

# The influence of age in usability testing

Andreas Sonderegger<sup>\*</sup>, Sven Schmutz, Juergen Sauer

Department of Psychology, University of Fribourg, Fribourg, Switzerland

## ARTICLE INFO

### Article history:

Received 27 July 2014

Received in revised form

5 January 2015

Accepted 10 June 2015

Available online 15 August 2015

### Keywords:

Usability

Age

User experience

Performance

Affect

## ABSTRACT

The effects of age in usability testing were examined in an experiment. Sixty users from two age groups ( $M = 23.0$  yrs,  $M = 58.1$  yrs) operated two technical devices (keyboard-based and touchscreen-based smartphones). In addition to various performance measures (e.g. task completion time, task completion rate), several subjective measures were taken (e.g. perceived usability, affect, and workload). The results showed better performance scores for younger adults than older adults for task completion time. For older adult users there was a mismatch between usability ratings and task completion time but not between usability ratings and task completion rate. Age-related differences in the importance of speed and accuracy in task completion point to the need to consider more strongly the factor user age in usability research and practice.

© 2015 Published by Elsevier Ltd.

## 1. Introduction

For a number of years, pervasiveness of technology and an ageing population have represented two important trends in industrialised countries. The fact that the two trends coincide poses several interesting challenges for the science of ergonomics. The increasing omnipresence of technology, often paralleled by high levels of automation, may result in usability problems for all users but will pose a particular challenge to older adults (Rogers et al., 2005). The significance of older adults as a specific user group is on the increase due to demographic changes and the growing importance of technology in society. Against this background, the present article aims to examine issues related to the design of interactive consumer products for older adult users, with a special emphasis being placed on the implications for usability testing as an important method in product design.

### 1.1. Older adult users and interactive technology

Ergonomics has typically aimed to design technical systems for a broad population of users but, at the same time, it acknowledged the existence of special user groups and the necessity to consider their special needs (Kroemer, 2006). Older adult users are one of

these special user groups. In industrialised countries, this group is a substantial part of society already and is expected to increase further in size due to demographic changes.

The terminology of what an older adult user constitutes is quite diverse. For example, Kroemer (2006) stated that in the US “[...] a ‘middle-aged’ person becomes an ‘older’ at 40 or 45 years of age, then ‘elderly’ at about 65 years, ‘old’ at 75 years, and ‘very old’ (or ‘old–old’) if one lives beyond 85 years” (p.128). However, this terminology is not used consistently in the literature and there may be disagreement with regard to the threshold above which a person is considered to be ‘old’.

More important than the definitions of terms surrounding advanced chronological age are the many changes that are associated with the process of ageing. When age-related issues are discussed, a large number of functional impairments have been mentioned, including poorer eyesight, deteriorating hearing, degradations in manual skills, and decline in memory performance (e.g. Kroemer et al., 2001; Matthews et al., 2000).

These numerous age-related changes at the cognitive, perceptual and motor level are highly relevant when interactive technology is being used. For example, for the use of computers and the web, cognitive capabilities or “fluid abilities” are of particular importance (Czaja and Lee, 2007; Garfein et al., 1988), including short term memory span, information processing speed, spatial abilities, and abstract reasoning (Hanson, 2009). Since these ‘fluid abilities’ are important for the adaptation to novel situations, including the adoption and use of new technology, this may result in age-related impairments (Pak et al., 2009; Hanson, 2009). In

<sup>\*</sup> Corresponding author. University of Fribourg, Department of Psychology, Rue de Faucigny 2, 1700 Fribourg, Switzerland.

E-mail address: [andreas.sonderegger@unifr.ch](mailto:andreas.sonderegger@unifr.ch) (A. Sonderegger).

addition, abilities such as spatial thinking or memorization are crucial in building correct mental models of interfaces (Brunsman-Johnson et al., 2015). These mental models are essential for using an interface like a website (Pak et al., 2006). As cognitive abilities decline with age, this seems to have an influence on the performance of older adult users when using interactive technology. Deteriorations in perceptual abilities can be a further reason for problems in using certain technological products. Poor visual scanning capabilities or a reduction in acuity, colour perception and contrast discrimination may lead to difficulties in perceiving information and reading text (Armbrüster, 2007; Brunsman-Johnson et al., 2015). For motor skills, there is evidence that older adult users show poorer performance than younger users in the form of lower peak velocities of movements, more non-productive sub-movements, and particular difficulties when using fine motor skills (Jacko et al., 2000; Keates and Trewin, 2005). Furthermore, reaction time slows down with increasing age (Stelmach et al., 1987) and the speed-accuracy trade-off changes towards accuracy in older adult users (Welford, 1976).

In addition to various decrements in functional abilities, attitudes towards technology are related to the age of users as well. Older adult users have often less positive attitudes towards new technology compared to younger users (Dyck and Smither, 1994; Chua et al., 1999). These positive attitudes were found to be linked to computer experience and to influence the frequency with which technology is used (e.g. Elias et al., 2012; Chua et al., 1999). A number of concepts were proposed to describe different facets of user attitude towards technology including technophobia (e.g. Rosen and Weil, 1995), computer anxiety (e.g. Heinssen et al., 1987), computer playfulness (Webster and Martocchio, 1992), and technology commitment (Neyer et al., 2012). One may presume that these facets of user attitudes towards technology are also influenced by a process of technological socialisation (i.e. the way the user has been introduced to technological changes). Users born after the advent of the internet and its main applications such as email have grown up with computerised technology and may differ from those who have learnt about such technology when they were already adults. This difference in technological socialisation was described by the terms 'digital natives' and 'digital immigrants' (Prensky, 2001). Having been immersed in technology all their lives may influence the ease with which users operate such technology. The idea of a technological generation gap has been discussed in various domains such as ergonomics (e.g. Kolikant, 2010), educational science (e.g. Bennett, et al., 2008) and sociology (e.g. Prensky, 2001).

As a result of the recognised need to compensate for decrements in the functional abilities of older adult users, guidelines have been developed to address these issues (e.g. AgeLight, 2001; Petrie et al., 2013; Zaphiris et al., 2007). These guidelines also include the needs of other special user groups, such as blind, oversized or wheelchair-based users. The guidelines contain very specific rules to be followed during product design. For example, for older adult users too small font sizes and information overload on webpages need to be avoided and instructions of how to use the website should be provided (Nahm et al., 2004). Other work has also successfully explored the utility of adaptive training devices to support the older adult users (Bruder et al., 2014).

Attitudes and functional limitations of older adult users are especially important when it comes to the adoption and use of new technology such as touchscreen devices. Touchscreen interfaces have become increasingly popular due to the flexibility in design and convenience in usage (Taveira and Choi, 2009). While a considerable amount of research has addressed issues such as key size and spacing in touchscreen design using typical user populations like younger adults (e.g. Schedlbauer, 2007; Scott and

Conzola, 1997; Colle and Hiszem, 2004), comparatively little is known about the usage of touchscreen devices by older adult users. Research has shown that the use of touchscreen devices for text entry purposes may lead to decreased performance compared to the use of traditional keyboards (Plaisant and Sears, 1992). Although their finding is based on data from younger adult users, Wright et al. (2000) reported similar results for older adults when comparing a touchscreen-based handheld device with a keyboard-operated device. For the design of touchscreen interfaces, research has also revealed that the size of buttons and other interactive features play an important role for performance and user satisfaction (Lee and Zhai, 2009; Jin et al., 2007). Fezzani et al. (2010) compared the influence of target size in touchscreen interaction for younger and older adult users and found that reduced button size affected older adult users more strongly than younger ones in the form of greater difficulties with accurate pointing, increased time on task, and higher mental workload. These age-related effects were also reflected in different recommendations for minimal button size for younger and older adults (e.g. 10 mm vs. 11.5 mm; Lee and Zhai, 2009; Jin et al., 2007). Interestingly, research revealed that both, younger and older adult users preferred touchscreen keypads to physical keypads for numeric data entry tasks since, due to the virtual keypad being situated on the screen, it required less effort for users to divide their attention between the input and the display function of the device (Chung et al., 2010). On the other hand, physical keys have the advantage of giving direct tactile feedback to the user whereas feedback provided by touchscreens is usually limited to the visual modality, which may result in lower performance (Caprani et al., 2012). Overall, the empirical work reviewed points to the need to consider the differential effects that the design of interactive devices may have for user groups of different ages.

