Available online at www.sciencedirect.com
ScienceDirect
Int. J. Human-Computer Studies 68 (2010) 508–524
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

Benefits and costs of adaptive user interfaces

Talia Lavie*, Joachim Meyer

Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer Sheva, Israel

Received 29 December 2008; received in revised form 20 January 2010; accepted 27 January 2010
Available online 7 February 2010

Abstract

The paper examines the positive and the possible adverse effects of adaptive user interfaces (AUIs) in the context of an in-vehicle telematic system as a function of four factors: (1) four different levels of adaptivity (ranging from manual to fully adaptive with intermediate levels); (2) different tasks; (3) routine (familiar) and non-routine (unfamiliar) situations; and (4) different user age groups. Both experiments included three sessions during which participants drove a simple driving simulator and performed tasks with the telematic system at one of the adaptivity levels. We measured task performance times and lane position variance. Adaptivity was not always equally beneficial, and its benefits depended on a number of factors, including the frequency in which the tasks were performed, the user's age, the difficulty of the task and the user's involvement in the task. In familiar, routine situations, a fully adaptive system was beneficial for all participants, particularly older ones. In unfamiliar situations, to which the AUI was not adjusted, cognitive workload increased substantially, adversely affecting performance. Intermediate levels of adaptivity keep users involved in the task and help them become more proficient when performing both routine and non-routine tasks. However, intermediate levels of adaptivity should also be implemented with care, because they may also have adverse effects when users encounter non-routine situations.

Keywords: Adaptive user interface; Levels of adaptivity; Task frequency

1. Introduction

Adaptive user interfaces (AUIs) are defined as systems that adapt their displays and available actions to the user's current goals and abilities by monitoring user status, the system state and the current situation (Rothrock et al., 2002). The use of AUIs supposedly helps to improve user interaction with systems by facilitating user performance, minimizing the need to request help, easing system use, helping users deal with complex systems and avoiding cognitive overload problems (Browne et al., 1990; Edmonds, 1993; Hook, 1998, 1999; Trumbly et al., 1994). These benefits are apt to disappear (or are minimal) when AUIs violate usability design principles. For instance, AUIs are almost inherently inconsistent over time i.e., their interface or functionality may change. For additional examples of the possible usability problems that may arise from adaptivity, see Jameson (2003), Hook (1999), Keeble, Macredie (2000), Kuhme (1993) and Shneiderman (1997).

In spite of major progress in AUI research, we still lack a methodology for determining when and how adaptivity should be implemented. These decisions should be based on understanding the conditions in which benefits from adaptivity outweigh possible costs. In this paper we propose that AUI properties cannot be evaluated in isolation. Instead, one must consider the circumstance in which the system is used, the user population and other factors. For instance, the same algorithms for adaptation and the same interface may be very efficient in some contexts in which the system is always used the same way, and they may be very inefficient in other contexts in which system use varies more. We refer to this complex set of variables as the ecology of the system. Rather than focusing on one specific factor for determining whether adaptivity will be beneficial, we maintain that it is necessary to look at the whole ecology of system use. By looking at the wide range of relevant factors, system

*Corresponding author. Tel.: +972 8 6472216; fax: +972 8 647 2958.
E-mail addresses: tlavie@bgu.ac.il (T. Lavie), Joachim@bgu.ac.il (J. Meyer).

1071-5819/$ - see front matter
doi:10.1016/j.ijhcs.2010.01.004

designers will be in a better position to provide users with the systems they truly need and which serve their interests.

1.1. Conditions for successful adaptivity

Only few studies have attempted to map the conditions under which adaptivity can be beneficial, as opposed to the conditions in which it will have an adverse effect. Instead, most studies presented examples of successful and unsuccessful adaptation methods without looking at the causes for their success or failure (Gajos et al., 2006).

Findlater and Mcgrenere (2004), for example, compared static, customizable and adaptive versions of split menus, and they found somewhat different results under different circumstances. For example, subjects better understood the benefits of customization (i.e., placing the frequently used items near the top) when they were given the static or adaptive menu before the customizable menu, compared to when the customizable menu appeared first. In a more recent study, Findlater and Mcgrenere (2008) have shown that AUIs are more beneficial when screen real estate is constrained, and that the adaptive accuracy conditions were better in the small screen displays compared to the desktop sized displays. Tsandilas and Schraefel (2005) examined the impact of accuracy of two adaptation techniques (highlighted menu items vs. reduced font size of non-suggested items) applied to lists of textual selections. They concluded that the effectiveness of different adaptation techniques varies according to the accuracy of the prediction mechanism.

Gajos et al. (2006) addressed more explicitly which aspects of AUIs affect their success. They found the predictive accuracy of the AUI to have a significant impact on user performance. Based on previous research they also claim that the frequency of adaptation largely impacts the relative weights users assign to the different costs and benefits of adaptation. Slow paced adaptation, as applied by Sears and Shneiderman (1994), provided benefits (compared to the non-adaptive baseline), while fast paced adaptation, as used by Findlater and Mcgrenere (2004), leads to negative results. Finally, they demonstrated that the frequency of the interaction with the interface and the cognitive complexity of the task affected the aspects the users find relevant. Recently Gajos et al. (2008) showed that in addition to accuracy, the predictability of the AUI increased user satisfaction.

1.2. Factors involved in the interaction with AUIs

We suggest four key factors that determine the value of adding adaptivity to a system (in addition to the aspects described above)—the task, the user, the situation and the level of adaptivity.

1.2.1. The task

Tasks vary in many aspects, including their particular characteristics, difficulty level and what they require from the user. An AUI, for example, may be more beneficial for performing difficult rather than easy tasks. Additionally, some tasks require motor skills, while others may necessitate various cognitive skills. The outcome of using an AUI may change according to particular task requirements. Our research will look at different levels of task difficulty and how the difficulty affects the interaction with AUIs.

1.2.2. The user

Users differ in a wide range of variables, including demographic characteristics, background, education, personality, cognitive skills and preferences. Users' motivation, goals and moods also vary. Different users may interpret command names and icons differently and will attend to different aspects of computer displays (Benyon and Murray, 1993). One of the major challenges in the study of human–computer interaction is the question how to deal effectively with individual differences, preferences and experience. Different users may benefit differently from AUIs. We focus in our study on age, a user characteristic that is relevant in a wide range of situations. In our experiments we assess the differences between younger and older users in the benefits they derive from AUIs.

1.2.3. Routines versus non-routines

AUIs monitor user behavior over time, and they adjust the interface or the system functioning to the frequency of the tasks the individual user performs. Consequently, AUIs usually provide good support when users perform routine, frequent tasks. For example, an adaptive menu may provide users with easier access to frequently used menu items by displaying them at the beginning of the menu. However, at times users may need to perform infrequent, non-routine tasks. In such cases, when, for example, the user needs to search for a rarely- or never-used menu item, the AUI ceases to provide any benefits. It may, therefore, be valuable to differentiate between routine and non-routine tasks when examining AUIs. Only few studies have looked at the effects of task frequencies in the context of AUIs (see Bunt et al., 2004, for example). In our study we examine the effects of AUIs on the performance of routine and non-routine tasks.

1.2.4. Levels of adaptivity

We maintain that AUIs should not be viewed merely as having or not having adaptive features. Instead, AUIs can employ different levels of adaptivity. Levels of adaptivity relate to the question whether the system solely controls adaptation or whether adaptation is a co-operative process between the user and system. Several researchers have examined the question of allocating the control of the interaction between the user and system and offered different levels of adaptivity (see Dieterich et al., 1993; Kobsa et al., 2001; Kuhme, 1993). "Mixed-initiative interaction" is an additional term used in the context of controlling adaptation. Researchers of this approach are trying to make computers collaborate with users. The challenge is to define computation models of how initiative is or should be controlled in a human–computer dialog (Curry, 1998). A number of researchers have examined AUIs in a mixed-initiative interaction context (e.g., Schiaffino and Amandi, 2004). Bunt et al. (2004), advocating a mixed-initiative approach to interface adaptation, suggest integrating adaptable and adaptive systems. According to their view, the system should intervene and assist the user when difficulties arise in effectively customizing the system to user needs and abilities.

Level of control has also been examined extensively in the context of automation, specifically in the form of levels of automation (LOA). LOAs were suggested as a strategy for improving the functioning of the overall human–machine system by integrating the human and automated system in a way that allows the human to function effectively as part of the system (human-centered automation). Several LOA taxonomies have been proposed (e.g., Endsley, 1987; Endsle and Kabe, 1999; Kaber and Endsley, 2004; Sheridan and Verplanck, 1978). A design approach that focuses on utilizing intermediate levels of automation that integrate the human and the automated system in substantially different ways should be most beneficial (Wickens and Hollands, 2000). We advocate a design approach that looks at different levels of adaptivity, ranging from entirely manual (without adaptivity) to fully adaptive, with intermediate levels in which the user and system jointly control the interaction.

In summary, we believe that all four factors need to be considered to determine the value of adding adaptivity to a system. Some factors may remain constant in particular systems, such as the task and user, while other factors, such as the situation, can change over time. Moreover, not just the value of adaptivity, but also the appropriate level of adaptivity may depend on the particular task, situation and user. This study presents only an initial examination of these factors, so we will look at one property representing each factor, namely the difficulty level of the task, the user's age, the frequency of the task and the adaptivity levels.

