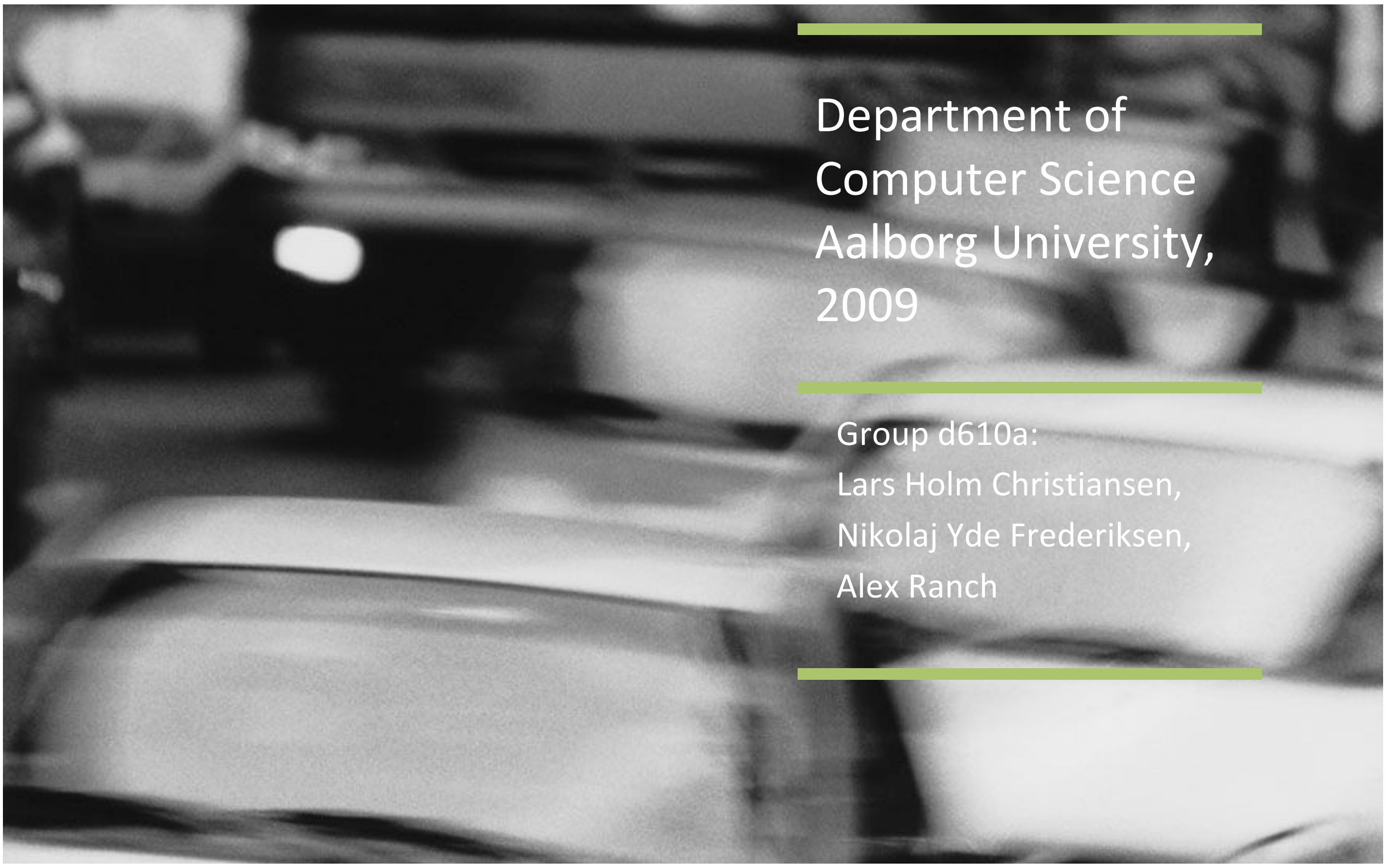
Investigating the effects of in-vehicle systems on driving behaviour and attention



Department of Computer Science
Aalborg University, 2009

Group d610a:
Lars Holm Christiansen,
Nikolaj Yde Frederiksen,
Alex Ranch

Title: Investigating the effects of in-vehicle systems on driving behavior and attention

Project period: Spring 2009

Project group: d610a

Group members:

____

Alex Ranch

____

Nikolaj Yde Frederiksen

____

Lars Holm Christiansen

Supervisor: Mikael B. Skov

Editions: 5

Number of pages: 34

Finished: 11-06-2009

As more and more technologies find their way into the car, it is becoming an increasingly important area to focus attention on. In light of this, it is important to determine how in-vehicle systems affect the driver.

In our master's thesis we seek to investigate what effects two different in-vehicle systems have on driving performance. The first system is an advance warning system intended to aid the drivers in the primary driving task; the second system is a music player and thus focuses on the secondary driving task.

Through two experiments, one conducted in a controlled environment, the other in a simulated driving setting, we investigate the effects of the aforementioned systems on driving behavior.

Preface

This report was written during the 10th semester of the Informatics study at the Department of Computer Science, Aalborg University and constitutes our master's thesis. It represents work from both the 9th and 10th semester. Our second research paper was initially written during our 9th semester, in collaboration with Brit Susan Jensen and Nissanthen Thiruravichandran, but has been revised into the current version in this report (changes noted in table below). The first research paper was written during 10th semester.

Section:	Change:
Introduction	Modified
Related Work	Rewritten
In-vehicle System	Modified
Experiment	Rewritten
Data Analysis	Rewritten
Results	Modified
Discussion	Modified
Conclusion	Modified

Acknowledgements

We would like express our thanks to a couple of individuals who have helped us throughout this project.

First of all we would like to thank our supervisor Mikael B. Skov for competent guidance throughout our master's thesis.

Furthermore we would like to thank our participants in our experiments, along with the staff at the driving facility in "Nørresundby", for their help and cooperation.

Table of contents

Introduction

When exploring the interaction between humans and computers (HCI), it is both interesting and relevant to study how this interaction fares in contexts that impose certain restrictions, where the information system being used is not necessarily the primary focus of the user. This is interesting because it presents unique challenges to designers and evaluators of such systems and forces them to rethink existing solutions or develop entirely new ones. It is also relevant because these interaction situations often occur in safety-critical contexts, where usability is not just important, but can make a vital difference.

In-vehicle systems

An example of a widespread safety-critical use context is the car, which millions of people find themselves in every day. This is an interaction environment that is becoming increasingly prevalent within the field of HCI. And with good reason, as the general technological progress and ubiquity of information systems means that cars today are becoming increasingly equipped with information systems of all kinds. In fact, the in-vehicle systems market has grown ten-fold in the last generation and is expected to increase exponentially in the next [9]. These systems, collectively known as in-vehicle information systems or simply in-vehicle systems have many different forms and objectives, but their common denominator is that they are secondary to driving the vehicle.

Car accidents are responsible for the deaths of thousands of people each year, reaching an estimated 37,313 deaths in 2008 in the US alone [27]. While the contributing factors are many, recent research on traffic safety indicates, that the use of in-vehicle systems is a contributor to the cause of traffic accidents [8, 15, 29]. One of the safety issues with in-vehicle systems relate to the fact that they are often used while driving, but not necessarily designed to be. Many of these systems have their origins outside the vehicle setting, where their design may be perfectly suitable and where the interaction with the system is the user’s primary concern. This is equally true for other dual-task conditions, as supported by an empirical study by Chewar et al. (2002) concerning dual-task performance on desktop computers. They found that established design approaches for primary task systems could not successfully be applied to systems used as a secondary task [11]. When applying the traditional guidelines for dual-task systems they found that users were unable to effectively perform the secondary task without a reduction in the performance of the primary task. And since the primary focus of the driver should always be the safe operation of the vehicle, this is something designers of in-vehicle systems must take into account.

In similar safety-critical contexts, such as aviation, where information systems also exist that are not the user’s primary focus, these systems are often subjected to more rigorous design processes and are required to adhere to strict rules and regulations. Another key difference is that the users are much better trained in the use of the systems. Drivers are only required to receive training in the safe operation of the car, not any in-vehicle systems that may be present in the car. And because the types and designs of in-vehicle systems are so varied and drivers equally so, drivers cannot be expected to be trained as well. The implications of this are that the design of the in-vehicle system must compensate for the lack of training.

Many in-vehicle systems are designed solely for entertainment or convenience purposes that are not directly related to driving the car. They include CD players, DVD players, game consoles and full-blown in-car entertainment systems and "carputers" that combine TV, radio, Internet access and office functionality in a single system. The trend is to add more and more features and complexity into these systems as the technology becomes available. As an example, take the car stereo, probably the oldest in-vehicle system. From being a simple radio receiver with a few dials, it has evolved and grown in complexity to a point where the original functionality is now only a small part of the full feature-set, and the dials have been replaced by numerous buttons or touch screens. Recent studies indicate that this increase in complexity of in-vehicle systems has a negative effect on driving safety. For instance, Chisholm et al. (2007) examined the effects of interacting with an iPod during driving. Their results show that drivers had more frequent collisions, slower response time as well as more and longer eye glances off the road when they interacted with the iPod [12].

Other systems, such as mobile phones, are not related to entertainment and are only present in the car due to their inherent mobility. Mobile phones use has often been accused of contributing to car accidents [1, 7, 26]. A study by Strayer and Drews (2007) found that drivers who were talking on a mobile phone, were less likely to memorize objects in the driving environment, even though they were staring directly at them [30]. However, Shinar and Tractinsky (2004) showed that there is a learning effect associated with in-vehicle mobile phone use, and that the negative impact of mobile conversations on driving performance may be lessened with continued practice [28].

Another increasingly popular in-vehicle system is the automotive navigation system, typically in the form of a GPS device, which assists the driver in navigating the vehicle to the desired location. Many GPS devices also have added functionality that utilizes the location information to provide the driver with information about Points of Interest (POIs) in the surrounding environment. Leshed et al. (2008) conclude that in-vehicle GPS systems have the potential of affecting the way drivers interact with their environment in complex ways. On the one hand, drivers become immersed more in the virtual world of the GPS, which affects their interaction with the physical world. On the other hand, they found that the added context information from the GPS, such as POIs, enriched the travel experience of the drivers [23].

The above-mentioned examples are in-vehicle systems that most drivers have some level of familiarity with, but there is a different and emerging breed of in-vehicle system that is not yet as familiar to the regular driver. These systems are not designed for entertainment or convenience, but for safety. They seek to support the driving task and actively improve safety either by adding previously unavailable information or presenting the existing information in a safer way. For example, head-up displays (HUDs) are being used to present information from the dashboard in the driver's main field of view, while parking sensors provide the driver with supplementary information to support maneuvering the vehicle. As an indicator of the emergence of these in-vehicle systems, car manufacturer Opel has included a new feature on the 2009 model Opel Insignia called Opel Eye that searches for and recognizes traffic signs along the road. These signs, for instance speed limit postings, are then presented to the driver as a warning symbol in the dashboard to increase the likelihood of the driver noticing the sign and adhering to it [31].

Driving behavior and attention

With the safety considerations in mind, we are interested in investigating exactly how interaction with in-vehicle systems affects the driving task and to what extent. Specifically, we want to study the effects on the behavior of the driver in terms of operating the vehicle and the impact on the driver's attention.

The main safety risk of in-vehicle systems is often associated with attention, and it is a concept that cannot be ignored when designing or evaluating in-vehicle systems. Attention can be defined as the ability to concentrate and selectively focus or shift focus between selected stimuli [13]. In relation to driving, these stimuli predominantly come from visual perception, although auditory and tactile perception also play their part. As a result, measures of driver attention are often associated with eye glance behavior [2]. Any disruption in the concentration or focus of attention is defined as distraction, which in the driving context would mean anything that takes attention away from the driving task.

In terms of in-vehicle systems, the problems regarding attention tend to lie in increasing the load on the driver's visual and cognitive attention as well as requiring the use of his/her hands. If the driver is distracted with operating the in-vehicle system, less attention is given to the driving the car safely in relation to the surroundings. It is widely accepted that drivers who are visually occupied with tasks other than driving are more likely be implicated in traffic accidents [24]. This is also supported in studies of in-vehicle systems, such as Lansdown, Brook-Carter and Kersloot (2004), who studied the effects of multiple in-vehicle information systems on driver distraction. Their results show that the participants' interaction with the in-vehicle system had an unfavorable impact on vehicle performance [21]. An important aspect of designing in-vehicle interaction must therefore be how to minimize the amount of attention it removes from the driving task.

On the other hand, there are systems that are designed to use the driver's attention as an advantage with the aim of improving safety. By attracting the attention, and in essence distracting the driver from the driving task, they aim to convey important or even vital information and hereby enable the driver to make more informed decisions. For instance, an in-vehicle system could warn the driver when too close to the vehicle in front or that a collision is imminent or draw attention to traffic signs like the aforementioned Opel Eye system.

Another measure of the effects of in-vehicle systems interaction is the impact on the driving behavior or driving performance. As driving must be considered the main job of the driver, it is often referred to as the primary driving task as opposed to the secondary driving task, which involves anything from talking to passengers to interacting with in-vehicle systems [14]. Primary driving task performance, therefore, is a measure of how the vehicle is controlled; the speed, acceleration, lane variability, steering wheel input, etc. This is closely related to attention, as distraction from vehicle control can cause degradation in primary driving task performance. However, attention and distraction can also be used to improve driving performance. By carefully controlling the attention of the driver, it is possible to enable better-informed decisions regarding how the vehicle is controlled. Hence, the relationship between attention/distraction and driving performance is two-sided, as attention has the potential to affect driving behavior in both a negative and positive direction.

An understanding of attention and driving behavior can therefore be exploited to both design safer interaction with in-vehicle systems as well as systems that actively encourage safer driving.

Research focus

The objective of this master's thesis is to explore what effects in-vehicle systems have on the driving task, in regards to driving behavior and attention, both for systems where a minimum impact is desirable and systems that aim to maximize their impact in a positive way. This is done in order to gain a better understanding of how to design safer interaction for the countless of systems already in the car and to uncover the potential benefits of emerging systems. Our research focus can be summarized into the following question:

- How do in-vehicle systems affect driving behavior and attention?

In order to answer this question, we study two types of in-vehicle systems that have two different purposes; one seeks to support the primary driving task and one is only concerned with the secondary driving task. We treat these separately by defining two sub-questions corresponding to the two research papers that are the main contribution of this master's thesis:

1. How does an in-vehicle system intended to support the primary driving task affect driving behavior and attention?
2. How do different combinations of input and output methods for in-vehicle systems used as a secondary task affect driving performance and attention?

The first question is answered in the first research paper and concerns the effects of an in-vehicle system that is designed to encourage safer driving by supporting the primary driving task. A dashboard-based in-vehicle advance warning sign system designed to warn the driver of upcoming traffic and road conditions is evaluated and compared to a no-warning baseline condition, using an instrumented vehicle on a closed circuit. The effects of the system are then analyzed with regards to impact on driving behavior and eye glance behavior.

The second question concerns the effects of interaction with in-vehicle systems that are used as a secondary driving task. Two input and two output methods for an in-vehicle music player system are evaluated in a driving simulator and their effects on primary driving task performance, secondary driving task performance and the eye glance behavior of the driver are compared. This gives us the opportunity to study the separate effects of input and output on driving behavior and attention. This experiment therefore contributes with knowledge on how to design safer interaction with systems used as a secondary driving task.

In the following we will first present our research contributions by describing how we have answered the two research questions. This is only a short description of our two research papers, and reading them in their entirety is recommended (they can be found in the appendices). Subsequently we will discuss and reflect upon the work described in the papers, before presenting our conclusions.

Research contribution

In the following we describe our research contributions in the form of our two research papers. The first paper represents our work studying the effects of an advance warning in-vehicle system indented to aid drivers in the primary driving task. The second paper concerns our study of the effects of different interaction techniques with an in-vehicle music player used as a secondary driving task on driving behavior.

Research contribution # 1

The tendency is that cars are filled with more and more in-vehicle systems, moving the attention of the driver away from the primary task of driving. But a new kind of in-vehicle systems is intended to utilize the attention of the driver to actively support the driving task. They seek to do this by adding previously unavailable information or presenting existing information in a safer way. However, it is unclear exactly what the effects of these systems are on driving behavior and attention, and what type of information is beneficial. Our aim is therefore to evaluate the effects of an in-vehicle advance warning system on driver behavior and attention, for a different road and traffic conditions.

We constructed an advance warning system designed to issue warnings to the driver 75m in advance of five different road and traffic conditions or incidents; left curve, right curve, slippery road, speed limit and traffic jam. When the incident is reached, a new warning is emitted indicating this. The warnings appeared as symbol and were accompanied by earcons. In order to evaluate the advance warning system, we conducted an experiment with an instrumented car on a closed circuit. Using a within-subject design, twelve participants (seven male and five female), drove eight laps in two different conditions, one with advance warnings, one without. The two conditions were counterbalanced to limit learning effects. The participants were instructed to drive the car as they would in real life traffic and obey normal rules and regulations. We directed the participants around the training facility, so that each of the participants was exposed to the same conditions and route. In the advance warning condition, the participants were given advance warnings regarding curves, slippery road, speed limit and traffic jam incidents. The warnings were triggered using the Wizard of Oz approach. During the experiment, the speed of the participants was logged automatically each second, and a video camera was used to record eye glances.

After our experiment we analyzed our test data by looking at the logged speed and analyzing the videos to identify eye glances. Our primary focus was to calculate how much the participants reduced their speed between the advance warnings was given until they reached the incident. The results of our experiment show that the speeds were remarkably similar across the two conditions, and in general the advance warnings did not cause any significantly larger reduction in speed. However, there was a more pronounced effect for some incidents, most notably for speed limits and slippery roads, where the reduction in speed was significantly larger with warnings. The results of our eye glance analysis showed that the participants had significantly more glances at the system with warnings than without, and that the participants glanced down for a vast majority of the glances. Based on the results from our evaluation, we conclude that the overall effects of in-vehicle advance warning systems are limited.

Research contribution # 2

With the increasing number of systems present in cars today, drivers have to divert more of their attention from the primary task of driving unto these secondary in-vehicle systems. This undoubtedly has considerable consequences on the driving performance, which is an issue developers of in-vehicle systems need to address. A lot of research exists that seeks to introduce some novel method of interaction that tend to these characteristics, but there is a tendency in the existing research to mainly focus on input techniques, and not to the same extent, on output techniques. We wish to address this by comparing different configurations of in-vehicle systems with an equal emphasis on both input and output methods. Our aim is to restrict system variables with regards to input and output to be able to perform a strict comparison.

