

A user study of auditory, head-up and multi-modal displays in vehicles



Grega Jakus^{a,*}, Christina Dicke^b, Jaka Sodnik^a

^a Faculty of Electrical Engineering, University of Ljubljana, Trzaska 25, Ljubljana 1000, Slovenia

^b Quality and Usability Lab, Telekom Innovation Laboratories, TU Berlin, Ernst-Reuter-Platz 7, 10587 Berlin, Germany

ARTICLE INFO

Article history:

Received 13 September 2013

Accepted 4 August 2014

Available online 21 August 2014

Keywords:

Auditory display

Head-up display

Vehicle

ABSTRACT

This paper describes a user study on the interaction with an in-vehicle information system (IVIS). The motivation for conducting this research was to investigate the subjectively and objectively measured impact of using a **single- or multi-modal IVIS while driving**. A hierarchical, list-based menu was presented using a windshield projection (head-up display), auditory display and a combination of both interfaces. The users were asked to navigate a vehicle in a driving simulator and simultaneously perform a set of tasks of varying complexity. The experiment showed that the interaction with visual and audio-visual head-up displays is faster and more efficient than with the audio-only display. All the interfaces had a similar impact on the overall driving performance. There was no significant difference between the visual only and audio-visual displays in terms of their efficiency and safety; however, the majority of test subjects clearly preferred to use the multi-modal interface while driving.

© 2014 Elsevier Ltd and The Ergonomics Society. All rights reserved.

1. Introduction

Head-up displays (HUDs) are the current state-of-the-art solution to reducing driver errors originating from distractive interfaces, such as on-board entertainment displays or in-vehicle information systems (IVIS). Compared to head-down displays (HDDs), which are integrated into the vehicle's control panel, the frequency and duration of glances towards the display are reduced by presenting information directly on the windshield in the driver's field of vision (Ablassmeier et al., 2007). Consequently, the response time to unanticipated road events is reduced when information is displayed on a HUD instead of a HDD (Horrey and Wickens, 2004; Liu and Wen, 2004; Sojourner and Antin, 1990), and the number of collisions is significantly reduced (Charissis et al., 2008). Using HUDs also leads to a, generally, more consistent speed control and reduced mental workload (Liu and Wen, 2004), reduced navigational errors (Burnett, 2003), and smaller variances in lateral accelerations and steering wheel angle (Yung-Ching, 2003).

However, cognitively switching between two sources of information, i.e., the HUD and traffic, still poses a problem, especially in high-workload situations. The so-called cognitive or attention

capture, i.e., when the driver's attention subconsciously shifts away from the road and becomes focused on processing the information presented by the HUD, has been identified as one of the disadvantages of HUDs (Gish and Staplin, 1995; Prinzel and Risser, 2004; Tufano, 1997). The resulting perceptual tunnelling may lead to a delayed reaction or a complete absence of response to situational changes in the environment (Haines, 1991; Thomas and Wickens, 2001).

Researchers have proposed to address this issue by utilising auditory interfaces (Lai et al., 2001; Sodnik et al., 2008). They report on the greater attention capturing properties of the auditory channel and the desire to keep attention focused longer on a complex auditory task to prevent a loss of information from the driver's working memory. Wickens et al. (2005) addressed this apparent contradiction in the context of a visual on-going task and an auditory interrupting task. They found that the auditory input improved the interrupting task performance when compared to a visual interrupting task, but degraded the on-going task performance because of an abrupt shift of attention at the onset of the auditory message (see, in addition, Biever, 2002; Horrey and Wickens, 2004).

A combination of auditory, especially speech-based, information presentations and text-based information presentations through a HUD may offer the attention and safety benefits of an eyes-free approach with the higher information-processing rate of a visual information presentation. The study presented in this paper was conducted to evaluate the assumption that allowing users to freely

* Corresponding author. Tel.: +386 1 4768 938.

E-mail addresses: grega.jakus@fe.uni-lj.si (G. Jakus), christina.dicke@tu-berlin.de (C. Dicke), jaka.sodnik@fe.uni-lj.si (J. Sodnik).

switch their attention between visual and auditory information presentations may have less impact on the performance of the primary task than with a single-modal information presentation.

1.1. Related work

Liu (2001) conducted a simulator-based study comparing a visual display centrally mounted above the dashboard to an audio-only (digitised human female voice) and a multi-modal (audio and visual) display of traveller information with varying complexity (icons and text). Three to five information units were displayed in the low-complexity condition and between nine and fourteen information units in the high complexity condition. Liu measured the perceived workload, navigational performance, and information detection under both high- and low-load driving conditions. Both the auditory and multi-modality displays impacted the driving task but lead to a better performance in terms of response times, total number of correct turns, and subjective workload ratings when compared to the visual-only display. Liu summarised that the driving performance was the least affected by the multi-modal display.

Sodnik et al. (2008) compared three types of interfaces for in-vehicle information systems; an HDD was compared to two auditory interfaces: one with a monophonic output playing from different spatial positions and one with multiple, simultaneously playing sounds. The results of the user study showed significant differences between the visual and auditory interfaces: the interaction with the visual interface was faster and more efficient, while the auditory interface was significantly safer and preferred by the majority of users. A comparison of the two auditory interfaces showed no advantage of the spatial over non-spatial audio in the vehicles. On the contrary, the use of simultaneous spatial sounds increased the mental workload of drivers and resulted in an increased disturbance instead of an improved interaction. The total mental workload proved to be comparable in all three conditions.

Weinberg et al. (2011a) compared the impact of interacting with combinations of a HDD and text-to-speech (TTS) system, a HUD and a TTS system, and an audio-only TTS representation of lists of choices. A menu representing three domains of content (Navigation, Music and Contacts) with various subdomains and functions was used, and participants could traverse the hierarchical menu by using a steering-wheel-mounted jog dial device or speech input. Weinberg et al. (2011b) measured task completion times, perceived cognitive loads, and personal preferences while participants followed a leading vehicle in a driving simulator (the primary, ongoing task) and solved secondary tasks using either of the IVIS display types. They also reported on the driving performance as indicated by the lateral position, speed, steering and throttle and the following distance (Weinberg et al., 2011a). Generally, the HUD and audio-only interfaces had similar impacts on the driving performance and mental workload with a slight advantage found using the audio-only version. However, the task completion times were significantly longer when using the audio-only interface. Participants rated the auditory menu as the least distracting version, while they found both the audio-only and the HUD to be the easiest to use and the most desirable.

