



U.S. Department
of Transportation
National Highway
Traffic Safety
Administration



DOT HS 808 533

February 1996

Technical Report

Inappropriate Alarm Rates and Driver Annoyance

This publication is distributed by the U.S. Department of Transportation, National Highway Traffic Safety Administration, in the interest of information exchange. The opinions, findings and conclusions expressed in this publication are those of the author(s) and not necessarily those of the Department of Transportation or the National Highway Traffic Safety Administration. The United States Government assumes no liability for its contents or use thereof. If trade or manufacturers' name or products are mentioned, it is because they are considered essential to the object of the publication and should not be construed as an endorsement. The United States Government does not endorse products or manufacturers.

	2. Government Accession No.	3. Recipient's Catalog No.	
Title and Subtitle Inappropriate Alarm Rates and Driver Annoyance	5. Report Date 2/96		6. Performing Organization Code
	8. Performing Organization Report No.		
Author(s) Lerner, N.D., Dekker, D.K., Steinberg, G.V., and Huey, R.W.		10. Work Unit No. (TRAILS)	
i. Performing Organization Name and Address COMSIS Corporation 8737 Colesville Road Suite 1100 Silver Spring, Maryland 20910		11. Contract or Grant No. DTNH22-91-C-07004	
		13. Type of Report and Period Covered Technical Report	
2. Sponsoring Agency Name and Address Office of Crash Avoidance Research National Highway Traffic Safety Administration U.S. Department. of Transportation Washington, DC 20590		14. Sponsoring Agency Code	
		5. Supplementary Notes	
6. Abstract <p>Future in-vehicle crash avoidance warning systems will inevitably deliver inappropriate alarms from time to time, caused for example, by situations where algorithms have correctly identified an object but pose no threat or danger to the driver. The current state of knowledge does not permit an estimate of how many inappropriate alarms users find unacceptable, and how that rate may vary with factors like the type of signal generated by the system (i.e., tone versus voice), or extended experience with the warning system itself. The purpose of this study is a direct comparison of drivers' subjective annoyance towards inappropriate alarms as a function of rate of occurrence and the type of signal generated in naturalistic, on-road driving conditions. Test equipment to generate and present signals, and to collect driver response was installed in fifteen participants' personal vehicles for a nine week period. Signals were presented at random times while the participants engaged in their normal, daily driving routines. In order to simulate future operating conditions where actual alarm warnings will require the driver's attention and reaction, "appropriate" alarms to which the driver had to make a simple motor response, and "inappropriate" alarms to which the driver did not have to make any response, were presented. Inappropriate tonal alarms were presented at four different frequencies of occurrence, including averages of four per hour, one per hour, one per four hours, and one per eight hours of driving time. In addition, a voice warning condition was included, at a rate averaging one per hour. Participants made daily and weekly ratings of the degree of annoyance that resulted from the nuisance alarm schedule. The 4/hour-tone and the 1/hour-voice were significantly more annoying, and less acceptable, than the other conditions. Participants showed a wide range of annoyance sensitivity, but the two most annoying conditions appear to be unacceptable, while the less frequent rates do appear potentially reasonable for functional systems.</p>			
17. Key Words ITS, crash avoidance warnings, alarms, annoyance		18. Distribution Statement	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages	22. Price

TABLE OF CONTENTS

ABSTRACT	i
1.0 Introduction.....	1
1.1 Determinants of Annoyance Response to Inappropriate Alarms.....	2
1.2 Ability to Simulate Real-World Annoyance Factors in Experimental Studies.....	3
2.0 Method.....	7
2.1 Participants.....	7
2.2 Apparatus	7
2.3 Procedure	10
3.0 Results	15
3.1 Analytic Method.....	15
3.2 Noticeability Ratings.....	15
3.3 Annoyance Ratings.....	16
3.4 Acceptability Ratings.....	19
3.5 Factors Influencing Annoyance.....	21
4.0 Discussion.....	23
5.0 References.....	26
Appendix A: Daily Driver Questionnaire.....	A-1
Appendix B: Weekly Driver Questionnaire.....	B-1
Appendix C: Traffic Conditions Rating Form.....	C-1

ABSTRACT

Future in-vehicle crash avoidance warning systems will inevitably deliver inappropriate alarms from time to time, caused for example, by situations where algorithms have correctly identified an object, but pose no threat or danger to the driver. The current state of knowledge does not permit an estimate of how many inappropriate alarms, users find unacceptable, and how that rate may vary with factors like traffic conditions, the type of signal generated by the system (i.e., tone versus voice), or extended experience with the warning system itself. The purpose of this study is a direct comparison of drivers' subjective annoyance towards inappropriate alarms as a function of rate of occurrence and the type of signal generated in naturalistic, on-road driving conditions. Test equipment to generate and present signals, and to collect driver response was installed in fifteen participants' personal vehicles for a nine-week period. Signals were presented at random times while the participants engaged in their normal, daily driving routines. In order to simulate future operating conditions where actual alarm warnings will require the driver's attention and reaction, "appropriate" alarms to which the driver had to make a simple motor response, and "inappropriate" alarms to which the driver did not have to make any response, were presented. Inappropriate tonal alarms were presented at four different frequencies of occurrence, including averages of four per hour, one per hour, one per four hours, and one per eight hours of driving time. In addition, a voice warning condition was included, at a rate averaging one per hour. Participants made daily and weekly ratings of the degree of annoyance that resulted from the nuisance alarm schedule. The 4/hour-tone and the 1/hour-voice were significantly more annoying, and less acceptable, than the other conditions. Participants showed a wide range of annoyance sensitivity, but the two most annoying conditions appear to be unacceptable, while the less frequent rates do appear potentially reasonable for functional systems.

INAPPROPRIATE ALARM RATES AND DRIVER ANNOYANCE

1.0 Introduction

Vehicles of the near future will be equipped with both informational and crash avoidance warning systems to assist the driver with safer and improved vehicle operations. These systems may convey certain messages to the driver with an auditory component for the purposes of alerting the driver or to deliver a warning alarm signal. In the case of crash avoidance or extra-vehicular object detection systems, “false alarms” will inevitably occur from time to time. While the term “false alarm” implies a hardware failure of some type, it has also been used to describe situations where the algorithms have inappropriately identified a non-threatening object as a hazard, thereby creating a “false” or inappropriate alarm. Included in this broad class of alarms are nuisance warnings, premature warnings, warnings triggered by weather conditions or other inappropriate environmental characteristics, and other undesirable or non-useful warnings. For example, a headway collision avoidance system may find an obstacle in the projected path of travel directly ahead of the vehicle, but the road curves and the obstacle is on the far side of the lane; or an obstacle has been detected to which the driver has already begun to respond.

Design recommendations for crash avoidance warnings, as well as more general human factors guidelines, generally caution designers to “minimize” the number of false alarms. Excessive inappropriate alarms result in two problems: a) they are annoying, resulting in poor product acceptance, reluctance to use the product, or intentional defeating of the system; and b) they degrade the user’s response, resulting in slower and less reliable reactions to valid warnings. Although the potential problem of diminished user response to inappropriate alarms is well recognized, the general admonition to “minimize” inappropriate alarms is not very useful. The designer must make trade-offs between sensitivity and speed of detection for potentially dangerous events and the likelihood of an inappropriate (false) warning. Resolution of this trade-off is not easy because in many situations, the consequences of a missed real signal may be more costly than the operator’s reactions to false alarms. Depending on the criticality of missed real alarms, designers may tend toward a liberal detection criterion for alarm systems.

The concern regarding operator responses to inappropriate alarms has been discussed in a variety of operational settings, for example, power plant control rooms, aircraft cockpits, railroad locomotive cabs, and medical situations, to name a few. In the vehicle crash avoidance warning area, the question has sometimes been cast as “given that there may be some inappropriate alarms, how many inappropriate alarms are acceptable?” The state of knowledge right now does not even permit an order-of-magnitude estimate of what acceptable or unacceptable inappropriate alarm rates might be, or how that rate may vary with such factors as traffic conditions. The purpose of the following study is to provide an order-of-magnitude, or bracketing estimate of the acceptability, and to evaluate various situational factors for their influence on drivers’ subjective annoyance to inappropriate alarms. There are very serious methodological limitations in studying driver response to crash warning false alarms (discussed

below), and this study must be viewed as preliminary. However, it does provide the initial systematic data on actual on-road annoyance, and also explores new methodologies for addressing the problem.

1.1 Determinants of Annoyance Response to Inappropriate Alarms

As an initial step in assessing annoyance due to inappropriate alarms, it is necessary to consider what factors contribute to the annoyance response. The physical attributes of the stimulus (e.g., the acoustic signal) are only one aspect of the problem. The actual annoyance that a given stimulus causes a motorist will also be determined by a variety of other factors. Research on driver response to inappropriate alarms should incorporate the range of relevant factors to whatever degree practical. The following factors are presumed to be general determinants of the annoyance response to an inappropriate alarm:

Physical Stimulus: This refers to the physical and temporal dimensions of the signal. For acoustic signals, this would include intensity (loudness), sound frequency spectrum, onset and offset dynamics (abruptness), and various possible changes over time (pulsing, swept signals, alternating signals, burst patterns, etc.).

