# Barriers of Information Access in Small Screen Device Applications: The Relevance of User Characteristics for a Transgenerational Design

Katrin Arning and Martina Ziefle

Department of Psychologie, RWTH Aachen University,
Jägerstr. 17-19, 52066 Aachen, Germany
{Katrin.Arning, Martina.Ziefle}@psych.rwth-aachen.de

**Abstract.** The proliferation of ubiquitous computing is accompanied by the development of devices, which promise to facilitate the daily living activities of people. However, the question if mobile devices address the usability demands of older users is still unsettled. This paper reports the findings of a series of studies that examined the performance of younger and older adults when using personal data management software applications of a PDA. In order to learn about the ageing impact, the influence of user characteristics like spatial and memory abilities, the subjective technical confidence and computer-expertise on performance outcomes were analysed. Beyond quantitative performance measurements, the major shortcomings in interface design were identified. Results showed that older users reached lower performance outcomes. Even young adults did not reach a perfect performance level, hinting again at shortcomings in the design of PDA applications. Overall, the findings demonstrate the need to include user characteristics in interface design.

**Keywords:** experimental, older adults, spatial ability, expertise, menu navigation, personal data management applications, PDA.

## 1 Introduction

Imagine the following scenario [1]:
*An old couple, Old Bird and Gramps wants to visit relatives on the canary island La Palma. Unfortunately, traveling is no „walk in the park" for them any more, because of sight and hearing losses and problems to find their way in unfamiliar surroundings. Coincidentally, they come across a new offer of the travel agency „Charly's Magic Travel", which particularly addresses the travel demands and needs of older or disabledpeople. Charly Magic Travel offers informational support while traveling about timetables, traveling routes, means of travel, etc. via mobile devices. On the day the journey starts, Old Birds PDA lets the couple know, when they have to leave their home. At the bus station it informs the seniors in how many minutes the bus will arrive and at which platform they will have to wait. The PDA also gives notice about the length of the journey and the arrival time, so that Old Bird and Grampy have enough time to prepare for leaving the bus. During their journey the PDA offers information about restaurants, accommodations and attractive sights. (…) The device is*

C. Stephanidis and M. Pieper (Eds.): ERCIM UI4ALL Ws 2006, LNCS 4397, pp. 117 – 136, 2007.

*context-aware, adaptive and able to save the individual user settings. Based on this user profile the PDA offers Old Bird and Grampy convenient services, e.g. it informs Grampy at the next day around noon, that an excellent café close to their hotel offers the best café latte of the whole island. Charly's Magic Travel concept and Old Birds PDA allow the old couple to travel independently and to take care for themselves.*

This scenario does not describe visionary dreams of the future, but gives an example how intelligent computational systems may accompany our lives. The rapid proliferation of ubiquitous computing and continuously accessible information and services refers to the development of a new computing paradigm known as ambient intelligence (AI). AI-systems will provide embedded, sometimes even invisible computer applications, which are supposed to support and facilitate the daily living activities of people. As target population of ambient intelligence not only young people and those interested in technology are addressed. Ambient computing aims at a broader and more heterogeneous group of users. Especially for older adults, who represent the fastest growing group in western societies, pervasive and adaptive devices can offer an enormous benefit in various areas of application. Ambient computing does not only provide information similar to the scenario mentioned above, it also enables the ageing population to have an independent and unrestricted life. A typical field of future standard applications is the "intelligent house", in which technical facilities (e.g. heating) and service features (drawing the curtains, shopping of food, etc) are operated by mobile computers like PDAs. In the aforementioned example "Old Bird" and "Granny" also use a PDA in order to retrieve relevant information for their journey. Further applications, which are currently developed especially for older adults are for example medical monitoring, wayfinding aids for blind or deaf people, communication or memory aids, as well as personal data management applications (e.g. a diary with a reminder for medical appointments or birthdays).

The utilisation of ambient computing devices is supposed to happen "along the way", i.e. the user does not have to pay much attention to it. In a typical user scenario, users are expected to integrate the utilisation of computer devices into their daily life, while being engaged with other activities. However, the benefit depends on the ease of use while interacting with those devices. Therefore, beyond feasibility demands of technology and mobile services, one of the most important aspects in the development of ambient computing devices is the ergonomic aspect.

Ambient computing devices should be developed in a way that users want to use it, not only "for technologies sake" [2]. However, it is important to take a step forward: technology has to be created in a way that users want to use it – and that they are able to use it. As long as interface designs are not easy to use and learn, the technical innovations will not have sustained success. Thus, users' needs and cognitive demands have to be carefully studied and considered. Even though the interaction of users with desktop computers is well studied, little is known about the ability of users to interact with small screen devices, like mobile phones, pagers, communicators and PDAs. The mobile character of these devices represents a higher cognitive demand compared to conventional computer work. Many different problem areas might hamper the successful interaction with these devices: The first problem refers to the utilisation in a multiple task environment. Ambient computing devices will be used in a context, where several actions can be executed simultaneously, e.g. retrieving information from the PDA while watching out for traffic. Second, problems of understanding can arise, because the interaction with a device requires a understanding of technology, which is not present in every user. Computers in the AI-paradigm do not resemble the conventional desktop computers any more, they will become physically and mentally invisible, e.g. when they are embedded in furniture or walls [3]. Third, a major problem is concerned with the interaction with menu-based software applications. As the complexity of functions implemented in devices is increasing, it is important to ensure that users are able to accomplish the navigation in a complex menu. Due to a limited screen size, the majority of menu functions is not visible. Thus, users are urged to memorize, a) which functions are implemented and b) at which location in the menu [4, 5], which imposes considerable cognitive load. The described problems do not affect all users in the same way, but especially aged users or those with little computer-related experience are disadvantaged. Therefore, it is necessary to obtain knowledge about specific usability demands of those user groups.