In order to investigate the effects of different input and output methods while driving, we developed an in-vehicle music player that allowed us to select the desired combinations of input and output modes. Based on Bach et al. [4] we chose traditional touch screen and touch screen-based gesture interaction as input forms and visual and audio-based output forms, which gave us the following system configurations; touch/visual, touch/audio, gesture/visual and gesture/audio. In order to evaluate these configurations, we conducted an experiment in a medium-fidelity driving simulator with 32 participants (16 male, 16 female). Using a between-subjects design, each of the four configurations were evaluated using eight participants. The participants were instructed to drive as they would in real-life traffic and obey the rules of the road. While they were driving, they were given 32 short tasks to be solved using the music player (e.g. “Play song no. 7”, “Turn up the volume”).

The sessions were all recorded on video from different angles, allowing us to analyze the data according to the following metrics; primary driving task performance (lateral and longitudinal control errors), secondary driving task performance (interaction errors and total task completion time) and eye glance behavior (number of glances below 0.5 seconds, between 0.5 and 2.0 seconds, above 2.0 seconds, as well as the total number of glances). The results of the experiment revealed that gesture input resulted in significantly fewer eye glances, but also worse primary driving task performance as well as longer task completion times compared to touch input. Concerning output, audio caused a dramatically lower number of glances and fewer interaction errors than visual output, while visual output had the most interaction errors but shorter task completions times. For the specific configurations of input and output, the results revealed that the gesture/audio configuration had significantly fewer total glances than any of the others, while simultaneously having more longitudinal control errors and longer total task completion times. Based on the results of the experiment, we argue that studying input and output on equal terms is important in any effort to understand in-vehicle interaction and its influence on driving performance and attention.

Discussion

During our work on our master's thesis we have identified different topics that are especially relevant to the area of research we are engaged in, which we wish to elaborate on. The following chapter consists of several different topics that we have found to be of special interest in relation to our work within the domain of in-vehicle systems, and their effects on the primary and secondary task.

Use context

When trying to determine the effects of in-vehicle systems on driving behavior, it is useful to look at the context in which it is indented to work. This is especially important in a use context where the users have other and more important tasks to focus on. This is why we in the following wish to elaborate on the importance of understanding the use context. Initially we wish to address what possible effects the driving setting, in which the systems are evaluated, has on driving behavior. Finally, we turn to the subject of reliability. This is relevant since it influences how the driver reacts to the information provided, potentially affecting driving behavior.

Driving setting

Since our aim is to investigate the effects of in-vehicle systems on different aspects of driving, it is relevant to look at the setting in which the driving takes place since this has the potential of affecting the outcome. In our work we have used two different kinds of driving settings, the potential effects of which we would like to elaborate on in the following. When conducting experiments with in-vehicle systems and depending on the aim of the research, there are four different settings to chose from; No driving, simulated driving, controlled driving and real traffic driving [2], with several different aspects influencing this choice. During our research we have conducted both simulator and controlled driving based experiments, which we would like to discuss in the following.

In our experiment regarding interaction with an in-vehicle system as a secondary driving task, we chose to use simulated driving since the objective of the research was to investigate factors relating to the secondary driving task. The experiment was aimed at testing different kinds of interface input and output techniques, and as such the task of driving was not the main focus of the experiment. According to Kemeny and Panerai (2003), when researching general dashboard ergonomics, a category in which input and output techniques belong, is one of the cases where simulated driving can be used efficiently [20].

When discussing simulated driving experiments, the subject of realism is unavoidable. There are obvious differences between real world driving and simulated driving, which have been discussed by numerous authors [4, 10, 19, 20]. Generally, the participants in our simulated driving experiment expressed satisfaction with the level of realism in the simulator, many of them, however, commented on the lack of peripheral information normally obtained through side windows, when for instance negotiating an intersection. Similar to the findings of Bach et al. (2007), where participants also seemed to lack sensory information leading to difficulties in controlling the car while interacting with a in-vehicle system [3]. In line with Bach et al., we tried to engage the participants as much as possible in the simulation by inducing a series of requirements such as observing traffic laws and driving within a given speed range, in order to make the task of driving as close to real world driving as possible.

In our experiment regarding advance warnings, we conducted our research using controlled driving on a closed circuit. The controlled environment was primarily chosen over real traffic driving due to ethical considerations, and since we did not have access to a simulator that would allow us to change road and traffic conditions. However, the controlled environment gave us the possibility to evaluate in more demanding road conditions, with a likelihood of failure that is unacceptable on public roads, which is also one of the advantages of a controlled environment, as stated by [10]. Conducting experiments in a controlled closed environment also has its disadvantages, however. As noted by Bach et al (2007), controlled driving cannot be considered real traffic driving in the sense that many of the objects and dangers that make real world driving challenging, are missing in the controlled environment [3]. We have tried to counter this by having another car drive around the circuit to induce some unpredictability, which according to [3] is one of the ways in which controlled driving can be made more realistic. The other car acted as an autonomous agent, whom the participants had to observe and react to, as they would in real life traffic. Furthermore we had imposed a set of rules that the participants had to follow. For instance, a general speed limit of 60 km/h was imposed and the participants were instructed to obey the sign postings around the track, as they would in real life. In line with Bach et al., we also experienced that controlled driving was more complicated in terms of data collection, as the quality of the collected data was inferior to the data collected in our simulated driving experiment, in terms of video quality, viewing angles etc.

Reliability of in-vehicle systems

When working with in-vehicle systems designed to assist drivers with their primary task, it is relevant to explore the basic prerequisites of these kinds of systems. Therefore, when trying to aid drivers in their primary task, it makes sense to take into account the reliability of the information these provide, since the information in turn influences the driver's decision-making, making it an important aspect to consider. This is no less true when the information is related to the primary task of driving, for instance information about traffic or road conditions regarding risks to the driver. There are several different ways of warning drivers, for instance using audio/visual information, either alone or in combination, or even haptic warnings. Which of these methods is the most effective way of conveying information to the driver is the subject of ongoing research. One of the concerns of this research is the reliability of in-vehicle warning systems, since the reliability can have a large effect on driving performance, safety and to no small extent, drivers' willingness to accept advance warning system, and the information they provide [16, 18].

In our advance warning experiment we investigated the effects of advance warnings on driving speed. The advance warnings we used were a combination of auditory and visual signals, where the auditory warning sounded the same for all types of advance warnings. The focus of the experiment was not to evaluate the influence of reliability, and therefore the person controlling our in-vehicle system tried to do so as accurately and reliably as possible. Hence we tried to make the warnings 100% reliable in the sense that we never intentionally produced false, missing or incorrect warnings. Other researchers have investigated how imperfect warning systems influence drivers. In an experiment investigating the effects of multiple alarms (different sound for every type of warning) vs. single alarms (same sound for every warning) on distraction and driver performance, Ho et al. (2006) subjected their participants to a series of warnings during simulated driving in which the reliability of the warnings varied between low reliability, with a ratio between true positive and false positive of 1:3, and high reliability with a ratio between true positive and false positive of 3:1 [17]. Ho et al.'s results show that correct (either ignoring a false positive or reacting accordingly to a true positive alarm) response to warnings was significantly lower in the low reliability condition. As a result, in the low reliability condition the participants reacted correctly in 58% of the cases, compared to 86% correct response rate for high reliability. This is a surprisingly high number, which is even higher when only looking at the participants' ability to react correctly to a false alarm that drops from 98.3% for high reliability to 60.7% for low reliability. On account of these numbers, the question arises whether a warning system with low reliability actually worsens driver performance, compared to no warnings. At the very least, false positives are dangerous in the sense that they have the potential of prompting the driver to react to a nonexistent danger. This could in turn lead to real danger, if for instance the driver swerves to avoid an obstacle, which is not there, only to find him- or herself in the wrong lane facing oncoming traffic. This and other dangers posed by unreliable warnings systems leads Ho et al. to speculate that drivers might be better off having no warnings at all, compared to unreliable ones. However, other researchers [5, 18, 25] have shown that even imperfect in-vehicle collision avoidance warning systems (IVCAWS) can have a beneficial effect on driving behavior, with their results showing participants adopting safer headway distance even under imperfect IVCAWS conditions. Bliss and Acton (2003) even found that drivers receiving warnings from a 50% reliable IVCAWS had fewer collisions compared to 75% and 100% reliable IVCAWS [6]. However, they explain that the higher collision avoidance rate of the 50% reliable IVCAWS is due to participants adopting a strategy where they confirm the existence of imminent threats before acting. This is a somewhat odd result, since it seems very illogical that less reliable warning systems should somehow produce better results – however, according to [6], the urgency of alarms with the 100% reliable IVCAWS seemed to cause drivers to initially react correctly, but then by account of distraction, often to overcompensate or reverse their action.

How the warnings from an IVCAWS affect driving behavior and driver acceptance of the system, is not only a question of whether the warnings are true or false. At least not according to Lees and Lee (2007), who distinguish between false alarms, which occur as a random activation of the IVCAWS when no threat is present, and unnecessary alarms which occur when the IVCAWS assesses a situation as being dangerous, but the driver does not [22]. Lees and Lee's results indicated that designers of IVCAWS should focus more attention on limiting false alarms, than unnecessary alarms, since the unnecessary alarms, in association with the context in which they occur, actually enhances trust in and compliance with the IVCAWS, whereas false alarms diminish trust and compliance. Lees and Lee's argue that this is linked to the drivers' ability to work out how the system works, and therefore trust it. In the case of unnecessary alarms, drivers get a better understanding of how the system works, and it also alerts them to potential dangers since they occur in relation to something in the driving context that might be a threat.

Even though our warning system was designed to be 100% reliable, one of our participants, when asked about her thoughts on the system, expressed concerns about not knowing how accurate and reliable it was. The same participant believed that the advance warnings had a positive effect in terms of reducing speed during curves. Answering the same question, another participant expressed concerns with coming to rely blindly on the warnings and therefore not paying due attention on the road. In fact, she later stated that she did not feel that she was as focused on driving as she usually was. This tells us that driver acceptance of in-vehicle systems depend on several variables. Several participants noted that they found it annoying or distracting when the system warned them about curves that they could clearly see coming, corresponding to an unnecessary alarm in Lee and Lee's research. However, none of the participants expressed any loss of trust in the system as a consequence of the unnecessary warnings.

Effects of advance warnings over time

When studying the effects of in-vehicle systems, it is relevant to determine not only what the net effects are, but also how the effects vary over time, both in order to determine where exactly the effects exist and to investigate what the long-term effects might be. In the following section, we present and discuss how the effects of our advance warning system vary over time, based on the results of our first experiment. We will discuss the speed development second by second as well as how the speed reductions and eye glances vary during the course of the experiment.

Speed development

After our first experiment, we analyzed how the speed of the participants developed over time. Since the speed was automatically logged once a second, we were able to calculate the average speed of the participants second by second for each incident they encountered. However, because the speed of the participants differed, the number of log entries also differed. This entailed that we needed to select the lowest common denominator in terms of the number of log entries, to ensure comparable data. For instance, if one of the participants had six log entries for a specific incident, and the other participants had eight entries, only the first six log entries of all participants' logs were used. The results of these calculations allowed us to compare how the speed developed for each passing second with and without advance warnings.

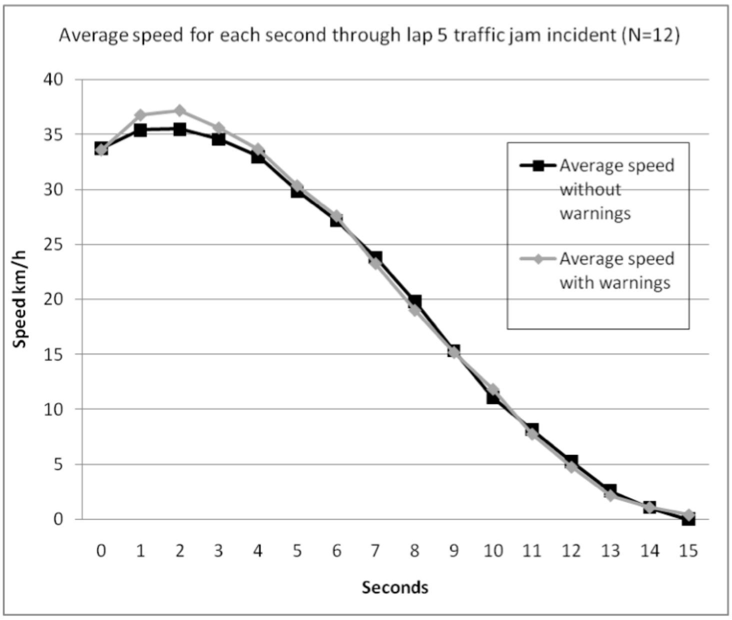


Figure 1. The speed development through the lap 5 traffic jam incident based on all common log entries (N=12)

While we did identify a few significant differences from these results, as described in the first research paper, it is equally interesting to note the remarkable resemblance in the speed between the two conditions under certain circumstances. For instance, Figure 1 illustrates the average speed second by second during the traffic jam incident in lap 5 of the experiment. Here, we see that their initial speed is virtually exactly the same, with a difference of only 0.17 km/h. Throughout the incident there are only slight differences in the speed between the two conditions, with the average variation being 0.23 km/h (SD=0.73). This also seems to be the case for other incidents, for instance the right curve encountered in lap 7 as seen on Figure 2, where we again see a very similar development of speed over time.

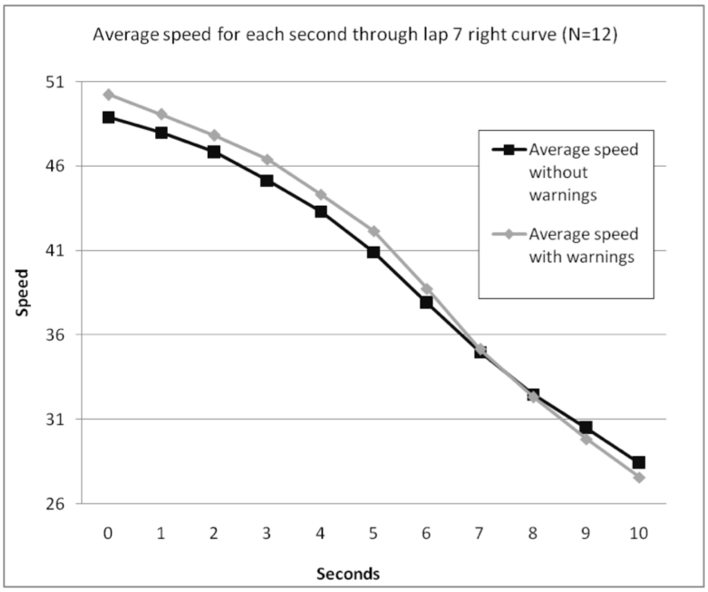


Figure 2. The speed development through the lap 6 right curve incident based on all common log entries (N=12)

These results clearly illustrate that for some of the incidents in our experiment, the advance warnings apparently had no considerable effect on the speed of the participants, at any point during the incidents. These very similar results indicate that the participants have a predetermined notion of an acceptable speed for the section they are travelling on, and an acceptable speed to continue towards the hazard ahead of them. This is also reflected in some of the statements by the participants, who express that they base their speed mainly on their visual assessment of the road and traffic conditions, and rely less on the advance information from the warnings.

Speed reduction and eye glance behavior variation

The results from our first experiment gave us an opportunity to investigate the variation in the effects of the advance warnings on driving behavior and eye glance behavior over the course of the experiment, and analyze the change over time.

Using the speed data collected, we investigated how the advance warnings affected the speed over time, in order to see if the effects increased or diminished. This was done by comparing the data for each occurrence of the individual incidents as they were encountered, e.g. by looking at the first, second and third occurrence of a specific left curve individually, and so on. For each occurrence, we then compared the reduction in speed between when the first warning is given to the incident is reached, to see what effect the warning had. This method was applied to all the incidents that occurred more than once, with and without advance warnings, respectively.

With warnings, looking across all types of incidents, the largest reduction in speed generally occurred at the first occurrence of each incident. At the subsequent occurrences, the reduction in speed diminished. An example of this is the first speed limit, which was encountered six times during each condition. As seen on Figure 4, the first occurrence accounted for the largest reduction in speed (*M*=13.83, *SD*=4.67), with the reduction diminishing for the following occurrences. For all these occurrences, the initial speed when the warning was given remained approximately the same, which entails that it was their speed when the incident was reached that caused the changes in the reductions.

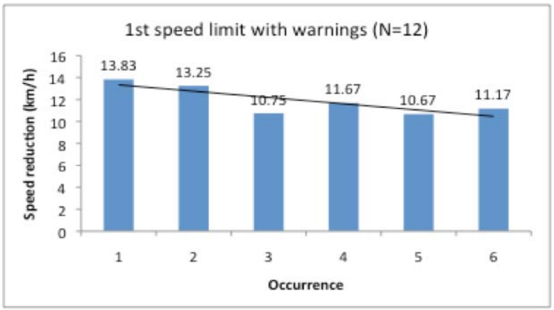


Figure 4. Reduction in speed for all occurrences of the first speed limit, with warnings (N=12).

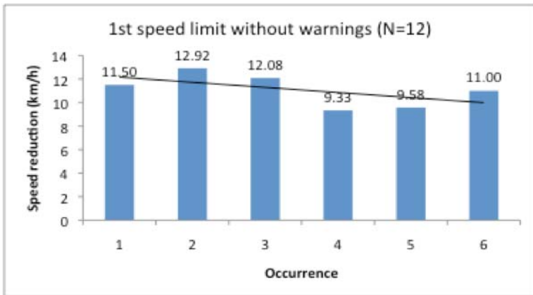


Figure 3. Reduction in speed for all occurrences of the first speed limit, without warnings (N=12).

Without warnings, on the other hand, there is no comparable tendency for any of the incident types. Here, the size of the reductions varies much more, showing no discernible pattern. Figure 3 shows the reductions without warnings for the first speed limit. Compared to Figure 4, the reductions are quite different, with the largest reduction being in the second occurrence, followed by a drop and then a rise again.

In addition to looking at the speed data, we also analyzed the eye glance results from the experiment, which allowed us to study how the eye glance behavior changed during the course of the experiment, to see what the effects the advance warnings had on eye glance behavior over time. This was done for the eye glances that were triggered by the state 1 warnings (issued in advance) and state 2 warnings (issued when the incident is reached), respectively. Figure 5 and Figure 6 show the percentage of state 1 and state 2 warnings in each lap where the participants glance at the system.