The results of previous studies (Sodnik et al., 2008; Weinberg et al., 2011a) provide a basis for the assumption that the display modality of an IVIS has a strong impact on both the primary and secondary tasks in a vehicle. Prior findings also support the assumption that a combination of auditory, especially speech-based information presentations and text-based information presentations through a HUD, may be the “best of both worlds” – the attention and safety benefits of an eyes-free approach with the higher information processing rate of a visual information

presentation. In the following study, we compare two single-modality interfaces, an audio-only and a video-only interface, to a multi-modal HUD (Weinberg et al. (2011a) compared an audio-only display to two different multi-modal displays). We asked participants to perform a series of tasks while driving. In contrast with previous studies, our tasks varied in complexity between a simple “search for and find a piece of information” task and a more complex “identify and interpret” task.

The experimental design, including the technical setup, is described in the following section. This is followed by an analysis of the gathered data and an interpretation and discussion of the findings. The paper concludes with recommendations for the design of list-based head-up displays to access multi-modal in-vehicle information and entertainment systems.

2. Methods

2.1. User study design

Three different IVIS interfaces were evaluated in regards to conducting typical (secondary) tasks while driving a simulated vehicle in different traffic conditions (“high-speed” and “low-speed” traffic was randomly distributed among the experimental conditions). Three different experimental conditions were compared based on these three different display techniques:

- a visual display/HUD (V),
- an auditory display (A) and
- a multi-modal auditory and visual display/HUD (AV).

The research questions addressed in this study are as follows:

- Which interface is more efficient in terms of task completion time and interaction activity?
- Which interface is less detrimental to driving performance and why (measured by driving behaviour)?
- How do the different interfaces affect the user experience (measured by NASA TLX workload ratings, user experience (UEQ) and personal ratings collected in a post-hoc questionnaire and an informal interview)?

2.2. Design of the in-vehicle information system (IVIS)

All modern in-vehicle computer systems combine numerous functionalities related to navigation (e.g., traffic reports, navigation assistance), entertainment (e.g., audio, video, communication) and vehicle control (e.g., air conditioning, system information, cruise control). The IVIS used in the experiment simulated the majority of these common features and was accessible through a hierarchical menu structure. The top-most level of the structure was called the “Main menu”. It consisted of five main categories that were further structured into sub-menus of various depths (cf. Fig. 1). Each level of the individual sub-menu consisted of up to eight options. At each level, the user could freely navigate between all the available options, select one of them and enter the corresponding sub-menu or exit and return to the previous menu. Navigation at a given level was “non-circular”, with a virtual barrier at both ends of the menu. Therefore, when users reached the first or the last item of a menu, they could only continue by moving in the opposite direction.

2.3. The visual display

The visual display (V) used in the experiment is depicted in Fig. 2. The recommendations given in (Tretten et al., 2011; Wickens

- **Heating and Cooling**
 - Temperature
 - 18 degrees
 - 20 degrees
 - 22 degrees
 - 24 degrees
 - 26 degrees
 - Seat Warmers
 - Off
 - Low
 - High
 - Fan Speed
 - Off
 - 1
 - 2
 - 3
 - 4
 - 5
- **Entertainment**
- **Communications**
- **Navigation and Traffic**
- **Trip Computer**

Fig. 1. The menu structure of the in-vehicle information system used in the experiment.

et al., 2005; Wittmann et al., 2006; Yoo et al., 1999) were taken into account for the positioning and design of the HUD. Its size was approximately 20 cm × 20 cm, and it was projected to the right-central position of the windshield. When choosing the HUD's position, particular caution was taken to avoid (partly) covering visual information required to steer the vehicle and to physically separate the two sources of information to avoid, or, at least, reduce, cognitive tunnelling effects. However, the HUD was kept in the peripheral field of view so that participants were not required to



Fig. 2. The design and positioning of the HUD used in the visual (V) and audio-visual (AV) conditions.

move their heads to access the information displayed. The items of the menu structure were shown in a high-contrast yellow colour. The current item was highlighted using a slightly increased font size and the colour red. A green coloured title on the top of the HUD indicated the current sub-menu. The background of the HUD was completely transparent.

The HUD was implemented as a “sliding window”, displaying a maximum of five items of the currently active menu or sub-menu. If more than five items were available at a particular level, they became visible by scrolling up or down, thus moving the window accordingly.

2.4. The auditory display

For the auditory display (A), each menu item was automatically played when the user moved to it within the current menu level. The current item could also be replayed on the user's command or skipped if the user moved quickly to the next or previous item. When an individual item was selected, a speech confirmation was played and a new level of the menu was loaded.

The auditory display utilised pre-recorded sound clips generated by the AT&T Labs TTS renderer (AT&T, 2014). A male voice called “Mike” was used for the main menu items, while several other voices were used for various other tasks (voicemail messages, traffic report service, etc.).

2.5. The audio-visual menu

The audio-visual condition (AV) was a combination of the auditory and visual display technique. The menu and each selected item were simultaneously presented to the user as a visual HUD and an auditory display as described in the previous two sections.

2.6. Tasks

Each user was given a set of four tasks for each of the three experimental conditions. To complete an individual task, the user was required to speak the answer aloud to the experimenter. The tasks were grouped into three randomly chosen, low-effort, conventional tasks and one high-effort, complex task consisting of two related subtasks. The second subtask required a user to remember and use a specific piece of information retrieved during the first subtask. The difference between the low- and high-complexity tasks was therefore in the physical activity (i.e., the number of steps) and the cognitive workload required for completing the tasks. The complex tasks were designed based on Lee et al. (1995) decision-making elements. The complete list of low-effort tasks is given in Table 1, while the high-effort tasks are listed in Table 2.

