Proceeding of the IEEE
International Conference on Information and Automation
Shenzhen, China June 2011

Assessing Age-Related Performance Decrements in User Interface Tasks

Xiaolei Zhou
School of Information
Capital University of Economics and Business
Huaxiang Zhangjialukou121,Beijing 100070 China
zhouxiaolei.cueb@gmail.com

Shengdong Zhao
Department of Computer Science
National University of Singapore
Singapore
zhaosd@comp.nus.edu.sg

Mark Chignell
Mechanical and Industrial Engineering
University of Toronto
Toronto, Canada
chignel@mie.utoronto.ca

Xiangshi Ren
The School of Information
Kochi University of Technology
Tosayamada- cho,Kami-shi, Kochi, Japan
ren.xiangshi@kochi-tech.ac.jp

Abstract **- As the computer and internet generations age, there is an increasing need to develop appropriate interfaces for the elderly that can accommodate age-related changes in manual dexterity, visual acuity, and cognitive abilities. Assessment of age effects is typically a necessary first step in designing age-appropriate interfaces, but assessment of age related effects may be complicated by a bias towards accuracy in the elderly or by other differences in how the tradeoff between speed and accuracy is handled by different people. In this paper, we attempt to investigate the effects of aging on performance difference in interacting with computer interfaces. An experiment was conducted to examine age related effects in a steering task. In order to assess the impact of a possible speed-accuracy tradeoff, performance was observed under three different instructional sets i.e., *accuracy (A)*, *neutral (N)*, and *speed (S)* when steering on a circular track. Experimental results showed that the elderly group performed significantly less accurately for all three instruction sets. The younger subjects were more influenced by instructions to perform faster, or with more accuracy. Cluster analysis of the empirical data individually for both the old and younger participants showed that variability among subjects was much greater in older users than younger users. Implications for user interface design for older users, and for the evaluation of age effects in HCI generally, are discussed.**

Index Terms - Human performance, speed-accuracy tradeoff, user group, steering task, aging effects.

I. Introduction

In the past, elderly users have tended to have low computer literacy and design of information technologies for the aged has received little emphasis. However, the situation is set to change as the baby boomer generation reaches retirement age, where large numbers of retiring people, particularly in developed countries, will both highly educated and experienced computer users. In spite of their greater technical proficiency, baby boomers and subsequent generations will still be subject to the physiological and psychological changes that occur with aging, including reductions in manual dexterity, visual acuity, hearing sensitivity and cognitive complexity, etc., which affect control of user interfaces in particular [4], [20], [22].

In order to meet the needs of aging computer users, novel interfaces are required for the elderly so that they can contribute to and function in a society that is increasingly dominated by information technology.

In this paper, we attempt to investigate the effects of aging by a bias towards speed or accuracy in the elderly or by other performance differences in how the tradeoff between speed and accuracy is handled by different people. An experiment was conducted to examine age related effects under three different instructional sets i.e., *accuracy (A)*, *neutral (N)*, and *speed (S)* when steering on a circular track. The reason that we chose the steering task was that this was a task that has been extensively studied in past research, and as [2] pointed out, straight and circular steering tunnels are two basic and representative steering tasks in HCI. Examples of steering task in HCI include navigation through a cascade menu, drawing, writing, or steering through a 3D space, etc. These are also tasks that require fine motor control and where age effects may be expected to occur. According to the experimental results, implications for user interface design for older users, and for the evaluation of age effects in HCI generally are discussed.

II. Related Work

The study of age effects in HCI is complicated by a tendency for the elderly to focus more on accuracy at the expense of speed [17]. This suggests that potential speed-accuracy tradeoffs [11], [21] need to be considered when age effects are examined in HCI.

The research reported in this paper focuses on assessing age effects on a common HCI task in visual interfaces: navigating, or steering the cursor through a 2-dimensional tunnel, which can be modeled using steering law [1]. Many studies have been performed in the past on the verification of the steering law for various input devices [2], scales [3], and parameters of steering motion [10]. However, the participants in previous studies were young adults, mostly involving

978-1-61284-4577-0270-9/11/$26.00 ©2011 IEEE

students from universities. The present study sought to address this deficiency in the literature by providing findings on the effect of aging on the steering task, while also serving as a case study of how to examine the impact of speed-accuracy tradeoffs in HCI evaluations.

