A COMPARISON OF THREE SUBSIDIARY TASKS USED AS DRIVER DROWSINESS COUNTERMEASURES

by

Helen Lenora Hardee

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Industrial Engineering and Operations Research

APPROVED:

W. W. Wierwille, Chair

J. G. Casali

R. D. Dryden

P. T. Kemmerling

E. S. Geller

December, 1985
Blacksburg, Virginia

A COMPARISON OF THREE SUBSIDIARY TASKS USED AS DRIVER DROWSINESS COUNTERMEASURES

by

Helen Lenora Hardee

Committee Chair: Dr. Walter W. Wierwille
Industrial Engineering and Operations Research

(ABSTRACT)

Two previous studies performed at Virginia Tech have shown that it is feasible to detect drowsy drivers using driving performance and physiological measures. Therefore, assuming that drowsiness can be detected, it becomes important to develop methods (countermeasures) by which drivers can regain and maintain alertness. The current study was thus undertaken in an attempt to evaluate three subsidiary tasks which differed only in regard to input modality (auditory, tactual, or visual) in terms of: 1) the degree to which they aided the driver by maintaining or restoring alertness; and 2) the degree to which the responses to these tasks could be used to detect drowsiness. Subjective measures of drowsiness were also obtained to provide an additional source of verification of level of drowsiness.

To accomplish these objectives, a total of 12 male and female driver-subjects drove a moving-base simulator continuously from 12:30 a.m. to 3:00 a.m. During this time, the subjects performed each of the subsidiary tasks for a 30-minute period; they also drove for a 30-minute period during which no subsidiary task was performed. During the simulated, nighttime, highway driving scenario, 20 driving performance, behavioral, and physiological measures were collected for each 3-minute driving interval, along with 5 subsidiary task measures and subjective alertness ratings.

The experimental results indicated that none of the three subsidiary tasks provided an effective means of maintaining driver alertness. However, the results of a second series of discriminant analyses did indicate that driver impairment due to drowsiness could be reliably detected with linear combinations of subsidiary task and driving measures. In fact, promising discriminant models for the auditory and visual tasks were identified which employed a subsidiary task response measure of the number of correct responses to the subsidiary task during each 6-minute driving interval as well as a physiological measure of the subject's heart rate variance; these models showed overall classification error percentages as low as 3% and 8%. Finally, the analyses of the subjective alertness ratings indicated that subjects' ratings were not significantly affected by either the type of subsidiary task performed or time-on-task.

ACKNOWLEDGMENTS

I am deeply grateful to Dr. Walter W. Wierwille for the support, guidance, patience, and expertise which he offered during the course of this experimental endeavor. His suggestions, knowledge, and encouragement were instrumental to the successful completion of this project. Special thanks are also due my committee members, Dr. John G. Casali, Dr. Robert D. Dryden, Dr. E. Scott Geller, and Professor Paul T. Kemmerling, for their helpful comments and insights.

This research was sponsored by the Transportation Research Department of General Motors Research Laboratories. I thus extend my thanks not only to GM for providing the funding necessary to conduct this research, but also to , , and for their support and helpful suggestions.

Additionally, I would like to extend special thanks to and , fellow graduate students and dear friends, for the help they gave me during the course of this research. Thus, thanks go to for developing the software to collect and reduce data, and for assisting me with data collection. Also, I appreciate the aid provided by sleep depriving subjects. More importantly however, I am thankful to both and for the emotional support and friendship which they extended to me during the past year.

Finally, words cannot express the deep gratitude and love I feel for my parents, for the encouragement and support they have unselfishly given to me during my graduate career. Without that support, I could have never undertaken and completed this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Driving an automobile is a very complex task which requires that the operator constantly attend and respond to the driving environment to maintain proper control of the vehicle. The complexity of this task is one which is very often taken for granted, however, since for most drivers the task is overlearned and commonplace. Nonetheless, given the importance of driver vigilance and control to proper and safe vehicle operation, there are many factors which may adversely affect driver performance and which may consequently result in injury or death. One such factor is driver drowsiness.

While some researchers maintain that driver drowsiness plays a very important causal role in vehicular accidents (Hulbert, 1972), the extent to which this factor is involved is extremely difficult to assess. First, drowsiness is generally difficult to define operationally. Erwin (1976) has defined drowsiness or intrusive sleep as

> an undesirable decrease in arousal level to the point that satisfactory performance is no longer possible. This altered state must by definition be unwanted by the person affected, and intrude upon his activities despite efforts to reverse the process (p. i).

While this represents a step in the right direction toward identifying and defining drowsiness, it also points to a second problem associated with the study of the effect of driver drowsiness on accident rates, namely, the assessment of driving performance decrements associated with driver drowsiness. As will be pointed out later, driver drowsiness may not necessarily be always associated with a decrement in driving performance. In fact, it is quite possible for the driver to maintain vehicular control while actually suffering from short bouts of sleep as long as the operator maintains sufficient control to keep the vehicle on the roadway. Thus, potential accidents due to drowsiness may often go unreported. This would certainly serve to deemphasize the contributory role of driver drowsiness in traffic accidents.

Nonetheless, some researchers have estimated that drowsiness, along with fatigue, may directly cause as many as 35-50% of all highway fatalities (Hulbert, 1972; Kearney, 1966). In Virginia alone in 1983, 18% of the total number of highway crashes were attributed to driver inattention (Virginia Department of State Police, 1983). It is quite certain that some of these accidents were caused by inattention due to drowsiness (Jones, Kelly, and Johnson, 1978).

That drowsiness is a problem for a large percentage of the driving population has been verified by Tilley, Erwin, and Gianturco (1973) in a driver's survey conducted by Duke University in Durham, North Carolina. These researchers administered questionnaires dealing with driving habits to 1500 successful license renewal applicants. The results revealed that 64% of those sampled reported that they had become drowsy while driving. Additionally, of those who had become drowsy while driving, 69% said that drowsiness while driving was associated with eating, 70% said that they could identify a diurnal pattern of drowsiness while driving, and 91% reported that drowsiness while driving was associated with sleep loss (Tilley et al., 1973).

Tilley et al. were also able to assess the extent to which drowsiness while driving may be a factor in auto accidents or near-accidents. For instance, of those respondents who reported experiences of drowsiness while driving, nearly 10% reported that they had had near auto accidents due to drowsiness or falling asleep. More than 10% of this group also reported that they had actually been involved in accidents as a result of drowsiness or falling asleep at the wheel.

There were other findings of the Duke University driving study which are particularly relevant to the current study. First, of those drivers who reported experiencing drowsiness while driving, 31.2% indicated that the onset of drowsiness occurred before they were aware of it. This finding has been verified in other experimental settings (Jones et al., 1978; Safford and Rockwell, 1967). Second, 93.7% of these drivers believed that they could do something to alert themselves without stopping the vehicle if they became drowsy while driving.

The results of the Duke University driver population study (Tilley et al., 1973) have been further supported by the findings of Jones et al. (1978). In an extensive emissions experiment involving 352 volunteer drivers at the GM Proving Ground, Jones et al. reported that, of the 219 respondents who completed sleep questionnaires, 87% admitted that they became drowsy while driving. Of these, 37.5% also reported that they were unaware of drowsiness onset. Once again, a substantial percentage of the respondents (88.3%) felt that they could do something to alert themselves without having to stop the car.

Given these results on subjective reports of the occurrence of drowsiness during driving and the inability of many drivers to recognize sleep onset (Jones et al., 1978; Tilley et al., 1973), the conclusion is that driver drowsiness represents an immediate hazard. These findings thus support the need for further research which would allow for the identification of sleep onset and provide a means by which drivers could be kept alert.

Drowsiness Detection Devices

Drivers have been forced to deal with drowsiness and its detrimental effects on driving behavior and safety for many years. As a result, several types of drowsiness countermeasures have been developed or suggested. These include: 1) federal regulations for professional drivers; 2) simple remedies such as singing or chewing gum; and 3) mechanical and electronic detection and warning systems.

Federal regulations for professional drivers. Many countries, including the United States, have recognized the safety hazards associated with drivers who may travel for a prolonged time in their vehicles. To decrease the risk associated with prolonged driving, many countries have imposed federal regulations which limit the maximum number of hours which can be driven per day by professionals (McDonald, 1981). As McDonald pointed out, however, these regulations may not be very effective since they are seldom enforced and hence, seldom obeyed.

Simple remedies. Other, simpler remedies have also been offered as effective countermeasures of drowsiness.

For example, Harris (1967; cited in Forbes, 1972, p. 292) lists "singing, chewing a pack of gum, taking off the right shoe, and sitting on something hard" as some methods of alleviating drowsiness. The problem with these simple solutions to drowsiness lies in the fact that, while these countermeasures may work for some individuals, no reliable evidence has been collected which substantiates the value of these methods (Harris, 1967).

