the ability to successfully process spatial information is related to the formation of an appropriate model of the menu structure. A $2 \times 2 \times 4$-MANOVA with the factors age (younger and older users), age-adjusted spatial ability after a median-split (high versus low), and the type of mental model was conducted to examine the effects of user characteristics and mental models on performance. As expected, a main effect 'age' was found ($F(4,11) = 7.5$, $p < 0.05$). Confirming again the impact of spatial abilities on navigation performance, the threefold interaction of age and spatial ability and the presence of having an adequate mental model was statistically significant ($F(4,11) = 5.9$, $p < 0.05$).

4. Discussion and conclusion

The present study was conducted to provide a deeper understanding of PDA menu navigation performance in younger and older users. A total of $n = 32$ younger and older adults accomplished four common tasks with digital diary of a PDA. In order to analyse individual factors that may differentially affect user's performance, spatial ability, verbal memory, technology-related self-confidence, users' computer-expertise, and the quality of their mental model of the PDA-menu-structure were surveyed and related to performance outcomes.

4.1. *Impact of user characteristics on performance*

The study confirmed the strong impact of user characteristics on PDA menu navigation performance. A first influential factor found in the analyses was users' age. However, with respect to the nature and the underlying reasons for the performance deterioration, the consideration of age alone does not have much explanatory power. Age is only a 'carrier' variable that involves many variables which change over the life span. In order to clarify the question which cognitive and/or personal variables play the major role in menu navigation performance, spatial abilities were found to be the best predictor. Spatial abilities had the most significant impact on PDA menu navigation performance. But how do spatial abilities actually facilitate performance? One objection might be that the performance superiority of users with high spatial abilities is caused by more intelligent navigation behaviour (e.g. the ability to reason about the structure and features of the system). However, this argument can be rejected, because spatial ability and general reasoning ability are not related with respect to menu navigation performance (Ziefle and Bay 2004): Students with high spatial abilities showed a distinctly better performance than those with low spatial abilities, but both groups did not differ in their reasoning abilities. Hence, it has to be assumed that spatial abilities facilitate the development of an appropriate mental representation of the system structure that, in turn, supports orientation within the menu.

Significant insights can also be derived from verbal memory, which was shown to be distinctly lower for older compared to younger users. Having high verbal memory abilities was also found to advantage performance. We assume that the influence of both cognitive abilities, i.e. verbal memory *and* spatial abilities, reflects different strategies while navigating through the menu. While users with high spatial abilities navigate through information structures based on their spatial representation of the menu structure, users with high verbal abilities refer to the memorisation of functions' names and the correct sequence of navigational steps. Although the conceptualisation of nodes as landmarks in information structures does not reach the high frequency and semantic density of landmarks in natural spaces, the utility of the landmark concept for electronic spaces structures was found in a number of studies, where landmark and route knowledge had a significant impact on menu navigation performance (e.g. Thorndyke and Goldin 1983, Goodman *et al*. 2004, Ziefle and Bay 2004, Bay and Ziefle 2008). However, future studies should be conducted in order to analyse the relationship between cognitive abilities and different navigation strategies.

Regarding personal factors, it was found that the extent of technical self-confidence was less pronounced in older adults than in the younger group. One interesting finding is that younger male users reported the highest technical confidence compared to female users of the same age, but without showing a better performance. This confirms that gender effects in subjective reports of technical self-confidence do not account for performance differences. When the effects of this personal/emotional style on performance outcomes are considered, technical self confidence was found to affect effectiveness and efficiency only to a lesser extent. This result underlines the necessity to survey performance measures in usability studies, not only to validate the outcomes of subjective ratings, but also to evaluate the predictive power of user ratings.

Contrary to other findings (e.g. Czaja and Sharit 1993, Rodger and Pendharkar 2004) computer experience was associated with neither age nor navigational performance. At first sight, this finding implicates that prior computer-expertise does not account for PDA performance differences between the age groups. However, in this context two methodological concerns regarding the validity of subjective ratings have to be considered. First, subjective ratings are often biased by social desirability, fear of failure and misinterpretation of ones own performance. Second, the items used to measure computer experience in this study (frequency and length of computer usage) might lack content validity. These items refer to quantitative aspects of computer usage and do not necessarily reflect qualitative aspects of computer-expertise, i.e. computer-related knowledge. An experienced user is familiar with basic and advanced concepts and the related actions, which are conventionally used in a computer system. An example for such a basic concept is that a computer document can be 'closed' by clicking on an icon (a cross in a little square), normally located in the upper right corner, before it disappears from the screen. It is thus conceivable that the ratings regarding temporal aspects (length, frequency) of computer usage do not, at least not necessarily, reflect the aptitude level regarding the knowledge of computer-related concepts. Therefore – instead of using ratings of temporal aspects of computer usage – the knowledge of task-relevant concepts should be assessed. A test to measure task-relevant theoretical and practical knowledge about computers was developed, validated and compared to the quality of ratings of computer experience in recent studies (Arning and Ziefle 2008).