Aging effects have been examined in other HCI tasks. For instance, it has been reported that older users position the cursor much more slowly than younger users and have great difficulty making targeted movements to small targets [19]. Novel interaction techniques have been proposed (e.g., use of proxy targets [9], area cursors and sticky icons [19]) to overcome this difficulty and improving the accessibility of user interface or web usability for older people. Other recommended changes to user interfaces involve changes to content aspects such as font type and size for enhanced legibility [5]. Moffatt et al. [12], [13], [14] conducted a controlled laboratory experiment to examine target acquisition difficulties across the lifespan (younger, pre-old, and old people) during two tasks: multi-dimension tapping and menu selection, and attempted to address these difficulties by using appropriate interactive techniques for older people. In addition, adaptive interface for older people or motor impaired person [6], [7], and systematic theory framework research for web design or user interface design [20] were also investigated. However, research investigating age related effects in steering tasks that also considers the possible impact of age-related differences in the speed accuracy tradeoff has yet to be carried out. The following experiment was designed to address that deficiency.

III. EXPERIMENT

A. *Speed and Accuracy*

In steering tasks, speed is typically represented by the time spent to accomplish a task, or movement time (*MT*). Accuracy may be measured as the standard deviation of sampled points in a trajectory made by a user, or by the number of points that are outside the specified area be steered within. We used both measurements for accuracy in this research. In this paper, the effect of three different instructional sets on performance in the circular steering tasks is contrasted for younger vs. older users. The following instructional sets were used: accuracy emphasis (*A*) where users were asked to focus on accuracy only; neutral (*N*) where users were asked to focus on both speed and accuracy; speed emphasis (*S*) where users were asked to focus on speed only.

B. *Participants*

12 younger participants (3 females and 9 males; 20 to 27 years old, mean age 21.3; all right-handed), and 12 older participants (4 females and 8 males, 61 to 72 years old, mean age 65.8; all right-handed) were recruited to participate in this experiment. The younger people were students, and the older people were educated (two graduated from college, and others from middle school) and from a "Older People's Center". The older people investigated in this paper are healthy older people, who don't appear disabled, but their functionality, needs and wants are different from those they had when they were younger [8].

C. *Apparatus*

The experiment was conducted on an IBM ThinkPad X41 Tablet PC running Microsoft Windows XP tablet edition, using a stylus as input. The screen size was 12.1 inches (1024×768 pixels resolution). The experimental software was developed in Java 6.0.

D. *Task*

Fig. 1 illustrates the experimental task for both user groups. Users were asked to perform steering tasks in a circular tunnel from the start line to the end line. *R* was the radius of the circular tunnel and *W* was its width. The movement amplitude *A* was equal to the circle circumference $2\pi R$.

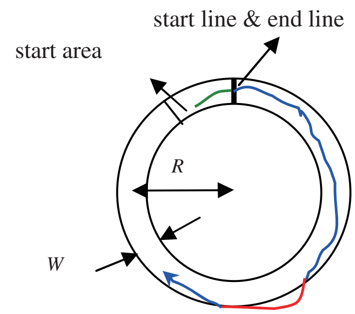


Fig. 1 Circle tunnel steering.

E. *Design*

The experiment used a mixed design. User group was the only between-subject factor with two levels (young vs. old). The three within-subject factors were: Amplitude (300, 600, 800 pixels), Width (20, 30, 40, 50, and 60 pixels), and Instructional set (*A*, *N*, and *S*, as defined above). The direction of the circular steering task was clockwise. Similar designs were used previously [2], [3].

Each subject repeated the experiment three times with the different instructional sets, i.e., *A*, *N*, and *S*. Instructions corresponding to each instructional set were given by the experimenter before each experiment.

The order of the three instructional sets, *A*, *N*, *S*, was balanced using a Latin square. The order of the 15 amplitude and width combinations was presented in random order to the participants within each instructional set. Each subject performed 3 strokes for each Amplitude/Width combination within each instructional set of the circular steering tasks. Subjects completed the experiment in one session of about 30 minutes. In summary, the experiment design involved 24 subjects x 3 (tunnel amplitudes) × 5 (tunnel widths) × 3 (strokes) × 3 (instructional sets) = 3240 for the total number of trials.