Mechanical and electronic detection and warning systems. One other approach to counteracting the effects of driver drowsiness on performance involves the use of a detection system which will not only detect driver impairment due to drowsiness but will also alert the driver to his or her drowsy condition. These types of devices are not new. In fact, some of them have been in use for nearly 25 years. However, there are problems associated with most of these which may preclude their effectiveness as warning devices. Several examples of these devices are given below, along with a brief discussion of their limitations.

1. The Electronic Transistor Safety Alarm. This is a plastic device which wraps around the driver's ear and buzzes when the driver's head nods. Unfortunately, the operation of this device is based upon the false assumption that the head nods as soon as alertness is lost. Therefore,

this device may sound an alerting signal too late (Hulbert, 1972).

2. Button Steering Wheel Alarm. This device requires the driver to maintain constant pressure on a button mounted on the steering wheel to prevent an alarm from sounding. Since this device requires constant pressure by a finger or thumb, fatigue may result. Additionally, inadvertant activation of the alarm may result when the hands are moved to execute a turn (Hulbert, 1972).

3. ALERTMASTER. This device consists of a pedal to the left of the clutch which must be continuously depressed by the left foot. If pressure is released, the horn sounds. The assumption upon which the effectiveness of this device relies is that the left foot of the driver relaxes as drowsiness occurs (Williams, 1966; cited in Hulbert, 1972).

4. Alert-O-Matic. This device produces a series of three alerting signals of increasing severity. First, a light flashes on; the alerted driver turns it off by lightly tapping the horn. If the light fails to get the driver's attention after five seconds, the device activates the car horn. If, after three more seconds the driver has still not been aroused, the device turns the ignition on and off for five seconds. After five additional seconds, the ignition is shut off completely. At any point in the sequence, the

cycle may be terminated by depressing the horn. At least two problems are associated with this system. First, no systematic attempt was made to validate the variable times associated with each of the system's stages. Second, since the light flashes on every 60 seconds it is possible that the driver may adapt to the system (Frederick, 1966; cited in Hulbert, 1972).

5. Electronic Advisor. More recently Nissan (Onboard electronic adviser, 1984) has developed an electronic warning device which detects impaired driver performance based upon the driver's steering behavior. The system monitors driving performance for the initial 10 minutes of driving and then uses these data as a baseline to which all subsequent driving performance is compared. If erratic steering behavior which is typical of drowsiness or fatigue is detected, both visual and auditory warning signals are activated. There is at present no empirical evidence which supports the validity or reliability of this detection device. Moreover, the literature which deals with the use of steering performance as an indicator of driver drowsiness offers mixed support. Therefore, it seems that further research is indicated before this warning device can be confidently endorsed as a reliable and effective drowsiness countermeasure.

While these aforementioned drowsiness countermeasures have been offered by their inventors as a means of increasing driving safety by decreasing the probability of the automobile driver falling asleep behind the wheel, it seems that these devices (except perhaps the one developed by Nissan) do not take into account the many dependent variables which have more recently been studied as possible predictors of driving impairment due to drowsiness. The following report thus provides a review of the literature dealing with drowsy and fatigued drivers and the many measures of driver impairment which have been investigated.

The literature is reviewed according to the type of dependent measure used: 1) actual driving performance measures; 2) physiological measures of drowsiness; 3) subsidiary task performance measures; and 4) subjective ratings of drowsiness. It should be noted at the outset that the interest in subjective ratings and physiological measures is due largely to their potential use as predictors of actual driving performance. Since performance measures have generally shown promise as predictors of driver impairment and since these measures can be practically implemented in conjunction with a monitoring and alerting device, the following report emphasizes the use of subjective ratings and physiological measures of drowsiness as correlates of drowsiness-impaired performance measures. Similarly, since a system which monitors subsidiary task performance could also be easily implemented, those studies employing subsidiary task measures are reviewed with a special emphasis placed on the validation of the use of subsidiary task performance measures as an indicator of driver impairment. Additionally, a discussion of the merits of using subsidiary tasks as active drowsiness countermeasures is offered.

The following review includes both those few studies dealing specifically with driver drowsiness and those dealing with the closely related topic of driver fatigue. (It is noteworthy that, while most of these reported studies defined drowsiness as simply being associated with sleep deprivation, for the purposes of the current study, drowinesss has been defined according to Erwin's (1976, p. ii) definition as "that moment when persistent eyelid closures," or slow, ramp closures occur. The operationalization of fatigue as time on the driving task is consistent with the reviewed literature, however.) Additionally, some factors which have been shown to affect performance on vigilance tasks and which may be applicable to drowsy driver research are discussed. Finally, a brief note concerning the validity of automobile simulators as an alternative to on-the-road driving is offered.

LITERATURE REVIEW

Driving Performance Measures

Perhaps the problems associated with attempts to assess the effects of fatigue and drowsiness on driving behavior were best expressed by Crawford (1961) when he said:

> It's generally agreed that performance can be impaired by driving for too long a period but it has proved extremely difficult to define what is meant by driving performance and to develop adequate techniques of measuring it. (p. 143-144)

Today, 25 years later, researchers are still struggling with the problem of how to define driving behavior. A look at the literature reveals almost as many performance variables as there are research articles dealing with the issue. For instance, Ryder, Malin, and Kinsley (1981) list several vehicle control variables which have been examined, such as:

1. Subjective measures of driving performance;
2. Steering behavior--including rate of steering wheel reversals, steering shaft angle, and lane position variability; and
3. Longitudinal control--including accelerator and brake applications, accelerator noise (standard deviation of acceleration), and speed variability.

In addition to these performance measures, reaction time to emergency events such as lead car deceleration and braking for obstacles has also been studied. (It should be noted that the studies reviewed in this section deal only with driver reaction time to specific, driving-related events and not to reaction times on tasks which are subsidiary to driving. Studies involving subsidiary tasks are discussed at length in a later section.) Finally, a behavioral index of increasing driver fatigue and drowsiness, namely, increased restlessness and movement, has also been noted. The research results concerning each of these variables are discussed below.

Subjective measures of driving performance. One study which used subjective evaluations of performance by trained observers was conducted by Brown in 1967. Brown tested six drivers in the morning and afternoon on a route in city traffic. On one test day subjects drove continuously for seven hours between the tests, while on the other day they performed their normal jobs. Experienced drivers rated subjects' driving performance based on competency of using controls, anticipating traffic changes, and courtesy to other drivers. Performance on the afternoon tests were rated significantly worse. A further analysis of these data indicated that this observed decrement in performance occurred mostly in perceptual skills and courtesy toward other drivers rather than in motor skills.

Steering behavior. Several researchers have investigated the effects of sleep deprivation and fatigue on the steering behavior of drivers. These investigators have employed several measures of steering behavior, including the number of steering wheel reversals, lane position variability, and steering wheel position, with mixed results. The inconclusiveness of these results is exemplified by the findings associated with the use of steering wheel reversals as an indicator of driving performance. For instance, Brown (1965) employed patrolmen who generally spent eight hours per day driving a vehicle to study the effects of prolonged driving on several performance variables, including steering wheel reversals. The results indicated no significant differences in the number of steering wheel reversals before and after extended periods of driving. Similar results were also obtained by Brown (1966) and Brown, Simmonds, and Tickner (1967) in a 12-hour, virtually continuous driving task also involving a vigilance task. Additionally, Riemersma, Sanders, Wildervanck, and Gaillard (1977) found no significant changes in the frequency and amplitude of steering during prolonged, on-the-road driving.

Other investigators have been more successful in their use of steering wheel reversals as an index of fatigue or drowsiness resulting from prolonged driving, however. Safford and Rockwell (1967) conducted a study in which drivers were asked to continue with a driving task for as long as they could up to 24 hours. Several vehicle control variables were monitored during this task. The results indicated that the number of steering wheel reversals correlated well with time-on-task. There was indication of a high degree of inter-subject variability, however.

In another study involving the use of simulators, Sussman, Sugarman, and Knight (1971) found a significant decrease in the number of steering wheel reversals over a four hour driving period. Similarly, Muto (1981), Wierwille and Muto (1981), and Muto and Wierwille (1982) found a significant decrease in small steering wheel reversals over prolonged driving periods of 30, 60, and 150 minutes. Drivers did exhibit a significant increase in the number of large steering wheel reversals, however.