2. The research

The research reported here continues a study that examined the advantages and limitations of different levels of adaptivity in various driving scenarios for routine and non-routine situations. In the study participants performed tasks with the telematic system at one of four adaptivity levels (manual, two intermediate and fully adaptive) while driving in a simple simulator. Task performance times and driving performance (lane position variance) were examined. The study demonstrated that in routine situations AUIs enhanced performances, that in non-routine situations high levels of adaptivity adversely affected performances, and that intermediate levels of adaptivity were generally more beneficial. However, the study was limited in a number of aspects. First, it used only a small sample (36 participants) of young drivers. Second, it was limited in terms of the conditions examined, and, for example, examined only simple driving conditions. Finally, improvements were made in the experimental system based on a number of problems that emerged in the previous study. Thus, this study aims to further examine the initial results obtained in the previous study by using a bigger sample, more diverse users in terms of their age, and different driving conditions.

Two experiments were carried out to examine the consequences of using AUIs with different levels of adaptivity in various conditions. The main independent variables in these experiment were: (1) different levels of adaptivity, ranging from entirely manual to fully adaptive with intermediate levels; (2) task characteristics, which relate to the different operations required to perform the tasks with the telematic system (reading, typing, etc.) and the difficulty of the driving; (3) the familiarity of the specific situation in which the tasks are performed (routine familiar tasks and non-routine, unfamiliar tasks); and (4) user age.

We examine AUIs in the context of in-vehicle telematic devices. These systems offer a variety of functions: driver assistance (such as access to navigation systems and traffic advice); entertainment (CD, radio, or other media players); communications (including cellular phone, text messages, email access, web access, etc.), and the control of other systems in the car (such as climate control). Designing in-vehicle telematic systems is particularly challenging because here, unlike in desktop environments, system use is secondary and must interfere as little as possible with the primary task, which is driving safely.

AUIs offer a solution to the challenges that arise in designing telematic systems. Indeed various AUIs have been introduced to support drivers operating such telematic systems. Examples are systems that reallocate telephone calls for reducing driver workload (Piechulla et al., 2003), adaptive route advisor systems (Fiechter and Rogers, 2000), and adaptive place advisors (Thompson et al., 2004). In the future, similar applications will probably be introduced into cars, and several researchers claim that AUIs are the key element for creating usable interfaces in automotive environments (e.g., Rogers et al., 2000). So far, however, very few AUIs have been implemented and tested in actual vehicles, and additional research is needed before vehicular AUIs can be widely adopted.

3. Experiment with younger participants

The first experiment examined the performance of younger participants in routine and non-routine situations using four levels of adaptivity varying between entirely manual and fully adaptive, with two intermediate levels. The experiment also examined different levels of driving difficulty.

The participants interacted with a mimicked, adaptive telematic system while driving. The mimicked system presented situations users encounter at two points in time: before the system adapts to the user and after adaptation, assuming a certain predefined user profile. We could thereby examine the consequences of using an AUI in various situations without having to develop a specific user profile based on prolonged and repeated interactions with the system.

The research hypotheses are:

(1) *In routine situations,* the fully adaptive system will be more beneficial, compared to the manual and intermediate adaptive systems. This will be indicated by shorter performance times with the telematic system, reduced lane deviation and better subjective evaluations for the fully adaptive system.
(2) *In non-routine situations,* the manual and intermediate systems will be more beneficial, compared to the fully adaptive system. The fully adaptive system will impair performance in particular in difficult driving conditions, as indicated by prolonged performance times with the telematic system and increased lane deviations.

3.1. Method

3.1.1. Participants

Sixty-four engineering students from Ben-Gurion University of the Negev with an average age of 25.7 (29 female and 35 male) participated in this experiment. They received course credit for their participation.

3.1.2. Apparatus

The experimental system consisted of two subsystems—a driving simulator and a simulator of an in-vehicle telematic system. The PC-based system displayed a road scene on a 21-in monitor located in the center of the participant's visual field (see Fig. 1). The driving simulator showed a two-lane curved road without additional traffic. A steering wheel controlled the car's lane position. The car traveled at one of two preset speeds, approximately 30 and 45 km/h, which participants could not change. The lane had two levels of curvature: high and low. With the high curvature, the car also drove at a higher speed, and with the lower curvature it drove at a lower speed. Thus the combination of speed and road curvature created two levels of driving difficulty—easy and difficult. The in-vehicle telematic system was simulated by a visual display (16 cm wide $\times$ 9 cm high) that was presented on a separate 15-in screen to the right of the driving simulator screen. The telematic system included three subsystems: a communication system (including text messages, email access, news updates); an entertainment system (including radio and CD); and a navigation system (including traffic updates). Fig. 1 also shows an example screen of the news update task. Participants controlled the telematic system with buttons located on the steering wheel (left, right buttons for navigation in the telematic system and an additional button for selection). The integrated system was connected to an output data file that recorded data on driving performance (the driver's steering actions and lateral lane position were recorded every 200 ms) and on task performance with the telematic system.

3.1.3. Procedure

Participants were requested to drive along the simulated road while performing a number of tasks with the telematic system. Because the telematic system only mimicked adaptivity, the participants were given a scenario that specified their preferences and the daily routine of tasks they were to perform with the telematic system while driving the car. The experiment began with a preliminary drive to familiarize participants with the system and the tasks in the manual mode, followed by two routine driving sessions and one non-routine driving session. During the first routine driving session, participants performed twelve tasks—four tasks that were repeated three times. During the second routine driving session, participants performed eight tasks—four tasks that were repeated twice each.

Routine tasks refer to tasks the participants performed repeatedly during the experiment. In the non-routine drive, non-routine (new) tasks were introduced, and the

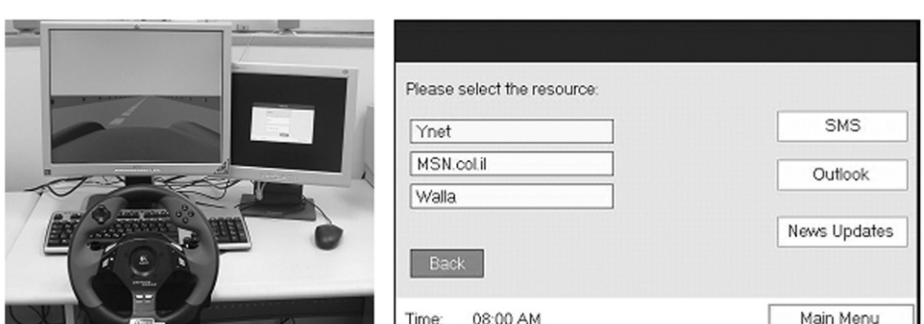

Fig. 1. The experimental system (left panel) and the news update screen as an example of the telematic system (right panel).

participants performed two routine and two non-routine tasks. The non-routine tasks required responses that differed from the responses required for the routine tasks. Prior to their drive, participants received detailed instructions about routine use of the telematic system. There were four routine tasks (see Appendix A for detailed descriptions of the tasks):

(1) *Reading a text message (Short Message Service [SMS] message) and sending a reply:* the participants were informed that they had received a text message they were required to read. They were told to reply "I am driving", their usual response when driving.
(2) *Reading an e-mail message:* participants received notice that they had received a new email message that they were required to read.
(3) *Receiving news updates:* participants were requested to read news updates from an Israeli news site, Ynet.
(4) *Changing to CD:* participants were told they were losing radio reception and should therefore change from radio to CD and select CD #3, track #3.

The non-routine driving session included four non-routine (new) tasks the participants did not perform previously, including:

(1) *Reading a text message (Short Message Service [SMS] message) and following instructions:* participants received a text message asking them to send as a reply the name of the city where they study (Beersheva).
(2) *Receiving news updates:* participants received a text message asking them to check CNN news, which they had not checked before.
(3) *E-mail check:* participants were asked to check the e-mail messages of a different user. To do so, they needed to change the user name in the e-mail system by entering the user's name and password.
(4) *CD change:* when participants tried to change from radio to CD, the system informed them that CD 3 was empty and that they should now change to CD #4, track #4 instead.

Each participant used the system with one of the four adaptivity levels:

(1) *Manual:* without adaptation—the participant performed all actions.
(2) *User selection (US):* the system initiated the action and presented the participant with the preferred alternative and additional alternatives. The participant selected the preferred alternative and implemented the decision.
(3) *User approval (UA):* the system initiated the action and notified the participant about the alternative he or she had chosen. The participant could then either approve the system choice or choose another alternative.
(4) *Fully adaptive (FA):* the system initiated the action, did not present the participant with other alternatives, decided and implemented the action. The participant could veto and abort the process at any stage.

We will demonstrate the levels of adaptivity on the example of the response to the text message (SMS). When the system notified the participant that he or she had received an SMS message, the participant had to reply with the message "I am driving". In the manual condition, participants were requested to reply manually by typing their reply on a virtual keyboard. In the US condition, because the system had presumably acquired the participant's frequent responses, it presented three most frequent responses the participant had supposedly used ("I am driving", "Talk to you later" and "Cannot talk now, call you later"). The participant was instructed to select the first option. In the UA situation, the system presented the participant with the most frequent reply, which is "I am driving", and the participant was requested to consent to the system sending that reply. In the FA state, the system automatically sent the participants' most frequent response. In the non-routine tasks, all participants had to override the adaptivity of the system and eventually performed the tasks manually. The tasks were announced through text messages that appeared in the upper part of the screen displaying the telematic system. The text messages were always accompanied by an auditory message describing the task (for example "you have received a new SMS message"). The appearance of the next task was conditioned on the successful completion of the previous task. All four driving sessions (preliminary, 2 routine and 1 non-routine driving sessions) took place in one experimental meeting lasting about 90 min, with 5-min rest periods between drives and time for completing the questionnaires.