F. *Procedure*

Warm-up trials were performed before each instructional set was used for the first time by each subject, leading to three sets of warm-up trials.

For each trial, subjects were instructed to trace a circular path from the start line to the end line in one clockwise motion. The trajectory of the stylus' movement was displayed in real time as feedback to users. The color of the trajectory was green if the stylus was inside the start area and had not entered the tunnel, blue if the movement of the stylus had crossed the start line and was inside the tunnel, and red if the stylus moved outside the path boundaries. Users completed the entire circle by passing the stylus across the end line from left to right, after which the tunnel disappeared. During this task, the tip of the stylus was required to stay in contact with the touchscreen. The same trial was repeated if the pen lifted off the touchscreen surface during this process.

G. *Measurements*

The position of the stylus was sampled every 10 milliseconds for each trial. The movement time (*MT*) required for users to trace the entire circle, from beginning to end, along with standard deviation (*SD*) of the distances from the center of the circular tunnel to the sampled points in pixel units, and the out of path movement (*OPM*), measured as the percentage of sample points outside the tunnel border were measured.

IV. Results

Since the focus of this study was on the speed-accuracy tradeoff, the effects of amplitude and width are not reported below. The main effect of instruction set (as expected) was significant on movement time, *SD*, and *OPM* ($p<0.001$ in all the three cases), demonstrating that the subjects changed their performance in response to the instructions provided to them. In the following subsections the results of more detailed analysis will be reported.

A. *Movement Time (MT)*

ANOVA analysis showed a significant interaction effect of age group and instruction set on movement times ($F(1.25, 44)=4.43$, $p<0.05$)[1], but no main effect of age group ($p>0.05$). As can be seen in Fig. 2, while movement times vary with instructional set for both age groups, the change is more dramatic for the younger group.

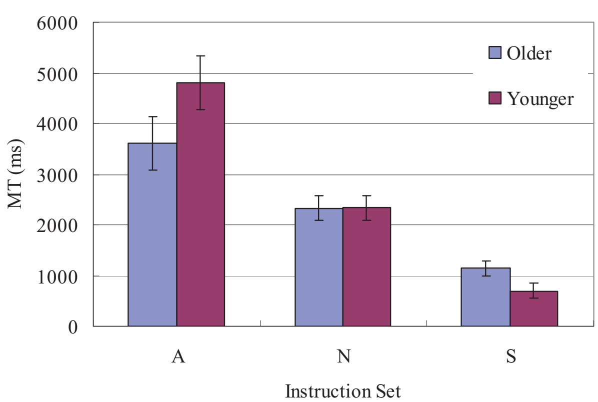


Fig. 2. Mean MT in instruction set A, N, and S for both older and younger user groups (with standard error bars).

B. *Standard Deviation* (*SD*)

[1] Note that non-integer values for degrees of freedom indicate that the sphericity assumption was violated and that the Huyn-Feldt adjustment to the effect degrees of freedom was used, as is recommended practice in such cases.

ANOVA analysis showed no significant interaction between instruction set and age group on *SD* ($F<1$), but there was a significant main effect of Age group on *SD* ($F(1,22)=4.94$, $p<0.05$). On average the younger group was more accurate for all three of the instructional sets (Fig.3).

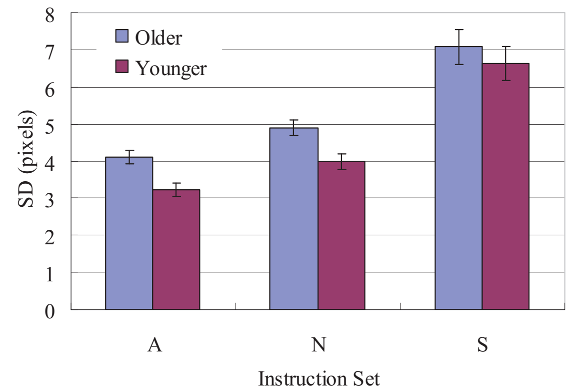


Fig. 3. Mean SD in instruction set A, N, and S for both older and younger user groups (with standard error bars).