The results of two other studies seem to indicate that steering wheel reversals do not provide a reliable measure of performance decrement due to drowsiness. In a study designed to assess the effects of sleep deprivation and alcohol on driving performance, both singly and in combination, Huntley and Centybear (1974) found no significant change in the number of fine or coarse steering reversals associated with 29 hours of sleep deprivation. Similarly, Erwin (1976) found that the number of steering wheel reversals does not change as drivers become more and more drowsy. Drivers do, however, cease to make steering movements entirely when their eyes close in sleep.

Perhaps the use of steering reversals as a measure of driving performance was best summed up by McLean and Hoffman (1975). After conducting an extensive literature review, these authors concluded that:

> 1) Steering control frequency characteristics are more reflective of steering task difficulty than of absolute steering performance.
>
> 2) In real-world driving studies, steering reversal rate can be a valid criterion for detecting differences in drivers or conditions. However, caution is required in using them to interpret the nature of such differences. (p. 256)

Those studies employing lane position variability as a measure of driving performance have found consistent results. In 1966, Mast, Jones, and Heimstra conducted a simulated driving task in which subjects were required to drive for four or six hours under various driving conditions. The most obvious performance decrements as measured by tracking error occurred in the 6-hour conditions where performance during the first hour was significantly better than that during the last hour. Similar results have also been found in studies using simulators; for instance, Dureman and Boden (1972) and Sussman et al. (1971) found a significant increase in lane tracking error (as measured by steering errors) in four-hour simulated driving tasks. These results were also obtained by Muto (1981), Wierwille and Muto (1981) and Muto and Wierwille (1982).

Additionally, Boadle (1976) found that subjects who drove an automobile simulator and tracked the taillights of a lead car for two hours exhibited a significant decrease in mean absolute tracking errors at the beginning of the drive and a significant increase toward the end. The results also indicated that signed error scores (which took into account the direction of and distance between the position of the taillights and their center of travel) were large at the start of the two hour period but showed a marked cyclic effect during the second hour.

This phenomenon has been associated with sleep deprivation as well. Skipper, Wierwille, and Hardee (1984) found a significant increase in the lane position variability of subjects who had been sleep deprived for 19 hours and who then operated an automobile simulator for one and one half hours. This increase in variability was associated with both sleep deprivation and time-on-task.

On-the-road driving experiments have also yielded identical findings. Riemersma et al. (1977) found a significant increase in the standard deviation of lane position of subjects who drove continuously from 2200 to 0600 hours. This increase was also a function of both sleep deprivation and time-on-task.

Similarly, O'Hanlon and Kelley (1977) had subjects drive on one of three different routes for approximately four hours at night. An analysis of these data indicated that lane position as measured by lane drift (the number of times the vehicle exceeded lane boundaries) varied markedly with the experimental route. Nonetheless, the number of lane drifts tended to increase as a function of time-on-task in all three driving conditions.

Longitudinal control. The longitudinal control variables which have been most often examined in conjunction with drowsiness and fatigue are accelerator and brake applications (including reversals and activations) and speed variability (longitudinal acceleration and velocity maintenance). A review of this literature indicates that there is no significant change over time in the number of brake activations or accelerator reversals as a function of eight (Brown, 1965) or twelve hours (Brown, 1966; Brown et al., 1967) of on-the-road driving. While Safford and

Rockwell (1967) found that both number of accelerator reversals and brake usage correlated highly with time, these correlations exhibited great inter-subject variability. Finally, Huntley and Centybear (1974) demonstrated that neither the number of accelerator reversals nor brake usage increased significantly with 29 hours of sleep deprivation.

The results of studies employing speed variability as a measure of performance decrement due to sleep loss or fatigue have indicated mixed results. For instance, Mast et al. (1966) found significant differences between subjects' abilities to maintain constant velocity during the first and last hours of both four- and six-hour simulated driving tasks. In three real world, on-the-road driving studies however, Brown (1965, 1966) and Brown et al. (1967) found no significant differences in velocity maintenance associated with either eight- or twelve-hour driving tasks. Likewise, Safford and Rockwell (1967) found no increases in speed variability as a result of 24 hours of driving. However, in a study of changes over a night of driving, Riemersma et al. (1977) found that speed variability increased significantly, but only during the second half of the night.

<u>Multivariate models of driver drowsiness</u>. One possible alternative to the use of the aforementioned performance variables singly has been investigated by Attwood and Scott

(1981). These investigators have attempted to combine performance measures (e.g., lane position median/range and accelerator position) into a multivariate prediction model of driver drowsiness. The results of their study suggest that it is possible to discriminate successfully between fresh and sleepy drivers in this way. While the data analysis reported by Attwood and Scott is based on the data of only one subject, these results do offer hope that multivariate modeling is a promising method of reliably detecting impaired driving due to drowsiness.

Alternatively, Skipper et al. (1984) were able to identify driving signatures which were characteristic of drowsy drivers and which were associated with obvious incidences of driving impairment, namely, lane exceedences. These investigators were able to identify two driving stereotypes associated with lane exceedences; when they exceeded lane boundaries, drivers tended to display either oscillatory build-up or nonoscillation and steering hold prior to the exceedence followed by a correction or steering burst. These stereotypes accounted for approximately 65% of the total number of lane exceedences. Thus, Skipper et al. (1984) used two steering variables, steering velocity and burst, and one lane position variable, oscillation, to characterize impaired driving. This provides further support for the use of multivariate models of driver performance and impairment.

Driver reaction time. Those investigators who have examined the effects of sleep deprivation and task duration on the performance of driving-related reaction time tasks have generally found that not only reaction time, but also reaction time variability increases with time-on-task or with sleep loss (Davies and Parasuraman, 1981; Ryder et al., 1981; Wilkinson, 1965). For instance, Lisper, Dureman, Ericsson, and Karlsson (1968; cited in Lisper, Dureman, Ericsson, and Karlsson, 1971) found a significant increase in brake reaction time to light signals on the side of the road as a function of time-on-task. Likewise, subjects in a study by Laurell and Lisper (1978) showed a significant increase in their reaction times to the detection of roadside obstacles as a function of time.

Muto and Wierwille (1982) also found that subjects' reaction times to an emergency situation involving reacting to the sudden deceleration of a lead car in a car-following scenario were significantly greater after they had driven an automobile simulator for 30, 60, or 150 minutes as compared to their reaction times both before prolonged driving and on baseline runs. This increase in reaction time was associated with only the first two or three emergency trials; however, after the presentation of several response trials the decrement normally associated with fatigue was eliminated or reversed. These results thus lead the investigators to conclude that repeated response trials may not provide valid indications of fatigue-induced performance decrements. Hence, these results should not be generalized to on-the-road driving performance under normal driving conditions. These results also seem to provide support for the hypothesis by Brown et al. (1967, p. 665) that "prolonged driving leads to greater automatization of control skills, which increases the time available for the perceptual requirements of the driving task."

Mixed results have also been found by some investigators. For instance, Mast et al. (1966) found that reaction times on a simulated driving task significantly decreased in two of six experimental conditions while, in the remaining four conditions there were no significant differences even though performance was better during the last hour of the task compared to the first hour. These experimenters thus concluded that tasks may differ in their sensitivity to fatigue.

In 1976, Yajima, Ikeda, Oshima, and Sugi demonstrated no increase in simple reaction time by drivers who completed seven to ten hours of driving on a driving circuit or nearly twenty-four hours on-the-road. As for the effect of prolonged driving on choice reaction time, there was great inter-subject variability--some subjects experienced an increase while others experienced a decrease. Nonetheless, in all cases there was an increase in reaction time variability as a function of time-on-task. This increase in reaction time variability has been explained as resulting from a pattern of behavior in which the individual performs at normal or near normal levels of efficiency with periods of inefficiency interspersed (Davies and Parasuraman, 1981; Wilkinson, 1965).

Changes in reaction time as a function of sleep loss have also been investigated. Buck (1973), using a step input tracking task known as a stressalyzer, found that the simple reaction times of subjects who had been sleep deprived for 48 hours were not significantly different from their reaction times during "fresh" conditions. There was, however, a significant increase in mean movement time (the interval between movement initiation and final alignment with the target) for sleep deprived subjects. Williams, Lubin, and Goodnow (1959), on the other hand, found a significant increase in choice reaction time for subjects who had been sleep deprived up to 98 hours.

Driver behavioral measures. One other interesting measure of driver drowsiness or fatigue which does not relate directly to performance but which may be predictive of driving impairment involves the increase of certain driver behaviors with sleep loss and time-on-task. For instance, Yajima et al. (1976) noted a general increase in driver restlessness as a function of driving time. A similar finding was also reported by Wierwille and Muto (1981); there was a significant increase in body movements per unit time which was associated with time-on-task. Finally, Skipper et al. (1984) noted that several behaviors (e.g., upper body movements, hand to head movements, and general "mannerisms") changed significantly as a function of sleep-deprivation and time-on-task. Thus, it appears that certain driver behavior measures may be used to predict fatigue-induced driver impairment.