3.1.4. Experimental design

A $4 * 2 * 3 * 4$ mixed between-within experimental design was employed. The level of adaptivity was manipulated as the between-subjects variable. The within-subjects variables included: the driving difficulty (easy and difficult driving conditions); the driving session (two routine and one non-routine drive); and four tasks (reading a text message, reading news updates, checking e-mail and changing a CD). Half the participants performed each task first in the easy driving condition and then in the difficult one: Drive1: easy then difficult; Drive2: difficult then easy; Drive3: easy then difficult. The other half performed the tasks in reversed order to counterbalance the driving difficulty variable: Drive1: difficult then easy; Drive2: easy then difficult; Drive3: difficult then easy.

The dependent variables in both studies included: (1) *Driving performance:* performance was measured by calculating the root mean square deviation (RMSD) from a participant's mean lane position for a certain segment of the route. Measures were computed for the times the participants performed the tasks with the telematic system. (2) *Performance with the telematic system:*

performance time was measured in milliseconds for all actions with the telematic system from the moment the message appeared on the screen until the participant completed the task.

3.2. Results and discussion

A possible problem with the use of AUIs is the need to override the system adaptivity when it becomes necessary to perform a new action to which the system did not adapt. Ideally, a system that is adequately designed will support such actions and enable users to override the adaptivity with minimal effort and perform the action they wish to perform. In our experimental system, the more adaptive the system is, the more actions it performs automatically. For the SMS task, for example, in the manual condition participants perform all tasks by selecting the desired functions from the menus. In the UA condition the users are automatically transferred to the message after selecting the "SMS" option, and in the FA condition, they are automatically transferred to the new SMS message, shortly after receiving notice of a new message. Therefore, when participants want to perform a non-routine task (as we requested in the third drive) they must override the system adaptivity. This may be more difficult in the more adaptive conditions.

Incorrect adaptation is costly for two reasons. First, it may be necessary to leave the screen or the settings into which the adaptation has put the system, so that it becomes possible to control the system manually. In addition, the adverse effect of incorrect adaptation may be caused by people not being used to perform tasks manually, and therefore needing more time to perform the task when adaptation has failed.

To examine the time needed to override incorrect adaptivity, i.e., the first cost, we measured the minimal time required to perform each non-routine task with each level of adaptivity. The need to override the adaptivity only prolonged the SMS task (SMS—Manual: 33, US: 35, UA: 37, FA: 39 s), and even there the effect is relatively small. In the other tasks, overriding did not prolong performance times in the FA condition, compared to the other conditions. We concluded from this that, although some overriding was necessary for the SMS task, it required relatively little time, and it should not greatly affect the results.

Performance time and driving performances were analyzed using a four-way ANOVA with the adaptation level (Manual, US, UA and FA); the driving session (three driving sessions); task type (News updates, SMS, E-mail check and CD change); and driving difficulty (easy and difficult) as independent variables. The latter three were within-subject variables. The first repetition of the four tasks in the first driving session was considered a learning phase and was not analyzed.

For performance times we found a main effect of the adaptivity level, $F(3,56) = 21.2$, $p < 0.01$, (manual condition: 26.6, US: 20.4, UA: 20.2, FA: 17.8 s). Performance times decreased with the level of adaptivity and were fastest for the FA condition, but Tukey post-hoc analyses showed significant differences only between the manual condition and three adaptive conditions (US, UA and FA, all $p < 0.01$). Performance times also increased significantly with the driving difficulty, $F(1,56) = 138.39$, $p < 0.01$, (22.9 s in the difficult driving condition and 19.6 s in the easy driving condition). A main effect was also found for the drive, $F(2,112) = 572.8$, $p < 0.01$, (first drive (routine tasks): 16.4, second drive (routine tasks): 14.7, third drive (non-routine tasks): 32.7 s), which was caused by significantly longer performance times for the third drive ($p < 0.01$). Finally, performance times differed between the tasks, $F(3,168) = 587.7$, $p < 0.01$, (News updates: 14.6, SMS: 35.4, Email: 23.2, CD change: 11.7 s).

Driving performance was significantly better in the easy driving condition, $F(1,56) = 205.6$, $p < 0.01$ (RMSDs of 1.5 for the easy driving condition and 2.3 for the difficult driving condition), and significantly poorer for the third driving session, $F(2,112) = 75.4$, $p < 0.01$ (RMSDs of 1.7 for the first drive, 1.5 for the second drive and 2.5 for the third drive). We found no significant differences for the other variables.

Fig. 2 presents the interaction Driving session * Task type * Adaptivity level for performance time and driving performances, $F(18,336) = 36.61$, $p < 0.01$ and $F(18,336) = 2.67$, $p < 0.01$ respectively. Planned comparisons (summarized in Table 1) revealed that performance times in the first and second drives (the routine drives) for all tasks were longer in the manual condition, compared to the FA condition. Additionally, in all tasks, except for the News updates task, performance times were longer in the manual condition compared to the two intermediate adaptive conditions (US and UA). With regard to driving performance the planned comparisons revealed no significant differences between the two intermediate levels of adaptivity and the first and second drive (routine drives) in both performance time and lane deviation for all tasks and adaptivity conditions. The only significant effect was better performance in the FA condition than in the manual condition for the SMS task.

In the third drive (the non-routine drive), performance times increased and driving performance decreased for most of the tasks as a function of the level of adaptivity, and the change was greatest in the FA condition. Here, participants had trouble finding the required functions needed for completing the tasks. In performance time the difference between the manual and US conditions was significant for all tasks, except for the CD task. No significant differences were found between the two intermediate adaptivity levels (US and UA), and the difference between the UA and FA conditions was only significant for the SMS task. Thus, with the exception of the SMS and CD tasks, the major increase in performance time was observed between the manual and intermediate adaptive conditions.

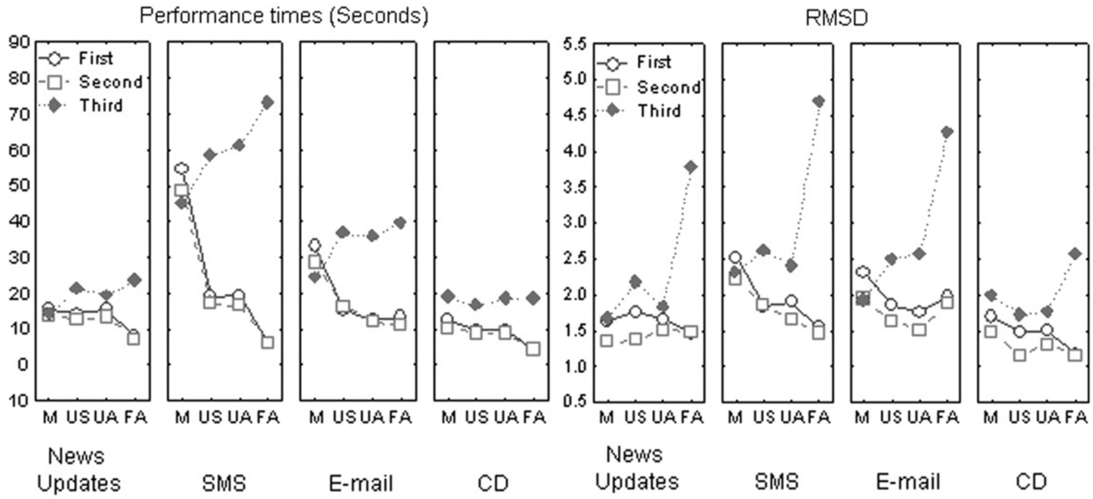

Fig. 2. Mean time to perform the tasks and RMSD for the four levels of adaptivity in the three drives. The first two drives include the routine tasks and the third drive includes the non-routine tasks (M is manual, US is user selection, UA is user approval, FA is fully adaptive).

Table 1
F values (df = 1,56) for planned comparisons between the adaptivity levels for all tasks for all three driving sessions in Experiment 1 (younger participants). The first two drives include the routine tasks and the third drive includes the non-routine tasks.

		News updates	*SMS*	*email*	*CD*
Performance time Drive 1	Manual–US	N.S.	277.2**	54.4**	14**
	Manual–UA	N.S.	272.7**	72.4**	11.8**
	Manual–FA	62.1**	518.7**	67.9**	115.1**
Performance time Drive 2	Manual–US	N.S.	185.7**	37.4**	12.1**
	Manual–UA	N.S.	192.5**	64.9**	7.6**
	Manual–FA	20.9**	338.6**	73.2**	191**
Performance Drive 3	Manual–US	13.90**	7.79**	20.05**	N.S
	US–UA	N.S.	N.S.	N.S.	N.S.
	UA–FA	N.S.	6.56*	N.S.	N.S.
RMSD Drive 1	Manual–US	N.S.	6.3*	N.S.	N.S.
	Manual–UA	N.S.	5.4	N.S.	N.S.
	Manual–FA	N.S.	12.9**	N.S.	4.7*
RMSD Drive 2	Manual–US	N.S.	N.S.	N.S.	N.S.
	Manual–UA	N.S.	N.S.	N.S.	N.S.
	Manual–FA	N.S.	164.4**	N.S.	N.S.
RMSD Drive 3	Manual–US	N.S.	N.S.	N.S.	N.S.
	US–UA	N.S.	N.S.	N.S.	N.S.
	UA–FA	12.4**	25.77**	14.67**	6.9*

*Significant at $p < 0.05$.
**Significant at $p < 0.01$.

