BEHAVIOUR & INFORMATION TECHNOLOGY, 2002, VOL. 21, NO. 4, 235–248

Taylor & Francis
Taylor & Francis Group

The effects of adverse condition warning system characteristics on driver performance: an investigation of alarm signal type and threshold level

NITIN GUPTA*, ANN M. BISANTZ* and TARUNRAJ SINGH†

*Department of Industrial Engineering, University at Buffalo, State University of New York, Amherst, NY 14260, USA

†Department of Mechanical and Aerospace Engineering, University at Buffalo, State University of New York, Amherst, NY 14260, USA

Abstract. This study addresses the issues concerning the design of adverse condition warning systems (ACWS). ACWS are designed to sense adverse road and weather conditions as well as system states that can negatively impact driving performance leading to skids or accidents, and alert drivers to these conditions. In this case, an ACWS was designed to sense when a car was likely to skid. A virtual-driving environment was used to test two levels of alarm sensitivity (low and high) and two types of auditory alarm signal (Binary ON/OFF and Graded) along with a no-alarm control group. Dependent measures reflected driver performance, response to the alarm signal and trust in the alerting system. Results indicated that participants had fewer skids in the low sensitivity and graded alarm signal condition compared to some other alerting system configurations. Participants in the graded alarm signal condition also had a greater degree of lateral control over the vehicle. Additionally, trust was found to be lower for the high vs. low sensitivity alarm condition, indicating a reduction in trust when the alerting system activated more often, perhaps because participants did not feel the system was accurately reflecting a dangerous condition. This simulator-based research emphasizes the fact that while ACWS may provide an advantage in terms of vehicle control, characteristics of both the alerting signal and system configuration should be considered.

1. Introduction

Advanced Traveller Information Systems (ATIS) are a type of Intelligent Transportation System that provide real-time, in-vehicle information to drivers regarding navigation and route guidance, motorist services, roadway signing and hazard warnings. In-Vehicle Safety and Advisory Warning Systems (IVSAWS) are a component of ATIS that provide warnings to the drivers regarding impending obstacles and inclement weather conditions (Mast 1998). Adverse weather conditions such as icy and foggy conditions are often the cause of accidents, skids and even fatal crashes. The development of hazard warning signals such as adverse condition warning systems, a type of IVSAWS, can help reduce personal injury and monetary loss.

Researchers in human factors have investigated variables such as display format and information reliability, which may affect driver performance for some components of the Intelligent Transportation Systems (ITS). However, these studies have been primarily concerned with collision avoidance systems or in-vehicle navigation and traveller information systems (Horowitz and Dingus 1992, Dingus *et al.* 1997, 1998, Lee and Kantowitz, 1998, Wheeler *et al.* 1998). The following study focuses on another, equally important component of ITS: adverse condition warning systems (ACWS).

Adverse condition warning systems for slippery road conditions attempt to provide a driver with an alerting warning whenever there is a possibility of a skid or a rollover. This possibility is gauged by deducing the coefficient of friction between the road and the vehicle tires. In-vehicle warning systems can utilize yaw sensors attached to the automobile system, which assist in the

Behaviour & Information Technology
ISSN 0144-929X print/ISSN 1362-3001 online © 2002 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0144929021000013473

evaluation of the driving situation of the vehicle. Sobottka and Singh (1996) have developed a simulation of emergency braking action on a road surface using a time varying adhesion coefficient. Yaw sensors in conjunction with the slip sensors can define when the automobile may spin out of control. Yaw sensors can be used to measure roll, in order to gauge the potential for the automobile to roll over. In-vehicle displays can then be used to convey this information to the drivers.

Dingus and Hulse (1993) have emphasized the need for research addressing issues of timing, modality, false alarms and potential operator reaction to in-vehicle warning or alerting systems. Pritchett (2001) has defined an alerting system in the context of aircraft cockpits as an 'attention-director'. She claims that alerting alarms should act as a trigger to the pilot to start the diagnosis and, if necessary, the resolution processes. To effectively incorporate adverse condition warning systems in vehicles, we must be able to determine the most effective configurations for triggering drivers to take actions. This requires investigation of possible system configurations and their effect on human performance and behaviour. These configurations involve comparisons of different alerting modalities, levels of information reliability, presentation formats and, finally, human acceptance of such systems.

1.1. *Alarm modality*

Although there are no standard guidelines for selection and design of appropriate alarm modalities for ITS, specific modalities may prove to be more effective depending upon the application under design. At least three types of modalities (visual, audio and tactile) and their combinations are possible for any alarm design.

Displays serve to communicate information and information is needed for decision making and performing the proper control action in a timely manner (Kantowitz and Sorkin, 1983). Stokes *et al.* (1990), in their research on visual and auditory displays in aircrafts, summarized that pilots prefer visual over auditory warnings when there is enough time to react. However, if the visual channel is overloaded, there are obvious advantages in allocating some task to other sensory channels. Driver workload may be impacted by alarm modality. Much of the information used by drivers is presented visually; therefore, it may be preferable to use alarms that access a different modality (Srinivasan and Jovanis 1997). For example, Srinivasan and Jovanis (1997) found that performance increased, in terms of decreased reaction times, for audio over visual displays of navigation information.

Stokes *et al.* (1990) suggest that replacing traditional visual indicators with auditory signals, such as bells, beepers and electronic tones, reduce the need for visual instrument scanning, thereby allowing the user to devote more attention to other visual tasks. As in aircraft cockpits, Bois (1982) says that automobile dashboards are also growing in complexity as the number of vehicle functions increases. In this case, auditory signals may be the most effective means of unburdening the visual channel (Doll *et al.* 1984). Auditory displays also have the advantage that they do not require the user, once alerted, to adjust his or her gaze in order to receive the message. Thus, they would be valuable in situations where the user must maintain his or her eye fixation at a particular point in order to perform most effectively. In slippery road conditions where a vehicle might be prone to a skid, auditory alarms would help the user keep his or her fixation on the road in what is described as a 'eyes busy' task (Scott and Kee 1987). Auditory displays do possess certain limitations that also need to be considered. While the use of auditory information may help to alleviate the visual clutter, auditory displays—by their very nature—can be intrusive and distracting (Stokes *et al.* 1990). Drivers may get startled, annoyed or both by auditory warnings especially for non-emergency situations such as 'low fuel' or 'low windshield fluid'. However, in contrast with aircraft requirements, audio alerting signals might be best suited for automobiles since an automobile driver needs virtually constant eye contact with the road in order to maintain proper lane position (Zwahlen 1985).

In automobiles, apart from audio alarms, visual alarms such as flashing lights or messages on the dashboard are a possibility, as are tactile interfaces such as vibrations in the steering wheel. Auditory displays may be more immediately salient than visual displays; however, they may also be more disrupting or annoying to the driver, increasing the likelihood that they will be turned off (Dingus *et al.* 1998). Mollenhauer *et al.* (1994) found that, in a system that presented road sign information to drivers either visually or aurally, drivers performed worse with auditory displays, and also felt the auditory displays were more distracting. However, instructive alarms ('brake', 'slow down'), scaled to the urgency of the situation (Dingus *et al.* 1998) could be provided with an auditory interface. Dingus *et al.* (1997) also studied alarm modality for a collision warning system and found that a combined visual and auditory system provided some advantage in terms of increasing following distance over a solely visual or auditory display, under certain traffic conditions. Compared to an auditory alarm, a tactile interface may be similarly salient with less disruption, but also may be less informative since instructive messages cannot be given.