Task Interference, Response Requirements: A meaningful signal requires some action on the part of the listener, and attending to the signal and executing the response will interfere with ongoing activity. Task disruption is itself annoying. Therefore annoyance will be related to the baseline task (e.g., some driving scenario) and what is required by the signal (e.g., immediate visual search, vehicle control actions).

Emotional Response to Meaning: A meaningful warning of an imminent dangerous event may generate emotional responses that may themselves be aversive, such as fear, startle, panic, etc.

Environment/Setting (Physical): The physical context in which a signal occurs may interact to influence the annoyance of the event. The context (environment, weather, roadway type, off roadway events, etc.) potentially could influence perception of the signal, driver activity upon which warning response requirements are imposed, and the status of the driver. Under this heading, one might also include physical and emotional state of the driver (fatigue, alertness, impairment, frustration, mood).

Social, Personal Setting: Much of the research on human response to alarms has been conducted for what might be considered “workplace” settings: control rooms, cockpits, decks, heavy vehicles, military applications, and so forth. Although formal comparisons have not been made, anecdotal experience suggests that people may be more annoyed by events that intrude into their own personal space and time, and upon their own chosen activities. This would certainly include the personal vehicle, and there may be various activities collateral to personal driving (e.g., use of the entertainment system, cellular

phone use, pleasure driving, reverie, conversation) that might be intruded upon. The presence of other people might also be a contributing factor, since there may also be concern for their disturbance, potential embarrassment, and interference with conversation.

Relation to Actions, External Events: Some inappropriate alarms may occur independently of what the listener is doing or what the surrounding situation is. However, often the inappropriate alarm will be correlated with other activities and events. For example, false positives or overly conservative triggering criteria might be related to the maneuver the driver is making or to the actions of surrounding vehicles. Some inappropriate alarms may be highly predictable, and even controllable, by the driver. While it seems highly likely that the relationship to actions and events may color the reaction to an inappropriate alarm, the nature of this influence is not known.

Permanence/Transience: The duration of exposure to an inappropriate alarm situation may lead to changes in the annoyance response over time. Furthermore, if the person experiencing inappropriate alarms perceives the situation to be very transient, greater tolerance might be anticipated.

Perceived Benefits (mitigator): A warning or alerting signal is presented because it has the potential to provide some benefit to the user. Even though a given presentation of the signal may be inappropriate, the user might perceive overall benefits for the warning system, which may influence the acceptability of inappropriate events. Perceived benefits would serve to mitigate the other negative effects of the warning.

1.2 Ability to Simulate Real-World Annoyance Factors in Experimental Studies

To study the annoyance generated by inappropriate alarms, a conceptually ideal study is straightforward: install a functioning collision warning device in a vehicle; program the device to generate various types of inappropriate alarms; and then measure annoyance subjectively (self report) or behaviorally (e.g., degree of system use) as a function of the nature and frequency of the inappropriate alarms. Unfortunately, such an ideal experiment is currently not technically feasible (given the performance of existing devices) and is certainly not ethically acceptable. Driver safety could be compromised by either the reaction to a (false) crash avoidance alert or by the degrading of the response to a legitimate warning. In practical terms, three general research methods appear to be more viable for addressing the problem: (a) “what-if” subjective judgments; (b) driving simulator experiments; and (c) in-vehicle exposure without collision avoidance aspects. None of these techniques is ideal. “What-if” procedures would ask the individual to make a judgment about how annoying an inappropriate alarm would be under various conditions. This method might be useful for comparing a range of alternative alarms with one another in terms of relative annoyance. However, there is no basis to believe that people could accurately project their actual degree of annoyance under different conditions. Driving simulators and in-vehicle exposure are each able to more broadly address the range of

factors that contribute to the annoyance response. Each method has different strengths and weaknesses, and they may be seen as complementary.

Table 1 summarizes the ability of driving simulators and in-vehicle exposure to accurately capture the effects of factors identified as important in the previous section. The table characterizes each method as good, fair, or poor for each factor; these qualitative judgments represent the opinion of the authors, based on the discussion that follows, but there is no quantitative basis.

Both in-vehicle studies and simulator studies can accurately reproduce the physical qualities of the signal, as well as the noise context in which it occurs. The in-vehicle method captures vehicle acoustics and the full range of noise conditions as actually experienced by the driver; simulator studies must define the conditions and simulate them with whatever fidelity is possible on the system. The simulator can impose driver response requirements reasonably similar to an actual collision avoidance warning. That is, the signal can require a rapid visual search, situation evaluation, and vehicle maneuver (assuming a valid alarm). For safety reasons, an in-vehicle system cannot require this, since it would be potentially dangerous in real traffic. However, it is possible to devise a required response to the signal that mimics many of the aspects of an actual crash avoidance warning (interruption of task, need for reasonably rapid visual search, decision regarding validity of signal, need for response to valid signal).

Table 1. Determinants of annoyance response to in-vehicle nuisance alarms.

Sources	Own Vehicle	Simulator
Physical Stimulus	Good	Good
Task Interference, Response Requirements	Good/Fair	Good
Emotional Response to Meaning	Poor/Fair	Fair? Poor?
Environmental/Setting (Physical)	Good	Fair
Social, Personal Setting	Good	Poor
Relation to Traffic, Maneuver	Poor	Good
Permanence/Transience	Fair	Poor
Perceived Benefits (Mitigator)	Fair	Fair

Neither simulators nor in-vehicle methods are able to capture the emotional response to an actual crash avoidance alarm. The threat of physical injury to oneself, passengers, or others will presumably contribute to the intensity of the event. There is debate within the highway research community as to the extent to which immersion in the simulator environment can lead to realistic emotional responses to threat when no possibility of harm is objectively present. The in-vehicle method cannot address physical threat at all; however, it can endow the signal with some emotional/motivational status, based on some reward/penalty structure.

If an in-vehicle method is used with the driver's own vehicle, then the study will perfectly capture the physical setting in which driving occurs. This includes environmental factors, roadway types, and so forth. The varying physical and emotional state of the driver might also be included here (fatigue, alertness, impairment, mood). The simulator can generate some or all of these scenarios to varying degrees depending on the sophistication of the simulator, but the distribution of settings will not be representative of any particular individual's driving.

One potentially important limitation to a simulator approach is that the intrusive alarms do not occur in the driver's own personal space. Paid individuals may be tolerant of inappropriate alarms for the duration of an experimental session in the experimenter's lab; it may be quite a different matter if the alarm occurs in one's own vehicle. Other passengers will be exposed as well, including small children who may be disturbed. Conversation, listening to the entertainment system, cellular phone use, or other real-life collateral activities might be disrupted. In-vehicle methods capture this part of actual experience.

The simulator is ideal for generating any desired traffic scenarios and can safely present signals with respect to what the driver and other vehicles are doing. Without a great deal of sensing capability and intelligence, in-vehicle systems could not do this and some situations may rarely occur in real driving.

The annoyance of repeated inappropriate alarms may change over time and tolerance for them may be different if the test exposure is of short duration. In-vehicle procedures can be left in place for extended periods of time as desired. Simulator exposure is more structured, restricted, and as an expensive resource (for high-grade simulators), will normally not be dedicated to long term exposure for a single individual.

The annoyance response to an inappropriate signal may be mitigated by the perceived benefits associated with the system. The extent to which people perceive a benefit from a real crash avoidance warning device could vary greatly depending on the device, so there is no objectively "correct" level to be simulated. Simulators can perform in a manner comparable to some chosen device, and the individual may perceive some benefit, but there is no actual benefit since there is no objective danger. In-vehicle systems (as well as simulator methods) can endow a signal with perceived benefits through the use of monetary incentives.

Having considered alternative procedures for addressing the inappropriate alarm issue in vehicular operations, it is apparent that no practical research methodology provides a close

simulation of the actual process across all of the relevant factors. Since user annoyance has proven to be a critical aspect of system success or failure in other warning application settings (i.e., cockpits, control rooms), it would be a mistake to conclude that the absence of a perfect technique should preclude research for driving situations. Available methods can be helpful in scoping the magnitude of the problem and the critical ranges for acceptability. Ultimately, a combination of in-vehicle and simulator methods might help yield a reasonable picture. For purposes of the present research, this study has chosen to employ in-vehicle methods. The rationale is that such a study can provide a conservative estimate of the types of rates that may be tolerable by the public. Since the alarms will not generate the urgent, emotionally laden response of actual crash warnings, and although installed for a relatively long time will still not be permanent, it may be assumed that there will be some underestimation of actual annoyance. Thus if a given condition of inappropriate alarm occurrence is found to be unacceptable under these conditions, it will help indicate preliminary performance levels that should not be exceeded.

The purpose of the following study was to compare the level of driver annoyance generated by various rates and conditions of inappropriate alarms. The study exposed drivers to a variety of different rates of occurrence of inappropriate alarms in the course of their normal driving, over an extended period of time. Although the alarms were not safety related, they did require a visual response and confirmation process as a real warning might. Occasional “valid” trials of alarms occurred amid some programmed rate of random inappropriate alarms. (In this experiment, a “valid” alarm indicated a situation in which the participant had to make a particular response in order to earn a monetary bonus). The study was designed to compare four different rates of inappropriate alarms as well as comparing acoustic tone and voice alarms.

