

Studying Driver Attention and Behaviour for Three Configurations of GPS Navigation in Real Traffic Driving

Brit Susan Jensen, Mikael B. Skov, Nissanthan Thiruravichandran

HCI Lab, Department of Computer Science, Aalborg University

Selma Lagerlöfs Vej 300, 9220 Aalborg East, Denmark

britjensen@gmail.com, dubois@cs.aau.dk, nissanthan@gmail.com

ABSTRACT

Global Positioning System (GPS) navigation systems were amongst the top selling consumer technologies in 2008 and research has indicated that such technologies could affect driving behaviour. In this paper, we study how different output configurations (audio, visual and audio-visual) of a GPS system affect driving behaviour and performance. We conducted field experiments in real traffic with 30 subjects. Our results illustrated that visual output not only causes a substantial amount of eye glances, but also led to a decrease in driving performance. Adding audio output decreased the number of eye glances, but we found no significant effects on driving performance. Although the audio configuration implied much fewer eye glances and improved driving performance, several participants expressed preference for the audio/visual output.

Author Keywords

GPS, navigation guide, in-vehicle systems, eye glances, output modalities, driving, field experiment

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Human Factors

INTRODUCTION

The market of in-vehicle systems has grown exponentially over the last three decades – a proliferation, which was set in motion by a reduction in hardware costs as well as innovation in communication and information technology [2, 4]. In-vehicle systems serve a variety of purposes, e.g. navigational guidance, media players, climate controls, and communication. The development of in-vehicle systems has initiated debates and inspired research on road and driving safety. In-vehicle systems may provide compelling means

to enhance mobility [4, 23], but research has shown that these systems may distract the driver and hereby divert focus from the primary task of driving, which could lead to driving accidents [2, 11, 13]. Also, in-vehicle systems have become increasingly sophisticated due to the integration of advance functionalities, novel interaction techniques, and emerging wireless network infrastructures. This progression may present the driver with tasks that are unrelated to the driving task (secondary tasks), which in hand may require high attentive interaction.

Global Positioning System (GPS) navigation systems were amongst the top selling consumer technologies in 2008. In the recent years the GPS has become a subject of research. Current research on GPS systems has shed light on the way in which the utilization of these navigation systems may alter driving practices and affect the way people understand the environment in which they navigate [22]. Studies have also addressed important usability aspects by evaluating the learn ability and memorability of a GPS system in order to identify problems associated with first-time and infrequent use [18]. Furthermore, additional research studies focus on how navigational information should be presented in order to enhance user comprehension or satisfaction [16, 23].

However, we still have a rather limited understanding of GPS navigation systems on their effect on driver attention and driving behaviour [13, 14]. The inherent challenge of dividing attention of the driver between road information and GPS navigation directions calls for further experiments and studies how to design such GPS systems. Particularly we need to understand how navigational outputs influence driving performance [16] and especially if we consider how secondary task performance can lead to various kinds of accidents [3, 11]. Thus, we need research that investigates how such systems are used in natural contexts, i.e. as field studies involving real traffic driving [8].

In this paper, we investigate different output modalities of a GPS navigation guide on their effects on driving behaviour and performance by comparing three output configurations; audio, visual and audio-visual. The paper is structured as follows. First, we present previous research on in-vehicle systems research. Secondly, we outline our experiment with the adapted measures used to investigate driver behaviour and driving performance. Next, we present results from the experiment and finally, we discuss the results.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.

Copyright 2010 ACM 978-1-60558-929-9/10/04....\$10.00.

RELATED WORK

The literature on in-vehicle systems research is particularly concerned with effects of such systems on driver attention, driver distraction, and driving behaviour [1, 8, 14, 17, 25]. Research studies separate between different types of driver distraction where withdrawal of attention from the driving scene is the most commonly used. According to Brown [4], withdrawal of attention can either be general or selective. The general withdrawal of attention (or *eyes-off-the-road*) refers to the insufficient (visual) perception of the driving information, e.g. when drivers operate a media player while glancing at it [13]. Selective withdrawal of attention (or *mind-off-the-road*) is a more subtle type of distraction as it involves e.g. perceptual interpretation or decision selection [15]. Horberry et al. [17] verify that the level of complexity of a given secondary task is correlated with the level of driver distraction. Drivers are required to have their eyes on the road in order to ensure safe driving; hence secondary tasks involving visual attention can induce safety risks.

Research studies have attempted to reduce driver distraction through novel means of interaction to support secondary tasks. Geiger et al. [11] conducted a comparative study in which they evaluated the use of a tactile interface and a gesture-based interface for secondary tasks. Their findings showed that the use of the tactile interface entailed higher task completion times and lower recognition performance. However, the gesture-based interface enabled the drivers to perform secondary tasks more accurately, and they further perceived the gesture-based interface to be less distracting.

Bach et al. [2] compared three interaction techniques for a media player – a tactile interface, a touch interface, and a gesture-based interface. The interaction techniques were compared in two complementary experiments. The findings showed an inclination towards the gesture-based interface. The touch interface allowed drivers to complete secondary tasks with significantly fewer eye glances in comparison to the tactile and touch interfaces. The tactile interface lacked intuitiveness; the system demanded perceptual resources in order to be operated and hereby diverting attention from the primary task. The touch interface introduced a lower task completion time and fewer interaction errors, in comparison to the other two interfaces.