## 1.2. The factor age in usability testing

In consumer ergonomics, an important aspect of product development is to examine the usability of a product by using usability evaluation methods (e.g. Nielsen, 1993). A widely used and very effective usability evaluation method is the usability test. It aims to evaluate the product by setting up a realistic task scenario for product usage involving prospective users. Despite the abundant literature on developmental changes in older adults, little is known about how such age-related changes affect the outcomes of usability tests. This applies to measures typically taken in usability tests (e.g. performance, perceived usability) but also to measures not that widely used (e.g. affect and workload).

In usability testing, performance measures are divided into indicators of effectiveness and efficiency (Jordan, 1998). Effectiveness refers to the degree to which a task was successfully carried out (e.g. task completion rate) whereas efficiency is concerned with the ease with which the task is carried out (e.g. task completion time, error rate). While the use of measures of perceived usability is standard practice in usability tests, affect have only more recently been considered to be a relevant aspect of usability with fun and pleasure gaining in importance for product usage. Subjective workload is rarely used in usability testing, which is surprising given that already in the 1990ies it has been argued that a measure of workload should be taken (Jordan, 1998) and given that measuring perceived workload is standard practice in work ergonomics (e.g. Wickens and Hollands, 2000). The rare use of affect and workload as outcome measures does not only apply to the testing of older adult users but to usability testing in general.

Experimental research generally showed that during the use of artefacts user performance was generally lower for older adult users than for younger ones (e.g. Armbrüster et al., 2007;

Jastrzemski, et al., 2005; Smith et al., 1999; Sonderegger et al., 2014). Interestingly, when the artefacts were subsequently redesigned with a view to increasing support to older adults (e.g. by using simpler terminology or removing unnecessary text), younger users benefited from these changes, too (Chadwick-Dias et al., 2003; Worden et al., 1997). A comparison of different computer devices (i.e. mouse and touchpad) across age groups also revealed age-related differences (Hertzum and Hornbæk, 2010). When using a mouse, younger adult users completed the tasks more quickly than older adults but there was no difference between groups with regard to error rate. When using a touchpad, both groups slowed down but older adults did so more strongly than younger adults. Adolescents as a third age group were also tested in Hertzum and Hornbæk's study and generally emerged between younger and older adults for the different performance measures, indicating that age-related influences are not necessarily of a linear nature. Research also showed that older adult users benefited more from practice on the device (i.e. touchpad and track point) than younger adult users did (Armbrüster et al., 2007). Interestingly, when employing a computer keyboard typing accuracy of older adult users was similar to the performance of younger adults due to the use of compensatory strategies, while older adults showed lower performance for typing speed (Salthouse, 1984; Bosman, 1993; Wright et al., 2000). Participants' age was also shown to have an influence on outcome measures other than performance. Visual appeal and readability of a website were rated more positively by older participants than younger participants (Chadwick-Dias et al., 2003). Interestingly, this pattern was not matched by performance since older adult users showed lower performance than younger participants (due to increased difficulties in reading text). Despite these problems, older adults provided higher ratings for perceived usability than younger ones (Chadwick-Dias et al., 2003). A similar dissociation between subjective usability ratings and performance measures within older and younger test participants was reported by Sayers (2004).

### 1.3. The present study

The literature review revealed that the empirical work addressing the issue of participants' age in usability testing largely focused on objective performance (e.g. speed and accuracy of task completion) while little is known about perceived usability as another critical parameter in usability testing (see also Piqueras-Fiszman et al., 2011). The present study aimed to measure both types of data by adopting a more comprehensive methodological approach, which addressed user experience from a broader perspective, including subjective user ratings which are not typically measured in usability testing (e.g. affect, and perceived workload). Using a broader methodological approach helps gain a better understanding of the multiple effects and trade-offs in usability testing.

Participants from two age groups were recruited for the study. One group represented younger adults, which were brought up with computerised technology. The other group consisted of older adult users who learnt about such technology when they were already at the age of an adult. A smartphone was chosen as a product to be examined in usability tests, which allowed a broad range of tasks to be completed (e.g. entering text, taking pictures).

In one condition, user product interaction took place by using a keyboard whereas in the other condition a touchscreen was employed. Two input devices were chosen for testing to be able to determine the effects of age on more than one interface design option. Therefore, the interaction between input device and age was of great relevance in the present study since it provides an indication of whether age-related effects are constrained to a

certain technical device or of a more general nature. In contrast, the main effect of input device was of lower relevance because we were not interested in any differences in usability between devices per se. Therefore, we did not put forward any hypotheses related to the main effects of input device and will not discuss any differences that may be found.

We predicted that younger adult users would show better performance than older ones. This would apply to speed as well as accuracy of performance, though the effects were expected to be stronger for the former. With regard to perceived usability, no prediction was made since there is little guidance in the literature about the difference in subjective usability ratings as a function of age. It is conceivable that older adult users give lower usability ratings than younger users because the former perform less well. Conversely, it is also conceivable that older adult users give higher ratings than younger users because they have lower expectations about the usability of technology and may attribute poorer performance to their own functional limitations rather than inadequate product design or even do not evaluate their performance adequately because of a lack of experience in using smartphones and therefore they do not perceive the usability as poor.

## 2. Method

### 2.1. Participants

Sixty participants (50% female), aged between 19 and 73 years ( $M = 40.6$ ;  $SD = 18.2$ ), took part in this study. Owing to the research question, a bimodal age distribution was created, with the group of 30 older participants ranging from 52 to 73 yrs ( $M = 58.13$  yrs;  $SD = 5.43$ ) whereas the group of 30 younger participants was aged between 19 and 29 yrs ( $M = 23.0$  yrs;  $SD = 2.39$ ). All participants were recruited among friends and family of members of Fribourg University. Participants were not paid for their participation. The sample is described in more detail in section 2.3 by outlining the various user characteristics that were measured as covariates.

### 2.2. Experimental design

A 2 x 2 between-subjects design was employed in this quasi-experiment, with age of participant (old vs. young) and input device (touchscreen vs. keyboard) as between-subjects variables. Participants were randomly assigned to a group using either touchscreen or keyboard. There was no difference in age between the two user groups for the input device conditions ( $M_{\text{keyboard}} = 41.43$  yrs;  $SD = 19.81$ ;  $M_{\text{touchscreen}} = 39.70$  yrs;  $SD = 16.73$ ;  $t(58) = .36$ ,  $p > .05$ ).

### 2.3. Material

The HTC Touch Pro 2 was chosen for this study because it was equipped with a QWERTY keyboard as well as a 3.6" TFT touchscreen interface. The touchscreen keyboard was very similar to the physical keyboard with regard to the positioning and the size of the keys. The operating system installed on the device was Windows Mobile 6.1 Professional. For further analysis of the performance data, the interaction of participants with the smartphone was logged via screen recorder (Pocket Controller Pro) on an HP Pavillon dv3550ez Notebook, which was connected with the smartphone via Bluetooth.

### 2.4. User tasks

Test participants had to complete two tasks: a text entry task and a menu selection task. In the text entry task, participants were

asked to type on the smartphone ten short phrases that are easy to remember. The phrases were chosen from a set of sentences that was developed for performance testing on text entry tasks (e.g. MacKenzie and Soukoreff, 2002). Examples of phrases are 'We currently eat tomato soup' and 'Where did I leave my glasses'. Before data collection began, participants were asked to type in three phrases to become familiar with the interface. After the practice trial, participants were asked to enter ten further phrases, which were presented one after the other on a sheet of paper.

The menu selection task consisted of three subtasks: taking a picture with the built-in camera of the smartphone, adding a new contact to the address book, and sending a short text message ('I am done') to a phone number stored on the smartphone. The instructions for the menu selection task were presented on a sheet of paper.

## 2.5. Measures and instruments

### 2.5.1. Dependent variables

**2.5.1.1. Perceived usability.** Perceived usability was measured by the well-established Post Study System Usability Questionnaire (PSSUQ; Lewis, 1995). The questionnaire uses a 7-point Likert scale as response format, ranging from 'strongly agree' to 'strongly disagree'. It comprises 19 items and three subscales which are 'system usefulness', 'information quality' and 'interface quality'. The psychometric properties of the PSSUQ are very good, with a Cronbach's alpha ranging from .90 to .97 (Lewis, 1995). For this experiment, the questionnaire was slightly modified by replacing the expression "system" with "mobile phone" (e.g. "it was simple to use this mobile phone").