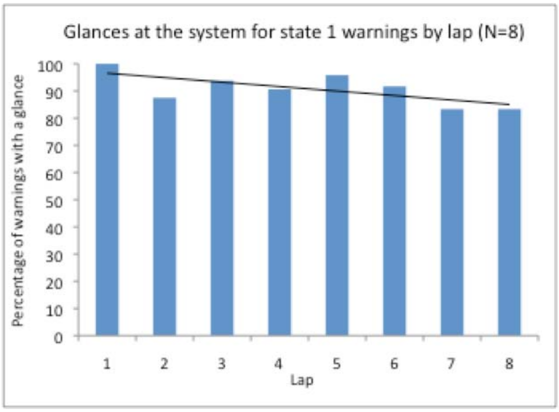


Figure 5. The percentage of state 1 warnings that have an associated glance, distributed over the 8 laps (N=8)

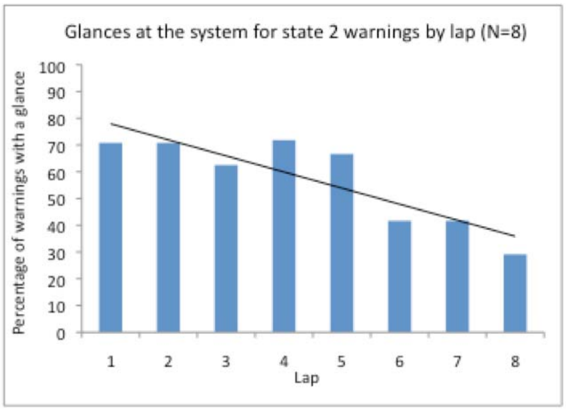


Figure 6. The percentage of state 2 warnings that have an associated glance, distributed over the 8 laps (N=8)

Overall, the percentage of warnings associated with a glance at the system for the state 1 warnings varies very little, as seen on Figure 5. In lap 1, they glance at the system every time there is a warning. In the following laps, the percentage of glances varies marginally, but there are signs of a small drop. By lap 7 and 8, the glances reach a low point at 83%. For the state 2 warnings, there is a much more distinct drop in glances at the system during the course of the experiment, ranging between 62% and 72% in the first five laps, then dropping to 42% in lap 6 and 7, and finally in lap 8 only 29% of the state 2 warnings have an associated glance.

Based on these results, there is a clear indication that the effects of the advance warnings changed over time. In the speed data, we see that the participants reduce their speed less and less each time they encounter the same incident, even though their approach speed stays about the same. This could indicate that the impact of the advance warnings wears off as the participants get more accustomed to them. This is partially confirmed by the results of Ben-Yaacov, Maltz and Shinar (2002), who investigated the effects of an in-vehicle collision avoidance warning system on short- and long-term driving performance [5]. They found that the system continued to have a positive effect on the speed of the participants, even after six months, but that the effect did diminish during that time.

Our findings could also be a result of the design of our experiment, which has the participants drive through the same incidents several times. This could in turn cause the participants to experience less and less uncertainty when they encounter the incidents, because they know from their previous encounters what to expect. If that is the case, then our results indicate that advance warnings are less effective when warning about conditions that are familiar to the drivers, which is also supported by statements made by several of our participants, who feel that some of the warnings would be less relevant in familiar settings.

The glances at the system stay more or less at a constant level for state 1 warnings, which indicates that the participants continue to take notice of the advance warnings throughout the experiment. For state 2 warnings, on the other hand, the results show that the participants pay less and less attention to them. This implies that the state 2 warnings are not as useful to them as the state 1 warnings are, which is also noted by several of the participants in the post-test interview. They stated that they did not notice the state 2 warnings or that they did not find them useful when they did, either because they were preoccupied with driving the car through the incident or because they found the warning obvious, as they were already at the incident.

Subjective measurements

When introducing new technologies into areas where they have not been present earlier, it is important to take into account the opinion of the potential end-users. Thus, focusing attention on this helps in gaining an understanding of how these types of systems affect their users. Therefore, in an effort to do so, we interviewed the participants of our advance warning experiment after each session, in order to gather their opinions and attitudes towards the system. Furthermore, we asked the participants to fill out a questionnaire consisting of a series of statements, regarding the advance warning system, in which they had to state their agreement using a five-point Likert scale, which the following section will present the findings from. All of the participant's answers can be seen in Table 1, along with the most common answer and standard deviation.

	Statement									
Subject	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**
1	4	4	3	4	5	5	4	3	4	4
2	4	3	4	4	4	3	3	2	4	2
3	4	4	4	4	2	1	2	3	4	3
4	3	3	3	5	4	1	1	4	4	4
5	4	4	4	4	5	1	3	2	4	1
6	5	4	4	5	5	3	2	2	4	2
7	4	3	4	4	4	3	2	3	4	2
8	4	3	3	4	4	4	2	3	4	2
9	5	5	5	5	5	5	2	1	5	2
10	4	4	3	2	5	1	4	2	4	2
11	4	3	3	4	3	4	3	3	3	2
12	1	3	4	4	4	1	2	2	2	4
Mode	4	3	4	4	5	1	2	3	4	2
SD	1,03	0,67	0,65	0,79	0,94	1,61	0,90	0,80	0,72	1,00

Table 1. Subjective measurements – results from questionnaire.

The scale consisted of the following five degrees of agreement: (1)"Strongly disagree", (2)"Disagree", (3)"Neither agree nor disagree", (4)"Agree", and (5)"Strongly agree". The results of this questionnaire revealed a generally positive attitude toward the advance warning system, as only one of the ten questions were answered predominantly negatively.

The first statement the participants responded to was, "Understanding the difference between when the warnings flashed and were constantly lit was easy", to which the most common answer (mode) was 4 corresponding to "agree" (*SD*=1.03). Only one of the participants answered (1)"Strongly disagree" to this statement, the rest answered from (3)"Neither agree nor disagree" and up the scale. In our post interview a participant furthermore expressed uncertainty in understanding the difference between when the warning flashed, and when it was constantly lit.

The second statement, “The system aided me in my driving”, was answered with a most common score of 3 “Neither agree nor disagree” (*SD*=0.67) with one of the participants noting that the warnings were a good reminder, another noting that the warnings were nice to have.

The third statement was, “The system would help me feel safer in traffic”, to which the participants answered with a most common answer of 4 “Agree” (*SD*=0.65). One of the participants stated that the warnings would be helpful, but that he would never trust them 100%, since they could be wrong, or not appear at all. Another noted that the warnings would be especially helpful in traffic jam and slippery road conditions.

“The warnings were intuitive and easy to understand” was the fourth statement, which the participants rated with a most common score of 4 (*SD*=0.79). One of the participants noted that different colored warnings could be a possibility, another noted the warning symbols were too similar. During the interview, a third participant stated that he recognized the traffic signs from real life driving.

The fifth statement was, “Being warned by a sound was nice”, to which the participants reported most common answer of 5 corresponding to “Strongly agree” (*SD*=0.94). A quote from our post-experiment interview furthermore supports this: “the auditory warnings are actually good”. A participant noted that only slippery road and traffic jam warnings ought to be accompanied by a sound, the remaining should be without sound. Another participant noted that the sounds were annoying and hard to tell apart, while a third participant noted that the sound warnings of the warnings had to be right, and that the ones we had chosen were satisfactory.

Statement number six was the only statement to which the participants were predominantly negative in their assessment. The statement was: “Understanding the difference between the two warning sounds was easy”. Here the participants reached a most common answer of 1 (*SD*=1.61) corresponding to “Strongly disagree”. The comparably large standard deviation of 1.61 tells us there was a large spread in the answers, and indeed answers ranged from 1 to 5, telling us that the participants did not all agree with the negative assessment. Four participants noted that they had not noticed the different sounds, with one of them noting that this was something you could learn over time, and would in turn help you not to have to look down at the system.

The seventh statement was: “When the warnings flashed it had a distracting effect on my driving” where the most common answer from the participants was 2 – “disagree” (*SD*=0.90). One participant stated that the warning flashing state worked well, another disagreed stating the opposite. Another participant remarked that the flashing warnings were disrupting to driving in some cases.

Statement number eight ”The system would *not* make me a safer driver” gave a most common answer of 3 “Neither agree nor disagree” (*SD*=0.80), with participants noting that warnings about traffic jams and slippery roads were useful, and could have an effect on their driving. In the post-experiment interview, a participant furthermore stated that the system would make her more attentive while driving.

The answers to the ninth statement, “I found the warnings from the system relevant”, produced a most common score of 4 – “agree” (*SD*=0.72). Several participants commented on this statement, with one of them stating that different categories of warnings relating to the degree of curves would be preferable. Another noted that sometimes the course of the circuit was obvious, and that the warnings therefore were less relevant. Yet another noted that only left/right curve and speed limit warnings were relevant, and furthermore they should not be accompanied by auditory warnings. Yet another participant noted that the warnings would be especially relevant when encountering unexpected curves on small roads.

The tenth statement: “The system provided me with too much information to an extent where it confused me” was answered with a most common answer of 2 – “disagree” (*SD*=1), with one participant noting that the system sometimes caused attention to be shifted away from the road and towards the system.

Conclusion

The number of information systems in our cars has increased rapidly in recent years and more advanced systems are appearing all the time. While these systems arguably raise the standard for what is possible in cars, designers have to consider the ramifications on traffic safety. While existing in-vehicle systems designed for entertainment and convenience tend to have a negative effect on safety, in-vehicle systems are emerging that are designed to proactively enhance safety. Regardless of the purpose of the in-vehicle system, a better understanding of their impact on driving is needed in order to design safer in-vehicle systems.

In this master's thesis we have sought to shed light on the effects of in-vehicle systems on driving, in particular driving behavior and attention, summarized by the following question:

- How do in-vehicle systems affect driving behavior and attention?

This question is answered by answering our two sub-questions that deal with in-vehicle systems that have different purposes. The first sub-question deals with in-vehicle systems that are designed to assist the driver in the primary driving task:

1. How does an in-vehicle system intended to support the primary driving task affect driving behavior and attention?

In order to answer this question, we evaluated the effects of an in-vehicle advance warning system on driver behavior and attention in an experiment conducted on a closed circuit. This enabled us to study how advance warnings about various road and traffic conditions influences the speed and eye glance behavior of drivers. The experiment showed that the in-vehicle advance warning system had an overall limited effect on the driving behavior of the participants. The advance warnings did not cause a significant reduction in speed, as the speeds were often remarkably similar across the two conditions. However, the advance warnings had a larger effect on the speed for some road and traffic conditions compared to others. In particular, advance warnings about speed limit zones and slippery road conditions proved to have a more pronounced effect. On the other hand, the advance warnings gave rise to significantly more eye glances as the participants glanced at the warning system, diverting their attention away from the road more often. In particular, when warnings were emitted the participants shifted attention to the warning system. In answer to the question posed, we found that an in-vehicle system designed to support the primary driving task overall had a limited effect on driving behavior, and a more pronounced effect on attention.

As most in-vehicle systems are not concerned with the primary driving task, but rather secondary driving tasks, we ask this second question:

1. How do different combinations of input and output methods for in-vehicle systems used as a secondary task affect driving performance and attention?

We answered this question by developing an in-vehicle music player with four different combinations of input and output mechanisms and evaluating them using a driving simulator. We then analyzed the primary driving task performance, secondary driving task performance and eye glance behavior for each combination. The two input methods used were touch and gesture, and the output methods visual and audio. Our results showed that gesture-based input resulted in significantly fewer eye glances when compared to touch-based input, but the results also show that gesture-based interaction entailed inferior primary driving task performance and longer task completion times. Audio output caused the participants to make more errors in maintaining speed compared to visual output, and had a significantly longer task completion time. Visual output proved to cause significantly more interaction errors with the system and imposed a drastically higher number of eye glances away from the road. When looking at the individual input/output configurations, our results show that the gesture/audio configuration has by far the fewest number of eye glances, but also a longer task completion time and more speed maintenance errors when compared to any other configuration. The answer to the question must be then, that different combinations of input and output methods affect driving performance and attention differently, and that when addressing the design of in-vehicle systems used as a secondary driving task, it is important to consider input and output as separate concepts.

Limitations

In our second experiment, regarding different input/output techniques we used a between-subject design. There were several reasons for this, the primary being that we had four different configurations in the experiment, thus using a within-subject design would have had the potential of leading to a significant learning effect. Since the tasks did not vary between configurations, reuse of participants would arguable have caused them to perform better the fourth time they were asked to solve the same set of tasks. However, this carryover effect could have been minimized by careful task design, for instance by allowing task solutions to change dynamically. Doing so could have caused the data to vary less on account of individual variance, since some of the participants behaved quite differently. Furthermore, having each participant try all four configurations would have added the possibility of getting their comparative opinion. In the experiment concerning advance warnings, on the other hand, we used a within-subject design. One of the reasons for this was that the comparison between the two conditions would have a lesser error variance due to individual differences, as the results are paired for each participant. As a disadvantage, there is the potential of a carryover effect in terms of learning between the two sessions, which could have caused the participants to perform differently in the second of the two sessions.

During our experiment regarding advance warnings, we issued 24 warnings per session, which corresponded to about two warnings per minute. This might have been too often, causing the participants to anticipate the warnings. Therefore it would be interesting to investigate if there are any change in the effects on drivers if the interval of the warnings is much less, and perhaps only for certain road hazards. Similarly, the long-term effects of advance warnings would be a relevant research topic in order to establish what, if anything, happens to the effects on driver behavior over time. Furthermore, we issued the advance warnings 75 m before the incident, which could have been too late, as the participants might already have perceived the upcoming road condition and already started to react accordingly. We based the distance before the warning on guidelines from the Danish transport authorities, who recommend placing traffic signs 50 m before the object or situation they are intended to warn about when speeds are within 30-60 km/h, plus a two second reaction time to allow drivers to read and decode the traffic sign. We set this two-second reaction time to be 25 m, based on an average of the reaction times recommended for speeds of 30, 40, 50 and 60 km/h. This might have been too little since the participants who actually traveled at 60 km/h and ideally should have had 33 m to react to the sign, only had 25 m to react, corresponding to a reaction time of 1.5 seconds.

As mentioned earlier the age spread in participants was somewhat limited at 24-30 years of age, which is not representative of the driving population. This has the potential of influencing our results in different ways. Younger drivers, who have less driving experience may react more pronouncedly to the warnings, could have caused our results to indicate a more significant effect of the advance warnings. In continuation of this, having older and more experienced drivers could likewise affect the effects of the advance warnings. The potential increase in driving experience for older drivers could arguably effect the reaction to advance warnings in the way that they rely more on their experience and less on the advance warnings, leading to a decrease in the effects of advance warnings. Furthermore there is arguably a difference in the participants approach to in-vehicle system with respect to age.

Further work

As mentioned in our limitations, we suspect that our warnings might have been issued to late, which could have reduced their effect. Therefore it could be interesting to conduct another experiment with an increased distance between advance warning and road or traffic incident. Our thoughts regarding the warnings being displayed too early, are shared with one of our test participants, who expressed that she would like the warnings too appear even earlier: *”Would you use the system if you had it your own car? – Yes, most definitely, if the warnings came 100 meters before…”*. If the warnings were issued earlier then 75 meters in advance, it could entail that drivers using an advance warning system would reduce their speed earlier, because the increased distance would cause them to rely on the warnings instead of the context information, as they would with no warnings.

Studying advance warnings over a longer duration, with warnings being emitted even earlier, could furthermore uncover the long-term effect of advance warning systems, in regards to possible changes in speed, determining which traffic and road conditions are most useful to warn about and if the number of eye glances decreases with prolonged use. An observation over a longer duration would result in an approximate insight of how advance warning systems could perform in everyday driving.

References

1) Alm, H., & Nilsson, L. (1995). The effects of a mobile telephone task on driver behaviour in a car following situation. Accident Analysis & Prevention, 27(5), 707-715.
2) Bach, K. M., Jæger, M., Skov, M. B., & Thomassen, N. G. (2007). A Classification of In-Vehicle Systems Research: Understanding, Measuring and Evaluating Attention. *HCI Lab Technical Report no. 2007/2* .
3) Bach, K. M., Jæger, M., Skov, M. B., & Thomassen, N. G. (2008). Evaluating Driver Attention and Driving Behavior: Comparing Controlled Driving and Simulated Driving. *HCI2008 - Culture, Creativity, Interaction.* British Computer Society.
4) Bach, K. M., Jæger, M., Skov, M. B., & Thomassen, N. G. (2008). You Can Touch, but You Can't Look: Interaction with In-vehicle Systems. *Proceeding of CHI '08* , pp. 1139-1140.
5) Ben-Yaacov, A., Maltz, M., & Shinar, D. (2002). Effects of an in-vehicle collision avoidance warning system on short- and long-term driving performance. *Human Factors , 44* (2), 335-42.
6) Bliss, J., & Acton, S. (2003). Alarm mistrust in automobiles: how collision alarm reliability affects driving. *Applied Ergonomics , 34*, 499–509.
7) Brookhuis, K. A., De Vries, G., & De Waard, D. (1991). The effects of mobile telephoning on driving performance. Accident Analysis & Prevention, 23, 309-316.
8) Brooks, C. & Rakotonirainy, A. (2007). *In-vehicle technologies, Advanced Driver Assistance Systems and driver distraction: Research challenges*. In: I.J. Faulks, M. Regan, M. Stevenson, J. Brown, A. Porter & J.D. Irwin (Eds.). *Distracted driving*. Sydney, NSW: Australasian College of Road Safety. Pages 471-486.
9) Broy, M. (2006). Challenges in Automotive Software Engineering in ICSE. ACM Press.
10) Caird, J., Chisholm, S., & Lockhart, J. (2008). Do in-vehicle advanced signs enhance older and younger drivers' intersection performance? Driving simulation and eye movement results. *International Journal of Human-Computer Studies. 66.* Academic Press.
11) Chewar, C. M., McCrickard, D. S., Ndiwalana, A., North, C., Pryor, J., & Tessendorf, D. (2002). Secondary task display attributes: optimizing visualizations for cognitive task suitability and interference avoidance. *Proceeding of the symposium on Data Visualisation (VisSum '02), Eurographics Association* , pp. 165-171.
12) Chisholm, S., Caird, J., Lockhart, J., Fern, L., & Teteris, E. (2007). DRIVING PERFORMANCE WHILE ENGAGED IN MP-3 PLAYER INTERACTION: EFFECTS OF PRACTICE AND TASK DIFFICULTY ON PRT AND EYE MOVEMENTS. *PROCEEDINGS of the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.*
13) Eysenck, M. W. (2001). Principles of Cognitive Psychology (2nd ed.). Psychology Press.

