ADVISORY AND ALARM STIMULI OPTIMIZATION FOR A
DROWSY DRIVER DETECTION SYSTEM

by

Sarah E. Fahey

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
Master of Science
in
Industrial and Systems Engineering

APPROVED

Dr. Walter W. Wierwille, Chair

Dr. Robert D. Dryden

Dr. Dennis L. Price

July, 1995
Blacksburg, Virginia
ADVISORY AND ALARM STIMULI OPTIMIZATION FOR A
DROWSY DRIVER DETECTION SYSTEM

by
Sarah E. Fahey
Dr. Walter W. Wierwille, Chair
Industrial and Systems Engineering

(ABSTRACT)

An experimental study was carried out to identify effective advisory and alarm
stimuli to be used in a drowsy driver detection system. The envisioned system has three
stages. In the first stage, previously developed detection algorithms would compute on-
line drowsiness levels. If a driver's drowsiness level exceeds a predetermined threshold
the system would proceed to stage two. At this point an initial advisory tone and a voice
message would be played. If the driver does not respond, he or she would experience a re-
alerting alarm. The third stage of the system would give the driver an option of using a
drowsiness countermeasure to help maintain the re-alerted state.

The goal of the present research was to determine the effectiveness of possible
stimuli to be used in the second and third stages of the envisioned system. Eight initial
advisory tones, two voice messages, eight alarm sounds, and five peripheral stimuli were
investigated as part of stage two. In addition, six drowsiness countermeasures to be used
in stage three were investigated. Eight graduate students in the Human Factors Engineering
program at Virginia Tech volunteered as subjects. Subjects drove the automobile simulator
throughout the experimental session. Data were collected using subjective opinion, paired
comparisons, and effectiveness ratings.
This study succeeded in answering many question regarding stimuli to be used in a drowsy driver detection, advising, and alerting system. The results of the study indicated very effective stimuli to be used in the advising and alerting stages of the envisioned system.
ACKNOWLEDGMENTS

My major advisor, Dr. Walter W. Wierwille, deserves thanks and appreciation for his wisdom, advice, and guidance in every phase of this project as well as for opening my eyes to a lot more than just Human Factors. I would also like to thank the other members of my committee, Dr. Dennis L. Price and Dr. Robert D. Dryden, for their helpful suggestions.

Special thanks go to my parents, John and Audrey Fahey, and my brother, Peter, for their love and support not only during the past two years but all my life. Without it, many accomplishments would not have been possible.

I'd also like to thank the many friends here at Tech that without whom my graduate experience would not have been nearly as enjoyable. I'd especially like to thank my fellow lab mates Vicki Neale, Steve Wreggit, and Chris Kirn for sharing their experiences and advice with me. Thank you to Mark McMulkin for his continuous patience and willingness to share his knowledge of statistics. And last but definitely not least, I'd like to thank Terry Fairbanks for being a forgiving sounding board, giving a sturdy shoulder to lean on, and being the all around great friend that he is. Without him I would have given up long ago.

This research was sponsored by the National Highway Traffic Safety Administration (NHTSA). Dr. Ron Knipling served as Contract Technical Representative.
# TABLE OF CONTENTS

INTRODUCTION .............................................................................................................. 1

LITERATURE REVIEW ................................................................................................. 4

Physiological Measures of Drowsiness ................................................................. 4

   Eyelid closure ....................................................................................................... 4

   Eye movement .................................................................................................... 5

   Muscle activity ................................................................................................ 5

   Brain wave activity ......................................................................................... 6

   Skin potential .................................................................................................. 6

   Heart rate variability ...................................................................................... 7

Subjective Ratings of Drowsiness ........................................................................ 7

Performance Measures as Indicators of Drowsiness ........................................ 7

   Lateral control measures .............................................................................. 8

   Driver response measures ......................................................................... 10

   Longitudinal control measures ................................................................... 11

Countermeasures to Drowsiness .......................................................................... 12

   Simple countermeasures ............................................................................ 12

   Chemical Countermeasures ......................................................................... 13

   Mechanical and Electronic Countermeasures ............................................ 14

Vehicle-Based Drowsy Driver Detection .............................................................. 17

   Recent Research ............................................................................................ 17

   Research at Virginia Tech ............................................................................ 17

      Validation of performance measures ..................................................... 18

      Validation of driver performance and behavioral measures ............ 19

      Operational definition development ..................................................... 19

      Algorithm development ......................................................................... 20
Algorithm validation .............................................. 22
Usability ................................................................... 22
The Method of Paired Comparisons ............................... 25
METHODS OF ALERTING ........................................... 27
Warning Systems for Motor Vehicles .............................. 27
  General warning device design ................................... 27
  Driver alertness warnings ......................................... 29
    Auditory displays .................................................. 29
    Tactile displays .................................................... 30
    Visual displays ..................................................... 30
    Termination of warnings ......................................... 31
Related Research ..................................................... 31
PRESENT STUDY ....................................................... 34
Research Objectives .................................................. 34
  Stage one - Initial detection of reduced alertness level (drowsiness) .......... 34
  Stage two - Re-alerting the driver ............................... 38
  Stage three - Maintaining alertness ............................. 39
    Seat vibration ...................................................... 39
    Scent ............................................................... 39
    Lane-minder ....................................................... 39
    A/O task ......................................................... 40
    Fresh air ......................................................... 40
    Glasses ........................................................... 40
Method .................................................................. 41
Subjects ............................................................... 41
Apparatus ............................................................. 43
  Simulator ............................................................ 43
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>System Flow Diagram for Initial Detection of Drowsiness (Stage 1)</td>
<td>35</td>
</tr>
<tr>
<td>1B</td>
<td>System Flow Diagram for Re-alerting Driver (Stage 2)</td>
<td>36</td>
</tr>
<tr>
<td>1C</td>
<td>System Flow Diagram for Maintaining Alertness (Stage 3)</td>
<td>37</td>
</tr>
<tr>
<td>2</td>
<td>The Effectiveness Rating Scale Used in the Experiment</td>
<td>42</td>
</tr>
<tr>
<td>3</td>
<td>Mean Effectiveness Ratings for Initial Advisory Tones</td>
<td>53</td>
</tr>
<tr>
<td>4</td>
<td>Mean Effectiveness Ratings of Alarm Sounds Alone and with Peripheral Stimuli</td>
<td>62</td>
</tr>
<tr>
<td>A1</td>
<td>Driver’s Control Panel for Alertness Monitor</td>
<td>81</td>
</tr>
<tr>
<td>D1</td>
<td>Sketch of Vibrator Assembly</td>
<td>89</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1: Summary of Average Sound Pressure Level Measurements ..................43
2: Summary Table of Initial Advisory Tone Data ..................................54
3: Results of the Maxwell Scaling Procedure on the Frequency of Preference ...56
4: ANOVA Summary Table For Scaled Values of Alarm Sound Preferences ....57
5: ANOVA Summary Table of Stimulus Condition (SC) x Alarm Type (AT) ...59
6: Summary of the Analysis of the Effectiveness Ratings of Alarm Sounds Alone...60
7: Average Effectiveness Ratings and Results of Post-hoc Comparisons Between Each Alarm Type Alone and that Alarm Type with Different Stimulus Conditions .................................................................61
8: Average Effectiveness Ratings and Results of Post-Hoc Comparisons Within Each Peripheral Stimulus Condition .................................................................63
B1: Average Sound Levels for Individual Vehicles ..................................84
F1: Characteristics of Initial Advisory Tones ..........................................101
F2: Presentation Order of Initial Advisory Tones .....................................101
G1: Sound Pressure Levels for Initial Advisory Tones ..............................103
G2: Sound Pressure Levels for Voice Messages ......................................103
G3: Sound Pressure Levels for Alarm Sounds ......................................104
G4: Sound Pressure Levels for Ambient Noise in Laboratory and Simulator Running .................................................................104
H1: Preference Matrix for Paired Comparison Data ..................................106
H2: Proportion Matrix for Paired Comparison Data ..................................106
H3: Matrix of z-deviates used in Maxwell’s scaling technique .....................107
H4: Mean z-deviates and corresponding scaled values ..............................107
I1: Average Amplitude and Effectiveness Ratings for the Voice Messages ....109
J1: Average Effectiveness Ratings for Peripheral Stimuli ..........................111
K1: Average Effectiveness Ratings for Countermeasures ..........................113
INTRODUCTION

Driver drowsiness is a major concern of transportation researchers today, and with good reason. Drowsy drivers are a major cause of motor vehicle accidents. Today’s automobile designs concentrate on comfort and luxury. Reduced road noise and vibration are viewed as quality. However, these qualities in a vehicle can add to the problem of driver drowsiness. Operating a motor vehicle requires constant monitoring of the surrounding environment. When a driver becomes drowsy it becomes increasingly difficult to be attentive to the driving task.

There have been several reports on the severity of the drowsy driver problem. Over the course of one year 50% of all fatalities on the Ohio Turnpike were attributed to accidents involving drowsiness (Kearney, 1966). In a 1973 survey reported by Seko (1984), it was found that 9% of all accidents were due to fatigue and that driver drowsiness was responsible for 45% of all automobile fatalities. In a 1980 survey also reported by Seko (1984), 75% of all drivers surveyed reported having experienced drowsiness while driving. Planque, Chaput, Petit, Tarriere, and Chabanon (1991) reported 26% of fatal accidents on the motorways in France were caused by drowsy drivers.

In 1991, the Office of Crash Avoidance Research (OCAR), which is a group within The National Highway Traffic Safety Administration (NHTSA), compiled a summary of crash statistics. A summary of the findings was as follows:

- There are 72,000 crashes annually in which driver fatigue/drowsiness involvement is cited on the police report.
• 14,000 (2.9% of all) accidents involving serious injuries are due to drowsy drivers.
• 1550 drowsy driver accidents involve fatalities. This makes up 3.4% of all fatalities.
• Actual involvement of driver fatigue/drowsiness is most likely much greater due to under reporting.

A 1973 study conducted at Duke University found that a significant number of drivers experienced drowsiness while driving and that drowsiness was a significant cause of automobile accidents (Tiley, Erwin, and Gianturco, 1973). This study found that 64% of 1500 drivers who were questioned were in agreement with the statement "I become drowsy while driving". This study also found that 31.2% of those who had become drowsy while driving believed their drowsiness had occurred before they were aware of it. Also, 93.7% reported the drowsiness process was reversible; that is they felt they could take actions, while continuing to drive, to reduce their level of drowsiness (Tiley et al., 1973). Two points can be drawn from these last findings. First, people often do not recognize they are drowsy until after they begin to exhibit symptoms. Second, the process is felt to be reversible suggesting that countermeasures could be quite effective.

In light of the fact that drivers usually don't realize they are drowsy until after they have displayed symptoms, methods of operationally defining drowsiness have been investigated. Recent studies conducted at Virginia Polytechnic Institute and State University, funded by The National Highway Traffic and Safety Administration, have investigated both physiological and performance measures. Many physiological measures have been shown to be good predictors of driver drowsiness. However, problems arise when trying to implement these measures because they may be intrusive to the driving task or annoying to the driver. For these reasons performance measures
have been the focus of more recent research. Drowsiness detection algorithms utilizing performance measures have been developed, refined, and validated (Wierwille, Wreggit, Kirn, Ellsworth, and Fairbanks, 1994). The focus of research must now turn to the events which must take place after a drowsiness condition is detected. Some means of advising and alerting a driver and maintaining that re-alerted state are needed.

The goal of this research was to develop an appropriate warning system to produce a fast acting drowsy driver detection and alerting system for use with the detection algorithms described in Wierwille et al. (1994). The purpose of the warning signals is three-fold: first, to advise the driver that a drowsiness condition has been detected; second, to re-alert the drowsy driver; and third, to maintain his or her alertness for a period of time, making it possible to find a safe place to pull off the road and refresh. Because of the many factors involved in this study, a heuristic type evaluation process appeared to be the most efficient way to gather information necessary to design an optimal advising/alarming system.
LITERATURE REVIEW

Physiological Measures of Drowsiness

Several researchers have attempted to define drowsiness over the years. Torsvall and Akerstedt (1988) defined drowsiness as the "state during which sleep is perceived as difficult to resist, the individual struggles against sleep, performance lapses occur, and sleep eventually ensues." Much research has been conducted to operationally define drowsiness to aid in the detection of drowsy drivers. Six physiological measures have been shown to be good predictors of drowsiness. These are eyelid closures, eye movement, muscle activity, brain wave activity, skin potential, and heart rate variability. These six measures were examined in a literature review conducted by Wierwille, Wreggit, and Mitchell (1992). A summary of the findings is presented here.

Eyelid closure. Studies have suggested that slow eyelid closure is a very reliable indicator of drowsiness (Erwin, 1976; Skipper, Wierwille, and Hardee, 1984). Erwin (1976) looked at several physiological measures and found slow eyelid closures to be the best predictor of the onset of sleep and degraded task performance. Eyelid closures can be easily and unobtrusively recorded with a low-light camera. The image is viewed by an experimenter who operates a potentiometer indicating the degree to which the driver's eyes are open. Two measures used by Ellsworth, Wreggit, and Wierwille (1993) and supported by research conducted by Wreggit, Kirn, and Wierwille (1993) are PERCLOS and EYEMEAS. These are described below:

PERCLOS - the percentage of time the eyes are 80% to 100% closed over a one-minute interval.
EYEMEAS - the mean-square of the percent eyelid closure signal sampled over a one minute interval. (EYEMEAS is more heavily weighted as eye closure increases).

Although these two measures have been shown to be good indicators of drowsiness, they ignore the possibility that some individuals may become drowsy without exhibiting eyelid closures. For this reason other physiological measures of drowsiness have been examined.

**Eye Movement.** Planque et al. (1991) reported that the characteristics of normal eye movement undergo changes with the onset of drowsiness. These changes can be measured through the use of electrooculography (EOG) which measures electrical potentials by placing electrodes in contact with the skin around the eyes. As a person becomes drowsy, he or she exhibits slow, rolling, lateral eye movements as opposed to normal eye movements which are quick and voluntary. The slow eye movements (SEMs) are characterized on the EOG signal by slow deflections of more than a second and amplitudes of at least 100 microvolts. As the level of drowsiness continues to increase the occurrence of SEMs also increases (Torsvall et al., 1988). It has also been reported that the amplitude of the EOG increases with the degree of drowsiness (Santamaria and Chiappa, 1987). Additionally, attentiveness, as well as alertness, has been shown to decrease with drowsiness (Endo, Inomata, and Sugiyama, 1978). Specifically, Endo et al. (1978) found that the number of voluntary lateral eye movements normally displayed during the driving task, such as looking in the rear view mirror and side mirrors, decreased with the occurrence of drowsiness.

**Muscle Activity.** A reduction in muscle activity has been suggested to accompany drowsiness, especially those muscles controlling one’s facial expressions
(Yabuta, Iizuka, Yanagishima, Kataoka, and Seno, 1985). A study conducted by Hauri (1982) showed a decrease in muscle potentials as an individual neared the first stage of sleep. Erwin (1976), on the other hand, argued that muscle activity has poor predictive value because many minutes of sleep are possible before a significant change in muscle tone occurs.

**Brain Wave Activity.** Brain wave activity, measured using an electroencephalogram (EEG), exhibits changes during sleep. However, it is not clear whether brain wave activity can be used as a predictive measure of drowsiness. Erwin (1976) found no alteration in brain wave activity prior to eyelid closure. However, upon eyelid closure Erwin (1976) found a rapid shift in brain wave patterns. Planque et al. (1991) found indications that sharp changes in the frequency of brain wave activity occurred as an individual moved from a state of alertness to that of sleep. Specifically, an increase in alpha waves and a decrease in beta waves accompanied the decline in performance. Seko (1984) states that alpha waves appear during times of decreased alertness. Seko (1984) cites the research of Kuroki, Kitakawa, and Oe (1974) which showed alpha waves were not present at the beginning of a driving session although they were easily detected as the driving session continued and the drivers’ alertness levels decreased.

**Skin Potential.** The skin potential level (SPL) is measured by attaching electrodes to the skin and looking at the difference in potential between the outermost layer of skin and the layer directly beneath it. Research conducted by Erwin, Hartwell, Volow, and Alberti (1976) showed a correlation between SPL and stages of arousal. A decrease in negativity of SPL accompanies all EEG-defined sleep although a decreased negativity of SPL was also found during times of wakefulness (Erwin et al., 1976). These findings suggest that a reduced SPL is a prerequisite for sleep but a decrease in SPL is not necessarily followed by the onset of sleep.
**Heart Rate Variability.** A correlation between heart rate variability and drivers’ fatigue levels has been found in several studies (Wierwille and Muto, 1981; Sugarman and Cozad, 1972; Riemersma, Sanders, Wildervanck, Gaillard, 1977). Not all of these researchers, however, feel their findings are conclusive. Riemersma et al. (1977) suggest that the changes in heart rate variability that were recorded could have been caused by the subjects’ adaptation to the experimental situation.