With regard to the SMS task, performance times in the third drive (non-routine) increased between the intermediate and FA conditions. Notably, participants in all three conditions (UA, US and FA) performed the same task, i.e., read the message and typed a reply. In the FA condition in the first two drives (routine), the system performed the entire SMS task for the user. Thus, the user was not involved at all in performing the task. In contrast, in the news update and e-mail tasks, participants performed part of the task themselves. Apparently this kept them somewhat involved in the task, limiting the adverse effect of non-routine events. It appears that the level of involvement is less important for avoiding negative effects of non-routine events than the fact that participants remain somewhat involved in the task. This is also indicated by the lack of differences between the intermediate levels of adaptivity (that differed in the level of user involvement).

Another factor that may explain the difference between the SMS task and e-mail and news updates tasks is the task difficulty. The SMS task was the most difficult task (requiring typing a name of an Israeli city—Beersheva). No significant difference was found between the four adaptivity levels for the CD task. This may be because the CD task was a relatively easy task that required only selecting from lists (of CDs and songs), and the structure was very clear to the participants, with a known structure hierarchy of the lists (unlike in the news updates task, where the hierarchy structure was less trivial and required more knowledge).

Driving performance only differed significantly between the UA and FA conditions for the four tasks. It seems that differences in task performance times do not necessarily parallel differences in driving performance, in which the significant decrease was between the intermediate levels of adaptivity and the FA condition, and not between the intermediate and manual condition. Still, with the FA system performance times were longer and lane deviations greater than in the other conditions. Apparently the effect of participants' skill degradation in the non-routine condition was reflected very strongly in their driving performance.

There was a significant interaction Driving difficulty * Driving session * Adaptivity level for driving performance, $F(6,112) = 3.99$, $p < 0.01$ (see Fig. 3). Although driving speed and road curviness influenced lane deviations in all conditions, planned comparisons showed that significant differences between the easy and difficult driving conditions appeared only in the third drive (non-routine), $F(3,56) = 4.36$, $p < 0.01$. Still, the difference between UA and FA was significant in both easy and difficult driving conditions [easy–$F(1,56) = 15.38$, $p < 0.01$; difficult—$F(1,56) = 18.31$, $p < 0.01$].

For performance times we found an interaction between Driving difficulty * Task type * Adaptivity level, $F(9,168) = 4.56$, $p < 0.01$, see Fig. 4. The difficulty of the driving task influenced performance times with the telematic system for all tasks. Driving difficulty had the greatest impact on performance time in the manual condition, where the differences between performance time in the easy and difficult driving conditions were significant for all tasks [News update—$F(1,56) = 19.67$, $p < 0.01$; SMS—$F(1,56) = 14.77$, $p < 0.01$; e-mail—$F(1,56) = 11.87$, $p < 0.01$; CD change—$F(1,56) = 26.64$, $p < 0.01$]. Higher levels of adaptivity were beneficial in both easy and difficult driving conditions, and no significant differences were found between the easy and difficult driving conditions for all tasks.

In addition to the reasons mentioned above, the different results in the third drive (non-routine) may be attributed to *skill degradation*. According to Jameson (2003), users may become less skillful in performing a task manually after getting used to the system performing the task for them. In the manual condition, performance times in drive three did not increase significantly, because the participants were more proficient in performing all steps of the task. In contrast, participants in the adaptive conditions were not used to performing parts of the task manually and were therefore less proficient. This is especially true for the FA condition, which usually does not require participants to perform any actions.

Finally, although overriding the adaptivity required sometime in the SMS task, the increase in performance times in the two driving conditions in the FA condition was still much greater, compared to the other conditions (manual, US and UA). In the other tasks, the overriding did not require additional time. We conclude from this that

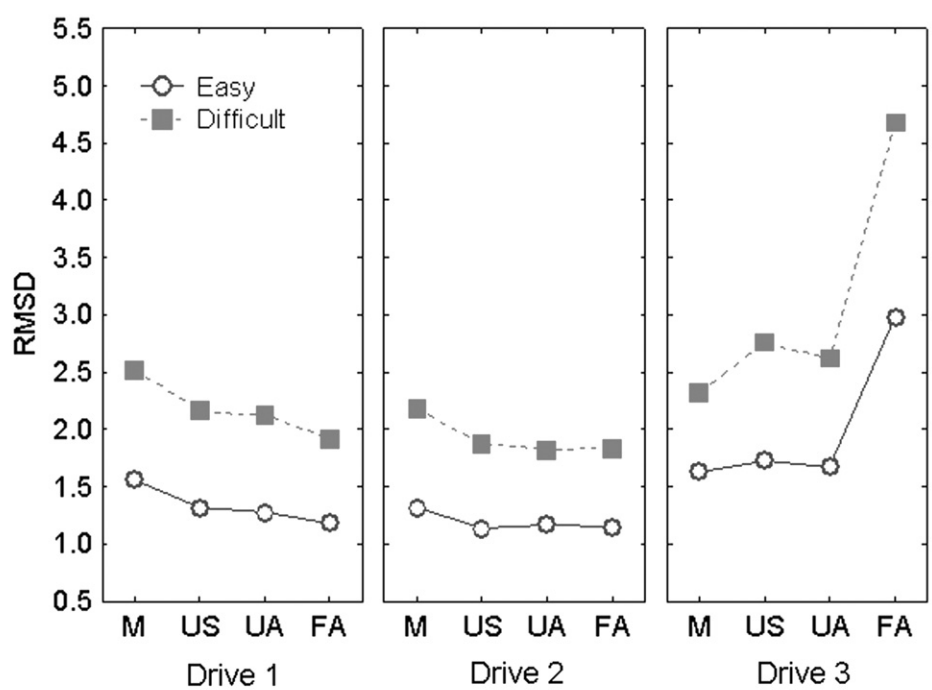

Fig. 3. RMSD for the three drives for the four levels of adaptivity in the easy and difficult driving conditions. The first two drives include the routine tasks and the third drive includes the non-routine tasks.

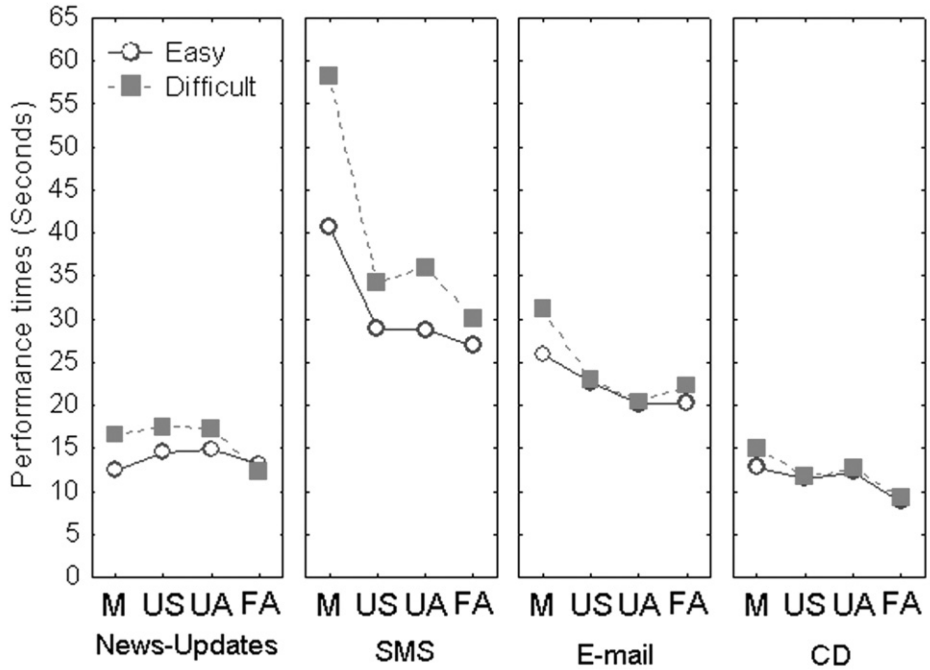

Fig. 4. Mean time to perform the tasks for the four levels of adaptivity in the easy and difficult driving conditions.

although overriding had a time cost in the SMS task, this cost is not large enough to explain the increase in performance time in the FA condition.

Overall, the results of both performance time and driving performance confirmed the hypotheses: (1) the FA system was more beneficial than the manual and intermediate systems in routine tasks; (2) the manual and intermediate systems were more beneficial than the FA system in the non-routine tasks. Nevertheless, in routine drives, adaptivity did not assist users more in the difficult driving conditions, compared to easy driving conditions. Adaptivity had, however, a negative effect in the non-routine drive in both driving conditions, and particularly in the difficult driving condition.

4. Experiment using older participants

In the second experiment we examined older participants. Older participants may differ in cognitive and motor abilities, which may affect their interaction with AUIs. Planek (1972) identified three interconnected areas of deficiency among older drivers: (1) sensory reception; (2) neural processing and transmission; and (3) motor response. Later studies emphasized the degradation of cognitive aspects of driving related to decision-making (Hakamies-Blomqvist, 1996), vision (Shinar and Schieber, 1991), divided attention, especially of the same output modality (Brouwer et al., 1990, 1991), skill acquisition (Craik and Jacoby, 1996) and responses to unexpected events (Falduto and Baron, 1986).

With respect to driving, it appears that age has only a small effect on response times to hazards on the road. It seems that older drivers are able to brake rapidly when they encounter a problematic situation as long as they expect it (Korteling, 1990; Lerner, 1993; Olson and Sivak, 1986).

Verwey (2000) examined the potential determinants of driver visual and mental workload and found that elderly drivers have less processing resources than younger drivers. Verwey (2000) also examined adapting message presentation to the type of road situation as a means of preventing driver overload. Results showed that adaptation does not prevent overload caused by complex messages for experienced drivers, whether young or middle-aged. However, adaptation may still help to prevent mental overload among inexperienced and older drivers and for more demanding tasks.