2.7. Participants

A total of 30 test subjects (9 female and 21 male) participated in the study. The subjects ranged in age from 21 to 56 years old

Table 1

A complete list of simple, low-effort tasks.

Task 1	Change the fan speed to 3.
Task 2	Load the game Packman.
Task 3	Check the fuel level.
Task 4	Check the battery level.
Task 5	Set the temperature to 24°.
Task 6	Change the radio station to FM 90.6.
Task 7	Disable the 3G network.
Task 8	Check your average speed.
Task 9	Change the seat heating to “on”.

Table 2

A complete list of complex, high-effort tasks.

Task 1a	You want to travel to New York City. Please check the traffic report for New York City and state the name of the street mentioned in the report to the experimenter.
Task 1b	Set your navigation system to navigate you to New York City. When asked, dictate the name of the street mentioned in the traffic report to the navigation system.
Task 2a	Go to your mailbox and find the message from your friend Mark. Retrieve the name of the song he suggested that you listen to.
Task 2b	Find the suggested song in your music player and play it.
Task 3a	Please retrieve the email from the person named John Smith, and retrieve the name of the restaurant where you should meet him.
Task 3b	Call your friend Kim and leave a message on her mailbox with the suggested name of the restaurant.

($M = 28.9$ years, $SD = 3.5$ years). All the participants had a valid driving license and an average of 11 years of driving experience. All the participants reported normal or corrected-to-normal sight and hearing. To avoid sequence effects and confounds, the study was fully randomised.

2.8. Experimental procedure

After being briefly introduced to the study, the participants were asked to provide information concerning their age, gender, hearing and sight disabilities, and driving experience as well as any prior experience with simulators. The participants were then given a thorough introduction to the experimental environment. Finally, the participants were asked to familiarise themselves with the driving simulator with a 20-minute warm-up drive. They were also given five minutes to freely interact with the IVIS and explore the three different interfaces.

After the participants became accustomed to the experimental environment, they proceeded with the experiment in the sequence determined for their group. Each task was read to the participants aloud before they were given the objective to start solving the tasks as quickly as possible while simultaneously attending to traffic rules and focusing on driving safely. The timing of the second tasks was therefore controlled by the experimenter to provide stable and comparable driving and traffic conditions during all the drives.

During the experiment, the data on the task completion times, physical activity (i.e., the single interaction made with the input device), and driving behaviour was automatically logged for each participant. The measurements began when the participants started performing the tasks and were stopped when the task was successfully completed. The entire user study was recorded with a digital video camera. The recordings were used for a post-hoc evaluation of the driving performance.

After each condition, the participants completed the electronic version of the NASA Task Load Index (NASA-TLX) followed by the User Experience Questionnaire (UEQ) (Laugwitz et al., 2008). After the participants completed all three conditions, they were asked to complete a short, post-study questionnaire. The experiment concluded with an informal interview during which the participants expressed their personal impressions of the experiment.

2.9. Technical setup

A core component of the experimental environment was the City Car Driving version 1.2 driving simulator (Multisoft, 2013). All



Fig. 3. Projection of the City Car Driving simulator used in the experiment.

the participants drove a left-handed Peugeot 206 CC with an automatic transmission. The “Motorway” route with a 10% traffic intensity was used for the “high-speed” route, while the “Modern district” route with a 50% traffic intensity was used as the “low-speed” route. The average driving speed of all the participants was 33 km/h (21 mph) for the “low-speed” route and 72 km/h (45 mph) for the “high-speed” route. A Thrustmaster RGT Force Feedback CLUTCH Racing Wheel was used as the driving simulator’s input device. The driving scene, i.e., the frontal view, was projected on a large screen as depicted in Fig. 3.

The simulated engine sounds and environmental sounds generated by the simulator were played through two Genelec 8030A bi-amplified monitoring system loudspeakers placed on both sides of the projected image. The readings of the menu title and individual menu items were played through two computer speakers. They were placed at the sides to imitate in-vehicle speakers, which are usually mounted in vehicle doors.

The participants could interact with the IVIS by using a small computer mouse attached to the backside of the steering wheel as shown in Fig. 4.

The mouse was placed at a position that allowed the participants to maintain control of the steering wheel while using the mouse without having to look at it or move their hands. Two buttons and a scrolling wheel were available for the user input: The buttons were used to confirm the selected menu item or to exit the current sub-menu and move up one level in the menu structure. The scrolling wheel was used to navigate among the items at a certain level of the menu structure. The functions of both buttons

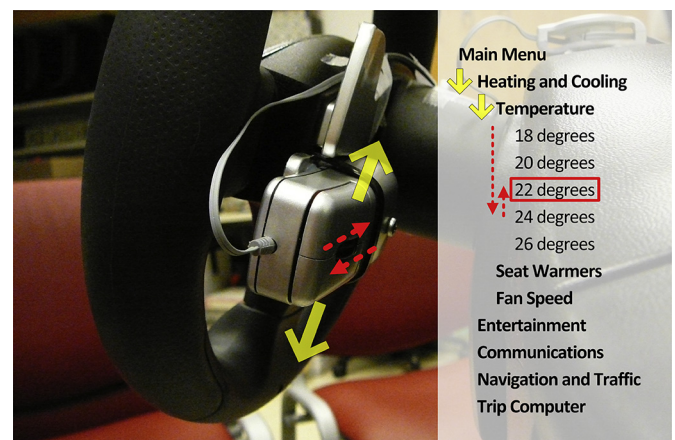


Fig. 4. The hardware device used for interacting with the IVIS menu.