2.0 Method

Driver annoyance to four different rates of occurrence of inappropriate alarms was tested over a nine week period. Experimental equipment was installed in each participant's own personal vehicle and left in place for the duration of the nine week period while the participants continued to drive in their customary driving routines. At random times while driving, warning alarms sounded; the rate at which inappropriate alarms occurred varied on a weekly basis. Subjective ratings of annoyance and noticeability toward the frequency of the alarms were made once a day and at the end of each week.

2.1 Participants

Fifteen drivers participated in this study, 8 females and 7 males. Their ages ranged from 26 to 56 years. Participants were sought who ordinarily drove their vehicles at least eight hours per week. This minimum level of driving would allow participants to adequately experience each inappropriate alarm rate. Participants were informed that the study would last nine weeks, during which time equipment would be installed in their personal car. All participants were paid for their participation on a base pay plus bonus system; they received a fixed weekly payment for allowing the test equipment to be in their vehicle, and for responding to daily and weekly questionnaires. They also had the opportunity to earn bonus payment for correctly performing the experimental task while driving.

2.2 Apparatus

The entire experimental control and data collection system was housed within the participant's personal vehicle. Computer logic, video recording equipment, and a power supply to operate the equipment were located out of sight in the trunk or cargo area of the vehicle. Input and output devices (driver response button, miniature camera, alarm indicator light, and an acoustic signal generator) were located in appropriate locations in the passenger compartment of the vehicle's interior. The full system is described below.

A Z-World miniature controller with 64 kilobytes of erasable/programmable read-only memory (EPROM) controlled the timing of alarm rate presentation as well as the operations of the other experimental equipment. The software program to manage the operations of the equipment was developed in the Dynamic C programming language and permanently stored on the EPROM chip installed in the miniature controller. The controller recorded in random access memory (RAM)

¹ Participants were primarily the sole driver of their vehicle. Because the participants were actually driving their vehicles, the terms "participant" and "driver" are used interchangeably throughout this report.

the time of day that an alarm signal occurred, the type (appropriate or inappropriate) of alarm, and the driver's response (e.g., "hit" or acknowledgment of an appropriate alarm, "miss" of an appropriate alarm, "false hit" to an inappropriate alarm) to every alarm signal presented to the driver.

During each occurrence of an alarm signal, a video cassette recorder (VCR) recorded video data 25 seconds prior to, 10 seconds during, and 15 seconds after the presentation of the alarm. A clock with voice output marked the current time on the VCR tape segment during each alarm. This audio time stamp was later used to link the video segment data with driver response data stored on the controller. The miniature controller, VCR, and alarm clock were housed in an attache style case situated in the trunk or cargo area of each participant's vehicle. This case was secured with tie-down straps to prevent bumping or movement from vehicle motion. The case was "locked" with a plastic cable to discourage anyone but the technician from accessing the equipment. A typical setup of a vehicle trunk with the test equipment is illustrated in Figure 1.



Figure 1. The controller housed in a plastic case along with the VCR, relay board, and voice clock.

A 12-volt sealed rechargeable battery powered all equipment used for the study. An accessory fuse from the inside of the vehicle was tapped and wired to an electro-mechanical relay which triggered the miniature controller. Each time the vehicle was turned on, the controller was activated, eliminating the need for the participant to turn the test equipment on or off; activation of the test equipment was linked to the operation of the vehicle. All wiring was inconspicuously routed from the input/output device to the case in the trunk. A schematic electrical diagram is provided in Figure 2.

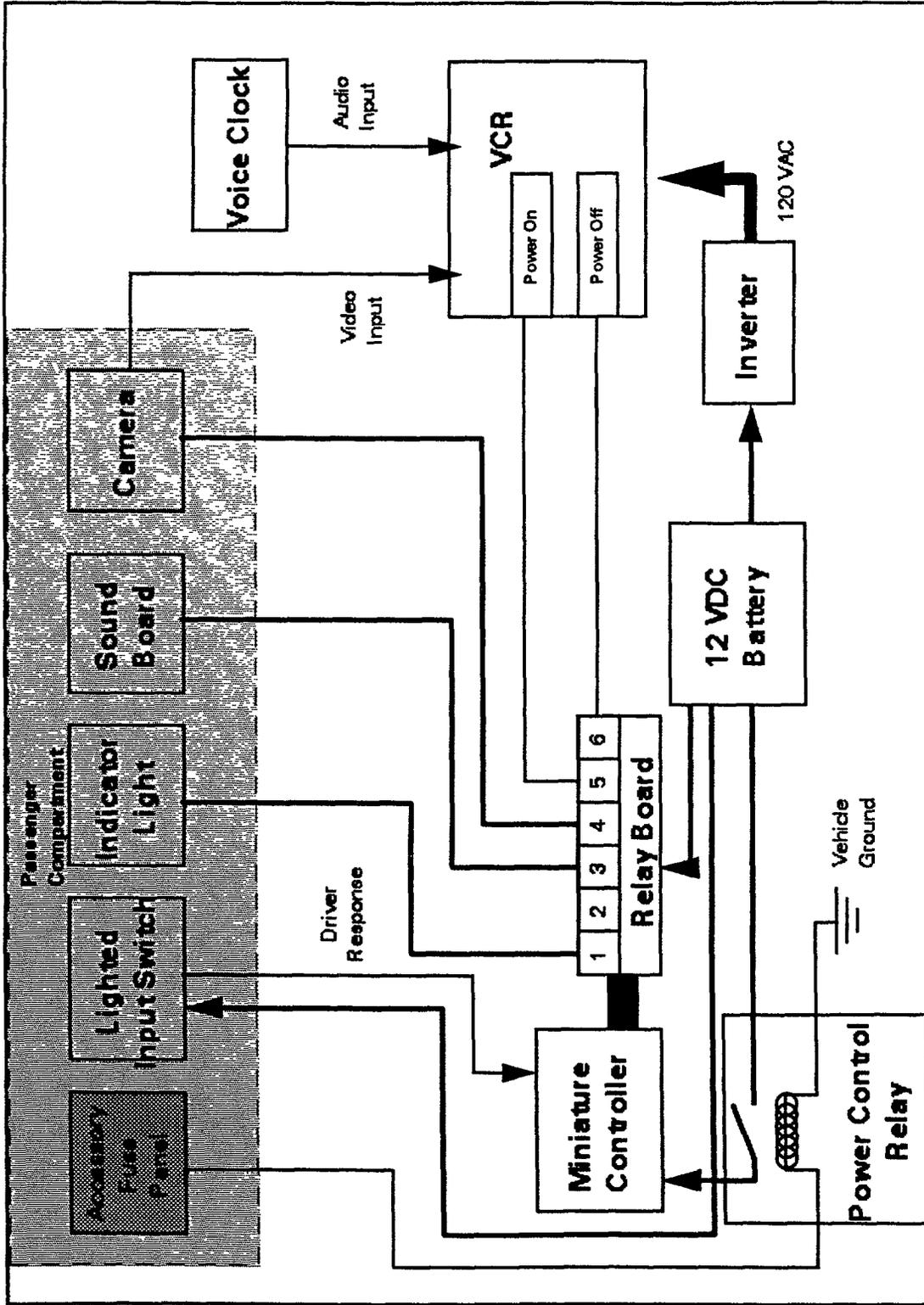


Figure 2. Schematic wiring diagram for in-vehicle alarm presentation and data collection.

In the passenger compartment of each vehicle, a blue indicator light was mounted on the passenger's side A-pillar to differentiate "real," or appropriate from "false," or inappropriate alarms. A second blue, back-lit response button was mounted in a convenient and accessible location for the driver; the exact location varied according to the design requirements of each vehicle. Typically it was near the driver's right leg, within easy reach. Participants pressed this button each time a "real" alarm was seen.

A sound output box containing a miniature amplifier, 3.5 inch speaker, and a digital sound chip was mounted underneath the dashboard on the passenger side. Two kinds of digital sound chips were available for installation depending on the experimental condition: one chip featured a rapidly beeping tone (low-fuel aircraft warning), and the other played a digitized male voice that repeated the words "check light. " The tone stimulus was tested in a previous study (COMSIS, 1995) and was selected for use in this study based on the premise that it was a likely candidate for actual in-vehicle warning devices. A voice stimulus similar to the "check light" voice was also pretested in the prior referenced study, with the exception that the original recorded message was "danger."

The primary volume adjustment of the sound output level from the box was made when the digital sound chips were recorded with the audio stimulus. All chips had the same digital pattern "burned in" to 75 dB(A) during the preparation of the test equipment. When the equipment was installed in the participant's vehicle, the sound output level was set to 75 dB(A) with an audiometer that was held at approximately the driver's ear level when seated in the front seat. Fine-tune volume adjustments were made at installation and during the weekly site visits to ensure proper volume levels.

A miniature video camera with a 92 degree field of view was used to record traffic conditions for a fifty second interval surrounding both real and inappropriate alarms. To reduce the threat of theft, the video camera was hidden on the bottom surface of an ordinary in-vehicle notepad mounted on the lower center windshield with suction cups. A schematic drawing of the passenger compartment equipment layout is provided in Figure 3.