**Subjective Ratings of Drowsiness**

For the most part, studies of drowsiness that have involved ratings have relied on individuals rating their own level of drowsiness. Observer ratings were investigated in a study carried out by Carroll, Blisewise, and Dement (1989). The study was conducted in several nursing home settings. The results of this study suggest that the use of observer ratings is a valid approach to studying drowsiness.

In 1993, Ellsworth et al. conducted a study in which raters viewed sleep deprived subjects on videotape. Observers were briefed on the rating scale which based drowsiness on eyelid closures, mannerisms, and facial tone. The results showed that informed raters rated consistently. Intrarater and test-retest reliability tests produced correlation values of 0.88 and 0.81 respectively. The intrarater reliability also showed a high correlation value of 0.81. These results show that raters appear to be consistent within themselves and also tend to agree with one another. Ellsworth et al. (1993) also found subjects’ self ratings to be correlated with eye closure measures.

**Performance Measures as Indicators of Drowsiness**

Although research has pointed to several physiological measures as predictive indicators of drowsiness, there has been interest in identifying driver performance
measures. Monitoring performance is usually unobtrusive to the driver since the measures are activities which normally occur during driving. Specifically, these measures are lateral control measures, driver response measures, and longitudinal control measures (Wierwille et al., 1992). These three measures are described below.

**Lateral Control Measures.** Three categories of lateral control measures which have been studied are lane-related measures, steering-related measures, and heading and lateral acceleration measures. Many components of driving are over-learned and thus taken for granted by an impaired driver. Maintaining correct lane position is one of these components (Dingus, Hardee, and Wierwille, 1985). Other studies have shown that lane position errors increase over prolonged periods of time (Mast, Jones, and Heimstra, 1966; Sussman, Sugarman, and Knight, 1971).

Wierwille et al. (1992) found five lane-related measures to be accurate and reliable measures for driver drowsiness detection. All five are feasible for on the road as well as in simulator use. The five measures are described here.

- **LANEDEVM** - Lane deviations which are heavily weighted for lane exceedences.

- **LANESTD** - The standard deviation of the lane position.

- **LANEDEV** - The global maximum lane deviation.

- **LANEDEVSQ** - The mean square of the lane deviation.

- **LATPOSM** - The mean square of the high pass lateral position (heavily weighted for rapid changes in lateral position).
Steering behavior tends to change as a driver becomes impaired due to drowsiness (Wierwille et al., 1992). Specifically, the changes can be noted in the frequency and type of steering reversals exhibited. Hulbert (1963) found that sleep deprived subjects made fewer small steering reversals than rested drivers. This causes lane deviations followed by large steering reversals once the driver has become aware of the deviation. As a result, large steering reversals tend to increase with drowsiness, and small steering reversals tend to decrease with driving time (Wierwille et al., 1981; Khardi and Vallet, 1994).

Four steering-related measures have been shown to be reliable predictors of drowsiness. Again, these measures are feasible for on the road as well as in simulator use. A brief description of these measures follows (Wierwille et al., 1992).

STVELM - The steering velocity weighted heavily for fast maneuvers.

STEXCEED - The number of times steering velocity exceeds a criterion (150 degrees/second over a three minute interval).

STVELV - Steering velocity variance (calculated over a three minute interval).

LGREV - The number of times the steering wheel position increment exceeds five degrees (after steering velocity passed through zero).

Heading is a very important aspect of driving that needs to be constantly monitored to ensure safe driving, especially at high speeds. A deviation of only one
degree can cause a car to go out of the lane boundary in a matter of seconds. From this it is easy to see that the monitoring of heading could be quite beneficial from a safety standpoint. Heading related measures have also been investigated and found to be reliable measures for the detection of drowsiness. The three most promising are described below (Wierwille et al., 1992).

**YAWDEV** - The global maximum yaw deviation.

**YAWVAR** - The yaw deviation variance (calculated over a three minute interval).

**YAWMEAN** - The mean yaw deviation (calculated over a three minute interval).

**Driver Response Measures.** As one would expect, reaction times on task performance have been shown to increase with sleep loss and time on task (Ryder, Malin, and Kinsley, 1981). Therefore, it follows that driver performance measures could be used in the detection of drowsiness (Hardee, Dingus, and Wierwille, 1985). Driving task response measures and subsidiary task response measures are two types of driver response measures that have been studied. Unfortunately, the stimuli associated with driving task response measures, such as roadside obstacles and lead cars, are not always present in the driving environment.

Subsidiary task response measures were studied by Hardee et al. (1985). Sleep deprived subjects were exposed to various types of stimuli. The correct number of responses in a three minute period was recorded and found to have potential as a drowsiness predictor. This measure is referred to as NUMHIT. Hardee et al. (1985)
suggest that this measure be used as a second level detector.

**Longitudinal Control Measures.** Longitudinal measures are those relating to speed, acceleration and braking behavior, and vehicle following-related measures. Many longitudinal measures would be difficult to obtain in a real vehicle situation, and they have given only mixed results in the literature.

A literature review conducted by Hardee et al. (1985) found no data to suggest a correlation between accelerator-related measures and drowsiness or time on task. However, research conducted by Safford and Rockwell (1967) found accelerator pedal reversals to be correlated with time over a twenty-four hour driving task.

Brake/Deceleration-related measures have not been found to change significantly after 29 hours of sleep deprivation (Huntley and Centybear, 1974) or as a function of eight or twelve hours of driving (Brown, 1965; 1966; Brown, Simmons, and Tickner, 1967).

Speed variability has been found to increase significantly during the course of night driving (Riemersa et al., 1977). Kirn (1994) found a moderately high correlation between speed variability and level of drowsiness for subjects in a "no task" condition. This study will be discussed in further detail later. Conversely, several studies have found no significant variability in speed over time during driving tasks (Brown, 1965; 1966; Brown et al., 1967; Safford et al., 1967).

Wierwille et al. (1981) researched vehicle-following behavior in a simulator study. Reaction times were found to be significantly greater after subjects drove the simulator for 30, 60, and 150 minutes compared to baseline runs. Many problems arise with the use of this measure. In a real driving environment a lead car is not always present and the ease with which a car could be implemented with a device capable of measuring this behavior is questionable.
Countermeasures to Drowsiness

Because of the scope of the drowsy driver problem, many measures have been researched in hopes of developing a system that would detect a decrease in alertness before the driver is aware of his or her condition. Much of this research has shown promising results. However, problems still remain. Research is necessary to determine how to notify the driver of his or her condition and also how to counteract the drowsiness for a period of time so that the driver can reach a safe stopping area.

Simple countermeasures. Haworth and Heffernan (1989) showed that inadequate sleep or rest, prolonged hours of driving, and the intake of certain foods and drugs contribute to the onset of fatigue. Simple avoidance of these factors could be one way of counteracting drowsiness. However, this is not always possible.

Haworth and Vulcan (1990) classified countermeasures into three types. These are driver-oriented, vehicle-based, and environmental countermeasures. The driver-oriented countermeasures consist of educating drivers to recognize the signs and dangers of drowsiness while driving. Two ways to do this, suggested by Haworth et al. (1990), are driver training courses and mass media campaigns. Another fairly simple way to reduce the risk of drowsy driving is to avoid driving for extended periods of time after a long interval of work. However, this scenario is one that is not easily avoided in today’s society. Many people commute an hour or more to and from work everyday. So, although this may seem like a simple way to reduce drowsy driving it is not very practical.

Vehicle-based countermeasures consist of listening to the radio and providing ventilation, possibly by opening the window. Mackie and O’Halon (1977) provide evidence of cool air having a preventive effect. However, there is no evidence of cool air alleviating already existing fatigue. Fatigue monitors are another type of vehicle-based countermeasure. Examples of these will be discussed in some detail later.
Environmental countermeasures include rest breaks and pavement treatments. There are conflicting results in research investigating the effectiveness of rest breaks. Harris and Mackie (1972) reported that rest breaks may have a refreshing effect before fatigue has developed but once fatigue has set in they have little beneficial value. Harris et al. (1972) also concluded that rest breaks of less than 20 minutes produced no recovery effect in driver performance while Lisper, Eriksson, Fagerstrom, and Lindholm (1979) found no difference in the effects of 15 and 60 minute rest breaks. Pavement treatments, specifically rumble strips, have been shown to be effective countermeasures to drowsiness (Haworth et al., 1990; Wood, 1994). Pavement treatments will also be discussed in some detail later.

Harris (1967) mentions several other simple methods of counteracting drowsiness. These include chewing gum, singing along with the radio, sitting on something hard, and taking off the right shoe. These methods may work for some, but unfortunately, they do not work consistently or for all people.

Chemical Countermeasures. Caffeine is considered by many as a favorable way to keep themselves awake while driving for extended periods of time. Haworth et al. (1990) cites many studies which examined the effects of caffeine on driving. Reaction times for a subsidiary task while driving were significantly shorter for a period of 1.5 hours for those drivers that had been given caffeine before a three hour drive (Lisper, Tornros, and Van Loon, 1981). Baker and Theologus (1972) found caffeine significantly reduced attention lapses. Childs (1978) suggests the normal daily intake of caffeine influences the overall effect it will have on reaction times. It was found that subjects who did not normally consume three cups of coffee a day exhibited increased reaction times in short term visual scanning when given 400 mg of caffeine. In contrast, those subjects who commonly drank larger amounts exhibited a decreased reaction time to a visual stimulus. Childs (1978) concludes that a large dose of caffeine may disrupt visual
attention in those who do not normally consume much caffeine, while enhancing visual attention in those who normally consume a large amount of caffeine.

Haworth et al. (1990) cites many studies that have shown stimulants, such as amphetamines, to improve tracking, concentration, and attention. The effect of these drugs is shown to be greater when the subject is sleep deprived. This suggests that they could be used to counteract fatigue. However, stimulants have also been found to increase risk taking behaviors. Because of this, they would probably not prove to be good countermeasures to use while performing a vigilance task such as driving (Haworth et al., 1990).

Nicotine has been supported in the literature as having a positive effect on the performance of vigilance tasks (Wesnes and Warburton, 1978). However, when considering any of the mentioned stimulants one must realize that they are drugs and might have serious health effects. For this reason, and because of individual differences in peoples' reactions to these drugs, they probably should not be considered viable sources of drowsiness countermeasures. A recently identified approach to combating drowsiness is the use of stimulating scents (Kaneda, Iizuka, Ueno, Hiramatsu, Taguchi, Tsukino, 1994). This research will be described further under the section entitled Recent Research.

**Mechanical and Electronic Countermeasures.** There are several devices that have been developed as attempts to measure and in some cases counteract drowsiness in drivers. These devices include eyelid closure monitors, head nodding monitors, and reaction time monitors.

One device which monitors eyelid closures is the Onguard developed by an Israeli company, Xanadu Ltd.. This battery operated device, which consists of a small infrared sensor and an electronic processor, can be fitted to a standard pair of eyeglasses. The device directs a beam of infrared light at the eye and measures the light reflected back.
When the eye is closed the amount of light reflected back is reduced. When the eye is closed for longer than 0.5 second an alarm is activated. Although this device has been shown to detect long eyelid closures, one flaw should be noted. Haworth et al. (1990) found that eyeglasses have a tendency to slip down the nose and as a result the alignment of the device may be in need of constant readjustment.

Head nod detectors have also been marketed to detect drowsiness in drivers. The Electronic Transistor Safety Alarm and Dozer’s Alarm are two examples of this type of detection device. These units, each of which consist of an angular rotational detector, are placed over the top of the ear and buzz loudly when the head nods forward past a certain angle. Dozer’s Alarm was studied by Haworth et al. (1990) and the results were less than satisfactory. Subjects found the device "Very Annoying." Thorpy and Ledereich (1990) suggest that using a head nod detector gives drivers a false sense of security. Another problem, which Hulbert (1972) points out, is the possibility that an individual may become extremely drowsy and even fully asleep before his or her head falls forward.

Finally, there are reaction time monitors. Roadguard is one such device. Once installed, the device is activated when a car is put into high gear. A timer stops at random periods of 4-14 seconds and a small red light is illuminated on the dash board. This light must be deactivated by the driver within three seconds or an alarm will sound. Haworth et al. (1990) found this device to be the most reliable of the three types studied: Onguard, Dozer’s Alarm, and Roadguard. However, there still exist limitations with the device. Lisper, Laurell, and Van Loon (1986) found that drivers could fall asleep behind the wheel for periods of less than two seconds. This suggests that the three second threshold of the Roadguard device may allow drivers to fall asleep and be involved in an accident before the alarm is activated (Haworth et al., 1990). Although the preceding devices have been shown to detect drowsiness in some cases, none has been shown to alleviate drowsiness for any significant length of time or lessen the decrease in driver performance.
(Haworth et al., 1990).

The Alert-O-Matic, developed by Frederik in 1966, is also a reaction time monitor (Hulbert, 1972). This device consists of a series of three alarms of increasing severity. A light is illuminated and the driver is able to turn it off by lightly tapping the horn. If the driver fails to do so within five seconds the device activates the car horn. If the driver still fails to tap the horn after three seconds of the horn sounding the device turns the ignition on and off for a period of five seconds. If there is still no response from the driver after an additional five seconds the ignition is shut off completely. Again the findings of Lisper et al. (1986) stated above suggest that a driver could fall asleep and crash before these alarms are activated. An additional problem with the Alert-O-Matic is that the light on the dashboard is illuminated at constant intervals of 60 seconds. Oswald (1962), cited in Harris (1967), noted that subjects exhibited a sleeping and waking pattern in which they woke up only to make a necessary response and then fell right back asleep. It seems likely that a driver could quickly learn the pattern of responding to the light every 60 seconds without being fully alerted (Harris, 1967).

The ALERTMASTER and Button Steering Wheel Alarm are similar devices. The ALERTMASTER requires the driver to apply constant pressure on a pedal located to the left of the clutch. When constant pressure is not maintained the horn sounds. The effectiveness of this device relies on the assumption that when a driver becomes drowsy the left foot will relax and not maintain the pressure needed (Hulbert, 1972).

The Button Steering Wheel Alarm operates on the same principle. In this device the button is located on the steering wheel. As with the ALERTMASTER, an alarm is activated when constant pressure is not maintained on the button. Fatigue may be a result of the constant pressure required by a driver's finger or thumb. Also, steering maneuvers and control adjustments on the instrument panel may warrant removal of the finger or thumb from the steering wheel for short amounts of time (Hulbert, 1972).
Most of the drowsiness detection devices that have been discussed here are quite intrusive to the driver. All are driver based or require the driver to actively participate in the detection process in some way. A large part of the success or failure of a detection device lies in whether or not those who could benefit from it actually use it. Therefore, public acceptance is an important issue. Because of this, much of the recent research has focused on vehicle-based drowsy driver detection, which involves ways of detecting drowsiness without attaching any equipment to the driver or requiring the driver to do anything that is not part of the normal driving task.

Vehicle-Based Drowsy Driver Detection

Recent Research. As stated previously, many physiological measures have been shown to be very good predictors of drowsiness. The correlation of these measures with other non-contact definitional measures has been investigated.

Kaneda et al. (1994) developed a detection system which uses image processing of the driver's face to detect diminished alertness. The image of the driver's face is processed to locate the eyes and then determine the degree to which they are open. This information was shown to have a moderately high correlation (0.77) with an alertness index which was based on a combination of brain wave measurements, blinking rates, and facial expressions rated by observers.

Artaud, Planque, Lavergne, Cara, Lepine, Tarriere, and Gueguen (1994) also looked at facial image processing as a drowsiness indicator. Attempts to relate facial images to driving behavior, such as steering and respiratory signals, are underway.

Research at Virginia Tech. Many studies involving drowsiness detection have been conducted in the Vehicle Analysis and Simulation Laboratory at Virginia Tech. As has been shown in the literature review, there are several physiological measures that have been proven time and time again to be good predictors of driver drowsiness.
However, there are problems with recording these measures because the necessary instrumentation is either intrusive upon the driving task or annoying to the driver. For these reasons, performance measures as predictors have been the topic of numerous research projects at Virginia Tech. Over the past several years algorithms for the detection of drowsy drivers have been developed, refined, and validated. A summary of this research follows.

*Validation of Performance Measures.* Skipper et al. (1984) conducted a study to investigate performance measures as potential predictors of drowsiness in drivers. The study looked at drowsiness resulting from both sleep deprivation and time on task. A measure of eyelid closure (EYEMEAS) was used as the primary indicator of drowsiness.