## 2 The Older User and Technology

For at least two reasons it is necessary to focus on the specific characteristics and requirements of older users when designing ambient computing systems. The first one is the rapidly growing number of senior citizens in western societies. The second reason comes from the increasing proliferation of technology into professional and private areas. Applications like word processor programs, web browser, ATMs, ticket vending machines or online library catalogues are deeply integrated into daily life. Although these technologies are supposed to be accessible to everyone, a gap between those, who are "computer-literate" and those who are not (predominantly older users) is emerging. Although the knowledge about the older user and the utilisation of ambient computing devices is rather limited, the literature shows that older users usually face larger difficulties than younger adults in learning and using new computer applications [6, 7, 8]. Contrary to current stereotypes, older users are in fact interested in the acquisition of computer skills [9, 10]. Up to now, new applications are often designed without considering the demands of this user group. In order to include the potential user group of older adults, a transgenerational design is needed, which is supposed to allow users of all age groups and ability levels to interact with new technical applications. To achieve this goal, interface designers need to know about age-related changes and how they may affect the interaction with computer systems.

### 2.1 Characteristics of the Older User

Several different factors account for age-related differences in computer utilisation and performance. A first important factor is computer experience. Studies have shown that older adults possess a lower computer experience compared to younger users [11, 12]. This "experience-gap" can be explained by cultural factors and a different upbringing. Furthermore, when people have retired, they no longer keep in intensive contact with new technology for work purposes and the acquisition of new computer applications becomes less attractive. As a result, computer-related knowledge and concepts or models about how to operate technical devices is limited in the older group. This might increase the difficulty of computer-related tasks and therefore accounts for differences in computer-based performance. But the pure factor computer experience alone is not sufficient to explain age differences in computer performance [6]. A further explanation for age-related computer performance differences is the decline in cognitive, psychomotor and sensory functioning over the lifespan [13, 14, 15, 16, 17]. Age-related changes in the cognitive system lead to a decline in working-memory capacities, a slowing-down in processing speed and a reduced ability to distinguish relevant from irrelevant information [18]. As a result, older learners face greater difficulties in extracting relevant information from user manuals or they are overwhelmed with displays with a high information density. A reduced working memory capacity becomes critical when task demands are high, e.g. when using novel or complex technical devices. The reduced processing speed is often considered as an additional decline, but it also serves as an explanation for the age-related reduction in other cognitive abilities [19]. This "slowing-down" results in longer execution times of single steps, especially in complex tasks [16, 20]. Another profound decline over the life span concerns spatial abilities [21, 22, 23] as well as spatial memory [24, 25]. Older users with reduced spatial abilities experience disorientation and the feeling of "getting lost" while navigating through the menu of software applications [26, 27, 28]. Regarding changes in sensory functioning, the auditory and visual abilities of older adults decline. Changes in the auditory system comprise an increase in the auditory threshold (especially for high-frequency tones), difficulties in the perception of speech and in the discrimination and the spatial location of tones [29]. Apart from an ongoing decrease in visual acuity, the eyes' ability to accommodate and adapt to light changes declines and the sensitivity to glare and reflections increases [30,31]. Finally, the ability to execute fast and accurate movements is also declining over the life span [32].

In the present experimental study, the focus was laid on the interaction of older users with personal data management software of a PDA. In order to learn about the age-specific impact, a group of young adults was examined as well. Thus, we wanted to find out if "typical" users of different age groups are able to successfully use the personal data management software. In order to find out, which specific demands have to be considered for a transgenerational design, users were surveyed with respect to their age, spatial ability, verbal memory, the technical self-confidence and the extent of computer experience. Furthermore, user ratings regarding usability problems while interacting with the device were assessed. The outcomes are not only important for the identification of age-related factors that should be considered in interface design, but also with respect to their impact on performance when working with software applications of a PDA-device. In general, the study aimed at two objectives:

1. The identification of user characteristics, which underlie the age-related differences and should be considered in a transgenerational design approach.
2. The identification of typical "barriers" in the interface design.

# 3 Method

## 3.1 Variables

As independent variable, user age was examined, contrasting the performance of older and younger participants. As dependent measures the effectiveness and efficiency were analysed according to the standard measures for usability [33]. For task effectiveness, the percentage of successfully solved tasks was summed up. As efficiency measure the time needed to process the tasks was measured. As user characteristics, spatial ability, verbal memory, technical self-confidence (STC) and the reported experience with computers were surveyed. Additionally, the perceived ease of use concerning the interaction with the PDA and the usefulness as well as usability problems were surveyed.

### 3.2 Participants

A total of 96 participants took part in the study. As three typical PDA applications were experimentally examined, three samples of 32 users were selected. They were balanced by age (16 adults per age group) as well as by gender (16 females and 16 men per group). The younger group had a mean age of 23.2 years (s = 2.4), the older group of 58.2 years (s = 6.0). There were no age differences in the groups of the three samples (see Table 1). A benchmark procedure was pursued regarding the recruitment of older adults: we aimed at the "younger and healthy older adults". All of them were active parts of the work force, mentally fit and not hampered by stronger age-related sensory and psychomotor limitations. The younger group consisted of students of different academic fields (engineering and social sciences).. Within their age group the three samples did not differ in their educational background (see Table 1)[1]. Contrasting the educational background of younger and older adults across the three samples, the younger adults were found to possess a significantly higher educational degree than older adults ($M_{young} = 4.1$ $M_{old} = 3.2$, $t = 3.5$, $p < 0.05$).

Younger participants were recruited at the university and fulfilled a course requirement; older participants were reached by an advertisement in a local newspaper. The older adults came from different professions (engineers, administrative officers, secretaries, (high school) teachers, nurses, architects, physiotherapists, physicians and psychiatrists). Ruling out further confounding variables, a careful screening of

**Table 1.** Mean and standard deviations of age and educational background for both age groups in the three experimental samples

| | | Diary | | To-Do-List | | Directory | | |
|---|---|---|---|---|---|---|---|---|
| | | M | SD | M | SD | M | SD | p |
| Age (in years) | Young adults | 23.8 | 2.8 | 23.4 | 2.0 | 22.2 | 2.0 | 0.13 |
| | Older adults | 56.4 | 6.8 | 59.9 | 5.9 | 58.2 | 5.0 | 0.26 |
| Educa-tional level | Young adults | 4.1 | 0.4 | 4.0 | 0.9 | 4.1 | 0.5 | 0.86 |
| | Older adults | 3.0 | 1.5 | 3.3 | 1.8 | 3.3 | 1.4 | 0.81 |

[1] The educational background was assessed on the following scale: 1 = ger. "Haupt/Volksschulabschluß" (6-8 years of school), 2 = ger. "Realschulabschluß" (10 years of school), 3 = ger. "Fachhochschulreife" (12 years of school), 4 = ger. "Abitur" (13 years of school), 5 = ger. "Hochschulabschluß" (academic degree).

participants' abilities was undertaken. All participants were in good physical and mental health conditions. Visual acuity was normal or corrected to normal and no history of eye-illnesses was reported.