While novel input techniques are important contributions to the field of in-vehicle systems research, it is also essential to consider potential opportunities and limitations of output modalities [2, 16, 23]. Green et al. [16] evaluated four configurations of a GPS navigation guide – auditory, auditory with landmarks, visual and visual with landmarks – in a simulated driving setting. The aim of the study was to clarify how much attention the systems required and how the participants perceived the use of the systems. The participants watched a video recording of a driver's view and received route guidance information. They were asked to press a button when they saw the intersections described

by the GPS navigation guide. The study showed that the auditory configuration required less attention and gave the lowest driver reaction time when compared to the visual configuration. When supplementing the audio and visual configurations with information on landmarks no penalties to driver attention incurred. The participants predominantly favoured configurations that included landmarks as opposed to the ones without landmarks. They generally expressed a slight inclination towards visual output.

A recent empirical study by Moldenhauer and McCrickard [23] partly confirms the results by Green et al. [16]. They investigated trade-offs involving information conveyance by evaluating four information modalities – audio, audio with overhead map, visual and visual with overhead map – in a driving simulator. Their results showed that the visual modality with an overhead map resulted in the highest number of driving errors and highest reaction time, whereas participants expressed that the information provided by the audio-based modalities were more difficult to comprehend. They acknowledged that a more immersive setting could provide further validation of the results.

While the above studies provide insights into the potential opportunities and limitations of different output modalities of in-vehicle systems, we still need further investigations on such modalities and especially we need studies conducted in real traffic driving and not only driving simulators. Bach et al. [1] found that a considerable part of in-vehicle system research is conducted in either driving simulators or as controlled driving situations. Especially, we need studies on GPS navigation guides focusing on output modalities, as these systems are highly output-oriented towards the driver and less input-oriented. Leshed et al. [22] conducted a study on the use of GPS guides in real traffic driving where they argue that GPS navigation sometimes can disengage people from their surrounding environment or context, but they also argue that these systems have the potential to open up novel ways to engage with the context. While their study provides design implications, the focus is on content of GPS guides rather than interaction and output modalities.

Inspired by previous research we evaluated the navigational output provided by a GPS system. As stated above, the majority of in-vehicle systems research has been conducted in controlled and simulated settings [1, 16, 23]. We chose to conduct the experiment in the field (in real traffic driving) due to the inadequacy of immersive settings within the area of in-vehicle research.

EXPERIMENT

The purpose of our experiment was to study how different output modalities of a GPS navigation guide affect drivers and driving performance during real traffic driving. We included three output modalities namely audio, visual, and audio-visual. In the following, we outline the experiment, the participants, GPS system, procedure, and data analysis.

Participants

30 people ranging between 21 – 38 years of age ($\bar{x}$ = 25.2, SD = 2.65) participated in the experiment. All participants (7 women and 23 men) carried valid driver licenses and had between 3 and 19 years of driving experience ($\bar{x}$ = 6.85, SD = 2.71). They drove by their own estimate between 0 – 40,000 kilometres per year ($\bar{x}$ = 7598.33, SD = 8557.9). On the basis of self-assessment – 9 participants indicated that they had poor knowledge of Aalborg and its surroundings, 15 indicated basic knowledge, and finally 6 claimed that they had good knowledge of Aalborg. The average amount of kilometres driven per year was equalled out between the three participant groups. Each participant group indicated equal acquaintance with the greater Aalborg area.

Experimental Design

We utilized a between-subject experimental design using output modality configurations (audio, visual, audio-visual) as independent variables and dependent variables included primary driving task performance (longitudinal control, lateral control, traffic violations), secondary driving task performance (navigational errors and task completion time) and eye glance behaviour (below 0.5 seconds, 0.5 - 2.0 seconds, above 2 seconds).

GPS Navigation Guide

The GPS navigation guide used in the experiment was a TOMTOM GO 930, which at the time of the experiment was a state-of-the-art model that had received favourable reviews [7]. The GPS came pre-installed with maps and a POI (Points of Interests) covering Europe.

The main input interface is a 4.3" touch screen with a 480 x 272 pixel resolution for visual output and an internal speaker for audio output. With participants having Danish as their native language, we selected the language setting to match this for both visual and audio output. The three test configurations (audio, visual, audio-visual) consisted of different combinations of output modalities. The audio configuration consisted exclusively of the system's audio output, while the visual configuration consisted exclusively of the system's visual output. The configuration on audio-visual comprised both output modalities.

Audio output consists of navigational instructions presented through pre-recorded speech (hence no speech output for street names was available) using a female voice. Nixon et al. found that female voices are easier to hear in noisy environments [24]. Each instruction included an estimated distance and a direction – for example *'after 200 meters, turn left'* – followed by a repetition of the direction. If there was a need to perform a sequence of turns (within 200 meters of the first turn) this would be included in the instructions – for example *'after 200 meters, turn right and then turn left'*. On longer stretches of road (over 500 meters) the system would add an additional reminder. Visual output in the GPS system consisted of the 'driving

view' (illustrated in figure 1). This screen consists of a 3D map showing the current part of the route. The selected route is marked with red line, manoeuvres are illustrated with green icons and the current position of the vehicle is shown with a blue arrow. The bottom part of the screen shows navigation instructions including distance, estimated arrival time, signal strength among others.

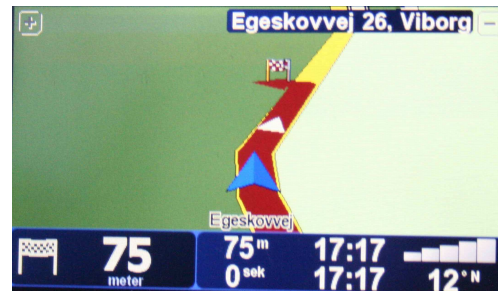


Figure 1. Example of the GPS in driving view. The map is illustrated in 3D. Included in the map are street names, POI icons and zoom options. Navigation and system information are shown on the bottom part of the screen.