C. *Out of Path Movement* (*OPM*)

ANOVA analysis showed no significant effect on *OPM* is found for user group ($p >0.05$). Mean *OPMs* for older and younger groups were 3.3% and 2.0% respectively. No significant interaction effect between instructional set and user group was observed ($p >0.05$). These showed that both older and younger users' behaviours could equivalently follow the requirement (tunnel width) the task set. A significant interaction between instructional set and *W* was observed on *OPM* ($F(8, 136) = 49.79$, $p <0.01$) (Fig. 4). The effect of *W* on *OPM* is larger in condition S (when speed is the only concern) than that in condition *A* and *N*.

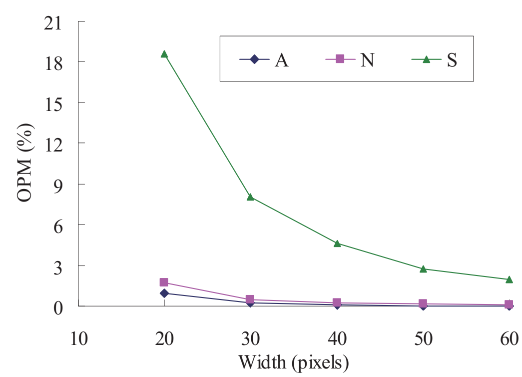


Fig. 4. *OPM* vs. *W*(Width) for instructional sets A, N and S – with older and younger user groups averaged

D. *Cluster Analysis of Individual Differences*

Individual differences are known to increase with age [16]. Here we investigated the individual difference for both older and younger people by the cluster analysis.

We found that the older group were overall less accurate in terms of the *SD* measure, and were also less responsive to instructional set. In order to better understand the nature of these effects, cluster analysis was carried out to determine the extent to which individual differences occurred within the two age groups, and how strongly such differences might have affected the results obtained.

Since there was both a strong age effect and a strong instruction set effect on accuracy, cluster analysis was carried out to examine how individual differences may have mediated the observed relationships between age, instruction set, and accuracy.

The average accuracies in terms of pixel *SD* were calculated across all combinations of the 24 participants and the three instructional sets. The accuracies were then converted into z-score units with the normalization being carried out for the data pooled across all instructional sets and participants. K-means analysis clustering was then carried out, with two, three, and four cluster solutions being examined. The four cluster solution was chosen for further study based on its interpretability.

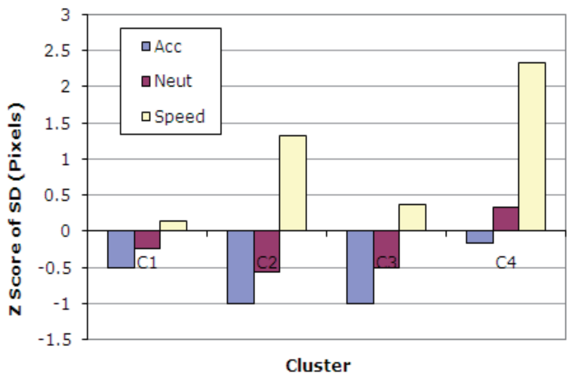


Fig. 5. Cluster centre values for z-score of SD across the four clusters (clusters one through four in left to right order).

Accuracy values across the three instruction sets are shown for each of the clusters in Fig. 5. Members of cluster 1 consisted of three older people and one younger people. They displayed very little change in accuracy in response to instruction set (with accuracy remaining within one half standard deviation of the experiment average across all three instruction sets). Cluster 4 contained only older (five) people. The people in this cluster were generally less accurate, and particularly so in the speeded instruction set (with average *SD* in this case being over two standard deviations above the average for the experiment). Eight of the nine people in clusters 1 and 4 were older people. In contrast, six of the eight people in cluster 2 were younger, and five of the seven people in cluster 3 were younger. Clusters 2 and 3 both had high accuracy when instructed to be accurate and a fairly good level of accuracy in the neutral condition, but they differed in that cluster 3 subjects retained an accuracy level close to the experimental average, whereas *SD* increased to almost 1.5 standard deviation units above the average for cluster 2 subjects.

The clusters described above were identified based on the accuracy data, could they also predict differences in movement time across the different instruction sets?

This question was addressed by turning the four clusters into four levels of a corresponding "Cluster" pseudo factor and again running mixed ANOVA, but this time using the Cluster factor in place of the age group factor. There was a significant interaction between cluster and instruction set on movement time ($F(4.78, 40)=4.135$, $p<0.01$).