2.3 Procedure

Participants maintained their normal daily driving routine for a nine week period. At random times during the course of driving, auditory alarms sounded from the test equipment. Since the timing of the alarms was random and in no way related to any outside traffic events, a protocol was devised to simulate driver response aspects likely of future crash avoidance warning systems. In future functional systems of this kind, drivers must respond to all signals to detect whether an alarm is valid. All alarms will initially require that the driver engage in a rapid visual search to confirm and locate the potential threat that the warning system has detected.

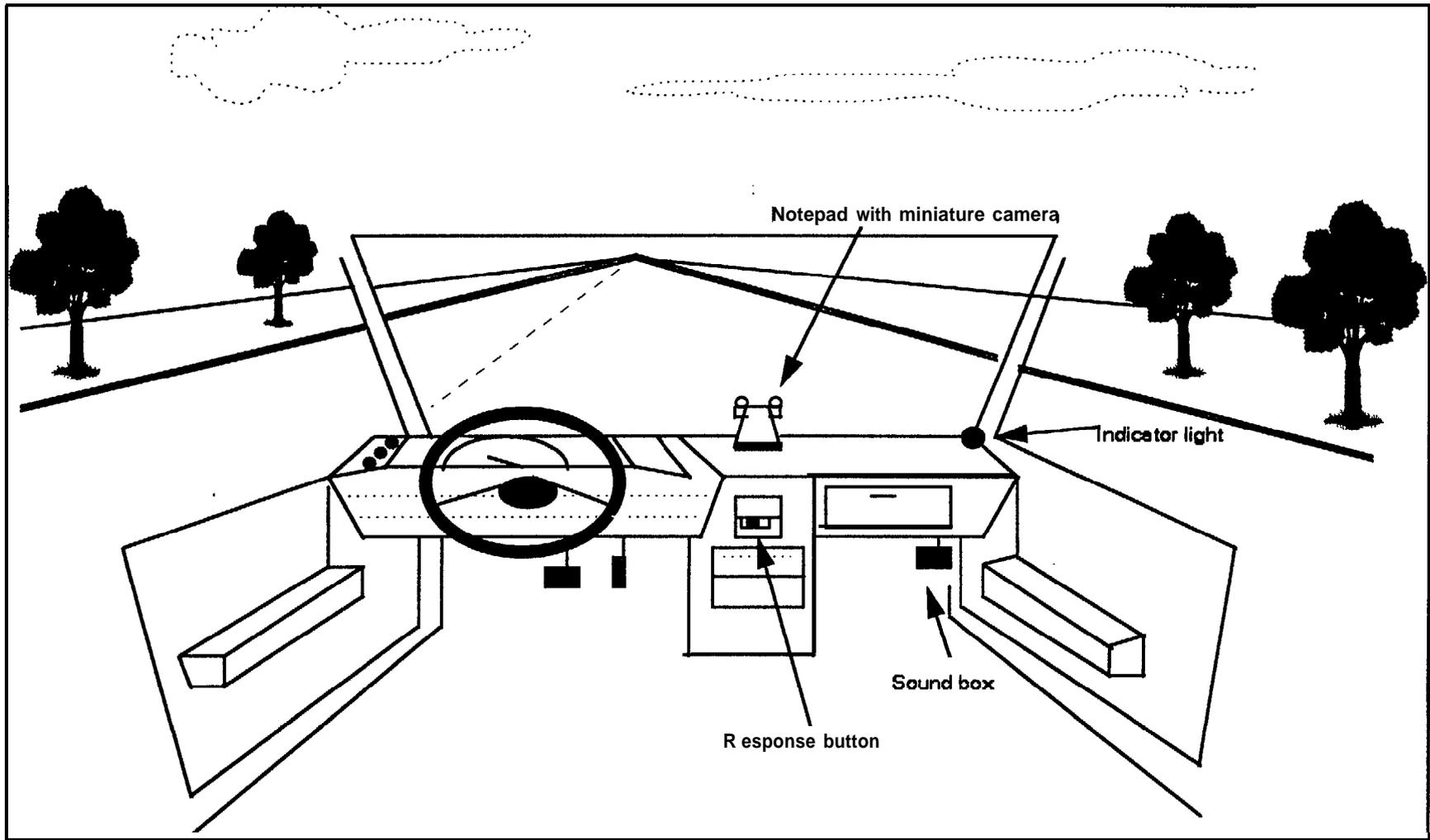


Figure 3. The sound box, blinking indicator light, lighted response button, and miniature camera (beneath notepad) were installed in the passenger compartment.

In the event of an actual threat, some type of overt response is required, such as an evasive driving maneuver. An inappropriate alarm would not require any response on the part of the driver; in fact, an overt response might be dangerous under some conditions. This processing sequence must happen in a rapid time frame. The procedure used in this experiment was designed to include this need to disrupt normal visual scanning and required time-restricted search for the sources of the warning, along with an appropriate rapid decision to respond or not respond.

Real and inappropriate alarms also were used in the experimental protocol to maintain driver motivation and attention to the task. In a fully functioning system, there is inherent reward for detecting real alarms, since by doing so a potential traffic conflict may be avoided. Since there was no inherent reward in the non-functioning test system, a monetary bonus payment scheme was used. All real alarms that were correctly detected and responded to in a timely fashion resulted in a bonus payment of \$4.00 to the driver. Inappropriate alarms required no response on the part of the driver; to deter drivers from indiscriminately responding to all alarms, a penalty of the loss of \$1.00 off any bonus payment was incurred for responding to an inappropriate alarm.

All alarms delivered by the test equipment consisted in part of the audio stimulus. The distinction between “real” and “inappropriate” alarms was defined as: a) “real” alarms occurred whenever the audio stimulus was accompanied by a blinking sequence of the small light situated at the passenger’s side A pillar; b) “inappropriate” alarms were defined as the audio stimulus only and the light at the A pillar remained off. Checking the experimental light situated at the passenger’s side A pillar simulated the physical confirmation response that a functional crash avoidance system is likely to elicit. The brightness of the lamp was such that during daylight operation it was not highly conspicuous, so that directed search was required. For a real alarm sequence, drivers were to press a response button either during the ten second period while the A pillar light blinked on and off, or within ten seconds after the light extinguished. The latter period allowed drivers to safely complete any complex driving maneuvers or situations they might be in when a real alarm sounded, while still maintaining an opportunity to indicate the detection of a real alarm. Inappropriate alarms required PZO response to the stimulus. Drivers in this study had ten seconds in which to check and confirm whether an alarm was a valid one. Actual crash avoidance warnings may have greater urgency and require a more rapid response than this, but the time constraint represented a reasonable compromise between response urgency and driver safety considerations for the purposes of this study.

The frequency of occurrence of inappropriate alarms and alarm type (tone or voice) defined the experimental conditions for this study. Across all experimental conditions, three real alarms were presented at random times during the first eight hours of driving every week. Five inappropriate alarm conditions were superimposed upon this, as follows:

- 1) four inappropriate alarms per hour of driving, one-week duration, tone stimulus;
- 2) one inappropriate alarm per hour of driving, one week duration, tone stimulus;
- 3) one inappropriate alarm per hour of driving, one week duration, voice stimulus;

- 4) one inappropriate alarm per four hours of driving, two week duration, tone stimulus;
- 5) one inappropriate alarm per eight hours of driving, three week duration, tone stimulus.

To reduce novelty or practice effects, condition 2 was presented during the first week for all participants, and then was repeated at a later time. Including this repeated condition, the combined durations of all conditions totaled nine weeks. Some conditions lasted longer than one week due to infrequent inappropriate alarm rates. The two- and three-week durations of conditions 4 and 5 were devised so that the participant would get a good subjective sense for the infrequent presentation rate over time. The order in which the participants experienced each experimental condition was counterbalanced using a Latin square design to control for learning, practice, or other potentially biasing effects; participants were randomly assigned to presentation orders.

The initial experimental session was devoted to installation of the equipment in the participant's vehicle. This procedure lasted between 1.5 and 2 hours, depending on the make and model of the vehicle. While the equipment was being installed by a technician, a second researcher explained the objectives of the study to the participant. Participants were told that the purpose of the study was to assess the effectiveness of a signal device for capturing drivers' attention. It was explained that similar devices could be used in the future for actual hazard warning and information service purposes. However, it was stressed that the test equipment was not a functioning system and that alarms would be presented at random intervals whenever the vehicle was turned on. Informed consent was obtained, and the purpose and function of each piece of equipment was explained. The experimenter demonstrated real and inappropriate alarms, and ensured that the participant understood the meaning and task associated with each kind of alarm presentation.

The daily questionnaire was explained, with special emphasis that participants should complete it each day rather than waiting until the end of the week. Participants were instructed to drive in their usual manner and not to make any special trips for the sake of the experiment. Participants were told that it would be their responsibility to warn anyone else who might drive their car that test equipment was installed. Other drivers were to be instructed to ignore all alarms and not to press the button. A small placard was placed on the bottom of the notepad that held the miniature camera to remind other drivers of the presence of test equipment. Space was provided in the daily questionnaire to record whether another driver had operated the car each day.