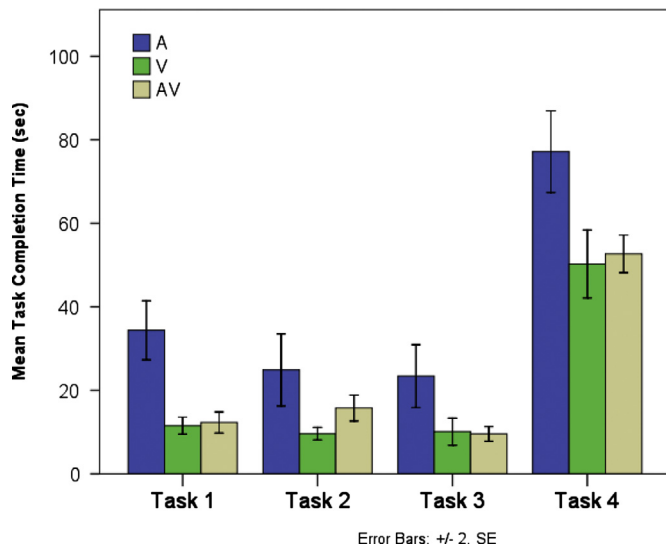


Fig. 5. Average task completion times across all conditions and for all four tasks.

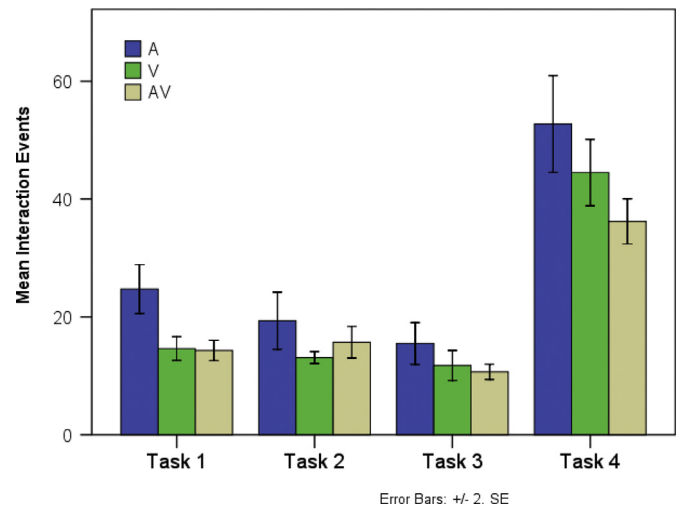


Fig. 6. Average physical interaction events across all conditions and for all four tasks.

and the scrolling direction were adjusted based on the preference of each test subject. When using the auditory display, the scrolling wheel could also act as a third button that caused the title of the current sub-menu to be repeated. This was analogous to the menu in the visual and multi-modal conditions, where the title was displayed on top of the menu at all times.

3. Results and analysis

3.1. The efficiency of the display techniques

In total, 360 trials were analysed (30 participants using three conditions with four tasks each). The average task completion times, as depicted in Fig. 5, were analysed using a Kruskal–Wallis test to confirm the significance of the differences in the task completion times between the three conditions. A non-parametric test was selected instead of a more commonly used ANOVA because of the non-homogeneity of the variances in the data groups. The results are presented in Table 3.

The test confirmed that significant differences existed between the conditions. A post-hoc Games–Howell test with a 0.05 limit on the family-wise error rate confirmed the task completion times to be significantly longer for the auditory display (A) compared to the other two displays (V and AV). No significant difference was found between the V and AV conditions. When separating tasks by their difficulty, the analysis found significant differences between the same interfaces as in the case when tasks were analysed regardless of their difficulty. Therefore, identical conclusions regarding the task completions times also apply separately to the simple and complex tasks.

Fig. 6 shows the average sum of the interaction events (button clicks and wheel turns) required to complete the set of four tasks

for all three conditions. As seen, more effort was required to complete the set of tasks when using the auditory display (A) than in the case of the other two displays (V and AV). The results of a Kruskal–Wallis test and a post-hoc Games–Howell test with a 0.05 limit on the family-wise error rate are given in Table 4. The post-hoc test confirmed the existence of a significant difference between the auditory and the other two displays; however, no significant difference was found between the V and AV conditions.

Similar results were obtained when only simple tasks were analysed. For complex tasks, the AV display proved to be significantly more efficient than the other two interfaces, while no significant difference was found between the A and V conditions.

The results of the task completion times and the interaction activity confirm our initial expectations. The audio-only (A) condition proved to be less efficient than the other two conditions. Generally speaking, the combined auditory and visual information display showed no significant advantage over the visual-only display. An exception is found in the interaction activity required to complete the complex tasks, where the AV display was found to be significantly more efficient than the other two displays.

3.2. Driving performance

The driving performance was evaluated based on the deviation logs created by the simulation software as well as on an additional analysis and revision based on the video recordings of the driving simulations. The video recordings were primarily used to verify the logged deviations and to eliminate those caused purely by other vehicles and the simulated traffic participants as opposed to by the secondary task-solving activities.

Table 3

Results of the Kruskal–Wallis test and the post-hoc Games–Howell test for the task completion times.

	Kruskal–Wallis test	Games–Howell test
Overall	$H = 40.279$, 2 d.f., $p < 0.001$	$p < 0.001$ (A–V, A–AV)
Simple tasks	$H = 45.693$, 2 d.f., $p < 0.001$	$p < 0.001$ (A–V, A–AV)
Complex tasks	$H = 20.787$, 2 d.f., $p < 0.001$	$p < 0.001$ (A–V, A–AV)

Table 4

Results of the Kruskal–Wallis test and the post-hoc Games–Howell test for interaction activity.