It can be seen in Fig. 6 that the an increasing instruction set for accuracy slowed movement times down in all four clusters, but that this effect was more pronounced for the "younger" clusters, i.e., clusters 2 and 3. By comparing Fig. 5 and 6 it can also be seen that subjects in cluster 2 responded more aggressively to the instructions, slowing down more in the accurate condition (but with no benefit to accuracy as compared with that obtained in cluster 3) and speeding up more in the speeded condition (but at a relatively high cost to accuracy). Similarly it can be seen for the "older clusters (1 and 4)" that the fast movement times achieved in cluster 4 came at the expense of a substantial loss of accuracy.

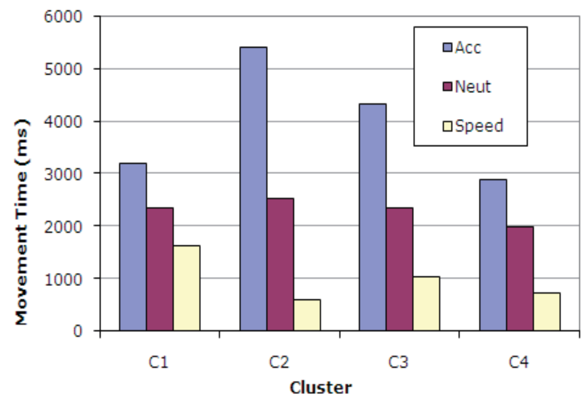


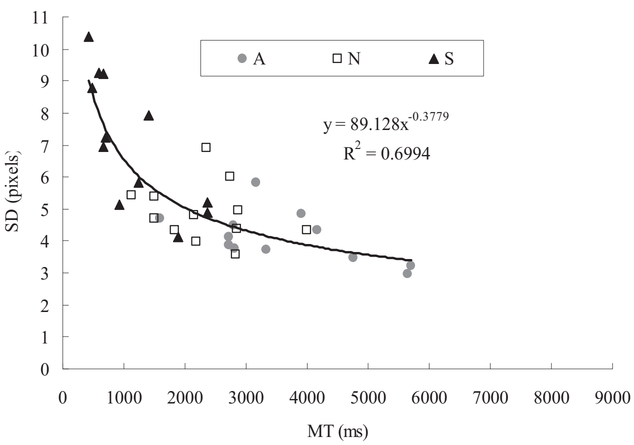
Fig. 6. The effect of instruction set on average movement time across the four clusters.

E. Discussion

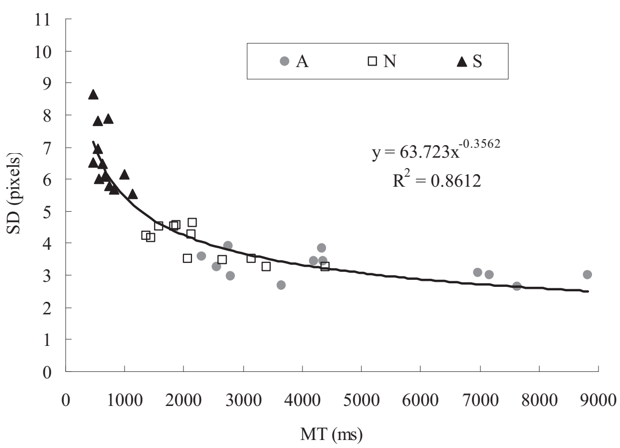
In general the more accurately a task is performed, the longer it takes and vice versa, with a characteristic s-shaped curve often being observed [15]. The relationship between *SD* and *MT* was fitted as a power function separately for both older and younger groups. As indicated by the power curve goodness of fits shown in Fig. 7, there was a more consistent speed-accuracy relationship for the younger subjects (R^2 = 0.861 vs. 0.699 for the older subjects). The effect of instruction set can be seen vividly in Fig. 7, where the triangles (speeded set) tend to be above and to the left of the squares (neutral) which in turn tend to be above and to the left of the circles (accurate set). It can also be seen that the speed-accuracy tradeoff is more strongly defined for the younger subjects.

Individual differences tend to increase with age [16]. We found that the older group were overall less accurate in terms of the *SD* measure, and were also less responsive to instructional set. When analyzing the data individually for both the old and younger participants, we found that variability among subjects was much greater in older users than younger users. Using cluster analysis based on *SD* over the three different instruction sets, we found that four of the elderly group showed a pattern of performance that was characteristic of the younger group, whilst only one of the younger group exhibited "older" performance.