### 3.3 Experimental Tasks

The experimental tasks simulated standard software applications, implemented in commercially available PDAs: the digital diary of a PDA, the to-do-list and, as a third application, the digital directory. Participants had to accomplish four prototypic PDA tasks, where they had to create a new entry or change an existing one.

A flowchart of the task procedures for the two task types ("create a new entry" and "change an existing entry") can be seen in Figure 1 on the following page. Participants had a time limit of five minutes per task.

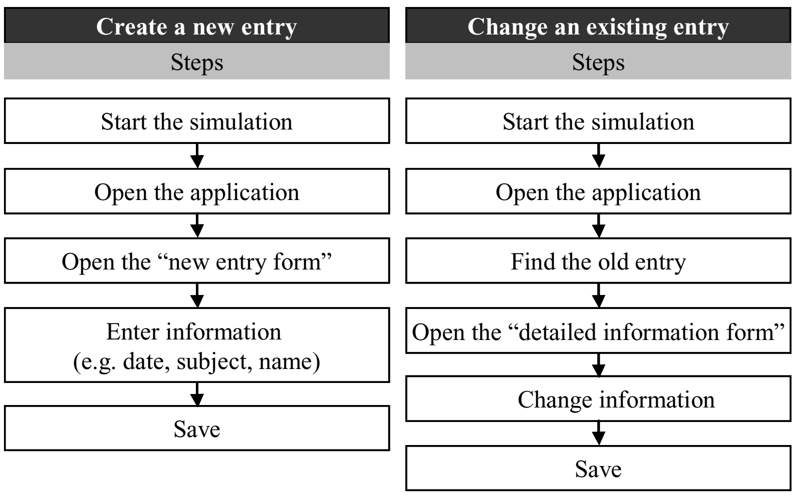


**Fig. 1.** Flowchart of the task procedure for a "new entry" and a "change entry" task

### 3.4 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 an LCD-screen (Iiyama TXA3841, TN, 15'', with a display resolution of 1024 x 768 Pixels). The software emulation exactly corresponded to the real device, in size (chassis 80 x 125 mm), display size (3.5''), font size (9 pt for functions, and 11 pt (bold) for category headers), menu structure and operational keys.

### 3.5 Procedure

Before participants started to work on the tasks, they completed a computer-based questionnaire. In order to familiarize participants with the experimental setting and to control that older adults were 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). All participants, even the older group, were highly familiar in using the mouse. 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 utilisation ("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). 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". After completing the computer-based questionnaire, the handling of the PDA-simulation was explained by the experimenter. All participants worked on four experimental tasks in each study. The order of tasks was kept constant and was presented in the following sequence: (1) create a new entry, (2) change an existing entry, (3) create a new entry, (4) change an existing entry. Finally, the ease of use ("Using the PDA 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) in a computer-based questionnaire. Participants were also asked to rate usability-problems they encountered during the interaction with PDA. Participants sat on a height-adjustable chair in a comfortable seating position. In order to optimize viewing conditions, participants were allowed to individually adjust the viewing distance and the inclination of the TFT-monitor, where the PDA-emulation was displayed. Corrective lenses, if necessary, were worn throughout the experiment.

### 3.6 Materials

In order to measure *spatial ability*, participants completed the Paperfolding-test, a spatial visualisation test [34]. 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 . The maximum score to be reached was 20.

To assess *verbal memory abilities*, a verbal memory test as used in earlier experiments [35] was conducted. In order to diminish semantic processing of the stimuli and further elaborative processes, participants had to memorize Turkish nouns, which are completely unfamiliar to native speakers of the German language. The stimuli were presented successively for three seconds on a computer screen. Immediately after the presentation of the 20 nouns (delay < 0.5 sec), participants had to recognize each target word among three phonologically and visually similar distractors. Here, the maximum score that could be reached was 15 (Figure 2).

tatil

□ tafil □ tarak □ tital □ tatil

**Fig. 2.** Item example of the verbal memory test

The *subjective technical confidence* (STC) measures the subjective confidence of a person regarding his/her competency when using technology [34]. Participants 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). The maximum score to be reached was 100. The reliability of the STC short version is high (Cronbach´s alpha = 0.89). Several studies assessing the validity of the STC showed satisfactory results and proved the construct of STC as a technology-related personality trait [36].

For the assessment of *ease of use* and *usefulness* the original items of Davis’ Technology-Acceptance-Model were used [37]. The validity and reliability of the items has been proven by Davis et al. and by a multiplicity of other empirical studies [38]. For each item, a 5-point Likert-scale was used and participants were asked to give a response ranging from 1 (strongly disagree) to 5 (strongly agree).

To get insights into sources of usability problems when working with the PDA applications, participants rated items, which referred to navigational disorientation, menu complexity, icon-design (pictorial transparency), icon-meaning (semantic transparency) and icon-naming (labeling). The items had to be answered on a 5-point Likert-scale ranging from 1 (strongly disagree) to 5 (strongly agree).

## 4 Results

Results were analysed by bivariate correlations, multivariate and univariate analysis of variance and multivariate regression analysis, with a level of significance set at 5%. The significance of the omnibus F-Tests in the MANOVA analyses were taken from Pillai values. In order to determine associations between performance outcomes and user characteristics, correlations were carried out (Pearson values for interval-scaled, Spearman-Rho values for ordinal-scaled and Eta values for nominal-scaled data).

As the experimental tasks share the same semantic context, the performance in the four tasks of an application was comprised and the means for the performance in the diary, to-do-list and directory are reported.

### 4.1 User Characteristics of the Sample

A 3x2 MANOVA was conducted in order to find out if the samples and the two age groups differed according to their cognitive abilities (spatial visualisation and verbal memory), their STC and computer expertise. No differences were found ($F(8, 174) = 0.7$, $p = 0.7$), therefore we regard the three samples as comparable and merge them for an overall-analysis of performance.