Summary. Given the mixed results associated with the use of both longitudinal control and steering behavior variables, with the exception of lane position variability, it would seem that these variables do not necessarily provide reliable measures of performance impairment associated with drowsiness or fatigue. Moreover, as Safford and Rockwell (1967) point out, given the high degree of inter-subject variability associated with some variables (e.g., accelerator reversals, mean velocity, velocity variance, and steering wheel reversals) with the passage of time, it may be that these types of measures cannot be accepted as universal measures of fatigue. Rather, it may be necessary to know specific characteristics of the driver in order to interpret or predict performance.

On the other hand, it seems that reaction time to driving events may provide a reliable measure of fatigue-induced performance decrement. However, choice reaction times appear to be more sensitive to the effects of fatigue or sleep loss than do simple reaction times. Additionally, since reaction time variability generally tends to increase with sleep loss or fatigue, this measure may be a valuable one for use as a predictor of performance impairment. Finally, given the correspondence of certain behavioral indices, such as driver restlessness, to sleep loss and time-on-task, it seems plausible to use measures of these behaviors as predictors of driver drowsiness and/or fatigue. In any event, the use of *multivariate* models as predictors of driver impairment shows promise and thus warrants further research.

Physiological Measures of Drowsiness and Fatigue

There has been a general interest within the scientific community in attempting to define drowsiness and fatigue through the use of physiological measures. More specifically, researchers have attempted to pinpoint the exact moment of the occurrence of sleep by monitoring various physiological measures such as EEG and heart rate. Additionally, others have been interested in studying the effects of driving-induced fatigue on the physiological responses of drivers. The following sections of this report thus review the research dealing with the physiological correlates of both drowsiness and driving-induced fatigue. Additionally, those studies which have found correlations between physiologic responses and driving performance measures are discussed. Finally, very brief descriptions of a few physiological alerting devices are given and evaluated.

Physiological measures of drowsiness. For years scientists have been interested in the study of drowsiness and its physiological correlates; however, the results of many of the medical studies which have been conducted in this area are of limited value because they are descriptive of subjects with pathological sleep disorders such as narcolepsy and hypersomnia (Fagerstrom and Lisper, 1978).

Since most drivers who experience drowsiness and/or intrusive sleep while driving do not possess pathological sleep disorders, the general driving population has been largely ignored (Erwin, 1976).

There has been one group of studies on drowsiness, however, in which both narcoleptics and normals representative of the general population were used. Ten laboratory studies conducted at Duke University were devoted specifically to the study of drowsiness in an attempt to achieve a greater understanding of the phenomenon of intrusive sleep. The results of these investigations have been summarized in a final report by Erwin (1976), and an overview of these findings is presented here.

One major problem which must be overcome before beginning any sleep research is defining when sleep occurs. While it is generally agreed that there is a continuum of decreasing arousal from full alertness, through drowsiness, to sleep, the exact moment when sleep occurs has been difficult to pinpoint (Erwin, 1976). In the past, the electroencephalogram (EEG) has been universally accepted as the most precise measure of an individual's arousal state. Through the use of the EEG, at least five stages of sleep have been identified; of these stages, Stage I represents light sleep. Unfortunately, Stage I sleep is very similar to being awake (Stage W) and is thus difficult to identify. Also, there is some debate among researchers as to whether Stage I actually represents sleep.

Given these inherent problems with the use of the EEG as an index of drowsiness and sleep onset, intrusive sleep was defined as "that moment when persistent eyelid closure occurs" (Erwin, 1976, p. ii) for the purposes of the Duke University studies. This operationalization of intrusive sleep is one which is particularly attractive and useful to those who study driver drowsiness. Since the driver must respond to dynamic visual cues to operate an automobile successfully, the onset of intrusive sleep (i.e., when persistent eyelid closure occurs) represents that point in time when driving performance is most likely to be severely impaired.

Given the aforementioned operationalization of intrusive sleep, the Duke University studies investigated several physiological measures in regard to their correlation with drowsiness onset. It should be noted that these were laboratory and not on-the-road studies. The eight physiological measures which were investigated were: plethysmography, respiration rate, electroencephalography, skin resistance level, skin potential level, electromyography, heart rate variability, and eyelid position. Each of these is discussed in turn with regard to its usefulness to drowsiness research.

1. Plethysmography - Erwin (1976) reported that he and other investigators (Volow and Erwin, 1973) were unsuccessful in obtaining stable, long-term (>30 minutes) records even under the most optimal laboratory conditions. The major problem associated with measures of heart rate, however, is the amount of noise related to body movements and changes in position.

2. Respiration - Erwin (1976) also reported "no direct correlation between respiratory rate and the onset of drowsiness and/or sleep" (p. vi). Alterations in respiratory rate did not develop until the onset of Stage II sleep; this stage of sleep does not generally begin until 10 to 15 minutes after the onset of Stage I sleep.

3. Electroencephalography - EEG changes were found to occur following eye closure; alpha activity decreased within 0.5 seconds and was replaced by the lower voltage activity associated with Stage I sleep. There were no significant changes in EEG measures prior to eye closure however (Erwin, Hartwell, Volow, and Alberti, 1976; Erwin, Volow, and Gray, 1973). These finding are not consistent with those of other researchers who have demonstrated a decrease in alpha activity associated with increasing sleep deprivation

(Davies and Parasuraman, 1981; Wilkinson, 1965; Williams, Lubin, and Goodnow, 1959). Erwin (1976) also noted that electrode application and poor signal to noise characteristics are always a problem even in controlled laboratory settings. Therefore, EEG seems to be of limited practical use in a vehicle.

4. Skin electrical characteristics - Both skin potential level (SPL) and skin resistance level (SRL) have been studied extensively. In general, it was found that SPL tended to decrease in negativity with sleep onset as measured by EEG (Erwin et al., 1973; 1976). SRL was also demonstrated to decrease with sleep onset (Erwin et al., 1976). However, Erwin (1976) pointed out that these measures may be less than practical for use as predictors of drowsiness for two reasons. First, individuals may give evidence of lowered autonomic arousal without showing other manifestations of drowsiness or sleep. Second, there is evidence of great inter- and intra-subject variability. Nonetheless, some researchers (Kousmens, Tursky, and Solomon, 1968) believe that the SPL shows greater potential for discriminating between a sleeping and awake state than the SRL measure (Erwin, 1976; Erwin et al., 1976; Erwin et al., 1973; Volow, Erwin, Cipolat, and Hartwell, 1976).

5. Electromyography - This measure offered no predictive measure of either drowsiness or sleep. Sleep can occur over several minutes before any real change takes place in the EMG (Erwin, 1976).

6. Heart rate variability - Volow and Erwin (1973) indicated that they were unable to document any significant changes in the variance of the interbeat interval during the course of spontaneous sleep onset. Erwin (1976) also reported great inter-subject variability.

7. Eyelid position - Erwin (1976) found that the eye aperture in the open position measures 13 mm ± 1.5. This measure was found to be remarkably stable across subjects. Additionally, eyelid closure as a physiologic signal of drowsiness offers several unique characteristics: 1) there is relatively little inter-subject variation in the aperture of opened eyelids and no intra-subject variation over time; 2) the significance of the phenomenon is unambiguous; clearly, a hazardous state exists when the eyes are closed, regardless of the reasons for the closure. As a matter of fact, Erwin (1976) states that eye closures greater than one second are hazardous since, at night, as much as 80 feet of roadway may go undetected during a one-second closure if the driver is traveling at a speed of 55 miles per hour; 3) eyelid closure was the most highly

correlated physiologic parameter associated with defective performance; and 4) the eyelids are lowered at a much slower rate than can be achieved when an individual is in an alert state (Erwin, 1976). For these reasons, Erwin states that "eyelid position offers the simplest and most direct measure predictive of subject performance" (pp. xii-xiii).

Erwin (1976) also described two reliable methods by which eyelid closure may be monitored. One technique, the infrared photoreflective technique, requires that modulated infrared light be directed at the upper lid. The contrast between the upper lid and the pupil is then detected by a photo-diode. A second technique involves DC potential recording. The eye serves as a dipole which is negative at the retina and positive at the cornea. As the eye closes the eyeball rotates upward, resulting in a change in the orientation of the dipole. While both of these techniques provide reliable measures in the laboratory setting, Erwin (1976) realized the problems inherent with their use in the field. Nonetheless, eyelid closure data provides a reliable index of sleep onset which could be used in research settings in an attempt to correlate this physiological measure of drowsiness with other, more easily measured driving performance variables.