**2.5.1.2. Affect.** Users' affective states were measured by means of the German version of the *positive and negative affect schedule* (PANAS) (Krohne et al., 1996). The scale measures two distinct dimensions of affect, positive affect (PA) and negative affect (NA). Each dimension is measured by ten items describing specific affective states (e.g. active, interested, sad, and angry) (Watson et al., 1988). As response format, a 5-point Likert scale was used, ranging from "not at all" to "extremely". The German version of the scale enjoys good psychometric properties (Cronbach's  $\alpha = .85$  for PA and  $\alpha = .84$  for NA; Krohne et al., 1996).

**2.5.1.3. Subjective workload.** Subjective workload is a variable also considered to be relevant in the field of consumer product design as an objective indicator for usability (e.g. Jordan, 1998; van Schaik and Ling, 2009), though it is more typically applied in the context of work (e.g. Wickens and Hollands, 2000). It was measured by the NASA Taskload Index (TLX). The scale, originally developed by Hart and Staveland (1988), measures six items (task-related mental, physical and temporal demands as well as performance, effort and frustration) on a 20-point Likert scale ranging from "very low" to "very high".

**2.5.1.4. User performance.** User behaviour was recorded during the whole experiment using a screen recorder. Due to inherent differences in the nature of the experimental tasks, different performance measures were taken. For task 1 (data entry task), three efficiency measures as indicators of user performance were calculated: writing speed (words per min), number of clicks, and number of typing errors. For task 2 (menu selection task), task completion rate was recorded as an effectiveness measure.

### 2.5.2. Covariates

**2.5.2.1. Experience in using technology.** Several measures of participants' experience with technical devices were taken.

Participants rated their experience of mobile phones, computers, QWERTY keyboards, touchscreen keyboards, and mobile phone keypads. For each item, a picture of the respective device was presented in the questionnaire (to ensure that there were no misunderstandings about the nature of the technical device), with the items being worded as follows: 'How experienced are you in using...?' A Likert scale ranging from 1 ("little experienced") to 5 ("highly experienced") was used for each item. Furthermore, participants were asked for how many years they had owned a mobile phone and how often they used it a day. Among the covariates examined, it emerged that compared to younger users older adult users owned a mobile phone for longer, had less experience with QWERTY keyboards, and less experience with alphanumeric mobile phone keypads (see Table 1). For the other covariates no difference was found between age groups. A comparison of possible differences between input devices revealed no such differences for any of the covariates. Please note that the touchscreen experience of some participants largely stemmed from devices other than mobile phones (e.g. ticket vending machines or touchscreen computers) because at the time of the study mobile phones were typically operated by a keyboard.

**2.5.2.2. Attitudes towards computers.** Rosen and Weil's (1992) Measurement Technophobia Instrument (MTI) was used to assess participants' attitudes towards computers. The MTI consists of three separate instruments designed to assess different facets of technophobia: Computer Anxiety Rating Scale (CARS-C), Computer Thoughts Survey (CTS-C), and General Attitudes Toward Computers Scale (GATCS-C). In the CARS-C, 20 situations of computer usage are described (e.g. 'Applying for a job that requires some computer training'). Participants are asked to indicate on a 5-point Likert scale (ranging from "not at all" to "very much") how frightened or nervous they would feel in each situation. The CTS-C consists of 20 items describing positive or negative cognitions that users may have when interacting with computers ('I am going to make a mistake' or 'This will be fun'). The cognitions are rated with regard to their experienced frequency when using a computer on a 5-point Likert scale (ranging from "never" to "very often"). In the GATCS-C, 20 statements expressing attitudes towards the use of computers and technology (e.g. 'Computers can ruin interpersonal relationships') are rated by participants on a 5-point Likert scale ("I do not agree at all" to "I completely agree"). Good psychometric properties of the CARS-C and the CTS-C were reported (with Cronbach's  $\alpha > .80$ ) whereas they were less satisfactory for the GATCS-C (Cronbach's  $\alpha < .40$ ; Rosen and Weil, 1992; Anthony et al., 2000). For the covariates measured, the analysis only revealed a significant age-difference for computer anxiety (see Table 2). None of the other

**Table 1**

Measures of mobile phone usage and subjective experience with technology as a function of age.

Age group	Younger adults	Older adults
	M (SD)	M (SD)
Ownership of mobile phone (yrs) <sup>1</sup>	8.57 (2.52)	10.87 (5.28)
Frequency of mobile phone usage (No/day)	23.33 (16.6)	14.03 (23.0)
Experience in using mobile phones <sup>2</sup> (1–5)	3.47 (1.31)	2.8 (1.42)
Computer experience <sup>a</sup> (1–5)	3.5 (1.36)	3.07 (1.68)
QWERTY experience <sup>a</sup> (1–5) <sup>2</sup>	3.93 (1.51)	2.77 (1.78)
Touchscreen experience <sup>a</sup> (1–5)	2.73 (1.62)	2.43 (1.63)
Mobile phone keypad experience <sup>a</sup> (1–5) <sup>3</sup>	4.83 (.38)	3.33 (1.54)

<sup>1</sup>Significant difference  $t(41.49) = 2.16, p < .05$ .

<sup>2</sup>Significant difference  $t(56.51) = 2.74, p < .01$ .

<sup>3</sup>Significant difference  $t(32.51) = 5.18, p < .01$ .

<sup>a</sup> Concept was measured with a 5-point Likert scale with the scale range being presented in parenthesis.



**Table 2**  
Measures of user characteristics as a function of age.

Age group	Younger adults	Older adults
	M (SD)	M (SD)
Computer anxiety (CARS-C) <sup>1</sup>	1.71 (.54)	2.34 (1.05)
Computer thoughts (CTS-C)	3.86 (.53)	3.64 (.65)
General attitude toward computers (GATCS-C)	3.28 (.34)	3.22 (.33)
Neuroticism	1.51 (.64)	1.43 (.51)
Extraversion	2.7 (.49)	2.55 (.49)
Openness to experience <sup>2</sup>	2.78 (.49)	2.45 (.51)
Agreeableness	2.69 (.47)	2.8 (.45)
Conscientiousness <sup>3</sup>	2.68 (.64)	3.04 (.47)

Each concept was measured by a 5-point Likert scale, with the scale ranging from 1 to 5.

<sup>1</sup>Significant difference  $t(43.35) = 2.91, p < .01$ .

<sup>2</sup>Significant difference  $t(58) = 2.59, p < .05$ .

<sup>3</sup>Significant difference  $t(58) = 2.45, p < .05$ .

covariates showed a significant difference between the two age groups. There was also no effect of input device.

**2.5.2.3. Personality.** Personality characteristics were measured with the short version of the Neo Personality Inventory (NEO-PI-R, Costa und McCrae, 1992), which was translated into German by Ostendorf and Borkenau (2008). Sixty statements reflecting the five dimensions of personality (neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness) were rated by participants on a 5-point Likert scale. Reliability of the NEO-FFI (Form S) is good (Cronbach's  $\alpha$  ranging from .68 for agreeableness up to .86 for neuroticism; Costa und McCrae, 1992). The analysis revealed a significant difference between age groups for openness to experience and conscientiousness but not for the three other personality dimensions (see Table 2). There was no effect as a function of input device.

## 2.6. Procedure

The experiment was carried out in a usability laboratory at the University of Fribourg. The experimenter welcomed the participants and explained to them that the purpose of the study was to evaluate the usability of a smartphone. It was checked beforehand that participants had no previous experience with the particular smartphone used in the study. Participants were instructed that they had to carry out several tasks with a smartphone with a view to gaining feedback about its usability. It was emphasised that the aim was to evaluate the smartphone and not the participant. Before operating the technical device, participants rated their current mood by means of the PANAS scale. Then, they completed the two experimental tasks. After task completion, participants rated their mood with the PANAS scale again, evaluated the usability of the smartphone using the PSSUQ and rated their subjectively experienced workload with the NASA-TLX. Finally, the experimenter administered three questionnaires assessing computer anxiety (CARS-C, CTS-C, and GATCS-C) and the personality inventory NEO-FFI. The experiment took about 60 min to complete.

## 2.7. Data analysis

The data were analysed with a two-factorial analysis of covariance, using the following covariates: openness to experience, conscientiousness, computer anxiety (CARS-C), and expertise in using QWERTY and alphanumerical keyboards. A user variable was entered as a covariate if significant differences between the two age groups were found. Effects of covariates are only reported in the results section if they reached statistical significance.