Tactile interfaces may also be preferable to drivers, as the alarms are less obvious to others in the car, and therefore less embarrassing to the driver (Dingus *et al.* 1998). In both cases (auditory and tactile), the alarm could be used to alert the driver to a dashboard message regarding adverse conditions and driving recommendations. Allowing the driver to choose the modality is another option (Wheeler *et al.* 1998).

1.2. *System reliability and sensitivity*

Important design considerations with respect to warning systems is the degree to which they perform reliably and the degree to which drivers believe that an alarm from the system indicates an impending dangerous situation. An alerting system can lack reliability if it fails to indicate a legitimate hazard, or if it activates without a true need. In the latter case, when the system produces a high number of false alarms, it is likely that the system will be deemed useless and annoying by the driver and lose its effectiveness (Horowitz and Dingus 1992). Pritchett (2001) has reported instances of aircraft pilots viewing alerting systems as *nuisances* due to their high false alarm rates. Horowitz *et al.* (1992) also state that in case of frequent warnings, the driver might ignore the warning since it would be perceived as a false alarm ('crying wolf') or useless/redundant information.

This variability in information reliability leads to the driver having his or her own perception of the usefulness of the system (Lee and Kantowitz 1998). For the adverse condition warning system, drivers' perceptions of system utility may be influenced by characteristics of the alerting system. Because the sensing system will identify conditions which are only tentatively linked to accident causation, a sensor based alert will not imply an impending accident or loss of control 100% of the time: false alarms and missed events may occur. The potentially negative impact of false alarms on the utilization of warning systems has been discussed for ITS systems such as collision avoidance systems (Horowitz and Dingus, 1992, Dingus *et al.* 1997, 1998, Lee and Kantowitz, 1998; Wheeler *et al.* 1998). If the criterion for the alarm activation is set at too low a threshold (e.g. when the probability of an accident is low), the alarm may conflict with other cues available to the drivers (e.g. visible road conditions, other driver's behaviour) and the driver may learn to ignore, discount and possibly disable the alerting system. Conversely, if the threshold for generating an alert is too high, potentially hazardous conditions may be communicated to the driver, and an activated alert may not provide any information that is not already available to the driver from other cues. Thus, the most effective alerting level should lie somewhere between these extremes.

Horowitz *et al.* (1992) have proposed a design suggestion for collision avoidance systems that would help overcome the previously mentioned issues. They suggest a graded sequence of warnings from mild to severe, as a function of the time to collision, T. The longer T is, the milder the warning would be. If T is shorter than a critical time c, no warning would likely help the driver. Similarly, a concept of Likelihood Alarm Display (Sorkin *et al.* 1988) has been suggested for situations when false alarms are likely. In this case, the alarm signal is presented in a graded format based on the likelihood of the event. These recommendations are consistent with the design for aircraft caution and waning signals made by Thompson (1981) and Patterson (1982). Patterson (1982) suggests using a gradual signal onset type sounds and the careful adjustment of warning signal power spectra. This provides a deviation from the traditional binary (e.g. either ON or OFF) warning systems. There are certain factors such as traffic condition (e.g. heavy or light), driving type (e.g. highway or city) or degree of adverse weather (heavy or light snow/fog), which may make accidents more likely. In the context of collision-avoidance systems, Wheeler *et al.* (1998) suggests allowing drivers to adjust the sensitivity or threshold of the system to reflect the prevailing conditions, or the drivers' own experience.

1.3. *User acceptance and trust in the alerting system*

The extent to which the alerting system reliably indicates dangerous situations along with the degree of danger, may impact the degree to which drivers trust in and utilize the system. There is a history of research on human trust from both sociological and human-machine perspectives that may apply to systems such as alerting systems. For instance, researchers have suggested that trust can affect how much people accept and rely on increasingly automated systems (Lee and Moray 1992, Parasuraman *et al.* 1993, Muir and Moray 1996, Sheridan 1988). Generally, research from both social science and engineering perspectives agree that trust is a multi-dimensional, dynamic concept capturing many different notions such as predictability, dependability, faith (Rempel *et al.* 1985), competence, responsibility, reliability (Muir and Moray 1996), robustness, familiarity, understandability, explication of intention, usefulness and dependence (Sheridan 1988).

Empirical results have shown that people's strategies with respect to the utilization of an automated system may be affected by their trust in that system. For example, Muir and Moray (1996) and Lee and Moray

(1994) studied issues of human trust in simulated, semi-automated pasteurization plants. These studies showed, among other results, that operators' decisions to utilize either automated or manual control depended on their trust in the automation and their self-confidence in their own abilities to control the system. Additionally, results showed that trust depended on current and prior levels of system performance, the presence of faults and prior levels of trust. For example, trust declined, but then began to recover after faults were introduced (Lee and Moray 1992). Lerch and Prietula (1989) found a similar pattern in participants' confidence in a system for giving financial management advice: confidence declined after poor advice was given, then recovered, but not to the initial level of confidence.

In a different environment, research has investigated driver trust and self-confidence in a simulated, in-vehicle decision aid which provided drivers with traffic information of varying degrees of reliability (Kantowitz *et al.* 1997). Drivers could request information about traffic congestion for different segments of their potential routes in order to make route-planning decisions. Drivers expressed less trust in the aiding system for conditions when the information was less reliable, particularly under familiar settings. Similarly, Bonsall and Parry (1991) used an artificial traffic network and investigated the effect of quality of advice on user acceptance. They concluded that for advice providing an optimal route to their destination, user acceptance declined with decreasing quality of advice. However, this relationship was based on the unfamiliarity of the traffic network. With increasing familiarity, they found that the users were less likely to accept advice from the system. Allen *et al.* (1991) provided conflicting results by investigating the effect of familiarity on a real traffic network. They found that familiarity did not affect route choice behaviour.

Thus, information reliability, experience, driver trust and driver self-confidence have been investigated in the context of an automated traffic information system (Kantowitz *et al.* 1997), but not for an in-vehicle system alerting system specifically.

1.4. *Study goals*

In summary, there are a variety of questions regarding the implementation of in-vehicle alerting system which warn drivers about imminent dangers such as collisions or skids. Because human operators must make control decisions based on the output from these sensing and alerting systems, it is necessary to consider the impact of different system design considerations, such as alarm modality, system reliability and sensitivity, and alarm characteristics, on driver behaviour. Therefore, a study was conducted in a virtual driving environment to answer the following questions regarding an adverse condition warning system that warned drivers of impending skids:

- Can the presence of Adverse Condition Warning Systems help improve driver performance and safety?
- Do drivers perform better under adverse snow conditions when the alerting signal is set at a lower sensitivity (nearer to skid) than at a higher sensitivity (farther away from the skid)?
- Does a graded alarm signal based on a scale of urgency induce better driver performance and response as compared to a binary ON/OFF alarm signal?
- What would be the impact of this system and its various configurations on user trust and acceptance?

2. Method

2.1. *Overview*

This study investigated the effect of an immediate response warning system on driver performance and behaviour, and the factors that influence it. The system tested provided an alerting signal regarding an imminent skid or rollover. Based on the literature previously reviewed, the variables selected for the study were types of alerting formats and levels of alarm sensitivities. The dependent measures assessed driver performance and behaviour, as well as drivers perceptions of trust in the alerting system.