The analysis of the individual representation of the PDA menu structure – users' mental model – was an effective approach to uncover the mental representation, which was guiding users on their individual menu navigation paths. A first insight can be retrieved by considering the type of model that was chosen by participants. None of the users chose the onion-shaped model, the only structure without connections between the concentrically arranged levels. Apparently, the possibility of wandering around within the menu is not compatible with this model. Interestingly, even though the PDA is also hierarchically structured, the hierarchy model turned out to be counterproductive, especially for older users. It provoked many detour steps, accompanied by longer processing times and lower task effectiveness. However, this finding clearly demonstrates the effect of a mental model on navigation behaviour. Whenever users had developed a hierarchical representation of the PDA information structure, the execution of detour steps and going back in the menu when re-orientating is quite 'logic', as the reorientation in hierarchical systems is possible by going back, automatically returning to a higher menu level. Given that mobile phones and computer hard discs are exclusively hierarchically structured and that these structures are well known by many users, the ergonomic demands with respect to training matters become apparent. Users who are familiar with hierarchical structures have to face enormous switching costs when they simply transfer their understanding from one technical device to another. Users with a network type representation of the PDA-menu were distinctly more effective and efficient than users that had an inappropriate model or even no mental model at all. The significance of mental models on navigation performance acknowledges the importance of spatial abilities once more (Gilbert *et al.* 2004, Ziefle and Bay 2006). Strikingly, older adults were guided more frequently by an inappropriate model or even no representation at all in their navigation. Even if this occurred in the younger group, too, it did not have a comparable negative impact on performance, presumably, because younger adults are able to compensate the lack of a adequate model with their higher cognitive abilities or computer experience. Older users with a correct mental model were strongly advantaged; they even reached the performance level of the younger users, thus overcoming their 'age handicap'. This outcome is of ergonomic significance as it shows that technical devices can be used by all users rather successfully if they are supported adequately (Gray 1990, Gischer 1991, Gilbert *et al.* 2004).

4.2. Potential applications of the study

The findings show that certain user characteristics advantage (or disadvantage) menu navigation performance. The knowledge about which factors contribute to the ageing-disadvantage in technical device usage is helpful regarding future research, especially for interface and instructional design. The results clearly showed that a correct mental model was associated with a high task performance. However, the conclusion that users will only have to be 'informed' about the network-like data structure is an oversimplification. An appropriate mental model is not only surface knowledge, which can be simply instructed, but it already represents a cognitive product of deeper understanding. Therefore, the impact of different training types and instructional formats, that may benefit performance, should be examined in future studies. A first type of training should address the knowledge transfer of an adequate mental model by showing the participants a spatial map of the system (Dee-Lucas and Larkin 1995, McDonald and Stevenson 1998, Ziefle and Bay 2006, Bay and Ziefle 2008). This instruction type would impart conceptual knowledge about the system's spatial structure (menu breadth and width) and the navigation paths to special functions. We expect this type of instruction to be especially effective for users with low spatial abilities, because it provides – like a street map – information about the size of the system, branching points, and possible ways to go, and therefore reduces navigational disorientation. Second, although the prior computer-expertise was not a predictor of menu navigation performance in this study, the instruction of basic knowledge principles and concepts of PDA usage (the meaning of icons, the concept of opening and closing documents, the concept of saving, etc.) should be examined as well. Older adults might require explicitly formulated concepts and rules, because they are less experienced in the interaction with computers. Furthermore, it was shown that older adults have greater difficulties in reasoning and abstraction abilities (Salthouse 1992). Therefore it is more difficult for them to derive concepts and principles while interacting with a device. Instructions which contain information about these concepts and principles might support older adults in handling the device. A third type of training refers to the instruction of sequential route knowledge (Thorndike and Goldin 1983). Accordingly, participants will be informed about the sequence of single steps to reach and execute a certain function in the menu. This training type, which is frequently used in conventional instruction manuals, does not provide structural knowledge, and is therefore not expected to have a sustained effect on performance. A final type of training refers to participants commentaries in the experiment. They stated that they usually learn to handle an unknown device by observing someone else demonstrating the usage. A training, which includes the observation of an expert acting with the device, corresponds to model learning. Future studies will have to specify which instructional format is the most helpful and address carefully the cognitive antecedents necessary to build up a correct mental model.