The GPS navigation guide was placed at the lower centre of the windscreen (as recommended in the instruction booklet) for all configurations (shown in figure 2). This ensured that participants using visual output unobstructed view of the visual output. For the sessions involving audio output, the internal speaker volume was set to 75%. None of the participants found it necessary to adjust the volume during the experiment. Furthermore, the GPS guide was slanted in a manner, which ensured that the participants were only able to receive auditory output.

Tasks

The experiment comprised four scenario-driven tasks – the tasks involved driving to predetermined locations and collecting associates of the University – for example "Collect Lisa Nielsen who lives on Poseidonvej 15, 9210". By applying scenario-driven tasks, we sought to promote a natural setting for the field trials. The GPS system served as an optional component allowing participants to approach a given task unassisted.

Procedure

We conducted our experiment as a field experiment in real traffic. All participants drove vehicles (C-segment – small family cars) equipped with manual transmissions. The field trials were conducted during daytime and in good weather conditions. As such, we strived for consistency between the sessions and to minimize safety hazards. We conducted two pilot tests to test and adjust the equipment.

The participants were classified as one of the following user types – GPS system users or non-users – the decisive factor for this classification process was based on whether or not the participants had used GPS systems previously. The participants were randomly assigned to a configuration until

an equal distribution of GPS system users (15) and non-users (15) were attained – we assigned five GPS system users and five non-users to each of the three configurations (which constitute three groups of ten). We ensured that each group had at least one female participant (audio = 3, visual = 3, audio-visual = 1). We collected demographic data of the participants through an interview.

Before the driving sessions, participants were introduced to the experiment. First, we introduced the GPS system and the car and we stressed the fact that they had to adhere to the general traffic rules. Participants were told that they were supposed to drive around the greater Aalborg area to pick up different people. They were given oral instructions of addresses to visit and had to type in these addresses in the GPS system. Secondly, participants were allowed to take a test drive to familiarize them with the car. Also, we stressed that we were testing different configurations of the GPS system and not their performance and they were told that they could pull over at any time while driving if they needed a break. Finally, participants were asked to sign an information and consent form.



Figure 2. Experimental setup with the placement of the GPS system in the lower-middle area of the windscreen (as seen from the driver's point of view).

All field trials started at the Computer Science Department at Aalborg University. The participants were not given tasks during driving – each task was presented prior to the associated driving segment. The estimated length of the entire route was 16 kilometres – the segments comprised both rural and densely populated areas in order to expose the participants to varied traffic environments and areas of Aalborg, which they may either be familiar or unacquainted with. The permitted speed limit ranged from 30 – 80 kilometres per hour in the four driving segments – we avoided motorways due to safety concerns.

All field trials were filmed using two camcorders – one of the camcorders was mounted on the dashboard in order to capture eye glances. The second camcorder was affixed on the front passenger seat to record lateral and longitudinal control errors, and driver view (as seen through the front

windscreen). We refrained from asking questions during the field trials. Dialogues only took place when participants initiated a conversation. The test manager was sitting next to the participant and a logger was sitting at the back seat. The test manager ensured that the experiment proceeded as intended while the logger collected qualitative data, e.g. verbal comments from participants.

The participants were debriefed as a concluding segment of the experiment. We conducted a semi-structured interview where participants were asked about subjective matters.

Data Analysis

Inspired by previous research studies [1, 2, 12, 20, 27], we integrated several measures for understanding driving behaviour and driving performance. Our measures included 1) primary driving task performance (longitudinal and lateral control errors, traffic light and directional indicator violations), 2) secondary driving task performance (task completion times and navigational errors), and also 3) eye glance behaviour (below 0.5 seconds, 0.5 – 2.0 seconds, above 2.0 seconds).

We analyzed the 30 video recordings (one for each session) both individually and collaboratively. One recording was omitted (only for the eye glance measure analysis) due to incomplete data collection caused by a technical error in the camera. The data analysis was done in three steps. First, we initially analyzed three randomly chosen video recordings collaboratively in order to establish guidelines and metrics for the subsequent individual video analysis. Secondly, two authors analyzed 16 video recordings individually reporting incidents on the selected measures. We randomly selected ten of the recordings for analysis of both authors to ensure procedural consistency. The produced incident lists of these ten recordings were compared and an inter-rate reliability test (weighted Cohen's Kappa) gave $K=0.75$ corresponding to a substantial agreement according to Fleiss et al. [10]. Thirdly, we compared and merged all incident lists into one combined list. If disagreeing, the authors would analyze the recordings once more to determine whether the concerned incident was valid or not. In the following, we outline the included measures in more detail.

1) The first measure was primary driving task performance and involved several variables [1, 2, 27]. We integrated longitudinal and lateral control errors, traffic lights crossing and directional indicator violations. Longitudinal control errors were defined as speeding violations in three different categories (inspired by current Danish legislation) namely speeding level 1, 2, and 3. Speeding level 1 was defined as participants exceeding the prescribed speed limit by three kilometres per hour (implies a speeding fine), speeding level 2 was when participants exceeded the speed limit by 30 percent (implies endorsement of license – one penalty point), and finally speeding level 3 was when participants exceeded the prescribed speed limit by 60 percent (implies

loosing the driving license). Lateral control errors denote loss of lateral vehicle control, i.e. lane excursions. We also recorded incidents where participants did not adhere to the stop signals assigned by traffic lights and incidents where participants failed to activate the directional indicator as required by Danish traffic regulations.

2) Secondary driving task performance denotes incidents in which participants diverged from the specified route due to misinterpretation of the navigational information provided by the GPS system. We also recorded the completion times for each of the four tasks.