14) Green, P. (2001). Variations in Task Performance Between Younger and Older Drivers: UMTRI Research on Telematics in *Association for the Advancement of Automotive Medicine Conference on Aging and Driving, Southfield, Michigan.*
15) Green, P. (2004). *Driver Distraction, Telematics Design, and Workload Managers: Safety Issues and Solutions*. University of Michigan Transportation Research Institute. SAE International -21-0022.
16) Gupta, N., Bisantz, A., & Singh, T. (2001). INVESTIGATION OF FACTORS AFFECTING DRIVER PERFORMANCE USING ADVERSE CONDITION WARNING SYSTEMS. *In Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting.*
17) Ho, A., Cummings, M. L., Wang, E., Tijerina, L., Kochhar, & D.S. (2005). Integrating Intelligent Driver Warning Systems: Effects of Multiple Alarms and Distraction on Driver Performance. *TRB 2006 Annual Meeting CD-ROM* .
18) Jonsson, I., Nass, C., Endo, J., Reaves, B., Harris, H., Le Ta, J., et al. (2004). Don't Blame me I am Only the Driver: Impact of Blame Attribution on Attitudes and Attention to Driving Task. ACM.
19) Kaptein, N., Theeuwes, J., & Van Der Horst, R. (1996). *Driving simulator validity: some considerations.* Transportation Research Record 1550.
20) Kemeny, A., & Panerai, F. (2003). Evaluating perception in driving simulation experiments. *TRENDS in Cognitive Sciences , 7* (1), 31-37.
21) Lansdown, T. C., Brooks-Carter, N., & Kersloot, T. (2004). Distraction from Multiple In-Vehicle Secondary Tasks: Vehicle Performance and Mental Workload Implications. *Ergonomics , 47* (1), pp. 91-104.
22) Lees, M., & Lee, J. (2007). The influence of distraction and driving context on driver response to imperfect collision warning systems. *Ergonomics , 50* (8), 1264–1286.
23) Leshed, G., Velden, T., Rieger, O., Kot, B., & Sengers, P. (2008). In-Car GPS Navigation: Engagement with and Disengagement from the Environment. *CHI 2008 Proceedings.* ACM.
24) Liu, Y. (2001). Comparative Study of the Effects of Auditory, Visual and Multimodality Displays on Drivers in Ergonomics, 2001, Vol. 44, No.4, Pages 425-442. Taylor & Francis Ltd.
25) Maltz, M., & Shinar, D. (2004). Imperfect In-Vehicle Collision Avoidance Warning Systems Can Aid Drivers. *Human Factors , 46* (2), 357–366.
26) McKnight, A. J. & McKnight, A. S. (1993). The effect of cellular phone use upon driver attention. Accident Analysis & Prevention, 25(3), 259-265.
27) NHTSA. (2009). *Early Estimate of Motor Vehicle Traffic Fatalities in 2008.* Administration, National Highway Traffic Safety. NHTSA's National Center for Statistics and Analysis.
28) Shinar, D., & Tractinsky, N. (2004). *Effects of Practice on Interference From an Auditory Task While Driving: A Simulation Study.* National Highway Traffic Safety Administration.
29) Stevens, A. (2000). Safety of Driver Interaction with In-Vehicle Information Systems in Proceeding of The Institution of Mechanical Engineers - Parr D -

Journal of Automobile Engineering, Vol. 214. Issue 6, Pages 639-644. Professional Engineering Publishing.

30) Strayer, D. L., & Drews, F. A. (2007). *Cell-Phone Induced Driver Distraction.* University of Utah.

31) TechWatch: Opel Eye Warns Drivers About Speed Limits, 2008, http://www.edmunds.com/insideline/do/News/articleId=127368

Appendix A: Research paper 1

Investigating the Effects of Advance Warning In-Vehicle Systems on Driving Behavior and Attention

Lars Holm Christiansen, Nikolaj Yde Frederiksen, Alex Ranch
Department of Computer Science
Aalborg University
Selma Lagerlöfs Vej 300, 9220 Aalborg East
Denmark
{larshc, nyf, ranch}@cs.aau.dk

ABSTRACT
Information systems are invading our cars, along with concerns about their impact on safety. But there are systems that aim to improve safety, such as advance warning systems. The effects of these systems, however, is a subject of ongoing research. We conducted an experiment on a closed circuit, to determine how advance warnings about various road and traffic conditions affected driver behavior and attention. Our results show that the advance warnings overall had a limited effect on the speed of the drivers, although they had a positive effect for some road conditions. The drivers had significantly more eye glances at the system. In particular, they glanced at the system when a warning was issued.

Author Keywords
Advance warnings, driving behavior, attention, in-vehicle systems.

INTRODUCTION

Car drivers today have more information systems in their vehicles than ever before, and more advanced systems are on their way. The purpose of these in-vehicle systems varies greatly, and includes providing entertainment, offering navigational aid and assisting mobile phone communication. These systems also concern different aspects of driving. Specifically, a distinction is often made between systems dealing with so-called primary driving tasks (related to maneuvering the vehicle, such as GPS systems) and secondary driving tasks (not directly related to maneuvering the vehicle, such as car stereo systems) (e.g. [3, 11].

With the increase in in-vehicle systems as well as the general technological progress, it is possible to present an ever-increasing amount and diverse type of information to the driver. Using the right information in the right manner, it would be possible to assist the primary driving task by supporting the decision-making process and contribute to improving traffic safety. In-vehicle advance warning sign systems are designed to do just this. Their purpose is to provide the driver with up-front information, about conditions ahead, in order to enable the driver to perform more accurate decisions about upcoming maneuvers [6].

The development of the technology that allows cars to gather and react to information about the driving context is well underway [19]. Already a consumer product is available that senses and relays information about the surroundings to the driver. The Opel Insignia car comes with the Opel Eye system that recognizes traffic signs along the road and draws the driver's attention to them via a dashboard-based warning sign system [17].

Although the topic of advance warning signs has been explored, it remains to be determined exactly how the added context information benefits the driver and what effects it has on driving behavior in the long term [6]. At the same time, it is unclear precisely which type of context information promotes traffic safety and which is merely a further distraction [5].

In this paper we investigate the effects of five different in-vehicle advance warning signs on driver behavior and attention. In an experiment on a closed circuit, participants are exposed to several different road and traffic conditions with and without advance warning signs, in order to see how they affect vehicle speed and eye glance behavior. First, we present related work on in-vehicle advance warnings signs. Then we introduce the system used in the experiment and describe the experiment itself. Finally, we present and discuss our results.

RELATED WORK

There have been several previous studies on the potentially beneficial effects of advance warning systems [8, 13, 16]. In-vehicle collision avoidance warning system in particular have received a lot of research attention (e.g. [4, 12]). Lee et al. (2002) examined the effects of collision avoidance systems on driver performance by conducting two experiments in a high-fidelity driving simulator. The results from the first experiment showed that advance warnings helped distracted drivers react more quickly, than with no warnings, and reduced the number of collisions by about 80%. The second experiment showed that undistracted drivers also benefit from the collision warning system, allowing them to brake faster [12]. These results are supported by Ben-Yaacov et al. (2002), who also find beneficial long-term effects of in-vehicle collision avoidance warning systems [4].

Luoma and Rämä (2002) set out to investigate driver acceptance of in-vehicle traffic sign information. In their study, subjects were exposed to four different configurations of traffic sign information in the car, while driving on a real road. This included a visual sign, visual sign with auditory message, visual sign and auditory feedback based on driver behavior and visual sign with elaborate auditory instructions. Luoma and Rämä's study showed that the test subjects accepted the integration of traffic sign information in the car, and generally found it useful in terms of improving traffic safety [13]. Specifically, the visual sign information was rated most useful. However, many of the drivers encountered problems such as unintentional speed decreases and late detection of other road users and obstacles. Similar side effects were found by Hanowski et al (1999) who examined the benefits of a dashboard-based in-vehicle information system that included an advance warning system. Drivers were warned 5 seconds in advance about incidents such as crash ahead, car entering from hidden entrance and emergency vehicle approaching from behind. Despite the minor side effects, the advance warnings were found to indicate a clear benefit to drivers [8].

Intersections are another driving situation where in-vehicle advance warnings are believed to have an impact. Staplin and Fisk (1991) performed a series of studies to determine, if advance information about left curves improved decision performance. These studies were conducted in a laboratory setting, where the participants were faced with decisions about whether to turn or wait at left curve intersections, both with and without advance sign information. The results revealed that the test participants made faster and more accurate decisions when advance information was available [16].

Similarly, Caird et al. conducted a study to determine whether in-vehicle advance warnings could improve the intersection performance of both younger and older drivers. Using a driving simulator, test subjects were exposed to two different Head-Up-Display-based advance warning signs, warning them of upcoming intersections. The results were then compared to baseline drives without any advance warning signs. During all the drives late yellow light changes were randomly interspersed at the intersections. The data recorded included vehicle speed before, during and after the intersections, the number of test subjects that stopped or ran the yellow light, eye movement behavior, as well as the subjects' response time at the late yellow light changes. The results showed that the advance warning signs caused an overall increase in the number of test subjects, who stopped at the late yellow lights. Furthermore, the intersection approach speed for all test subjects was decreased. Caird et al. argues that this caused the test subjects to make more accurate decisions regarding intersection traversal. The primary side effect of the advance warning signs was determined to be a tendency among the drivers to reduce speed in advance of intersections. Based on the results of their study, Caird et al. conclude that drivers who are inattentive or distracted as they approach intersections may benefit from in-vehicle advance sign systems. Similarly, drivers who look, but do not see the intersection ahead, may find advance warning signs helpful. Both Caird et al. and Luoma and Rämä conclude that while in-vehicle advance warnings signs look promising, more research is needed to investigate the effects of this kind of system on driver behavior, especially outside the laboratory [6, 13].

The experiment in [6] was conducted in a driving simulator. While this approach has its advantages, it arguably entails some limitations in the degree of realism associated with the driving and context, a claim that is supported by e.g. Bach et al. (2008), who find that the lack of sensory feedback from the vehicle and context, in simulated driving, causes problems with speed maintenance [2]. Since in-vehicle advance warning signs are so closely related to the driving context, it might therefore be worth studying their effects in a more realistic setting. Motivated by this, we aim to investigate the effects of a variety of context-based advance warning signs on driving performance in a real vehicle on a closed circuit, which should provide a higher degree of realism in terms of contextual information.

ADVANCE WARNING IN-VEHICLE SYSTEM

The system is designed to present the driver with necessary basic functionality in the dashboard to the driver, displaying a speedometer, fuel gauge, trip meter and the cars operating temperature (as depicted on Figure 1) .

The basic functionalities have a central position on the screen and are always visible, whereas the warning signs only are visible, when the driver approaches a given road or traffic condition. The system shows the current speed of the

Figure 1. System interface with two warnings active.

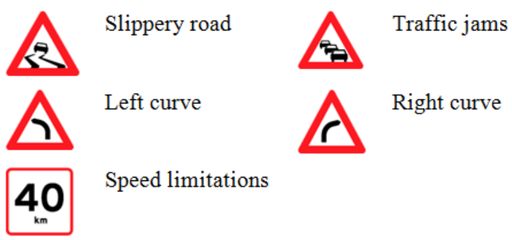


Figure 2. Types of advance warnings used, based on real signs used on Danish roads.

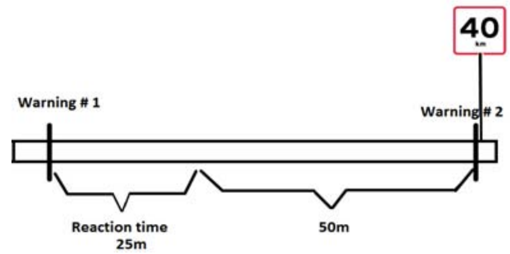


Figure 3. The first warning is displayed 75m before the incident, the second warning at the incident.

car, by means of a Holux GPS unit, placed in the windscreen of the car. The GPS unit receives the current speed every second and through a Bluetooth connection sends it to a laptop which then displays it on the system. The speed is represented in 3x2 cm white digital numbers, on a black background, resulting in easily read numbers in high color contrast. The advance warnings are, when visible, depicted as 1.5 x 1.5 cm icons above the speed representation. The fuel gauge, trip meter and operating temperature are depicted in the lower part of the screen. The system was developed in C# using Microsoft Visual Studio .NET 2008. Besides the aforementioned visual information, the system prompts the driver with audio (earcons) and visual warnings (see Figure 2), regarding the incident ahead. The system displays advance warning signs which are recognizable since they have been chosen to mimic those from real life driving. The advance warnings appear in the dashboard, 75 m before the incident occurs where they are constantly lit, and again, when the incident occurs, when they start to flash. Besides the visual warnings, the system furthermore warns the driver with two different earcons one for each state of warning.

We issued the advance warnings 75m before the incident occurred, based on guidelines from the Danish road directorate, who recommend placing traffic signs 50 meters before the incident for speeds below 60 km/h, plus two seconds of reaction time to allow drivers to read and decode the traffic sign. We set this reaction distance to be 25m, based on an average of the recommended distances at 30, 40, 50 and 60 km/h. As an example see Figure 3 depicting the distance before a 40 km/h sign [18].

EXPERIMENT

The purpose of our experiment was to investigate the effects advance warnings, produced by our in-vehicle system have on participants driving behavior.

Experimental Design

In the experiment, a within-subject approach with twelve participants was adopted. In order to minimize learning effects, the participants were counterbalanced such that they alternated between the two conditions of the experiment, namely switching between starting with advance warnings from the system, and starting without (baseline). The dependent variable of the experiment was primarily driving speed and secondarily eye glance behavior. The independent variables were the aforementioned two conditions of driving with and without advance warnings.

Participants

Our twelve participants consisted of seven male and five female participants with ages ranging between 24 and 30 years (*M*=26.17, *SD*=1.85). All test subjects carried valid drivers' licenses' and had done so for an average of 7.5 years (*SD*=0.65). The participants driving experience, in terms of how many kilometers they drove pr. year, ranged from 50 to 25000 km/year (*M*=6354, *SD*=7205). All of the participants stated that they, to some degree, had prior experience with in-car systems, mostly GPS systems.

Setting

The experiment was conducted on a closed circuit primarily used for driving courses for training future drivers (see Figure 4). The course consisted of various sections, three of which were equipped with water sprinklers along both sides, and were asphalted with a special epoxy/asphalt blend which, when wet, made the road slippery (the white sections on Figure 4). The sprinklers could be turned on and of enabling us to vary the driving condition.

For added realism, another car drove around the circuit in order to simulate real life driving, in the sense that the participants had to be aware of the other car, and act accordingly. The vehicle used in the experiment had been fitted with our advance warning prototype system in such a way that the cars original speedometer was covered and no

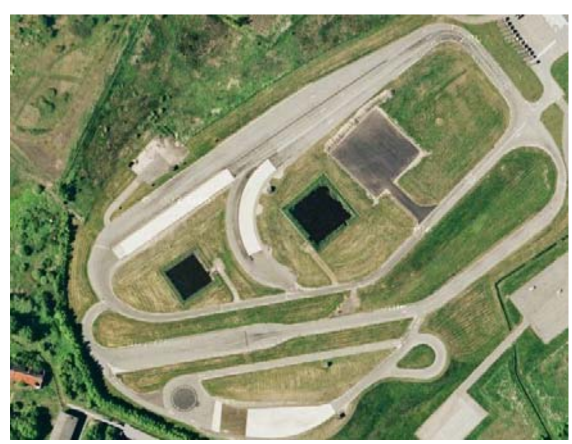


Figure 4. Experimental setting at a closed circuit.

longer visible, as seen in Figure 5.

Procedure

At the onset of each session the participants were seated in the driver's seat of the car, and asked to adjust their seating position according to preference. The supervisor of the experiment then read an introductory text aloud explaining what was going to happen, and which configuration they were going to start with. If the participants were to start with advance warnings engaged, they were shown how the warning system worked, otherwise this was done just before the second part of the session. Furthermore, the participants were instructed to obey the normal traffic regulations with a general speed limit of 60 km/h, and to otherwise drive the car, as they would normally do. Afterwards the participants were given an opportunity to familiarize themselves with the car before starting the experiment.

When the participants felt they were ready, the supervisor instructed them to enter the course and throughout the session instructing the participants which direction to drive. Each session was divided into two parts; (each part around 6.5 km) 8 laps with advance warnings – 8 laps without. The two parts followed the same eight predetermined routes in the same order. During each of the two parts, the water sprinklers were initially turned off on the sections equipped with these, such that they were not slippery the first time around. The water sprinklers were then turned on without the participants knowing, making the two sections slippery, the second time the participants encountered them. The experiment was in part based around a Wizard of Oz approach, as the advance warnings were controlled by an observer, seated in the back of the car. To ensure uniformity we had placed inconspicuous markings, signifying to the observer when to turn on the warnings, around the track.

As previously mentioned one of the authors of this paper drove another car around the circuit during the test in order to simulate real life traffic, and to initiate the traffic jams, which the participants encountered during the experiment. One of the authors of this paper was also responsible for turning the water sprinklers on and off at the right time.

After the driving part of the session was concluded, the participants were interviewed by the supervisor following a semi-structured exploratory interview guide, to collect the participant's thoughts on the use of advance warnings. Finally, the participants were asked to fill out a questionnaire.

Figure 5. Experimental setup.

Data Logging

In our advance warning prototype, we had implemented the possibility to automatically log the information currently in use by the system for instance; current speed, lap number, and advance warning description etc. As previously explained the data logging was initiated 75m before the road or traffic incident, were the observer turned on the warning. When the participants reached the area they had been warned about, 75m later, they were once more warned with a different sound and a flashing icon. In the data logging we distinguished between these two types of warnings, enabling us to analyze the effects of the advance warnings, on the driver's behavior prior to and during the part of the track which the participants had been warned about. The data logging was stopped when the participants exited the section the advance warning had concerned. This procedure amounted in around 650 lines of log data pr. participant with one line of data being logged pr. second a warning was displayed.

Data Analysis

During the experiment we collected two types of data, namely the speed with which the participants drove, and video material to determine eye glance behavior.

Vehicle Speed

Firstly, we analyzed the speed data gathered by the data logging system. As stated earlier, the logging started when an advance warning was first activated (referred to as *state 1* of the incident) through the second state of the warning (*state 2*) until the incident is passed and the warning is turned off (*incident exit*). The equivalent data was also logged without warnings.

The primary focus of the speed analysis was to determine how much the participants reduced their speed between when they received the first warning, to when they reached the incident in question, and then compare this to when they drove without warnings (baseline). In order to do this we needed to determine the average speed, at state 1 onset and state 2 onset for each incident, and then calculate the difference. Additionally, we computed the average speed at incident exit, and the reduction in speed from state 2 onset to incident exit, in order to examine if the warnings had an effect after the participants had reached the incident.