4.3. Limitations of the study

Some final remarks deal with potential limitations of the study.

(1) One objection is that the performance outcomes, especially of the older users, might not be representative for the older population. As we pursued a benchmark testing procedure, a very homogenous "older" group was selected.

The older participants in this study were comparably young, healthy, had good visual abilities and were experienced in the usage of the mouse as common input device. Furthermore, they were quite experienced in technology and had a higher education. Compared to more heterogeneous samples of older adults the outcomes at hand represent an underestimation of menu disorientation problems in PDA usage. Moreover we assume that a more heterogeneous user group might show different interaction styles and learning characteristics. For example inexperienced older might show a cautious interaction style, because they want to avoid damaging the technical system. Therefore a differential approach, which accounts for individual differences among older users, should be pursued, in order to learn which support is appropriate for inexperienced older users.

(2) Another concern is related to problems of oversimplification and a limited external validity, due to the employment of a PDA-emulation instead of a real device. The usage of a simulated device is certainly less demanding than the usage of a real device regarding cognitive and motor load (e.g. holding the device with one hand and pointing with a stylus with the other hand, or encoding of objects on a reflecting glass surface). Therefore, when compared to the demands when using a real device, the findings – again – might be an underestimation.

(3) Future studies should therefore examine other PDA-applications and tasks beyond the digital diary application. This is of considerable impact as future devices will possess a higher complexity in both number and variety of functions. Apart from this, the present study provided helpful insights into the underlying factors causing performance differences between older and younger adults in menu navigation performance. Based on the significant importance of mental models, which was proven in this study, different instructional formats regarding the usage of mobile devices will be examined, in order to facilitate the usage of mobile devices for older adults.

(4) A final cautionary note is concerned with the way we measured the adequateness of the mental representation of the information structure adopted. The assumption that users form a mental representation while interacting with a specific device type is, of course, arbitrary. In addition, the uncovering of the cognitive model, which may be responsible for the specific interaction, is methodologically highly sophisticated, because users are not necessarily aware that they have a specific representation and, therefore, mental models cannot be simply reported.

Acknowledgements

The authors thank the older participants for courageously taking part in the study, sometimes facing frustrating feelings in the interaction with the technical device. Moreover, thanks are due to Dominique Ziegelmayer, Lutz Böhnstedt, André Calero Valdez and Mahir Kilic who programmed the software to display the PDA on computer screens and the logging tool for recording of user actions. Thanks are also due to Judith Strenk and Sarah Hatfield for their research support.

References

Arning, K. and Ziefle, M., 2006. What older users expect from mobile devices: an empirical survey. *In*: R.N. Pikaar, E.A. Konigsveld, and P.J. Settels, eds. *Meeting Diversity in Ergonomics. Proceedings of the 16th World Congress on Ergonomics*. [CD-ROM]. Amsterdam: Elsevier.

Arning, K. and Ziefle, M., 2007a. Barriers of information access in small screen device applications: the relevance of user characteristics for a transgenerational design. *In*: C. Stephanidis and M. Pieper, eds. *User Interfaces for All – Universal Access in Ambient Intelligence Environments*. Berlin: Springer, 117–136. Lecture Notes in Computer Science 4397.

Arning, K. and Ziefle, M., 2007b. Understanding age differences in PDA acceptance and performance. *Computers in Human Behavior*, 23 (6), 2904–2927.

Arning, K. and Ziefle, M., 2008. Development and validation of a computer expertise questionnaire for older adults. *Behaviour & Information Technology*, 27 (1), 89–93.

Bandura, A., 1986. *Social Foundations of Thought and Action: A Social Cognitive*. Englewood Cliffs, NJ: Prentice Hall.

Bay, S. and Ziefle, M., 2003. Design for all: user characteristics to be considered for the design of devices with hierarchical menu structures. *In*: H. Luczak and K.J. Zink, eds. *Human Factors in Organizational Design and Management*. Santa Monica: IEA Press, 503–508.