The objectives of this experiment were to examine different levels of adaptivity in routine and non-routine tasks with older participants and to assess whether the same pattern of results emerged as in the previous experiment. This experiment replicates the experiment with younger participants with two variations:

1. The previous experiment showed no significant difference between the two intermediate levels of adaptivity (US and UA), and therefore only one intermediate level of adaptivity was examined in this experiment. Thus, in this experiment we examined three levels of adaptivity: manual, US and FA.
2. Only the easy driving condition was utilized. We made this decision for two reasons. First, we established in the first experiment that the task difficulty affects driving performance, as well as performance with the telematic system. Additionally we assumed that older participants

will have trouble performing the tasks in the more difficult condition. Their performance in the easy condition was already clearly inferior to the younger drivers' performance.

We hypothesized that older users will benefit more from higher levels of adaptivity in routine situations, but suffer more from the adverse effects that arise in the non-routine situations. These hypotheses will be examined in both performance times and driving performance.

4.1. Method

The same experimental system was used in this experiment as in the experiment with younger participants, and therefore the next sections will only describe the changes between the two experiments.

4.1.1. Participants

Twenty-four participants, 12 females and 12 males participated in this experiment. Their average age was 58.6. The participants were professionals from a wide range of different fields. All participants volunteered to participate in the experiment.

4.1.2. Apparatus

The car drove at a fixed speed of approximately 30 km/h, and the lane position had only one level of curvature (the one used in the easier driving condition in the previous experiment).

4.1.3. Experimental design

A 3 * 2 * 3 * 4 mixed between-within experimental design was employed. The level of adaptivity was manipulated as the between-subject variable. The within-subject variables included: the driving session (two routine and one non-routine driving session); four tasks (reading a text message, reading news updates, checking e-mail and changing a CD); and repetition (first and second repetition of each task in each drive). The dependent variables were: (1) driving performance and (2) performance with the telematic system.

4.2. Results and discussion

Performance time and driving performances were analyzed using a 4-way ANOVA with the adaptation level (manual, US and FA); driving session (three drives); task type (four tasks); and repetitions (two repetitions) as the independent variables. All variables except the first one were within-subject variables. The first repetition of all the four tasks in the first drive was regarded as the learning phase and was not analyzed.

With regard to performance times we found main effects for all the variables examined. Performance times decreased with the adaptivity level, $F(2,21) = 9.32$, $p < 0.01$ (34.2 s for the manual condition, 28.3 s for the US and 23 s for the FA). However, Tukey post-hoc analyses revealed a significant difference only between the manual and FA conditions ($p < 0.01$). Performance times also varied according to the driving session, $F(2,42) = 78.1$, $p < 0.01$ (first drive (routine): 19.9 s, second drive (routine): 19 s, third drive (non-routine): 46.5 s), and were significantly higher for the third drive ($p < 0.01$). Performance times decreased for the second time the tasks appeared, $F(1,21) = 16.68$, $p < 0.01$ (29.5 s for the first repetition and 27.7 s for the second repetition). Finally, performance times differed between the tasks, $F(3,63) = 211.3$, $p < 0.01$ (with 17.5 s for News updates, 49.4 s for SMS, 31.5 s for email and 15.5 s for CD change).

In the driving performance significant main effects were found for the driving session, $F(2,42) = 63$, $p < 0.01$ (RMSDs of 1.5 for the first drive (routine), 1.3 for the second drive (routine) and 2.1 for the third drive (non-routine)), and driving performances were longest for the third drive. However, driving performances were significantly longer for the third drive only compared to the second drive ($p < 0.01$). Driving performance also improved for the second repetition, $F(1,21) = 7.1$, $p < 0.01$ (RMSDs of 1.7 for the first repetition and 1.6 for the second repetition). Finally, driving performance differed between the tasks, $F(3,63) = 16.9$, $p < 0.01$, (with RMSDs of 1.7 for News updates, 1.8 for SMS, 1.8 for email and 1.3 for CD change). No significant main effect was found for the adaptivity level.

An interaction Driving session *Task type * Adaptivity level was found for performance time and driving performance, $F(12,126) = 33.91$, $p < 0.01$ and $F(12,126) = 2.50$, $p < 0.01$ respectively (see Fig. 5). Planned comparisons conducted on the first two drives (routine drives) showed significant differences between the manual and FA condition and performances were worse in the manual condition, compared to the FA condition for all tasks, and for the SMS and email tasks (the more difficult tasks) with regard to driving performance (Table 2 summarizes the planned comparisons). These results demonstrate that older participants benefit from adaptivity in routine situations.

In the third drive, performance was poorer in the FA condition, compared to the manual condition for all tasks (except for the performance time in the CD task). Driving performance in the manual condition improved in most tasks in the third drive (non-routine) relative to the first two drives (routine). This indicates that it took older drivers some time to learn to maintain the lane position, and a different task did not impair that learning process. However, in the adaptive conditions, and especially in the FA condition, the novel tasks strongly affected older participants' ability to maintain lane position. Planned comparisons also revealed significant differences in performance times and driving performances between the manual and US conditions for all tasks expect for the CD task. Overall older participants' patterns of results resemble the results of younger participants.

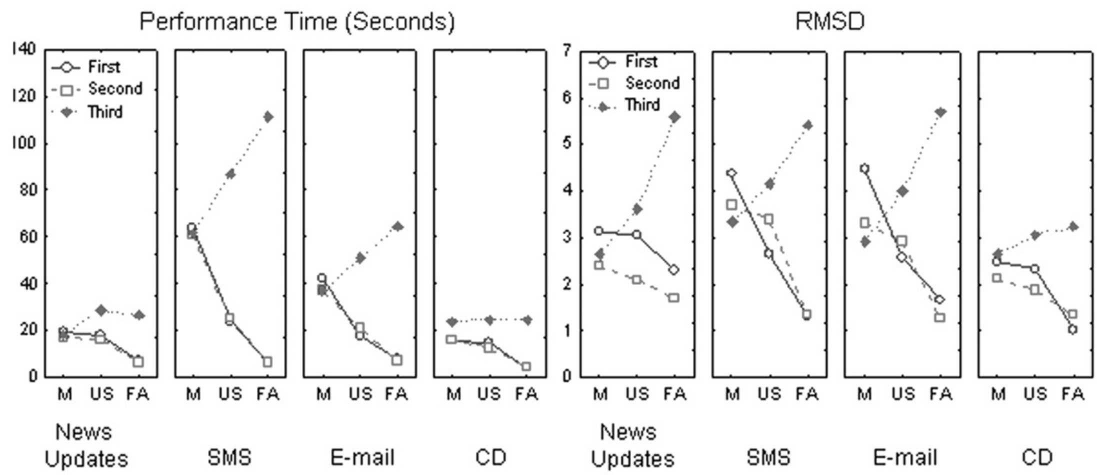

Fig. 5. Mean time to perform the tasks and RMSDs for the three levels of adaptivity in the three drives. The first two drives include the routine tasks and the third drive includes the non-routine tasks.

Table 2
F values (df = 1,21), for planned comparisons conducted between the adaptivity levels for all tasks for all three drives in Experiment 2 (older participants). The first two drives include the routine tasks and the third drive includes the non-routine tasks.

			News updates	*SMS*	*email*	*CD*
Performance time	Drive 1	Manual–US	N.S.	48**	37.8**	N.S.
		Manual–FA	53.5**	99.7**	75.9**	166.4**
	Drive 2	Manual–US	N.S.	52.2**	23.4**	13.8**
		Manual–FA	60.6**	120.9**	81.6**	181.5**
	Drive 3	Manual–US	12.49**	7.19*	197.87**	N.S.
		Manual–FA	8**	28.5**	14.8**	N.S.
RMSD	Drive 1	Manual–US	N.S.	7.6*	4.8*	N.S.
		Manual–FA	N.S.	23.8**	10.2**	5.4*
	Drive 2	Manual–US	N.S.	N.S.	N.S.	N.S.
		Manual–FA	N.S.	5.3*	5.9*	N.S.
	Drive 3	Manual–US	9.03**	6.1*	9.16**	N.S.
		Manual–FA	9**	6.1**	9.1**	N.S.

*Significant at $p < 0.05$.
**Significant at $p < 0.01$.

4.2.1. A comparison between younger and older participants

To examine the hypothesis that the effects of the different independent variables will be more pronounced for older participants, we compared the performance of younger and older participants. The following changes were made in the analysis to accommodate the differences between the studies:

(1) The UA condition existed only in the experiment with younger participants, and therefore it was not analyzed.
(2) The experiment with younger participants examined two driving conditions (easy and difficult), while the older participants only drove in the easy driving condition. Consequently, only the easy driving condition was analyzed.
(3) The experiment with younger participants included two versions that varied in the order of the driving difficulty while performing the task. Since we ignored the difficult driving condition and no significant difference was found between the versions, we analyzed both versions together. Each task was performed twice in each drive in this experiment. An average of the two repetitions was calculated and was used as the data for the younger participants.

We conducted a four-way ANOVA with the age of the participants (young and old); adaptation level (manual, US and FA); driving session (three driving sessions), and of the tasks (four tasks) as the independent variables. The latter two were within-subject variables. Performance times were significantly longer for the older participants (19.88 s for younger participants and 28.53 s for older participants, $F(1,42) = 58.34$, $p < 0.01$) and lane position variance was significantly greater for older participants (RMSDs of 1.5 for young participants and 2.9 for old participants, $F(1,42) = 23.46$, $p < 0.01$).