At the end of each day, participants completed a daily driving questionnaire regarding that day's driving experiences as well as their subjective impressions of the alarms they heard that day from the experimental equipment. Multiple copies of the questionnaire were combined into a small booklet that participants could keep inside the vehicle. Each page was predated to encourage participants to fill out the questionnaires on a daily basis. Responses to a question regarding the noticeability of the sounds were recorded on a nine-point Likert scale with the

anchors “very noticeable” and “barely noticeable” for the endpoints of the scale and an anchor of “somewhat” for the midpoint. Participants reported how annoying the sounds were by using a second nine-point Likert scale anchored by “not annoying at all” and “extremely annoying,” with a midpoint of “tolerable.” Participants were also asked to indicate the kinds of trips they made that day, e.g., “running household errands, ” and the kinds of activities they engaged in while driving, e.g., “talking to a passenger.” If a participant did not drive on any particular day, or if they did not hear any signals from the test equipment for an entire day, they were not required to complete a questionnaire for that day. A sample daily driver questionnaire form is provided in Appendix A.

During the weekly visit by an experimenter, participants also completed a weekly questionnaire. Identical questions to those in the daily driver questionnaire were asked concerning how noticeable and annoying the sounds were. However, these questions were worded so that the participant was asked to make an estimate with regard to the entire week. The participant also rated the acceptability of the inappropriate alarm rate they experienced during the week using a Likert scale anchored by “unacceptable” and “highly acceptable,” with a midpoint of “can’t tell.” Participants compared the annoyance of sounds that occurred in various driving situations, e.g., “driving by myself” compared to “driving with other passengers.” The weekly questionnaire is presented in Appendix B.

Other activities that occurred during the weekly visit included the downloading of data from the controller, swapping of videotapes in the VCR, and checking all equipment for proper functioning. Participants were paid their base rate plus any bonuses earned during the past week. Occasionally, a condition was repeated due to equipment failure. During the final visit, the test equipment was removed. Participants were fully debriefed on the nature of the study. A semi-structured interview was followed to gauge changes in feelings about the alarms over nine weeks and to evaluate the longitudinal study design.

Video data recorded during the week was manually coded for traffic and environmental conditions present during the occurrence of an alarm. A fifty-second video clip was recorded each time a real or inappropriate alarm was presented to the participant. The video segments of the traffic conditions recorded during the presentation of all alarms were coded by a trained rater according to pre-defined categories. These categories covered items such as roadway factors (e.g., local, arterial, freeway, etc.), outside lighting conditions (light or dark), visibility (clear, precipitation, other limited visibility), roadway surface conditions (wet or dry), general location (e.g., suburban residence, rural, etc.), traffic density (e.g., free flow, stable, interactions, high density, etc.), and vehicle maneuvers at the time of the alarm (e.g., forward or backward motion, left or right turn, at intersection, etc.). Roadway design standards (American Association of State Highway and Transportation Officials, 1990) were used as an initial basis to classify roadway factors (i.e., level-of-service) and traffic density; categories were combined and refined by the researchers to more adequately represent actual video footage gathered in this study. The coding scheme is shown in Appendix C.

3.0 Results

3.1 Analytic Method

Participants provided daily ratings of noticeability and annoyance, and weekly ratings of noticeability, annoyance, and acceptability. These were analyzed using analysis of variance (ANOVA) methods. All of the ANOVAs were two-way designs, with the alarm condition (frequency and type of alarm) as a within-subjects factor, and participant gender as a between subjects factor. For the analysis of weekly ratings, each participant's rating for a given alarm condition was represented by a single number: that is, it was the single weekly rating for those conditions that were in effect for a single week, and an average rating of the two or three weeks for those conditions that were in effect for more than one week. For ANOVA's based on daily ratings, the individual daily ratings for each day were retained as repeated measures for that participant. Where appropriate, post-hoc pair-wise comparisons between alarm conditions were made using Tukey's studentized range test. All of the ANOVAs used local error terms for computing the F-ratio, which is more conservative than using a global error term. As a result, a number of analyses were reported as non-significant, which would have been statistically discriminable with a less conservative procedure.

In none of the ANOVAs was gender, or the interaction of gender with the alarm condition, found to be statistically significant. Therefore gender effects are not further discussed.

In all of the ANOVAs, and in the tables and figures that follow, there were six levels of alarm condition considered in the ANOVA. The initial week, labeled "practice," presented inappropriate acoustic alarms at the average rate of one per hour. This was treated as a separate condition from the "1/hour" condition that occurred later within the random sequence of conditions. Based on limited pilot data collected in preparation for this study, an effect of novelty was anticipated. While there may have been a slight indication of this in the experiment (ratings for the practice condition were a little lower for annoyance and a little higher for acceptability), there was no statistically significant difference between the practice and 1/hour conditions. ANOVAs were also run with "week" as a variable to determine whether there were any meaningful trends over the nine weeks. These analyses were not significant either, nor was there any apparent trend in inspection of the data. Changes in ratings over time are therefore not discussed further in this section.

3.2 Noticeability Ratings

There was no statistically significant effect of the alarm condition on noticeability ratings, for either daily or weekly analyses. The average noticeability ratings for each inappropriate alarm condition for the weekly data are shown graphically in Figure 4. Because all alarms occurred at the same loudness level, regardless of sound type or frequency of occurrence, the absence of differences in noticeability ratings is not surprising. The overall mean ratings were very similar for the daily and weekly measures, averaging 6.7 and 6.6 (on the g-point scale), respectively. This confirms that the sounds were of sufficient loudness to be readily noticed by the participants

under their typical driving conditions. However, it should also be noted that participant comments, from the questionnaires or from the debriefing procedure, indicated that a number of participants found that there were occasional circumstances during which the alarm was more difficult to notice. For example, these included radio use, wind or vibration noise (e.g., lowered windows, bike rack on roof), conversation, and environmental noise (e.g., airplane, freeway noise).

3.3 Annoyance Ratings

Average annoyance ratings generally tended to increase as the frequency of inappropriate alarms increased. Means for the daily ratings are listed in Table 2. Mean ratings are shown graphically in Figure 5 for the weekly ratings and in Figure 6 for the daily ratings. It may also be seen in these figures that the 1/hour-voice alarm had a higher annoyance rating than the 1/hour-tone alarm. The alarm condition had a statistically significant main effect ($p = 0.001$) on the daily annoyance ratings. Although quite similar, the weekly ratings (which did not have repeated measures for each participant) were not statistically significant ($p = 0.12$; although the main effect was statistically significant using a less conservative statistical model using a global error term).

Table 2 presents the results of post-hoc pair-wise comparisons among alarm conditions. The mean annoyance ratings (g-point scale) ranged from 2.33 for the practice condition, up to 3.85 for the 4/hour condition. The 4/hour condition was rated significantly more annoying than all other conditions, with the exception of the 1/hour-voice condition. The voice condition also differed from all conditions other than 4/hour, including the identical frequency of acoustic alarm occurrence (1/hour-tone). The greater relative annoyance of the voice alarm was also reflected in comments volunteered by the participants.

The group mean annoyance ratings appear to be relatively low, the most annoying being 3.85 on a 9-point scale. This should not be taken to mean that annoyance was minimal. One issue concerns how participants used the rating scale. However, it is also the case that the group mean rating is not necessarily typical of individual participants. Some participants were generally unresponsive to the alarms, giving a rating of "1" nearly every day. One-third of the participants had overall mean ratings (across all alarm conditions) of 2.0 or less. This subgroup of "non-responders" had the effect of lowering the overall group mean rating and reducing differences between alarm conditions. It is not known whether their low ratings reflect a genuine indifference to the alarms, or suggest a different interpretation of the subjective meaning of the rating scale values. Across all participants and alarm conditions, a little over half of all ratings (54%) were a "1" or "2." Nonetheless, a substantial number of annoyance ratings were quite high. Figure 7 shows the percentage of all daily annoyance ratings that were "5" or greater; Figure 8 shows the same information for the weekly ratings. As the figures show, 40- 50% of the ratings for the more annoying conditions exceeded "5," and even for the least annoying, about 15% of daily ratings and 25-30% of weekly ratings exceeded "5." While no one used the most extreme rating of "9," many ratings were well beyond the point of the scale labeled "tolerable" (rating of 5).

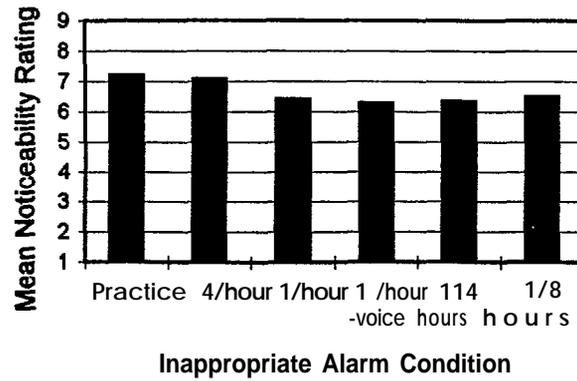


Figure 4 Average weekly noticeability ratings for each inappropriate alarm frequency rate condition.