The task environment was a simulated driving task under icy road conditions with randomly situated hazards. These hazards consisted of patches of icy road, manipulated by changing the coefficient of friction between the icy road surface and the tires.

2.2. *Apparatus*

A virtual driving simulator featuring simulated icy road conditions was used to conduct the laboratory experiment. The simulator utilized a network of two computers and control circuit for the generation of the auditory alarm signal. An SGI Onyx 2 computer (client), which provided a 21″ screen for the simulation was connected to an Optiplex series computer (server). The server collected data through a National Instruments PCI-6024E data acquisition board. The Microsoft Sidewinder Force Feedback wheel package (consisting of steering wheel, brake and throttle pedals) was use to allow control over the simulated vehicle, and to provide force feedback through the steering wheel. An audio alarm device was constructed which produced an audio signal of the range of 250Hz to 500Hz. Lab View and World Tool Kit (WTK) were used to develop the simulation scenarios (Singh *et al*. 2000).

A simplified bicycle model of an automobile was used to simulate vehicle dynamics, and was instantiated in a computer simulation linked to a suitable virtual reality display. The bicycle model is a reduced degree of freedom model, which includes frictional behaviour at each wheel of the bicycle model, yaw motion of the vehicle and its longitudinal and lateral motion (Sobottka and Singh 1996). The tire model used was developed by Szoatak *et al*. as reported in the US Department of Transportation Report (NHTSA DOT–HS–805–271). This model of the tire provides lateral and longitudinal forces generated by the tire under different conditions of slip, slip angle and the force that can be generated by the tires.

The track simulated for the experiment was a two-lane road with no vehicular traffic on either side. Two yellow lines separated the two lanes. There were no road signs or traffic signals. The total length of the track was approximately 3600 m; the layout is shown in figure 1. An appearance of snow on the road and surrounding area was created using a white surface and simulated snow texture. Participants saw a view through the windshield along with a dashboard display on the computer screen (see figure 2).

2.3. *Independent variables*

Two different alerting levels (low sensitivity and high sensitivity), and two different types of alarm signals (step signal and ramp signal) were generated to obtain four experimental conditions. There was also a fifth, no-alarm control condition.

2.3.1. *Alerting Level*: Alarm activation was based on a threshold value, which depended on the coefficient of friction between the road surface and car tires, and a

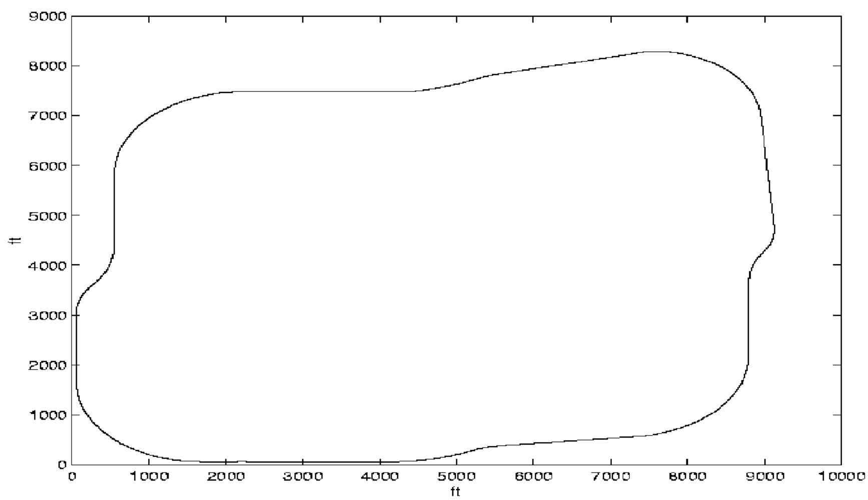

Figure 1. Layout of the simulated track.

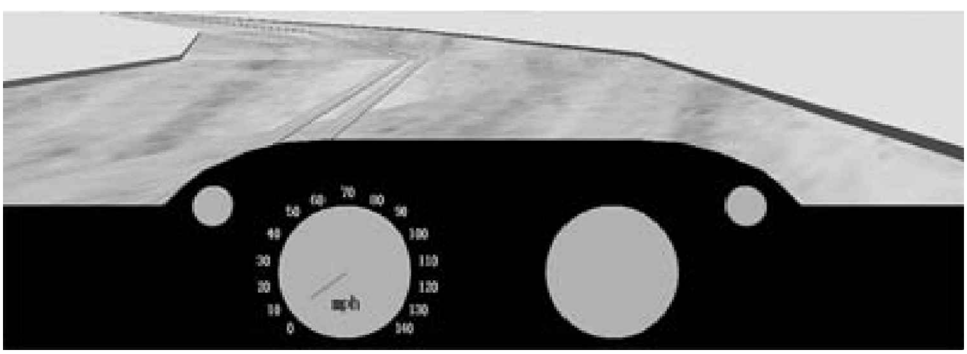

Figure 2. Through the windshield display seen by the participants.

gain parameter. The gain could be manipulated to generate different threshold values, which would consequently change the point at which the alarm activated. At the highest level of gain, the alarm would not activate until the point at which the car actually started to skid. Two different values of the threshold were used which gave rise to two conditions within the alerting level:

- Alerting level 1 (low sensitivity): the first threshold value was set to simulate a low sensitivity alerting level. For this condition, the value of the sensitivity gain was set at 0.8. Theoretically, this alerting level was 80% of the actual skid threshold.
- Alerting level 2 (high sensitivity): the second threshold value was set to simulate a high sensitivity alerting level. For this condition, the value of the sensitivity gain was set at 0.3. Theoretically, this alerting level was 30% of the actual skid threshold.

The high sensitivity alerting level produced more alerts, and thus may have resulted in a perception of greater false alarms, than the low sensitivity alerting level.

2.3.2. *Alarm Signal*: An auditory alarm was used in the study. There were two types of alarm signals: a step or Binary ON/OFF signal and a ramp or graded signal. In the step signal condition, the auditory alarm was activated at a constant level of amplitude (loudness) when the threshold resistive force was reached and at any point thereafter. In the ramp signal condition, the amplitude of the alarm signal was a linear function of the resistive force: the amplitude increased from 0 at the threshold force level to its maximum level at the force where a skid would occur. Thus, the ramp signal condition provided an indication of the level of danger.

2.4. *Participants*

Twenty-five participants from the general student population at the University at Buffalo were recruited. The participants were required to have a valid driving license, driving experience of two years or more and an average driving pattern of 5000 miles (8000 km) or more per year. Each received a financial compensation of $7.00 per hour for their participation.

2.5. *Experimental design and procedure*

Alarm sensitivity and type of alarm signal were between-subjects factors. Additionally, there was a no-alarm control condition, giving a total of five between-subjects conditions. Participants were randomly assigned to conditions.

Eighteen driving scenarios were created. In order to incorporate some uncertainty about road conditions, a randomly varying coefficient of friction from 0.2 to 0.8 was set over the length of the track. Additionally, wind gusts were simulated by dividing the track into fifteen equal sections, introducing one gust at a random point in each section. The duration of the gust varied randomly from 0.1 to 1 seconds. The magnitude (force) of the wind gusts was dependent on the value of the coefficient of friction between the road surface and the car tires. Coefficients of friction and the occurrence and characteristics of the wind gusts were varied across the 18 scenarios.