3) Eye glance behaviour is a commonly applied measure for driver attention [12]. We adapted the coding of Bach et al. [2] where eye glances were measured from three categories namely below 0.5 seconds (not eye fixation [29]), between 0.5 and 2.0 seconds, and above 2.0 seconds (drivers are usually reluctant to continue without roadway information for more than 2 seconds [25]). Our coding involved counting recorded video frames where category one (below 0.5 seconds) corresponds to less than 14 frames, category 2 (0.5 – 2.0 seconds) corresponds to 14 – 50 frames, and category 3 (above 2.0 seconds) is more than 50 frames.

RESULTS

In this section we present the results from the experiment. First, we will present primary driving task performance including longitudinal and lateral control errors, next we present secondary driving task performance and finally eye glance behaviour. The results were subjected to one-way independent-samples ANOVA tests and Tukey HSD post hoc tests. The results are presented in tables 1, 2, 3, and 4.

Primary Driving Task Performance

We integrated metrics for measuring primary driving task performance primarily longitudinal control errors (speeding violations) but also lateral control errors (lane excursions), directional indicators, and traffic light violations.

Our results showed that visual and audio-visual participants in general had many more longitudinal control errors (i.e. speeding violations) than audio participants and that they had more lateral control errors (i.e. lane excursions) during driving. In total, we identified 647 violations concerning primary driving task performance across all configurations (Audio = 110, Visual = 265, Audio-Visual = 272). Of the 647 primary driving task violations, 522 are classified as longitudinal control errors (illustrated in table 1). When assessing speeding violations we identified some major differences between the three configurations.

Audio participants had significantly fewer speeding level 1 violations (> 3 km/t) than both visual and audio-visual participants. Our experiment showed that participants using the audio configuration on average had 8.8 violations during the trails whereas visual participants on average had 17.9 violations and audio-visual participants had 19.3

violations. An ANOVA test showed significant difference among the three configurations, $F(2,27) = 6.67$, $p < 0.01$. A Tukey HSD post hoc test showed that audio-visual had significantly more speeding violations than audio ($p < 0.01$) and audio had significant fewer violations than visual ($p < 0.05$). A comparison between the visual and audio-visual configurations showed no significant difference.

When looking at speeding level 2 violations (exceeding the allowed speed limit by more than 30 percent), we found a similar pattern as with speeding 1 violations. Audio-visual participants had more than 6 times as many speeding 2 violations than audio participants. The audio configuration had on average 0.6 violations while visual participants had 2.2 and participants in the audio-visual configuration had 3.3 violations, $F(2,27) = 5.78$, $p < 0.01$. Not surprisingly, a post-hoc test confirmed that audio-visual participants had significantly more violations than audio participants ($p < 0.01$). On the other hand, while visual participants had more than 3 times as many eye glances as the audio participants this difference was not significant. Finally, we found no statistically significant differences between the visual and audio-visual configurations.

We identified only one speeding level 3 violation incident (exceeding the prescribed speed limit by 60 percent) in the visual configuration. Thus, no significant or key differences were found for this speeding violation.

	Audio (N=10)	Visual (N=10)	Audio-Visual (N=10)
Speeding (> 3 km/t) (N=460)	8.8 (4.16) -	17.9 (8.03) +	19.3 (8.17) +
Speeding ($> 30\%$) (N=61)	0.6 (0.97) -	2.2 (1.81)	3.3 (2.31) +
Speeding ($> 60\%$) (N=1)	0.0 (0.0)	0.1 (0.32)	0.00 (0.00)
Total (N=522)	9.4 (4.81) -	20.2 (8.87) +	22.6 (9.03) +

Table 1. Speeding violations (longitudinal control errors) for the three configurations. Standard deviations are given in parentheses. A plus denotes a significant difference at the 1% or 5% significance levels to the configuration marked with a minus.

The three speeding violation levels showed key differences between the three configurations of the GPS navigation guide and hence, this is also illustrated in the total number of speeding violations. First, we saw that participants using the audio configuration had fewer speeding violations in total in comparison to participants using the visual and audio-visual configurations. This may not come as a major surprise when the results from the experiment show that the visual (accounts for 202 speeding violations) and audio-visual configurations (accounts for 226 speeding violations) both have more than twice the speeding violations in comparison to the audio configuration (accounts for 94 speeding violations).

Our participants using the audio configuration had an average of 9.4 speeding incidents, participants using the visual configuration had 20.2, while finally the participants using the audio-visual had 22.6, $F(2, 27) = 8.09$, $p < 0.01$. A Tukey HSD post hoc test showed significant difference at the 5% level between audio and visual participants as well as strong significant difference at the 1% level between audio and audio-visual participants. When re-assessing the incidents related to longitudinal control we see that there are no significant differences between the audio-visual and visual configurations – both these configurations constitute almost an equal number of speeding violations.

	Audio (N=10)	Visual (N=10)	Audio-Visual (N=10)
Lateral control errors (N=77)	0.4 (0.7) -	4.3 (3.77) +	3.0 (3.06)
Directional Indicators (N=40)	1.0 (0.94)	1.7 (1.83)	1.3 (1.06)
Traffic Light (N=8)	0.2 (0.42)	0.3 (0.67)	0.3 (0.48)

Table 2. Additional primary driving task performance measures for each output configuration. Standard deviation is given in parentheses. A plus denotes a significant difference at the 1% or 5% significance levels to the configuration marked with minus.

Besides speeding violations, we further analyzed additional primary driving task performance measures including lateral control errors (lane excursions), directional indicator errors, and traffic light crossing violations (see table 2). In total, we identified 77 lateral control errors. When comparing the lateral control errors, we see a difference as participants in the visual and audio-visual configurations collectively constituted 95% of all errors. Participants in the audio configuration accounted for only two errors.