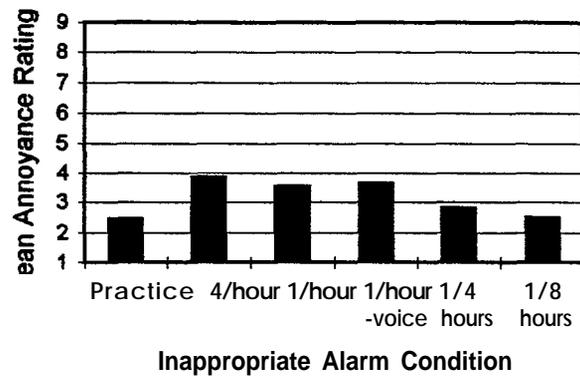


Figure 5 Average weekly annoyance ratings for each inappropriate alarm frequency rate condition.

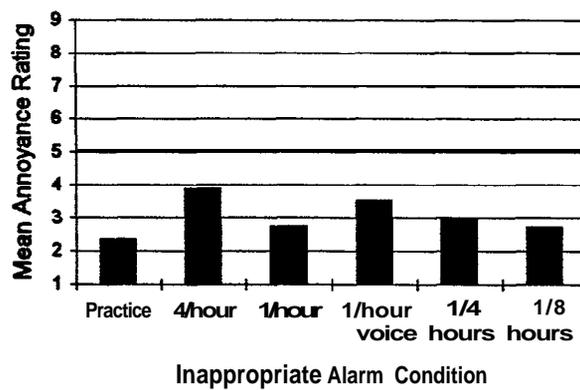


Figure 6 Average daily annoyance ratings for each inappropriate alarm frequency condition.

Table 2. Average daily annoyance ratings for each inappropriate alarm rate frequency condition, and pair-wise comparisons between conditions.

Condition Number	Condition	Average Annoyance Rating	Post-Hoc Pair-wise Comparisons				
			1	2	3	4	5
0	Practice	2.33	*		*	*	
1	4 / Hour	3.85		*		*	*
2	1 / Hour - Tone	2.70			*		
3	1 / Hour - Voice	3.49				*	*
4	1 / 4 Hours	2.96					
5	1 / 8 Hours	2.69					

A "*" denotes a statistically significant difference between conditions at the .01 level of significance, using Tukey's studentized range test for the post-hoc comparison tests.

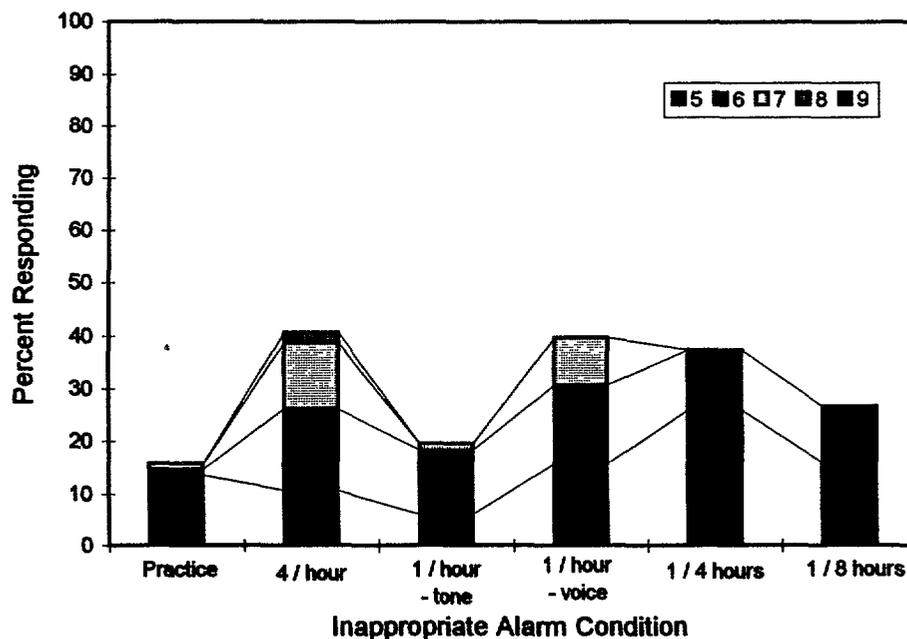


Figure 7 Percent of all daily subjective annoyance ratings that were 5 or greater for all inappropriate alarm rate conditions.

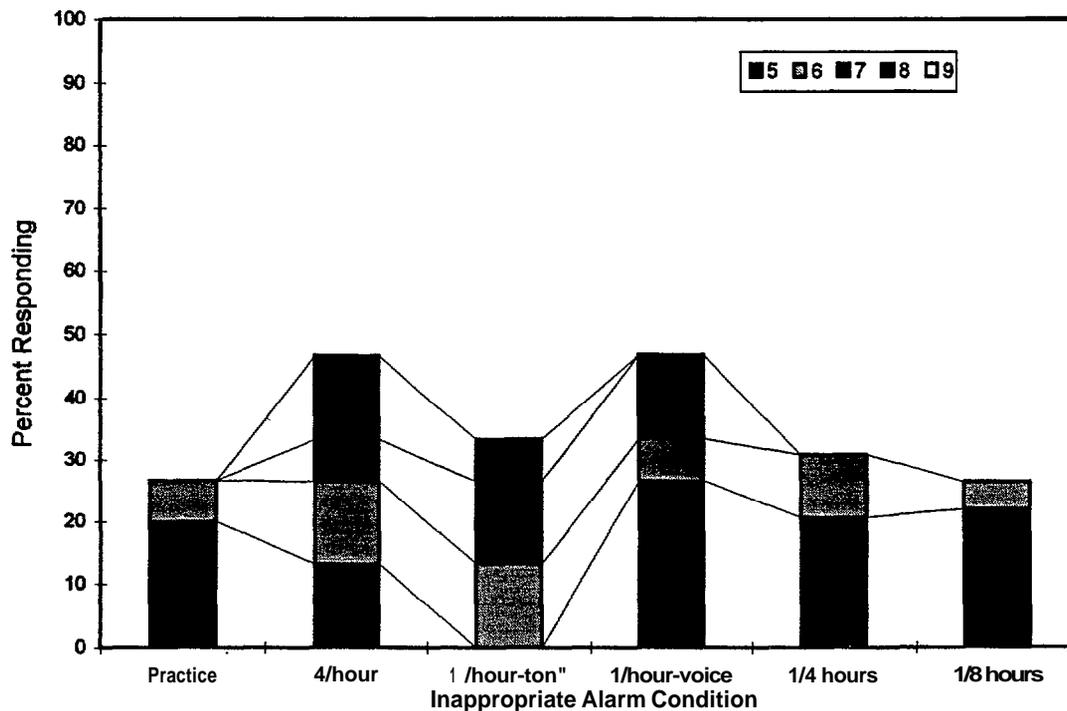


Figure 8 Percent of all weekly subjective annoyance ratings that were 5 or greater for all inappropriate alarm rate conditions.

3.4 Acceptability Rating

On a weekly basis only, participants rated how acceptable a “smart car” warning system would be if it had the false alarm rate just experienced. Mean acceptability ratings for each alarm condition are shown graphically in Figure 9 and are listed in Table 3. The ANOVA main effect of alarm condition was statistically significant ($p < 0.01$). As expected, these acceptability judgments were generally the inverse of the annoyance ratings, and in fact, there was a statistically significant correlation between them ($r = -0.39$, $p < 0.001$). On the bipolar rating scale running from “unacceptable” (1) to “highly acceptable” (9), only the two most annoying conditions (4/hour and 1/hour-voice) were rated below the mid-point of the scale. The other conditions ranged from 5.87 to 6.79. Post-hoc pair-wise comparisons indicated that the 4/hour and 1/hour-voice were significantly different from the 1/4 hours and 1/8 hours conditions. Although once again the 1/hour-voice appeared considerably less acceptable than the same frequency of the tonal signal, this failed to reach statistical significance.

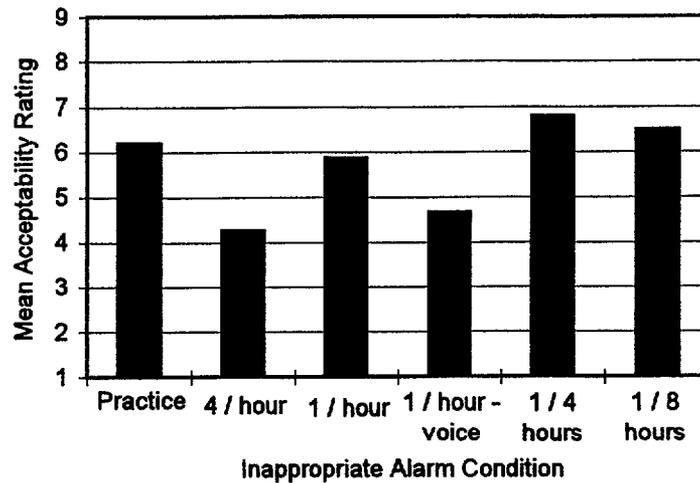


Figure 9 Average weekly acceptability ratings for each inappropriate alarm frequency rate condition.

Table 3. Average weekly acceptability ratings for each inappropriate alarm rate frequency condition, and pair-wise comparisons between conditions.