	Kruskal–Wallis test	Games–Howell test
Overall	$H = 20.581$, 2 d.f., $p < 0.001$	$p = 0.009$ (A–V), $p < 0.001$ (A–AV)
Simple tasks	$H = 25.353$, 2 d.f., $p < 0.001$	$p < 0.001$ (A–V, A–AV)
Complex tasks	$H = 12.414$, 2 d.f., $p = 0.002$	$p = 0.002$ (AV–A) $p = 0.018$ (AV–V)

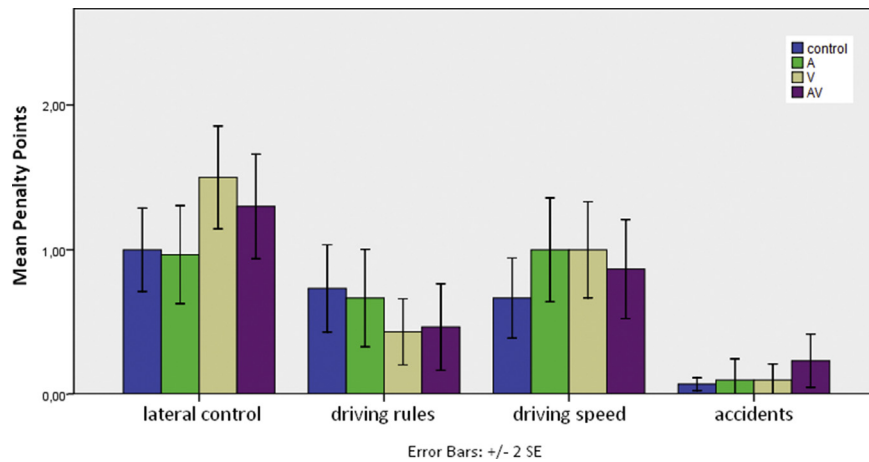


Fig. 7. Mean penalty points for all four categories.

The evaluation was performed for four different categories reflecting four different aspects of driving (examples of deviations are stated for each category):

1. lateral control over the car (e.g., variance in lateral acceleration, abrupt lateral manoeuvres, etc.);
2. driving speed anomalies (e.g., sudden and unnecessary speed reductions, inappropriate speed for a certain driving environment, etc.);
3. violation of traffic rules (e.g., not giving way to pedestrians, driving through red light, etc.); and
4. causing accidents (e.g., crashing into other vehicles, running over pedestrians, etc.).

The performance within each category was rated on a scale of 0–3 points based on the following criteria:

- 0 points – no deviations were observed;
- 1 point – a total of one or two deviations were observed;
- 2 points – a total of three to five deviations were observed; and
- 3 points – a total of more than five deviations were observed (very unsafe driving).

Fig. 7 illustrates the mean penalty points for the four different rating categories and the control group across all the conditions. The control group was defined for the purpose of evaluating the driving performance whereby each subject acts as his or her own control. The data collected for the control group are the driving data collected during the time when the participants did not perform tasks. In this way, the driving performance of the control group is normalised against traffic conditions and learning curves.

The results reveal a significant difference in the total number of penalty points for lateral control (i.e., winding) ($F(3,116) = 3.492$; $p = 0.018$ – the ANOVA tests were used here because of the homogeneity of the variances in the data groups.). A post-hoc Bonferroni test showed a significant difference between the control group and the visual-only (V) condition ($p = 0.049$) and between the audio-only (A) and the visual-only (V) conditions ($p = 0.032$). In both cases, the visual display proved to be less safe and caused lower lateral stability.

3.3. Workload

The perceived workload during the study was measured using the NASA-TLX rating procedure. The NASA-TLX is a multi-

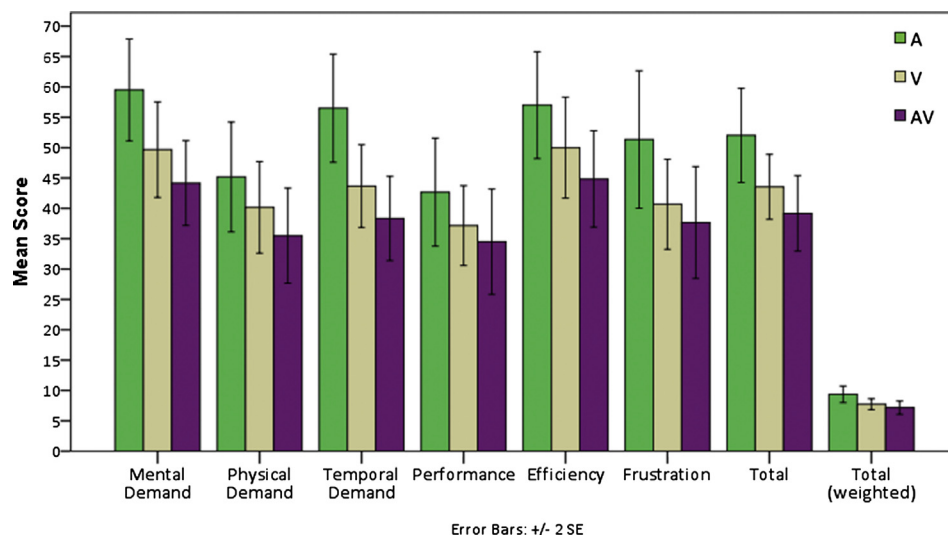


Fig. 8. Mean workload scores for all three conditions. Higher values indicate higher workload and lower performance.

Table 5

Results of the Mann–Whitney test pairwise comparisons of the NASA TLX data.

		A compared to	W	U	r	Z	p
TLX subscales	Mental demand	AV	742.5	277.5	–0.33	–2.56	0.01
	Temporal demand	AV	713.5	248.5	–0.39	–2.99	<0.01
	Temporal demand	V	767.5	302.5	–0.28	–2.19	0.03
	Effort	AV	775.5	310.5	–0.27	–2.07	0.04
TLX total	Total	AV	757.0	292.0	–0.30	–2.34	0.02
	Total, weighted	AV	760.0	295.0	–0.30	–2.29	0.02

dimensional scale designed to obtain subjective workload estimates (Hart and Staveland, 1988). The procedure derives an overall workload score on the basis of a weighted average of ratings on six subscales, which can be found in Fig. 8.

Fig. 8 shows the unweighted subscale and the total workload scores as well as the weighted total score for all the display conditions. Because of the non-homogeneity of the variances in the data groups, a Kruskal–Wallis test was conducted to evaluate the differences in the perceived workloads among the three conditions. The test indicated significant distinctions existed between the conditions. Several follow-up Mann–Whitney *U* tests were conducted to evaluate the pairwise differences among the three conditions, checking for Type I errors across tests by using the Bonferroni approach.

The results of these tests show the differences between the A and AV conditions and between the A and V conditions, as listed in Table 5.