Our experiment showed that participants using the audio configuration on average had 0.4 lateral control errors, participants using the visual configuration had 4.3, while participants using audio-visual had 3.0, $F(2,27) = 4.92$, $p < 0.05$. Using a post-hoc test we found that this difference lies between the audio and visual configurations where the participants using the visual configuration constitute a total of 43 lateral control errors, which is significantly more than the number of lateral control errors committed by audio participants ($p < 0.05$). Somewhat surprisingly there is no significant difference between audio and audio-visual despite the fact that the audio-visual participants had more than seven times as many incidents.

We further compared the results on violations related to traffic light (not adhering to the caution and stop signals assigned by traffic lights) and the directional indicators (failing to activate the directional indicators). We found no significant differences between the three configurations where participants using the audio configuration had an average of 1.0 violations, participants using a visual configuration had 1.7 violations while participants using the

audio-visual configuration had 1.3 violations, $F(2,27) = 0.69$, $p > 0.51$. When assessing traffic light violations, the results show that the number of incidents is almost equally distributed amongst the three configurations.

Secondary Driving Task Performance

When assessing secondary driving task performance, we used the following measurement variables; task completion time and navigational errors. Our results revealed that the three configurations were relatively alike and participants were rather alike within each configuration especially for task completion time (shown in table 3). Task completion times for each of the four tasks showed an average of 24.13 minutes ($SD=1:44$) for audio users, 22.55 minutes ($SD=0:58$) for visual and 23.05 minutes ($SD=1:06$) for audio-visual users. We also identified 34 navigational errors (Audio = 10, Visual = 14, Audio-Visual = 10). The results do not reveal any significant differences when compared across the three configurations.

	Audio (N=10)	Visual (N=10)	Audio-Visual (N=10)
Navigational Errors (N=34)	1.0 (1.25)	1.4 (1.84)	1.0 (0.82)
Task Completion Times (mm:ss)	24:13 (1:44)	22:55 (0:58)	23:05 (1:06)

Table 3. Navigational errors and task completion times as mean values. Standard deviation is given in parentheses.

Eye Glance Behaviour

Eye glances were categorized into three categories based on previous research [2, 26, 27]. These include category one glances (below 0.5 seconds), category two glances (0.5 – 2.0 seconds) and category 3 glances (above 2.0 seconds). Due to technical problems during the trials, the audio-visual configuration included one participant where eye glances could not be identified.

Our results revealed that participants in both the visual and audio-visual configurations had many more eye glances than audio participants (as shown in table 4). While this is not surprising in itself, the rather high number of glances for visual and audio-visual is surprising. We identified a total of 3977 glances within all the field trials.

For the category one glances (below 0.5 seconds), our experiment showed that participants on average had 6.6 in the audio configuration, 45.4 in visual and 51.67 glances in audio-visual. A one-way ANOVA test confirms that there is a significant differences among the three configurations, $F(2,26) = 14.65$, $p < 0.001$. Subjecting these results to a Tukey HSD post hoc test showed that participants using the audio configuration have significantly less category one glances than participants in the visual and audio-visual configurations, $p < 0.01$. This may not come as a surprise when the results show that audio participants only account for 7% of the recorded glances in the category one variable. The remaining occurrences in category one are almost

equally distributed between the visual and audio-visual configurations; hence we found no significant differences.

	Audio (N=10)	Visual (N=10)	Audio-Visual (N=9)
< 0.5 sec (N=985)	6.6 (4.55) -	45.4 (22.67) +	51.67 (26.21) ±
0.5 – 2 sec (N=2905)	2.8 (2.35) -	178.3 (33.35) +	121.56 (38.11) +
> 2 sec (N=87)	0.0 (0.0) -	6.7 (4.37) +	2.22 (2.86) -
Total (N=3977)	9.4 (5.93) -	230.4 (37.47) +	175.44 (64.24) ±

Table 4. Eye glance behaviour for the three configurations. Standard deviation is given in parentheses. A plus denotes a significant difference at 1% or 5% significance levels to the configuration marked with minus.

When looking at category two glances (0.5 - 2 seconds), our experiment revealed a surprising total of 2905 glances (Audio = 28, Visual = 1783, Audio-Visual = 1094) and we can see that participants in both the visual and audio-visual configurations had an extremely high number of glances. The audio participants had on average 2.8 glances, 178.3 for participants using visual and 121.56 for participants using audio-visual, $F(2,26) = 95.93$, $p < 0.001$. A post-hoc test showed significant differences at the 1% level between the visual configuration and the two other configurations. This is also reflected in the results where participants using the visual configuration accounted for 73% of all category two glances, whereas audio participants accounted for less than 1%. The post-hoc test also revealed a significant difference between audio and audio-visual, where the audio configuration (28 incidents) has nearly none compared to audio-visual (1094 incidents), $p < 0.01$.

When assessing category three glances (above 2 seconds), we found a total of 87 glances (Audio = 0, Visual = 67, Audio-Visual = 20), where audio-visual participants on average had 2.33, participants using visual had 6.7 and audio had 0.0. A one-way ANOVA test showed that the difference among the configurations is significant, $F(2,26) = 12.71$, $p < 0.001$. The visual configuration participants accounted for 76% of all category three glances. A post-hoc test showed that participants using visual have significantly more category three glances than participants in the audio configuration ($p < 0.01$). A comparison of the visual and audio-visual configurations reveals a significant difference, where participants using visual have more category three glances than audio-visual ($p < 0.01$). Interestingly, we saw approximately the same glance ratio between the visual and the audio-visual configurations as for category two glances, participants using the visual configuration again accounted for three times as many glances as audio-visual participants.