Bay, S. and Ziefle, M., 2008. Landmarks or surveys? The impact of different instructions on children's performance in hierarchical menu structures. *Computers in Human Behaviour*, 24 (3), 1246–1274.

Beier, G., 1999. Kontrollüberzeugungen im Umgang mit Technik [Locus of control when interacting with technology]. *Report Psychologie*, 24, 684–693.

Birren, J.E., Schaie, K.W., and Warner, K., eds. 2006. *Handbook of Psychology and Aging*. New York: Academic Press.

Brosnan, M.J., 1998. The impact of computer anxiety and self-efficacy upon performance. *Journal of Computer Assisted Learning*, 14, 223–234.

Busch, T., 1995. Gender differences in self-efficacy and attitudes toward computers. *Journal of Educational Computing Research*, 12, 147–158.

Cherry, K.E. and Park, D.C., 1993. Individual difference and contextual variables influence spatial memory in young and older adults. *Psychology and Aging*, 8, 517–525.

Chi, R., Glaser, R., and Farr, M.J., 1988. *The Nature of Expertise*. Hillsdale, NJ: Lawrence Erlbaum.

Christensen, H., MacKinnon, A., Jorm, A.F., Henderson, A.S., Scott, L.R., and Korten, A.E., 1994. Age differences and interindividual variation in cognition in community-dwelling elderly. *Psychology and Aging*, 3, 381–390.

Conklin, E.J., 1987. Hypertext: an introduction and survey. *IEEE Computer*, 20, 17–41.

Craik, F.I.M. and Salthouse, T.A., eds. 1992. *Handbook of Aging and Cognition*. Hillsdale, NJ: Lawrence Erlbaum.

Czaja, S.J. and Sharit, J., 1993. Age differences in the performance of computer-based work. *Psychology and Aging*, 8, 59–67.

Czaja, S.J., 1997. Computer technology and older adults. *In*: M.E. Helander, T.K. Landauer, and P. Prabhu, eds. *Handbook of Human-Computer Interaction*, 2nd edition. New York: Elsevier, 797–812.

Davies, S.P., 1994. Knowledge restructuring and the acquisition of programming expertise. *International Journal of Human-Computer Studies*, 40, 703–726.

Dee-Lucas, D. and Larkin, J.H., 1995. Learning from electronic texts: effects of interactive overviews for information access. *Cognition and Instruction*, 13, 431–468.

Dillon, A. and Watson, C., 1996. User analysis in HCI – the historical lessons from individual differences research. *International Journal of Human-Computer Studies*, 45, 619–637.

Dorn, B., Zelik, D., Vepadharmalingam, H., Ghosh, M., and Adams, S.K., 2004. Designing a user interface for a PDA-based campus navigation device. *In*: *Proceedings of the 48th Annual Meeting of the Human Factors and Ergonomics Society*, Santa Monica. CA, Santa Monica: Human Factors Society, 861–865.

Downing, R.W., Moore, J.L., and Brown, S.W., 2005. The effects and interaction of spatial visualization and domain expertise on information seeking. *Computers in Human Behaviour*, 21, 195–209.

Edwards, D.M. and Hardman, L., 1989. "Lost in hyperspace": cognitive mapping and navigation in a hypertext environment. *In*: R. McAleese, ed. *Hypertext: Theory into Practice*. Oxford: Intellect, 105–125.

Ekstrom, R.B., French, J.W., Harman, H.H., and Dermen, D., 1976. *Manual for the Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.

ISO. EN ISO 9241-11:1998. *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs), Part 11: Guidance on Usability*. Berlin: Beuth.

Evans, A.W., Jentsch, F., Hitt, J.M., Browers, C., and Salas, E., 2001. Mental model assessments: is there convergence among difference among different methods? *In*: *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, Santa Monica. CA, Santa Monica: Human Factors Society, 293–296.

Evans, G. and Simkin, M., 1989. What best predicts computer proficiency? *Communications of the ACM*, 32, 1322–1327.

Fisk, A.D. and Rogers, W.A., 1997. *Handbook of Human Factors and the Older Adult*. San Diego, CA: Academic Press.

Gentner, D. and Stevens, A.L., eds. 1983. *Mental Models*. Hillsdale, NJ: Lawrence Erlbaum.