As a secondary aim, we also analyzed how the speed of the participants developed over time, by calculating the average speed for each second based on all the speed readouts logged during each state. Because the speed of the participants, and therefore number of log entries, differed, we needed to select the lowest common denominator in terms of the number of entries. For instance if one of the logs contained five entries for a specific incident, and the remaining logs had seven entries, only the first five entries of all logs were used, ensuring comparable data. Where this method is applied, we refer to the data with the postfix "common", e.g. state 2 onset common.

Finally, we determined the mean speed throughout state 1 and state 2 separately for all incidents, by calculating the average speed readout, using all log entries in each state.

The above-mentioned procedure was carried out for all individual incidents, and then compiled according to incident type. The results were then subjected to two-tailed paired Student's t-tests to reveal any significant differences between the two conditions.

Eye Glance Behavior

Eye glances were identified from video recorded during the experiment, in order to evaluate the effects of advance warnings on eye glance behavior. By analyzing the video from the experiment frame by frame, eye glances were identified and categorized according to the following three categories according to duration, inspired by [3]: (1) 0.5 seconds and below, (2) between 0.5 and 2.0 seconds, (3) above 2.0 seconds.

We defined the duration of an eye glance as the time between when a participant moved his or her gaze away from the road and onto the system, and back onto the road. During the video analysis, eight videos were reviewed (due to quality issues four videos were discarded) by two reviewers who each reviewed five videos causing an overlap of two videos, which were analyzed by both reviewers. Initially the two reviewers corporately analyzed one of the videos in order to ensure consistency in the analysis. An inter-rater reliability test of this analysis (using weighted Cohen's Kappa) gave α =0.84, corresponding to an excellent agreement according to [10].

Reaction Time

We also calculated the participants' reaction time, i.e. the time it took for them to react to a warning, which we defined as the time it took from a warning was emitted, and the participant had glanced down at the system to when the participant had his or her eyes back on the road. This would also enable us to calculate how far they drove during this time, using the logged speed data. Our approach was to count the number of frames in the video during this interval, convert the number of frames to seconds, and then crosscheck the number of seconds with the speed and warnings from the log to determine the reaction time and distance travelled. Eye glances that occurred within two seconds of a warning were assumed to correlate to that warning. This way we were also able to check if the participants were reacting to the warnings at all, and if that reaction changed over time.

RESULTS

In this section we will present the results of our data analysis. First we will present our analysis of the speed data collected by the system, organized according to warning type. Then we present the results of the eye glance and reaction time analyses.

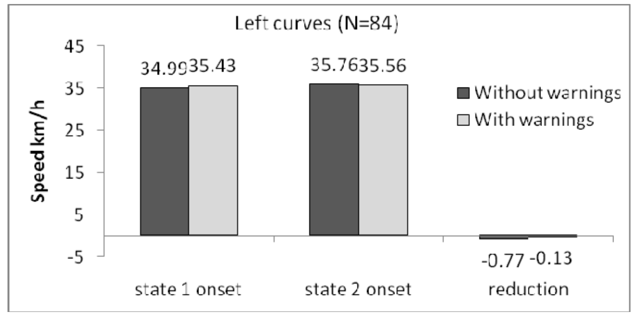


Figure 6. Speed at state 1 onset, state 2 onset and the reduction in speed between the two for left curves (N=84).

Vehicle Speed

In this section we will present the results of our analysis of the vehicle speed data. Emphasis is put on the reduction in speed from state 1 onset to state 2 onset, and any other results will only be presented in-depth if they are of interest. Each warning type is addressed separately by presenting the results for all occurrences of that incident. If relevant, specific incidents in each type will be presented separately.

Left Curve Warnings

The experiment contained seven left curves in each condition (i.e. N=84). As seen on Figure 6, the advance warnings had very little effect for the left curves. The speeds are very similar with and without warnings at state 1 onset, and reductions in speed between state 1 onset and state 2 onset also do not differ notably. Similarly, the mean speeds through state 1 and state 2 differ only marginally. The participants increase their speed before encountering state 2 onset for both conditions, i.e. the reduction is negative. This is caused by a specific left curve on the track that is encountered four times in all, where the entry speed is relatively low due to the layout of the track. However, removing these four left curves from our data, does not affect the results significantly.

Right Curve Warnings

The participants encountered a total of five right curve incidents in each condition (i.e., N=60). As seen on Figure

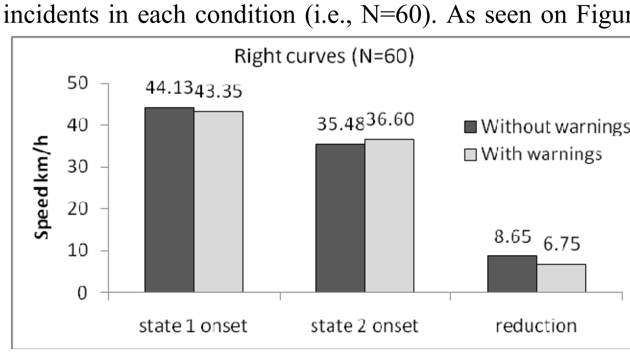


Figure 7. Speed at state 1 onset, state 2 onset and the reduction in speed between the two for right curves (N=60).

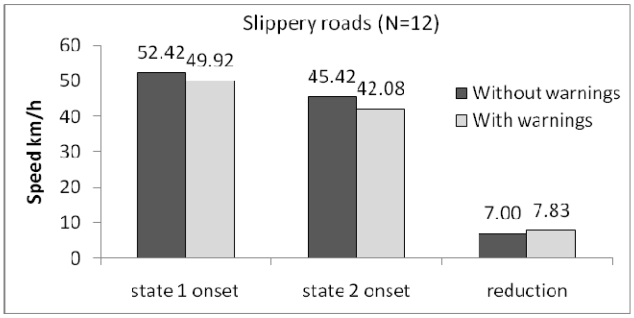


Figure 8. Speed at state 1 onset, state 2 onset and the reduction in speed between the two for slippery roads (N=12).

7 the advance warnings had negative effect here, as the participants reduce their speed less between state 1 onset and state 2 onset with the warnings than without, which amounts to 6.75 km/h (*SD*=5.03) with warnings and 8.65 km/h (*SD*=5.10) without warnings, which a Student's t-test reveals as a significant difference, t = 1.68, p = .026. Comparing their speed at state 2 onset, the average speed with warnings is 36.60 km/h (*SD*=5.88) and 35.48 km/h (*SD*=6.41) without warnings, which reveals signs of a trend, t = 1.68, p = .058. The mean speeds throughout state 1 and state 2 do not differ notably for the right curves.

The track contained two unique right curves, the first of which was encountered four times in each condition (hence N=48). The speeds at state 1 onset for this specific right curve only, are still quite similar, but the speed reduction between state 1 onset and state 2 onset is larger without warnings at 9.81 km/h (*SD*=4.78), than with warnings at 7.31 km/h (*SD*=5.30), which represents a significant difference, t = 2.01, p = .016. This indicates that the difference must lie at state 2 onset, where the average speed with warnings is 34.31 km/h (*SD*=3.43) versus 33.13 km/h (*SD*=4.11) without warnings, which indeed indicates a trend, t = 2.01, p = .054. If we then look at the average speed reduction between state 2 onset and incident exit for this right curve, we also find a significant difference, t = 2.01, p = .026. This time, however, the decrease is larger with warnings at 1.67 km/h (*SD*=5.02) than it is without warnings at just 0.06 km/h (*SD*=4.35).

Slippery Road Warnings

The advance warnings had an effect for slippery road incidents, as shown on Figure 8. The participants drive 2.5 km/h slower with warnings at state 1 onset, than they do without warnings. This is not statistically significant however, t = 2.20, p = .307. The reduction in speed between state 1 onset and state 2 onset differs only marginally. At state 2 onset, however, the average speed with warnings is 42.08 km/h (*SD*=7.48) versus 45.42 km/h (*SD*=8.51) without warnings, which suggests a trend, t = 2.20, p = .054. The pattern continues at incident exit, where the average speed with warnings is 34.75 km/h (*SD*=4.56) and 37.25 km/h (*SD*=6.11) without warnings. A Student's t-test reveals this difference of 2.50 km/h to be significant, t = 2.20, p = .046.

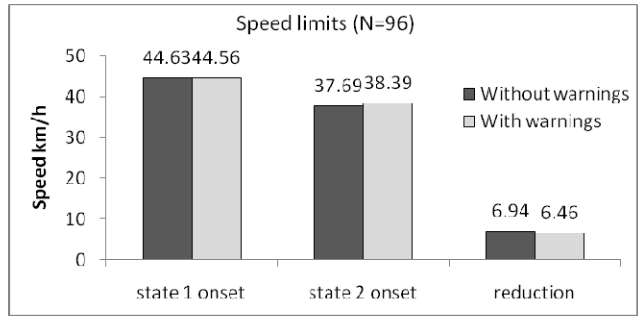


Figure 9. Speed at state 1 onset, state 2 onset and the reduction in speed between the two for speed limit (N=96).

Looking at the common log entries for incident exit, we see a strong significant difference between the average speed with warnings at 35.67 km/h (*SD*=5.15) and without warnings at 40.58 km/h (*SD*=6.85), t = 2.20, p = .004. However, the reduction in speed between state 2 onset common and exit common does not quite constitute a significant difference, t = 2.20, p = .086, with an average reduction of 6.42 km/h (*SD*=3.73) with warnings, and 4.83 km/h (*SD*=3.49) without warnings.

Throughout state 1 the mean speed is lower with warnings than without warnings, but even more so through state 2, where the mean speed is 36.56 km/h (*SD*=6.29) with warnings, and at 41.55 km/h (*SD*=7.61), somewhat higher without warnings. This difference of almost 5 km/h is statistically strong significant, t = 2.20, p = .004.

Speed Limit Warnings

For all eight speed limit incidents (N=96), we see that the average speeds at state 1 onset are very similar, differing only by 0.07 km/h (see Figure 9). The reductions in speed between state 1 onset and state 2 onset are also very similar. At state 2 onset, the speed is lower with warnings at 37.69 km/h (*SD*=3.94) than it was without warnings at 38.39km/h (*SD*=3.62). The difference suggests a trend, t = 1.99, p = .073. The results for the mean speeds throughout state 1 and 2 of do not reveal any noteworthy differences.

As the experiment contained two different speed limit sections, we can look at the results for the first speed limit separately (N=72), where the effects of the advance warnings were more substantial. Here, the decrease in speed between state 1 onset and state 2 onset is 11.89 km/h

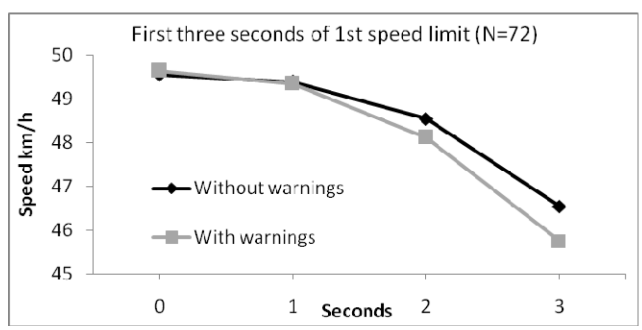


Figure 10. Speed for the first four seconds of state 1 of the first speed limit incident (N=72).

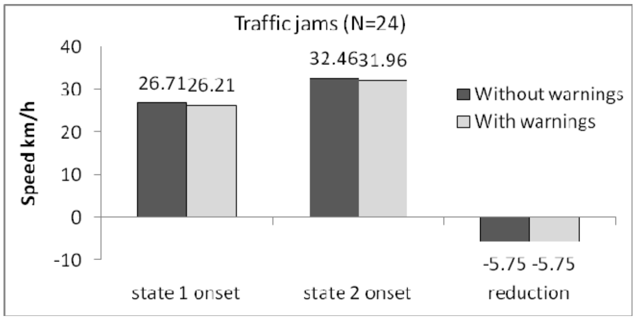


Figure 11. Speed at state 1 onset, state 2 onset and the reduction in speed between the two for traffic jam (N=24).

(*SD*=4.56) with warnings and 11.07 km/h (*SD*=6.08) without warnings, a difference that implies a trend, $t = 1.99$, $p = .063$. At state 2 onset the speed is significantly slower with warnings at 37.75 km/h (*SD*=3.74) compared to 38.86 km/h (*SD*=3.82) without warnings, $t = 1.99$, $p = .013$.

Figure 10 shows the average speed of the participants for the first four common log entries of the first speed limit incident. It shows that while the initial speeds are almost identical, the speeds develop differently over the course of the following three seconds. With warnings the participants reduce their speed earlier than they do without warnings. Indeed, there is a significant difference, $t = 1.99$, $p = .018$, in the average reduction in speed between state 1 onset and the last common speed reading, with an average decrease of 8.13 km/h (*SD*=5.45) with warnings and 6.63 km/h (*SD*=5.54) without warnings. Similarly, when looking at the speed after three seconds, there is a strong significant difference, $t = 1.99$, $p = .0012$, between the average speed of 41.50 km/h (*SD*=4.26) with warnings, and 42.91 km/h (*SD*=3.78) without warnings.

Throughout state 2 of the first speed limit the mean speed with warnings is 37.59 km/h (*SD*=3.72) while it is 38.53 km/h (*SD*=4.06) without warnings, which indicates a trend, $t = 1.99$, $p = .053$. The mean speeds through state 1, on the other hand, do not differ notably.

Traffic Jam Warnings

During the experiment the participants encountered two different traffic jam incidents. Looking at the results of both traffic jams collectively (N=24), the average speeds at state 1 onset are remarkably similar across both conditions, and the average reduction in speed from state 1 onset to state 2 onset is exactly the same across the two conditions, as seen on Figure 11. Similarly, there are no noteworthy differences in the mean speeds throughout state 1 and state 2.

If we look at the first traffic jam incident separately, we see that the advance warnings had a negative effect on the speed reduction. The participants increase their speed from state 1 onset to state 2 onset by 2.58 km/h (*SD*=3.75) with warnings and decrease it by 0.50 km/h (*SD*=3.73) without, which represents a significant difference, $t = 2.20$, $p = .048$.

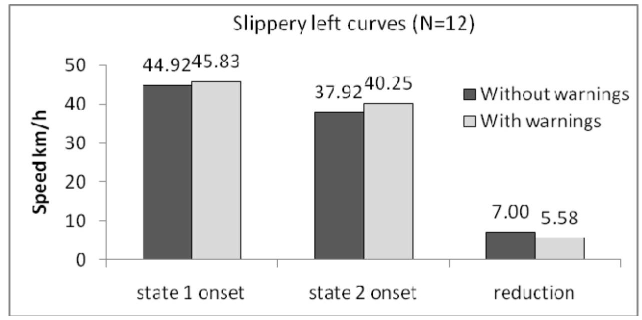


Figure 12. Speed at state 1 onset, state 2 onset and the reduction in speed between the two for slippery left curves (N=12).

Combination of Slippery Road and Left Curve Warnings

Our experiment contained one incident where a left curve was combined with a slippery road and where two warnings were activated simultaneously. Here, none of the results show any significant differences. At state 1 onset, the average speeds were quite similar for both conditions (see Figure 12). The reduction in speed from state 1 onset to state 2 onset is 7.00 km/h (*SD*=5.74) with warnings, which is somewhat larger than without warnings at 5.58 km/h (*SD*=6.89). A difference that is only marginal, $t = 2.20$, $p = .462$. With warnings the average speed at incident exit is slightly higher than without warnings at 32.50 km/h (*SD*=3.09) and 31.42 km/h (*SD*=3.37), respectively. While this difference is not significant, we do see a trend, $t = 2.20$, $p = .090$.

Eye Glance Behavior

We classified the eye glances according to their duration; below 0.5 seconds, between 0.5 and 2.0 seconds and above 2.0 seconds.

From the analysis of the 16 video sessions, we identified a total of 1363 eye glances. Of these, 451 were categorized as being shorter than 0.5 seconds, 911 were categorized as being between 0.5 and 2.0 seconds in length and just one eye glance was longer than 2.0 seconds. Without warnings had the least number of glances of the two conditions, with 494 in all, compared to 869 with warnings, which is an increase of nearly 76%. As seen in Table 1 the average number of glances is higher with warnings than it is without. This amounts to a strong significant difference, t =

	Without warnings (N=8)	With warnings (N=8)
> 0.5 s.	28.5 (20.79)	27.88 (16.03)
0.5 - 2.0 s.	**33.13 (18.29)**	**80.75 (24.34)**
> 2.0 s.	0.13 (0.35)	0 (0)
Total glances	**61.75 (34.69)**	**108.63 (19.43)**

Table 1. Means (standard deviations) for eye glance behavior. Statistically significant differences at the 95% confidence level are highlighted.

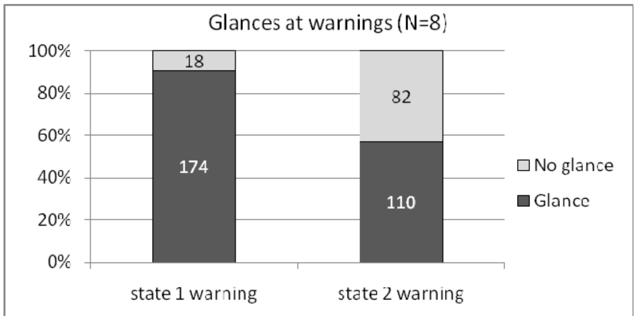


Figure 13. The total number of times participants glance at the system when there is a warning (N=8).

2.36, p = .0065.

The number of glances below 0.5 seconds in duration is very similar for the two conditions, with 228 without warnings and 223 with warnings. Our results show that it is in the number of glances between 0.5 and 2.0 seconds that the real difference lies. With warnings accounts for 646 eye glances in this category, compared to just 265 without warnings. A Student's t-test reveals this to be a strong significant difference, t = 2.36, p = .0028. In the last category, glances above 2.0 seconds, we found just one without warnings and none with warnings.

We also looked at the number of times the participants looked down when the warning system displayed a warning, for both state 1 and state 2. Figure 13 illustrates how many times they glanced at the system when there was a warning, for both warning state 1 and warnings state 2. For more than 91% of the warnings, the participants glance at the system following the state 1 warning. The same holds true for about 57% of the warnings at state 2. A Chi-square test shows this difference to be extreme significant, $\chi^2(1, N=8) = 55.38, p < .0001$.

For those warnings where the participants looked down at the system, we also measured their reaction time, i.e. the time that passed from when the warning was emitted to when the driver had glanced down and back up. The average reaction time was calculated to be 0.92 seconds (*SD*=0.19). Based on their speed during these glances, we were also able to calculate how far they drove during the reaction time. This distance varied between 1.67 m and 30.61 m, and was on average 10.30 m (*SD*=1.79) for N=8.