Not surprisingly, we identified large differences on total number of glances between the three configurations. From the total of 3977 glances, 94 glances occurred in the audio configuration, 2304 in visual and 1579 in audio-visual.

Participants using the audio configuration had an average of 9.4, participants in visual had 230.4 and audio-visual participants had 175.44 glances, $F(2,26) = 74.49$, $p < 0.001$. The post-hoc test showed that the visual configuration – which accounted for 58% of all the glances – has a significantly higher number of glances compared to the audio configuration, which only accounted for 2.3% of the glances ($p < 0.01$). When comparing the audio and audio-visual configurations, we found that the audio configuration has a significantly lower number of glances than audio-visual ($p < 0.01$). Further, the difference between the visual and audio-visual configurations showed significance at the 5% level. Audio-visual participants accounted for 39.7% of all glances – nonetheless, participants in the visual configurations accounted for significantly more.

GPS Versus Non-GPS Users

We divided the participants between the configurations according to whether they had previous experiences with GPS navigation guides. Each configuration had five GPS users and five non-GPS users. We identified no significant differences between GPS and non-GPS users. However, we did find that GPS participants on average had fewer eye glances than non-GPS participants on all three categories 140.33 against 124.80, but this difference was not significant. However, GPS users had a higher number of speeding violations and driving errors.

DISCUSSION

Previous research indicates that the use of GPS systems can result in decreased driving performance. We set out to evaluate three configurations of output modalities in order to shed light on how drivers are affected by such a highly output oriented device. Through our evaluation we also sought to identify potential design implications. We will initiate the discussion by focusing on eye glance behaviour, since eyes-off-the-road time is known to affect primary driving task performance [3, 15].

Eye Glance Frequency

Not surprisingly, our results showed a correlation between GPS visual output and eye glance frequency. Although we expected visual output to stimulate eye glances, we were surprised to see how often participants diverted their visual attention from the driving scene to the GPS system. Green found in a comparable field study a glance frequency of one glance every 8.5 seconds for a satellite navigation system [13]. This corresponds well with our results as our audio-visual configuration participants had an average glance frequency of one glance every 7.8 seconds while visual participants glanced at the GPS system every 5.9 seconds on average. Thus, adding audio output seems to decrease eye glance frequency. However, if we consider that all participants in the experiment were equally successful in completing the navigational tasks, a frequency of one eye glance every 7.8 seconds is still extremely high. Having

more than one output source available did not seem to be an advantage. Visual output only seemed to make participants take their eyes off the road perhaps due to the fact that found driving tedious and felt entertained by the visual output.

Since driving performance usually decreases when visual demand increases [3], such high eye glance frequencies are problematic for off-the-shelves GPS navigation systems. But the high glance frequency in our experiment could be explained by the fact that several of our participants belong to a relatively young age group. Similar studies indicate that younger and older drivers differ in driving behaviour [17] and Green et al. found that younger drivers on average have a higher glance frequency compared to older drivers [16]. Finally, the eye glance frequency did not appear to be influenced by the context. Participants repeatedly looked at the system in both densely populated areas as well as on longer rural segments.

Driving Performance While Using GPS Systems

We found no significant differences on driving performance for the visual and audio-visual configurations even though eye glance frequency differed significantly between the two configurations. In spite of these consequences and previous research findings [18], our study did not seem to confirm these. This is quite surprisingly as lack of visual attention usually decreases primary driving task performance [1].

While glance frequency did not seem to affect participants driving performance – the presence of visual output did. We recorded several incidents (e.g. running red lights, missing turns or speeding violations) while participants glanced at the GPS system. Our results showed that audio only participants performed better in relation to primary driving task performance than the other participants. Other research studies confirm that audio configurations would be most ideal in terms of road safety [16, 23]. Nonetheless, it is also important to consider that driving performance differences between the configurations could be explained by the difference in how visual and audio information is acquired. Cautious driving could be a result of participants relying on the GPS system to guide them and not being in control of when and where they receive instructions. This level of uncertainty and alertness could cause an increased cognitive workload of drivers. This matter is further emphasized by a behavioural pattern seen in all audio participants, where they drastically decrease speed when presented with auditory instructions. Studies on voice instructions reveal that a decrease in speed is one of the most significant indicators of increased cognitive workload [28].

The lack of attention on the driving activity resulted in participants missing turns when attempting to relate the map provided by the GPS to their surroundings. Two of the participants became so engulfed by the system that the vehicle nearly came to a standstill while traversing a roundabout. We further found that all participants using the

visual configuration repeatedly looked at the system while performing driving manoeuvres, e.g. while making a turn or driving through a roundabout. This behavioural pattern only occurred in two field trials, which involved the audio-visual configuration. This could be explained by the difference in how visual and audio information is provided. Visual information is readily available allowing users to retrieve information whenever they deem it necessary while the audio configuration only provides information in selected situations. Since GPS navigation guides serve as navigators within unfamiliar environments, drivers using only visual output need to confirm their manoeuvres and driving while the addition of audio output seems to alleviate this need.

Using GPS Instructions and other IVS

One of the known challenges or limitations of GPS guides is precision and accuracy of the guidance, e.g. maps have to be updated regularly. One visual configuration participant intentionally diverted from the route assigned by the GPS system. This was motivated by the fact that the participant knew a better route. The participant defined a better route as involving less turns and traffic lights, while permitting a higher speed level. The tendency to ignore GPS instructions while driving in familiar areas is also identified in a study by Leshed et al. [22]. This study describes how users would still utilize the system in order to feel in control by locating and orienting themselves on the map. We also observed this pattern as the participant repeatedly looked at the system in order to see if his chosen route was shorter than the one recommended by the GPS system.