## 3. Results

### 3.1. Performance

#### 3.1.1. Writing speed (task 1)

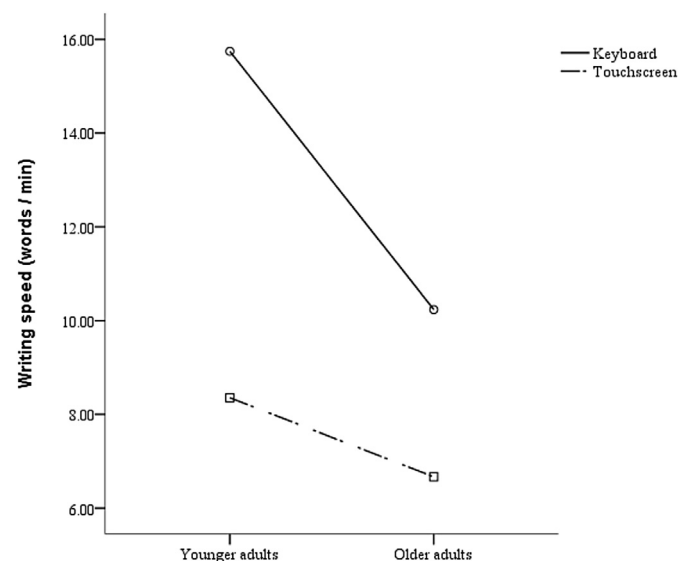
This efficiency variable measured the speed with which the data entry task was completed (words/min). The results indicated that younger participants were more efficient ( $M = 13.04$ ;  $SD = 4.95$ ) than older participants ( $M = 7.46$ ;  $SD = 2.92$ ). This difference between age groups was significant ( $F(1, 51) = 12.28, p < .001$ ;  $\eta^2_{\text{partial}} = .19$ ). In addition, the interaction between age and input device was significant ( $F(1, 51) = 6.00, p < .05$ ;  $\eta^2_{\text{partial}} = .11$ ). This was because when younger participants operated the keyboard higher performance levels were recorded than in the three other experimental conditions (see Fig. 1). Finally, analysis of covariance revealed a main effect of input device, with participants using the touchscreen showing poorer performance ( $M = 7.85$ ;  $SD = 2.32$ ) than those operating the keyboard ( $M = 12.65$ ;  $SD = 5.65$ ;  $F(1, 51) = 52.18, p < .001$ ;  $\eta^2_{\text{partial}} = .51$ ).

#### 3.1.2. Number of clicks (task 1)

This efficiency measure counted the number of clicks needed by participants to complete the text entry task. The results revealed no difference between age groups ( $M_{\text{young}} = 409.6, SD = 36.8$ ;  $M_{\text{old}} = 423.9, SD = 57.8$ ;  $F < 1$ ). No interaction was found either. However, the analysis showed that participants using the touchscreen needed more clicks to complete the task ( $M = 435.30, SD = 52.57$ ) than participants employing the keyboard ( $M = 398.17, SD = 36.47$ ;  $F(1, 51) = 7.86, p < .01$ ;  $\eta^2_{\text{partial}} = .13$ ).

#### 3.1.3. Number of typing errors (task 1)

Like the preceding measure, the number of errors represents an efficiency measure and is related to accuracy in performance. It was unaffected by age-related factors, neither showing a main effect of age ( $M_{\text{young}} = .63, SD = 1.1$ ;  $M_{\text{old}} = 1.87, SD = 3.0$ ;  $F < 1$ ), nor an interaction ( $F < 1$ ). Finally, no performance difference between input devices was found ( $F < 1$ ). The analysis of covariance revealed a significant negative association ( $r(60) = -.54$ ;  $p < .001$ ) between alphanumerical experience and the number of typing errors ( $F(1, 51) = 7.98, p < .01$ ;  $\eta^2_{\text{partial}} = .14$ ). We also converted this measure of



**Fig. 1.** The efficiency indicator writing speed (words/min) as a function of age and input device.

efficiency into an effectiveness measure by computing whether participants completed the tasks without committing a typing error (i.e. only in this case it was scored as effectiveness in task completion). The analysis revealed a similar pattern as above. Neither the main effects nor the interaction showed significant differences between groups (all  $F$ 's < 1.63;  $p$  > .20).

### 3.1.4. Speed-accuracy trade-off

In order to determine whether there were changes in the speed-accuracy trade-off for participants, we created the following index: 'number of words'/'(number of errors' + 1). The analysis of variance showed a main effect of age group, with younger adults ( $M = 9.86$ ) showing a stronger propensity towards speed than older adults ( $M = 4.96$ ;  $F(1, 51) = 20.6$ ,  $p < .001$ ;  $\eta^2_{\text{partial}} = .27$ ). An effect of input device was also recorded, with touchscreens producing a stronger relative tendency towards speed at the expense of accuracy than keyboards ( $M_{\text{touchscreen}} = 9.86$ ;  $M_{\text{keyboard}} = 5.73$ ;  $F(1, 51) = 9.63$ ,  $p < .005$ ;  $\eta^2_{\text{partial}} = .27$ ). No interaction was observed  $F(1, 51) = 1.58$ ,  $p < .05$ ;  $\eta^2_{\text{partial}} = .03$ ).

### 3.1.5. Task completion rate (task 2)

This variable represents an effectiveness measure emphasising the accuracy of task performance. There was little difference between age groups for successful task completion ( $F(1, 51) = 2.56$ ,  $p > .05$ ). However, older participants working with the keyboard were considerably less effective with regard to task completion compared to the other three experimental conditions (see Fig. 2). As Fig. 2 shows, a significant interaction between age and input device was found ( $F(1, 51) = 4.25$ ,  $p < .05$ ;  $\eta^2_{\text{partial}} = .08$ ). There was no main effect of input device ( $F(1, 51) = 2.25$ ,  $p > .05$ ).

## 3.2. Perceived usability

Analysis of covariance revealed a significant cross-over interaction between age and input device on the overall PSSUQ measure ( $F(1, 51) = 9.33$ ,  $p < .01$ ;  $\eta^2_{\text{partial}} = .16$ ). As Fig. 3 shows, usability ratings of older participants were higher for the touch keyboard compared to the physical keyboard while usability ratings of younger participants were higher for the physical keyboard than for the touch keyboard. No significant difference between age groups was observed ( $F < 1$ ). There was no effect of input device

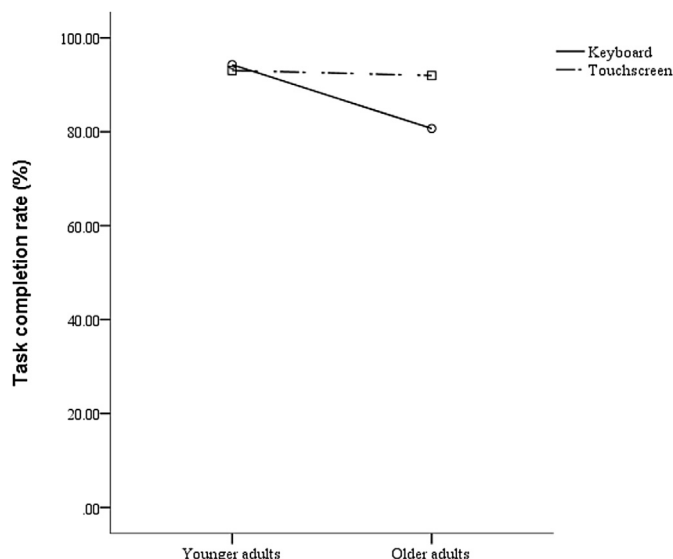


Fig. 2. Task completion rate (%) as a function of age and input device.

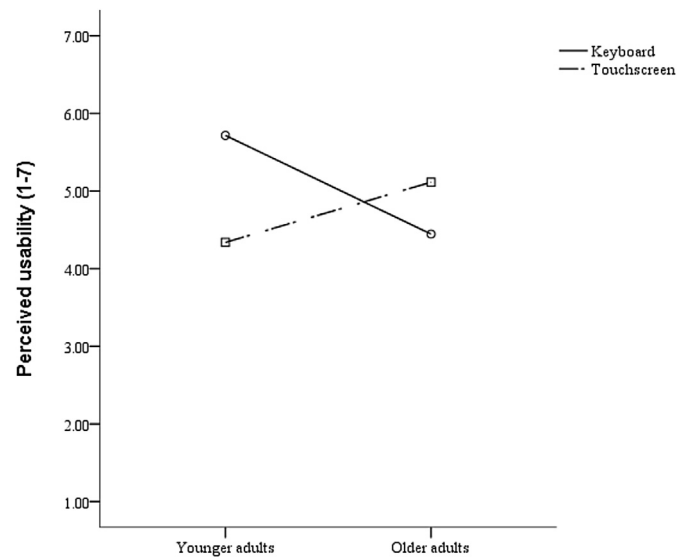


Fig. 3. Perceived usability as a function of age and input device.