Subjects drove in a high fidelity moving base simulator for 1.5 hours. During their "drive" low level steering wheel torque and front wheel disturbances were applied to the simulation at approximately one-minute intervals. The driver's task was to respond to these stimuli to maintain the simulated vehicle's lane position. Nine of the seventeen driver-vehicle measures were taken in the six-second time period directly following the presentation of the stimuli. The remaining eight measurements were taken during the 50 seconds of "normal driving" immediately preceding the introduction of the stimuli.

Results showed that the driver's condition, whether sleep deprived or well rested, and time on task did indeed have an effect on the measures. The measures of LANESTD, LANDEV, and STVELM were highly correlated with EYEMEAS during "normal driving" between stimuli. The measures of LANDEV and YAWMEAN were found to be highly correlated with EYEMEAS during the six-second time period directly following the presentation of the stimuli.

Periods where lane boundaries were exceeded were also recorded. Analysis of
this information showed that drivers had a tendency to oscillate between the lane boundaries before actually crossing over them. Also, drivers exhibited a tendency to hold the steering wheel relatively still while lateral deviation increased and then produce a large steering input to correct lane position. The most promising performance measures, found in the Skipper et al. (1984) results, were then combined into multivariate linear discriminant models. These models suggested these performance measures could predict drowsiness with some degree of success.

*Validation of Driver Performance and Behavioral Measures.* A subsequent research project conducted by Dingus et al. (1985) looked at both driver performance and behavioral measures as predictors of driver impairment due to drowsiness and/or alcohol. Nineteen measures in all were examined. Seven measures of eyelid closures and lane deviation were examined for their predictive value, and the remaining twelve measures were evaluated for reliability in impairment detection. Data collection took place at three- and six-minute intervals. After a series of analyses, YAVVAR, SEATMOV, LANEDEVSQ, YAWMEAN, and STEXCEED were found to contain a significant amount of independent information. SEATMOV is the measure of driver "restlessness" which was measured by transducers in the seat pad and backrest. The other measures have been described earlier. The results showed the six-minute interval data provided better discrimination and had fewer false alarms, however, the miss percentage was lower for the three-minute intervals (Dingus et al., 1985).

*Operational Definition Development.* Research conducted by Ellsworth et al. (1993) had two objectives. The first was to determine if informed individuals were able to consistently and accurately rate levels of drowsiness. This study, which was described previously, resulted in validation of the average observer rating measure (AVEOBS) which was used as both a predictor and detector of drowsiness in the second part of the study by Ellsworth et al. (1993).
A second objective in the Ellsworth et al. (1993) study was to develop one or more operational definitions of drowsiness that reflect physiological changes normally associated with drowsy driving. Through correlational analysis, physiological measures that reliably detected performance impairment were found. Multiple regression analysis was conducted on these physiological measures to determine linear relationships between measures to predict driver impairment due to drowsiness. The results showed that a good definition of drowsiness could be obtained using the measures of eyelid closure, simple EEG, and simple heart rate. This new definition of drowsiness is called NEWDEF. In addition to the development of NEWDEF, the results showed that linear combinations of detection measures could reliably detect performance impairment due to drowsiness.

Algorithm Development. Wreggit et al. (1993) conducted a study which examined the physiological measures identified as reliable measures in the Ellsworth et al. (1993) study, along with performance measures that have been supported in the literature as potential drowsiness indicators. The purpose of this research was to develop algorithms that could be used to detect drowsiness in drivers.

Sleep deprived subjects were asked to begin driving the simulator shortly after 12:00 A.M. They were instructed to maintain a speed of 60 mph and stay within the boundaries of the simulated lane. The "drive" continued for 2.5 hours. This is the time frame that was found by Wierwille et al. (1992) to be the most realistic scenario for fatigued drivers. Physiological and performance measures, including subsidiary task responses, were recorded.

Analysis of the collected data showed five dependent measures which could be accurately predicted by various combinations of driver-related and performance-related measures. These operational definitions of drowsiness are described below.
AVEOBS - the average drowsiness rating of three observers for each one minute interval. These data were collected via videotape and viewed after the experimental session.

EYEMEAS - the mean square of the percentage of the subjects' eye closure.

PERCLOS - the proportion of time that a subject's eyes were closed 80% or more.

NEWDEF - the operational definition of drowsiness developed by Ellsworth et al. (1993):

\[ Y = 18.45722(\text{PERCLOS}) - 0.01569(\text{MNALPHA}) + 0.020173(\text{MNTTHETA}) - 0.00549(\text{MNBETA}) + 0.000698(\text{MNSQHRT}) \]

MASTER - the sum of the standardized values of AVEOBS, EYEMEAS, NEWDEF, and PERCLOS.

Six-minute averages were calculated to produce a drowsiness level value for each minute. These levels were classified as "Awake", "Questionable", or "Drowsy". Discriminate analysis resulted in a low number of large misclassifications, thus indicating that the algorithms were quite accurate in detecting drowsiness. However, Wreggit et al. (1993) point out that false alarms (2-3%) are likely to occur. A two-stage detection system, first using performance measures and then subsidiary task response measures, is suggested to eliminate the problem of false alarms (Wreggit et
Algorithm Validation. A second phase of the Wreggit et al. (1993) study was conducted to test the validity of the newly developed drowsiness detection algorithms by applying them to a new set of data collected from a new group of subjects. Several algorithms were chosen for validation. One class employed steering and lateral accelerometer measures while the other class employed steering, lateral accelerometer, and lane-related measures. These algorithms were chosen for two reasons. First, they contained both reliable and attainable measures, and secondly they allowed for a step-up, step-down procedure. That is, if lane related signals are not being picked up accurately, the detection system can “step-down” to the algorithm employing only steering and lateral accelerometer measures. Conversely, when valid lane related signals become available the system would then “step-up” to the algorithm employing all three of the measures (steering, lateral accelerometer, and lane-related measures). Results from the validation of these algorithms using driver-vehicle performance measures showed there was no significant loss in drowsiness prediction accuracy when they were applied to the new data (Wreggit et al., 1993).

These algorithms are ready to be implemented into a drowsy driver detection system. Now the focus of research must move to the sequence of events which will occur after drowsiness has been detected in order to re-alert the driver.

Usability

Usability is a very important consideration when designing systems for use by the general public. A user must be able to use the system correctly and with few to no problems. Usability tests can be conducted at all stages of development of a system. They are used to identify and correct problems throughout the iterative design process (Nielsen, 1993). There are several different ways the usability of a system can be tested.
A brief overview of usability testing is presented below.

A common practice in usability testing is the use of verbal protocol. When using verbal protocol, a subject is asked to "think aloud" as he or she uses the system in question. This allows the experimenter to follow the subject's train of thought and possibly obtain more useful information than if the subject were not speaking. There are two types of verbal protocol, concurrent and retrospective. During concurrent verbal protocol the subject verbalizes his or her thoughts while performing the tasks at hand. Retrospective verbal protocol requires the subject to verbalize his or her thoughts after the task has been completed, possibly while watching his or her performance of the task on video. In a recent study concurrent protocol was shown to facilitate task performance in software evaluation (Wright and Converse, 1992). Another study found subjects in the retrospective verbal protocol condition produced more valuable statements than those subjects in the concurrent verbal protocol condition (Ohnemus and Biers, 1993).

Three methods of usability testing were compared in a study conducted by Virzi, Sorce, and Herbert (1993). The three types were:

- Heuristic Evaluation - experts critique the user interface.
- Think-Aloud Evaluation - naive subjects comment on their thought processes as they use the system.
- Performance Test - error rates and task completion times are recorded as the subject interacts with the system.

The results of this study showed that the heuristic evaluation uncovered 81% of the problems present in a prototype of a voice mail system, while the performance test uncovered 46% and the think-aloud evaluation uncovered 69%. The heuristic evaluation also was found to be less expensive and less time consuming.
Heuristic evaluation is defined by Nielsen (1993) as a "systematic inspection of a user interface design for usability". The compliance of an interface with recognized usability principles is judged by evaluators. These principles, or heuristics, include using simple dialogue, minimizing user memory load, maintaining consistency, and having proper feedback. Subjects in the present study were graduate students in Human Factors Engineering. This is to say that they possessed a more detailed knowledge of warning systems and human machine interfaces, such as that of a drowsy driver detection system, than the average person. In a sense they can be considered "experts".

It is possible that one evaluator could uncover the major problems in a design. However, it has been reported that only 35% of usability problems were discovered by a single evaluator and approximately 87% were discovered when ten evaluators were employed (Nielsen, 1993). Because different evaluators will uncover different problems it is best to have several evaluators and compile their findings (Nielsen, 1993).

A study conducted by Virzi (1992) investigated the appropriate number of subjects in usability testing. The results revealed that most severe problems are uncovered with the first few subjects and 80% of all usability problems are uncovered with 4 or 5 subjects. The results also showed there was little benefit to having many more than 5 subjects since the amount of new information revealed by each subject decreased with each additional one.

Lewis (1994) conducted a study which examined additional considerations concerning sample size for usability testing. Lewis (1994) refutes the results of the Virzi (1992) study which concluded that 80% of the problems associated with a system were uncovered by 4 or 5 subjects. Results showed that the likelihood of problem detection effects the percentage of problems found (Lewis, 1994). In the Virzi (1992) study the likelihood of problem detection was 0.32 to 0.42. However, Lewis (1994) found that with an average likelihood of problem detection of 0.16, a sample size of 10 subjects
would be needed to uncover 80% of the problems. This finding suggests that it is important that usability evaluators are aware of the likelihood of problem detection for the particular system they are testing, so that they may more accurately estimate sample size needed.

The Method of Paired Comparisons

There are many instances in research when subjective measures, such as personal preferences and ratings, are sought. These measures are often difficult to interpret because there is no physical scale on which to place them. Many times psychological scales and psychophysical laws are used to scale and interpret these subjective data.

The method of paired comparisons is considered to be one of the most useful of all the commonly used scaling techniques (David, 1963). It is used when stimuli being compared have no measurable characteristics, other than subjective ratings, which would indicate which stimuli are judged to be greater on some psychological dimension (Gescheider, 1985). In the method of paired comparisons, subjects are presented with all possible pairs of stimuli and the subject's preference between the two stimuli in each pair is sought. There are several objectives which should be met when using this method. Every stimulus should appear in each presentation position as often as the other stimuli. No stimulus should be presented in two successive pairs. The position of each stimulus should be alternated between pairs. In other words, if stimulus A is presented first in the first pair, it should be presented second in the next pair in which it appears.

This procedure is derived from Thurstone's Law of Comparative Judgment (Guilford, 1954). This law states that the psychological scale values for two stimuli can be calculated from the proportion of times one stimulus is preferred over the other.
(Gescheider, 1985). Certain assumptions are made based on this law. First, it is assumed that subjects’ responses follow a normal distribution. Second, the stimuli being compared do not effect a subject’s response to one particular stimulus. Finally, the stimuli are equally easy to place on the scale of the attribute being measured.
METHODS OF ALERTING

The goal of a warning is to change a person's behavior. To be effective, warnings must be sensed, received, understood, and heeded (Sanders and McCormick, 1993). Thus, the warning must attract a person's attention, convey the correct message, and suggest the correct action to be taken.

Warning Systems for Motor Vehicles

When designing warning systems for use in motor vehicles, as opposed to aircraft or other systems, there are specific aspects which need to be examined carefully. When designing crash avoidance warnings for vehicles, the time available to react must be considered, and the alarm must be designed so that it conveys the appropriate level of urgency. Other aspects that should to be considered include variations in vehicles and drivers. The warning system must be compatible with many different types of vehicles. Another variation lies in the potential users of these warning systems. People of all ages drive. They possess different physical, sensory, and cognitive abilities. Some drive alone and some drive with passengers. These differences will impact the effectiveness of an alarm, and should be considered when designing a warning system.

General warning device design. Preliminary human factors guidelines for crash avoidance warnings have been developed (Lerner, Kotwal, Lyons, and Gardner-Bonneau, 1993). These guidelines suggest a general picture of an ideal warning system, and a summary of these guidelines is presented below.

An effective warning signal must be intrusive and convey a sense of urgency. However, warning systems in vehicles must not be so intrusive and urgent that they startle the driver, and possibly put him or her in more danger, or annoy the driver to the point that he or she will deactivate the system. At the same time, the warning must not be
so conservative that it fails to result in the desired effect of alerting a driver of an approaching danger.

Multiple levels of warnings should be used to alleviate the activation of urgent false alarms. The most urgent alarms should be saved for imminent crash warning signals, with less urgent and less disturbing signals signifying earlier stages. The urgent alarms used for imminent crash warnings should also be unique. A certain frequency or pitch should be reserved for use only in a case of immediate danger.

Alarms signifying immediate danger should be presented through two sensory modalities to improve their likelihood of being received. Because of differences in drivers' perceptions and differences in driving environments, dual modality alarms are likely to be more effective. For example, an auditory tone combined with a haptic stimulus is likely to be more effective over a broad range of drivers than just a tone.

Warning systems should ordinarily be activated automatically every time the vehicle is started. However, in the case of motor vehicles, it is probably necessary for the driver to activate the detection/warning system for legal reasons. Nevertheless, drivers should be able to deactivate the system since it is possible that there will be unforeseeable situations when the system is not necessary, and if left engaged, would produce false or nuisance alarms.

Built-in diagnostic tests should be implemented with the detection/warning system. These tests should be engaged every time the system is activated. If the built-in diagnostic test detects a system failure the driver should be notified immediately. However, this information should not be displayed in such a way that the driver could mistake the warning for one requiring a more immediate response.

Multiple settings should be available on warning alarms. These settings should be adjustable. When the vehicle is started the current settings should be obviously displayed or the settings should default to a predetermined setting. It is very likely that multiple
drivers will drive one vehicle and possibly in different environments. For example, settings may be changed for night time driving or driving in noisy environments, such as driving in a convertible with the top down or with the radio at a high volume. In situations like these it is important that the driver be aware of the alarm settings and assure that they are in the most effective mode for the specific driver and environment.

**Driver alertness warnings.** There are many possible ways to alert a driver in a drowsy condition. These include auditory displays such as tones, buzzers, rumble sound, and speech. Other possibilities include vibrations of the steering column or driver's seat. The maintenance of the driver's alertness is also necessary. Possible methods of maintaining a driver's alertness level are seat vibration, supply of fresh air, driver involvement in a secondary task, and the use of a stimulating scent, such as peppermint (Kaneda et al., 1994). Another possible method of maintaining a driver's alertness is the use of a lane-minder. This is a concept developed by Wierwille in 1989 which consists of sensors located in the vehicle which are able to sense the boundaries of the lane. If the vehicle exceeds those boundaries an alarm is activated.

Warning systems specifically designed for alerting drowsy drivers should follow the general guidelines presented above, as well as additional specifications which are presented below.

**Auditory displays.** Auditory displays are generally preferred for their effectiveness in alerting (Horowitz and Dingus, 1992). However, special precautions must be taken to avoid startling or distracting the driver. Therefore, onset rates of 10 dB/msec and higher should be avoided. Also, it is suggested that sounds coming from a single area be avoided unless they are consistent with the direction of the hazard (Lerner et al., 1993).

The fundamental frequencies used in acoustic warnings should be in the range of 500-3000 Hz, and frequencies easily masked by the ambient noise of the environment
should be avoided (Lerner et al., 1993). Edworthy, Loxely, and Dennis (1991) identified sound characteristics which increase the perceived urgency of a warning signal. These characteristics include a high repetition rate, high intensity, high fundamental frequency, and large frequency oscillations (warbling, for example).

When considering speech warnings there are several areas of concern. Voice displays are not capable of conveying a message as quickly as other modes of alerting. Also, care must be taken so that the warning voice is discernible from that of a passenger or radio messages. For these reasons, speech warnings may be most effectively used as part of a combination of two or more types of warnings.

_Tactile displays._ Subjects participating in driving simulator studies have shown a greater reduction in their level of alertness than subjects involved in road studies. Simulators, for the most part, have a smoother ride and less road vibration than an actual car. This suggests that vibration may have an alerting effect and for this reason, should be considered a viable warning signal for driver drowsiness. It is suggested that tactile displays, such as vibration, be located in the driver's seat or the steering column. The vibration frequencies of such warnings, according to Lerner et al. (1993), should be in the range of 100-300 Hz. However, these frequencies are much higher than vibration frequencies normally considered to affect operators. Therefore, lower frequencies may be appropriate.