The results demonstrate that the audio-only display (A) had a stronger negative effect on the participants' perceived workload as indicated on three of the subscales when compared to the audio-visual display (AV). These subscales are mental demand, temporal demand and effort. These subjective ratings support the insights gained from the objective measures of the task completion times and the interaction activity; here, the audio-only display necessitated increased task completion times and higher interaction activity as well. A significant difference between A and V, however, could only be found for the temporal demand subscale.

The cumulative workload scores (total and total, weighted) reflect the afore-mentioned significant differences between A and AV with the audio-only display having a significantly stronger negative impact on the perceived workload.

3.4. User experience

To quantify the participants' subjective experience of using the different displays apart from the perceived workload, the

participants were asked to complete the User Experience Questionnaire (UEQ) after each condition. The UEQ is intended to be a user-driven assessment of software quality and usability. It consists of 26 bipolar items, each to be rated on a seven-point Likert scale (1–7). The UEQ allows the experience to be rated using the following six subscales: attractiveness, perspicuity, efficiency, dependability, stimulation and novelty of the display technique.

Fig. 9 shows the UEQ score for each subscale and the overall score across all display techniques. The average scores of the three displays were compared using a Kruskal–Wallis test and a post-hoc Games–Howell test. The analysis showed no significant differences between the displays on the overall scale. However, a number of differences were found in the dependability ($H = 5.84$; 2 d.f.; $p < 0.05$) and stimulation subscales ($H = 5.93$; 2 d.f.; $p < 0.05$). The post-hoc test revealed a significantly higher dependability of the AV display when compared to the A display ($p = 0.049$). This means that the participants found the audio-visual HUD to be more predictable, supportive, and secure and met their expectations more so than the audio-only display.

Furthermore, UEQ results for the stimulation subscale show a significant difference between the V and AV conditions ($p = 0.03$). The participants found the audio-visual display to be more valuable, exciting, interesting and motivating than the visual-only HUD. A general trend towards higher ratings for the AV display can be seen in all the subscales and hence in the overall score. However, no significant differences were found between the displays in the overall UEQ score.

3.5. Subjective preference ratings

In addition to the user experience evaluation, the users were also asked several questions regarding different aspects of the driving simulator and the three different displays. While the UEQ was completed immediately after each condition, the questions listed in the following section were answered after the participants completed all three conditions.

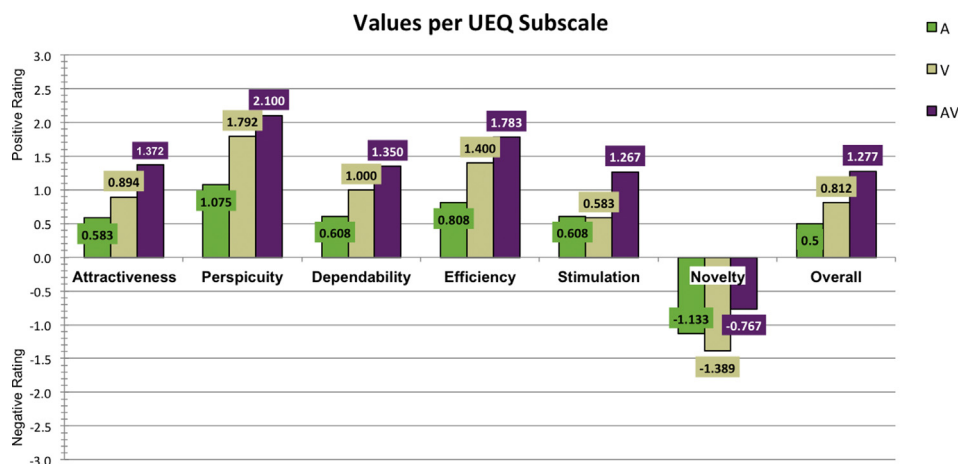


Fig. 9. Mean user experience ratings for all conditions. Positive values indicate affirmative ratings.

Table 6

Median values for the subjective ratings of several aspects of the experiment gathered in a post-hoc questionnaire. Ratings greater than 3.5 show affirmative answers (“I totally agree.”).

	Item	Median
1	I found the driving simulator realistic.	4.5
2	I found the tasks realistic.	6
3	I was able to perform the tasks in a straightforward manner.	6
4	I found it easy to learn to navigate the system.	6
5	I could clearly read all the menu items.	6
6	I found the playback of the menu distracting.	2
7	I found the scrolling wheel attached to the steering wheel easy to use.	5
8	I found the buttons (up and down) attached to the steering wheel confusing.	3
9	Using the simulator made me feel sick.	1

When asked “Which display was the easiest to use?”, the majority ($N = 18$) of the participants chose the audio-visual HUD, while seven chose the auditory, and five the visual display. This distribution of subjective preferences is in line with the user experience ratings, which show a trend towards better ratings for the AV condition. A participant explained why he preferred the AV display with the following statement:

“The third display [the audio-visual HUD] is very easy to use, because you can watch the road all the time, you don't have to look at the menu to see where you are in the menu.”

While the efficiency of the different displays only differs significantly between the A and the other two (V and AV) conditions, the participants' preferences exhibit a clear distinction between the visual and the audio-visual displays.

In addition, the participants were asked to rate the experimental setup and selected tasks on a seven-point Likert scale (1–7). Table 6 lists the items and their median values.

Overall, the participants found the tasks performed in the experiment realistic and performable in a straightforward manner. The simulated HUD received good readability scores, and the menu structure was rated as easy to learn. A number of participants commented on the structure and suggested we should further simplify it, e.g., by offering shortcuts to common tasks, limiting the depth of the structure to only two levels or by adding icons.

Most participants did not feel distracted by the playback of the menu. Among the suggestions for improvements of the audio-only display was the use of spatial sound or the enhancement of the pronunciation of the text-to-speech engine.

The hardware interaction device used for the user input was rated as mostly easy to use; however, the participants noted that it could still be improved in a number of ways. Several participants suggested using a bigger scrolling wheel, removing one button or integrating the device into the steering wheel.