Five participants drove in each of the condition. Each participant drove six scenarios each day for three consecutive days, for a total of 18 scenarios. The first three scenario runs were used as training runs and were not considered in the analysis. Participants were not aware that data from these runs would be discarded. In the case of a skid or spin-out which ended a scenario, participants were instructed that they would have to begin the scenario again from the start; they were not given any other instructions regarding driving strategy. To determine the participants' perception of trust in the adverse condition warning system, a 12-item questionnaire (Jian *et al.* 2000) was filled out by the participants after the first and third day. Questionnaire items are shown in table 1.

2.6. *Dependent measures*

Dependent measures included task performance measures, participant response measures and responses to the questionnaires. Except for the questionnaire responses, measures were extracted from log files that

Table 1. Items in the trust questionnaire.

The system is deceptive
The system behaves in an underhanded manner
I am suspicious of the system's intent, action or output
I am wary of the system
The system's action will have a harmful or injurious outcome
I am confident in the system*
The system provides security*
The system has integrity*
The system is dependable*
The system is reliable*
I can trust the system*
I am familiar with the system*

*Positively framed questions.

captured simulation parameters at 0.025 second intervals during the scenario runs. The following dependent measures were collected:

- *Number of skids*: the number of skids, defined as instances when the car went out of control (spun around) and could not be recovered, was recorded.
- *Longitudinal velocity*: this variable represented the velocity of the vehicle in the longitudinal direction, making it an indicator of the vehicle speed.
- *Yaw angle*: yaw angle is the angle between the direction the vehicle is pointing and the direction it is travelling. A measure of the yaw angle helps assess the vehicle handling capability of the driver: a smaller yaw angle implies better vehicle control.
- *Steering angle*: steering angle is the angle between the longitudinal axis of the car and the direction in which the wheel rim points. This angle is induced by the direct response of the driver on the steering wheel of the vehicle.
- *Slip angle*: slip angle refers to the direction the wheel rim is pointing vs. the wheel's path over the surface of the road. Slip angle helps assess the handling of the vehicle by the drivers. A smaller slip angle implies better control.
- *Alarm signal*: an indication of the alarm signal activating was collected in order to determine the drivers' responses around and at that instance.
- *Lateral acceleration*: this score represented the deviation of the vehicle in the lateral direction. Lateral acceleration is the force that is imparted on the vehicle in the direction of the steering wheel movement. This force is dependent on the vehicle speed and the adhesion friction (between road surface and vehicle tires). The outward force is resisted by a restoring moment applied through friction forces at the tires. Under constantly changing friction (as in snowy conditions), the driver has to adjust the steering wheel movement and the speed in order to avoid rollovers. Higher lateral acceleration indicates a higher chance of a rollover and greater instability.
- *Action on the brake pedal*: response of the driver on the brake pedal during the entire duration of the simulation was collected. This was a measure of the pedal displacement. The pedal displacement spanned from 0 to 1, where 1 represented the pedal being completely pressed. This value were multiplied by the maximum braking torque as follows:- maximum braking torque: 3400 lbf-ft (4610 Nm; 1984 Honda Accord Specification).
- *Action on the throttle pedal*: response of the driver on the throttle pedal during the entire duration of the simulation was collected. Similar to the brake pedal, the displacement from 0 to 1 was multiplied by the maximum throttling torque (1000 lbf-ft; 1360 Nm). This value was used as per the 1984 Honda Accord specification.
- *Response to questions*: each participant filled a questionnaire after the first and the third sessions of the experiment. This was a 12-item questionnaire where each question required a response on a 7-point scale ranging from 'not at all' to 'very much'.

3. Results

Analyses of variance were used to investigate differences in these measures between conditions. A two-way ANOVA with alarm sensitivity and type of alarm signal treated as between-subjects factors was used. An initial analyses that included session as a within-subjects factor indicated few significant effects of session. Inspection of the significant effects of session (on slip angle, response on the throttle, and velocity five seconds after the alarm) indicated differences due to random variations in session composition rather than any trends over time. Thus, to increase statistical power for tests of the variables of primary interest (alerting level and alarm signal type), session was excluded from the final analyses for measures related to performance and driver response. For the trust questionnaire responses, session was analysed as a within-subjects factor. Additionally, in order to compare results from the four alarm conditions (two levels of sensitivity and two levels of alarm type) to the no-alarm control condition, a five-level one-way ANOVA was used. Statistical results, discussed below, are summarized in table 2.

There were no main effects of alarm type or sensitivity level on the number of skids. There was, however, a significant two-way interaction between alarm type and sensitivity level. A one-way ANOVA comparing the control group to the four alarm conditions was significant, and *post hoc* tests showed that there were significantly fewer skids in the low sensitivity, ramp alarm type condition than the high sensitivity, ramp alarm type condition, but that there were no other differences. These results are plotted in figure 3.

Results for yaw angle, slip angle and lateral acceleration were all similar to each other. There were significant main effects of alarm type for yaw angle and lateral acceleration. Values of these measures were higher for the step than ramp alarm signal, indicating better control over the vehicle in the latter case. The trend was similar for slip angle (mean angle of 0.032 radians in the step condition, and 0.026 radians in the ramp condition) but the effect was not significant ($F_{1,}$

Table 2. Summary of ANOVA Results.

Measure	Sensitivity	Type of alarm	Sensitivity alarm type x	Comparison to control
Number of skids	n.s.	n.s.	$F_{(1, 296)} = 7.60$**	$F_{(4, 370)} = 2.47$ *
Yaw angle	n.s.	$F_{(1, 296)} = 15.4$**	n.s.	$F_{(4, 370)} = 9.43$**
Slip angle	n.s.	n.s.	n.s.	n.s.
Lateral accel.	n.s.	$F_{(1, 296)} = 10.02$**	n.s.	$F_{(4, 370)} = 5.62$**
Overall velocity	n.s.	$F_{(1, 296)} = 19.29$**	$F_{(1, 240)} = 19.41$**	$F_{(4, 370)} = 26.48$**
Final run vel.	n.s.	$F_{(1,296)} = 11.65$**	$F_{(1, 296)} = 12.36$**	$F_{(4, 370)} = 11.43$**
Vel. at skid	$F_{(1, 296)} = 3.63^{\#}$	n.s.	n.s.	$F_{(4, 108)} = 4.025$**
Throttle response	n.s.	$F_{(1, 296)} = 4.82$*	n.s.	N/A
Vel. after alarm	$F_{(1, 296)} = 4.49$*	$F_{(1, 296)} = 29.44$**	n.s.	N/A
Steering reaction	$F_{(1, 240)} = 4.62$*	n.s.	n.s.	N/A

The rightmost column provides results for the 5-level, one-way ANOVA which compared the experimental conditions to the control condition. *$p < 0.05$; **$p < 0.01$; $^{\#}p = 0.06$

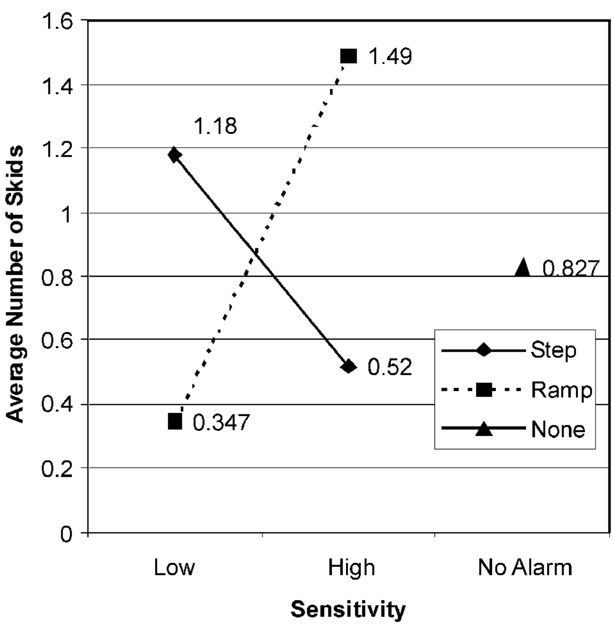

Figure 3. Effect of alarm sensitivity and signal conditions, as well as the no-alarm control condition, on the average number of skids.