( $F < 1$ ). A separate analysis of covariance was also carried out on the three subscales of the PSSUQ (system use, information quality, and interface quality). It revealed highly similar effects and confirmed the pattern found for the overall scale.

## 3.3. Affect

### 3.3.1. Negative affect

The data of change in participants' affective states (after product use compared to the initial baseline measure) indicated that older participants' negative affect increased ( $M = .16$ ,  $SD = .53$ ) whereas for younger participants showed a decrease ( $M = -.08$ ,  $SD = .32$ ). This difference between age groups was significant ( $F(1, 51) = 5.22$ ,  $p < .05$ ;  $\eta^2_{\text{partial}} = .09$ ). There was no significant interaction between age and input device ( $F(1, 51) = 1.50$ ,  $p > .05$ ). A comparison of input devices revealed that in the touchscreen condition participants' negative affect increased ( $M = .15$ ,  $SD = .48$ ) while negative affect decreased in the keyboard condition ( $M = -.07$ ,  $SD = .40$ ;  $F(1, 51) = 4.30$ ,  $p < .05$ ;  $\eta^2_{\text{partial}} = .08$ ).

### 3.3.2. Positive affect

In contrast to the findings for negative affect, the analysis of data did not reveal any age-related effects for positive affect, neither as a main effect ( $M_{\text{young}} = .12$ ,  $SD = .51$ ,  $M_{\text{old}} = .10$ ,  $SD = .52$ ,  $F < 1$ ) nor as an interaction ( $F(1, 51) = 2.23$ ,  $p > .05$ ). No difference between input devices was found ( $F < 1$ ).

## 3.4. Subjective workload

As indicated in Fig. 4, older participants rated workload on the NASA TLX higher when they worked with the keyboard than with the touchscreen. Conversely, younger participants experienced higher workload when using the touchscreen compared to the keyboard. This cross-over interaction between age and input device was significant ( $F(1, 51) = 13.28$ ,  $p < .001$ ;  $\eta^2_{\text{partial}} = .21$ ). There was no significant main effect of age ( $F < 1$ ). There was no significant difference between input devices ( $F < 1$ ). A separate analysis of the effects of the experimental manipulations for each subscale of the NASA TLX largely confirmed the pattern found for the overall scale (i.e. the scales frustration, effort, mental demand and performance

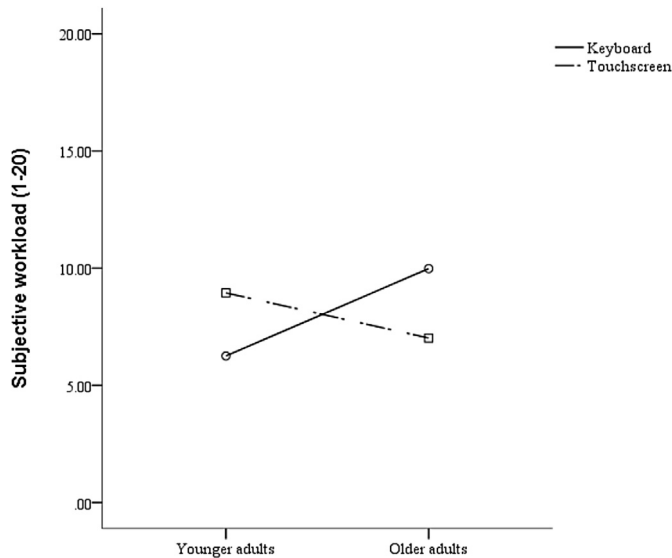


Fig. 4. Ratings of workload as a function of age and input device.

showed the same pattern whereas the subscales physical and temporal demand failed to be significant).

Of particular interest is the subscale 'performance' since it allows a comparison between self-rated performance and various measures of actual performance. Self-rated performance also showed a significant interaction between age and input device ( $F(1, 51) = 8.68, p < .01; \eta^2_{\text{partial}} = .15$ ), with older participants rating their performance higher when working with the touchscreen than with the keyboard whereas the inverse pattern was observed for younger participants (see Fig. 5).

### 3.5. Correlational analysis of data

In addition to the comparisons between experimental conditions using analyses of covariance, the size of correlations between variables may provide further insights. This point has notably been made by Hornbæk and Law (2007) arguing that usability studies

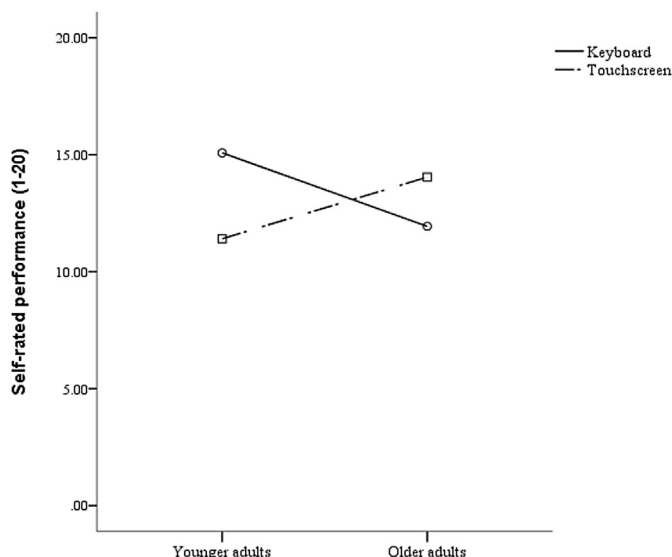


Fig. 5. Ratings of performance as a function of age and input device.

should report such correlations in order to facilitate interpretation and comparison of outcomes of usability evaluations.

Since there were considerable differences between younger and older adult users, correlation coefficients were computed for both groups (see Table 3). Most interesting are the relationships between the different measures of objective performance, self-assessed performance and perceived usability. For older adult users, there are higher correlations between self-assessed performance and the two accuracy measures (typing errors and task completion rate) than for younger users. In contrast, there is little difference between age groups with regard to the relationship between self-assessed performance and the speed-related measure. With regard to the relationship perceived usability and performance, for older adult users there is a significant correlation with task completion rate as an effectiveness measure but not with the three efficiency measures. For younger adult users, no such difference between effectiveness and efficiency was found for the correlation coefficients between perceived usability and the different performance measures. Finally, there were significant correlation coefficients between age and user errors as well as between age and task completion rate for the group of older adult users. This shows that within this group of older adult users, for very old adults decrements were even found for accuracy-related performance measures.

## 4. Discussion

The results showed overall age-related differences for speed-related performance, with better scores being observed for younger adults than for older ones while no such difference was found for accuracy. The study also revealed discrepancies between objective measures and user perceptions. Although older adult users took much longer to complete the typing task when using a touchscreen, their subjective perception of the device was rather positive. The results for perceived workload also showed some dissociation with performance measures, with workload being similar in its pattern to perceived usability.

The results showed no age-related effect for effectiveness measures such as task completion rate. The performance parameters that showed a decrement for older adult users were all efficiency measures (e.g. word processing speed) which refer to the time and cognitive resources needed for task completion and are different from effectiveness measures (Jordan, 1998). Speed of task completion is a measure of task efficiency, an aspect of performance that is impaired by increasing age. This progressive decline of response speed as a typical pattern associated with ageing has been found in numerous other studies (for an overview, see Matthews et al., 2000). However, these effects may not only be due to a generalised slowing of mental abilities, there may also be changes in strategy with increasing age. Older adults are more cautious and concentrate more on accuracy than speed, representing a substantial shift in the speed-accuracy trade-off towards accuracy (Welford, 1976). This finding was also supported in the present study by accuracy-related measures (e.g. error rate, task completion rate), which showed no main effect of age. Indeed, an additional analysis based on an index reflecting the speed-accuracy trade-off (i.e. 'number of words' divided by 'number of errors') confirmed the relative propensity of older adult users towards accuracy at the expense of speed compared to younger adults. This focus towards accuracy among older adult users was also reflected in the findings for user personality traits, which showed that older adults were more conscientious than younger adults.