DISCUSSION

Overall, our results show that the advance warnings had some effects on driving behavior, but that they were limited. When comparing the participants' speed for the two conditions, they are remarkably alike, often with almost indistinguishable differences. However, we did identify several situations where the advance warnings had a positive effect on the speed, though not for all warning types.

For instance, at the first, and most prevalent, speed limit incident, the results show that the speeds develop differently for the two conditions in the first three seconds. With warnings the participants decrease their speed significantly more than without warnings. Another substantial result is the mean speed through state 2 of the slippery road incident. Here, the participants drove almost 5 km/h slower with warnings, than they did without. Other results also show that the advance warnings had a positive effect on the participants' speed for the slippery road.

On the other hand, the effect of the advance warnings are negated in the right curves, since there is a significant difference in the average reduction in speed, which is higher in the no warning condition. While this may seem curious, it is perhaps an indication, that the many of the factors that influence the behavior of drivers are complex in nature.

Our results seem to question the findings of similar studies [1, 5, 6]. Caird et al., for instance, found that advance warnings significantly reduced the speed adopted by drivers through intersections [6]. Their findings may be different from ours because their experiment was conducted in a simulator, which does not give the participants the same amount of context information as real life or controlled driving does. This could make the participants more likely to react to advance warnings regardless of whether or not it is necessary, since the sensation of speed provided by the context is not present. This is supported by Bach et al., who compared simulated and controlled driving, and found that that the lack of sensory feedback from the vehicle and context in simulated driving, caused problems with perceiving driving speed [2].

Furthermore, the fidelity of simulators means that participants are able to see less of their surroundings, which in turn may cause them to rely more on warnings. In contrast, participants in controlled driving are able to rely more on the available context information and therefore the warnings may not have the same effect. However, Kemeny and Panerai state that in driving simulators with a large enough field of view, speed can be estimated correctly by visual information [9]. But Kemeny and Panerai also note that recent studies have shown that vestibular information has a more important role than previously assumed. Additionally, Kemeny and Panerai state that experiments regarding driver alertness, as in Caird et al., can be carried out in driving simulators without the need for absolute simulation fidelity. Moreover, the amount of risk perceived by the participants is arguably bound to be higher in real life driving. A study by Boyle and Mannering into the impact of travel advisory systems on driving speed, suggests that while the average speed can be reduced by in-vehicle system advisory messages, drivers tend to try and make up for lost time by increasing their speed when the warning/advisory message is no longer relevant, which questions the net safety effects of advisory messages [5].

Our results indicate that the primary influence on the driving behavior adopted by the participants is the context in which they drive. The fact that the speeds are so similar in the two conditions, could be an expression of the participants' prior driving experience and training, which helps them establish a reasonable speed to adopt. This means that their behavior could also be explained by how far in advance the warning was issued. If the warning is issued too late, the participants could have already made decisions based on context information and started to react accordingly, rendering the advance warning useless. As described earlier, we issued the warnings 50m in advance plus an additional two-second or 25m reaction time. This might have been too little, since the participants who actually traveled at 60 km/h ideally and should have had 33m to react, only had 25m (or 1.5 s). However, our analysis of the reaction time of the participants revealed that they use an average of 0.92 s (10.30m) to react to a warning, which still leaves on average, about 65m to react appropriately. Still, our analysis is based on visual reaction time only and therefore we cannot estimate how long it took them to mentally react to the warnings, although it arguably adds to the overall reaction time.

The incidents warned about were in some cases rather obvious. For instance, the slippery parts of the course were clearly distinguishable from the rest of the course due to their color. Had the road conditions been less obvious, the results could have been different. For example, had the experiment been conducted on public roads during winter with occasionally treacherous road conditions, the participants presumably would have benefitted more from advance warnings, as any unsafe road conditions would be less visible. This is in line with the conclusions of Luoma et al., who found that warnings about black ice conditions, which are harder to spot, has a greater effect on speed and the amount of headway between road users, compared to warnings in snowfall conditions where the hazardous road conditions are clearly evident [14].

The results of the eye glance behavior analysis reveal that the advance warnings attract significantly more glances. This result is perhaps unsurprising given that the visual nature of the warnings arguably attracts more visual attention. The results also show that when the participants are presented with an advance warning, they look down at the system in 91% of the cases, which indicates that the participants do detect the warnings when they are issued. Similarly, Caird et al. found that drivers gazed at 75.4% of all warnings. However, there are also downsides to diverting attention toward the system. As de Waard et al. notes, whenever an in-vehicle system issues a notification, attention has to be allocated towards processing the information, and subsequently also towards behavioral adaptation. De Waard et al. gives the example of a speed violation message, which can only be prevented by driving below the speed limit and thereby allocating more attention to checking the speedometer more frequently [7]. Significantly fewer glances accompany the state 2 warnings. This difference indicates that the participants are preoccupied with driving the car through the incident safely, and therefore do not notice the state 2 warnings or choose not to divert attention to them. Many of the participants also stated afterwards that they in general did not notice the state 2 warnings or did not find them useful when they did, because they warned about conditions that were obvious.

The age group of the participants in our study (24-30 years of age) was selected because young people usually do not yet have a lot of driving experience, and advance warnings could therefore have a more pronounced effect on them. Still, the fact that the age of our participants is not completely representative of the entire driving population, has potentially affected our results. In an experiment regarding drivers' ability to divide attention between driving a car (simulated driving) and counting the amount of dots appearing on a screen, Ponds et al. conclude that young (mean age = 27.5) and middle-aged (mean age = 46.7) adults did not differ in the ability to divide attention [15]. However, elderly drivers showed a significantly decreased ability to divide attention. In a study by de Waard et al., the reactions towards an in-car enforcement and tutoring system were measured in young (M=37, SD=4.5) and elderly drivers (M=66, SD=3.8). De Waard et al. found that the degree of acceptance of the system was higher for the elderly group of drivers who were pleased with the system. Even though the group of younger drivers believed the system to have a positive effect on traffic safety, they nevertheless disliked it. The findings of de Waard et al. and Ponds et al. indicate that a more balanced and representative age spread in our experiment may have produced different results. For instance, de Waard et al. note that research has shown that elderly drivers overlook traffic signs more often, and speculate that these violations could occur out of inattention [7].

Limitations

Due to the ethical reasons, we chose to conduct our experiment at a closed training circuit instead of in real traffic. While real traffic undoubtedly would have added substantial realism, we found it to be ethically unsound. Testing at a closed circuit furthermore meant that the participants learned the course of the track and the location of the slippery areas, which could entail that participants reacted on behalf of this knowledge.

The experiment was conducted over a period of three days, where we were issued a different car each time. This could have affected our data, as the different cars differed somewhat from one another regarding the amount of wear on the mounted tires, which resulted in one of the cars having less traction compared to the other two, making it perform worse on the slippery parts of the track.

The warnings were manually displayed in the dashboard using the Wizard of Oz technique, which could be cause of inaccuracy, due to the fact that the warnings were displayed when one of the authors, sitting in the back of the car, manually pushed a button when the car passed certain points on the track, indicated by traffic cones.

CONCLUSION

While in-vehicle systems are designed for entertainment and convenience, other systems, such as in-vehicle advance warning systems, are emerging that seek to exploit technology in an attempt to actively improve safety. It is however unsure how these systems affect driving behavior, and what type of warnings are beneficial. In this paper, we have therefore sought to shed light on the effects of in-vehicle advance warning systems on driving behavior and attention. By conducting an experiment on a closed circuit, we were able to compare speed and eye glance behavior data for a series of road and traffic conditions with and without advance warnings.

Overall, our experiment showed that the advance warnings had a limited effect on the behavior of the participants in experiment. The speeds at the different road and traffic incidents were remarkably similar across the two conditions. We did however see some beneficial effects of the advance warnings at particularly the speed limit and slippery road incidents. The use of the advance warnings caused the participants to make significantly more eye glances diverting attention away from the road and onto the in-vehicle system. In particular, they glanced at the system when the warnings were emitted.

There is no doubt that in-vehicle advance warning systems are coming, and that the next couple years will see an increase in products related to this type of system. As context and setting are such important factors for these systems, further research should be conducted that evaluates them in as realistic a setting as possible.

REFERENCES

1. Alm, H., & Nilsson, L. (2000). Incident warning systems and traffic safety: a comparison between the PORTICO and MELYSSA test site systems. Transportation Human Factors, 2 (1), 77-93.
2. Bach, K. M., Jæger, M., Skov, M. B., & Thomassen, N. G. (2008). Evaluating Driver Attention and Driving Behavior: Comparing Controlled Driving and Simulated Driving. HCI2008 - Culture, Creativity, Interaction. British Computer Society.
3. Bach, K. M., Jæger, M., Skov, M. B., & Thomassen, N. G. (2008). You Can Touch, but You Can't Look: Interaction with In-vehicle Systems. Proceeding of CHI '08, pp. 1139-1140.
4. Ben-Yaacov, A., Maltz, M., & Shinar, D. (2002). Effects of an in-vehicle collision avoidance warning system on short- and long-term driving performance. Human Factors, 44 (2), 335-42.
5. Boyle, N., & Mannering, F. (2004). Impact of Traveler Advisory Systems on Driving Speed: Some New Evidence. Transportation Research Part C, 12, 57-72.
6. Caird, J., Chisholm, S., & Lockhart, J. (2008). Do in-vehicle advanced signs enhance older and younger drivers' intersection performance? Driving simulation and eye movement results. International Journal of Human-Computer Studies. 66. Academic Press.
7. de Waard, D., van der Hulst, M., & Brookhuis, K. (1999). Elderly and young drivers' reaction to an in-car enforcement and tutoring system. Applied Ergonomics, 30, 147-157.
8. Hanowski, R., Dingus, T., Gallagher, J., Kieliszewski, C., & Neale, V. (1991). Driver response to in-vehicle warnings. Transportation Human Factors, 1 (1), 91-106.
9. Kemeny, A., & Panerai, F. (2003). Evaluating perception in driving simulation experiments. TRENDS in Cognitive Sciences, 7 (1), 31-37.
10. Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33, 159-174.
11. Lansdown, T. C., Brooks-Carter, N., & Kersloot, T. (2004). Distraction from Multiple In-Vehicle Secondary Tasks: Vehicle Performance and Mental Workload Implications. Ergonomics, 47 (1), pp. 91-104.
12. Lee, J., McGehee, D., Brown, T., & Reyes, M. (2002). Collision Warning Timing, Driver Distraction, and Driver Response to Imminent Rear-End Collisions in a High-Fidelity Driving Simulator. Human Factors, 44 (2), 314-334.
13. Luoma, J., & Rämä, P. (2002). Acceptance of traffic sign information provided by an in-vehicle terminal. Proceedings of the Ninth World Congress on Intelligent Transportation Systems.
14. Luoma, J., Rämä, P., Penttinen, M., & Antilla, V. (2000). Effects of variable message signs for slippery road conditions on reported driver behaviour. Transportation Research Part F , 3, 75-84.
15. Ponds, R., Wiebo, B., & van Wolffelaar, P. (1988). Age Differences in Divided Attention in a Simulated Driving Task. Journal of Gerontology: PSYCHOLOGICAL SCIENCES, 43 (6), 151-156.
16. Staplin, L., & Fisk, A. (1991). A cognitive engineering approach to improving signalized left-turn intersections. Human Factors, 33, 559–572.
17. TechWatch. (2008). Opel Eye Warns Drivers About Speed Limits. Retrieved from http://www.edmunds.com/insideline/do/News/articleId=127368
18. Vejregelrådet. (2001). Færdselsregulering: Færdselstavler. Vejdirektoratet.
19. Vidales, P., & Stajano, F. (2002). The Sentient Car: Context-Aware Automotive Telematics. UBICOMP - Adjunct Proceedings

Appendix B: Research paper 2

Don’t Look at Me – I’m Talking to You: The Effects of Separating Input and Output in In-Vehicle Systems

Lars Holm Christiansen, Nikolaj Yde Frederiksen, Alex Ranch
Department of Computer Science
Aalborg University
Selma Lagerlöfs Vej 300, 9220 Aalborg East
Denmark
{larshc, nyf, ranch}@cs.aau.dk

ABSTRACT
As more information systems make their way into cars, designers of these systems are faced with unique challenges in order to minimize the negative effects on driving performance. Previous research into this field tends to focus on input. In our study, we sought to put equal emphasis on both input and output in order to examine their separate effects on driving performance and eye glance behavior. To this end we evaluated four combinations of input methods (touch and gesture) and output methods (visual and audio) in a driving simulator. Our results showed that the separation of input and output is a non-trivial one. Gesture input resulted in significantly fewer eye glances compared to touch input, but also worse primary driving task performance. Using audio as output caused a drastically lower number of eye glances, but significantly longer task completion times and inferior primary driving task performance compared to visual output.

Author Keywords
Gesture interaction, touch interaction, attention, eye glances, auditory feedback, in-car systems.

INTRODUCTION
As an increasing amount of technology is incorporated into new cars, research into the design of in-vehicle interaction becomes increasingly relevant, as these non-essential systems become present in the driving context. With the driver’s attention being divided between devices such as GPS-systems, onboard computers, music players etc. several studies have been conducted regarding driver distraction, which find that driver distraction is problematic and can lead to accidents [9, 4].

One of the problems is that the driver has to remove attention from his primary task of driving the car, in order to perform secondary tasks such as changing the radio station. As the amount and complexity of in-vehicle systems increase, so does the demands on the driver’s attention. This presents a series of challenges when attempting to make this interaction as quick and efficient as possible in order to minimize the amount of time the driver has to remove attention from the road, in particular the driver’s visual attention [2, 3].

One specific interaction technology repeatedly seen in a variety of in-vehicle systems is the touch screen. The flexibility in its application capabilities, low price, and utilization of a more natural way of interaction, makes it an obvious choice for in-vehicle systems, with its increasing presence in new cars and aftermarket GPS-units. But the nature of a touch screen also means that it competes for the visual attention of the driver.

In order to evaluate the effects of in-vehicle interaction, we create a touch screen-based system to be used in an experiment where we separate input and output techniques, so as to evaluate their separate influence on driving performance and attention.

In this paper we present this experiment and our results. The paper is structured as follows; initially we present previous research on in-vehicle systems, and secondly we introduce the interaction techniques. Then we describe the experiment and the results are presented. Finally, the results are discussed.

RELATED WORK
When researching the field of vehicle safety and the use of in-vehicle systems, it is crucial to address the concepts of attention and distraction. Attention can be defined as the ability to concentrate and selectively focus or shift focus between selected stimuli [13]. In the vehicle domain, the driver’s attention is mainly focused on monitoring the environment and executing maneuvers, also called the primary driving task [2, 5, 9, 11]. Disruption of attention is defined as distraction. Green describes distraction in the vehicle domain as anything that grabs and retains the

attention of the driver, shifting focus away from the primary driving task [2, 4, 9].

A specific focus within the field of in-vehicle research, concerning distraction and attention, is the dynamics between the primary driving task and secondary driving tasks, which includes use of in-vehicle systems. This is significant since research identifies the use of in-vehicle systems as a cause of traffic accidents [4, 9]. Green points out that most drivers will go to great lengths to complete a given secondary task and rarely abandon a task after initiation [9]. With a critical primary task, like driving, this seemingly irrational behavior and distribution of attention between the primary and secondary task, can in worst case scenarios endanger the safety of the driver and the surroundings. Lansdown et al. acknowledges this troubling tendency concerning in-vehicle systems, in a study focusing on distraction imposed by in-vehicle secondary systems [11].

A tendency within in-vehicle interaction research involves attempts to identify an interaction technique that surpasses the capabilities of the traditional tactile interface. In a comparative study, Geiger et al. set out to evaluate the use of dynamic hand movements (gestures) in order to operate a secondary in-car system and compare it to a traditional haptic (tactile) interface [8]. The parameters used for comparison, were errors related to driving performance, tactile/gesture recognition performance and the amount of time drivers didn't have their hands on the steering wheel. The experiment showed that use of the tactile interface resulted in high task completion times and the system lacked in recognition performance when compared to the gesture interface. The gesture interface allowed users to perform the primary task appropriately, while the users also found the gesture interface more pleasant and less distracting. A recent study by Alpern & Minardo supports these findings [1]. They set out to evaluate gestures through an iterative development of an interface for performing secondary tasks. In the final iteration of their experiment, they noted that users made fewer errors compared to a traditional tactile radio interface. Findings from both studies indicate that gestures could be a viable alternative for secondary in-car systems.

Bach et al. sought to shed light on how perceptual and task-specific resources are allocated while operating audio systems in a vehicle [3]. Three system configurations – a conventional tactile car stereo, a touch interface and an interface that recognizes gestures as input – were evaluated in two complementary experiments. The experiments suggest an overall preference for the gesture-based configuration, as it enabled the users to reserve their visual attention for controlling the vehicle. The conventional car stereo on the other hand lacked an intuitive interface; consequently the system requires additional perceptual and task-specific resources to be operated, thereby affecting primary task performance. The touch interface introduced a reduction in overall task completion time and interaction errors when compared to both the conventional tactile and gesture interfaces.

While the future prospect of using gestures as an input method for in-vehicle systems seems promising, little attention is given to the possible influence of output methods. In order to address this it would be necessary to distinguish between input and output to clarify how combinations of different output and input methods might affect the interaction and primary task performance. The need to separate output from input in relation to in-vehicle systems is acknowledged by Bach et al. as a limitation in their study, and the need for further research is recognized. Their primary research focus was on system input as opposed to output, which meant the output mechanisms differed for each configuration. The variation in output could have affected the findings – the results do not show which output mechanism is suitable for in-vehicle systems. This suggests the need for an elaborate study on output methods in order to investigate how they influence primary and secondary task performance in the vehicle domain.

The aim of our study is to compare different configurations of in-vehicle systems with an equal emphasis on both input and output mechanisms. We aim to limit the amount of variables influencing input and output, and hereby approximate a comparative study – in order to address the limitations of Bach et al. [3]. We intend to accomplish this through a study of visual and auditory output in combination with either touch or gesture input. The rationale behind this combination is the duality in the interaction possibilities of touch screens, which support both touch and gesture interaction and the polarity in the two different sensory channels of output.

IN-VEHICLE SYSTEM

By distinguishing between two input methods and two output methods, we have four different configurations of our in-vehicle system; touch input with visual output, touch input with audio output, gesture input with visual output and gesture input with audio output. These configurations will hereafter be referred to as <input>/<output>, e.g. touch/visual. In order to evaluate these configurations with regards to their effect on attention, we chose a well-known in-vehicle system as our case; the music player or car stereo. This choice is also inspired by Bach et al. [3] as well as other studies [1, 14] and served as a simple platform for our evaluation.