_Visual displays._ A visual display is suggested to be used as an initial signal in a warning system. However, in a drowsy driver alerting system the use of visual displays is problematic. If a driver is drowsy and inattentive he or she may be less likely to perceive a visual warning in time to react appropriately. Also, most stimuli presented during driving are presented through the sensory channel of vision. Use of a visual display may overload this channel, and possibly distract the driver. If a visual signal is used as part of a warning, it should be presented within 15 degrees of the driver's normal
line of sight of the roadway (Lerner et al., 1993).

*Termination of warnings.* Warnings which are automatically triggered by a specific condition should be presented for at least one second and until the triggering condition no longer exists (Lerner et al., 1993). This time period should give the driver enough time to recognize the purpose of the alarm. Immediately following the termination of the alarm the system should be reactivated in order to detect any quickly reoccurring decrease in alertness. Some warnings may require manual termination by the driver. The mode of termination in these cases should not be too easily accessed. It is possible that the termination of warning signals could become habit, just as in pressing a snooze button on an alarm clock. For this reason, the termination control should require some physical motion on the part of the driver.

**Related Research**

Many accidents involving drowsy drivers occur when the driver allows the car to drift off the side of the road. These accidents have been labeled Drift-Off-Road (DOR) accidents. A study of police accident reports in the 1980's showed this type of accident was the leading contributor to the total number of accidents on the Pennsylvania Turnpike. Fifty-seven percent of all accidents on the Turnpike in 1986 were classified as DOR accidents (Wood, 1994). Because of the frequency of DOR accidents, alerting drivers who are drifting became an area of interest.

A strip of patterned pavement along the shoulder of a road called a rumble strip has been effective in reducing the number of DOR accidents. If a driver drifts off the road, onto a shoulder having a rumble strip, the tires on the grooved or uneven pavement produce a loud sound and vibration. The California Department of Transportation tested a rumble strip pattern on a monotonous road between Las Vegas and Los Angeles. A 45% reduction in DOR accidents was reported after the
installation of the rumble strip (Chaudoin and Nelson, 1985). The Pennsylvania Turnpike began installing their version of rumble strips, SNAPS (Sonic Nap Alert Patterns), in 1987. Data collected between 1 year and 3.5 years after the installation of the SNAPS showed a 70% reduction in DOR accidents (Wood, 1994).

Although these results indicate that rumble strips are quite effective, there are some negative issues which need to be addressed. The installation of rumble strips in every shoulder of every road would be very time consuming and expensive. Also, there is the problem of only one side of the road being equipped. Because of intentional lane changes (i.e. passing), the inside boundary of the lane is more difficult to equip with rumble strips. However, a simulated rumble strip effect (a combination of vibration and a rumble sound) produced from within the vehicle when a drowsy driver has been detected may prove to be a very effective countermeasure to drowsiness.

This idea has been investigated in recent research. Daimler Benz (1994) has developed an image processing system for lane position recognition. This process takes into consideration a vehicle's initial position, steering angle, and speed to determine the path the vehicle is predicted to follow if no adjustments are made. The Time to Line Crossing (TLC) is then calculated. If this time is one second or less a warning signal is activated.

Three types of warnings were investigated: acoustic, haptic, and corrective haptic. The acoustic warning used was a simulated rumble strip sound emanating from either side of the vehicle depending on which side of the lane the vehicle was close to exceeding. The haptic warning was steering wheel oscillation, and the corrective haptic warning was steering wheel oscillation combined with a corrective pull of the steering wheel. In other words, if the vehicle was moving toward the right side of the lane the steering wheel would be pulled slightly to the left to initiate
correction.

The Daimler Benz (1994) study resulted in an overall positive judgment concerning the acoustic warning by subjects. Subjects also found the acoustic warning "enlivening". This suggests an acoustic warning is not only effective at immediately alerting a drowsy driver but it may have a sustained alerting effect as well.

The use of scents as a countermeasure to drowsiness was investigated in a study conducted by Kaneda et al. (1994). The refreshing effect of four scents, lavender, lemon, jasmine, and peppermint, were compared. The results showed that peppermint had the greatest refreshing effect. Kaneda et al. (1994) attributed this finding to the menthol found in peppermint. The stage of alertness at which the scent was introduced had an effect on the lasting refreshing effect of the scent. When the scent was introduced to a subject who was already experiencing a low level of alertness, its effect was brief in comparison to when the scent was introduced immediately following the first signs of a decrease in alertness. When a buzzer was sounded immediately before the release of the scent, the refreshing effect was extended from approximately 3 minutes to approximately 11-16 minutes (Kaneda et al., 1994). Although the Kaneda et al. (1994) results are impressive, there was no statistical analysis of the data, and the experimental method used was not clearly described. Therefore, further investigation into the use of scents as a method of alerting drowsy drivers is needed.
PRESENT STUDY

Research Objectives

This study was directed at evolving the potentially best configuration of advising/alarming stimuli to be used in a drowsy driver detection, advising, and alarming system. The proposed system consists of three stages. This research focused on optimizing the second and third stages of the system. A detailed diagram of the system is shown in Figure 1.

Stage one - Initial detection of reduced alertness level (drowsiness). The first stage involves the use of performance algorithms to detect a decrease in a driver's alertness level. The algorithms used would be those developed and validated by Wierwille et al. (1994). "Step-up/step-down" procedures would be used, thereby allowing the detection system to continue to function during intervals when not all parameters are available. For example, two algorithms could be used, one employing steering and lateral accelerometer measures, and the other employing steering, lateral accelerometer, and lane-related measures. When lane-related measures are available the algorithm using all three types of measures would be used. If lane-related measures were not available for some reason, such as lane markers not being on all sections of road, the detection system would "step-down" to the algorithm that does not utilize the lane-related measures and would thereby remain effective.

Once the system is engaged the algorithms would compute an alertness level using six-minute averages updated every minute. The first twelve minutes of driving would be used for establishing a baseline. Thereafter, alertness would be computed and evaluated each minute (using six-minute moving averages). As long as there is no reduction in alertness level of the driver the system would remain in stage one. Once the system detects that the driver's alertness level has fallen below a predetermined threshold,
Figure 1A: System Flow Diagram for Initial Detection of Drowsiness (Stage 1)
Figure 1B: System Flow Diagram for Re-alerting Driver (Stage 2)
Figure 1C: System Flow Diagram for Maintaining Alertness (Stage 3)
the system will progress to stage two. In the mean time, however, stage one processing
would continue.

Stage two - Re-alerting the driver. The second stage would begin with an
auditory stimulus informing the driver that a drowsiness condition has been detected. A
full alarm would then be activated unless the driver manually depresses the reset button.
The advisory stimulus would consist of an audible tone followed by a voice message. If
cruise control is engaged at this point it would be disengaged by the detection system.
One of the objectives of this research was to determine the optimal tone and voice
message to be used for this advisory stimulus. The option to reset the system would give
the driver the opportunity to avoid unnecessary exposure to full alarms. However, the
reset button should be placed in an area that would not allow deactivation to occur too
easily. When the driver depresses the reset button the initial alert would be disengaged
for six minutes. After the reset button is depressed the algorithms will most likely still be
detecting a noticeable reduction in alertness level. Therefore, the six-minute delay in the
system is intended to avoid an immediate re-activating of the initial alerting tone and
voice message after the driver has depressed the reset button.

If drowsiness is detected in the first stage and the driver does not reset the system
after the initial advisory tone and voice message (at the beginning of the second stage), a
subsequent full alarm would be sounded. There are many possible configurations for this
alarm. Therefore, another objective of this research was to determine the optimal alarm
to be used at this stage of the advising/alarming system. Characteristics of auditory
displays, specifically sounds with a wide variety of wave shapes, frequencies, and
modulations, were investigated. The effects of vibration of the steering wheel and the
driver's seat, and a simulated braking effect were also investigated.

The alarm would continue until it is manually deactivated by the driver. Again,
deactivation would be effected through the depression of the reset button. When the reset
button is depressed the system would again delay re-activating of the initial advisory tone and voice message at the beginning of stage two for six minutes. Whenever the driver depresses the reset button, stage two would end and stage three would begin.

**Stage three - Maintaining alertness.** At the beginning of this stage, a voice message would advise the driver to engage an alertness aid to help maintain the driver's re-alerted state. If the driver feels that an alertness aid is needed to help remain alert while looking for a safe rest area, there may be several methods from which to choose. Possibilities would include seat vibration, introduction of a secondary task, activation of a lane-minder, use of eye glasses which measure slow eye closures, introduction of a stimulating scent, and directing fresh air toward the driver's face. The driver would select countermeasures by means of a control panel. Appendix A contains a description of its functions.

**Seat vibration.** When the seat vibration button is depressed momentarily the button would be illuminated and the seat would vibrate. When the button is depressed again the vibration would cease. It is possible to have vibration in the seat back and seat pan. The current study was directed at determining the most effective combination.

**Scent.** When the scent button is depressed the fresh air blower would be activated and a scent would be discharged into the air in front of the driver. This button would be illuminated along with the fresh air button. The blower would continue to operate for a predetermined length of time to disperse the scent. The subjective effectiveness of the use of a scent was tested in the current study.

**Lane-minder.** When the lane-minder button is depressed it would be illuminated and the lane-minder would be activated. If the driver allows the vehicle to exceed the lane boundaries an alarm would be activated. The lane-minder would only be able to be activated when the A/O task (see next page) is deactivated and vice versa. The lane-minder would be able to be deactivated by depressing the lane-minder button again.
Lane-minder effectiveness was also tested in the current study.

**A/O task.** When the A/O task button is depressed the button would be illuminated and the A/O task would begin. This task involves presentation of recorded words presented aurally to the driver. The driver responds using yes/no push buttons mounted on the steering wheel spokes. If the presented word contains the letters "a" and/or "o", the yes button is to be depressed. Otherwise the no button is to be depressed. During this task the monitoring system would employ the A/O task algorithms developed by Wreggit, Kern, and Wierwille (1993). If a decrease in the driver's alertness level is detected the A/O task would cease and the system would revert back to the beginning of stage two and the initial advisory tone and voice message would be sounded. The subjective effectiveness of the A/O Task was also tested in the current study.

Although this research concentrates on the investigation of these aforementioned countermeasures, there are other possible ways to help maintain a driver's alertness level, as described below.

**Fresh air.** The fresh air button would activate the fresh air blower.

**Glasses.** An ordinary pair of eyeglass frames equipped with a source which produces a beam of infrared light and sensors which detect when the beam of light is interrupted by a slow eyelid closure would be used. The measure of slow eyelid closure, PERCLOS, would be calculated, and if a high degree of slow eyelid closure is detected the system would revert back to the initial advisory tone and voice message at the beginning of the second stage. PERCLOS would continue to be calculated until detection is made or the countermeasure is deactivated by the depression of the glasses button on the driver's control panel.

The purpose of these alertness aids is to maintain the increased alertness level of the driver, achieved in stage two, until the driver can safely pull off the road and refresh.
Method

There are many unanswered questions regarding the interface of such a system. The purpose of this study was to answer as many of these questions as possible. Because of the many variables involved in a study of this kind, subjective usability testing, specifically a type of heuristic evaluation, utilizing rating scales and the method of paired comparisons appeared to be the best approach. As stated in the literature review, heuristic evaluation has been shown to be the most effective in uncovering major problems. Heuristic evaluation is also the least time consuming and least expensive with respect to other types of usability testing (Virzi et al., 1993).

The method of paired comparisons was used to allow the subjects to rate the different alarm stimuli. This method was chosen for several reasons. The subjects were presented with auditory stimuli which are likely to remain in memory for only a short period of time. There are statistical manipulations that may be performed on data obtained through the method of paired comparisons which allow the data to be analyzed as scaled values. Also, it was of some interest as to whether paired comparisons would result in the same rank order of the eight alarm sounds as straight effectiveness ratings.

Because subjects were driving the simulator while they were asked to rate the different stimuli, they gave their responses orally. A copy of the effectiveness scale can be seen in Figure 2. The scale was illuminated and located on the lower right side of the simulator's dash throughout the experiment. Subjects were asked to rate each stimulus as "Not Effective", "Slightly Effective", "Moderately Effective", "Very Effective", or "Extremely Effective". They were also permitted to rate stimuli as falling between these descriptors.

Subjects. Eight graduate students (four males and four females) in the Industrial and Systems Engineering Department at Virginia Tech were used as subjects. Each had
Figure 2: The Effectiveness Rating Scale Used in the Experiment.
completed at least one semester of graduate work in Human Factors Engineering. Subjects were selected from this group because of their familiarity with behavioral methods and because they possessed expertise related to the experimental topic. All subjects were licensed drivers.

**Apparatus.**

*Simulator.* The moving-base automobile simulator at Virginia Tech, validated by Leonard and Wierwille (1975), was used in the study. The simulator handles like a mid-sized rear wheel-drive vehicle. It is computer controlled and is hydraulically powered. It has four degrees of physical motion (roll, yaw, lateral translation, and longitudinal translation). The ambient noise level in the simulator is adjustable. However, during the experiment it was set at approximately 75.5 dBA with the sound level meter 5 inches from the right ear of the driver (73.1 dBA with the sound level meter 3 inches from the right ear of the driver). This level was based on a small road study which determined the ambient noise level of five different vehicles. The average noise level at 65 MPH was found to be 76.2 dBA, while the average at 55 MPH was found to be 74.6 dBA (Table 1). The instructed speed in the experiment was 60 MPH. Further explanation of this road study can be found in Appendix B.

<table>
<thead>
<tr>
<th>Speed</th>
<th>dBA</th>
</tr>
</thead>
<tbody>
<tr>
<td>45 MPH</td>
<td>72.53</td>
</tr>
<tr>
<td>55 MPH</td>
<td>74.61</td>
</tr>
<tr>
<td>65 MPH</td>
<td>76.22</td>
</tr>
</tbody>
</table>
Initial Advisory Tone. The initial advisory tone was produced from a function generator. The length of this tone was controlled by the experimenter through an interval timing device which allowed the adjustment of the length of the tone from 0.3 to 2.75 seconds.

Voice Message. The voice messages were recorded on a custom designed digital voice recording/playback system. The system is composed of four channels, each capable of recording and playing 10 seconds of audio information. Each of the channels can be played back on command and can be repeated indefinitely without delay.

Alarm Sounds. The audible sounds of the full alarm were produced from various function generators. These alarm sounds were recorded on and played back from a Sound Blaster 16 sound card installed on an IBM 433 DX/S PC.

The audio signals from the initial advisory tone, the voice message system, and the sound card were combined in a stereo audio mixing console. The outputs of the console were passed through dual power amplifiers and applied to small, high quality speaker enclosures containing 4-inch woofers and 1-inch tweeters with a frequency response range of 100 Hz - 20 kHz. These speakers were located at mid dash height, approximately 35° to the right and left of the straight-ahead position.

Vibration. Vibration was produced in the driver's seat and the steering wheel. Eccentrics (unbalanced rotational masses) located in the seat pan and the seat back were driven by high-quality servomotors to produce the vibration of the driver's seat. The frequency of these vibrations was adjustable. The steering vibration was produced by a periodic signal applied to the active servo system that provides steering feel. A more detailed description of the vibration mechanisms as well as the other peripheral stimuli is presented in Appendix C.

Simulated Brake Pulse. A simulated braking effect was achieved by activating a switch at the investigator's station. When the switch was activated, the simulator
momentarily lurched backward to produce the feeling of braking. The speed also decreased somewhat as would be the case in an actual vehicle when the brake pedal is momentarily depressed.

**Lane Minder.** The simulator was equipped with a lane sensing device which was capable of activating an alarm when the simulated vehicle exceeded the lane boundaries. The alarm consisted of a chirping/beeping sound produced by piezo buzzers.

**Scent.** The peppermint scent was supplied through an aerosol spray can. A small amount of peppermint oil was placed in a pressurized, reusable (repressurizable) can. Vapors from the can were then dispensed in front of the driver.

**Experimental Design.** The study was a within subject design. All subjects were exposed to all conditions. Stimuli were tested in three areas. In the second stage of the advising/alarming system (first phase of this research), the presence of a stimulus is necessary to indicate to the driver that he or she has been detected as exhibiting drowsiness. This stimulus was a tone followed by a voice message. The voice message was “Possible drowsiness has been detected; press reset now”. General features of the initial advisory tone and voice message that were tested were as follows:

**Tone:**

- Frequency (high, mid-high, mid-low, and low)
- Wave shape (rectangular and sinusoidal)
- Amplitude
- Duration

**Voice message:**

- Gender
- Amplitude
Eight candidate alarm sounds were also investigated. A detailed description of each alarm sound used is presented in Appendix D. These alarms varied in several characteristics including:

- Center frequency
- Wave shape
- Maximum amplitude
- Modulation characteristics

In addition to the alarm sounds, various tactile (haptic) stimuli were investigated. A detailed description of these stimuli is presented in Appendix C. They included the following:

- Steering wheel vibration
- Seat back vibration
- Seat pan vibration
- Brake pulse

In the second phase of the study, the effectiveness of four drowsiness countermeasures was investigated. These countermeasures were:

- Seat vibration
- Lane minder
- A/O Task
- Scent

(It should be noted that the levels and frequencies of seat vibration were lower for the countermeasures phase than for the alarm phase. See Appendix C.)