Physiological measures of driving-induced fatigue. While Erwin (1976) and his associates were conducting their laboratory research pertaining to the physiological measures associated with drowsiness and sleep onset, other investigators were interested in determining what, if any, physiological measures might correlate with, and thus be predictive of, driving-induced fatigue. The following review of this research is organized as was the previous review of the Duke studies--according to the physiological measure of interest.

1. Plethysmography - Those researchers who have examined the effect of prolonged driving on heart rate have generally found either a decrease or no change in heart rate over time. For instance, Boadle (1976) and Dureman and Boden (1972) found no significant changes in heart rate over two- and four-hour simulated driving periods.

On the other hand, significant decreases in drivers' heart rates have been observed in on-the-road driving (Fagerstrom and Lisper, 1977; Lisper, Laurell, and Stening, 1973; O'Hanlon and Kelley, 1977; Riemersma et al., 1977). Riemersma et al., however, cautioned that this decrease may not be the result of fatigue but may be due instead to adaptation, decreasing stress, or, in their case, to the effect of diurnal rhythms. Two other studies have indicated, however, that there is no significant change in heart rate after 12 (Brown, 1967) or after 7-10 or 24 hours of prolonged driving (Yajima et al, 1976).

2. Respiration - Several researchers have shown an interest in the effect of prolonged driving on respiration rate. These investigators have found that respiration rate significantly decreases after two (Boadle, 1976) and four hours (Dureman and Boden, 1972) of simulated driving. Similarly, Lisper et al. (1973) found that, in an on-the-road driving study, respiration rate decreased slightly for both inexperienced and experienced drivers, but that there was not a significant difference in respiration rates for these two groups.

3. Electroencephalography - The EEG has also been employed as a predictor of driving-induced fatigue. The results of these studies have shown that alpha activity tends to increase after four hours of simulated driving (Sussman et al., 1971) and after prolonged on-the-road night driving (O'Hanlon and Kelley, 1977). Additionally, Lemke (1982) found a change in EEG data associated with prolonged driving under monotonous conditions. These results are in opposition to those associated with sleep deprivation, namely, a decrease in alpha activity with increasing loss of sleep (Davies and Parasuraman, 1981; Williams et al., 1959).

5. Skin electrical characteristics - The studies dealing with the use of skin electrical characteristics as predictors of driving-induced fatigue have demonstrated mixed results. For instance, Dureman and Boden (1972) found that drivers experienced a decrease in skin resistance level but no change in skin resistance response during four hours of simulated driving. O'Hanlon and Kelley (1977), however, found that skin resistance response increased slightly over three to five hours of on-the-road driving. Finally, Yajima et al. (1976) found that drivers who operated a vehicle on-the-road for 7-10 or for 24 hours exhibited an increase in GSR early on, followed by a decrease later.

5. Electromyography - One study has been reported which investigated electromyography (EMG) as a predictor of driving-induced fatigue. These results indicated no change in EMG activity during four hours of simulated driving (Dureman and Boden, 1972).

6. Heart rate variability - Studies dealing with changes in heart rate variability as a function of time-on-task have generally found that heart rate variability increases significantly during prolonged driving. Muto (1981) found a significant increase in heart rate variability for subjects who had driven an automobile simulator for 60 or 150 minutes, but not for those who had

driven for 30 minutes. Similarly, Wierwille and Muto (1981) demonstrated that subjects who were exposed to emergency situations evidenced an increase in heart rate variability during 30-, 60-, and 150-minute simulator driving. Furthermore, the results of on-the-road experiments have yielded identical results; significant changes in heart rate variability were found to be associated with prolonged night driving (O'Hanlon and Kelley, 1977; Riemersma et al., 1977).

7. Eye closures and/or movements - Much research time has been devoted to changes in eye closures and/or movements as a function of both sleep deprivation and time-on-task; however, only three studies are reviewed here. First, Kalugar and Smith (1970; cited in Forbes, 1972) studied the effect of prolonged driving on the eye movements of subjects who had and had not been sleep deprived for 24 hours. They found that the sleep deprived drivers: 1) fixated in close and to the right of the highway; 2) exhibited pursuit eye movements approximately 5% of the time (while control subjects exhibited no such movements); and 3) frequently closed their eyes for one to three seconds.

Yajima et al. (1976) monitored saccadic eye movements during prolonged on-the-road driving and found a significant decrease in movements as a function of time-on-task.

Finally, Skipper et al., (1984) measured the eyelid closures of both rested and sleep deprived subjects during a 90-minute driving simulation task. An analysis of these data indicated that the three different measures of eyelid closure which were used were sensitive to both sleep deprivation and time-on-task. More specifically, the percent eyelid closure increased from the first to the third 30-minute driving segment in both sleep deprived and rested subjects.

Correlations of physiological and performance measures. While some physiological measures do show promise as indices of drowsiness or driving-induced fatigue (e.g., eye closures), the practicality of using physiological monitoring equipment in moving vehicles is limited (Erwin, 1976) due to the increase in measurement noise contributed by the automobile as well as to the unwillingness of drivers to wear the necessary leads or equipment. Therefore, research has been conducted which attempts to correlate physiological measures of drowsiness and/or fatigue with performance measures in the hope that driver performance may be monitored and decrements detected.

Several researchers have attempted to correlate physiological measures of drowsiness and/or fatigue with driving performance measures. For instance, Dureman and

Boden (1972) had subjects perform a four-hour driving simulator task during which they performed an auditory subsidiary task. An analysis of the data revealed that changes in EMG, heart rate, and respiration rate were all positively correlated with an increase in steering errors over time. Similarly, Erwin (1976) found that two physiological measures were associated with performance decrement on a tracking task which simulated automobile driving: prolonged or ramp-like eye closures and a decrease in skin potential level. Skipper et al. (1984) also demonstrated high correlations of eyelid closures with both lane deviations and steering velocity.

The EEG has also been demonstrated to be a reliable predictor of driving performance. Lemke (1982) conducted a prolonged driving study in which both mechanical signals (steering angle velocity, accelerator pedal velocity, yaw angle velocity, and brake and seat activity) and physiological indices (EEG, ECG, pulse rate, blood pressure, and flicker fusion frequency) were measured. A factor analysis and canonical correlation showed a change in EEG data which was accompanied by a loss of precision in the control behavior of drivers. Additionally, the redundancy of EEG signals against the mechanical signals implied that the mechanical signals could be used to predict decreased driver performance due to decreased vigilance. Lemke thus suggested that EEG measures be used as a criterion measure in a warning device for detecting decreasing driver vigilance.

Finally, other researchers have been less fortunate in their findings. For instance, Fagerstrom and Lisper (1977) employed an auditory subsidiary task in an attempt to relate reaction time to the task with heart rate. These researchers found no significant relationship between these two measures, however.

Physiological monitoring devices. The idea of monitoring physiological events in an attempt to detect and alert drivers to lowered states of arousal is not a new one. Several researchers have tried various methods with limited success. For instance, Travis and Kennedy (1947) attempted to employ neck EMG's in a physiological alerting device but found the system to be impractical.

One other type of physiological monitoring device takes advantage of the findings by Erwin (1976) concerning eyelid closures and measures eyelid droop or closure in normal driving conditions. To accomplish this, however, a special pair of eyeglass frames which contain a sensor assembly and an amplifier-detector unit must be worn (Technology Associates, Inc., 1978). The disadvantages associated with this type of device are fairly apparent. First, those drivers who must wear prescription glasses cannot use this device. Second, those drivers who do not normally wear glasses may be inconvenienced by having to wear eyeglass frames. Finally, the frames require eight fine adjustments before eye closure can be detected accurately. Thus, it seems that, at least for now, the monitoring of eyelid closure during actual driving is impractical.

Summary. The results of these studies dealing with the use of measures of physiological responses as indices of driver drowsiness and/or fatigue provide two important findings. First, of all the measures investigated, eye closures or droop seems to be the most reliable index of drowsiness onset. Second, while the idea of using physiological monitoring devices in driver "alerting" systems is an attractive one, their impracticality and potential intrusiveness to the driver renders physiological monitoring devices unsuitable for such use. Instead, future research should focus upon the use of physiological measures (particularly eyelid closures) as predictors or correlates of impaired driver performance.