$_{296} = 2.83$, $p = 0.093$). There were no significant main effects of alarm sensitivity on these variables, and the interaction was not significant.

Comparison of the alarm conditions to the no alarm conditions showed a significant main effect for yaw angle and lateral acceleration; *post-hoc* tests indicated that the control group participants had a similar yaw angle to participants in the ramp alarm conditions, and a significantly higher lateral acceleration than participants in the high sensitivity, ramp alarm condition. Thus, patterns of results were similar across alarm conditions. However, while the no-alarm condition resulted in worse performance compared to the alarm conditions when measured by lateral acceleration, performance without the alarm was similar to the best alarm condition performance when measured by yaw angle. These results are illustrated in figure 4. There was no main effect across the five conditions on slip angle.

The overall velocity during each participant's sessions provided an indication of the drivers' behaviour in selecting the speed at which to drive. There was a significant main effect of type of alarm signal as well as a significant interaction between alarm type and sensitivity level as shown in figure 5. *Post hoc* analyses indicated that the velocity for the step alarm type, high sensitivity condition was greater than the velocities for the two low sensitivity conditions, which in turn were greater than the ramp alarm, high sensitivity condition. Comparison of all conditions, including the control condition, indicated a significant difference among the five conditions; results are shown in figure 5. *Post hoc* tests revealed that the no-alarm condition had significantly higher velocity than all other conditions: participants drove faster in the absence of the alarm. Velocities during the final (no-skid) runs for each session were analysed separately and results were similar, indicating that although the experimental condition affected speed, speeds were consistent across attempts at each session: attempts with skids did not tend to have higher or lower speeds than those without.

The velocity at the time of a skid was also analysed. Results indicated that there was a marginal effect of alarm sensitivity: average velocity was greater for the low levels of alarm sensitivity (13.11 m/sec for the step alarm, and 14.88 m/sec for the ramp) than high levels of alarm sensitivity (11.91 m/sec for the step alarm, and 12.17 m/sec for the ramp). There was no effect of alarm type, or interaction. Similar to overall velocity, comparison across all five conditions showed significance, and post hoc tests showed that the velocity at the instance of a skid (16.3 m/sec) was higher in the

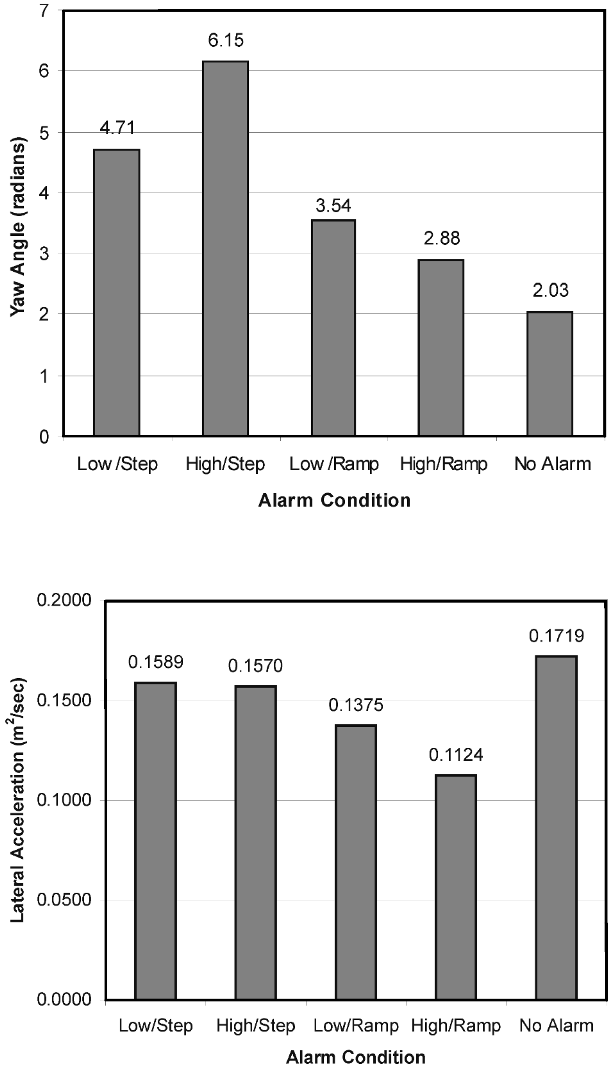

Figure 4. Effect of alarm sensitivity and signal conditions, as well as the no-alarm control condition, on yaw angle and lateral acceleration.

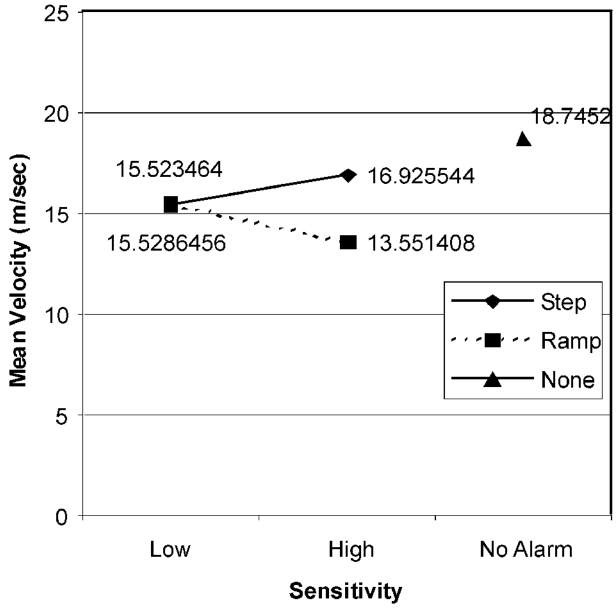

Figure 5. Effect of alarm sensitivity and signal conditions, as well as the no-alarm control condition, on mean overall velocity.

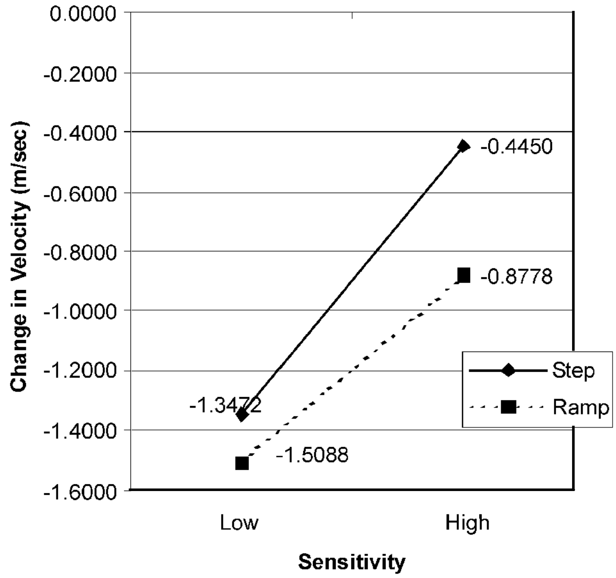

Figure 6. Effect of alarm sensitivity and signal conditions on the change in velocity 5 seconds after the onset of the alarm.

no-alarm condition than all but the low sensitivity/ramp alarm condition.