**Procedure.** Each subject was instructed to awaken at 7 am the morning of the experiment. He or she was asked to go about normal daily activities with the exception of taking naps. At 6 pm the subject was picked up at home by an experimenter and taken out to dinner. The subject was allowed to eat whatever he or she wanted. However, the
subject was not allowed to drink caffeinated or sugared beverages with his or her meal.

After dinner the subject was brought to the Vehicle Analysis and Simulation Laboratory and asked to read and sign an informed consent form. The subject was also asked to read the instructions to the study (Appendix E). A simple auditory test was then given, consisting of an experimenter and subject standing 3 feet apart with their backs facing each other. The experimenter spoke five or six words in a normal speaking voice and the subject was asked to repeat the words back to the experimenter. This test ruled out any severe hearing loss. Any questions or concerns the subject might have were addressed at this point. The subject was asked to remain in the lab until the experiment began. The subject could study or watch television.

At 11:45 pm the subject was asked to get into the simulator. The subject was given some final instructions and then asked to drive the simulator for approximately 5 minutes to become familiar with the simulator. After this period the subject was asked to exit the simulator and walk around in order to become acclimated to the simulator. The subject then re-entered the simulator and the possible advising and alarming stimuli were explained in more detail. Each step of the advising/alarming system was explained to the subject, including where it fit into the envisioned system and the desired effect of each particular stimulus. The subject was asked to maintain a speed of 60 MPH and continue driving while performing all tasks.

The initial advisory tone was played first. The subject was asked to optimize and rate eight different tones. The eight tones consisted of combinations of four different frequencies and two wave shapes. The wave shapes used were a rectangular wave and a sinusoidal wave. The eight tones were counterbalanced across subjects (Appendix F). Each subject was asked to set the duration of the tones, based on the first tone. Because of counterbalancing this was a different tone for each subject. Each of the eight tones was played and the subject was asked to adjust the amplitude of each to a preferred level.
The subject was permitted to adjust the duration and the amplitudes a second time in a repeat presentation. The tones were then played at the subject's preferred amplitude and duration, and the subject was asked to rate their effectiveness in advising a driver of a subsequent alarm.

The voice message was played next. There was a male and a female voice. Four of the subjects received the male voice first, and the other four received the female voice first. The appropriate amplitude for each voice was first chosen by the subject. He or she was then asked to rate the effectiveness with which each voice conveyed the message.

Once the initial advisory tone and voice message were set, the experiment focused on the full alarm. There were eight alarm sounds. The subject was first asked to set the amplitude for each alarm sound. The first alarm sound was played and the subject was asked to adjust the amplitude. The next alarm sound was played and again the subject was asked to adjust the amplitude. The amplitudes for all eight alarm sounds were set in this fashion. This process was repeated a second time to give the subject a chance to readjust the final chosen amplitude for each alarm sound. Once the amplitudes were set the method of paired comparisons was used to determine the preferred alarm sound. There were 28 possible pairs. The alarm sound pairs were presented in a different order for each subject so as to minimize possible order effects. Each alarm sound was played at its preferred amplitude during the comparisons. Once the paired comparisons were completed the study concentrated on possible peripheral stimuli. (Peripheral stimuli are those that could be combined with an alarm sound to produce a more effective alerting or awakening stimulus.)

Peripheral stimuli included steering vibration, seat back vibration, seat pan vibration, and a simulated brake pulse. The subject was exposed to each stimulus individually and asked to rate the effectiveness of the stimulus at re-alerting the driver. Next, the subject was re-exposed to the alarm sounds and he or she was asked to rate the
effectiveness of these sounds (by themselves) at re-alerting a drowsy driver. Once the subject had been exposed to both the peripheral stimuli and the alarm sounds, the alarm sounds were played in combination with each individual stimulus. The subject was asked to rate the effectiveness of each alarm sound combined with each peripheral stimulus. Each particular subject was presented with the same order of alarm sounds each time they were played during his or her experimental session. However, the order of presentation was different for each of the 8 subjects. The order in which the peripheral stimuli were produced also differed for each of the eight subjects in order to correct for any systematic effects.

In the next phase of the experiment, the subject was asked to focus attention on possible drowsiness countermeasures. These countermeasures included possible combinations of seat back vibration and seat pan vibration as well as the use of a lane-minder, the introduction of a secondary task (A/O task), and exposure to a peppermint scent. The subject was exposed to each of the five countermeasures individually and then asked to rate the effectiveness of each countermeasure in maintaining the driver's re-alerted state. It should be noted that the seat vibration countermeasure cues used lower amplitudes and frequencies than those used for the peripheral alarm stimuli (Appendix C).

Data Analysis Overview

Four types of data were collected during the experiment. These included data obtained from the effectiveness rating scales and data collected from the paired comparisons, as well as volume and duration data. All subjects' responses were recorded manually by the experimenter. All volume data were originally recorded as values from 0 - 10 on a volume control. Sound level measurements of these values were taken after subject participation and the corresponding dBA levels were used in the analyses of the
volume data. A summary of the sound pressure levels for all stimuli used in this study is presented in Appendix G.

Assumptions. Several assumptions underlie the use of analysis of variance. These include the assumption of normality and homogeneity of variance. These assumptions are often violated in the real world (Vasey and Thayer, 1987). Fortunately the analysis of variance $F$ statistic is known to be robust with regard to violations of these assumptions (Lindquist, 1953). A Monte Carlo study (a study which looked at the frequency distribution of the $F$ statistic over several experiments) conducted by Norton (1952), reported by Lindquist (1953), showed when using a 5% rejection region that even the most deviant population distributions resulted in a rejection region no larger than $\alpha = 0.08$. From this study it may be concluded we need only be concerned with flagrant violations of the assumption of normality when $F$ values fall extremely close to the start of the rejection region.

Monte Carlo studies have also investigated the effect of violating the assumption of homogeneity of variance. The results of these studies indicate that even quite large differences among variances have little effect on $F$ values (Keppel, 1982). Rogan and Keselman (1977) suggest that a ratio of the largest to smallest variance of as large as 10 might cause a researcher to be concerned if the $F$ value falls close to the chosen critical value. Yet another study conducted by Glass, Peckham, and Sanders (1972) found that non-normality skewness and heterogeneous variances had very little effect on the level of significance of the $F$ test. In light of these findings, an analysis if variance was used to analyze the effectiveness rating scale data collected during this experiment.

$T$-test. T-tests were used to determine if there was a significant difference between the first and final chosen durations of the initial advisory tones. A t-test was also used to test the significance of the volume level and the effectiveness ratings of the two voice messages. Alpha levels were set at 0.05.
**Analysis of variance.** Analyses of variance (ANOVAs) were used to test for significant differences with respect to the volumes of the tones and alarms as well as for all of the effectiveness ratings ($\alpha = 0.05$). The effectiveness rating, excluding those of the voice messages, were taken using the scale shown in Figure 2. Subjects were permitted to select points in between descriptors if they wished. All variables were within subject, and a Greenhouse-Geisser correction was used when calculating the ANOVAs. All reported p-values are Greenhouse-Geisser corrected.

**Paired Comparisons.** The frequency of preference for each of the eight alarm sounds was recorded for each subject (Appendix H). The frequency data were converted into a proportion matrix. Maxwell's scaling technique was then used to transform the proportion data into an interval scale (Maxwell, 1974). This was achieved by converting the proportion data into $z$ -deviates of the normal distribution. The mean $z$ -values were then transferred to a scale on which the value of zero corresponds to the lowest mean $z$ -deviate. These data were then analyzed using a one-way analysis of variance with alarm being the single factor with eight different levels ($\alpha = 0.05$) (Maxwell, 1974).
RESULTS

Initial Advisory Tone

Three aspects of the initial advisory tones were investigated in this study. These included the preferred length and volume of the tone as well as how effectively it attracted the attention of a driver and directed it to a forthcoming voice message. Figure 3 graphically depicts the average ratings for each of the eight tones. Table 2 shows each tone's wave shape and frequency as well as the average chosen dBA level and effectiveness rating (0 being "Not Effective" and 4 being "Extremely Effective"). The results of the Newman-Kuels post-hoc tests for amplitudes and effectiveness ratings are shown in the columns directly to the right of those values. Significant differences are reported between tones having different letters within these columns. A Pearson's correlation was performed on the tone data. A high negative correlation ($r = -0.93$) was found between the chosen amplitudes and the effectiveness ratings.

Duration. A $t$-test was performed on the two groups of duration data (Subjects' first chosen duration compared to their final chosen duration). No significant differences ($p > 0.05$) were found between the two groups. The final average tone duration chosen by the subjects was 0.80 second.

Amplitude. A one-way analysis of variance was performed on the preferred sound levels of the initial advisory tones. Significant results were found ($p = 0.0001$). Newman-Kuels post-hoc tests were conducted to identify specific differences. The results show that there were many cases where there were significant differences in the chosen amplitudes of specific advisory tones.

Ratings. A one-way analysis of variance was performed on the numerical rating scale data, with a significant result ($p = 0.003$). The results of the Newman-Kuels post-hoc tests show several significant differences between the ratings of the eight different tones.
Figure 3: Mean Effectiveness Ratings for Initial Advisory Tones
Table 2: Summary Table of Initial Advisory Tone Data. Tones with same letter are not significantly different.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Wave Shape</th>
<th>Frequency (Hz)</th>
<th>Avg. dBA</th>
<th>N-K Test Amp.</th>
<th>Avg. Effect. Rating (0 - 4)</th>
<th>N-K Test Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone 8</td>
<td>rectangular</td>
<td>1000</td>
<td>65.150</td>
<td>A</td>
<td>3.121</td>
<td>A</td>
</tr>
<tr>
<td>Tone 6</td>
<td>rectangular</td>
<td>700</td>
<td>64.512</td>
<td>A</td>
<td>3.120</td>
<td>A</td>
</tr>
<tr>
<td>Tone 4</td>
<td>rectangular</td>
<td>400</td>
<td>65.800</td>
<td>A</td>
<td>2.706</td>
<td>B</td>
</tr>
<tr>
<td>Tone 7</td>
<td>sinusoidal</td>
<td>1000</td>
<td>71.975</td>
<td>B</td>
<td>2.665</td>
<td>B C</td>
</tr>
<tr>
<td>Tone 5</td>
<td>sinusoidal</td>
<td>700</td>
<td>71.700</td>
<td>B</td>
<td>2.460</td>
<td>B C D</td>
</tr>
<tr>
<td>Tone 2</td>
<td>rectangular</td>
<td>200</td>
<td>71.088</td>
<td>B</td>
<td>2.455</td>
<td>D</td>
</tr>
<tr>
<td>Tone 3</td>
<td>sinusoidal</td>
<td>400</td>
<td>73.288</td>
<td>B</td>
<td>1.874</td>
<td>E</td>
</tr>
<tr>
<td>Tone 1</td>
<td>sinusoidal</td>
<td>200</td>
<td>79.750</td>
<td>C</td>
<td>1.414</td>
<td>F</td>
</tr>
</tbody>
</table>

Voice Message

Two types of data were collected on the voice messages. These were preferred amplitude levels and effectiveness ratings. The results of the two analyses are reported below. A table of the average amplitudes and effectiveness ratings is presented in Appendix I. The possibility of a gender bias was also investigated. Half of the subjects were presented with the same gender voice message first while the other half was presented with the opposite gender voice message first. It was found that 3 out of the 4 male subjects rated the male voice as more effective and 3 out of the 4 female subjects rated the female voice as more effective. This result went against the expectation that the opposite gender voice would be rated as more effective than the same gender.

Amplitude. A t-test was conducted to determine if there was a significant difference between the preferred volume level of the male and female voice messages. The result was not significant (p > 0.05).
**Ratings.** A *t*-test was also conducted on the effectiveness ratings. Again, the result was not significant (*p* > 0.05).

**Alarm Stimuli**

Several groups of data were collected in this segment of the experiment. The preferred volume levels were collected for each of the eight alarm sounds. Data on the alarm sounds by themselves were collected through the method of paired comparisons. Effectiveness ratings were collected for the peripheral stimuli by themselves, the eight alarm sounds by themselves, and the alarm sounds coupled with each of the peripheral stimuli. The results of these analyses are presented in Tables 3, 4, 5, 6, 7, and 8.

**Amplitudes.** A one-way analysis of variance was conducted to determine if any significant differences existed between preferred dBA levels of the alarm sounds. Significant results were found (*p* = 0.0001).

**Paired comparison.** A one-way analysis of variance was conducted on the paired comparison data after it was converted into scaled values. These scaled values along with results of the Newman-Kuels post hoc tests are shown in Table 3. The scaled values were checked for additivity by comparing the statistic, $X^2 = \frac{f ns^2}{4} (f = 1/2(k-1)(k-2)$ degrees of freedom), to the chi-squared distribution. The results of this test gave a value of 29.946 which when compared to the chi-squared value for 21 df is not significant (*p* > 0.05). This result indicates that the scale possesses additivity and the usual variance ratio test can be used to test for significance. The F-test did produce significant results (*p* < 0.05) indicating that subjects had significant differences in preference among the eight alarms. The ANOVA summary table is presented in Table 4.

**Ratings.** A one-way analysis of variance was conducted on the ratings of the peripheral stimuli. No significant differences (*p* > 0.05) were found. The mean effectiveness ratings for the five peripheral stimuli presented by themselves are shown in
Table 3: Results of the Maxwell Scaling Procedure on the Frequency of Preference Data for the Alarm Sounds. Alarms with the same letter are not significantly different. Preference is from high (top) to low (bottom).

<table>
<thead>
<tr>
<th>Condition</th>
<th>Short name</th>
<th>Frequency of Preference</th>
<th>Scaled value</th>
<th>N - K Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alarm 2</td>
<td>On-Off Tone</td>
<td>50</td>
<td>4.0073</td>
<td>A</td>
</tr>
<tr>
<td>Alarm 1</td>
<td>Dual Tone Warble</td>
<td>42</td>
<td>3.2242</td>
<td>A B</td>
</tr>
<tr>
<td>Alarm 4</td>
<td>Freq. Swept Tone</td>
<td>38</td>
<td>3.0244</td>
<td>A B C</td>
</tr>
<tr>
<td>Alarm 7</td>
<td>Rapid Amp.-Mod. Tone</td>
<td>31</td>
<td>2.4316</td>
<td>B C D</td>
</tr>
<tr>
<td>Alarm 8</td>
<td>Gapped Freq.-Swept Tone</td>
<td>27</td>
<td>2.0404</td>
<td>C D</td>
</tr>
<tr>
<td>Alarm 6</td>
<td>Overmod. Mid-Freq. Tone</td>
<td>22</td>
<td>1.7343</td>
<td>D</td>
</tr>
<tr>
<td>Alarm 5</td>
<td>Overmod. Low-Freq. Tone</td>
<td>11</td>
<td>0.7083</td>
<td>E</td>
</tr>
<tr>
<td>Alarm 3</td>
<td>Continuous Tone</td>
<td>3</td>
<td>0.0000</td>
<td>E</td>
</tr>
</tbody>
</table>
Table 4: ANOVA Summary Table For Scaled Values of Alarm Sound Preferences

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Alarms</td>
<td>7</td>
<td>98.665</td>
<td>14.095</td>
<td>19.769*</td>
</tr>
<tr>
<td>Residual</td>
<td>21</td>
<td>14.9679</td>
<td>0.713</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>113.6329</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant test results (p < 0.05)
An 8 x 6 ANOVA was conducted to test the main effects and interactions of alarm type and peripheral stimulus condition (alone, with a brake pulse, with lower seat vibration, with upper seat vibration, with combination seat vibration, and with steering vibration).

Table 5 shows the summary table for this analysis. Significant results were found for the main effects of Alarm Type and Stimulus Condition as well as the interaction between the two. Newman-Kuels post hoc tests were used to isolate significant differences. Only logical comparisons are reported. (Unimportant comparisons such as a comparison between Alarm 3 with a brake pulse and Alarm 6 alone are not reported.)

Table 6 shows the break down of the significant main effect of Alarm Type. Significant differences are shown by different letters in the right hand column of the table. Table 7 shows the comparisons between each of the alarm sounds by themselves and combined with each of the peripheral stimuli. These comparisons were made to determine if the peripheral stimuli added to the effectiveness of the alarm sounds. As can be seen in the table, there are increases in several cases. The effects can also be seen in Figure 4. Comparisons were also made within each stimulus condition. These results are reported in Table 8.