An interesting finding was also the dissociation between usability ratings and objective performance, which emerged for older adult users during touchscreen operation. It showed that older

**Table 3**

Correlations between dependent variables and age (N = 30).

		Age	Self-rated performance	Perceived workload	Perceived usability	Writing speed	Number of user errors	Frequency of user input	Effectiveness	Positive affect	Negative affect
Age	Young	1	−.058	−.076	−.298	−.232	.230	.057	.000	−.246	.072
	Old	1	−.515 <sup>b</sup>	.192	−.139	−.290	.585 <sup>b</sup>	.116	−.367 <sup>a</sup>	−.083	.350
Self-rated performance	Young	1	1	−.377 <sup>a</sup>	.631 <sup>b</sup>	.234	−.284	.009	.247	.106	.082
	Old	1	.036	.176	.153	−.496 <sup>b</sup>	.059	.451 <sup>a</sup>	−.141	−.183	
Perceived workload	Young	1	1	1	−.415 <sup>b</sup>	−.251	−.052	.007	−.071	−.191	.055
	Old	1	1	1	−.556 <sup>b</sup>	.225	.091	−.014	−.317	−.555 <sup>b</sup>	.424 <sup>a</sup>
Perceived usability	Young	1	1	1	1	.253	−.084	−.135	.308	.395 <sup>a</sup>	.074
	Old	1	1	1	1	−.003	−.058	−.147	.476 <sup>b</sup>	.457 <sup>a</sup>	−.310
Writing speed	Young	1	1	1	1	1	.071	−.437 <sup>a</sup>	.101	.065	.143
	Old	1	1	1	1	1	−.320	−.442 <sup>a</sup>	.070	−.313	−.007
Number of user errors	Young	1	1	1	1	1	1	−.014	.112	−.103	.079
	Old	1	1	1	1	1	1	.006	−.376 <sup>a</sup>	.058	.464 <sup>b</sup>
Frequency of user input	Young	1	1	1	1	1	1	1	.034	−.118	.233
	Old	1	1	1	1	1	1	1	.087	−.122	.116
Effectiveness	Young	1	1	1	1	1	1	1	1	−.065	.059
	Old	1	1	1	1	1	1	1	1	.228	−.171
Positive affect	Young	1	1	1	1	1	1	1	1	1	−.154
	Old	1	1	1	1	1	1	1	1	1	−.237
Negative affect	Young	1	1	1	1	1	1	1	1	1	1
	Old	1	1	1	1	1	1	1	1	1	1

Values in bold indicate significant correlations.

<sup>a</sup> Correlation is significant at the .05 level (2-tailed).<sup>b</sup> Correlation is significant at the .01 level (2-tailed).

adults gave the touchscreen a better usability rating than it was justified from their performance. This finding appears to be contradictory at first sight but may be explained by low user expectations with regard to new technical devices such as touchscreens. The positive usability evaluation of the touchscreen among older participants despite their poor performance may be due to their increased anxiety towards new technology. As a consequence, they might have had a low level of perceived self-efficacy prior to using such a new device (please note that touchscreen-based phones had just become commercially available at the time of conducting the study). When older adult users realised that they were actually able to complete the tasks successfully with such a device, this unexpectedly positive experience might have caused them to increase their usability ratings. In addition, the judgement of older adult users might have also taken into account the potential usability of the device rather than the current usability because they expected that with increasing practice they would become more familiar with the new input device.

The argument that effectiveness may be more important for older adult users than efficiency was also corroborated by the analysis of correlations between perceived usability and performance. For older adults, there was a significant correlation between perceived usability and effectiveness but not between perceived usability and efficiency. For younger adult users, no such difference between correlation coefficients was found. This points to the importance of considering differences between age groups and type of performance measures (e.g. effectiveness vs. efficiency) when comparing results across studies. Since previous research generally provided low correlations between user performance and user satisfaction measures such as perceived usability (e.g. Frøkjær et al., 2000; Hornbæk and Law, 2007; Nielsen and Levy, 1994), it raises the question of whether differences in user age and type of performance measures may have contributed to the rather small correlation coefficients found (e.g. ranging from .164 to .247 in a recent meta-analysis by Hornbæk and Law, 2007). The role of these moderating variables needs to be examined in future meta-analyses of this kind.

The users' perception of their own performance also provides an interesting pattern. For older adult users, there are higher

correlations between self-assessed performance and accuracy-related performance measures than for younger users, suggesting that older adults base their own performance assessment on accuracy in task completion rather than speed. In contrast, there is little difference between age groups with regard to the strength of relationship between self-assessed performance and speed-related performance. This reiterates the argument that with increasing age, accuracy in task completion gains in importance over speed-related aspects. While this supports other research findings on age-related changes in speed-accuracy trade-off (Rabbitt, 1979; Salthouse, 1979; Smith and Brewer, 1995; see also Forstmann et al., 2011), it also suggests a need for usability research to make a distinction between speed and accuracy in performance. This distinction would extend the methodological framework by a further dimension and provide an important addition to the distinction between effectiveness and efficiency.

The pattern found for perceived workload matches the inverse pattern of perceived usability, that is, when users considered the usability of a device to be high, the workload associated with operating the device was low, and vice versa. This was confirmed by a significant correlation between the two measures, which was found for both age groups and was of similar magnitude. This finding corresponds to results from previous studies which showed substantial correlations between workload and satisfaction measures such as perceived usability (e.g. meta-analysis of Hornbæk and Law, 2007). The substantial association between workload and perceived usability may have also been increased by some similarity in the way the two concepts are defined and measured (Hornbæk and Law, 2007). For example, the definition of user satisfaction in the ISO standards includes aspects of workload, such as "overloading or underloading of the user's cognitive or physical workload" (ISO 1998, p. 5). Overall the association seems plausible since poorer usability is expected to lead to higher workload. More surprising may be the dissociation between workload and performance. Such dissociation between performance and subjective measures of workload have been previously observed and have been subject to extensive discussions in the human factors literature (e.g. Brookhuis and de Waard, 2002; Hancock et al., 1995; Vidulich and Wickens, 1986; Yeh and Wickens, 1988). Overall, our



findings suggest that task performance and perceived workload are distinct constructs and therefore need to be measured separately in usability evaluations (see also Schmutz et al., 2009; Hornbæk and Law, 2007).

While user age had an effect on several usability measures (i.e. performance, perceived usability, and workload), it is interesting to note that fewer age-related effects were observed for affect. Positive affect was not influenced by participants' age at all. However, older adult users reported an increase in negative affect after task completion while younger users experienced a slight decrease. This indicates that for older adult users, the participation in a usability test was a displeasing experience. The evaluative context of the usability test situation might have been experienced differently by older adult users compared to younger users. The testing situation might have been more stressful for older participants because they might have felt as if their technical skills and abilities had been evaluated rather than the technical device. Although it was explicitly emphasised in the test instructions that the aim of the usability test was to evaluate the smartphone rather than the abilities of the test participant, several older test users reported in the post-experimental debriefing session that they felt stressed because they had to interact with modern and unfamiliar technology. While this phenomenon was described before (e.g. Schrier, 1992), the present findings suggested that older participants might be more strongly affected and particular emphasis should be placed on the test instructions to reduce stress among older adult users. Older test participants need to be reassured that it is not their fault if they face problems during task completion. Instead, reporting usability problems is helpful, if not essential, for improving the device.

The measurement of relevant covariates was beneficial for our study since it allowed us to carry out more powerful data analyses (i.e. analysis of covariance instead of analysis of variance). While using analysis of covariance resulted in stronger effects than we would have obtained with a simple analysis of variance, it needs to be noted that the associations between the covariates and the dependent variables were generally small in size, with only one covariate showing a significant relationship with a dependent variable. This demonstrates that the difference in performance between older and younger users cannot simply be attributed to age-related differences in users' attitudes towards technology (e.g. computer anxiety) or to the behavioural consequence of such attitudes (e.g. frequency of mobile use or experience on mobile use). Since we controlled for the influence of these attitude-related factors, it suggests that it is age-related differences per se that need to be paid attention to in usability testing. For future research, it would be advisable to control for age-related factors such as eyesight, manual dexterity or memory performance to obtain a more detailed picture of the underlying factors leading to age-related effects in usability evaluation.

The study has several limitations. First, type of task and the effectiveness/efficiency distinction are confounded. This might have further increased the differences between effectiveness and efficiency measures. Second, the study participants may not be representative of the broader population. It cannot be excluded that the older age group had lower cognitive abilities than the student sample representing the younger age group because we do not know how many of the older participants attended university.