Subsidiary Task Performance Measures

As Brown (1962) pointed out, it is often difficult to tell a fatigued or drowsy driver from a fresh one by looking at his or her vehicle control manipulations alone. Because of this, researchers have often tried other means of assessing the effects of fatigue and/or drowsiness on drivers. One such method involves the use of subsidiary tasks. The logic involved with the use of subsidiary task performance as a measure of the driver's reduction in spare mental capacity due to fatigue or drowsiness is relatively simple. Basically, as the driver becomes more fatigued or drowsy, the primary task of driving becomes more demanding as a result of the decline in the driver's total effort or attention. Therefore, performance on a subsidiary task suffers (Ryder et al., 1981). Given this rationale, several researchers have attempted to demonstrate the effect of fatigue and/or drowsiness on subsidiary task performance. The following report summarizes their findings. To facilitate discussion, the reviewed literature has been organized according to the modality of the subsidiary task stimulus, namely, auditory and visual.

Auditory subsidiary tasks. Given the goal to measure the spare capacity of the driver as an indication of the driver's level of fatigue and/or drowsiness without further impairing driving performance, it seems obvious that the subsidiary task should involve the use of those sensory and response modalities which are least used in the primary task of driving. Accordingly, most subsidiary tasks used in driving studies have employed auditory stimuli (Wetherell, 1981) in conjunction with both vocal and tactual responses.

Brown and Poulton (1961) provided support for the validity of the use of subsidiary task performance as a measure of the driver's spare mental capacity. These researchers used both average and advanced drivers who drove around a test circuit which included both residential and shopping areas. While they drove, the average drivers performed an auditory task in which they listened to eight digits and then picked out and spoke the new digit (relative to the last eight digits). The advanced group performed a task involving mental arithmetic; they were also required to respond verbally. The results demonstrated that both groups of drivers made significantly more errors on the subsidiary task during driving in shopping areas than in residential areas. This provided validation for the use of subsidiary tasks as a measure of spare capacity since one would expect a reduction in spare capacity associated with driving in a busy shopping area and a subsequent decrement in subsidiary task performance. Furthermore, the subsidiary task did not substantially affect driving performance as measured by changes in vehicle velocity.

Other researchers who have used auditory subsidiary tasks requiring vocal outputs have also found a change in subsidiary task performance associated with time-on-task, even though the change has not always been in the expected direction. For instance, Brown (1962) used mobile patrol officers who normally drove eight hours a day in a study which employed two auditory tasks. One task, the attention task, required the driver to listen to random digits and respond "Now" when an odd-even-odd sequence was presented; the second memory task required the subject to listen to a sequence of 10 letters and pick out the one letter that occurred twice. The subjects were tested at approximately 4:00 p.m., either just before working the 4:00 p.m. to midnight shift or just after working the 8:00 a.m. to 4:00 p.m. shift. These data indicated that, contrary to expectation, performance on both auditory tasks as measured by the percentage of correct responses was better _after_ eight hours of driving. One explanation for these curious findings, which was offered by the investigator, was that the measures of subsidiary task performance taken before the workshift began may have shown the effect of general fatigue which was unrelated to driving (Brown, 1962). This general fatigue may have been a result of the patrolmen performing household chores before coming in to work at 4:00 p.m.

In 1965, Brown replicated the 1962 study using the same paradigm and subject pool while adding more control groups. An analysis of these data indicated that performance on the auditory tasks alone did not differ before or after prolonged driving. However, a comparison of performance on the subsidiary task alone with that while driving showed that performance on the attention task was significantly better after eight hours of driving than before. There was no significant difference in performance on the memory task as a function of driving time, however. Once again, Brown (1965) offered the explanation of generalized fatigue for this finding. However, he also suggested that fatigue may produce greater lability of performance so that when conditions are arousing (as in the present experiment) driving is better than normal, but when the conditions are monotonous (such as in motorway driving) driving is worse than normal.

In addition to this effect of driving on subsidiary task performance, there was also an effect of the subsidiary task on driving. Time to complete the driving circuit was significantly longer when the auditory tasks were performed. More specifically, it was found that: 1) the attention task distinguished significantly between performance at the beginning and end of eight hours of driving compared to the memory test which did not; 2) changes in subjects' speed of driving were significantly correlated with changes in subsidiary performance only on the attention task; and 3) seven of eight subjects found it necessary to reduce normal speeds of driving less when performing the attention task than when performing the memory task.

Given these three findings, Brown (1965) argued that the attention task is a better measure of the driver's spare mental capacity than the memory task because driving competed more effectively with the attention task and was more sensitive to fluctuations in perceptual load on the primary task of driving. Also, the memory task had a greater detrimental effect on driving. Thus, Brown concluded that subsidiary task performance measures are indeed sensitive to driver fatigue, and that attention task measures are more sensitive than memory task measures.

Wetherell (1981) was also interested in comparing the efficacy of several auditory-vocal subsidiary tasks as measures of mental load. To accomplish this, Wetherell had subjects perform eight subsidiary tasks (addition, verbal reasoning, attention, short-term memory, random digit generating, memory search, and white noise) while they drove along a test circuit located on private, rural roads. Several driving measurements (time taken to complete each circuit, deceleration rates at junctions, and the number of gear changes and brake operations per circuit) as well as subsidiary task performance measurements were recorded.

An analysis of the effects of the subsidiary tasks on driving showed that, for males there were no significant between-task differences in any of the driving variables measured. For females, however, there were significant between-task differences in trip times; no significant correlations were found between the changes in driving performance and the performance decrements on any of the subsidiary tasks. Additionally, both males and females exhibited significant between-subject differences in all driving variables. Finally, significant differences were found between male and female trip time variance ratios under several conditions (task free, addition, verbal reasoning, attention, memory search, and white noise).

These results also indicated that driving caused a decrement in subsidiary task performance (as measured by response time) for both males and females on the addition and verbal reasoning tasks. Additionally, male drivers suffered from performance decrements on the short-term memory task while females did not.

Wetherell (1981) concluded from these results that no single task appeared to be outstanding as a measure of driving-imposed mental load. He also pointed out that the female drivers in the study may not have given priority to the driving task as they were instructed to do since the females tended to reduce their driving speed in an attempt to decrease the processing demands of driving so that more attention could be paid to the subsidiary tasks. Therefore, performing the subsidiary task appeared to interfere with female, but not male, driving ability.

Other researchers have provided evidence of the validity of auditory-tactual subsidiary tasks as measures of driving performance decrements due to prolonged driving. For instance, Lisper (1966; cited in Lisper, Dureman, Ericsson, and Karlsson, 1971) had subjects react to the presentation of a tone by depressing a foot pedal. Subjects' simple reaction times to this subsidiary task was found to correlate with time to detect road signs. Similarly, other investigators have found a high correlation between reaction time on an identical subsidiary task and brake reaction time to light signals on the side of the road (Lisper, Dureman, Ericsson, and Karlsson, 1968; cited in Lisper et al., 1971) and to detection distance to roadside obstacles (Laurell and Lisper, 1978). Thus, these three studies indicate a parallel breakdown of driving and subsidiary task performance with increasing time-on-task.

The remaining studies which are reviewed here also used subsidiary tasks identical to that cited by Lisper (1966; cited in Lisper et al., 1971) and Lisper et al. (1968; cited in Lisper et al., 1971). The results of these studies almost without exception indicate that drivers experience a significant increase in simple reaction time to a subsidiary task as a function of time-on-task. For example, Lisper et al., (1971) had subjects drive for four hours in daylight, in darkness, and in daylight after one night of sleep deprivation. While there was a significant increase in reaction time associated with time-on-task, the major hypothesis that subjects would experience greatest increases in reaction time with sleep deprivation was not confirmed. Neither was the hypothesis that driving in darkness would yield greater simple reaction times.

In a similar study, Fagerstrom and Lisper (1977) tried to determine the effect of listening to a car radio on driving performance decrements associated with prolonged driving. To accomplish this, these experimenters had subjects drive for four hours in each of three conditions: with music, talk, or silence. As was predicted, drivers exhibited significant increases in their simple reaction times to a subsidiary task as a function of time-on-task. There was, however, a greater average increase in reaction time in the silent condition than in the the other two conditions, thus suggesting that music or talk aids the driver in maintaining alertness.

Additionally, Lisper, Eriksson, Fagerstrom, and Lindholm (1979) demonstrated in a study designed to assess the effect of diurnal rhythm on prolonged driving performance that simple reaction time to a subsidiary task increased as a function of time-on-task but that this increase in reaction time was unaffected by time of day.

Finally, Lisper et al. (1973) demonstrated that simple reaction time to subsidiary tasks may actually decrease as a function of time-on-task for some drivers. These researchers found that reaction time significantly decreased for experienced drivers while it increased for inexperienced drivers. Thus, Lisper et al. (1973) concluded that inexperienced drivers may be more adversely affected by long-term driving than experienced drivers. This issue, along with the potential effects of diurnal rhythm on driving performance will be discussed more fully in a later section.