Participants' responses to the activation of alarm were also analysed. One such measure was the change in velocity measured five seconds after the activation of the alarm compared to the velocity at the time of the alarm. There were significant main effects of both alarm sensitivity and type of alarm signal; no interactions were significant. As shown in figure 6, there was a reduction of velocity in all cases, but greater reduction for the ramp alarm condition, and for the low sensitivity condition.

The change in throttle displacement (a scaled measurement from 0 to 1000, where 1000 was fully depressed) one second after the activation of the alarm signal was analysed. Results indicated a significant effect of type of alarm signal; the mean for the step alarm was 23.95 while the mean for the ramp alarm signal was −4.54. The positive change for the step alarm indicated that participants actually increased the displacement of the throttle after the alarm activated, perhaps indicating that they were startled. There were no other significant effects or interactions. Braking response was also noted and was consistent across conditions: all participants had applied the brake by one second following the activation of the alarm signal.

Finally, response on the steering wheel was also analysed. The absolute differences between the steering wheel angle at the time of the alarm, and one second after alarm activation, were compared to see the extent of the response on the steering wheel following the alarm signal. There was a significant main effect of alarm sensitivity, with an average deviation of 0.016 radians for the low sensitivity condition and 0.046 radians for the high sensitivity condition, indicating a greater response for the high sensitivity condition. There were no other significant effects or interactions.

Since the activation of the alarm was a function of driver behaviour in addition to the scenario conditions, the number of alarms (points at which the threshold force was crossed) was also analysed. There was a main effect of both alarm type and alarm sensitivity ($F_{1, 296} = 20.03$, $p < = 0.000$; $F_{1, 296} = 329.28$, $p < = 0.000$). As expected, there were a greater number of alarms for the high vs. low sensitivity condition (means of 21.72 and 18.75 respectively); and there were also a greater number of alarms for the step vs. ramp conditions (means of 34.48 and 8.47, respectively). The interaction was not significant.

Finally, responses to the trust questionnaire were analysed. Average responses to the five negatively framed questions were analysed separately from responses to the seven positively framed questions; results are shown in figure 7. There was a significant main effect of alarm sensitivity on responses to both the negatively framed questions ($F_{1, 16} = 4.78$, $p = 0.05$) and positively framed questions ($F_{1, 16} = 8.392$, $p = 0.011$). Responses indicated a greater degree of trust for the low sensitivity condition. Additionally, responses were compared across time: the questionnaire was administered and the end of the first and third days. There was a significant effect of time on responses to both the negatively framed questions ($F_{1, 16} = 11.276$, $p = 0.004$) and positively framed questions ($F_{1, 16} = 13.146$, $p = 0.002$): responses indicated an increase in trust across time. There was no affect of alarm signal type, and no significant interactions. Thus, analysis of the participants' trust in these systems showed significantly higher positive feelings of trust and lower feelings of distrust for the low sensitivity alarm conditions. This was consistent with the performance results obtained and demonstrated the potentially negative effect of the higher sensitivity system, where the alarm activated at a lower threshold. This may have appeared to drivers to have a greater number of false alarms. While there was also a consistent increase in trust as session increased, the difference between sensitivity levels remained.

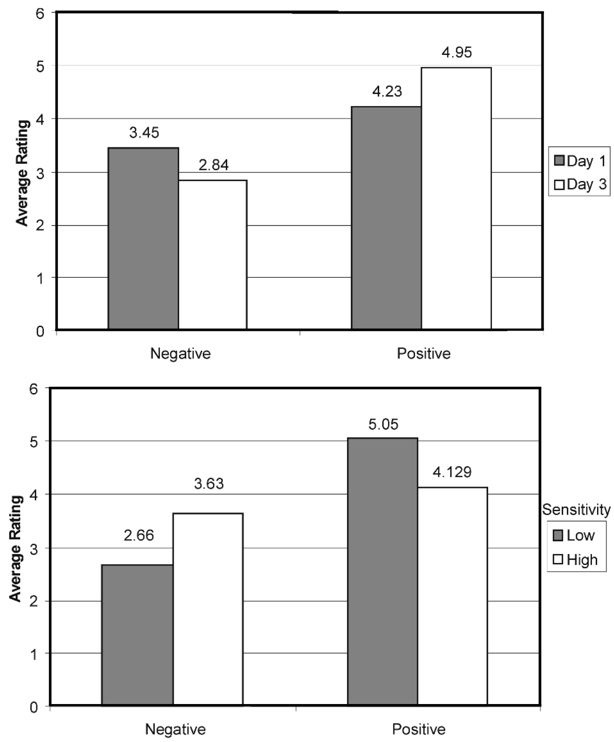

Figure 7. Results of the trust questionnaires, showing average responses to positively and negatively framed questions. The top graph shows the effect of session: the trust questionnaire was administered at the end of the sessions on the first and third days. The bottom graph shows the effect of alarm sensitivity.

4. Discussion

4.1. *Effects of alarm sensitivity*

In the case of the system under consideration, alarm sensitivity represented the point at which the warning system alerted the driver. This point was 30% of threshold for a high sensitivity alarm signal and 80% of the threshold for the low sensitivity alarm. Consequently, the high sensitivity alarm signal provided more time to react between the onset of the alarm and the threshold (point at which a skid would occur). Additionally, one would expect a greater sense of false alarms, or alarms occurring without justification, for the higher sensitivity alarm.

The results obtained indicate effects of alarm sensitivity for the velocity at the instance of skid, change in velocity 5-seconds following the alarm, and the steering wheel reaction 1-second following the alarm, as follows.

Velocity at the instance of a skid was higher for the low compared to the high sensitivity alarm. Additionally, participants in the low sensitivity condition had a greater reduction in velocity 5 seconds after the alarm onset, compared to the velocity at the time the alarm activated. An explanation for this apparently contradictory result is that there was greater time available for participants in the high sensitivity alarm condition to reduce their vehicle speeds before the onset of a skid. That is, since the alarm went off sooner, participants had more time to slow down after the alarm, and before the skid. Additionally, the analysis of the driver response on the steering wheel 1-second after the alarm was significantly greater for the high sensitivity alarm, indicating a greater deviation—participants turned the steering wheel more—for the high sensitivity alarm signal. Since steering wheel deviation is an important aspect of the drivers control over the car under slippery road conditions, a higher deviation for the high sensitivity alarm indicated decreased performance compared to the low sensitivity alarm. Additionally, participants in the high sensitivity condition had greater negative feelings of trust, and lower positive feelings of trust, than those in the low sensitivity condition. This result was consistent with hypotheses that participants would tend to reduce their trust in the system in the case when the alarm activated earlier, under conditions participants may not have believed to be hazardous. Continued distrust in an alerting system may result in drivers ignoring or disabling such systems. Overall, there was no indication that alarm sensitivity affected driver control as measured by yaw angle, slip angle, and lateral acceleration.