4. Discussion

The interaction with the auditory display proved to be significantly slower and necessitated a higher interaction activity than the interaction with the two display techniques utilising a visual representation of the menu. These results confirm prior findings reported in (Sodnik et al., 2008; Weinberg et al., 2011a, 2011b). The participants had to go through the list item-by-item to obtain an overview of all items in the list to find a specific item. Due to the temporal nature of sound, the participants had to retain the items on the list in their short-term memory in order to obtain and maintain an overview of all items. This contributed to longer task

completion times and was also reflected in the perceived workload ratings, especially in the sub-scores for *mental demand*, *temporal demand* and *effort*, where the audio-only condition showed significantly higher and thus worse rating scores when compared to the audio-visual and visual conditions respectively. During a post-hoc informal interview, a number of participants explained that they found the auditory display especially difficult to use for longer and more complicated tasks, e.g., when they had to recall details of messages from the first part of the complex task.

In contrast to experimental setups in which an experimenter initiates a task, our experiment was designed to simulate a realistic interaction scenario where the decision of when to interact with the IVIS is made by the driver. Consequently, we did not find proof of auditory pre-emption effect, i.e. when a driver is distracted by the sudden onset of auditory playback, which can have negative effects on primary task performance (cf. Wickens et al., 2005). However, it was found that the participants demonstrated reduced lateral stability in both conditions employing a visual display of information. This may be an indication of the cognitive tunnelling effect (cf. Gish and Staplin, 1995; Prinzel and Risser, 2004; Tufano, 1997; Wickens et al., 2005), whereby a driver focuses on the HUD, neglects the changes in the environment and consequently exhibits degraded driving performance. On the other hand, reduced lateral stability may also be a consequence of taking the eyes off the road when reading longer texts. Although the HUD was always visible in the peripheral vision, the details in the text could not be read without changing the focus to the visual information on the windshield. The exact location and dimension of the HUD are therefore very important and require further testing and evaluation.

The audio-visual display showed a clear advantage over the visual-only and audio-only displays in terms of secondary task performance for complex tasks. In the post-hoc questionnaire, the majority of participants reported that they relied heavily on the visual representation of the menu for these tasks. This subjective assessment is supported by the UEQ ratings; the audio-visual HUD was perceived to be more predictable, supportive, and secure than the audio-only display. A number of participants reported that they had focused on the auditory “overlay” of the multi-modal HUD in highly dynamic situations (e.g. when overtaking a vehicle) or in cases when the visual information density was high (e.g. when following unfamiliar routes in the city centre).

While the visual and the audio-visual HUDs showed little differences in the objective performance measurements, the participants clearly favoured the multi-modal display over the two other displays. 60% of the participants found the audio-visual display to be the most useful for solving tasks, while only 16.7% chose the visual and 23.3% the auditory displays respectively. Andre and Wickens (1995) summarize similar performance-preference dissociations found in other HCI related studies. They point out that factors such as colour, dimensionality, familiarity, and simulator fidelity may impact user preferences for interface type, which is supported also by the UEQ ratings.

Interestingly, the participants did not only seem to point out the versatility of the multi-modal display, but they also found it more valuable, exciting, interesting, and motivating as indicated by the UEQ results. It could be argued that these attributions may partly be the result of a novelty effect whereby the results are influenced by the participants' enthusiasm for new technology. However, this is in contrast with the low novelty ratings in the UEQ results for all the display types. The results may also indicate how important it is to include and meet situational requirements during the design process of in-vehicle information system. Also, the multi-modal HUD offers additional degrees of freedom compared to single-modality interfaces that would only support a binary “focus/no

focus” interaction style. The challenge to design information systems for very dynamic situations, such as driving a vehicle, may be best met by offering more freedom to drivers to adapt the systems to their current situational requirements rather than restricting them to a single-modality interface optimised for a narrow range of functionality and task complexity. Similarly, it can be expected that current restrictions on IVIS intended to retain driver safety will be loosened by linking technology, such as MirrorLink (CCC, 2014). By connecting smartphones or tablet computers with the IVIS, displays and GUIs that were not designed for in-vehicle usage may be made accessible in the car. Also in such situations, a multi-modal approach may offer better support for attending to secondary tasks in a broadened ecosystem of applications and information displays.

5. Conclusion

Three display techniques for IVIS were compared for their efficiency in secondary task performance, their impact on the primary driving task, and differences in the perceived user experience.

Overall, the visual and audio-visual head-up displays were more efficient for performing secondary tasks if compared to the audio-only display. We asked the participants to complete tasks by navigating a hierarchical menu structure, which they completed in less time and with less physical effort in both conditions when utilising a visual representation of the menu. All three display types had a similar impact on the primary task performance, i.e. driving a simulated car and abiding by traffic rules.

An important finding of this study is that the participants preferred the audio-visual display because it allowed them to adapt to highly dynamic situations while driving. During complex tasks in which the participants had to extract and memorise information, they could rely on the better-suited modality; they could for example focus on the audio presentation and keep their eyes on the road when the traffic required their visual attention. When the situation allowed, they could scan the text projected on the windshield to quickly extract information.

All interfaces used in this user study were designed and developed in accordance with generally accepted recommendations for this domain and based on similar implementations reported in the related work. However, the presented results must be considered in the context of specific design decisions made in this study. We believe this is a common issue of all user studies with innovative human–computer interfaces, as they require a certain degree of visual or auditory design to be usable. As Andre and Wickens (1995) point out, even minor design decisions can impact the user preference ratings dramatically. Therefore, it is always a challenge to completely eliminate the influence of a specific implementation of an individual interface and to generalize findings to a certain modality.

We used a driving simulator with simulated road users, pedestrians, traffic lights, and crossings to create an authentic driving scenario. However, important aspects that may impact primary and secondary task performance cannot be simulated and must be evaluated by an in-situ driving study. One of such issues is also the impact of the drivers' age on the overall performance. Future studies should also consider different variants of the HUD design and how drivers interact with the system.