The overall results showed that aging significantly diminishes performance on the steering task in terms of accuracy but not movement time. However, the speed-accuracy tradeoff induced by differences in instruction set was much stronger for the younger subjects.



(a) Older user group



(b)Younger user group

Fig. 7. Changes in speed (*MT*) and accuracy (*SD*) - for (a) older user group and (b) younger user group. Points corresponding to different instruction sets are labeled with differed shapes.

V. Implications for Interface Design

While older users tend to have accuracy bias, in this study older users produced larger *SD* (greater deviation from the centre of the tunnel) than did younger users. In the user interface design for older users, one obvious strategy for dealing with this age effect would be to use larger tunnel size or target size.

It is generally more difficult for older users to perform a trajectory tracing task. In our experiment we observed that slow but precise movement of the stylus required firm but stable grasp of the stylus, which was difficult for older users' due to hand tremor. Fast movement of the stylus required good hand dexterity, which was also difficult for older users.

The older subjects showed less variability in movement time in spite of the instruction set. When older users did attempt to speed up a lot, there was a disastrous loss of accuracy. The inability of at least some of the older subjects to perform both fast and accurately may be partly due to the effects of reduced hand dexterity.

Recently, there has been considerable interest in gesture-based interfaces for pen-based computing. Many of these interfaces designed for efficient performance, especially by expert users. However, our results suggest that the utility of an interaction technique may be influenced by age. Many innovative techniques are currently tested only with younger users and age effects are ignored. Since gestural interfaces often rely on steering tasks of one sort or another, the differences found in our results indicate that some of the advantages exhibited by these interfaces in a younger user group may not apply to older users. Empirical studies are needed to reassess the effectiveness of these interfaces for an older population. Some of these techniques may need to be redesigned or enhanced by special interactive techniques such as force feedback, and area cursors to make steering tasks more manageable for older users.

Observations made in our experiment also point out issues that need to be considered while conducting age related evaluation. Our experiment shows that relying on chronological age in studying age effects may be misleading. A distinction needs to be made between elderly users who perform like younger people and more typical elderly users who show the effects of age in their performance.

In addition, the effects of implied or explicit instruction sets on performance need to be carefully controlled in studies involving age effects. In general younger people may show a greater effect of instruction set on their performance. Thus depending on the instruction set younger users may appear to speed up or slow down (or become more or less accurate) relative to older subjects.

VI. Conclusion

As Salthouse has pointed out [18], as people age, their cognitive, perceptual, and motor abilities decline, with negative effects on their ability to perform many tasks. However, as the present results demonstrate, aging effects need to be evaluated carefully. While it seems clear that aging has a negative impact on the steering task (and likely on many other HCI tasks as well), the situation is complicated both by speed-accuracy tradeoff effects and also by the heterogeneous nature of elderly populations. In any tasks there are likely to be some elderly people that can perform like younger people, and for those people an interface design specifically for the elderly might be annoying and inefficient. With respect to speed versus accuracy, researchers need to be careful about how much they stress speed or accuracy in performing experimental tasks. It is possible that relatively subtle changes in instructional set may lead to radically different apparent age effects. Where possible, it may be useful, as carried out in the present paper, to examine explicitly speed-accuracy tradeoffs and the effect of individual differences on performance. It would seem that analysis of aging effects in HCI may require more careful and detailed analysis of the experimental data and the subtle patterns and effects obtained therein.

This paper shows how interpretation of age effects in HCI is complicated by the different way in which speed is traded off against accuracy in older people, the tendency of older people to stress accuracy at the expense of speed, and to have greater individual differences in performance, enriching our knowledge of how aging effects are likely to impact evaluations in HCI. In the future, we would like to perform further analysis to systematically investigate the individual effects of visual acuity, manual dexterity, and possibly other factors (such as hearing and cognitive abilities) on the steering task and other common user interface tasks.

ACKNOWLEDGMENT

We thank all the participants for their support on this research.

REFERENCES

[1] J. Accot,, and S. Zhai, *Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks*. In: Proc. CHI1997, pp. 295--302, ACM Press (1997).