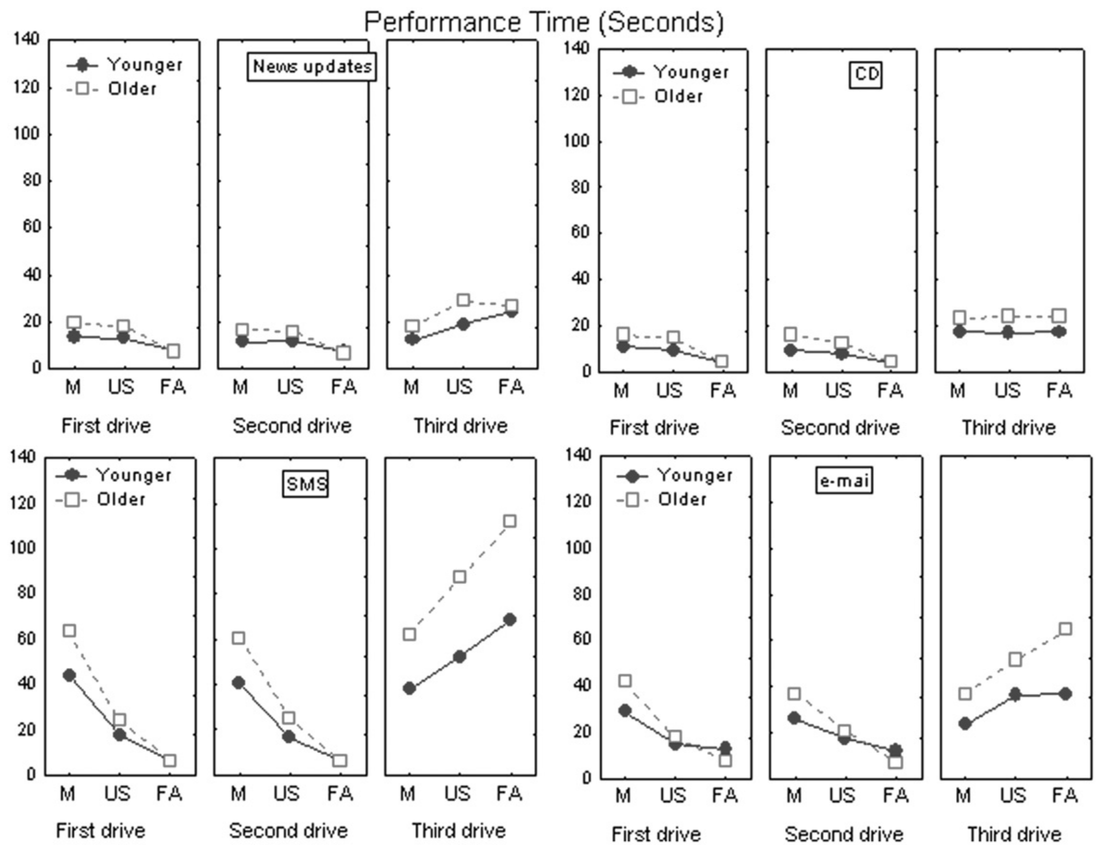

Fig. 6. Mean time to perform the tasks for the three levels of adaptivity in the three drives for younger and older participants. The first two drives include the routine tasks and the third drive includes the non-routine tasks.

The results correspond to the assertion that older participants will have more difficulty performing the tasks. This may be attributed to degraded motor skills, prolonged response times (Falduto and Baron, 1986), and degraded cognitive skills, such as difficulties in dividing attention between driving and an additional task (Brouwer et al., 1991). If this assertion is true, a difference should be found between the tasks, varying according to the difficulty level of the task, i.e., more difficult tasks will result in longer performance times. Indeed, the interaction Age ∗ Driving session ∗ Task type ∗ Adaptivity level was significant for performance time, $F(12,252) = 4.69$, $p < 0.01$. Fig. 6 presents the results.

Overall, similar result patterns were found for younger and older participants. In all tasks in the first two drives (the routine drives) older participants benefited more from a FA system than younger participants. As summarized in Table 3, planned comparisons revealed that older participants took significantly longer to perform all tasks in the manual condition for all tasks. However, no significant differences were found between younger and older participants in the FA condition. In the third drive, the cost of adaptivity was evident for both groups of participants. Performance time increased more in the FA condition for the older participants (for all tasks except for the news updates task). Moreover, different tasks yielded different results. In the simpler tasks – news updates and CD tasks – the difference between the performance of younger and older participants in the FA condition was minimal, and in the more difficult tasks – SMS and email – performance time increased substantially more for older participants. Thus older participants' performance is more impaired by non-routine situations than younger participants' performance.

Table 3
Results of planned comparisons conducted between older and younger participants for performance times for the manual, US and FA conditions for all tasks, $F(1,21)$. The first two drives include the routine tasks and the third drive includes the non-routine tasks.

		News updates	*SMS*	*Email*	*CD*
Drive 1	Manual	13.8**	19.6**	16.4**	24.2**
	US	10.6**	N.S.	N.S.	28.1**
	FA	N.S.	N.S.	N.S.	N.S.
Drive 2	Manual	25**	26.5**	10.2**	90.8**
	US	15.4**	4.4*	N.S.	45.1**
	FA	N.S.	N.S.	N.S.	N.S.
Drive 3	Manual	4.7*	11.5**	5.9*	15.1**
	US	14.3**	23.4**	7.7**	22.8**
	FA	N.S.	36.7**	25.4**	15.3**

*Significant at $p < 0.05$.
**Significant at $p < 0.01$.

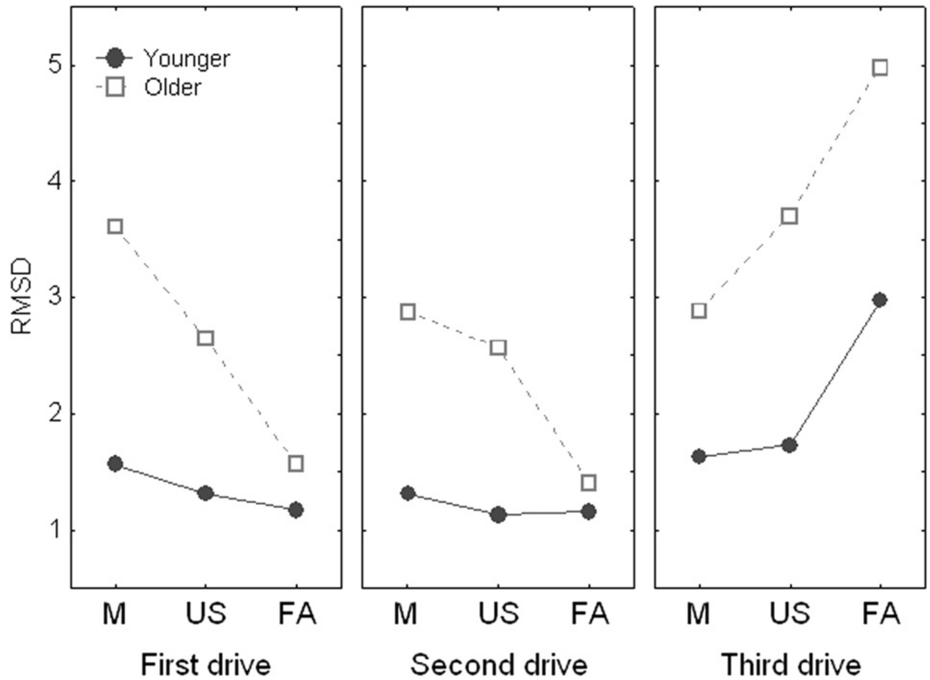

Fig. 7. RMSD for the three drives in the three adaptive conditions for younger and older participants. The first two drives include the routine tasks and the third drive includes the non-routine tasks.

For driving performance, a significant interaction Driving Session ∗ Age ∗ Adaptivity level was found, $F(4,84) = 8.63$, $p < 0.01$. The results (presented in Fig. 7) show that in the routine drives (first two drives) the largest differences between the older and the younger participants were in the manual condition (first driving session: $F(1,21) = 16.2$, $p < 0.01$, second driving session: $F(1,21) = 12.1$, $p < 0.01$). Differences were somewhat smaller, but still significant with the intermediate levels of adaptivity (first driving session US: $F(1,21) = 6.9$, $p < 0.05$, second driving session US: $F(1,21) = 10.1$, $p < 0.05$) and they were not significant for the FA condition. Thus, the FA system was very beneficial for older participants in terms of lane position variance, and the difference between the younger and older participants almost disappeared here. Apparently the FA system enabled the older participants to concentrate more on the driving task, resulting in decreased lane position variance.

On the other hand, performing tasks with the telematic system had less of a negative effect on younger participants, and therefore they benefited less from the adaptive system. In the non-routine drive (third drive), lane position variance increased significantly with the level of adaptivity for both younger and older participants, and for the older participants in particular (manual: $F(1,21) = 4.4$, $p < 0.05$, US: $F(1,21) = 11$, $p < 0.01$, FA: $F(1,21) = 11.3$, $p < 0.01$). In this case the adaptive system had a negative effect on performance for all participants, and especially for older participants.

The results confirm the hypothesis that AUIs are more beneficial for older participants in terms of performance times and driving performance in routine situations, and more costly in non-routine situations. The results of older participants are in line with those presented by Verwey (2000), who found adaptation to be important for older drivers, because it lowers their mental workload. The results also support what we know about the degradation of cognitive and motor skills among older drivers (i.e., Hakamies-Blomqvist, 1996; Brouwer et al., 1991), especially when two tasks require the same output modality, such as motor responses (e.g., Craik and Jacoby, 1996; Falduto and Baron, 1986; Korteling, 1990; Lerner, 1993; Olson and Sivak, 1986).