Visual subsidiary tasks. Brown et al. (1967) investigated the effect of 12 hours of car driving on a visual-vocal subsidiary task in an attempt to overcome some methodological problems associated with previous studies dealing with subsidiary task performance while driving. Subjects were given short driving tests at several points in time on two days. On the experimental day, the subjects drove continuously for 12 hours between driving tests; on the control day they carried on their normal work between driving tests. The subjects were also required to perform a subsidiary task while driving on main roads; this task involved detecting a light signal in the interior mirror and responding vocally. Once again, simple reaction times to the subsidiary task were found to improve significantly with time-on-task. These findings are consistent with those of Brown (1962, 1965). Brown et al. (1967) reiterated that these results can be explained by the automatization of control skills during prolong driving which increases the time available for the perceptual requirements of the driving task.

In another on-the-road test of driving performance during prolonged night driving, Riemersma et al. (1977) had drivers perform two visual subsidiary tasks--one requiring that the driver keep track of the kilometrage and report each time the counter reached a multiple of 20, and a second, simple reaction time task requiring the driver to react to changes in the color of a light by pressing the button normally used for the horn. The data from this study revealed that performance on both of these tasks was significantly worse during the last half of the prolonged driving run than during the first half.

Finally, Boadle (1976) attempted to examine the effect of prolonged simulator driving on subsidiary task performance. The subjects in this study drove an automobile simulator for two hours and tracked the taillights of a lead car while they also performed a visual subsidiary task in which they responded by pressing one of two buttons to indicate if they were "sure" or "not sure" that a signal on the dash had occurred. An analysis of these data indicated that, while the number of "sure" responses did not change significantly with time-on-task, the overall response rate declined due to a decline in the number of false detections. Additionally, when Boadle (1976) compared these results with previous studies where driving was the only task, two effects of the subsidiary task became apparent. First, there was a decline in the accuracy of tracking (driving) performance. Second, the size and variability of tracking errors were greater throughout the session. Moreover, since performance on the subsidiary task was more stable over time than the driving performance, Boadle (1976) concluded that drivers may have been maintaining subsidiary task performance at the expense of driving. Given the monotony of the task, it is possible that subjects were using the subsidiary task as a source of novelty.

Since the change in driving rather than subsidiary task performance is contrary to the results of other on-the-road studies (Brown, 1967; Brown and Poulton, 1961), Boadle (1976) suggested that this difference may be due to the lack of danger in the simulator and to other general differences between the simulated task and the real situation. Thus, Boadle cautioned that future "research using simulators needs to be based on clear evidence that the simulated task and the real one have common elements and a similar effect on performance" (p. 225).

Summary. The results of these studies taken together indicate that subsidiary task performance can be used as a measure of decreased spare mental capacity due to drowsiness and/or driving-induced fatigue. These results do not, however, provide a clear indication of which type of subsidiary task is most sensitive to driving performance decrements. While there is some scant evidence to suggest that attention tasks are more sensitive to changes in spare mental capacity than are memory tasks (Brown 1965), there is no real data upon which to base a decision concerning the most appropriate subsidiary task stimulus and response modality. Nonetheless, the researcher who is interested in employing subsidiary tasks as a measure of reduced mental capacity due to fatigue and/or drowsiness should probably make his or her choice of subsidiary tasks with the following restrictions in mind: 1) the performance of the subsidiary task itself should not be detrimental to the primary task of driving; and 2) performance decrements on the subsidiary task should correspond to decrements in other driving performance measures.

Subsidiary Tasks as Drowsiness Countermeasures

While the aforementioned research examined the use of subsidiary task performance measures as measures of driving impairment due to drowsiness and/or fatigue, these studies did not investigate the possibility of using subsidiary task performance as an alerting mechanism. As a matter of fact, of those studies which were reviewed, only two attempted to develop countermeasures to increase driver alertness. Fagerstrom and Lisper (1977) examined the alerting effect of listening to music or talk on the car radio while driving. These authors chose the car radio as a drowsiness countermeasure partly because many drivers report using the radio to keep them alert. The results of their research indicated that the radio did significantly increase driver alertness as measured by reaction time to a subsidiary task. Thus, the results of this study suggest that the car radio may provide an effective countermeasure to drowsiness.

While Fagerstrom and Lisper (1977) were able to demonstrate that drowsiness and/or fatigue can be successfully countered, they did not use a subsidiary task as a drowsiness countermeasure. Drory (1985), however, conducted a study in which he attempted to assess the relative effectiveness of two subsidiary tasks used as driver drowsiness and fatigue countermeasures. To accomplish this, subjects who were employed as truck drivers for a mining firm performed a simulated truck driving task for seven consecutive hours between 12:00 p.m. and 7:00 a.m. During the driving task, the subjects were also required to perform two subsidiary tasks--a vigilance and a voice communication task. The vigilance task was a choice reaction time task which required that the driver monitor a display of lights and turn them off when they switched on by manipulating the correct control. The performance of the voice communication task, however, required the driver to respond to a verbal question concerning his current position by reading the last two significant digits of the odometer. An analysis of these data indicated that, in general, the voice communication task condition yielded significantly better driving performance than both the basic driving task condition (during which no subsidiary task was performed) and the vigilance task condition. Hence, these results tend to suggest that it is possible to maintain driver alertness through the use of subsidiary tasks.

Even though this one study provides some support for the use of subsidiary tasks as drowsiness countermeasures, further research in this area is warranted for at least three reasons. First, the study reported by Drory (1985) represents the sole experimental investigation of the use of subsidiary tasks as drowsiness countermeasures which could be located by the current author. Second, while Drory's experimental results are encouraging, they may be somewhat suspect due to the experimental design employed by the researcher. For instance, the description of the experimental paradigm indicates that the two subsidiary tasks were both performed during the same 15-minute driving blocks, thus making it extremely difficult to independently assess the effect of the performance of each task on driving. Finally, even though Drory concluded that the voice communication task was superior compared to the visual vigilance task in terms of maintaining driving performance, the nature of these two tasks does not allow for a determination of what specific aspects of the tasks are responsible for increasing driver alertness. More specifically, it is not possible to determine if the voice communication task was superior due to the sensory modalities employed, the response required to complete the task, or both.

Given this gap in the driving literature, the currently reported research was undertaken in an attempt to investigate the effectiveness of using subsidiary tasks to reduce drowsiness and to thus increase driving safety. To accomplish this, three subsidiary tasks were employed which required similar responses by the subjects and which varied only in terms of stimulus modality; hence, performance on these tasks was compared to determine which stimulus modality, if any, was superior for use in a task which is subsidiary to driving. These tasks also required the subject to make choices before responding rather than to simply respond to the presence or absence of a stimulus. This type of choice reaction time task was deemed to be more appropriate than simple reaction time tasks for several reasons:

1. Simple reaction time tasks have been found to be less sensitive to time-on-task and drowsiness effects (Williams et al., 1959; Yajima et al., 1976).

2. Since choice reaction time tasks require cognitive processing, lack of attentiveness due to drowsiness and/or fatigue should induce longer reaction times and errors.
3. Finally, since choice reaction time tasks require cognitive processing, they should have an alerting effect on the driver.

For these reasons then, three choice reaction time subsidiary tasks which varied only according to stimulus input modality were used to determine which, if any, may serve as an active drowsiness countermeasure. These subsidiary tasks are described in great detail in later sections of this report.

Subjective Ratings of Drowsiness and Fatigue

Given the problems associated with the use of more objective measures of drowsiness and fatigue (i.e., measures of driving performance, subsidiary task performance, and physiological responses), researchers have turned to an investigation of the potential use of subjective ratings of fatigue and drowsiness. Their ease of implementation and low cost certainly make them attractive, and there is some evidence to suggest that subjective measures of drowsiness and fatigue do indeed correlate with performance decrements associated with fatigue and drowsiness (Ryder et al., 1981).

Subjective fatigue or drowsiness is usually measured through the use of a rating scale or a questionnaire which is completed or reported by the subject. Generally, these types of questionnaires or rating scales can be divided into three categories: non-dimensional single point measures, unidimensional rating scales, and multidimensional scales (Ryder et al., 1981). Each of these types of scales is discussed below.

Non-dimensional single point measures. These measures are the simplest of the three indicators; they require that the subject give just one single verbal report of fatigue or drowsiness while performing a task (Ryder et al., 1981). The studies which report the use of this type of measure have generally reported inconsistent findings. For instance, Pierson (1963) had subjects perform a reaction time task and indicate when they felt that their responses were becoming slower. While he did find a significant correlation between reaction time and this subjective measure of fatigue, he was unable to replicate these results. According to Kinsman and Weiser (1976), non-dimensional single point measures are ineffective because, in their simplicity they view fatigue as an all-or-none event, rather than as a relative occurrence which affects performance differently depending upon its degree or level.