The system is designed to fit an 8" touch sensitive screen, and the graphical user interface in all configurations is divided into the same output and input areas, to keep the interaction areas the same for all conditions. Furthermore, the output area of the screen is covered by a clear plastic shield to discourage deliberate input and prevent accidental input in this area.

Input

We distinguish between two input methods; conventional touch-screen based input with graphical buttons, and gesture-based input using the touch-screen as a drawing canvas.

The graphical layout of the two touch configurations is inspired by Bach et al. [3] and our goal was to keep it as simple as possible, while still providing the necessary basic functionality. To facilitate easy interpretation the icons on the buttons resemble icons common on music players. Furthermore, the buttons are grouped according to their functionality. The layout includes a "'Song info" button, which is only enabled in the touch/audio configuration, but is included in the touch/visual configuration to keep the design consistent. The size and spacing of the buttons is chosen based on previous research on touch screen layout [6, 16, 17]. Input is only possible by pressing the buttons, which work according to the click-on-release principle. This means that the buttons are activated only when the finger has left the button, which also means that nothing happens when a button is held.

The gesture-based systems have no buttons. Instead, the systems are controlled by gestures drawn directly on the screen using a finger. The gestures used are inspired by Pirhonen et al. [14] and Bach et al. [3] and allow for the same functionality as the touch buttons. The only gesture that is different is the "Song info" gesture, which is performed by drawing a line straight down followed by a line straight up, without the finger leaving the canvas. This was chosen to resemble the "i" often used as an icon for "information". The gestures can be executed anywhere in the input (grey) area of the screen, but not in the output (white) area.

Output

We use two different modes of output; visual output using icons and text and audio output using earcons and voice. Visual and audio output is not used simultaneously at any point. We distinguish between two kinds of output; feedback on input and information about the state of the system.

The visual feedback is implemented using visual cues to inform the user of the result of his or her actions. For the touch/visual system, this is done by changing the appearance of buttons to indicate they have been pressed. Furthermore, when the volume is all the way down, pressing the "Volume down" button will change its appearance to reflect a disabled state. The same principle applies to the "Volume up" button. For the gesture/visual system, the same icons are used to indicate a recognized gesture. The icon corresponding to the recognized gesture will be displayed in the middle of the input area for about a

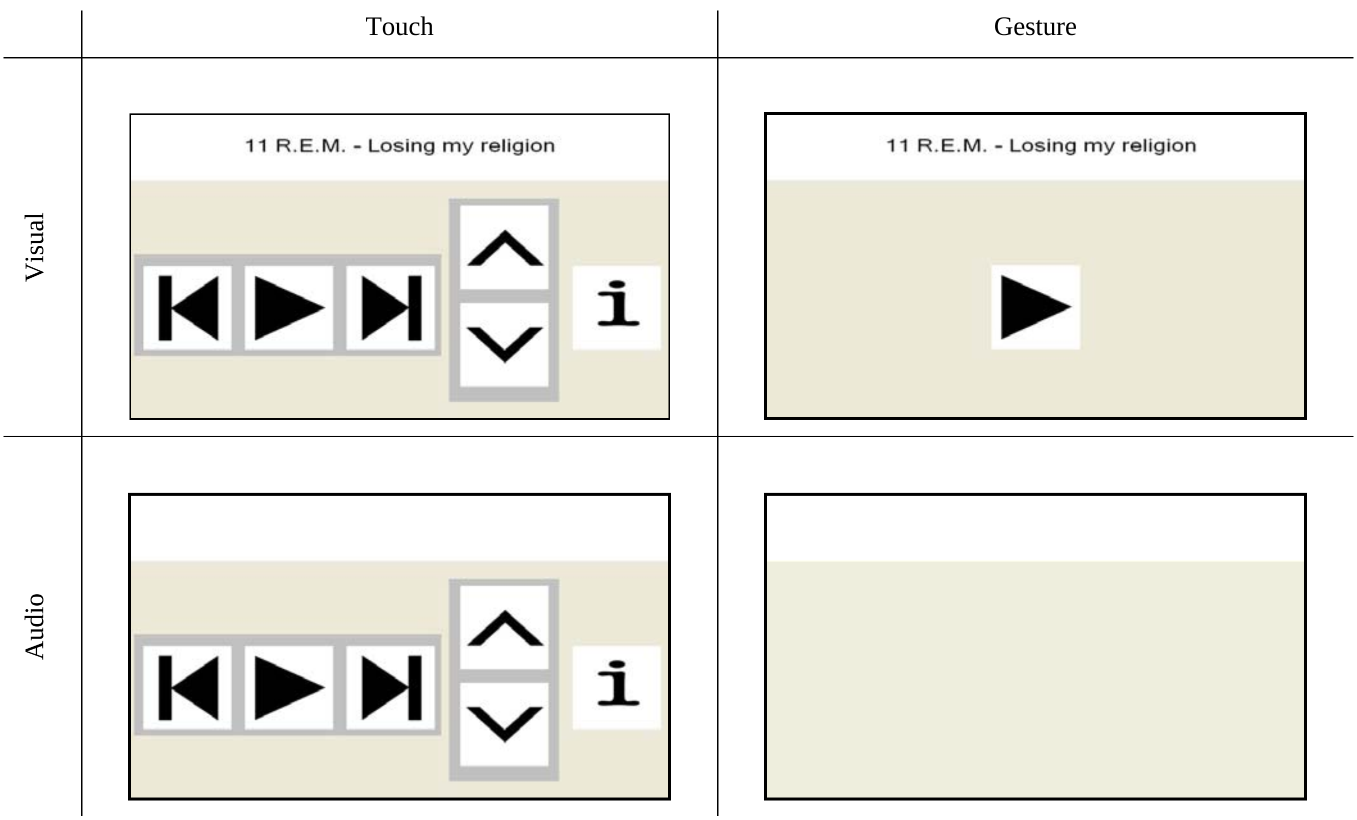


Figure 1. The graphical user interface for the four configurations. The top (white) part of the screen is reserved for output, while the grey area is for input. On the visual (top row) configurations the buttons are, from left to right, "Next song", "Play/pause", "Previous song", "Volume up" and "Volume down", "Song info". In the figure for gesture/visual, the user has just performed the "Play" gesture, causing the system to flash the "Play" icon.

second (as shown on Figure 1).

Audio feedback is implemented using earcons. In the touch/audio and gesture/audio systems, when the user either pushes a button or performs a gesture, the system will provide feedback in the form of a clearly audible “click” sound. Following the same principle that applies to visual feedback, any attempt to adjust the volume either up or down when it is fully up or down, will result in a “dong” sound.

Output regarding the state of the system consists of information regarding the current song; the song’s number in relation to the playlist, the artist and the title of the song. Visual output about the state of the system is provided by text in the output area of the screen and is available at all times. The equivalent audio output is implemented using playback of voice recordings containing the same information. Either pushing the “Song info” button or performing the “Song info” gesture plays these recordings.

EXPERIMENT

The purpose of the experiment was to compare the four different configurations of the system and consequently the different ways of interaction. In the following, we will describe how the experiment was conducted.

Experimental Design

In our experiment, we used a between-subject design with 32 participants, which were divided into four groups of eight corresponding to the four configurations, as shown in Table 1. Each group consisted of four male and four female participants and was assigned to one of the four configurations of our music player.

		Input	
		Touch (N=16)	Gesture (N=16)
Output	Visual (N=16)	N=8	N=8
	Audio (N=16)	N=8	N=8

Table 1. Experimental design

Participants

In our experiment 32 people participated (16 male and 16 female), which all stated that they were in good health. Their age ranged from 21 to 56 years (*M*=28.2, *SD*=9.2). All of the participants carried valid driver’s licenses and had done so for between 0.5 and 29 years (*M*=9.4, *SD*=8.7). Their driving experience was quite varied as it ranges between 100 and 30.000 km/year (*M*=6114.7, *SD*=7989.9). Two of the participants stated that they had previous experience with the computer game used in our simulator.

Setting

Our experiment was conducted in the HCI laboratory at Aalborg University, where we created a medium-fidelity driving simulator (as seen in Figure 2). The simulator consisted of two car seats, a force feedback enabled steering wheel with a brake and accelerator pedal, which controlled a car in the PC game *Test Drive Unlimited*. The game takes place in a realistic setting; on regular roads complete with traffic and road signs. The setup also included two sets of speakers; a set of 4.1 surround sound speakers, which played the sound from the game, and a set of 2.1 stereo speakers for music playback. The game was projected onto the wall in front of the participants (see Figure 2 middle). The speedometer and tachometer of the car was visible to the participants during the experiment as part of the projected image. The participants occupied the driver’s seat while the supervisor sat in the passenger seat during the experiment.

Tasks

The participants were asked to solve 32 tasks in all. The tasks were designed in such a way that half of them primarily focused on system input, and the other half on output. Furthermore, we attempted to create the tasks in such a manner that they did not favor any of the four configurations. The tasks were chosen to reflect realistic interactions one might perform with an in-car music player, e.g. “Change to the next song”. The instructions for each task were kept short and clear in order to interrupt the participants as little as possible. The tasks themselves varied in complexity, from simple ones like “Stop the music” to more demanding ones like “Find and play the song by Coldplay called Viva la Vida”. The tasks were all read aloud by the supervisor in sequence.

Figure 2. Driving setting in the simulator.

Procedure

All sessions followed the same basic procedure. First, the demographical data of the participants was gathered. The participants were then asked to take a seat in the simulator and make sure that the driving position was comfortable. The supervisor then briefed the participants by reading a text aloud, which told them what they were about to do. They were also shown how to operate the music player in the particular configuration they were to use during the experiment. After each instruction was demonstrated, the participants were asked to repeat it, in order to ensure they had understood how to operate the system. The participants were instructed to drive the car between 40 and 60 km/h, except when performing maneuvers like turning and braking, to stay in the right lane, and otherwise observe normal traffic regulations and drive as they would in a real car. The participants were then given a chance to familiarize themselves with the game and the steering wheel and pedals, as they were allowed to try the game prior to the experiment itself.

After the practice run the supervisor reset the game and the actual experiment began. The driving itself was divided into two parts. In the first part the supervisor instructed the participants where to turn, making sure they all followed the same predetermined route. In the second part the participants were instructed to drive anywhere they wished. The length of each part was determined by the tasks, which they were asked to solve while driving. The tasks were divided evenly between the two parts, with 16 tasks to be solved in each. The participants were instructed to start solving the tasks only when they felt ready to do so. The sessions were recorded on four different video cameras for later analysis.

DATA ANALYSIS

In order to assess the performance of the four configurations and their effect on the drivers, we chose to incorporate several parameters. Inspired by [2, 3], we examined the following in the analysis of our data:

- Primary driving task performance
- Secondary driving task performance
- Eye glance behavior

Primary driving task performance was measured as the number of errors in lateral and longitudinal control. A lateral control error was defined as a lane excursion where the participant failed to stay within the two lines denoting the right hand side lane of the road. Longitudinal control errors were defined as failure to maintain a speed within the instructed range of 40-60 km/h. A longitudinal error was noted each time the participants went above or below the speed range. Staying at a wrong speed for a period of time only counted as one error. Identifying these errors was done by reviewing the video, which also captured the in-game speedometer.

Secondary driving task performance we defined as interaction errors and task completion time. Interaction errors were defined as attempts to interact with the system that either had no effect or didn’t have the effect towards completion of the task that the participants expected. In order to identify these errors, one of the cameras recorded an up-close view of the participants’ interaction with the screen. Task completion time was measured from the time the participants started solving the task, defined by either moving their hand from the steering wheel, or moving their head/eye gaze towards the system, until the task was completed.

Eye glances were divided into three categories according to duration:

1. 0.5 seconds and below
2. between 0.5 and 2.0 seconds
3. above 2.0 seconds

The nature of these metrics, in particular the eye glance analysis, meant that it was necessary to view the videos frame by frame. In determining the length of an eye glance for instance, we knew that the each second of video contained 25 frames and a glance of 0.5 seconds or less therefore corresponded to 12.5 frames (in practice, 13 frames).

In order to ensure the highest possible consistency in the interpretations of the data, two of the 32 sessions were analyzed by all the authors cooperatively. This presented us with an opportunity to discuss the various types of incidents in the data, and subsequently to compile a set of directions to be followed in the following individual analyses. Each of the 32 sessions was reviewed by three of the authors of this paper. Each reviewer analyzed the video individually while logging and categorizing instances of all the abovementioned incidents. The resulting three logs were then compared and compiled into one final list containing all the incidents for that session. This was done by way of majority vote; if for instance only one reviewer had recorded a specific incident, which neither of the two other reviewers had recorded, the incident would not make it to the final list, and so forth. The same principle applied to categorization of eye glances. In situations were no majority vote could be secured, the video recording was reviewed again in order to reach the final verdict.

RESULTS

The results of the data analysis are presented in three sections; Primary Driving Task Performance, Secondary Driving Task Performance and Eye Glance Behavior. In each section we first compare the results for the two input methods (N=16), then the two output methods (N=16) and finally all four configurations (N=8). The results were subjected to either two-tailed unpaired Student’s t-tests or one-way repeated-measures ANOVA tests, as well as Tukey’s HSD post hoc tests where applicable. The results are listed in the tables below. Any statistically significant differences at the 95% confidence level are highlighted. An inter-rater reliability test of this analysis (using Fleiss' Kappa) gave $\kappa = 0.70$, corresponding to a substantial agreement according to [11].

Primary Driving Task Performance

The metrics for measuring primary driving task performance included lateral control errors (lane excursions) and longitudinal control errors (deviations from accepted speed range). Across the 32 sessions, we identified a total of 256 lateral control errors and 511 incidents of longitudinal control errors.

	Touch (N=16)	Gesture (N=16)
Lane excursions	7.19 (4.79)	8.81 (7.13)
Speed increases	6.31 (3.07)	6.69 (6.05)
Speed decreases	8.31 (7.42)	10.63 (5.02)
Total speed deviations	14.63 (8.50)	17.31 (7.67)

Table 2. Means (standard deviations) for primary driving task performance for input (N=16). Significant differences at the 95% confidence level are highlighted.

When comparing the primary driving task performance across the two input methods (Table 2), we see no significant difference between any of the metrics, although gesture input generally has a higher number of errors across all the metrics. As seen on Table 3, the results for the output methods, however, do reveal a significant difference in the number of speed increases, with visual having significantly fewer than audio, $t = 2{,}04$, $p < 0.05$. However, there are no significant differences in the number of total speed deviations, although it is worth noting that the number of speed decreases and total speed deviations is higher for audio output than for visual output.

	Visual (N=16)	Audio (N=16)
Lane excursions	8.63 (6.64)	7.38 (5.5)
Speed increases	**4.56 (3.41) -**	**8.44 (5.15) +**
Speed decreases	9.38 (2.04)	9.56 (6.79)
Total speed deviations	13.94 (5.78)	18.00 (9.63)

Table 3. Means (standard deviations) for secondary driving task performance for output (N=16). Significant differences at the 95% confidence level are highlighted.

	Touch/ visual (N=8)	Touch/ audio (N=8)	Gesture/ visual (N=8)	Gesture/ audio (N=8)
Lane excursions	7.63 (4.87)	6.75 (5.01)	9.63 (8.28)	8.00 (6.23)
Speed increases	6.00 (3.66)	6.63 (2.56)	**3.13 (2.59) -**	**10.25 (6.54) +**
Speed decreases	6.38 (5.13)	10.25 (9.11)	12.38 (5.68)	8.88 (3.83)
Total speed deviations	12.38 (8.79)	16.88 (11.67)	15.50 (8.27)	19.13 (10.37)

Table 4. Means (standard deviations) for primary driving task performance for the four configurations (N=8). Significant differences at the 95% confidence level are highlighted.

If we look at the results of the primary driving task performance and compare the four configurations (Table 4), we see a significant difference in the number of speed increases, $F(3, 28) = 3.95$, $p < 0.05$. A Tukey's HSD post hoc test revealed that there are significantly fewer speed increases in the gesture/visual configuration vs. the gesture/audio configuration ($p < 0.05$). The remaining measurements of primary driving task performance show no significant differences. But the results do show that the two audio configurations have the highest number of total speed deviations.

Secondary Driving Task Performance

For secondary driving task performance we measured the total task completion time and identified a total of 1018 interaction errors. Comparing just input methods (as seen on Table 5) the results show only marginal differences in the number of interaction errors and the task completion time, although gesture does show a higher task completion time than touch, $t = 2.04$, $p < 0.19$.

	Touch (N=16)	Gesture (N=16)
Interaction errors	29.38 (19.69)	34.25 (29.99)
Task completion time	271.00 (62.13)	308.81 (95.20)

Table 5. Means (standard deviations) for secondary driving task performance for input (N=16). Significant differences at the 95% confidence level are highlighted.

Whereas the input methods revealed no significant differences in secondary task performance, the results for output showed 77% more interaction errors for visual output compared to audio output (Table 6). A t-test shows that this is a significant difference, $t = 2.04$, $p < 0.05$. The task completion times, however, were significantly longer for audio output, $t = 2.04$, $p < 0.05$.

	Visual (N=16)	Audio (N=16)
Interaction errors	**40.69 (29.13) +**	**22.94 (16.82) -**
Task completion time	**256.94 (67.66) -**	**322.88 (82.40) +**

Table 6. Means (standard deviations) for secondary driving task performance for output (N=16). Significant differences at the 95% confidence level are highlighted.

As seen on Table 7, secondary driving task performance results reveal no significant differences in the number of interaction errors distributed among the four configurations, even though the average number of interaction errors for the touch/audio configuration is less than half that of the touch/visual and gesture/visual configurations, $F(3, 28) = 1.87$, $p < 0.16$. However, a significant difference does exist between the task completion times, $F(3, 28) = 3.06$, $p < 0.05$. A post hoc test showed that there is a significant difference between task completion times for the touch/visual and gesture/audio configurations ($p < 0.05$).

	Touch/ visual (N=8)	Touch/ audio (N=8)	Gesture/ visual (N=8)	Gesture/ audio (N=8)
Interaction errors	42.38 (19.72)	16.38 (7.46)	39.00 (37.72)	29.50 (21.27)
Task completion time	**249.88 (24.28) -**	292.13 (81.62)	264.00 (95.42)	**353.63 (75.66) +**

Table 7. Means (standard deviations) for secondary driving task performance for the four configurations (N=8). Significant differences at the 95% confidence level are highlighted.

Eye Glance Behavior

We identified a total of 2371 glances divided into 560 glances below 0.5 seconds, 1729 between 0.5 and 2.0 seconds and 52 above 2.0 seconds.