Countermeasures

A one-way ANOVA was used to look for significant differences among the effectiveness ratings of the six countermeasures (lower seat vibration, upper seat vibration, combination seat vibration, A/O task, lane-minder, and the peppermint scent). No significant differences (p > 0.05) were found. The mean effectiveness ratings of the six countermeasures are shown in Appendix K.
Table 5: ANOVA Summary Table of Stimulus Condition (SC) x Alarm Type (AT)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F-value</th>
<th>p-value</th>
<th>G-G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject (S)</td>
<td>7</td>
<td>51.852</td>
<td>7.407</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC</td>
<td>5</td>
<td>26.197</td>
<td>5.239</td>
<td>6.118</td>
<td>0.0004</td>
<td>0.0048</td>
</tr>
<tr>
<td>SC x S</td>
<td>35</td>
<td>29.974</td>
<td>0.856</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT</td>
<td>7</td>
<td>40.325</td>
<td>5.761</td>
<td>5.532</td>
<td>0.0001</td>
<td>0.0074</td>
</tr>
<tr>
<td>AT x S</td>
<td>49</td>
<td>51.024</td>
<td>1.041</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC x AT</td>
<td>35</td>
<td>5.274</td>
<td>0.151</td>
<td>3.415</td>
<td>0.0001</td>
<td>0.0104</td>
</tr>
<tr>
<td>SC x AT x S</td>
<td>245</td>
<td>10.811</td>
<td>0.044</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6: Summary of the Analysis of the Effectiveness Ratings of Alarm Sounds Alone. Alarm sounds with the same letter are not significantly different.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Short name</th>
<th>Avg. dBA</th>
<th>N-K Test Amp.</th>
<th>Avg. Rating (0 - 4)</th>
<th>N-K Test Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alarm 2</td>
<td>On-Off Tone</td>
<td>76.575</td>
<td>A</td>
<td>3.462</td>
<td>A</td>
</tr>
<tr>
<td>Alarm 1</td>
<td>Dual Tone Warble</td>
<td>83.387</td>
<td>C</td>
<td>3.219</td>
<td>A B</td>
</tr>
<tr>
<td>Alarm 4</td>
<td>Freq. Swept Tone</td>
<td>81.000</td>
<td>B</td>
<td>3.191</td>
<td>B C</td>
</tr>
<tr>
<td>Alarm 7</td>
<td>Rapid Amp.-Mod. Tone</td>
<td>80.775</td>
<td>B</td>
<td>3.039</td>
<td>C D</td>
</tr>
<tr>
<td>Alarm 8</td>
<td>Gapped Freq.-Swept Tone</td>
<td>81.413</td>
<td>B</td>
<td>2.913</td>
<td>D E</td>
</tr>
<tr>
<td>Alarm 6</td>
<td>Overmod. Mid-Freq. Tone</td>
<td>77.725</td>
<td>A</td>
<td>2.850</td>
<td>E F</td>
</tr>
<tr>
<td>Alarm 5</td>
<td>Overmod. Low-Freq. Tone</td>
<td>77.012</td>
<td>A</td>
<td>2.608</td>
<td>F</td>
</tr>
<tr>
<td>Alarm 3</td>
<td>Continuous Tone</td>
<td>88.900</td>
<td>D</td>
<td>2.392</td>
<td>G</td>
</tr>
</tbody>
</table>
Table 7: Average Effectiveness Ratings and Results of Post-hoc Comparisons Between Each Alarm Type Alone and that Alarm Type with Different Stimulus Conditions. Conditions with the same letter within one alarm type are not significantly different.

<table>
<thead>
<tr>
<th>Alarm Type</th>
<th>Stimulus Condition</th>
<th>Avg. Rating (0-4)</th>
<th>N-K Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Alone</td>
<td></td>
<td>3.080</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Brake Pulse</td>
<td>3.539</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Lower Seat Vibration</td>
<td>3.121</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Upper Seat Vibration</td>
<td>3.122</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Combination Seat Vibration</td>
<td>3.330</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Steering Vibration</td>
<td>3.121</td>
<td>A</td>
</tr>
<tr>
<td>2 Alone</td>
<td></td>
<td>3.415</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Brake Pulse</td>
<td>3.749</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Lower Seat Vibration</td>
<td>3.331</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Upper Seat Vibration</td>
<td>3.330</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Combination Seat Vibration</td>
<td>3.663</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Steering Vibration</td>
<td>3.287</td>
<td>A</td>
</tr>
<tr>
<td>3 Alone</td>
<td></td>
<td>1.537</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Brake Pulse</td>
<td>2.914</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Lower Seat Vibration</td>
<td>2.287</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Upper Seat Vibration</td>
<td>2.414</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Combination Seat Vibration</td>
<td>2.787</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Steering Vibration</td>
<td>2.412</td>
<td>A</td>
</tr>
<tr>
<td>4 Alone</td>
<td></td>
<td>2.829</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Brake Pulse</td>
<td>3.497</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Lower Seat Vibration</td>
<td>3.081</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Upper Seat Vibration</td>
<td>3.164</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Combination Seat Vibration</td>
<td>3.414</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Steering Vibration</td>
<td>3.162</td>
<td>A</td>
</tr>
<tr>
<td>5 Alone</td>
<td></td>
<td>2.039</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Brake Pulse</td>
<td>3.038</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Lower Seat Vibration</td>
<td>2.537</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Upper Seat Vibration</td>
<td>2.579</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Combination Seat Vibration</td>
<td>2.997</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Steering Vibration</td>
<td>2.456</td>
<td>A</td>
</tr>
<tr>
<td>6 Alone</td>
<td></td>
<td>2.205</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Brake Pulse</td>
<td>3.287</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Lower Seat Vibration</td>
<td>2.787</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Upper Seat Vibration</td>
<td>2.829</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Combination Seat Vibration</td>
<td>3.122</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Steering Vibration</td>
<td>2.870</td>
<td>B</td>
</tr>
<tr>
<td>7 Alone</td>
<td></td>
<td>2.622</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Brake Pulse</td>
<td>3.374</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Lower Seat Vibration</td>
<td>2.956</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Upper Seat Vibration</td>
<td>2.997</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Combination Seat Vibration</td>
<td>3.329</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Steering Vibration</td>
<td>2.954</td>
<td>A</td>
</tr>
<tr>
<td>8 Alone</td>
<td></td>
<td>2.496</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>With Brake Pulse</td>
<td>3.207</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Lower Seat Vibration</td>
<td>2.871</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Upper Seat Vibration</td>
<td>2.912</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Combination Seat Vibration</td>
<td>3.204</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>With Steering Vibration</td>
<td>2.789</td>
<td>A</td>
</tr>
</tbody>
</table>
Figure 4: Mean Effectiveness Ratings for Alarm Sounds Alone and With Peripheral Stimuli
Table 8: Average Effectiveness Ratings and Results of Post-Hoc Comparisons Within Each Peripheral Stimulus Condition.

<table>
<thead>
<tr>
<th></th>
<th>Alarm Type</th>
<th>3</th>
<th>5</th>
<th>8</th>
<th>6</th>
<th>7</th>
<th>4</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alarms with Brake Pulse</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N - K Test</td>
<td>A</td>
<td>A B</td>
<td>A B C D</td>
<td>B C D</td>
<td>B C D</td>
<td>C D E</td>
<td>D E</td>
<td>E</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Alarm Type</th>
<th>3</th>
<th>5</th>
<th>6</th>
<th>8</th>
<th>7</th>
<th>4</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alarms with Lower Seat Vibration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. Rating (0 - 4)</td>
<td>2.287</td>
<td>2.537</td>
<td>2.787</td>
<td>2.871</td>
<td>2.956</td>
<td>3.081</td>
<td>3.121</td>
<td>3.331</td>
<td></td>
</tr>
<tr>
<td>N - K Test</td>
<td>A</td>
<td>A B</td>
<td>B C</td>
<td>C D</td>
<td>C D E</td>
<td>C D E F</td>
<td>C D E F</td>
<td>F</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Alarm Type</th>
<th>3</th>
<th>5</th>
<th>6</th>
<th>8</th>
<th>7</th>
<th>4</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alarms with Upper Seat Vibration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. Rating (0 - 4)</td>
<td>2.287</td>
<td>2.537</td>
<td>2.787</td>
<td>2.871</td>
<td>2.956</td>
<td>3.081</td>
<td>3.121</td>
<td>3.331</td>
<td></td>
</tr>
<tr>
<td>N - K Test</td>
<td>A</td>
<td>A B</td>
<td>B C</td>
<td>C D</td>
<td>C D</td>
<td>C D</td>
<td>D</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Alarm Type</th>
<th>3</th>
<th>5</th>
<th>6</th>
<th>8</th>
<th>7</th>
<th>1</th>
<th>4</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alarms with Combination Seat Vibration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N - K Test</td>
<td>A</td>
<td>A B</td>
<td>A B C</td>
<td>B C</td>
<td>B C</td>
<td>B C</td>
<td>C D</td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Alarm Type</th>
<th>3</th>
<th>5</th>
<th>8</th>
<th>6</th>
<th>7</th>
<th>1</th>
<th>4</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alarms with Steering Vibration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. Rating (0 - 4)</td>
<td>2.412</td>
<td>2.456</td>
<td>2.789</td>
<td>2.870</td>
<td>2.954</td>
<td>3.121</td>
<td>3.162</td>
<td>3.287</td>
<td></td>
</tr>
<tr>
<td>N - K Test</td>
<td>A</td>
<td>A B</td>
<td>B C</td>
<td>B C D</td>
<td>B C D</td>
<td>B C D</td>
<td>C D</td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>
DISCUSSION

Initial Advisory Tone

The results of the analyses of the initial advisory tone data demonstrate that particular characteristics are more effective at capturing the attention of an alert driver. As can be seen in Table 2, the higher frequency, rectangular wave tones were placed at lower preferred volume levels and rated higher on the effectiveness scale. The 1000 Hz and 700 Hz rectangular wave tones were placed in the "Very Effective" to "Extremely Effective" range on the rating scales. The sinusoidal waves were placed at much higher sound levels and still were only rated in the "Slightly Effective" to "Moderately Effective" range on the rating scales. As shown by the high negative correlation found in the results, the most effective tones are not necessarily the ones with the highest sound level. The characteristics of the tone such as wave shape and frequency have a larger impact on the effectiveness of a certain tone. Tones with frequencies nearer 1000 Hz and with a rectangular wave shape were rated as more effective. Explanation for this is that the human ear is less sensitive to frequencies as they decrease from 1000 Hz, and the rectangular wave shape gives the subjective impression of having a sharper, crisper sound due to its strong harmonics, and therefore requires a lower amplitude to attract a driver's attention. Tone 2 (rectangular wave at 200 Hz) is the only rectangular wave for which the chosen amplitude was greater than that of the sine waves. This is most likely due to its low frequency.

Voice Message

The analyses of the voice message data resulted in no significant differences. The average preferred peak amplitude for the voice messages was found to be 83.4 dBA and 82.5 dBA for the male and female voices, respectively. The male voice on average was
rated slightly higher (2.75) than the female voice (2.58). However, this difference is not statistically significant. This non-significant difference could partially be due to the small sample size used in this study (N = 8). Both the male or female voice seem to be quite effective at conveying the necessary information to the driver. However, a same gender voice preference was found. The majority of female subjects rated the female voice as more effective than the male, and the majority of the male subjects rated the male voice to be more effective than the female. As was stated in the results section, this was unexpected and more research is needed to determine why this occurred. A possible confound in regards to the voice messages is that the voices used in both messages were those of individuals with which all subjects were familiar. Therefore, subjects may have recognized the voice and partially based their judgments of effectiveness on this.

Alarm Stimuli

The effectiveness of the alarm sounds was investigated in two different ways. The ANOVA performed on the paired comparison data and the ANOVA performed on the rating scale data both produced significant results. Not only were these results significant, but they placed the eight alarm sounds in the same ranked order. The order of this rank can be seen in Tables 3 and 6. In both cases alarms #2 (on-off tone) and #1 (dual tone warble) are considered the most effective alarms. Both were placed well into the "Very Effective" to "Extremely Effective" range on the rating scales. In the paired comparisons, alarms #2 and #1 were preferred over the other alarms 50 and 42 times, respectively, out of a possible 56 times. Alarms #2 and #1 share several characteristics. They are high-frequency, square waves with alternating periods of no sound or a sound at a lower frequency. Both have a repetition frequency of 3 Hz.

The 2-way ANOVA of Alarm Type and Stimulus Condition also produced significant results. Post-hoc tests revealed that in almost all cases the addition of a
peripheral stimulus increased the effectiveness of a particular alarm. The brake pulse added to the effectiveness of every alarm sound. This can be seen clearly in Figure 4. The combination seat vibration also added to the effectiveness of most of the alarm sounds. An increased effectiveness was observed for all alarms with the combination seat vibration over the alarms by themselves. However, this increase was not significant in alarms #1 and #2. This lack of significance may be due in part to the high ratings given to these two alarms by themselves. Subjects may have run out of room on the scales. Another interesting finding involving alarm #2 is that three of the peripheral stimuli; lower seat vibration, upper seat vibration, and steering vibration were reported to have detracted slightly from the effectiveness of this alarm (Figure 4). This again may be due to the initial high rating of this alarm by itself.

**Countermeasures**

As stated above, no significant differences were found among the effectiveness ratings of the countermeasures. Again, the non-significant differences may be partially due to the small sample size used in this study (N = 8). For the most part the countermeasures were only rated between "Slightly Effective" and "Moderately Effective". The lane-minder and the scent were rated slightly higher than "Moderately Effective" (2.121 and 2.165, respectively). A possible reason for the low rating of the seat vibration is that many subjects expressed that they found it relaxing. Many subjects also found the A/O task monotonous and something which they could perform in a very relaxed and inattentive state.

**Conclusions**

This study succeeded in answering a wide variety of questions regarding alerting/alarming stimuli to be used in a drowsy driver detection system. Conclusions
drawn from the different areas of analyses in this study are presented below.

**Initial Advisory Tone.** The purpose of the initial advisory tone is to capture the attention of a driver and direct his or her attention to a subsequent voice message. The results of this study indicate the preferred length of this tone to be 0.8 second. The recommended characteristics for the initial advisory tone are a rectangular shaped wave at a frequency of 700 Hz to 1000 Hz with a dBA level of approximately 8.1 dBA below the ambient noise level in the vehicle (The ambient noise level of the vehicle at 60 mph during this experiment was 73.1 dBA). Because of the concentration of high frequencies this tone is still audible even though it falls below the surrounding noise level.

**Voice Message.** As stated in the discussion section, there were no significant differences found between the male and female voice messages. However, the results do show that the voice message used in this study ("Possible drowsiness has been detected: press reset now"), presented at a peak amplitude of approximately 10 dBA above the ambient noise level in the vehicle (average amplitude of approximately 5 dBA below ambient noise level), was consistently rated from "Moderately Effective" to "Very Effective". Therefore, it may be concluded that either a male or a female voice may be used effectively as part of a drowsy driver detection system.

**Alarm Stimuli.** Both analyses of the alarm sounds showed alarms #2 and #1 as very effective. These alarms are both high-frequency, square waves with alternating periods of no sound or a sound at a lower frequency. Both have a repetition frequency of 3 Hz. The results of this study support increased effectiveness when dual modality is incorporated in an alarm. It can be concluded that the most effective way to re-alert a drowsy driver is to expose the driver to an alarm sound with characteristics similar to alarms #2 and #1 at a amplitude of 3.5 dBA and 10 dBA (respectively) above the ambient noise level in the vehicle. This audible alarm should be accompanied by a simulated brake pulse to achieve maximum effectiveness. If brake pulses can not be implemented,
combined seat vibration can be used as an effective substitute.

Countermeasures. The results of this study found no countermeasures that were significantly different from one another. In addition, no investigated countermeasure rated much higher than "Moderately Effective". This leads us to conclude that further research may be needed. However, in the absence of further research, a lane-minder or peppermint scent system might be implemented. Both were rated as "Moderately Effective", and there is no guarantee that further research on other countermeasures will produce better results.
REFERENCES


Report #8402). Blacksburg, VA: Virginia Polytechnic Institute and State University, IEOR Department.


ISE Department.


APPENDIX A

Descriptions of Functions in Driver's Control Panel
Lighted Push Buttons

"Monitoring System On" (Momentary: Push-on, Push-off)
Activates drowsiness detection system.
Begins performance monitoring.
Provides voice caution to driver to use directional signals when changing lanes.
Lights the "Monitoring System On" push button.
Initializes drowsiness detection system (new baseline established).