To conclude, although current research findings are promising, additional research is needed to improve multi-modal display techniques for in-vehicle information systems. This includes insights into how, when and for which tasks the drivers want to interact with an IVIS and which modality is preferred for each condition. In this study, we mainly addressed the research questions concerning the display of information; however, investigating the methods of user interaction with the system is equally important.

References

- AT&T, 2014. AT&T Natural Voices® Text-to-Speech Demo. AT&T. <http://www2.research.att.com/~ttsweb/tts/demo.php> (Accessed June 12).
- Ablassmeier, M., Poitschke, T., Wallhoff, F., Bengler, K., Rigoll, G., 2007. Eye gaze studies comparing head-up and head-down displays in vehicles. In: Proc. of ICME 2007, pp. 2250–2252.
- Andre, A.D., Wickens, C.D., 1995. When users want what's not best for them. *Ergon. Des. Q. Hum. Fact. Appl.* 3, 10–14.
- Biever, W.J., 2002. Auditory-based Supplemental Information Processing Demand Effects on Driving Performance. Master's Thesis. Virginia Tech.
- Burnett, G.E., 2003. A road-based evaluation of a head-up display for presenting navigation information. In: Proc. of the Tenth International Conference on Human-Computer Interaction. Lawrence Erlbaum Associates, pp. 180–184.
- CCC, 2014. “MirrorLink.” Car Connectivity Consortium. <http://www.mirrorlink.com/> (Accessed June 12).
- Charissis, V., Papanastasiou, S., Vlachos, G., 2008. Comparative study of prototype automotive HUD vs. HDD: collision avoidance simulation and results. In: In Proceedings of the Society of Automotive Engineers World Congress 2008.
- Gish, K.W., Staplin, L., 1995. Human Factors Aspects of Using Head up Displays in Automobiles: A Review of the Literature. Report No. DOT HS 808 320. Department of Transportation, Washington: U.S.
- Haines, R.F., 1991. A breakdown in simultaneous information processing. In: Obrecht, G., Stark, L. (Eds.), *Presbyopia Research*. Plenum, New York, pp. 171–175.
- Hart, S.G., Staveland, L.E., 1988. Development of a multi-dimensional workload rating scale: results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (Eds.), *Human Mental Workload*. Elsevier, Amsterdam, pp. 139–183.
- Horrey, W.J., Wickens, C.D., 2004. Driving and side task performance: the effects of display clutter, separation, and modality. *Hum. Fact. J. Hum. Fact. Ergon. Soc.* 46 (4), 611–624.
- Lai, J., Cheng, K., Green, P., Tsimhoni, O., 2001. On the road and on the web?: Comprehension of synthetic and human speech while driving. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems ACM, pp. 206–212.
- Laugwitz, B., Held, T., Schrepp, M., 2008. Construction and evaluation of a user experience questionnaire. In: *HCI and Usability for Education and Work*. Springer, Berlin Heidelberg, pp. 63–76.
- Lee, J.D., Morgan, J., Wheeler, W.A., Hulse, M.C., Dingus, T.A., 1995. Development of Human Factors Guidelines for Advanced Traveler Information Systems and Commercial Vehicles: ATIS and CVO Functional Description. Federal Highway Administration, Washington, DC.
- Liu, Y.-C., 2001. Comparative study of the effects of auditory, visual and multi-modality displays on drivers' performance in advanced traveller information systems. *Ergonomics* 44 (4), 425–442.
- Liu, Y.-C., Wen, M.-H., 2004. Comparison of head-up display (HUD) vs. head-down display (HDD): driving performance of commercial vehicle operators in Taiwan. *Int. J. Hum. Comp. Stud.* 61 (5), 679–697.
- Multisoft, 2013. City Car Driving 1.2 Description. Multisoft. <http://citycardriving.com/products/citycardriving12> (accessed April 16).
- Prinzel, L.J., Risser, M., 2004. Head-up Displays and Attention Capture. Tech. Rep. NASA/TM-2004-213000. NASA – Langley Research Center, Hampton.
- Sodnik, J., Dicke, C., Tomazic, S., Billingshurst, M., 2008. A user study of auditory versus visual interfaces for use while driving. *Int. J. Hum. Comp. Stud.* 66 (5), 318–332.
- Sojourner, R.J., Antin, J.F., 1990. The effects of a simulated head-up display speedometer on perceptual task performance. *Hum. Factors* 32 (3), 329–339.
- Thomas, L.C., Wickens, C.D., 2001. Visual displays and cognitive tunneling: frames of reference effects on spatial judgments and change detection. In: Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting, pp. 336–340.
- Tretten, P., Gärling, A., Nilsson, R., Larsson, T., 2011. An on-road study of head-up display: preferred location and acceptance levels. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 55(1), pp. 1914–1918.
- Tufano, D., 1997. Automotive HUDs: the overlooked safety issues. *Hum. Factors* 39 (2), 303–311.
- Weinberg, G., Harsham, B., Medenica, Z., 2011a. Evaluating the usability of a head-up display for selection from choice lists in cars. In: Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI), pp. 39–46.
- Weinberg, G., Harsham, B., Medenica, Z., 2011b. Investigating HUDs for the presentation of choice lists in car navigation systems. In: Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design (DAC), pp. 195–202.
- Wickens, C.D., Dixon, S.R., Seppelt, B., 2005. Auditory preemption versus multiple resources: who wins in interruption management?. In: Proc. of the Human Factors and Ergonomics Society 49th Annual Meeting, pp. 463–467.
- Wittmann, M., Kiss, M., Gugg, P., Steffen, A., Fink, M., Pöppel, E., Kamiya, H., 2006. Effects of display position of a visual in-vehicle task on simulated driving. *Appl. Ergon.* 37 (2), 187–199.
- Yoo, H., Tsimhoni, O., Watanabe, H., Green, P., Shah, R., 1999. Display of HUD Warnings to Drivers: Determining an Optimal Location. Tech. Rep. UMTRI-99-9. The University of Michigan, Transportation Research Institute (UMTRI), Ann Arbor.
- Yung-Ching, L., 2003. Effects of using head-up display in automobile context on attention demand and driving performance. *Displays* 24 (4–5), 157–165.