With respect to the main effect “age”, highly significant age differences were found ($F (4,87) = 33.9$, $p = 0.000$). Younger adults scored higher in spatial ability, verbal memory, STC and computer-expertise (see Table 2). No interaction between sample and age group was found.

**Table 2.** Cognitive abilities , STC and expertise for younger and older adults

| | Spatial ability [max = 20] | | Verbal memory [max = 15] | | STC [max = 100] | | Expertise [max = 160] | |
|---|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD | M | SD |
| Young adults | 14.3 | 2.7 | 11.4 | 3.0 | 76.3 | 14.7 | 95.1 | 48.5 |
| Older adults | 6.7 | 3.6 | 8.1 | 3.5 | 61.2 | 17.2 | 63.6 | 56.3 |

Controlling for gender effects it was found that male participants reported a significantly higher STC than female participants ($F (1, 95) = 13.5$, $p = 0.000$). Male participants reached an average score of 74.3, while female participants scored lower with 63.1 out of 100 points. No other effects of gender or interactions with age or sample group were found.

In order to find out how the cognitive variables, STC and expertise were associated, bivariate correlations were carried out for both age groups separately. In the younger group the cognitive variables were not strongly related. The only significant association was found for the STC and expertise ($r = 0.39$, $p < 0.05$), according to which younger adults with a high STC also reported to have a high computer experience.

A different pattern can be seen in Table 3 or the older group: Older participants with high spatial abilities also show higher scores in verbal memory, STC and expertise. Similar to the correlation pattern in the group of younger adults, higher reports of expertise of older adults correlated positively with STC-ratings.

**Table 3.** Correlations between cognitive variables, STC and expertise for older adults ($n = 48$), an asterisk indicates a significant relationship ($p < 0.05$)

| | Spatial ability | Verbal memory | STC | Expertise |
|---|---|---|---|---|
| Spatial ability | 1 | .35* | .34* | .36* |
| Verbal memory | | 1 | -.17 | -.01 |
| STC | | | 1 | .35* |
| Expertise | | | | 1 |

### 4.2 Performance of Younger and Older Adults Using a PDA

The performance of younger and older adults in the three software applications was compared by multivariate analysis. A 2 (age group) x 3 (application) MANOVA revealed a highly significant effect of age ($F (2, 180) = 72.4$, $p = 0.000$). The task effectiveness of older adults was nearly 50% lower than performance of younger adults when working with standard software applications (see Figure 4). Older adults solved only 47.9%, whereas younger adults accomplished 94.3% of the tasks. Older adults also needed more than twice as much processing time than younger adults ($M_{young}$ = 337.4 sec, $M_{old}$ = 812.5 sec).

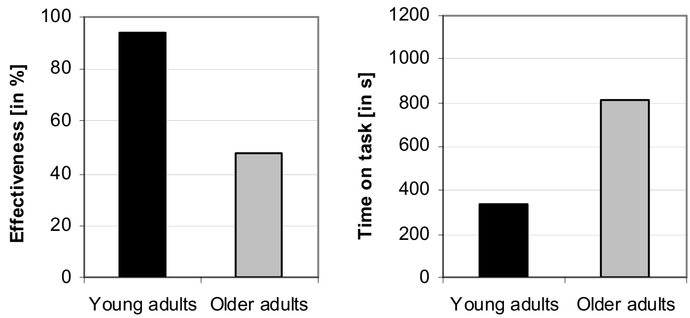


**Fig. 3.** Effectiveness (left) and time on task (right) for young and older adults

Furthermore, it is insightful whether older participants differ in their judgments with respect to the perceived ease of use and the perceived usefulness. These judgments were collected after the participants had completed the PDA tasks (see Figure 4).

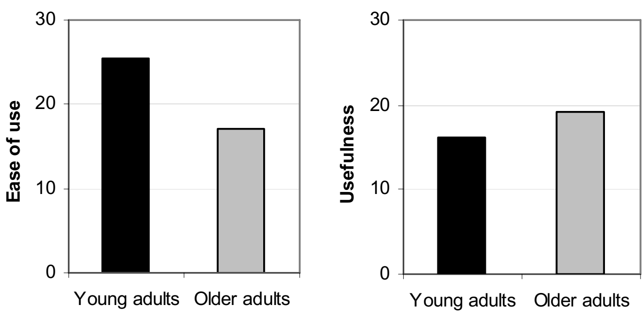


**Fig. 4.** Ease of use (left) and usefulness (right) for both age groups

As found, older participants reported significantly lower values with respect to the perceived ease of use and the perceived usefulness after working with the PDA ($F(2, 180) = 30.5$, $p = 0.000$). Older adults reported an ease of use of on average 17 (out of 30 points), whereas younger adults gave a rating of 25.5 points. The age group differences were less pronounced in the ratings of perceived usefulness: older adults gave an average rating of 16.1; the rating of the younger group was 19.2.

Comparing the performance of the two age groups in the three software applications, significant differences were found ($F(2, 180) = 12.9$, $p = 0.000$). The highest task effectiveness was reached in the digital directory application, where 84.4% of the tasks were solved, followed by the to-do-list with 66.4% and the diary with 62.5%.

Participants worked the fastest in the diary (M = 514.2 s), followed by the directory (M = 532.2 s) and the to-do-list (M =687.4 s). No differences between the three applications were found in the ratings of ease of use and usefulness. Furthermore, there was no interaction between the factors age and application.

**Table 4.** Performance in the three PDA-software applications

| | Diary | | To-Do-List | | Directory | |
|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD |
| Effectiveness in % | 62.5 | 40.2 | 66.4 | 39.5 | 84.4 | 27.5 |
| Efficiency | 514.2 | 301.2 | 687.4 | 364.2 | 523.2 | 259.9 |

In a next step we analysed the learnability regarding task effectiveness, i.e. if performance improved in the second trial of a task in comparison to the first trial (see Figure 5). A 2 (age group) x 2 (trials) x 3 (application) ANOVA with repeated measurement revealed that effectiveness increased in the second trial ($F(1, 90) = 10.0$, $p = 0.02$).