The findings of this study have two main implications. First, when interpreting the effects of age in usability testing, it needs to be taken into account that the present study adopted a conservative approach by not using a group of much older adult users. With an average age of only 58 yrs in the older adult group, we could already demonstrate differences between age groups. We expect that such age-related differences would have been even larger if

much older adult users had been examined. Evidence for this stems from the significant correlation coefficients between age and some performance indicators that were only found for older adults but not for younger ones. Second, some conceptual and methodological frameworks used in usability research have made a distinction between efficiency and effectiveness measures. The findings of the present study reiterate the importance of such a distinction since it seems to be insufficient to employ a one-dimensional concept of user performance to describe how well users operate a device. However, in addition to the distinction between efficiency and effectiveness, it may be necessary in usability research to distinguish between speed and accuracy in performance, too. This further distinction should be added to the theoretical framework, especially when user age is a variable of concern.

## Acknowledgements

We are very grateful to Daniel Imwinkelried and Tamara Ackermann for their help in collecting the data.

## References

- Anthony, L.M., Clarke, M.C., Anderson, S.J., 2000. Technophobia and personality subtypes in a sample of South African university students. *Comput. Hum. Behav.* 16 (1), 31–44. [http://dx.doi.org/10.1016/S0747-5632\(99\)00050-3](http://dx.doi.org/10.1016/S0747-5632(99)00050-3).
- AgeLight, 2001. Interface Design Guidelines for Users of All Ages. Available at: [www.agelight.com/webdocs/designguide.pdf](http://www.agelight.com/webdocs/designguide.pdf).
- Armbrüster, C., Sutter, C., Ziefle, M., 2007. Notebook input devices put to the age test: the usability of trackpoint and touchpad for middle-aged adults. *Ergonomics* 50 (3), 426–445.
- Bennett, S., Maton, K., Kervin, L., 2008. The “digital natives” debate: a critical review of the evidence. *Br. J. Educ. Technol.* 39 (5), 775–786. <http://dx.doi.org/10.1111/j.1467-8535.2007.00793.x>.
- Bosman, E., 1993. Age-related differences in the motoric aspects of transcription typing skill. *Psychol. Aging* 8, 87–102.
- Brookhuis, K.A., de Waard, D., 2002. On the assessment of (mental) workload and other subjective qualifications. *Ergonomics* 45 (14), 1026–1030.
- Bruder, C., Blessing, L., Wandke, H., 2014. Adaptive training interfaces for less-experienced, elderly users of electronic devices. *Behav. Inf. Technol.* 33 (1), 4–15.
- Brunsmann-Johnson, C., Narayanan, S., Shebilske, W., Alakke, G., Narakesari, S., 2015. Modeling web-based Information Seeking by Users who are Blind. *Disability and Rehabilitation: Assistive Technology*. Retrieved from [http://www.tandfonline.com/doi/abs/10.3109/17483107.2010.549897#vcm81\\_mNhQJ](http://www.tandfonline.com/doi/abs/10.3109/17483107.2010.549897#vcm81_mNhQJ).
- Caprani, N., O'Conner, N., Gurrin, C., 2012. Touch screens for the older user. In: Cheein, F.A.A. (Ed.), *Assistive Technologies*. Intech, pp. 96–118.
- Chadwick-Dias, A., McNulty, M., Tullis, T., 2003. Web usability and age: how design changes can improve performance. In: *Proceedings of the 2003 Conference on Universal Usability – CUU '03*. ACM Press, New York, New York, USA, pp. 30–37.
- Chua, S.L., Chen, D.-T., Wong, A.F.L., 1999. Computer anxiety and its correlates: a meta-analysis. *Comput. Hum. Behav.* 15 (5), 609–623. [http://doi.org/10.1016/S0747-5632\(99\)00039-4](http://doi.org/10.1016/S0747-5632(99)00039-4).
- Chung, M.K., Kim, D., Na, S., Lee, D., 2010. Usability evaluation of numeric entry tasks on keypad type and age. *Int. J. Ind. Ergonomics* 40 (1), 97–105.
- Colle, H.A., Hiszem, K.J., 2004. Standing at a kiosk: effects of key size and spacing on touch screen numeric keypad performance and user preference. *Ergonomics* 47 (13), 1406–1423.
- Costa, P.T., McCrae, R.R., 1992. Normal personality assessment in clinical practice: the NEO personality inventory. *Psychol. Assess.* 4 (1), 5–13. <http://doi.org/10.1037/1040-3590.4.1.5>.
- Czaja, S.J., Lee, C.C., 2007. The impact of aging on access to technology. *Univers. Access Inf. Soc.* 5, 341–349.
- Dyck, J.L., Smither, J.A., 1994. Age differences on computer anxiety: the role of computer experience, gender and education. *J. Educ. Comput. Res.* 10 (3), 239–248.
- Elias, S.M., Smith, W.L., Barney, C.E., 2012. Age as a moderator of attitude towards technology in the workplace: work motivation and overall job satisfaction. *Behav. Information Technol.* 31 (5), 453–467.
- Fezzani, K., Albinet, C., Thon, B., Marquie, J., 2010. The effect of motor difficulty on the acquisition of a computer task: a comparison between young and older adults. *Behav. Inf. Technol.* 29 (2), 115–124.
- Forstmann, B.U., Tittgemeyer, M., Wagenmakers, E.-J., Derrfuss, J., Imperati, D., Brown, S., 2011. The speed-accuracy tradeoff in the elderly brain: a structural model-based approach. *J. Neurosci.* 31 (47), 17242–17249.
- Frøkjær, E., Hertzum, M., Hornbæk, K., 2000. Measuring usability: are effectiveness, efficiency, and satisfaction really correlated? In: *CHI2000*, pp. 345–352.