Gilbert, D., Rogers, W., and Samuelson, M., 2004. Long-term retention of a spatial mental model for younger and older adults. *Experimental Aging Research*, 30, 217–224.

Gischer, G., 1991. The importance of models in making complex systems comprehensible. *In*: M.J. Tauber and D. Ackermann, eds. *Mental Models and Human Computer Interaction 2*. North-Holland: Elsevier, 3–36.

Goodman, J., Gray, P., Khammampad, K., and Brewster, S., 2004. Using landmarks to support older people in navigation. *In*: S. Brewster and M. Dunlop, eds. *MobileHCI 2004*. Lecture Notes in Computer Science 3160. Berlin: Springer, 38–48.

Gray, S., 1990. Using protocol analyses and drawings to study mental model construction during hypertext navigation. *International Journal of Human Computer Interaction*, 2, 359–377.

Haegerstrom-Portnoy, G., Schneck, M.E., and Brabyn, J.A., 1999. Seeing into old age: vision function beyond acuity. *Optometry and Vision Science*, 76, 141–158.

Hegarty, M., Kriz, S., and Cate, C., 2003. The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction*, 21, 325–360.

Kim, H. and Hirtle, S.C., 1995. Spatial metaphors and disorientation in hypertext browsing. *Behaviour & Information Technology*, 14, 239–250.

Larson, K. and Czerwinski, M., 1998. Web page design: implications of memory, structure and scent for information retrieval. *In*: *Proceedings of CHI 98*, New York. New York: ACM, 25–32.

Leonhardt, S., 2006. Personal health care devices. *In*: S. Mukherjee, E. Aarts, R. Roovers, and F. Widdershoven, eds. *AmIware: Hardware Technology Drivers of Ambient Intelligence*. Dordrecht: Springer, 349–370.

Lin, D.M., 2001. Age differences in the performance of hypertext perusal. *In*: *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, Santa Monica. CA, Santa Monica: Human Factors Society, 211–215.

Liu, L. and Grandon, E.E., 2003. How performance and self-efficacy influence the ease of use of object-orientation: the moderating effect of prior training. *In*: *Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS'03)*, January 2003, Big lsland, HI, 327–336.

Mayer, R.E., 1997. From novice to expert. *In*: M.G. Helander, T.K. Landauer, and P. Prabhu, eds. *Handbook of Human Computer Interaction*. Amsterdam: Elsevier, 781–795.

McDonald, S. and Stevenson, R., 1998. Effects of text structure and prior knowledge of the learner on navigation in hypertext. *Human Factors*, 40, 18–27.

Mitchell, T.J.F., Chen, S.Y., and MacRedle, R.D., 2005. Hypermedia learning and prior knowledge: domain expertise vs. system expertise. *Journal of Computer Assisted Learning*, 21, 53–64.

Morrell, R.W., Park, D.C., Mayhorn, C.B., and Kelley, C.L., 2000. Effects of age and instruction on teaching older adults to use Eldercomm, an electronic bulletin board system. *Educational Gerontology*, 26, 221–235.

Morrow, D., Miller, L.S., Ridolfo, H., Kokayeff, N., Chang, D., Fischer, U., and Stine-Morrow, E., 2004. Expertise and aging in a pilot decision-making task. *In*: *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, Santa Monica. CA, Santa Monica: Human Factors Society, 228–232.

Noyes, J.M. and Sheard, M.C.A., 2003. Designing for older adults – are they a special group? *In*: Stephanidis C., ed. *Universal Access in HCI: Inclusive Design in the Information Society*. Mahwah, NJ: Lawrence Erlbaum, 877–881.

Omori, M., Watanabe, T., Takai, J., Takada, H., and Miyao, M., 2002. Visibility and characteristics of the mobile phones for elderly people. *Behaviour & Information Technology*, 21, 313–316.

Oulasvirta, A., 2004. Task-processing demands and memory in web interaction: a levels-of-processing approach. *Interacting with Computers*, 16, 217–241.

Pak, R., 2001. A further examination of the influence of spatial abilities on computer task performance in younger and older adults. *In*: *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, Santa Monica. CA, Santa Monica: Human Factors Society, 1551–1555.

Park, D.C. and Schwarz, N., 1999. *Cognitive Aging*. Philadelphia: Buchanan.

Parush, A. and Yuviler-Gavish, T., 2004. Web navigation structures in cellular phones: the depth/breadth trade-off issue. *International Journal of Human-Computer Studies*, 60, 753–770.