GPS navigation guides comprise just one class of in-vehicle systems that are going to compete for the driver attention in the future. Other research studies have investigated how we can minimize driver distraction through different interaction techniques [2, 3]. Possibly, one should understand the use of GPS navigation guides in the context of other systems used in the car, e.g. climate control, audio players.

When considering the measurement variables, our results showed an inclination towards the audio configuration. To shed light on this matter we assessed the data collected through the post-task questionnaire in order to evaluate the participants' opinions of the audio configuration. When asked to assess the system instructions and output modality, the audio configuration was rated highest (rated from 'neutral' to 'very satisfied') compared to visual and audio-visual (several of the ratings were 'dissatisfied' and 'very dissatisfied'). The outcome of the questionnaire indicates an overall satisfaction amongst the audio participants. The responses given in the interview contradict these findings. Over half the participants expressed that they would prefer the presence of visual output (where two only preferred visual output). Interestingly, half of the visual participants expressed the desire to have both visual and audio output (three preferred audio only), while half of the audio-visual participants would have preferred visual output only. The

expressed opinions contradict each other, but a third of the participants stated that they would prefer to enable and disable the audio output in accordance to their own preferences. This seems to indicate that the audio-visual configuration would result in the highest user satisfaction, even though results indicated that it is less safe.

Design Implications

With the improvement of technologies and prize reduction, GPS navigation systems have become popular consumer technologies over the past years [7]. Most current GPS systems utilize touch-screen interfaces and rely extensively on visual interaction. Our findings suggest that producers should consider output modalities for future GPS systems.

Previous research studies have found that car drivers prefer navigation instructions in terms of e.g. landmarks, road numbers, or street names, whereas distance instructions are less desired [6, 22]. Several of our participants with audio output experienced problems relating such information to their surroundings and requested use of more landmarks and descriptive details in the instructions. We believe that this additional descriptive information would not only support drivers, but also alleviate confusion caused by technical limitations in GPS systems. More than half of our participants experienced problems with inconsistency in the GPS system's distance estimates – for example, delayed updates of visual maps or instructions due to unstable satellite signals or loss hereof. Areas, which had several navigational options, caused confusion amongst the drivers, which consequently led to navigational errors. Utilizing landmarks and additional descriptive information would further enable drivers to relate visual maps and instructions to their surroundings.

The audio participants also expected more guidance before they were to traverse a complex intersection or roundabout. Allowing the user to retrieve information whenever they deem it necessary could be a potential solution as it also maintains an auditory interface. On the other hand, designers should also consider how to restrict or omit visual output when the car is in motion.

During the audio configuration field trials, we observed that participants looked at the system during travel, even though no visual output was provided. Since most GPS systems are sold as independent consumer technologies, they have the disadvantage of giving the user a visual focal point that may attract their attention when presented with instructions. We believe that by integrating GPS systems into vehicles and utilizing the car stereo for audio output could eliminate the visual focal point. This concept is similar to many hands-free phone systems.

CONCLUSION

The market for in-vehicle systems has grown significantly over the last years and especially GPS navigation guides

have become very popular. We conducted an experiment with the aim to compare three output configurations of a GPS guide on drivers and driving performance. Our results indicated that visual output not only causes a substantial amount of eye glances, but also leads to a considerable decrease in driving performance. While the introduction of audio output in combination with visual output reduced the frequency of glances, the effects on driving performance were minimal. This could indicate that the presence of audio output may induce additional cognitive workload, nonetheless audio output is beneficial when considering eye glance behaviour and glance tendencies. A direction to be pursued is to design an audio output based navigation system, which accedes to user preferences, as our results already indicate that audio output is an adequate output modality in terms of road safety. Moreover, further studies are needed to fully understand the behavioural patterns emerging when using GPS systems.

Although we strived to approximate a natural setting, we cannot eliminate the possibility that participant behaviour was affected by the fact that they were being observed. During the interview, two participants expressed that they chose to follow the provided route despite disagreements with the given instructions, as they believed that a linear approach was necessary in order to complete the tasks. We also acknowledge that there is an imbalance between the visual and audio configurations in relation to the way the navigational information was provided. Participants in the visual output configurations had additional information e.g. street names, estimated arrival time or distance, which was not available through audio output. Finally, the setup of our experimental study could be problematic as drivers often fit the involvement of in-vehicle systems with the driving activity as pointed out by Esbjörnsson et al. [9]. This could also be the case for GPS navigation guides.

ACKNOWLEDGMENTS

We would like to thank our participants for the involvement in the experiment. Also, we thank anonymous reviewers for valuable comments on previous versions of this paper.