- Garfein, A.J., Schaie, K.W., Willis, S.L., 1988. Microcomputer proficiency in later-middle-aged and older adults: teaching old dogs new tricks. *Soc. Behav.* 3 (2), 131–148.
- Hancock, P.A., Williams, G., Manning, C.M., 1995. Influence of task demand characteristics on workload and performance. *Int. J. Aviat. Psychol.* 5 (1), 63–86.
- Hanson, V.L., 2009. Age and web access. In: *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibility (W4A) – W4A '09*. ACM Press, New York, New York, USA, p. 7. <http://dx.doi.org/10.1145/1535654.1535658>.
- Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: H., P.A., Meshkati, N. (Eds.), *Advances in Psychology*, 52. North-Holland, pp. 139–183. Retrieved from: <http://www.sciencedirect.com/science/article/pii/S0166411508623869>.
- Heinssen, R.K., Glass, C.R., Knight, L.A., 1987. Assessing computer anxiety: development and validation of the computer anxiety rating scale. *Comput. Hum. Behav.* 3 (1), 49–59.
- Hertzum, M., Hornbæk, K., 2010. How age affects pointing with mouse and touchpad: a comparison of young, adult, and elderly users. *Int. J. Human-Computer Interact.* 26 (7), 703–734.
- Hornbæk, K., Law, E.L.-C., 2007. Meta-analysis of correlations among usability measures. In: *Proceedings of the SIGCHI Conference on Human Factors in Comput. Systems – CHI '07*, San Jose, USA. ACM, New York, pp. 617–626.
- ISO 9241-11, 1998. *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)-Part 11, Guidance on Usability*, ISO.
- Jastrzebski, T., Charness, N., Holley, P., Feddon, J., 2005. Input devices for web browsing: age and hand effects. *Univ. Access Inf. Soc.* 4 (1), 39–45. <http://doi.org/10.1007/s10209-003-0083-5>.
- Jacko, J.A., Barrett, A.B., Marrnet, G.J., Chu, J.Y.M., Bausch, H.S., Scott, I.U., Rosa, R.H., 2000. Low vision: the role of visual acuity in the efficiency of cursor movement. In: *Proceedings of the 4th International ACM SIGAPH Conference on Assistive Technologies (ASSETS'00)*, pp. 1–8.
- Jin, Z.X., Plocher, T., Kiff, L., 2007. Touch screen user interfaces for older adults: button size and spacing. In: *Proceedings of Universal Access in Human-Computer Interaction: Coping with Diversity (UAHCI '07)*, pp. 933–941. Beijing, China, July 22–27.
- Jordan, P.W., 1998. *An Introduction to Usability*. Taylor & Francis, London.
- Keates, S., Trewin, S., 2005. Effect of age and Parkinson's disease on cursor positioning using a mouse. In: *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, MD, USA, October 09–12, 2005). Assets '05. ACM, New York, NY, pp. 68–75.
- Kolikant, Y.B.-D., 2010. Digital natives, better learners? Students' beliefs about how the Internet influenced their ability to learn. *Comput. Hum. Behav.* 26 (6), 1384–1391. <http://doi.org/10.1016/j.chb.2010.04.012>.
- Krohne, H., Egloff, B., Kohlmann, C., Tausch, A., 1996. Untersuchungen mit einer deutschen Version der "Positive and Negative Affect Schedule" (PANAS). *Diagnostica* 42 (2), 139–156.
- Kroemer, K.H.E., 2006. *Extra-ordinary Ergonomics*. CRC Press, Boca Raton.
- Kroemer, K.H.E., Kroemer, H., Kroemer-Elbert, K., 2001. *How to Design for Ease and Efficiency*. Prentice-Hall, New Jersey.
- Lee, S., Zhai, S., 2009. The performance of touch screen soft buttons. In: *Proceedings of Human Factors in Computing Systems (CHI '09)*, ISBN 978-1-60558-246-7, pp. 309–318. Boston, USA, April 4–9.
- MacKenzie, I.S., Soukoreff, R.W., 2002. Text entry for mobile computing: models and methods, theory and practice. *Human-Computer Interact.* 17 (2), 147–198. [http://dx.doi.org/10.1207/S15327051HCI172&3\\_2](http://dx.doi.org/10.1207/S15327051HCI172&3_2).
- Matthews, G., Davies, D.R., Westerman, S.J., Stammers, R.B., 2000. *Human Performance, Cognition, Stress, and Individual Differences*. Psychology Press, East Sussex.
- Nahm, E.S., Preece, J., Resnick, B., Mills, M.E., 2004. Usability of health web sites for older adults: a preliminary study. *Comput. Inf. Nurs.* 22 (6), 326–334.
- Neyer, F., Felber, J., Gebhardt, C., 2012. Entwicklung und Validierung einer Kurzsкала zur Erfassung von Technikbereitschaft. *Diagnostica* 58 (2), 87–99.
- Nielsen, J., 1993. *Usability Engineering*. Academic Press, Boston.
- Nielsen, J., Levy, J., 1994. Measuring usability: preference vs. Performance. *Commun. ACM* 37 (4), 66–75.
- Ostendorf, F., Borkenau, P., 2008. NEO-Fünf-Faktoren-Inventar nach Costa und McCrae (NEO-FFI); Manual. Hogrefe, Göttingen.
- Pak, R., Czaja, S.J., Sharit, J., Rogers, W.A., Fisk, A.D., 2006. The role of spatial abilities and age in performance in an auditory computer navigation task. *Comput. Hum. Behav.* 24 (6), 3045–3051.
- Pak, R., Price, M.M., Thatcher, J., 2009. Age-sensitive design of online health information: comparative usability study. *J. Med. Internet Res.* 11 (4), e45. <http://dx.doi.org/10.2196/jmir.1220>.
- Petrie, H., Kamollimsakul, S., Power, C., 2013. Web accessibility for older adults: effects of line spacing and text justification on reading web pages. In: *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, pp. 1–2.
- Piqueras-Fiszman, B., Ares, G., Alcaide-Marzal, J., Diego-Más, J.A., 2011. Comparing older and younger users' perceptions of Mobile phones and watches using data questions and preference mapping on the design characteristics. *J. Sens. Stud.* 26 (1), 1–12.
- Plaisant, C., Sears, A., 1992. Touchscreen interfaces for alphanumeric data entry. *Proc. Hum. Factors Ergonomics Soc. Annu. Meet.* 36 (4), 293–297.
- Prenksy, M., 2001. Digital natives, digital immigrants. *Horizon* 9 (5), 1–6.
- Rabbitt, P., 1979. How old and young subjects monitor and control responses for accuracy and speed. *Br. J. Psychol.* 70 (2), 305–311.
- Rogers, W.A., Stronge, A.J., Fisk, A.D., 2005. Technology and aging. *Rev. Hum. Factors Ergonomics* 1 (1), 130–171.
- Rosen, L.D., Weil, M.M., 1992. *Measuring Technophobia Manual* (Unpublished), Available: [Irosen@dhvx20.csudh.edu](mailto:Irosen@dhvx20.csudh.edu).
- Rosen, L.D., Weil, M.M., 1995. Computer anxiety: a cross-cultural comparison of university students in ten countries. *Comput. Hum. Behav.* 11 (1), 45–64.
- Salthouse, T.A., 1979. Adult age and the speed-accuracy trade-off. *Ergonomics* 22 (7), 811–821.
- Salthouse, T.A., 1984. Effects of age and skill in typing. *J. Exp. Psychol. General* 113, 345–371.
- Sayers, H., 2004. Desktop virtual environments: a study of navigation and age. *Interact. Comput.* 16 (5), 939–956.
- Schrier, J.R., 1992. Reducing stress associated with participating in a usability test. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 36. Human Factors and Ergonomics Society, pp. 1210–1214.
- Schedlbauer, M., 2007. Effects of key size and spacing on the completion time and accuracy of input tasks on soft keypads using trackball and touch input. In: *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, Baltimore, MD. Human Factors and Ergonomics Society, Santa Monica, CA, USA, pp. 429–433.
- Scott, B., Conzola, V., 1997. Designing touch screen numeric keypads: effects of finger size, key size, and key spacing. In: *Proceedings of the Human Factors Society 41st Annual Meeting*, Albuquerque, New Mexico, USA, pp. 360–364.
- Smith, G.A., Brewer, N., 1995. Slowness and age: speed-accuracy mechanisms. *Psychol. Aging* 10 (2), 238–247.
- Schmutz, P., Heinz, S., Métrailler, Y., Opwis, K., 2009. Cognitive load in eCommerce applications—measurement and effects on user satisfaction. *Adv. Human-Computer Interact.* 2009, 1–9. Article ID 121494.
- Smith, M.W., Sharit, J., Czaja, S.J., 1999. Aging, motor control, and the performance of computer mouse tasks. *Hum. Factors* 41 (3), 389–396.
- Sonderegger, A., Sauer, J., Eichenberger, J., 2014. Expressive and classical aesthetics: two distinct concepts with highly similar effect patterns in user-artefact interaction. *Behav. Inf. Technol.* 33 (11), 1180–1191.
- Stelmach, G.E., Goggin, N.L., Garcia-Colera, A., 1987. A movement specification time with age. *Exp. Aging Res.* 13, 39–46.
- Taveira, A.D., Choi, S.D., 2009. Review study of computer input devices and older users. *Int. J. Human-Comput. Interact.* 25 (5), 455–474.
- van Schaik, P., Ling, J., 2009. The role of context in perceptions of the aesthetics of web pages over time. *Int. J. Human-Comput. Stud.* 67 (1), 79–89.
- Vidulich, M.A., Wickens, C.D., 1986. Causes of dissociation between subjective workload measures and performance: caveats for the use of subjective assessments. *Appl. Ergonomics* 17 (4), 291–296.
- Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Personal. Soc. Psychol.* 54 (6), 1063–1070.
- Webster, J., Martocchio, J.J., 1992. Microcomputer playfulness: development of a measure with workplace implications. *MIS Q.* 16 (2), 201–226.
- Welford, A.T., 1976. Motivation, capacity, learning and age. *Int. J. Aging* 7, 189–199.
- Wickens, C.D., Hollands, J.G., 2000. *Engineering Psychology and Human Performance*, third ed. Prentice Hall, Upper Saddle River, NJ.
- Worden, A., Walker, N., Bharat, K., Hudson, S., 1997. Making computers easier for older adults to use: area cursors and sticky icons. In: *Proceedings of CHI 97*. ACM, New York, pp. 266–271.
- Wright, P., Bartram, C., Rogers, N., Emslie, H., Evans, J., Wilson, B., Belt, S., 2000. Text entry on handheld computers by older users. *Ergonomics* 43 (6), 702–716.
- Yeh, Y.-Y., Wickens, C.D., 1988. Dissociation of performance and subjective measures of workload. *Hum. Factors* 30 (1), 111–120.
- Zaphiris, P., Kurniawan, S., Ghiawadwala, M., 2007. A systematic approach to the development of research-based web design guidelines for older people. *Univ. Access Inf. Soc.* 6 (1), 59–75.