"Drowsiness Detected" (This is a light only)
Flashes yellow when performance monitoring system detects drowsiness and reset push button has not yet been depressed.
   When reset is depressed momentarily, it goes to dull yellow for six minutes and then extinguishes.
Flashes yellow when A/O task monitoring system detects drowsiness and reset push button has not yet been depressed.
   When reset is depressed momentarily, it goes to dull yellow for six minutes and then extinguishes.
Flashes yellow when lane-minder senses out of lane condition. Extinguishes when vehicle re-enters the lane.

"Reset" (Momentary: Push-on, Push-off)
Usually dull orange. Bright orange when momentarily depressed or when "drowsiness detected" light flashes.
Steps the flashing of the "drowsiness detected" light when depressed momentarily.
Prevents or stops the alarm system from sounding/vibrating for a period of six minutes.
Initiates message to select countermeasures when depressed momentarily.
Cancels all activated countermeasures when depressed momentarily.
"Seat Vibration" (Momentary: Push-on, Push-off)
Activates both the push button light and the seat vibration system.
Both the light and seat vibration are extinguished by depressing reset button.

"Scent" (Momentary: Push-on, Push-off)
Activates the scent push-button light and causes scent to be injected into air in front of the driver.
Activates the "Fresh Air" light and fresh air blower.

"Fresh Air" (Momentary: Push-on, Push-off)
Activates the fresh air push button and fresh air blower.
If scent light is lit, "Fresh Air" button has no effect when depressed.

"A/O Task" (Momentary: Push-on, Push-off)
Initiates A/O task and push-button light.
Causes monitoring system to activate A/O task algorithms.
When deactivated, it causes the monitoring system to rely on performance monitoring only.
Prevents lane-minder system from engaging.

"Lane-Minder"
Initiates lane-minder system and push-button light.
Prevents A/O task system from engaging.
Figure A1: Driver's Control Panel for Alertness Monitor
APPENDIX B

Average Sound Pressure Level Study
Measurement of Sound Pressure Levels

Method

An Extech Instruments digital sound level meter was used to measure the sound pressure levels in five different vehicles. The sound level meter was operated in its slow response mode. The measurements were taken on the A-scale at 45, 55, and 65 MPH.

All measurements were taken while driving on the bypass section of Rt. 460 between North and South Main Streets in Blacksburg, Virginia. All readings were taken while in the highest gear.

All windows were closed. Any fan or air conditioning was turned off. Radios remained off. Neither driver nor experimenter spoke while measurements were taken.

For each run, the sound level meter was placed approximately 5 inches from the driver's right ear. Seven measurements were taken at each of the three speeds. These seven measurements were used to produce an average sound pressure level for each speed. The five vehicles used and the sound levels measured are shown in Table B1.
Table B1: Average Sound Levels for Individual Vehicles

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>45 mph</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>73.7</td>
<td>75.2</td>
<td>78.2</td>
<td>72.1</td>
<td>72.1</td>
<td>73.6</td>
</tr>
<tr>
<td>72.6</td>
<td>76.1</td>
<td>77.9</td>
<td>71.8</td>
<td>72.5</td>
<td>73.1</td>
</tr>
<tr>
<td>71.8</td>
<td>75.8</td>
<td>77.6</td>
<td>69.9</td>
<td>71.9</td>
<td>72.5</td>
</tr>
<tr>
<td>73.0</td>
<td>76.3</td>
<td>77.5</td>
<td>70.1</td>
<td>70.9</td>
<td>72.6</td>
</tr>
<tr>
<td>72.5</td>
<td>75.9</td>
<td>77.8</td>
<td>69.7</td>
<td>72.8</td>
<td>71.8</td>
</tr>
<tr>
<td>72.9</td>
<td>76.0</td>
<td>78.0</td>
<td>69.1</td>
<td>72.2</td>
<td>72.8</td>
</tr>
<tr>
<td>72.6</td>
<td>76.4</td>
<td>77.7</td>
<td>69.2</td>
<td>71.3</td>
<td>74.0</td>
</tr>
<tr>
<td>Averages</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72.7</td>
<td>76.0</td>
<td>77.8</td>
<td>70.3</td>
<td>72.0</td>
<td>72.9</td>
</tr>
</tbody>
</table>


APPENDIX C

Specifications for Peripheral Alarm Cues and Countermeasures
SPECIFICATIONS FOR ALARM PERIPHERAL CUES AND COUNTERMEASURES

STEERING VIBRATION

A 10 Hz square wave is applied to the servo system, whose output is torque applied to the steering wheel shaft.

For the alarm cue, the torque causes a 6.3 mm (0.25") peak-to-peak vibration at the outside rim of the steering wheel. This measurement is made at 60 mph with the hands removed from the wheel. The wheel is assumed to be 38 cm (15") in diameter (to the outside edge).

For the countermeasures cue, the torque causes a 2.0 mm (0.079") peak-to-peak vibration. Conditions for this measurement are the same as those for the alarm cue described above.

SEAT VIBRATION

Vibrators are installed in the seat pan (bottom) and in the seat back (opposite the L-1 point of the spine for an average height male). The vibrators are composed of eccentric masses driven by small servomotors that are run open loop (Figure D1). The motors are mounted on brackets that are attached to flexible lucite sheets. The sheets are 7" by 7" by 3/16" thick and are lashed to the seat springs. For the seat pan installation, the lucite is between the springs and the foam padding of the seat (which is approximately 1 1/2" thick). For the seat back installation, the lucite is against the back side of the foam pad because the springs are embedded in the foam pad. (The foam pad thickness for the seat back is also 1 1/2" thick.)

For the alarm cues, the upper eccentric spins at 1380 RPM and the lower spins at 1440 RPM.

For the countermeasures cues, the upper eccentric spins at 875 RPM and the lower spins at 1030 RPM.

The eccentric for the upper vibrator is a rectangular solid made of mild steel. It is 1/2" by 1/2" by 2" long. The shaft hole for the drive motor and the set screw to hold the eccentric are 1 3/8" from one end of the eccentric.

The eccentric for the lower vibrator is also a rectangular solid made of mild steel. It is 1/2" by 1/2" by 2 3/16" long. The shaft hole for the drive motor and the set screw to hold the eccentric are 1 9/16" from one end of the eccentric.

Both motors are mounted on brackets with their shafts parallel to the lucite sheets. The shafts are 1 13/16" from the lucite sheets.

The upper vibrator is mounted in the seat back with the motor shaft vertical (actually, tilted rearward slightly) and the eccentric at the top. The lower vibrator is mounted with the shaft horizontal and aligned with the longitudinal axis of the vehicle. The eccentric is mounted at the front end.
BRAKE PULSE

The brake pulse alarm cue is estimated to be applied for 0.5 sec and results in a 6 mph drop in speed from a cruising speed of 60 mph.

(Note that if the brake pulse alarm cue is implemented in an actual vehicle, the rear brake lights of the vehicle should begin flashing during the advisory tone and voice message. When the brake pulse is applied and for a period of time thereafter, the brake lights should then be activated continuously.)

SCENT

A repressurized (one pint) spray paint can (Jennican Products, Ltd., Guilford, Surry, U.K.) is used. The interior plastic hose is removed so that the can will spray a vapor and not the liquid from the bottom. One and one-half ounces of peppermint oil and an equal amount of water are poured into the can. After sealing, the can is pressurized to 85 psi.

When the scent is to be administered as a countermeasure, the experimenter shakes the can to mix the ingredients. The can is then handed to the subject who is instructed to dispense a burst of the vapor by directing the nozzle from right to left in front of the face. The subject may then inhale as much of the vapor from the single burst as he or she deems appropriate.

LANE MINDER

The lane minder is a countermeasure that provides an audible warning that the (simulated) vehicle is partly out of lane. The alarm begins with a continuous tone as soon as the vehicle begins to edge out of lane. As the deviation continues, the sound level increases and then begins “beeping.” Maximum sound level and beeping occur when the vehicle is three feet or more out of lane.

If the vehicle goes out of lane on the left, the sound emanates from a piezo-electric transducer to the left of the driver. If the vehicle goes out of lane on the right, the sound emanates from a piezo-electric transducer to the right of the driver. The transducers are mounted at mid-dash height, approximately 35° off to the left and off to the right of the driver’s straight-ahead position.

The frequency of each transducer is 2800 Hz and the beep frequency is 3 Hz. Sound level at maximum excursion is 73 dBA.

A/O TASK

This countermeasure is a subsidiary task presented auditorially to the driver. A recorded male voice presents a single word at each 15-second interval. The driver is to determine if the word contains the letter “A” or the letter “O”, or both. If it does, the driver depresses the “yes” pushbutton mounted on the right spoke of the steering wheel at the usual right thumb position. If it does not, the driver depresses the “no” pushbutton mounted symmetrically on the left at the usual left thumb position.
Stimulus words are selected so that they do not contain both an A and an O. Furthermore, they are selected so that presence or absence of an A or O is easy to determine.
Figure C1. Sketch of vibrator assembly
APPENDIX D

Technical Description of the Eight Alarm Sounds Used in the Experiment
Technical Description of the
Eight Alarm Sounds Used in the Experiment

All eight sounds are composed of rectangular waves with equal positive and negative
going durations that are modulated in some way.* In other words, if the modulation were
temporarily removed, the resulting wave would be a simple rectangular (square) wave.

Alarm 1. Short name: Dual Tone Warble

The waveform is at 800 Hz for 167 msec and then switches to 400 Hz for 167 msec.
(Repetition frequency is 3 Hz).

*Alarm 3 is unmodulated
Alarm 2. Short name: On-off Tone

The waveform is at 1000 Hz for 167 msec and then is off for 167 msec. (Repetition frequency is 3 Hz.)

\[
\begin{array}{c}
\text{1000 Hz} \\
\text{167 msec} \\
\text{1000 Hz} \\
\text{167 msec} \\
\text{167 msec}
\end{array}
\]

Alarm 3. Short name: Continuous Tone

The waveform is a continuous tone at 440 Hz. (There is no modulation.)

Alarm 4. Short name: Frequency Swept Tone

The waveform sweeps sinusoidally between 70 Hz and 1040 Hz. The sweep time from 70 to 1040 Hz is 250 msec, and the sweep time from 1040 to 70 Hz is 250 msec. (Repetition frequency is 2 Hz.)

\[
\begin{array}{c}
\text{1040 Hz} \\
\text{70 Hz} \\
\text{1040 Hz} \\
\text{250 msec} \\
\text{1040 Hz} \\
\text{70 Hz} \\
\text{250 msec} \\
\text{250 msec} \\
\text{250 msec}
\end{array}
\]
Alarm 5. Short name: Overmodulated low-frequency tone

The waveform is a sinusoidally amplitude-overmodulated 100 Hz tone. Peak amplitudes are 200 msec apart. (Repetition frequency is 5 Hz.)

Alarm 6. Short name: Overmodulated mid-frequency tone

The waveform is a sinusoidally amplitude-overmodulated 400 Hz tone. Peak amplitudes are 333 msec apart. (Repetition frequency is 3 Hz.)
Alarm 7. Short name: Rapid, amplitude-modulated tone

The waveform is at 400 Hz with the amplitude at full amplitude for 31 msec followed by the amplitude at 5% of full amplitude for 31 msec. (Repetition frequency is 16 Hz.)

Alarm 8. Short name: Gapped frequency-swept tone

This waveform sweeps sinusoidally from 70 Hz and 1040 Hz and back again. Then there is a short gap before the process is repeated. The sweep time from 70 to 1040 Hz is 194 msec, and the sweep time from 1040 to 70 Hz is 194 msec. The gap before repeat is 112 msec. (Repetition frequency is 2 Hz.)
APPENDIX E

Introduction to the Study and Informed Consent
Introduction to the Study

The purpose of this research is to investigate different characteristics of advising/alarming signals for use in a drowsy driver detection system in an automobile and to determine one or more appropriate configurations. The study is being conducted in the Vehicle Analysis and Simulation Laboratory, Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg. The research team consists of Sarah E. Fahey and Terry Fairbanks. The two researchers are graduate students in the Department of Industrial and Systems Engineering. Dr. Walter W. Wierwille is the Principle Investigator and Paul T. Norton Professor in the Department.

Your task will be to sit in an automobile simulator and drive as you would normally. The simulator will move so as to mimic the motions of an actual automobile. You will be asked to make preference choices and rate several different stages of the advising/alarming system. Phase one of the experiment will consist of three stages in which you will be asked to give your input. Phase two of the experiment will consist of only one stage.

The first stage of phase one will consist of advisory tones and voice messages. You will be asked to optimize and select parameters associated with each. For example, you will be asked to select the volume of each as well as other parameters such as the length of the advisory tones. You will also be asked to select which tones and voice messages are more effective at advising a driver of a subsequent alarm. It is important to understand that the intended purpose of the initial advisory tone and voice message is not to re-alert a drowsy driver but, rather, to inform a driver of a subsequent alarm. Therefore, you should keep in mind while setting the volumes that they do not need to be so loud as to get the attention of a driver who is not at least moderately alert. The volume should be set at a moderate level so as to capture the attention of a driver.

In some applications of the drowsiness detection system, drivers may not respond to the advisory tone and voice message. Therefore, the second stage of phase one will involve the selection and optimization of alarm sounds to re-alert the driver. Again, you will be asked to set the volume of the alarm sounds and other parameters. You will also be asked to select the alarms which you feel are most effective at alerting a drowsy driver. It is important to keep in mind that the intended purpose of these alarm sounds is to re-alert a drowsy driver. Therefore, in setting the volume you should be aware that the alarm sounds need to be loud enough to alert a drowsy driver but not so loud that they would startle the driver.

The third stage of phase one will deal with additional peripheral stimuli. You will be asked to rate these stimuli, such as steering and seat vibration, on how much they improve the effectiveness of the alarm stimulus complement.

The second phase of the study will focus on stimuli used to maintain a driver's alertness level. The purpose of this maintenance is so that the driver can stay alert for ten to fifteen minutes while driving to a safe place to stop and refresh. Stimuli which you will have experienced earlier, such as vibration, will be presented. You will, once again, be asked to rate the effectiveness of these stimuli. However, the intended purpose of these countermeasures is to maintain the driver's re alerted state achieved by the previous alarm sounds. Therefore, you will be asked to rate the degree of effectiveness with which these countermeasures will help maintain a driver's re alerted state for a period of approximately ten minutes.

If you decide to participate in this study, you must awake at 7 AM or before and go through your normal daytime activities without resting or napping. Then, at about 6 PM, a member of the experimental team will pick you up at your residence. This team
member will buy you dinner at a fast-food restaurant. You may eat whatever you like, but you will not be permitted to drink caffinated or sugared beverages, such as coffee or coke. If you are a smoker, you will be permitted to smoke right after dinner, but not thereafter. You will then be taken to the laboratory where you will be allowed to read, study, watch TV (which will be provided), or listen to your own personal headset stereo. You will not be permitted to eat, smoke, or drink caffinated coffee, or caffinated soft drinks, since these may effect the outcome of the experiment. However, you will be permitted to drink water or non-caffinated, diet soft drinks.

Shortly after midnight the experimental session will begin. It will last roughly two (2) hours. You will be asked to drive the simulator and experience various stimuli. Your opinion will be sought in the volume adjustment of these stimuli and alarms. Your preferences between the various stimuli will also be sought.

After the completion of the experiment, you will be paid and any remaining questions will be answered. If you participate in this experiment you must agree to let one of the experimenters drive you home, since they will not be drowsy at this time.

Payment for the experiment will be $5 per hour between 6 PM and midnight, and $8 per hour after midnight. If you complete the experiment you will receive approximately $46. If you decide to withdraw during the experiment or simply cannot continue for whatever reason, you will be paid for the time actually spent. Since the simulator is a complex system, equipment failures do occasionally occur. If this happens it may be necessary for the experimenters to terminate the experiment, in which case you will similarly be paid for the time actually spent.

Once you are seated in the simulator, you must not attempt to leave the simulator until you have given the experimenters a chance to stop the simulator and guide you in exiting.

Initially, you will be asked to take a simple hearing test and fill out a brief questionnaire on your normal sleeping/waking patterns and your normal eating/drinking/smoking (if any) patterns. If you qualify, you will then be scheduled for the experiment.

There are some minor risks and discomforts to which you will be exposed in this experiment. They are outlined in the attached informed consent form, which you should read carefully.
Participant’s Informed Consent

1. You are being asked to volunteer to be a subject in a research project whose purpose and description are contained in the document Introduction to the Study, which you have already read.

2. There are some minor risks and discomforts to which you expose yourself in volunteering for this research. The risks are:

   The risk of possible interference with your next day’s activities caused by less than a full night’s sleep. This risk can be minimized by sleeping longer than usual the morning following your participation.