REFERENCES

1. Bach, K. M., Jæger, M., Skov, M. B., & Thomassen, N. G. (2009). Interacting with In-Vehicle Information Systems: Understanding, Measuring, and Evaluating Attention. In *Proceedings of the HCI 2009*, ACM Press, pp. 453-462
2. Bach, K. M., Jæger, M., Skov, M. B., G., Thomassen, N. G. (2008) You Can Touch, but You Can't Look: Interacting with In-Vehicle Systems. In *Proceedings of the 26th annual SIGCHI Conference on Human Factors in Computing Systems (CHI'08)*, ACM Press, pp. 1139-1148
3. Brooks, C. & Rakotonirainy, A. (2007) In-vehicle technologies, *Advanced Driver Assistance Systems and*

- driver distraction: Research challenges. Distracted driving. Australasian College of Road Safety.
4. Brown, I. (1994) Driver Fatigue in Human Factors. Human Factors and Ergonomics Society, Vol. 36, Issue 2, pp. 298 - 314
 5. Broy, M. (2006) Challenges in Automotive Software Engineering in ICSE, ACM Press.
 6. Burnett, G. (2000) Turn Right at the Traffic Lights - The Requirement for Landmarks in Vehicle Navigation Systems. The Journal of Navigation, Vol. 53, No. 3, pp. 499 - 510
 7. Cassavoy, L. (2008) TOMTOM GO 930 Review, PC World, November 2008.
 8. Devonshire, J., Flannagan, C., and Sayer, J. (2005) The Effects of Secondary Tasks on Naturalistic Driving Performance, The University of Michigan Transportation Research Institute, UMTRI-2005-29.
 9. Esbjörnsson, M., Juhlin, O., and Weilenmann, A. (2007) Drivers Using Mobile Phones in Traffic: An Ethnographic Study of Interactional Adaptation. International Journal of Human-Computer Interaction, Vol. 22, No. 1&2, pp. 37 - 58
 10. Fleiss, J. L., Levin, B. and Paik, M. Cho (2003). Statistical Methods for Rates and Proportions, 3rd ed., John Wiley, New York
 11. Geiger, M., Zobl, M., Bengler, K., Lang, M. (2001) Intermodal Differences in Distraction Effects while Controlling Automotive User Interfaces. In Proceedings Usability Evaluation and Interface Design, HCI 2001.
 12. Gellatly, A. W. (1997). The Use of Speech Recognition Technology in Automotive Applications. Faculty of the Virginia Polytechnic Institute and State
 13. Green, P. (1996) Customer Needs, New Technology, Human Factors, and Driver Science Research for Future Automobiles in Journal of the Society of Mechanical Engineers, University of Michigan Transportation Research Institute.
 14. Green, P. (1999) Visual and Task Demands of Driver Information Systems, The University of Michigan Transportation Research Institute.
 15. Green, P. (2001). Variations in Task Performance Between Younger and Older Drivers: UMTRI Research on Telematics. Association for the Advancement of Automotive Medicine Conference on Aging and Driving, Southfield, Michigan
 16. Green, P., Hoekstra, E., Williams, M., Wen, C., and George, K. (1993) Examination of a Videotape-based Method to Evaluate the Usability of Route Guidance and Traffic Information Systems, UMTRI-93-31.
 17. Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., and Brown, J. (2006) Driver Distraction: The Effects of Concurrent In-Vehicle Tasks, Road Environment Complexity and Age on Driving Performance. Accident Analysis and Prevention, Vol. 38, No. 1, pp. 185 - 191
 18. Jovanis, P. and Srinivasan, R. (1997) Effect of In-Vehicle Route Guidance Systems on Driver Workload and Choice of Vehicle Speed: Findings From a Driving Simulator Experiment. Ergonomics and safety of intelligent driver interfaces, pp. 97 -114
 19. Keith, O., and Burnett, G. (2008) Learning-oriented Vehicle Navigation Systems: A Preliminary Investigation in a Driving Simulator. Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services, pp. 119 - 126
 20. Lansdown, T. C., Brook-Carter, N. and Kersloot, T. (2004). Distraction from Multiple In-Vehicle Secondary Tasks: Vehicle Performance and Mental Workload Implications in Ergonomics, Vol. 47, No. 1., Taylor & Francis Group, pp. 91–104
 21. Lee, D., Caven, B., Haake, S., Brown, T. (2001) Speech-based Interaction with In-vehicle Computers: The Effect of Speech-based E-mail on Drivers' Attention to the Roadway, University of Iowa.
 22. Leshed, G., Velden, T., Rieger, O., Kot, B., Sengers, P. (2008) In-Car GPS Navigation: Engagement with and Disengagement from the Environment. Proceedings of ACM Conference on Human Factors in Computing Systems (CHI'08), ACM Press pp. 1675-1684
 23. Moldenhauer, D. and McCrickard, S. (2003) Effect of Information Modality on Geographic Cognition in Car Navigation Systems. Center for HCI and Department of Computer Science, Virginia Tech.
 24. Nixon, C., Anderson, T., Morris, L., McCavitt, A., McKinley, R., Yeager, D., and McDaniel, M. (1998) Female Voice Communications in High Level Aircraft Cockpit Noises part II: Vocoder and Automatic Speech Recognition Systems. Aviation, Space and Environmental Medicine, Vol. 69, No. 11, pp. 1087 - 1094
 25. Rockwell, T. H. (1988). Spare Visual Capacity in Driving – Revisited: New Empirical Results for an Old Idea in Gale, A.G., et al. Ed.: Vision in Vehicles II, Elsevier Science, North Holland, pp. 317 - 324
 26. Tsimhoni, O. and Green, P. (2001). Visual Demand of Driving and the Execution of Display-Intensive, In-Vehicle Tasks in Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting, pp. 1586 - 1590
 27. Tsimhoni, O., Smith, D. and Green, P. (2004). Address Entry While Driving: Speech Recognition Versus a Touch-Screen Keyboard. Human Factors and Ergonomics Society, Vol. 46, No. 4, pp. 600 - 610
 28. Uchiyama, Y., Kojima, S., Hongo, T., Terashima, R. and Wakita, T. (2002) Voice Information System Adapted to Driver's Mental Workload. Proceedings of the Human Factors and Ergonomics Society, pp. 1871 - 1875
 29. Wickens, C. D. and Hollands, J. G. (2000). Engineering Psychology and Human (3. ed.). Prentice Hall