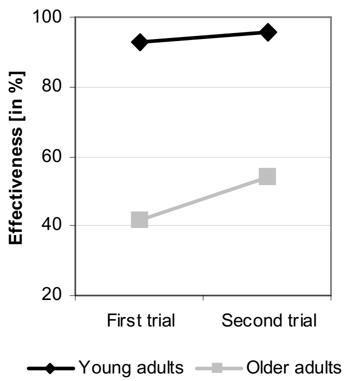


**Fig. 5.** Learnability effects in task effectiveness for both age groups

The task effectiveness increased from 67% in the first trial to 75% in the second trial. Interestingly, the learning progress was more pronounced for the older age group (from 41.7% in the first trial to 54.2% in the second trial) as compared to the younger group (from 92.7% to 95.8%), although this interaction missed statistical significance ($F(1, 90) = 3.5$, $p = 0.06$). It is remarkable, that not even the younger group was able to reach a “perfect” performance in the second trial, especially because the examined PDA tasks referred to standard applications. Apparently, the personal data management applications were not easy to use, even for a technology-prone student user group. Nevertheless, the high variances in task effectiveness indicate the effect of differential factors. Therefore the following section focuses on user characteristics, which might explain the sources of variability.

### 4.3 User Characteristics and Performance

In order to find out, which user characteristics are associated with a successful interaction with the PDA software applications, bivariate correlation analyses with performance variables and user characteristics were conducted (see Table 5). In the group of younger adults a successful performance was accompanied by high computer expertise (for task effectiveness) and high spatial abilities (for time on task). This means that young computer-experienced users were able to solve the tasks more effectively and those with high spatial abilities are able to accomplish the task more efficiently. The missing relationship between effectiveness and spatial ability might by explained by the high task performance by younger adults, i.e. the low variability in task effectiveness.

**Table 5.** Bivariate correlations between performance measures and user characteristics for younger adults. An asterisk indicates a significant relationship ($p < 0.05$).

| Younger adults (N = 48) | Spatial abil- | Verbal mem- | ST | Exper- |
|---|---|---|---|---|
| Effectiveness in % | -.01 | .05 | .19 | .30* |
| Efficiency | -.35* | .02 | -.23 | -.16 |

A similar pattern for the relationship between performance measures and user characteristics was found for the group of older adults, but here the associations were even stronger. Older participants with high spatial abilities were able to complete the tasks more effectively and efficiently. In contrary to the correlation pattern of younger adults, where expertise and effectiveness were positively related, in the group of older adults expertise was related with efficiency. This finding indicates that older adults' expertise did not enhance the success of interacting with the software applications, but was related to the time, which users needed to process the tasks.

**Table 6.** Bivariate correlations between performance measures and user characteristics for older adults. An asterisk indicates a significant correlation ($p < 0.05$), two asterisks a highly significant relationship ($p < 0.01$).

| Older adults (N = 48) | Spatial abil- | Verbal mem- | ST | Exper- |
|---|---|---|---|---|
| Effectiveness in % | .33* | .23 | .19 | .25 |
| Efficiency | -.31* | -.16 | - .27 | -.41** |

The results of the correlational analyses show that a successful interaction with the PDA is strongly associated with two user characteristics: spatial ability and computer experience. Participants of both age groups, which showed high levels of spatial ability and expertise, were able to work more successful with the personal data management software.

Summarising the findings of the user-centered analysis we found that the very same user characteristics (i.e. spatial abilities and expertise) are related to a high performance in both age groups, even though the young adults performed much better. However, the present study was not limited to a user-centered approach, that aimed on the identification of "optimal user characteristics", which promote a successful interaction with a device or which might compensate for suboptimal interface design. We also pursued a "device-centered" approach, in order to analyse the interface-design-factors, which account for performance decrements. Therefore, participants' ratings of usability problems were related to performance measures (see 4.4) and a qualitative analysis of interface shortcomings (see 4.5) was conducted.

### 4.4 Usability Problem Ratings

In a next step we wanted to explore the reasons for the performance differences between younger and older adults from the "device-perspective". First insights came from the ratings of usability problems. Nonparametric tests showed that older adults faced significantly higher usability problems regarding navigational disorientation ($M_{old} = 4.0$, $M_{young} = 3.0$; $z = -3.7$, $p = 0.00$) and a hampered memorization of functions in the menu ($M_{old} = 3.5$, $M_{young} = 2.5$; $z = - 3.0$, $p = 0.00$). Older participants reported larger disorientation problems when navigating through the PDA menu and also higher difficulties to remember the location of functions. Interestingly, the size of font or buttons was not seen as a severe usability problem in the older group. No age differences occurred in nonparametric t-Tests for usability topics like icon-design, icon-meaning, font- or button size or naming of functions (see Table 7).

**Table 7.** Bivariate correlations between usability problem ratings and performance measures for older adults. An asterisk indicates a significant correlation ($p < 0.05$), two asterisks a highly significant relationship ($p < 0.01$).

| Older adults (N = | Dis-orientation | Memorization of func- | Icon -Design | Icon-Meaning | Font- / Button size | Naming |
|---|---|---|---|---|---|---|
| Effectiveness in % | -.449** | -.356* | -.245 | -.193 | .183 | .064 |
| Efficiency | .306* | .384** | .247 | .232 | -.044 | .112 |

The reports about usability problems were correlated with performance measures in order to find out, which of the problem domains accompanied low task effectiveness and longer processing times. In the group of younger adults a reduced efficiency was related to navigational disorientation ($r = -.31$, $p < 0.05$), further significant associations were not present. However, in the group of older adults stronger associations were found: low task effectiveness and efficiency were accompanied by high navigational disorientation and problems in the memorization of the location of functions in the menu. The other usability problem domains like icon-design, icon-meaning, font- or button size and the naming of functions showed only weak relationships with performance measures. Accordingly, these aspects did not account for a lower performance in the group of older adults. The main problems, which occurred when participants reached low task effectiveness, were disorientation in the menu and difficulties in memorizing the location of functions in the menu.