Condition Number	Condition	Average Acceptability Rating	Post-Hoc Pair-wise Comparisons				
			1	2	3	4	5
0	Practice	6.20					
1	4 / Hour	4.27	*			*	*
2	1 / Hour - Tone	5.87	*	*			
3	1 / Hour - Voice	4.67	*	*	*	*	*
4	1 / 4 Hours	6.79	*	*	*	*	
5	1 / 8 Hours	6.49	*	*	*	*	*

A "***" denotes a statistically significant difference between conditions at the .01 level of significance, using Tukey's studentized range test for the post-hoc comparison tests.

3.5 Factors Influencing Annoyance

Several approaches were taken to help identify some of the factors that might contribute to the annoyance associated with any given nuisance alarm. These included: (1) analysis of the videotaped records that accompanied each alarm occurrence; (2) ratings for specific factors listed in the weekly rating form (see Appendix B); and (3) any comments made on the daily or weekly forms or in the course of the debriefing interview.

Analysis of the video records did not prove helpful in clarifying contributing factors. The original analysis plan called for attempting to develop a linear regression model, using situational factors from the video as predictor variables, which would account for the day-to-day variability in ratings. However, since there was very little variability from day-to-day, such a formal statistical approach was not feasible. Therefore an informal procedure was implemented, identifying “outlier” data points for individual participants and then attempting to relate these to elements of the video record. However, no obvious factors emerged from this. The video records, together with information from the daily rating forms, did at least prove useful as a means of describing the typical types of driving in which the participants were engaged while taking part in this experiment. Trips made by the group of participants occurred primarily on arterial roads (24.3%), freeways (20.4%) or divided multilane highways (15.8%). Participants drove in either rural (21.4%), suburban business (20.8%) or suburban residence (18.9%) locations. Conditions were usually dry, clear and during daylight hours. Some trips were taken during nighttime conditions but the video quality was poor and traffic conditions for that segment could not be properly coded (however, because the data collection portion of this study occurred during the summer months, many evening driving trips were easily coded because of longer daylight conditions). The majority of traffic density conditions were classified as either free flow (17.7%), stable flow (36.5%), or slight interactions (19.3%). Most participants were not in the process of a vehicle maneuver at the time of the alarm.

When noting their subjective annoyance ratings in the daily questionnaire, participants also noted the types of trips they made, what, if anything, they were doing in addition to driving at the time of alarms, and whether another driver drove the car during that day (see Appendix A). Participants reported driving their cars 90% of the days of the study. They reported not driving 7% of the days, and they did not report whether or not they drove on 3 96 of the days. These figures suggest frequent driving activity and a high degree of compliance with the instruction to use the daily question. Participants reported that another person drove the car on 5% of the reporting period, no other drivers drove the car on 93% of the days, and no data was reported regarding other drivers on 2% of the days. These percentages indicate that the vast majority of alarm exposure was experienced by the intended participant. On each day that a participant drove and an alarm occurred (65 96 of all days), the participant recorded information about the kinds of trips made that day. Commuting trips were made on 65% of such days, running household or other errands were made on 45 % of days, pleasure/leisure trips were made on 43% of days, and driving as part of the participant’s job occurred on 2% of days. On days that alarms were presented, participants also indicated the

secondary activities in which they were engaged in addition to driving. Activities of the driver while making these trips included listening to the radio or car stereo on 94% of such days; talking to another passenger on 35% of the days; eating, drinking or smoking on 14% of the days; and using the car phone on 2% of the days.

The weekly questionnaire form (Appendix B) asked participants to compare the relative annoyance of alarms under various conditions: driving alone vs. with passengers; with the radio/stereo system on vs. off; in light traffic vs. in heavy traffic; on minor roads vs. freeways and major highways; during leisure driving vs. other trip purposes; and in daytime vs. during night. Each of these comparisons was made using a 9-point scale, where a rating of 5 indicated indifference. No clear contributing annoyance factors emerged from these ratings. There was little agreement among participants, with ratings generally spanning both ends of the scale. The most extreme mean difference from 5.0 was for the stereo on vs. off comparison, and even here the mean rating was 4.45. Thus no consensus emerged regarding any of these factors as major contributors to annoyance, even though individual participants may have rated some as important.

During exit interviews, eight of the 15 participants mentioned that the voice alarm was more annoying, confirming the rating data. A few expressed very strong feelings about the voice. A number indicated that the alarms (especially those other than the voice) were not particularly unacceptable, or that they became routine. Two other categories of contributing factors emerged from the comments. One category was social aspects. Although the weekly ratings did not indicate that participants judged the sounds more annoying when there were passengers in the car, a few comments suggested that the influence on other people might be negative. This included other passengers mimicking the voice, as well as eliciting strange reactions from adjacent drivers when the windows were down. The other set of comments had to do with the driver's state. There were several comments that indicated the sounds were more annoying when the driver was sick or having a "bad day." While these various comments provide some insight, there was no factor, related to the driver or environment, that clearly emerged as important to most participants.

4.0 Discussion

This study was able to measure changes in driver annoyance to various nuisance auditory alarm conditions for participants experiencing the alarms over extended time in their own vehicles. Annoyance and unacceptability were greatest for the most frequent alarm condition, averaging four tonal alarms per hour, and for the condition that used a digitized voice message (averaging once per hour). Clearly, alarm frequency and alarm type can have substantial influence on the annoyance response. However, for the tonal alarms, the major effect of alarm frequency appeared to occur between the conditions of 4/hour and 1/hour. The 1/hour rate did not differ significantly from the less dense alarm conditions (1/4 hours, 1/8 hours), although the means for the weekly ratings did show some trend toward this.

One of the findings of this experiment was the wide range of individual differences in participants' judgments of the annoyance of the alarms. It is possible that some of this may be due to differences in how participants interpreted the scales, and methodological refinements may improve this in future work. However, based on participant comments, as well as the ratings, it is certainly the case that some people experienced little annoyance to the alarms (except perhaps for the voice alarm), while others reported meaningful annoyance. This dichotomy means that it may be misleading to base conclusions on group averages alone. A more extensive experiment would be required to better estimate the proportion of drivers experiencing greater annoyance, and to better anchor the meaning of the ratings with respect to ultimate behavioral acceptance.

In interpreting the findings of this experiment, it must be recognized that there was no real traffic safety consequence to the signals. The procedure was designed to simulate, in real life settings over extended time, the perceptual and behavioral effects of nuisance alarms, and the reaction to the acoustic properties of the signal and to the interruption of driving, with demands for visual search and overt response. To the extent that true crash avoidance system alarms engender strong emotional reactions, the negative response to a nuisance alarm may be increased. Whether a driver would accept a given system would also depend upon perceived benefits. But considering only the negative aspects, i.e., the annoyance generated by the signals, the findings of this experiment may somewhat underestimate the annoyance of more emotionally-laden signals.

Based on these preliminary findings, some implications for warning system design may be offered, recognizing they must be tentative pending further research or experience. First, tonal signals of about 75 dB(A) should not occur substantially more often than once per hour of driving time. There were some questions, based on participant comments, about whether 75 dB(A) was always sufficiently loud; we do not know the acceptability of louder, or of driver adjustable, volumes. At 1/hour, a substantial number of weekly annoyance ratings still exceeded the "tolerable" rating point; given this, plus the fact that true safety-related nuisance alarms may be more annoying, it is probable that 1/hour itself may not be broadly acceptable. However, it is encouraging for intelligent-vehicle applications that the major effect of alarm frequency was seen between 4/hour and 1/hour. Not much influence was seen between 1/4 hours and 1/8

hours, suggesting that drivers may accept warning systems even if nuisance alarms are not extremely rare.

Another implication is that voice messages may lead to substantially greater problems with user acceptance than comparably loud tonal signals. Although only one rate of nuisance alarm (1/hour) was tested with the voice message, annoyance was significantly greater than for the comparable frequency of the tonal signal. In fact, annoyance to the voice alarm was roughly equivalent to the annoyance generated by the tonal signal when it occurred four times as often. Participant comments during the debriefing confirmed the frequent feeling that the voice signals were more annoying. The acceptable rate of voice false alarms apparently will be lower than that for other signals, and an average rate of one per hour is clearly not acceptable.

This experiment was designed to measure driver annoyance as a function of experience (i.e., lengthy exposure) with some nuisance alarm condition. Participants judged annoyance retrospectively, integrating their exposure over time. This appears to be a reasonable measure for relating the experiment to people's willingness to tolerate nuisance alarms in extended real-life driving. The method, therefore, was not particularly designed to reveal what situational factors contributed to the annoyance reaction to any particular instance of a nuisance alarm. Nonetheless, efforts were made to identify any such factors. While no environmental or traffic operational factors clearly emerged, there was some indication that annoyance may be influenced by the presence of others (passengers, other drivers) or by transient states of the driver. Because alarms occurred at random times with respect to what the driver was doing, future in-vehicle warning systems will probably show a relationship of nuisance alarms to driving maneuvers that was not represented here. No relationship to maneuver was evident in this experiment, but the data were very limited in this regard. Other methods, such as simulator studies, would be more appropriate for addressing this issue.