Previous research has suggested that providing an alarm very close to the threshold (low sensitivity) in immediate response adverse condition warning systems might startle the driver and lead to increased stress, delay in action, and incorrect response (Dingus *et al.* 1997). In this study, however, the lower sensitivity alarm tended to lead to increased trust and better performance as measured by steering wheel deviation; thus, a high level of sensitivity is not necessarily preferable. It seems likely that the optimal level of sensitivity for a system implemented in a real vehicle, in terms of both performance, and operator trust, will depend on particular characteristics of the alerting system sensors and algorithms, and potentially operator preference. Therefore, this characteristic is a candidate for further study in more realistic vehicle simulators.

4.2. *Effects of type of alarm signal*

Significantly lower values of yaw angle and lateral acceleration, and a similar trend for the slip angle, indicate generally better performance controlling the vehicle in the ramp signal condition, when the system provided a graduated alarm signal. Additionally, under the ramp condition, the difference in throttle displacement during and after the alarm was negative, in contrast to the positive value obtained for the step signal. That is, in the step signal condition, participants actually increased their depression of the accelerator pedal immediately after the alarm signal onset, strongly indicating that they were startled by the more sudden onset of the step alarm. Finally, there was no impact of alarm signal on participants' trust in the alerting system.

These results are consistent with those of prior research. For instance, Stokes *et al.* (1990) describes the potential difficulties due to the startling effect of alarms. Results obtained in the present study support this claim, since better overall control (in terms of yaw and slip angles, and lateral accelerations) and driver responses (in terms of response on the throttle) were observed in the ramp signal. Additionally, the ramp alarm signal in some sense provides an indication of the imminence of a skid. This may be interpreted by participants as the chance that a skid is going to occur, since when a skid is less imminent (less of the skid threshold has been reached), participants have more opportunity to act to avoid a skid. Other research has shown that for situations where false alarms are likely, a display based on the chance of an event occurring may be beneficial (Sorkin *et al.* 1988). Finally, this condition has similarities to the graded sequence of warnings that Horowitz *et al.* (1992) suggest would lead to better performance.

Some of the impacts of alarm sensitivity were dependent on the alarm signal type. For instance, the mean overall velocity was lower in the low sensitivity condition than the high sensitivity, step alarm condition, but higher than the high sensitivity, ramp alarm condition. It is possible that drivers in the low sensitivity condition were more cautious than those in the high sensitivity condition, since they were less likely to receive an alert. However, those drivers in the high sensitivity, ramp condition may have made better use of the graduated alarm information than those in the step condition, and reduced their speed after the alarm activated. Additionally, low sensitivity alarm combined with the ramp alarm condition had fewer skids than the high sensitivity, ramp alarm condition. Since skids are due to a loss in lateral control over the vehicle, an explanation for this result may be found in the greater deviation in steering after an alarm in the high sensitivity condition.

4.3. *Changes in trust over time*

Prior research and theories regarding trust have suggested that trust in automated systems can change over time, due to people's increased experience with the automated system (Lerch and Prietula 1989, Lee and Moray 1994, Muir and Moray 1996). In many of those studies, however, the focus of trust has been on systems for automated control, rather than decision support systems or, more specifically, an alerting system such as the one investigated here. Results of this study are consistent with those previous outcomes, indicating that results and theories regarding trust in automated control systems are potentially generalizable to trust in vehicle based alerting systems. As noted above, ratings of trust increased with all alerting system conditions over time. There was no indication that differences in trust between alarm sensitivity conditions would be mitigated by time. However, an experiment with additional experience could potentially change that result.

4.4. *Overall impact of the alerting system*

Overall, the presence of the alerting system tended to have a positive effect on performance compared to the control group. Participants in the control group had a higher final run velocity and velocity at the instance of a skid than participants in the four alarm conditions, indicating more conservative driving choices. Additionally, participants in the control group showed decreased performance in terms of lateral acceleration. Unexpectedly, however, the analysis of yaw angle—another measure of vehicle control, showed least yaw displacement for the no-alarm control group. Finally, participants in the control group tended to have a higher number of skids than some of the alarm conditions, although this was not a significant difference. The tendencies for a positive impact of the alerting system suggest that such systems can provide important, additional cues to drivers regarding road conditions. Such a system can supplement visual information regarding road conditions. While the visual cues used in the simulation do not exactly mimic those available from a real road surface, there are certainly real world cases (e.g. wet patches that are really icy; icy surfaces covered with snow) where people have difficulty in judging the surface coefficient of friction based on visual information.

4.5. *Study limitations and implications for future work*

An important limitation of the current study was the level of fidelity of the vehicle simulator. While the simulator software provided a relatively high fidelity simulation of the vehicle dynamics and operation of the alerting systems, the study utilized a 21'' computer screen as a visual representation of the outside world during driving. In addition, participants did not have a feeling of motion during the study. In slippery road conditions, the driver may react to a situation by understanding the grip or stability of the tires on the road surface. Future studies could provide additional design insights by incorporating motion cues through the use of motion platforms that could simulate the various accelerations experienced by the driver. Additionally, user demographics such as age, experience and risk taking capacity may play an important role in the effective use of an ACWS Additional studies could help assess the importance of these factors.

A major issue in alerting system design is the selection of alerting modalities. While the current study investigated auditory alerting signals, further research on various other modalities such as visual, tactile or some combination of these modalities could help determine the modality best suited for such systems. Though auditory alarms are recommended for urgent situations to attract attention quickly, drawbacks such as startling the driver (as indicated in this study) might cause other modalities or combination of modalities to be more effective.

Finally, this study indicated that the level of alarm sensitivity can have a measurable affect on both performance and level of trust. Studies in more realistic simulation, combined with additional sensitivity levels, could further refine these results.

5. Conclusions

The alerting system conditions under which the participants drove invoked different responses and variable performance. The higher sensitivity alarm signal induced more conservative behaviour from the participants as they reduced speeds to a greater extent as compared to the participants in the low sensitivity alarm signal, possibly due to the greater response time available before reaching the threshold. However, in the high sensitivity condition, participants also produced greater steering wheel deviations: a trait that is not desirable in slippery road conditions. Additionally, the higher sensitivity condition resulted in greater negative feelings of trust, and less positive feelings of trust, than the low sensitivity condition, a difference that persisted across sessions. This analysis may indicate that the appropriate sensitivity level for the alarm lies somewhere between 30% and 80% of threshold, at a point where drivers would have enough time to respond, but not so much time that they overcompensate.

For the type of alarm signal, a graduated auditory warning based on distance to threshold resulted in better performance across a range of dependent measures. In particular, participants exercised more lateral control over the vehicle. Consistent with the literature, a graduated alarm based on urgency of the situation was better. Additionally, the ramp signal combined with the low sensitivity level provided the lowest number of skids. Overall, a low sensitivity alarm and a graded alarm signal would be the suggested design for adverse condition warning systems.

Acknowledgements

This work was supported by a grant from the Center for Transportation Injury Reduction at the CALSPAN/UB Research Center. The authors gratefully acknowledge the contributions of Yang Cao and Dr T. Kesevadas for their work in developing the simulator.