### 4.5 Qualitative Analysis of Usability Problems

Based on participants' navigational path, which was recorded online in logfiles, and participants' comments while interacting with the PDA, an analysis of frequent user errors was carried out. A phenomenon-based description of the main user errors and their possible reasons is follow in the next section:

*Choosing the right application:* Especially older participants had great difficulties to decide, which application to choose in order to accomplish the experimental tasks (see Figure 6). Even though the task instructions contained hints about the correct application, participants sometimes chose the wrong one. This especially occurred in the to-do-list application and the diary application, e.g. when participants opened the diary in order to enter a new task or vice versa. We assume that the participants were guided by a inappropriate mental model and were confused by the similar semantic context of the applications. From participants' commentaries it was derived, that they have a mental model that does not differentiate between a "diary" and a "to-do-list".

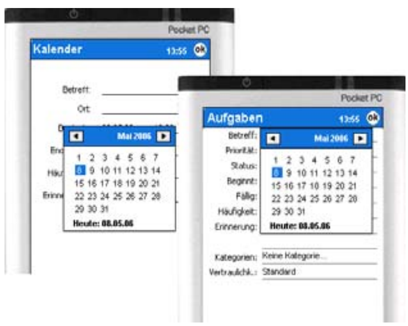


**Fig. 6.** Scaled-down screenshot of the date sheets in the digital diary (left) and the to-do-list-application

Participants reported to usually write their to do's or tasks into their diaries. Therefore, the differentiation between a diary and a to-do-list in the data management software was not comprehensible for them. Moreover, the interface-design features in both applications might have contributed to the difficulties in choosing the right application. Both, the to-do-list application and the diary application contain entry-fields where data information can be filled in by choosing the data from a date sheet (see Figure 6). Participants assumed that the calendars of both applications were linked and that information would automatically be transferred and exchanged between the two applications. In this case, the design-guideline to present operating elements like the date sheet across different software applications in a consistent way caused wrong assumptions about connections between the two applications. In order to overcome the described difficulties, users of personal data management applications should learn (e.g. in manuals or trainings) about the two different applications and their separate functioning.

*Differentiation between "Create a new entry" and "Change an existing entry":* Interestingly, in the experimental tasks where an existing entry had to be changed, participants tended to create a new entry instead of changing the given one. Based on our data we cannot interpret if participants could not find the "change"-button in order to edit the information in the already existing entry or if they did not know about the conceptual differences between creating a new entry and changing an existing one. Again, participants' commentaries about their mental model when changing an existing entry were highly insightful. When paper-based diaries or directories are used, an existing entry is "changed" by writing in the new one and by crossing out or erasing the old entry. When participants pursued the familiar procedure they knew from paper-based data management, they "only" created a new entry. In order to avoid double entries, which unnecessarily waste storage capacity and might cause confusion (e.g. when the warning signal rings for an old and a new appointment), the user manual should contain information about the location (in the menu and on the screen) and design of the "change"-button. Moreover, users should be informed about the two different concepts of creating a new entry and changing an existing entry in a training or user manual.

*Opening the "new-entry form":* Many participants had difficulties when they wanted to click on the "New-Button" in order to open the "new-entry form". Instead of clicking on the "new"-button they clicked on the "extras"-button, which was located right next to the "new"-button (see Figure 7).

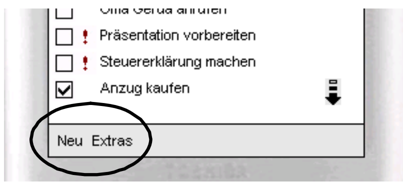


**Fig. 7.** Missing visual contour between the buttons "Neu" (left, engl. new) and "Extras" (right, engl. extras)

Based on the analysis of logfiles we assume that participants did not recognize that two separate buttons were implemented, because there was no visual contour between the two buttons. Instead of this, they perceived the field as one big button. The two buttons in the lower bar of the display should be re-designed, e.g. by encircling the icons or colouring the background.

*Saving the entries:* The most frequent error, which occurred in the group of older adults as well as in the group of younger adults, was the failure to save new entries or changes in given entries. At least two interpretations are possible to explain this error. First, we assume that participants were not acquainted with the concept of "saving", at least in the context of using a PDA device. Users of desktop computers are supposed to be acquainted to the fact that any data input has to be saved. However, users do not necessarily transfer this knowledge to software applications of mobile devices like a PDA. Instead, it was assumed that the digitally based management of personal data in digital diaries or directories activated a mental model of conventional paper-based personal data management. The personal data management software on the PDA shares conceptual and semantic features with the conventional way to organise personal data, but it also features considerable differences. Whenever new data is entered in a conventional paper-based directory, diary or to-do-list, people are used to simply write it down, and the writing process itself is – from a cognitive point of view – the concluding and confirming action. Especially older adults reported that they were sure that the information was saved after entering it in the PDA, because it was "written" on the PDA-display. Again, participants were misguided by a model or procedure, which was derived from the traditional paper-based data management, but not transferable to the utilisation of software applications. Another less probable explanation is that "saving" as the final step in the sequence of action was simply forgotten, because the motivation to finish the task and to reach the task goal was so high. In both cases, an automatic feedback, which reminds the user to save the entries before closing the application, seems to be appropriate. Interestingly, feedbacks like "are you sure that you want to close the document without saving" are highly common in desktop computers, but this feature was not applied to the design of the PDA-interface.

Summarising the findings in this section so far, the majority of usability problems found in the qualitative analysis stem from a mismatch between the users' models and the design of the software application. On the one hand, the knowledge and expectations of users about how the device may work is not appropriate and on the other hand, the device is not addressing these expectations. The implications for interface-design will be discussed in the following section.

## 5 Discussion

The present study surveyed the performance and usability ratings of 96 younger and older adults, which worked with the personal data management software of a PDA. It was found, that older adults generally show a lower performance in comparison to younger users. But it is noteworthy, that even younger users, which are much more familiar with computers and software applications, were not able to handle the software applications error-free. These findings clearly demonstrate that applications of current small screen devices like the PDA, which might be implemented in ambient computing devices, are not as "easy to handle" as they are considered to be. Even the "best" users of both age groups, i.e. healthy and comparably young older adults in the older group and students of a technical university in the younger group faced usability problems and were not able to perfectly solve the experimental tasks. However, we assume that the current findings still underestimate the "real" performance level of older users with age-related illnesses and constraints.