5. Summary of experiments

Overall the results demonstrate the potential advantage of using AUIs to change a system into an attentive assistant that supports the user by providing information or by performing actions in accordance with the user's preferences. However, the support of the AUI is only available as long as the user encounters situations and performs actions to which the system is adapted. When the user encounters an unfamiliar situation, the AUI is no longer adjusted and can no longer support the user. The user then has to override the actions of the AUI and must perform the task manually. The results show that this is more difficult for users who are accustomed to a fully adaptive system, since they probably experience skill degradation over time. Users who had used the intermediate levels of adaptivity also experienced difficulties in performing the unfamiliar task, but not as severely as users in the fully adaptive condition. These results indicate the advantage of using intermediate levels of adaptivity and keeping the user somewhat involved. If the user was not involved in performing a task during the first two drives, performance times increased when the user encountered a new version of the task (as in the third drive).

We have focused on routine and non-routine tasks, but we can extend the discussion to all cases in which AUIs suggest incorrect adaptations. In these cases the system may perform an irrelevant task. The user then needs to override the adaptivity and perform the task instead of the system. The system may also offer inadequate suggestions. Here the user will need to detect the system's inaccuracy and ignore the suggestion. The impact of the correctness of the adaptation has been previously demonstrated (Gajos et al., 2006, 2008; Findlater and Mcgrenere, 2008; Tsandilas and Schraefel, 2005). Some of the negative effects we showed in our research may also appear in these situations, in which incorrect adaptivity is offered by the system. In these cases intermediate levels of adaptivity may again benefit users.

Previous research has pointed to the importance of user control and involvement as influencing the interaction (e.g., Rouse and Morris, 1986; Sherman et al., 2003). An additional confirmation of the impact of user involvement comes from the difference between the intermediate adaptive systems and the fully adaptive systems. In non-routine situations, performance was generally poorer with the fully adaptive system, compared to the intermediate systems. Since users in both systems performed the same non-routine task, an explanation for these results is the level of involvement the user had in the routine tasks. The user, more involved with the system in the intermediate adaptive system, was better able to perform a new task with the system. On the other hand, a user who was not involved in the task at all, as in the case of the fully adaptive system, was less able to perform a new task. Additionally, no significant difference was found between the intermediate adaptive systems, suggesting that the level of involvement is less important than the involvement itself.

The experiment with older participants demonstrated that older users had more trouble operating the telematic system while driving. This difficulty was seen both in the time they needed to perform tasks with the telematic system and in their driving performance. It was evident in both routine and non-routine situations and was greater for more difficult tasks. Since older participants find it more difficult to perform the tasks, and particularly the difficult ones, they benefit from adaptivity more than younger participants. On the other hand, when adaptivity cannot support the task, as is the case in non-routine situations, the cost of using adaptivity is much higher.

To conclude, we found that all factors we examined affect the interaction with the AUI. Fig. 8 presents a schematic view of the relations among the factors we have

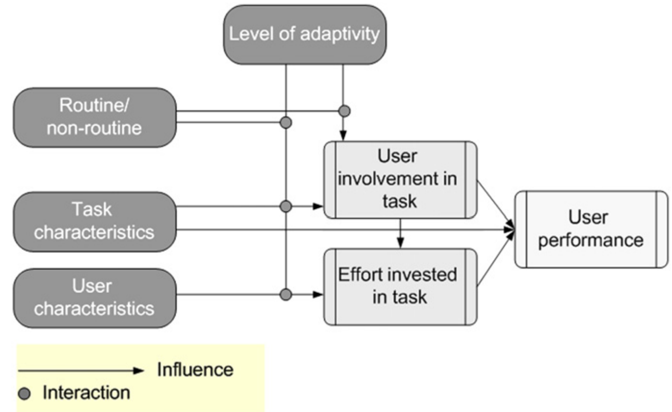

Fig. 8. A schematic representation of the model.

examined. It should be noted that some of the relations between the variables are more complex than could be presented in this schematic representation.

The arrows in the figure represent influencing factors and the dots represent the presence of interactions between the factors. The *user involvement in task* and *effort invested in task* are two factors affected by the level of adaptivity, task characteristics, user characteristics and routine vs. non-routine situations. User involvement relates to the degree to which the user is involved in performing the task with the system. The level of involvement depends on the level of adaptivity and the type of situation (routine vs. non-routine). For example, when performing a routine task using a manual system, the user is strongly involved in performing the task. In a fully adaptive system the user is less involved. On the other hand, when performing a non-routine task, users are equally involved in the task when they use the fully adaptive system and manual system. The level of involvement also depends on the task characteristics. Some tasks require more user involvement than others. The level of involvement in each task is influenced differently by the type of situation and the level of adaptivity. The effort invested in a task is influenced by the user involvement in the task and the interaction between the level of adaptivity and each of the other variables (task characteristics, user characteristics and routine vs. non-routine situations). For example, compared to a younger user, an older user invests more effort in performing a task with a manual system. On the other hand, older and younger users invest the same effort when using a fully adaptive system.

Additional factors, besides the ones studied here, are likely to affect performance with AUIs. For example, user characteristics, such as technology adaptation, or environmental driving conditions, such as night, fog and so forth may influence results. Also, people may use AUIs differently when they drive to work in the morning or when they are on a weekend trip. On a daily drive to work, they will probably perform more routine tasks, as opposed to what they might do on a leisurely drive through the countryside. Therefore an AUI that may be beneficial for the daily commute might be less useful on trips.

6. Conclusions and future work

The main objective of this research was to gain a systematic understanding of some of the consequences from the use of AUIs. Our research points to a number of important issues. First, adaptivity is not equally beneficial under all conditions, and since AUIs are usually used in complex, changing environments, it is important to consider the different variables affecting the interaction. We have shown that the preferred type of system depends on a number of factors, such as the frequency at which the tasks are performed, the user's age, the difficulty level of the task and the level of user involvement in the task.

We also found that it may be beneficial to consider intermediate levels of adaptivity, rather than seeing the introduction of adaptivity as an all-or-none decision. Intermediate levels of adaptivity keep users involved in the task and help them become more proficient when performing both routine and non-routine tasks. However, even when intermediate levels of adaptivity are implemented, one needs to consider the following aspects:

(1) *The proportion of routine and non-routine tasks:* AUIs utilizing intermediate levels of adaptivity may be more suitable as the proportion of non-routine tasks, relative to the routine tasks, increases. In situations in which only routine tasks are performed, intermediate levels of adaptivity may be less advantageous compared to the fully adaptive system.
(2) *The type of task:* users have to invest more effort to perform more complex tasks, so higher levels of adaptivity may be more beneficial for these tasks. Intermediate levels of adaptivity may not provide any benefits when performing easy tasks.
(3) *The users:* intermediate levels of adaptivity were less useful for older users than for younger users.

It should, however, be noted that our research only examined a limited range of properties relating to each factor. Additional variables related to the task, the user and the frequency of routine and non-routine tasks should be examined for a better understanding of their impact.

Of course, adaptivity is in reality more complex than the simplified laboratory situation in our experiment. However, AUIs are still not perfect (and perhaps will never be so). Our study contributes to the question whether we should develop and install such systems, and what should be considered when doing so. We propose to model different usage situations to determine beforehand whether adaptivity will benefit a system or not. This can save designers and users considerable time, money and inconvenience.

Appendix A

See Table A1.

Table A1
Description of the four tasks using the four adaptive conditions.

	Manual	*US*	*UA*	*FA*
News updates				
Initial system message and action	The user receives a notice to check news updates	The user receives a notice to check news updates	The user receives a notice to check news updates	The system notifies the user to check news updates and automatically presents the news updates
Actions required by the user	1. In main menu select "Communications" 2. Select "News updates" 3. Select "Local News" out of a list of news categories 4. Select "Ynet" out of a list of news sources 5. Press "Exit" after reading headlines	1. In main menu select "Communications" 2. Select "News updates" 3. Select "Ynet" out of a smaller list of news sources 4. Press "Exit" after reading headlines	1. In main menu select "Communications" 2. Select "News updates" 3. Press "Exit" after reading headlines	Press "Exit" after reading headlines
Reading a text message				
Initial system message and action	The user receives a notice he received a new text message	The user receives a notice he received a new text message	The user receives a notice he received a new text message	The system notifies the user of a new text message, automatically opens the message window and sends an automatic reply after a few seconds

Table A1 (*continued*)

	Manual	***US***	***UA***	***FA***
Actions required by the user	1. In main menu select "Communications" 2. Select "SMS" 3. Select "Reply" after reading the message 4. Type reply on virtual keyboard 5. Press "Send"	1. In main menu select "Communications" 2. Select "SMS" 3. Select "Reply" after reading the message 4. Select a reply from a list of reply options 5. Press "Send"	1. In main menu select "Communications" 2. Select "SMS" 3. Select "Reply" after reading the message 4. Press "Send" after the system presents the reply usually sent	No actions are required in part of the user except for reading the text message received
Reading an email message				
Initial system message and action	The user receives a notice of a new e-mail message	The user receives a notice of a new e-mail message	The user receives a notice of a new e-mail message	The system notifies the user that he received a new message and automatically presents the new message received
Actions required by the user	1. In main menu select "Communications" 2. Select "Inbox" 3. Enter user name and password 4. Press "Enter" 5. Select "Read Message" 6. Presses "Exit" after reading the message	1. In main menu select "Communications" 2. Select "Inbox" 3. Select "Read Message" 4. Press "Exit" after reading the message	1. In main menu select "Communications" 2. Select "Inbox" 3. Press "Exit" after reading the message	Press "Exit" after reading the message
Changing to CD				
Initial system message and action	The user receives a notice of bad reception suggesting switching to CD	The user receives a notice of bad reception suggesting switching to CD	The user receives a notice of bad reception suggesting switching to CD	Notifies the user of bad reception and automatically changes from radio to CD due to bad reception
Actions required by the user	1. In main menu select "Radio/CD" 2. Select "CD" 3. Select "CD 3" out of the CD list 4. Select "Track 3" from the list of tracks	1. In main menu select "Radio/CD" 2. Select "CD" 3. Select "CD 3" out of a smaller list of CDs 4. Select "Track 3" from the list of tracks	1. In main menu select "Radio/CD" 2. Select "CD" 3. Select "Track 3" from the list of tracks	No actions required

References

Benyon, D.R., Murray, D.M., 1993. Applying user modeling in human–computer interaction design. Artificial Intelligence Review 6 (3), 43–69.