Pitman, J.A. and Payne, S.J., 2006. Creating names for retrieval by self and others. *Behaviour & Information Technology*, 25, 489–496.

Rodger, J.A. and Pendharkar, P.C., 2004. A field study of the impact of gender and user's technical experience on the performance of voice-activated medical tracking application. *International Journal of Human-Computer Studies*, 60, 529–544.

Salthouse, T.A., 1992. *Mechanisms of Age-Cognition Relations in Adulthood*. Mahwah, NJ: Lawrence Erlbaum.

Sasse, M.-A., 1991. How to t(r)ap users' mental models. *In*: J. Tauber and D. Ackermann, eds. *Mental Models and Human-Computer Interaction 2*. North Holland: Elsevier, 59–79.

Schröder, S. and Ziefle, M., 2005. Semantic transparency of cellular phones menu. Comparing users from different age groups. *In*: B. Fisseni, H.-C. Schmitz, B. Schröder, and P. Wagner, eds. *Computer Studies in Language and Speech*. Frankfurt: Peter Lang, 302–315.

Sein, M.K., Olfman, L., Bostrom, R.P., and Davis, S.A., 1993. Visualization ability as a predictor of user learning success. *International Journal of Man-Machine Studies*, 39, 599–620.

Shiffler, G., Smulders, C., Correia, J.M., Hale, K., and Hahn, W.M., 2005. *Gartner dataquest market databook*. Available online at: www.gartner.com [Accessed 1 September 2006].

Siebert, H., 2001. Economic perspectives for aging societies. Available online at: www.uni-kiel.de/IfW/pub/siebert/pdf/econ.pdf (accessed August 2007). Revised version of the opening paper of the Kiel Week Conference on Economic Policy for Aging Societies, 18–19 June 2001.

Stanney, K. and Salvendy, G., 1995. Information visualization: assisting low-spatial individuals with information access tasks through the use of visual mediators. *Ergonomics*, 38, 1184–1198.

Tauscher, L. and Greenberg, S., 1997. Revisitation patterns in world wide web navigation. *In*: *Proceedings of CHI 1997*, New York. New York: ACM, 399–206.

Thorndyke, P.W. and Goldin, S.E., 1983. Spatial learning and reasoning skill. *In*: H.L. Pick and L.P. Acredolo, eds. *Spatial Orientation: Theory, Research, and Application*. New York: Plenum, 195–217.

Vercruyssen, M., 1997. Movement control and speed of behaviour. *In*: A.D. Fisk and W.A. Rogers, eds. *Handbook Human Factors and the Older Adult*. San Diego: Academic Press, 55–86.

Vicente, K.J., Hayes, B.C., and Williges, R.C., 1987. Assaying and isolating individual difference in searching a hierarchical file system. *Human Factors*, 29, 349–359.

Westerman, S.J., 1995. Individual differences in the use of command line and menu computer interfaces. *International Journal for Human Computer Interaction*, 9, 183–198.

Ye, N. and Salvendy, G., 1994. Quantitative and qualitative differences between experts and novices in chunking computer software knowledge. *International Journal of Human Computer Interaction*, 6, 105–118.

Ziefle, M. and Bay, S., 2004. Mental models of a cellular phone menu. Comparing older and younger novice users. *In*: S. Brewster and M. Dunlop, eds. *MobileHCI 2004*. Lecture Notes in Computer Science 3160. Berlin: Springer, 25–37.

Ziefle, M. and Bay, S., 2005. The complexity of navigation keys in cellular phones and their effects on performance of teenagers and older adult phone users. *In*: *Proceedings of the HCI International 2005, Vol. 4: Theories, Models and Processes in Human Computer Interaction*. [CD-ROM]. St. Louis, MO: Mira Digital.

Ziefle, M. and Bay, S., 2006. How to overcome disorientation in mobile phone menus: a comparison of two different types of navigation aids. *Human Computer Interaction*, 21, 393–432.

Ziefle, M., 2002. The influence of user expertise and phone complexity on performance, ease of use and learnability of different mobile phones. *Behaviour & Information Technology*, 21, 303–311.

Ziefle, M., Michel, T., Strenk, J., and Schroeder, U., 2007. How young and older users master the use of hyperlinks in small screen devices. *In*: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2007*, New York. New York: ACM, 307–316.