References

Allen, R. W., Ziedman, D., Rosenthal, T.J., Torres, J. and Halati, A. 1991, *Laboratory Assessment of driver route diversion in response to in-vehicle navigation and motorist information systems.* In SAE Technical Paper series (SAE No. 910701,), Warrendale, PA: Society of Automotive Engineers, pp. 1 – 25.

Barfield, W. and Dingus, T. A. 1998, *Human factors in Intelligent Transportation Systems* (Mahwah, NJ: Lawrence Erlbaum).

Bois, W. 1982, On board check system with speech synthesis. In *IEEE 1982 Workshop on Automotive Application and Microprocessors*, Institute of Electrical and Electronics Engineering, pp. 88 – 92.

Bonsall, P. W. and Parry, T. 1991, Using an interactive route-choice simulator to investigate driver's compliance with route guidance advice. *Transportation Research Record*, **1306**, 59 – 68.

Dingus, T. A. and Hulse, M. C. 1993, Human Factors Research Recommendations For The Development of Design Guidelines For Advanced Traveler Information Systems. In *Proceedings of the Human Factors and Ergonomics Society, 37th Annual Meeting*, Seattle, WA, October 11 – 15, pp. 1067 – 1071.

Dingus, T. A., Hulse, M. C. and Barfield, W. 1998a, Human-system interface issues in the design and use of advanced traveler information systems. In W. Barfield and T. A. Dingus (eds) *Human Factors in Intelligent Transportations Systems* (Mahwah, NJ: Lawrence Erlbaum), pp. 55 – 94.

Dingus, T. A., Jahns, S. K., Horowitz, A. D. and Knipling, R. 1998b, Human factors design issues for crash avoidance systems. In W. Barfield and T. A. Dingus (eds) *Human Factors in Intelligent Transportation System* (Mahwah, NJ: Lawrence Erlbaum), pp. 55 – 94.

Dingus, T. A., McGhehee, D. V., Manakkal, N., Jahns, S.K., Carney, C. and Hankey, J. M. 1997, Human factors field evaluation of automotive headway maintenance/collision warning devices. *Human Factors*, **39**, 216 – 229.

Doll, T. J., Folds, D. J., and Leinker, L. 1984, *Auditory Information System in Military Aircraft: Current Configurations versus the state of the art*. USAF School of Aerospace Medicine, Brook Air Force Base, TX.

Horowitz, A. D. and Dingus, T. A. 1992, Warning signal design: a key human factors issue in an in-vehicle front-to-rear end collision warning systems. In *Proceedings of the Human Factors and Ergonomics Society, 36th Annual Meeting*, Atlanta, GA, October 12 – 16, pp. 1011 – 1013.

Jian, J. Y., Bisantz, A. M. and Drury, C. G. 2000, Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, **4**, 53 – 71.

Kantowitz, B. and Sorkin, R. D. 1983, *Human Factors: Understanding People – System Relationships* (New York: Wiley and Sons).

Kantowitz, B. H., Hanowski, R. J. and Kantowitz, S. C. 1997, Driver acceptance of unreliable traffic information in familiar and unfamiliar settings. *Human Factors*, **39**, 164 – 176.

Lee, J. D. and Moray, N. 1994, Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human – Computer Studies*, **40**, 153 – 184.

Lee, J. D. and Moray, N. 1992, Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, **35**, 1243 – 1270.

Lee, J. D. and Kantowitz, B. H. 1998, Perceptual and Cognitive Aspects of Intelligent Transportation Systems. In W. Barfield and T.A. Dingus (eds) *Human Factors in Intelligent Transportation Systems* (Mahwah, NJ: Lawrence Erlbaum), pp. 31 – 54.

Lerch, F. J. and Prietula, M. J. 1989, How do we trust machine advice? In G. Salvendy and M. J. Smith (eds) *Designing And Using Human-Computer Interface And Knowledge Based Systems* (North-Holland: Elsevier Science Publishers), pp. 410 – 419.

Mast, T. 1998, Introduction to ITS. In W. Barfield and T.A. Dingus (eds) *Human Factors in Intelligent Transportation Systems* (Mahwah, NJ: Lawrence Erlbaum), pp. xv – xxi.

Mollenhauer, M. A., Klee, J., Cho, K., Hulse, M .C. and Dingus, T. A, 1994, Effects of sensory modality and information priority on in-vehicle signing and information systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Nashville, TN, October 24 – 28, pp. 1072 – 1076.

Muir, B. M. and Moray, N. 1996, Trust in automation: part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, **39**, 429 – 460.

Parasuraman, R., Molloy, R. and Singh, I. L. 1993, Performance consequences of automation-induced 'complacency'. *The International Journal of Aviation Psychology*, **3**, 1 – 23.

Patterson, R. D. 1982, *Guidelines for Auditory Warning Systems in Civil Aircraft*. Civil Aviation Authority paper 82017, London.

Pritchett, A. R. 2001, Reviewing the role of cockpit alerting systems. *Human Factors and Aerospace Society*, **1**.

Rempel, J. K., Holmes, J. G. and Zanna, M. P. 1985, Trust in Close Relationships. *Journal of Personality and Social Psychology*, **49**, 95 – 112.

Scott. B. L. and Kee, C. R. 1987, Speech Recognition in the automobile: coping with noise in the automotive environment. In *Recent Developments in Automotive Electronic Displays and Information Systems* (Warrendale, PA: Society of Automotive Engineers, pp. 83–85.

Sheridan, T. B. 1988, Trustworthiness of command and control systems. In *Proceedings of The IFAC/IFIP/IFORS/IEA Conference on Analysis, Design, and Evaluation of Man-Machine Systems*, Oulu, Finland, pp. 427–431,

Singh, T., Kesavadas, T., Mayne, R., Kim, J.-J. and Roy, A. 2000, *Design of Hardware/Algorithms for Enhancement of Driver-Vehicle Performance in Inclement Conditions Using a Virtual Environment*, SAE 2000 World Congress, March 6-9, Detroit, MI, SAE Technical Paper Series 2000-01-0322.

Sobbota, C. and Singh, T. 1996, Optimal fuzzy logic control for an anti-lock braking system. *Proceedings of the IEEE Conference on Control Application*s, pp. 49–54.

Sorkin, R. D., Kantowitz, B. H. and Kantowitz, S. C. 1988, Likelihood alarm displays. *Human Factors*, **30**, 445–459.

Srinivasan, R. and Jovanis, P. P. 1997, Effect of selected in-vehicle route guidance systems on driver reaction times. *Human Factors*, **39**, 200–215.

Stokes, A., Wickens, C. and Kite, K. 1990, *Display Technology: Human Factors Concepts*. Society of Automotive Engineers, Washington DC.

Thompson, D. A. 1981, Commercial aircrew detection of system failures: state of the art and future trends. In J. Rasmussen and W.B. Rouse (eds) *Human detection and diagnosis of system failures* (New York: Plenum Press).

Wheeler, W. A., Campbell, J. L. and Kinghorn, R. A. 1998, Commercial vehicle-specific aspects of intelligent transportation systems. In W. Barfield and T. A. Dingus (eds) *Human Factors in Intelligent Transportation Systems* (Mahwah, NJ: Lawrence Erlbaum), pp. 95–130.

Zwahlen, H. T. 1985, Driver eye scanning, the information acquisition process and sophisticated in-vehicle information displays and controls. In *Ergonomics International 1985 Proceedings*, pp. 508–510.