Brouwer, W., Ickenroth, J.G., Ponds, R.W.H.M., Van Wolfelaar, P.C., 1990. Divided attention in old age. In: Drenth, P., Sergeant, J., Takens, R. (Eds.), European Perspectives in Psychology, vol.2; 1990, pp. 335–348.

Brouwer, W., Waterink, W., van Wolfelaar, P., Rothengatter, T., 1991. Divided attention in experienced young and older drivers: lane tracking and visual analysis in a dynamic driving simulator. Human Factors 33 (5), 573–582.

Browne, D., Totterdell, P., Norman, M. (Eds.), 1990. Adaptive User Interfaces. Computers and People Series. Academic Press, Harcourt Brace Jovanovich Publishers, London.

Bunt, A., Conati, C., McGrenere, J., 2004. What role can adaptive support play in an adaptable system? In: Proceedings of the 9th International Conference on Intelligent User Interfaces, Funchal, Madeira, Portugal, 13–16.

Craik, F.I.M., Jacoby, L.L., 1996. Aging and memory: implications for skilled performance. In: Rogers, W.A., Fisk, A.D., Walker, N. (Eds.), Aging and skilled performance: advances in theory and applications. Erlbaum, Mahwah, NJ, pp. 113–137.

Curry, I.G., 1998. An analysis of initiative selection in collaborative task-oriented discourse. User Modeling and User Adapted Interaction 8 (3–4), 255–314.

Dieterich, H., Malinowski, U., Kuhme, T., Schneider-Hufschmidt, M., 1993. State of the art in adaptive user interfaces. In: Schneider-Hufschmidt, M., Kuhme, T., Malinowski, U. (Eds.), Adaptive User Interfaces Principles and Practice. Elsevier Science Publishers B.V, pp. 13–48.

Edmonds, E.A., 1993. The future of intelligent interfaces: not just "how?" but "what?" and "why?". In: Proceedings of the First International Conference on Intelligent User Interfaces, Orlando, Florida, U.S.A, 13–17.

Endsley, M.R., 1987. The application of human factors to development of expert systems for advanced cockpits, In: Proceedings of the Human Factors Society 31st Annual Meeting. Human Factors and Ergonomics Society, Santa Monica, CA 1388–1392.

Endsle, M.R., Kabe, D.B., 1999. Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics 42 (3), 462–492.

Falduto, L.L., Baron, A., 1986. Age-related effects of practice and task complexity on card sorting. Journal of Gerontology 41 (5), 659–661.

Fiechter, C.N., Rogers, S., 2000. Learning subjective functions with large margins, In: Proceedings of the Seventeenth International Conference on Machine Learning. Stanford University 287–294.

Findlater, L., Mcgrenere, J., 2004. A comparison of static, adaptive, and adaptable menus. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 89–96.

Findlater, L., Mcgrenere, J., 2008. Impact of screen size on performance, awareness, and user satisfaction with adaptive graphical user interfaces. In: Proceeding of the Twenty-sixth Annual SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 1247–1256.

Gajos, K.Z., Czerwinski, M., Tan, D.S., Weld, D.S., 2006. Exploring the design space for adaptive graphical user interfaces. In: Proceedings of the Working Conference on Advanced Visual Interfaces, New York, NY, USA, 201–208.

Gajos, K.Z., Everitt, K., Tan, D.S., Czerwinski, M., Weld, D.S., 2008. Predictability and accuracy in adaptive user interfaces. In: Proceeding of the Twenty-sixth Annual SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 1271–1274.

Hakamies-Blomqvist, L., 1996. Research on older drivers: a review. Journal of the International Association of Traffic and Safety Sciences 20, 91–101.

Hook, K., 1998. Evaluating the utility and usability of an adaptive hypermedia system. Knowledge Based Systems 10 (5), 311–319.

Hook, K., 1999. Designing and evaluating intelligent user interfaces. In: Proceedings of the International Conference on Intelligent User Interface, Los Angeles, CA, USA, 5–6.

Jameson, A., 2003. Adaptive Interfaces and Agents. In: Jacko, J.A., Sears, A. (Eds.), Human–Computer Interface Handbook. Erlbaum, Mahwah, NJ, pp. 305–330.

Kaber, D.B., Endsley, M.R., 2004. The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. Theoretical Issues in Ergonomics Science 5 (2), 113–153.

Keeble, R.J., Macredie, R.D., 2000. Assistant agents for the world wide web: intelligent interface design challenges. Interacting With Computers 12 (4), 357–381.

Kobsa, A., Koeneman, J., Wolfgang, P., 2001. Personalized hypermedia presentation techniques for improving online customer relationships. The Knowledge Engineering Review 6 (2), 111–155.

Korteling, J.E., 1990. Perception-response speed and driving capabilities of brain-damaged and older drivers. Human Factors 32 (11), 95–108.

Kuhme, T., 1993. A user-centered approach to adaptive interfaces. In: Proceedings of the International Conference on Intelligent User Interfaces, Orlando Florida, 243–245.

Lerner, N.D., 1993. Brake-perception reaction times of older and younger drivers. In: Proceedings of the Human Factors Society 37th Annual Meeting, Santa Monica, CA, 206–210.

Olson, P.L., Sivak, M., 1986. Perception-response time to unexpected roadway hazards. Human Factors 28 (1), 91–96.

Piechulla, W., Mayser, C., Gehrke, H., Konig, W., 2003. Reducing drivers' mental workload by means of an adaptive man–machine interface. Transportation Research Part F, 6(4): Traffic Psychology and Behavior, 233–249.

Planek, T.W., 1972. The aging driver in today's traffic: a critical review. In: Planek, T.W., Mann, W.A., Wiener, E.L. (Eds.), Aging and Highway Safety: The Elderly in a Mobile Society. North Carolina Symposium on Highway Safety, Chapel Hill, NC, pp. 1–38.

Rogers, S., Fiechter, C.N., Thompson, C., 2000. Adaptive user interfaces for automotive environments. In: Proceedings of the IEEE Intelligent Vehicles Symposium. Dearborn, MI, 662–667.

Rothrock, L., Koubek, R., Fuchs, F., Haas, M., Salvendy, G., 2002. Review and reappraisal of adaptive interfaces: toward biologically inspired paradigms. Theoretical Issues in Ergonomics Science 6 (2), 157–172.

Rouse, W.B., Morris, N.M., 1986. On looking in the blackbox: prospects and limits in the search for mental models. Psychological Bulletin 100 (3), 349–383.

Schiaffino, S., Amandi, A., 2004. User interface agent interaction: personalization issues. International Journal of Human–Computer Studies 60 (1), 129–148.

Sears, A., Shneiderman, B., 1994. Split menus: effectively using selection frequency to organize menus. ACM Transactions on Computer–Human Interaction 1 (1), 27–51.

Sheridan, T.B., Verplanck, W.L., 1978. Human and Computer Control of Undersea, Teleoperators. MIT, Technical report, Man–Machine Laboratory, Cambridge, MA.

Sherman, R., Alpert, J.K., Karat, C., Carolyn, B., Vergo, J., 2003. User attitudes regarding a user-adaptive e-commerce web site. User Modeling and User-adaptive Interaction 13 (4), 373–396.

Shinar, D., Schieber, F., 1991. Visual requirements for safety and mobility of older drivers. Human Factors 33 (5), 507–519.

Shneiderman, B., 1997. Direct manipulation for comprehensible, predictable and controllable user interface. In: Proceedings of the 1997 International Conference on Intelligent User Interfaces. ACM Press.

Thompson, C., Goker, M., Langley, P., 2004. A personalized system for conversational recommendations. Journal of Artificial Intelligence Research (JAIR) 21, 393–428.

Trumbly, J.E., Arnett, K.P., Johnson, P.C., 1994. Productivity gains via an adaptive user interface: an empirical analysis. International Journal of Human–Computer Studies 40 (1), 63–81.

Tsandilas, T., Schraefel, M.C., 2005. An empirical assessment of adaptation techniques. CHI '05 Extended Abstracts on Human Factors in Computing Systems, New York, NY, USA, 2009–2012.

Verwey, W., 2000. On-line driver workload estimation: effects of road situation and age on secondary task measures. Ergonomics 43 (2), 187–209.

Wickens, C.D., Hollands, J.G., 2000. Engineering Psychology and Human Performance third ed. Prentice Hall, Upper Saddle River, New-Jersey.