*Transgenerational Design*: Critical voices claim that a transgenerational interface design is not necessary, because usability-problems will become extinct, when the current generation of technology-experienced young users becomes older (cohort-explanation). From our point of view there are at least two arguments, which contradict this perspective. First, the cycles of technical innovations become faster and faster and lead to a development of novel and more complex devices, which might even be a problem for the current technology-prone generation. Second, the young generation will also grow older will be subject to cognitive ageing processes, which happen cohort-independently. This underlines the importance of research activities, which contribute to a transgenerational design, where even the “weakest” user is able to handle a technical device successfully.

*Ageing Impact:* However, we found strong evidence, that chronological age alone is not very informative for the deduction of guidelines for a transgenerational design. Hence, chronological age must be regarded as a “carrier-variable”. Based on our results the underlying reasons for performance differences are lower spatial abilities and a lower computer expertise. This finding has important implications for the interface design of ambient computing devices. Referring to current interface solutions, an effective interaction is restricted to user groups, which possess high levels of spatial abilities and / or computer expertise. However, it is highly improbable that these “optimal” characteristics are pronounced in broader and diverse user groups (e.g. as present in “old old adults”). On the one hand, “spatial abilities” decline with increasing age; therefore older adults need extra-support in order to diminish disorientation when navigating through menu-based applications.

*Typical barriers:* Beyond age-related performance differences, a final consideration is concerned with typical barriers, which hamper a proper handling of PDA personal data management applications. The detailed qualitative analysis showed that the majority of user errors could be ascribed to a lack of transferability between different mental models of the user and /or a lack of cross-platform consistency between traditional desktop computers and mobile devices. In this context, many pitfalls are present when the relation between the users’ model of the task procedure (represented as procedural knowledge) and the task structure implemented in different devices is not cognitively compatible. Nowadays, similar applications are implemented in different technical devices (e.g. personal data management applications can be found in mobile phones, smartphones, PDAs, desktop computers). Although the semantic content and the field of application are the same, the data structures or interaction rules of these technical devices are different. The user is urged to learn the similarities and differences to handle these applications in a tedious “trial and error”-way. This is especially problematical for the older user group. According to our experience, older users are highly motivated to use new technologies (Ziefle & Bay, in press), but they have immense difficulties to change over to new operating routines and they are very sensitive to a suboptimal interface design. According to Richter, Nichols, Gajos & Seffah, “we cannot expect to easily transfer our digital activities from one platform to another as our interaction context changes. The challenge for user interface design is to ensure that users can seamlessly move between interfaces and continue to accomplish their tasks, even when the interfaces are presented on devices with considerably different display and interaction mechanisms” [38]. Especially older users weigh up the costs and benefits before accepting and utilizing new technologies, and a low perceived ease of use and usefulness due to suboptimal interface design might lead to a rejection of these devices [39]. Therefore it is important that AI-research activities take the user-perspective into consideration. This refers to the identification and development of innovative areas of application but also to the inclusion of user abilities, which affect the interaction with ambient computing devices.

*Limitations of the Study:* Some final remarks deal with potential limitations of the study, with respect to methodological aspects and the generalizability of our findings. A first point refers to the dependent variables in our study. The binary definition of task effectiveness (task solved or not solved) might lead to an underestimation of performance. As observed in our study, younger and older adults did not execute the final "saving"-step. Although the task accomplishment was completely correct up to this final step, the task was assessed as "not solved", which completely ignores the performance in the rest of the task. Therefore, more sensitive measures should be applied, which more gradually assess task effectiveness. One possible solution is to analyse users' problem-solving activities while working on the tasks and to define sub goals, which have to be reached by the user. This allows a more sensitive assessment of performance on the hand and gives insights into more or less difficult elements (~sub goals) of the task. These elements can be addressed in an ergonomic redesign- or the support of users.

A second point deals with the assessment of computer expertise. In our study, computer expertise was assessed by subjective user ratings of frequency, ease and length of computer use, which were aggregated multiplicatively. However, subjective ratings are often biased by social desirability, fear of failure and misinterpretation of ones own performance. Furthermore, the items used to measure computer experience in this study (frequency and length of computer utilisation) might lack content validity. The items refer to quantitative aspects of computer utilisation and do not necessarily reflect qualitative aspects of computer expertise, i.e. computer-related knowledge. Therefore, the correlation between STC and computer-expertise in our study might be traced back on a method (self-report)-bias. In future studies, instead of using ratings of temporal aspects of computer utilisation, the knowledge of task-relevant concepts should be assessed.

In this study, we comprised the results of the three applications of the PDA, because they represent common standard applications of a PDA. Thus, the described usability problems or barriers are valid for the three application types. However, a more detailed analysis of the three applications should be carried out in future studies in order to explore performance differences in the three application types and to optimize the interface design of the three applications.

One could critically object that the PDA interface itself was not "intelligent", because it contains more inconsistencies and "logical" errors than other devices. However, in the huge number of studies dealing with shortcomings in interface design of desktop computers and mobile devices, it was consistently shown that the presence of cognitive incompatibilities is not specific for a certain device type, but for devices, which are developed without considering the human factor. The present study clearly shows that the gap, which still exists between technological genius and usability demands, might represent a serious obstacle for the acceptance of technology by broader user groups. However, if the knowledge of both, the technological and the human factors is incorporated into current design, the devices may meet the demands of users, of designers and manufacturers at the same time.

A last point refers to the experimental laboratory setting. Working with software applications in a laboratory is certainly less demanding than the utilisation in a real environment, where the workload is typically higher. Therefore, when compared to the demands of a multitasking context, the findings might be an underestimation of usability problems occurring in the real world.