[2] J. Accot,, and S. Zhai, *Performance Evaluation of Input Devices in Trajectory-based Tasks: An Application of the Steering law*. In: Proc. CHI1999, pp. 466--472, ACM Press (1999).

[3] J. Accot,, and S. Zhai, *Scale Effects in Steering Law Tasks*. In Proc. CHI2001, pp. 1--8, ACM Press (2001).

[4] S. Becker, *A study of web usability for older adults seeking online health resources*. ACM Transactions on Computer-Human Interaction, vol. 11, no. 4, pp. 387--406, (2004).

[5] M. Bernard, C. Liao, and M. Mills, *The Effects of Font Type and Size on the Legibility and Reading Time of Online Text by Older Adults*. In: Proc. CHI2001, pp. 175--176, ACM Press (2001).

[6] K. Gajos, J. Wobbrock, and D. Weld, *Automatically generating user interfaces adapted to users' motor and vision capabilities*. In: Proc. UIST2007, pp. 231--240, ACM Press (2007).

[7] K. Gajos, J. Wobbrock, and D. Weld, *Improving the performance of motor-impaired users with automatically-generated, ability-based interfaces*. In: Proc. CHI2008, pp. 1257--1266, ACM Press (2008).

[8] P. Gregor, A. Newell, and M. Zajicek, *Designing for Dynamic Diversity-interfaces for older people*. In: Proc. ASSETS2002, pp. 151--156, ACM Press (2002).

[9] F. Hwang, H. Batson, and N. Williams, *Bring the Target to the Cursor: Proxy Targets for Older Adults*. In: Proc. CHI2008, pp. 2775--2780, ACM Press (2008).

[10] S. Kulikov, I. Mackenzie, and W. Stuerzlinger, *Measuring the Effective Parameters of Steering Motions*. In: Proc. CHI2005, pp. 1569--1572, ACM Press (2005).

[11] I. Mackenzie, and P. Isokoski, *Fitts' Throughput and the Speed-Accuracy Tradeoff.* In: Proc. CHI2008, pp. 1633--1636, ACM Press (2008).

[12] K. Moffatt, and J. McGrenere, *Slipping and Drifting: Using older users to uncover pen-based target acquisition difficulties*. In: Proc. ASSETS'07, pp. 11--18, ACM Press (2007).

[13] K. Moffatt, S. Yuen, and J. McGrenere, *Hover or Tap? Supporting pen-based menu navigation for older adults*. In: Proc. ASSETS'08, pp. 51--58, ACM Press (2008).

[14] K. Moffatt, *Increasing the accessibility of pen-based technology: An investigation of age-related target acquisition difficulties*. In: Proc. CHI2008, pp. 2625--2628, ACM Press (2008).

[15] R.G. Pachella, *The interpretation of reaction time in information processing research*. In: B.H. Kantowitz (ed.), Human information processing: tutorials in performance and cognition. New York: Wiley (1974)

[16] P.M.A. Rabbitt, and C. Lowe, *Patterns of Cognitive Aging. Psychological Research*, 63, pp. 308--316, (2000).

[17] T. A. Salthouse, *Adult Age and the Speed-Accuracy Trade-off. Ergonomics* 22, pp. 811--821, (1979).

[18] T. A Salthouse,. Theoretical Perspectives on Cognitive Aging. Hillsdale, NJ: Erlbaum, (1991).

[19] A. Worden, N. Walker, K. Bharat, and S. Hudson, *Making Computers Easier for Older Adults to Use: Area Cursors and Sticky Icons*. In: Proc. CHI1997, pp. 266--271, ACM Press (1997).

[20] P. Zaphiris, S. Kurniawan, and M. Ghiawadwala, *A systematic approach to the development of research-based wed design guidelines for older people.* Universal Access in the Information SocietyJ., vol.6, no.1, pp. 59--75, (2007).

[21] S. Zhai, J. Kong, and X. Ren, *Speed-accuracy tradeoff in Fitts' tasks—on the equivalency of actual and nominal pointing precision.* Int'l J. of Human-Computer Studies, 61(6), pp. 823-856, (2004),

[22] M. Ziefle, U. Schroeder, J. Strenk, and T. Michel, *How younger and older adults master the usage of hyperlinks in small screen devices*. In: Proc. CHI2007, pp. 307--316, ACM Press (2007).