   The risk of injury if you attempt to leave the simulator without the help of the experimenters. Please inform one of the experimenters if you feel that you must leave the simulator. You will then be guided out of the simulator.

   The discomforts are:

   Possible discomfort associated with trying to drive while tired or drowsy.

   Possible minor motion sickness due to the movement of the simulator.

3. The data gathered in this experiment will be treated with anonymity. Shortly after you have participated, your name will be separated from your data.

4. While there are no direct benefits to you from this research (other than payment), you may find the experiment interesting. Your participation and that of other volunteers should aid in the determination of appropriate drowsiness warning and alerting alarms.

5. You should not volunteer for participation in this experiment if you have known hearing impairment, are under 18 years old, if you are pregnant, if you are not in good health, or if you have any other condition which would adversely affect you by being sleep deprived and staying up until approximately 2 AM.

6. You should know that the principle investigator of the research project (Dr. Wierwille) and the research team will answer any questions you may have about your participation, and you should not sign this consent form until you are satisfied that you understand all of the previous descriptions and conditions. (Dr. Wierwille’s phone number is 231-7952).

   You should further be aware that you may contact Dr. E. R. Stout, Chair of the University’s Institutional Review Board, if you have questions or concerns about this experiment. Dr. Stout’s phone number is 231-9359.

7. You should know that at any time you are free to withdraw from participation in this research program without penalty for any reason.
8. You will be paid at a rate of $5.00 per hour before midnight and $8.00 per hour thereafter. If you complete your participation you will be paid $46.00. Payment will be made shortly after you have finished your participation.

9. You agree to allow one of the experimenters to drive you home following the experiment.

I have read and understand the scope of this research and I have no other questions. I hereby give my consent to participate, but I understand that I may stop participation at any time if I choose to do so.

Signature ____________________________

Date __________________________

(A copy of this informed consent form shall be given to the participant)
APPENDIX F

Tone Characteristics and Presentation Order
Table F1: Characteristics of Initial Advisory Tones.

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave shape</td>
<td>sine</td>
<td>rect.</td>
<td>sine</td>
<td>rect.</td>
<td>sine</td>
<td>rect.</td>
<td>sine</td>
<td>rect.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave shape</td>
<td>sine</td>
<td>rect.</td>
<td>sine</td>
<td>rect.</td>
<td>sine</td>
<td>rect.</td>
<td>sine</td>
<td>rect.</td>
</tr>
</tbody>
</table>

Table F2: Presentation Order of Initial Advisory Tones

<table>
<thead>
<tr>
<th>Tones</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
</tr>
<tr>
<td>Subject 2</td>
</tr>
<tr>
<td>Subject 3</td>
</tr>
<tr>
<td>Subject 4</td>
</tr>
<tr>
<td>Subject 5</td>
</tr>
<tr>
<td>Subject 6</td>
</tr>
<tr>
<td>Subject 7</td>
</tr>
<tr>
<td>Subject 8</td>
</tr>
</tbody>
</table>
APPENDIX G

Sound Pressure Levels for Stimuli and Ambient Noise levels during experiment
Table G1: Sound Pressure Levels for Initial Advisory Tones. All measurements made with digital sound pressure level meter in slow response mode (Simulator not running).*

<table>
<thead>
<tr>
<th>TONE 1 (200Sine)</th>
<th>TONE 2 (200Rect)</th>
<th>TONE 3 (400Sine)</th>
<th>TONE 4 (400Rect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>level</td>
<td>dBA</td>
<td>level</td>
<td>dBA</td>
</tr>
<tr>
<td>6.5</td>
<td>76.2</td>
<td>2.5</td>
<td>65.0</td>
</tr>
<tr>
<td>7.0</td>
<td>77.8</td>
<td>3.0</td>
<td>69.2</td>
</tr>
<tr>
<td>7.5</td>
<td>79.6</td>
<td>3.5</td>
<td>70.9</td>
</tr>
<tr>
<td>8.0</td>
<td>80.7</td>
<td>4.0</td>
<td>72.3</td>
</tr>
<tr>
<td>8.5</td>
<td>81.1</td>
<td>4.5</td>
<td>73.4</td>
</tr>
<tr>
<td>10.0</td>
<td>81.9</td>
<td>5.0</td>
<td>74.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TONE 5 (700Sine)</th>
<th>TONE 6 (700Rect)</th>
<th>TONE 7 (1000Sine)</th>
<th>TONE 8 (1000Rect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>level</td>
<td>dBA</td>
<td>level</td>
<td>dBA</td>
</tr>
<tr>
<td>3.5</td>
<td>69.3</td>
<td>2.0</td>
<td>60.2</td>
</tr>
<tr>
<td>4.0</td>
<td>70.5</td>
<td>2.5</td>
<td>64.6</td>
</tr>
<tr>
<td>4.5</td>
<td>71.6</td>
<td>3.0</td>
<td>67.9</td>
</tr>
<tr>
<td>5.0</td>
<td>72.2</td>
<td>3.5</td>
<td>69.4</td>
</tr>
<tr>
<td>6.5</td>
<td>75.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table G2: Sound Pressure Levels for Voice Messages. All measurements made with digital sound pressure level meter in slow response mode (Simulator not running).*

<table>
<thead>
<tr>
<th>MALE</th>
<th>FEMALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>level</td>
<td>Peak values**</td>
</tr>
<tr>
<td>2.5</td>
<td>78.8</td>
</tr>
<tr>
<td>3.0</td>
<td>82.7</td>
</tr>
<tr>
<td>3.5</td>
<td>84.7</td>
</tr>
<tr>
<td>4.0</td>
<td>86.3</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Measurements were made with the sound level meter 3 inches from the right ear of a typical driver.

**All analyses done with peak values.
Table G3: Sound Pressure Levels for Alarm Sounds. All measurements made with digital sound pressure level meter in slow response mode (Simulator not running)*.

<table>
<thead>
<tr>
<th>ALARM 1</th>
<th>ALARM 2</th>
<th>ALARM 3</th>
<th>ALARM 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>level</td>
<td>dB</td>
<td>level</td>
<td>dB</td>
</tr>
<tr>
<td>1.5</td>
<td>75.6</td>
<td>2.0</td>
<td>69.6</td>
</tr>
<tr>
<td>2.0</td>
<td>79.6</td>
<td>2.5</td>
<td>73.9</td>
</tr>
<tr>
<td>2.5</td>
<td>82.5</td>
<td>3.0</td>
<td>77.9</td>
</tr>
<tr>
<td>3.0</td>
<td>86.3</td>
<td>4.0</td>
<td>81.1</td>
</tr>
<tr>
<td>3.5</td>
<td>87.6</td>
<td>4.5</td>
<td>81.5</td>
</tr>
<tr>
<td>4.0</td>
<td>89.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ALARM 5</th>
<th>ALARM 6</th>
<th>ALARM 7</th>
<th>ALARM 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>level</td>
<td>dB</td>
<td>level</td>
<td>dB</td>
</tr>
<tr>
<td>3.0</td>
<td>71.0</td>
<td>2.5</td>
<td>71.4</td>
</tr>
<tr>
<td>3.5</td>
<td>73.3</td>
<td>3.5</td>
<td>76.8</td>
</tr>
<tr>
<td>5.0</td>
<td>77.0</td>
<td>4.0</td>
<td>78.2</td>
</tr>
<tr>
<td>5.5</td>
<td>78.3</td>
<td>4.5</td>
<td>79.4</td>
</tr>
<tr>
<td>6.0</td>
<td>79.6</td>
<td>6.0</td>
<td>83.2</td>
</tr>
<tr>
<td>8.0</td>
<td>86.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table G4: Sound Pressure Levels for Ambient Noise in Laboratory and Simulator Running.*

| Ambient noise in lab = 51 dBA |
| Simulator up and running: dBA (digital, slow response) |
| 55 MPH | 72.5 |
| 60 MPH | 73.1 |
| 65 MPH | 75.1 |

* Measurements were made with the sound level meter 3 inches from the right ear of a typical driver.
APPENDIX H

Preference Matrix and Maxwell's Scaling Procedure for Paired Comparison Data
Table H1: Preference Matrix for Paired Comparison Data.

<table>
<thead>
<tr>
<th></th>
<th>Alarm1</th>
<th>Alarm2</th>
<th>Alarm3</th>
<th>Alarm4</th>
<th>Alarm5</th>
<th>Alarm6</th>
<th>Alarm7</th>
<th>Alarm8</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alarm1</td>
<td>3</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td></td>
<td>42</td>
</tr>
<tr>
<td>Alarm2</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td>Alarm3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Alarm4</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>Alarm5</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>Alarm6</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td></td>
<td>22</td>
</tr>
<tr>
<td>Alarm7</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td></td>
<td>4</td>
<td>31</td>
</tr>
<tr>
<td>Alarm8</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>7</td>
<td>6</td>
<td>4</td>
<td></td>
<td>27</td>
</tr>
</tbody>
</table>

Table H2: Proportion Matrix for Paired Comparison Data.

<table>
<thead>
<tr>
<th></th>
<th>Alarm1</th>
<th>Alarm2</th>
<th>Alarm3</th>
<th>Alarm4</th>
<th>Alarm5</th>
<th>Alarm6</th>
<th>Alarm7</th>
<th>Alarm8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alarm1</td>
<td>X</td>
<td>0.375</td>
<td>0.875</td>
<td>0.75</td>
<td>0.875</td>
<td>0.75</td>
<td>0.75</td>
<td>0.875</td>
</tr>
<tr>
<td>Alarm2</td>
<td>0.625</td>
<td>X</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>0.875</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Alarm3</td>
<td>0.125</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>0.125</td>
<td>0.125</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Alarm4</td>
<td>0.25</td>
<td>0.25</td>
<td>1</td>
<td>X</td>
<td>1</td>
<td>0.75</td>
<td>0.625</td>
<td>0.875</td>
</tr>
<tr>
<td>Alarm5</td>
<td>0.125</td>
<td>0</td>
<td>0.875</td>
<td>0</td>
<td>X</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
</tr>
<tr>
<td>Alarm6</td>
<td>0.25</td>
<td>0.125</td>
<td>0.875</td>
<td>0.25</td>
<td>0.875</td>
<td>X</td>
<td>0.125</td>
<td>0.25</td>
</tr>
<tr>
<td>Alarm7</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
<td>0.375</td>
<td>0.875</td>
<td>0.875</td>
<td>X</td>
<td>0.5</td>
</tr>
<tr>
<td>Alarm8</td>
<td>0.125</td>
<td>0</td>
<td>1</td>
<td>0.125</td>
<td>0.875</td>
<td>0.75</td>
<td>0.5</td>
<td>X</td>
</tr>
</tbody>
</table>
Table H3: Matrix of z-deviates used in Maxwell’s scaling technique.

\[ z_{ij} = \log_e \{ \frac{P_{ij}}{(1-P_{ij})} \} - \text{ OR } z_{ij}^* = \log_e \{ \frac{(n_{ij}+1/2)}{(n-n_{ij}+1/2)} \} \]

* if P’s are 1 or 0

<table>
<thead>
<tr>
<th></th>
<th>Alarm1</th>
<th>Alarm2</th>
<th>Alarm3</th>
<th>Alarm4</th>
<th>Alarm5</th>
<th>Alarm6</th>
<th>Alarm7</th>
<th>Alarm8</th>
<th>Total (X)</th>
<th>( \chi^2 )</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alarm1</td>
<td>0.000</td>
<td>-0.511</td>
<td>1.946</td>
<td>1.099</td>
<td>1.946</td>
<td>1.099</td>
<td>1.946</td>
<td>1.099</td>
<td>1.946</td>
<td>8.623</td>
<td>74.352</td>
</tr>
<tr>
<td>Alarm2</td>
<td>0.511</td>
<td>0.000</td>
<td>2.833</td>
<td>1.099</td>
<td>2.833</td>
<td>1.946</td>
<td>2.833</td>
<td>2.833</td>
<td>14.888</td>
<td>221.659</td>
<td>1.861</td>
</tr>
<tr>
<td>Alarm3</td>
<td>-1.946</td>
<td>-2.833</td>
<td>0.000</td>
<td>-2.833</td>
<td>-1.946</td>
<td>-2.833</td>
<td>-2.833</td>
<td>-2.833</td>
<td>-17.171</td>
<td>294.829</td>
<td>-2.146</td>
</tr>
<tr>
<td>Alarm4</td>
<td>-1.099</td>
<td>-1.099</td>
<td>2.833</td>
<td>0.000</td>
<td>2.833</td>
<td>1.099</td>
<td>0.511</td>
<td>1.946</td>
<td>7.025</td>
<td>49.344</td>
<td>0.878</td>
</tr>
<tr>
<td>Alarm5</td>
<td>-1.946</td>
<td>-2.833</td>
<td>1.946</td>
<td>-2.833</td>
<td>0.000</td>
<td>-1.946</td>
<td>-1.946</td>
<td>-1.946</td>
<td>-11.504</td>
<td>132.346</td>
<td>-1.438</td>
</tr>
<tr>
<td>Alarm6</td>
<td>-1.099</td>
<td>-1.946</td>
<td>1.946</td>
<td>-1.099</td>
<td>1.946</td>
<td>0.000</td>
<td>-1.946</td>
<td>-1.099</td>
<td>-3.296</td>
<td>10.863</td>
<td>-0.412</td>
</tr>
<tr>
<td>Alarm7</td>
<td>-1.099</td>
<td>-2.833</td>
<td>2.833</td>
<td>-0.511</td>
<td>1.946</td>
<td>1.946</td>
<td>0.000</td>
<td>0.000</td>
<td>2.282</td>
<td>5.209</td>
<td>0.285</td>
</tr>
<tr>
<td>Alarm8</td>
<td>-1.946</td>
<td>-2.833</td>
<td>2.833</td>
<td>-1.946</td>
<td>1.946</td>
<td>1.099</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.847</td>
<td>0.718</td>
<td>-0.106</td>
</tr>
</tbody>
</table>

789.319

Table H4: Mean z-deviates and corresponding scaled values.

<table>
<thead>
<tr>
<th></th>
<th>Alarm 3</th>
<th>Alarm 5</th>
<th>Alarm 6</th>
<th>Alarm 8</th>
<th>Alarm 7</th>
<th>Alarm 4</th>
<th>Alarm 1</th>
<th>Alarm 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean z</td>
<td>-2.1463</td>
<td>-1.438</td>
<td>-0.412</td>
<td>-0.1059</td>
<td>0.2853</td>
<td>0.87807</td>
<td>1.07784</td>
<td>1.86103</td>
</tr>
<tr>
<td>scaled</td>
<td>0.0000</td>
<td>0.7083</td>
<td>1.7343</td>
<td>2.0404</td>
<td>2.4316</td>
<td>3.0244</td>
<td>3.2242</td>
<td>4.0073</td>
</tr>
</tbody>
</table>
APPENDIX I

Average Amplitude and Effectiveness Ratings for the Voice Messages
Table II: Average Amplitude and Effectiveness Ratings for the Voice Messages.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>83.4</td>
<td>68.4</td>
<td>2.75</td>
</tr>
<tr>
<td>Female</td>
<td>82.5</td>
<td>67.5</td>
<td>2.58</td>
</tr>
</tbody>
</table>
APPENDIX J

Average Effectiveness Ratings for Peripheral Stimuli
Table J1: Average Effectiveness Ratings for Peripheral Stimuli

<table>
<thead>
<tr>
<th>Peripheral Stimuli</th>
<th>Avg. Effect. Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper Seat Vibration</td>
<td>2.496</td>
</tr>
<tr>
<td>Lower Seat Vibration</td>
<td>2.789</td>
</tr>
<tr>
<td>Combination Seat Vibration</td>
<td>2.997</td>
</tr>
<tr>
<td>Steering Vibration</td>
<td>2.581</td>
</tr>
<tr>
<td>Brake Pulse</td>
<td>2.914</td>
</tr>
</tbody>
</table>
APPENDIX K

Average Effectiveness Ratings for Countermeasures
Table K1: Average Effectiveness Ratings for Countermeasures

<table>
<thead>
<tr>
<th>Countermeasure</th>
<th>Avg. Effect. Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Seat Vibration</td>
<td>1.540</td>
</tr>
<tr>
<td>Upper Seat Vibration</td>
<td>1.289</td>
</tr>
<tr>
<td>Combination Seat Vibration</td>
<td>1.662</td>
</tr>
<tr>
<td>A/O Task</td>
<td>1.580</td>
</tr>
<tr>
<td>Lane Minder</td>
<td>2.121</td>
</tr>
<tr>
<td>Peppermint Scent</td>
<td>2.165</td>
</tr>
</tbody>
</table>