## References


1. Christoph, U., Krempels, K.-H., Spaniol, O.: Automatic mapping of information service interfaces on operational elements of mobile devices. Unpublished Master Thesis, RWTH Aachen University, Germany (2006)
2. Wyeth, P., Austin, D. & Szeto, H.: Designing Ambient Computing for use in the mobile healthcare domain. In: Online-Proceedings of the CHI 2001, (2001) http://www.teco.edu/chi2001ws/17_wyeth.pdf, 01.05.2005
3. Streitz, N.A.: The Disappearing Computer: From Human-Computer Interaction to Human-Artefact Interaction (Opening Keynote). In: M. Khalid, M. G. Helander, A. W. Yeo (Ed.): Proceedings of the 7th International Conference on Work With Computing Systems (WWCS 2004), Kuala Lumpur, Malaysia (2004)
4. Ziefle, M., Arning, K. & Bay, S.: Cross-platform consistency and cognitive compatibility: How important are users` mental models for the performance when interacting with mobile devices. In: The Many Faces of Consistency, Workshop CHI 2006, http://www.multipleu.org (2006) 1-5
5. Ziefle, M. & Bay, S.: How older adults meet complexity: Ageing effects on the usability of different mobile phones. Behaviour and Information Technology (2005) 375-389
6. Czaja, S.J. & Sharit, J.: The influence of age and experience on the performance of a data entry task. Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting, Santa Monica, Human Factors and Ergonomics Society (1997) 144-147
7. Kelley, C. L. & Charness, N.: Issues in training older adults to use computers. Behaviour & Information Technology (1995) 107-120
8. Westermann, S.J.: Individual Differences in the Use of Command Line and Menu Computer Interfaces. International Journal of Human Computer Interaction (1997) 183-198
9. Arning, K. & Ziefle, M.: What older users expect from mobile devices: an empirical survey. In: Proc. IEA 2006 (2006)
10. Morrell, R. W., Mayhorn, C. B. & Bennett, J.: A survey of World Wide Web use in Middle-aged and older adults. Human Factors (2000) 175-182
11. Czaja, S.J. & Sharit, J.: Age Differences in the Performance of Computer-Based Work. Psychology and Ageing (1993) 59-67
12. Mead, S.E., Sit, R.A., Rogers, W.A., Jamieson, B.A. & Rousseau, G.K.: Influences of general computer experience and age on library database search performance. Behaviour & Information Technology (2000) 107-123
13. Birren J.E, & Fisher L.M.: Ageing and speed of behavior: possible consequences for psychological functioning. Annu. Rev. Psychol. (1995) 329-353
14. Salthouse, T. A.: The processing-speed theory of adult age differences in cognition. Psychological Review (1996) 403-428
15. Fisk, A. D. & Rogers, W. A.: Handbook of Human Factors and the Older Adult. CA, Academic Press San Diego (1997)
16. Park, D.C. & Schwarz, N. Cognitive Ageing. Philadelphia, Buchanan (1999)
17. Ziefle, M.: Ageing, visual performance and eyestrain in different screen technologies. Proceedings of the human factors and ergonomics society 45th annual meeting (2001) 262-266

18. Welford, A.T.: Practice effects in relation to age: A review and a theory. Developmental Neuropsychology (1985) 173- 190
19. Salthouse, T. A., & Babcock, R. L.: Decomposing adult age differences in working memory. Developmental Psychology (1991) 763-776
20. Vicente, K. J. Hayes, B. C. & Williges, R. C.: Assaying and isolating individual difference in searching a hierarchical file system. Human Factors (1987) 349-359
21. Stanney, K. & Salvendy, G.: Information Visualisation: Assisting low-spatial individuals with information access tasks through the use of visual mediators. Ergonomics (1995) 1184-1198
22. Downing, R.W., Moore, J.L. & Brown, S.W.: The effects and interaction of spatial visualisation and domain expertise on information seeking. Computers in Human Behaviour (2005) 195-209
23. Cherry, K.E., & Park, D.C.: Individual differences and contextual variables influence spatial memory in young and older adults. Psychology and Ageing (1993) 517-525
24. Park, D. C., Cherry, K. E., Smith, A. D., & Lafronza, V. N.: Effects of distinctive context on memory for objects and their locations in young and older adults. Psychology and Ageing (1990) 250-255
25. Edwards, D.M. & Hardman, L.: Lost in hyperspace: Cognitive mapping and navigation in a hypertext environment. In R. McAleese (Ed.), Hypertext: theory into practice. Oxford: Intellect limited (1989) 105-125
26. Ziefle, M. & Bay, S.: How to overcome disorientation in mobile phone menu. A comparison of two different navigation aids. Human Computer Interaction (2006)
27. Kline, D.W. & Scialfa, C.T.: Sensory and perceptual functioning: Basic Research and Human Factors Implications. In: A. D. Fisk & W. A. Rogers (Eds.). Handbook of Human Factors and the Older Adult San Diego: Academic Press (1997) 27-54
28. Craik, F.I.M. & Salthouse, T.A. (Ed.): Handbook of Ageing and Cognition. Hillsdale, N.J.: Lawrence Erlbaum Associates (1992)
29. Haegerstrom, T. Portnoy, G., Schneck, M.E. & Brabyn, J.A.: Seeing into old age: vision function beyond acuity. Optom. Vis. Science (1999) 141-158
30. Vercruyssen, M.: Movement Control and Speed of Behavior. In A.D Fisk and W.A. Rogers (Eds.), Handbook Human Factors and the Older Adult. San Diego: Academic Press (1997) 55-86
31. EN ISO 9241-11: Ergonomic requirements for office work with visual display terminals. Part 11: Guidance on usability. Berlin, Germany: Beuth (1997)
32. Ekstrom, R. B., French, J. W., Harman, H. H, & Dermen, D.: Manual for the Kit of Factor-Referenced Cognitive Tests. Princeton, NJ: Educational Testing Service (1976)
33. Bay, S., & Ziefle, M.: Design for All: User Characteristics to be Considered for the Design of Devices with Hierarchical Menu Structures. In H. Luczak and K. Zink (Eds.) Human Factors in Organisational Design and Management Santa Monica: IEA. (2003) 503-508
34. Beier, G.: Kontrollüberzeugungen im Umgang mit Technik. [Locus of control in the interaction with technology] Report Psychologie (1999) 684-693
35. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly (1989) 319-340
36. Adams, D.A., Nelson, R., & Todd, P.A. Perceived usefulness, ease of use, and usage of information technology: A replication. MIS Quarterly (1992) 227-247
37. Richter, L., Nichols, J., Gajos, K., Seffah, A. The Many Faces of Consistency in Cross-Platform Design. In: Extended Proceedings of CHI'2006 (2006)
38. Arning, K. & Ziefle, M.: Cognitive and personal predictors for PDA navigation performance. Human Factors (submitted)