The methodology employed in this study appears promising, although some procedural refinements may be useful. It was proven feasible to install and maintain this equipment in people's vehicles over an extended period (nine weeks), with excellent cooperation maintained from the participants. The findings were interpretable and generally significant. However, some improvements may be suggested. It is not clear that the various rating scales were optimal, and changes should be piloted before any future implementations. Changes in the anchor words may be useful, as would the use of clearly bipolar scales with a zero point for indifference. Furthermore, it would be useful to develop comparative values for external events or conditions, so that annoyance could be expressed more clearly relative to meaningful anchors. For example, one might try to determine what rating scale value would correspond to the equivalent annoyance of some other nuisance or what level corresponds to the point at which participants are no longer willing to accept the system. Other methodological factors may relate to the duration of the alarm conditions. While nine weeks is very long compared to most driver behavior experiments, it is certainly less than "permanent." Furthermore, since the condition changed every week or two, there was still some novelty. Future work might consider the effect of a condition when it occurs over a much longer period (or if the participant at least thinks that it will). Changes to the reward/penalty structure of the method may have some effect on the

participant's response, and this might be explored and related to the magnitude of perceived benefits to warning systems. Finally, as was discussed in the introduction to this report, there is no practical ideal method for studying driver response to nuisance alarms, and the stimuli and methods of this type of procedure should be integrated with that of other methods, such as simulator studies. While generally successful, the present experiment must be viewed as an initial step, both in terms of findings and method.

In summary, this experiment used an innovative in-vehicle method to collect systematic data on driver annoyance to various nuisance alarm conditions. This is certainly a preliminary study, and refinements to the method may be useful. Furthermore, it is clearly recognized that in order to address the very difficult issue of acceptable rates for nuisance alarms, a variety of research methods will have to be used in a complementary fashion. Nonetheless, this experiment demonstrated that there is annoyance engendered by nuisance alarms, and that it is related to the frequency and type of alarms. Voice messages generated notably more annoyance than the tonal signal used here. Voice alarms occurring once per hour, on the average, and acoustic signals occurring four times per hour, on the average, resulted in significantly more annoyance than less frequent tonal signals. While the rate of nuisance alarms that will be tolerable to drivers will require much more research to specify, and while it will certainly be related to a host of complex factors, the findings of this experiment suggest promise for future in-vehicle crash avoidance warning systems. The finding that most of the substantial drop in annoyance came between the 4/hour and 1/hour rates, and the high "acceptability" ratings for the more sparse rates, suggests that nuisance alarms may not have to be very infrequent for drivers to accept warning products.

In the absence of research data or adequate field experience, there has been speculation within the intelligent transportation system (ITS) community that drivers will reject as too annoying any warning product that produces nuisance alarms except on an extremely rare basis. Drivers in this experiment typically drove at least eight hours per week, and reported spending a weekly average of about 13 hours driving. Thus even under the most sparse nuisance alarm condition tested, nuisance alarms routinely occurred once or twice weekly. Yet both ratings and subjective comments suggested that while some annoyance may have been measurable for those more sparse alarm conditions, it appears rather minimal and acceptable. While future research will have to confirm and refine the findings, the reactions of drivers to extended exposure in their own vehicles suggests that moderate rates of intrusive alarms may be acceptable.

5.0 References

American Association of State Highway and Transportation Officials. (1990) *A policy on geometric design of highways and streets*. Washington, DC: Author.

COMSIS Corporation. (1995). *Multiple attribute evaluation of auditory warning signals for in-vehicle crash avoidance warning systems*. Report under contract DTNH22-91-07004. Washington, DC: National Highway Traffic Safety Administration.

Appendix A

Daily Driver Questionnaire

Date:

Did anyone other than yourself drive the car today?

Yes No

If yes, please explain on bottom, and provide the time(s) during the day this happened

Did you do any driving today?

Yes No

Did you hear any sounds from the equipment today?

Yes No

If no to either question. do not fill out remainder of questionnaire for today

About how many times did you hear all the sounds today?

of times

Did the light and the sound ever come on at the same time today?

Yes No

Did the sound without the light ever come on today?

Yes No

Use to note anything unusual about your reaction to the sounds

Over all the trips you made today:

A) How noticeable were all the sounds from our equipment?

1 2 3 4 5 6 7 8 9
barely noticeable somewhat very noticeable

B) Considering the driving you did today, how annoying was it to have the sounds in your car?

1 2 3 4 5 6 7 8 9
not annoying at all tolerable extremely annoying

What kinds of trips did you make today? (Check all that apply)

- Commuting to and from work
Driving as part of my job
Running household errands
Driving for pleasure or leisure
Other:

In general at the time of all the sounds, were you doing anything else in addition to driving? (Check all that apply)

- Listening to radio/stereo
Using car phone
Talking to passenger(s)
Eating/drinking/smoking
Other:



Appendix B

Participant ID:

Date:

WEEKLY DRIVER QUESTIONNAIRE

Current Odometer reading _____

During the course of driving this past week, how noticeable were all the sounds from our equipment?

1	2	3	4	5	6	7	8	9
barely noticeable			some what			very noticeable		

During the course of driving this past week, how annoying was it to have the sounds in your car

1	2	3	4	5	6	7	8	9
not annoying at all			tolerable			extremely annoying		

Imagine your car is equipped with a future “smart car” system that can sense possible crash situations, such as another vehicle too close to you. How acceptable would you find this kind of system to be, if it also gave you false alarms at the rate you experienced this week?

1	2	3	4	5	6	7	8	9
unaccept-able			can't tell			highly acceptable		

Based on your experiences this week, how would you rate the annoyance of the sounds under these various conditions:

	less annoying		equally annoying			more annoying			did not occur	
Compared to driving by myself, the sounds that occurred while driving with other passengers were:	1	2	3	4	5	6	7	8	9	0
Compared to when the radio/stereo system was off, the sounds when the radio/stereo system was on were:	1	2	3	4	5	6	7	8	9	0
Compared to when the sounds occurred during light traffic, the sounds that occurred in heavy traffic were :	1	2	3	4	5	6	7	8	9	0
Compared to when I was driving on minor roadways, the sounds that occurred when I was on freeways/major highways were:	1	2	3	4	5	6	7	8	9	0
Compared to when I was driving for pleasure or leisure, the sounds that occurred when I was driving for other purposes were:	1	2	3	4	5	6	7	8	9	0
Compared to when I drove during daytime or light conditions, the sounds that occurred when I was driving during nighttime or dark conditions were:	1	2	3	4	5	6	7	8	9	0
Other: (please describe below)	1	2	3	4	5	6	7	8	9	0

Appendix C

TRAFFIC CONDITIONS RATING FORM

Participant ID _____ Week # _____ Rater ID _____

Roadway Factors

_____ Clear

Type:

Traffic Density/Levels of Service

_____ Freeway

_____ A - free flow

_____ Ramp

_____ B - stable flow

_____ Divided Multilane Highway

_____ C - interactions noticeable

_____ Undivided Multilane Highway

_____ D - high density, stable flow

_____ 2-Lane Highway

_____ E - at or near capacity

_____ Arterial

_____ F - breakdown, stop & go

_____ Collector

_____ not applicable (i.e., not in traffic)

_____ Downtown Street

Vehicle Motion/Maneuvers

_____ Non-roadway (e.g., parking lot)

_____ Car in forward motion

_____ Other: _____

_____ Car in backward motion

Location:

_____ Stopped at intersection

_____ Rural

_____ Stopped elsewhere (e.g.,

_____ Urban (business)

MacDonalds)

_____ Suburban

_____ Left turn

Surface Conditions:

_____ Right turn

_____ Dry

_____ Lane Change

_____ Wet

_____ Merge

Outside Visibility

_____ Anticipating a maneuver

_____ Precipitation falling

_____ Other: _____

_____ Other limited visibility (e.g. fog)

Alarm Time:

a.m. / p.m.

Appendix C-- Continued

Roadway Location Classification Scheme -- based on definitions per Highway Capacity Manual:

1. Freeway

- Divided highway facility having two or more lanes for the exclusive use of traffic in each direction and full control of access and egress.
- Completely “uninterrupted” flow; access and egress is only at ramps.

2. Ramp

- A length of roadway providing an exclusive connection between two highway facilities.

3. Multilane highways

- 4 basic classifications:
 - rural, divided
 - rural, undivided
 - suburban, divided
 - suburban, undivided
- Differ from arterials by:
 - roadside development not as intense as arterials
 - less dense traffic
 - signalized intersections greater than two miles apart

4. 2-lane Highways

- A 2-lane roadway having one lane for use by traffic in each direction.

5. Urban and Suburban Arterials

- Signalized streets.
- Provide access to abutting properties.

6. Collector

- Traffic access service within residential, commercial, and industrial areas.
- Not always dominated by traffic signals.

7. Downtown streets

- Provide access to local business.

Appendix C- Continued

Traffic Density Bating Criteria -- to be used for any type of roadway location

- A -- free flow, unaffected by others in traffic stream**
- B -- stable flow, presence of other users noticeable; slight decline in freedom to maneuver**
- C -- beginning stage in which operation of user becomes significantly affected by interactions with others in traffic stream**
- D -- high density, but stable flow; speed and maneuvers severely restricted, bordering on unstable flow**
- E -- conditions at or near capacity; speeds low but relatively uniform; freedom to maneuver extremely difficult**
- F -- forced or breakdown flow; queues forming; stop and go traffic**