	Touch (N=16)	Gesture (N=16)
< 0.5 s.	16.44 (13.85)	20.44 (12.09)
0.5 – 2.0 s.	**71.88 (19.35) +**	**36.19 (36.66) -**
> 2.0 s.	0.88 (1.36)	2.38 (3.74)
Total glances	**89.19 (19.10) +**	**59.00 (46.83) -**

Table 8. Means (standard deviations) for eye glance behavior for input (N=16). Significant differences at the 95% confidence level are highlighted.

Of the total glances, around 60% occurred with touch input, which amounts to a significant difference compared to gesture input, $t = 2.04$, $p < 0.05$ (see Table 8). Looking at the individual eye glance categories, the results show a strong significant difference in the number of glances between 0.5 and 2.0 seconds, with gesture input having substantially fewer, $t = 2.04$, $p < 0.01$. But in the two remaining categories touch has the fewest, although the difference is only marginal.

	Visual (N=16)	Audio (N=16)
< 0.5 s.	15.94 (11.85)	20.94 (13.88)
0.5 – 2.0 s.	**76.06 (24.34) +**	**32.00 (13.88) -**
> 2.0 s.	**3.19 (3.43) +**	**0.06 (0.25) -**
Total glances	**95.19 (30.14) +**	**53.00 (34.43) -**

Table 9. Means (standard deviations) for eye glance behavior for output (N=16). Significant differences at the 95% confidence level are highlighted.

The number of glances for visual output account for 1523 (64%) of the total number of glances across output types, which amounts to an extreme significant difference, $t = 2.04$, $p < 0.001$ (Table 9). There is also an extreme significant difference in the number of glances between 0.5 seconds and 2.0 seconds with audio being significantly lower than visual, $t = 2.04$, $p < 0.001$. Finally, there also exists a strong significant difference in the number of glances above 2.0 seconds, with visual again having more (with 51 glances vs. just 1 glance), $t = 2.04$, $p < 0.01$. On the other hand, audio output has more glances below 0.5 seconds than visual output, albeit only marginally.

	Touch/ visual (N=8)	Touch/ audio (N=8)	Gesture/ visual (N=8)	Gesture/ audio (N=8)
< 0.5	8.88 (4.19)	24.00 (16.20)	23.00 (13.02)	17.88 (11.34)
0.5–2.0	**86.50 (12.40) +(+)**	**57.25 (12.62) +(-)**	**65.63 (29.44) +**	**6.75 (5.70) -**
> 2.0	**1.75 (1.49) -**	**0.00 (0.00) -**	**4.63 (4.27) +**	**0.13 (0.35) -**
Total glances	**97.13 (18.08) +**	**81.25 (28.83) +**	**93.25 (46.73) +**	**24.75 (17.40) -**

Table 10. Means (standard deviations) for eye glance behavior for the four configurations (N=8). Significant differences at the 95% confidence level are highlighted. The (+) and (-) in the row for glances between 0.5 and 2.0 seconds indicates that a significant difference exists between these two values as well.

Across the four configurations, the touch/visual configuration accounts for around 32% of the total amount of glances, touch/audio for 27%, gesture/visual for 31% and gesture/audio for just 8% (see Table 10). A one-way repeated-measures ANOVA showed this difference to be extreme significant, $F(3, 28) = 13.59$, $p < 0.001$. Looking at these percentages, it is perhaps not surprising that the post hoc test revealed that the number of glances for the gesture/audio configuration was significantly lower than for any of the other configurations, $p < 0.01$.

Although touch/visual has substantially fewer glances below 0.5 seconds compared to e.g. touch/audio, this does not represent a significant difference, but a one-way repeated-measures ANOVA indicates that it is approaching significance, $F(3, 28) = 2.65$, $p < 0.07$. For glances between 0.5 and 2.0 seconds, however, an extreme significant difference exists, $F(3, 28) = 30.22$, $p < 0.001$. The results of the post hoc test showed that gesture/audio has significantly fewer glances in this category than any of the other configurations, $p < 0.01$. This is perhaps not surprising, as gesture/audio accounts for just 8% of all the glances in this category. The post hoc test also revealed a significant difference between the number of glances between 0.5 and 2.0 seconds for touch/visual and touch/audio, $p < 0.05$. In the last category, glances above 2.0 seconds, our results show an extreme significant difference in the number of glances, $F(3, 28) = 7.20$, $p < 0.001$. According to the post hoc test, gesture/visual has significantly more glances in this category than any of the other configurations, with $p < 0.01$ compared to touch/audio (0 glances) and gesture/audio (1 glance), and $p < 0.05$ compared to touch/visual.

DISCUSSION

The overall problem we set out to research was how to design in-vehicle systems that require as little visual attention from the driver as possible in order to avoid a decrease in driving performance, as current conventional techniques tend to do [12]. In the following we discuss and reflect on our results.

Separating Input From Output

Bach et al. state that they are unsure what effect it has that their interaction techniques differ both in input and output, and further studies are needed to address this issue [3]. This is what we have done in our work, where the results show that a distinction between input and output is indeed an important one to make. Our results show that there is a significant difference in the number of eye glances when comparing across output technique. This seems to imply that when conducting experiments with in-vehicle systems it is important to isolate and focus on both the input and output methods of the system.

Input

Our initial assumption was that touch input would require more eye glances than gesture input, since the participants presumably needed to visually obtain the position of the buttons before commencing interaction. This is also supported by our findings where we find a strong significant difference in glances between 0.5 and 2.0 seconds, and a significant difference in the total number of glances, which is in line with [1, 14]. In fact, the touch technique accounted for 51% more glances than the gesture technique, with respect to the total amount of eye glances. This number is even greater when viewing the glances between 0.5 and 2.0 seconds isolated, where touch input accounts for almost twice as many glances (98%) as gesture input. This is in line with Alpern & Minardo's findings which show that gesture interfaces, although not attention free, help drivers solve their task while allowing them to keep their eyes on the road [1].

The difference in eye glance behavior can perhaps in part be explained by the fundamental design of the systems. When interaction fails with a touch button based interface, or if several interactions have to be performed in quick succession, users might have a tendency to use more glances in order to ensure/reassure that the correct button is being pressed. Similarly one might suspect that with gesture input, the user only has to visually confirm the position of the screen before being able to issue one or more commands without looking, as opposed to finding the correct button on the screen. This could be part of the explanation for the difference in the number of glances.

Before conducting the experiment we also had the assumption that gesture input would have relatively more glances below 0.5 seconds compared to touch, the rationale being that the aforementioned visual confirmation of the position of the screen should not take long. However, none of our findings corroborate this assumption. In terms of the number of interaction errors, the two input techniques show no significant difference to each other. In line with the findings of [3] our results also show touch as the fastest of the two input forms, although not significantly.

Output

In the measurements of primary driving task performance there is some difference between audio and visual output. Only in the number of speed increases is this difference significant, in favor of visual output. However the total number of speed deviations is not significantly different, so what these results indicate, if anything, is unclear since the number of speed decreases is almost identical, and the total amount of speed deviations imply no significant difference.

When comparing task completion time for the two output techniques of our system, there is a significant difference between the two, with visual output being faster. We believe this is due to the nature of audio output. When solving tasks requiring audio output, the user first has to hear the audio message, which can be of arbitrary length, and then process the information they are presented with before being able to solve the task. With visual output the user only has to read the information before being able to answer, which presumably takes less time. Or perhaps the user has already seen the information while performing another task, which further decreases the time required to solve certain tasks with the visual output technique.

Another interesting finding is that there is a strong to extreme significant difference in the number of eye glances between visual and audio. We believe that there are several reasons for this difference: first and foremost, the nature of audio output gives less incentive for looking at the screen, since it does not contain any visual information, nor does it give any kind of visual feedback. Obviously, users of touch/audio have more motivation for looking at the screen, compared to gesture/audio, since they still need to locate the buttons on the screen. However, for both configurations it applies that when issuing commands to the system, nothing is gained from looking at the screen, since no feedback is presented there. This is clearly different from the configurations with visual feedback, where there is no way of obtaining feedback other than looking at the screen, which would explain the difference in the number of glances. As a result, audio output leads to a higher task completion time, but fewer eye glances compared to visual output. And, aside from a significant difference in the number of increases in speed, there is no overall significant difference in the primary driving task performance.

In terms of road safety it can be argued that the increase in task completion time is a favorable tradeoff if it comes with fewer eye glances, which in turn leads to more attention on the road. Our results do not however, show a link between the number of glances and primary driving task performance, which is similar to the findings in Bach et al. [3]. However, other studies state that a relationship between eye glance behavior and driving performance does exist [8, 15]. In line with Gellaty [7], it is not difficult to imagine that more visual attention on the road is preferable, since the driver's primary method of assessing danger signs in traffic arguably is through the eyes. However, drivers keeping their eyes on the road is perhaps not enough, as indicated in a study on the effects of hands-free mobile phone conversations on driving performance [18]. Here, Strayer & Drews state that even if drivers conducting a hands-free mobile phone conversation direct their gaze at the road, they often fail to notice objects in the driving environment, since their attention is occupied with conducting the mobile phone conversation. However, the results in [18] relate to mobile phone conversations, which they claim might differ qualitatively from other auditory tasks.

Although our results show that systems with audio output lead to distinctly fewer eye glances than systems with visual output, the results also seem to indicate that audio output comes at a price – namely an apparent drop in primary driving task performance. For instance, the number of speed increases and total number of speed deviations are marginally higher for audio output than for visual output. This could indicate that listening to audio output while driving causes an increase in the cognitive load of the driver, thereby drawing mental resources away from the task of driving. This would be in line with a recent study in the field of brain research, which showed that driving while comprehending language, i.e. listening to voice messages from a hands-free mobile phone, results in a deterioration of driving performance [10]. Cognitive workload is also discussed in Bach et al. [3] in relation to their gesture/audio system, but their setup does not allow them to see an explicit connection to the output method, which leads them to attribute it to memory load, e.g. the driver having to remember the gestures and the state of system. Another possible contributor to increased, or perhaps misaligned cognitive load, is the amount of the time the driver spends on solving a specific secondary driving task. As previously mentioned, our results show that the subjects receiving audio output spent significantly more time completing the tasks. Hence, while audio output might result in fewer glances, the driver is occupied with the task for a longer time, if only mentally.

Limitations

Some of our participants found the limited level of realism in the simulator problematic. They pointed to the absence of tire noise, lack of opportunity to orientate themselves through the side and rear windows and sensation of movement, as some of the factors they felt affected the realism and their driving performance. This was in part because these factors provide drivers with a sensation of movement, which helps them estimate speed, without having to look at the road ahead. This could imply that particularly longitudinal control performance suffers from simulated driving, which is also commented on by Bach et al. [3].

Our choice of case system represents a possible source of inaccuracy. The nature of the music player means that it will always give a form of audio feedback, regardless of which output methods we choose. For instance, pushing the "Play" button will cause music to be played; turning up the volume will cause the music to become louder, etc. This means that participants given visual output would not necessarily need to look at the screen to receive feedback.

CONCLUSION

As more and more systems are making their way into cars, and existing in-vehicle systems are becoming more advanced, research is needed to further illustrate how to design interaction techniques that consider the unique characteristics and requirements of the vehicle domain. There is a tendency in previous research to focus mainly on the input aspect of in-vehicle interaction. The aim of this paper was to address this issue by putting equal emphasis on input and output, in order to investigate their separate effects on driving performance and eye glance behavior. This was done by evaluating four different combinations of input and output techniques in a driving simulator.

The results of our evaluation show that when addressing in-vehicle systems design, separating input and output modes does make a difference. Using gesture input resulted in significantly fewer eye glances compared to touch input, but also inferior primary driving task performance and longer task completion times. Audio output caused the participants to make more longitudinal control errors compared to visual output, and had a significantly longer task completion time. Visual output, on the other hand, accounted for significantly more interaction errors and a drastically higher number of eye glances. Looking at the individual input/output configurations, our results show that gesture/audio by far has the fewest number of glances, but also a longer task completion time and more longitudinal control errors than any other configuration.

Our results did not indicate that fewer eye glances necessarily entails better primary driving task performance. On the contrary, audio output, which has the fewest eye glances by far, seems to cause worse primary driving performance as well as longer total task completion times compared to visual output. This could imply that audio output has an effect on the mental load of the driver, distracting their cognitive attention away from the primary task of driving the car. Further research might shed more light on this phenomenon.

REFERENCES

1. Alpern, M., & Minardo, K. (2003). Developing a Car Gesture Interface for Use as a Secondary Task. *CHI 2003: New Horizons* , pp. 932-933.
2. Bach, K. M., Jæger, M., Skov, M. B., & Thomassen, N. G. (2007). A Classification of In-Vehicle Systems Research: Understanding, Measuring and Evaluating Attention. *HCI Lab Technical Report no. 2007/2* .
3. Bach, K. M., Jæger, M., Skov, M. B., & Thomassen, N. G. (2008). You Can Touch, but You Can't Look: Interaction with In-vehicle Systems. *Proceeding of CHI '08* , pp. 1139-1140.
4. Brooks, C., & Rakotonirainy, A. (2007). In-vehicle Technologies, Advanced Driver Assistance Systems and Driver Distraction: Research Challenges. In I. J. Faulks, M. Regan, M. Stevenson, J. Brown, A. Porter, & J. D. Irwin (Eds.), *Distracted Driving* (pp. 471-486). Sydney, NSW: Australasian College of Road Safety.
5. Chewar, C. M., McCrickard, D. S., Ndiwalana, A., North, C., Pryor, J., & Tessendorf, D. (2002). Secondary task display attributes: optimizing visualizations for cognitive task suitability and interference avoidance. *Proceeding of the symposium on Data Visualisation (VisSum '02), Eurographics Association* , pp. 165-171.
6. Colle, H. A., & Hiszem, K. J. (2004). *Standing at a kiosk: Effects of key size and spacing on touch screen numeric keypad performance and user preference.* Wright State University.
7. Geiger, M., Zobl, M., Bengler, K., & Lang, M. (2001). Intermodal differences in distraction effects while controlling automotive user interfaces. *Proceedings Vol. 1: Usability Evaluation and Interface Design, HCI 2001* , pp. 236-267.
8. Gellatly, A. (1997). *The Use of Speech Recognition Technology in Automotive Applications.* Faculty of Virginia Polytechnic Institute and State.
9. Green, P. (2004). *Driver Distraction, Telematics Design, and Workload Managers: Safety Issues and Solutions.* University of Michigan Transportation Research Institute. SAE International.
10. Just, M. A., Keller, T. A., & Cynkar, J. (2008). *A decrease in brain activation associated with driving when listening to someone speak.* Carnegie Mellon University.
11. Landis, J. R. and Koch, G. G. (1977) "The measurement of observer agreement for categorical data" in *Biometrics*. Vol. 33, pp. 159--174
12. Lansdown, T. C., Brooks-Carter, N., & Kersloot, T. (2004). Distraction from Multiple In-Vehicle Secondary Tasks: Vehicle Performance and Mental Workload Implications. *Ergonomics , 47* (1), pp. 91-104.
13. *MedicineNet*. (n.d.). Retrieved from http://www.medicinenet.com/hemin_for_injection/glossary.htm
14. Pirhonen, A., Brewster, S., & Holguin, C. (2002). Gestural and Audio Metaphors as a Means of Control for Mobile Devices. *CHI Letters , 4* (1), pp. 291-298.
15. Rockwell, T. H. (1988). Spare visual capacity in driving - revisited: New empirical results for an old idea. In M. H. Freeman, P. Smith, A. G. Gale, S. P. Taylor, & C. M. Haslegrave (Eds.), *Vision in Vehicle II* (pp. 317-332). Elsevier Science.
16. Sears, A. (1991). *Improving Touchscreen Keyboards: Design issues and comparison with other devices.* University of Maryland.
17. Sears, A., Revis, D., Swatski, J., Crittenden, R., & Schneiderman, B. (1992). *Investigating Touchscreen Typing: The effect of keyboard size on typing speed.* University of Maryland.
18. Strayer, D. L., & Drews, F. A. (2007). *Cell-Phone Induced Driver Distraction.* University of Utah.

Appendix C: Summary

In our masters thesis we wish to study the research area of in-vehicle systems regarding driving performance and attention in cars. These in-vehicle systems are becoming more and more common in cars, due to the fact that the necessary technology has reached an acceptable price level and people are used to adopting new information systems into their lives. There are two predominant types of in-vehicle systems; systems that seek to improve safety by assisting the driver in the primary task of driving, and systems that are only used as a secondary task.

Due to the above-mentioned tendency, we have divided our master´s thesis into two main fields of focus on in-vehicle systems. The first field concerns the effects of systems that intend to support the primary driving task, the second concerns how different kinds of interaction with secondary task systems affect driving and driver attention. In order to address these fields of focus, we conducted two experiments, documented in two research papers.

The purpose of the first research paper was to investigate what effects an advance warnings system had on driver behavior and attention. We constructed an advance warning system designed to issue warnings 75m in advance of various road and traffic conditions (curves, speed limits, traffic jams, slippery roads). We then conducted a controlled driving experiment on a closed circuit using a with-subjects design with 12 participants. The participants followed a route where they encountered the various incidents, both with and without advance warnings.

Afterwards we analyzed the test data that consisted of the speed of the vehicle and the video material used to identify eye glances. Our primary concern was the reduction in speed from the warning was issued to the incident was reached. The results show that the advance warning system overall had a limited effect on the speed of the drivers, although the effect was larger for some incident types. The warnings caused the participants to direct more eye glances off the road and at the system.

Inspired by [4], the second research paper concerned the effects of different combinations of input and output techniques for an in-vehicle system used as a secondary driving task. We designed an in-vehicle music player that had four different configurations based on two input methods and two output methods. They were touch/visual, touch/audio, gesture/visual, and gesture/audio. The effects of these configurations on primary driving task performance, secondary driving task performance and eye glance behavior was then evaluated in an experiment using a driving simulator. We used a between-subjects design with 32 participants who were instructed to solve short tasks while driving a predetermined route. The experiment sessions were recorded on video for later analysis.

During the analysis of the experiment data, we measured the primary driving task performance (errors in driving the car), the secondary driving task performance (errors in system interaction and task completion time) and eye glance behavior (number of glances below 0.5 seconds, between 0.5 and 2.0 seconds and above 2.0 seconds). The results of the experiment revealed that gesture input resulted in significantly fewer eye glances, but also poorer primary driving task performance as well as longer task completion times compared to touch input. In relation to output, audio output caused a considerably lower number of eye glances and fewer interaction errors than visual output, while visual output had the most interaction errors but shorter task completions times. Concerning the specific combinations of input and output, the results showed that gesture/audio had a significantly lower number of total glances than any of the other combinations, while at the same time having errors in relation to driving the car and longer total task completion times. Based on the results of the experiment, we argue that studying input and output on equal terms is important when working with in-vehicle systems.