Unidimensional rating scales. Unidimensional rating scales generally provide a much more reliable measure of fatigue. These scales can be divided into two types: ordinal and interval. In a unidimensional ordinal scale, the subject selects one adjective which describes fatigue and ranks it in order of magnitude (Ryder et al., 1981). For example, Nunney (1963) had subjects rate their levels of fatigue on a nine-point subjective scale ranging from 1--no fatigue, pleasant experience to 9--extreme fatigue, unpleasant experience. He found that fatigue as measured by pulse rate was significantly related to subjective ratings of fatigue between experimental groups. This relationship was not significant within groups however, thus suggesting great individual variation.

Muto (1981) and Wierwille and Muto (1981) were able to demonstrate the effectiveness of a unidimensional ordinal subjective rating scale as a valid index of alertness decrements. These researchers had subjects drive a simulator for 30, 60, or 150 minutes and then rate their level of alertness before and after the driving task on a seven-point scale with rating values ranging from (-3) extremely tired to (3) extremely alert. The results showed that the mean ratings on the scale before and after prolonged driving were significantly different, indicating that subjects judged themselves to be less alert after the driving task. Additionally, subjects rated themselves as being more tired as driving time was increased from 30 to 60 minutes.

Unidimensional interval scales are similar to ordinal scales except that the intervals between points on the scale are equidistant (Ryder et al., 1981). One unidimensional interval scale which has been employed successfully in several studies is the Feeling-Tone Checklist by Pearson and Byars (1956; cited in Ryder et al. 1981). This scale consists of an adjective checklist arranged on a 10-point interval scale ranging from "better than" to "same as" to "worse than." Fatigue, as measured by this Feeling-Tone Checklist has been found to correlate significantly with task duration (Nelson, Laden, and Carlson, 1979; cited in Ryder et al., 1981). Kinsman and Weiser (1976) maintain that Pearson and Byars' (1956) checklist is the best subjective scale which is currently available.

One unidimensional scale which quantifies subjective changes in sleepiness is the Stanford Sleepiness Scale (SSS). This scale is composed of seven statements which correspond to feelings of alertness and vitality at one end and feelings of extreme sleepiness at the other. The statements and scale values of the SSS are listed below (Hoddes, Zarcone, Smythe, Phillips, and Dement, 1973):

1 - Feeling active and vital; alert; wide awake

2 - Functioning at high level but not at peak; able to concentrate

3 - Relaxed; awake; not at full alertness; responsive

4 - A little foggy; not at peak; let down

5 - Fogginess; beginning to lose interest in remaining awake; slowed down

6 - Sleepiness; prefer to be lying down; fighting sleep; woozy

7 - Almost in reverie; sleep onset soon; lost struggle to remain awake

Hoddes et al. (1973) found that SSS ratings were significantly higher after sleep deprivation as compared to baseline days. Additionally, an examination of the relationship between ratings on the SSS and mental test scores before and after sleep deprivation revealed that the SSS correlated with mental test scores if the tests were long, monotonous, and boring.

Multidimensional scales. According to Ryder et al. (1981), many researchers feel that "an effective measure of fatigue must take into account more than just quantitative changes in perception" (p. 14). Kinsman and Weiser (1976) pointed out that fatigue also consists of qualitatively different subjective states that vary with not only the task involved but also the individual. Some researchers have thus developed multidimensional scales which take into account both quantitative and qualitative changes. These scales are usually developed from factor or cluster analysis which identify groups of symptoms comprising qualitatively different symptom categories (Ryder et al., 1981). For example, through the use of factor analysis, Wolf (1967) designed a multidimensional scale which identified three categories of fatigue: nervous (tense, jumpy, keyed up, head tightness, feel dizzy, irritable); exhaustion (physically tired, aching muscles, exhausted, easily distracted, no energy, perspiring); and drowsy (mentally sluggish, want to fall asleep, lazy, drowsy, feel sleepy, tired).

One other multidimensional scale was developed through the use of cluster analysis by Kinsman, Weiser, and Stamper (1973). These researchers had subjects exercise on a bicycle ergometer and then rate their subjective levels of fatigue. Three categories of fatigue were subsequently identified: fatigue (describing bodily feeling states associated with prolonged exercise); task aversion (which appeared to measure general levels of discomfort and disinclination to continue with task); and motivation (describing general drive and vigor).

Summary. The previous discussion seems to indicate that the use of unidimensional and multidimensional scales do appear to provide a reliable methodology for operationally defining fatigue and drowsiness. Their continued use in research settings is certainly warranted to provide further verification of their correspondence to more objective measures of fatigue and drowsiness.

Other Variables Affecting Driving Performance

The literature which has been reviewed up to this point has focused solely on the effects of prolonged driving and/or sleep deprivation on the driver in terms of his or her physiological and subjective responses as well as in terms of vehicle control and subsidiary task performance decrements. However, there are several other variables or issues which should be considered when conducting driving research. Among these are driver personality, gender, experience level, motivation, diurnal rhythms, instructions to subjects, and drugs. Each of these is briefly discussed below.

Driver personality.

A considerable amount of research has been devoted to the study of subject personality, specifically introversion-extroversion, and its effect on vigilance task performance (Davies and Parasuraman, 1981). In general, it has been found that the performance of extroverts deteriorates more rapidly than that of introverts. This effect has been explained as a difference in arousal level between the two groups; extroverts tend to become more bored and distracted more easily in monotonous situations. This effect has been empirically validated by Fagerstrom and Lisper (1977). These researchers found that extroverted drivers exhibited significantly greater reaction times to a subsidiary task during prolonged on-the-road driving than did introverted drivers. Additionally, in situations where the car radio was used for stimulation, only the extroverts benefited. This is consistent with Davies and Parasuraman's (1981) finding that "introducing varied auditory stimulation into a vigilance situation improves the performance of extroverts but not that of introverts" (p. 116).

Driver gender. Of the studies reviewed in this report, only one reported any differences in driver performance due to gender. Wetherell (1981) reported that performing a subsidiary task while driving tended to interfere with female, but not male, driving ability. Nonetheless, it would seem wise to control for gender differences whenever possible.

Driver experience level. There exists evidence which suggests that the experience level of the driver is an important factor which moderates the effect of prolonged driving on performance. Lisper et al. (1973) found that experienced drivers (> 20 000 km/year) exhibited a significant decrease in reaction times to a subsidiary task during three hours of driving while inexperienced drivers showed a significant increase. Fagerstrom and Lisper (1977) found that the increase in subsidiary task performance reaction times was greater for inexperienced (2500 km during prior six months) than experienced (10 000 km during prior six months) drivers who drove on-the-road for four hours, but that this difference was not significant. These researchers did find, however, that inexperienced drivers benefited more from a car radio that did experienced drivers.

Motivation. For those researchers who are interested in studying the effects of fatigue and/or drowsiness on driving, one variable which should be controlled (insofar as it is possible) is subject motivation. At least two groups of investigators (Mast et al., 1966; Safford and Rockwell; 1967) concluded that specifying the number of hours which the subject must drive may increase the subject's motivation and result in goal-setting. This may thus prevent fatigue effects from occurring as soon as they might otherwise. Hence, the prudent researcher should refrain from telling subjects how long they must drive.

Diurnal rhythms. A review of traffic accident rates plotted over the hours of the day shows a marked diurnal rhythm with one critical period being between 2 and 6 a.m. Some sources suggest that many of these accidents are caused by drivers falling asleep at the wheel (Lisper et al., 1979). Hoffman, Mayer, Grundei, and Meier (1971-72) presented a physiological basis for these findings when they provided evidence that the level of physiological arousal during highway driving was higher after midday than after midnight. Harris and Mackie (1972) further demonstrated a marked diurnal variation in physiological arousal with the lowest level occurring between 2 and 7 a.m. While these results do suggest that traffic accidents may result from lowered physiological arousal due to diurnal rhythms, these studies do not deal with diurnal rhythm as a single factor. Instead, diurnal rhythm in the performance capabilities of drivers has often been confounded with one or more additional factors (e.g., driving in daylight and darkness, driving in different traffic densities, and the number of hours since the last sleep period) (Lisper et al, 1979).

In an attempt to control these confoundings, Lisper et al. (1979) had subjects drive for three hours beginning at different times of day (0300, 0900, 1500, and 2100 hours). Since the experiment took place in Sweden, the experimenters were able to take advantage of the midnight sun so that subjects were always driving during daylight. The results showed that, while the typical diurnal rhythm was shown, and while the expected deterioration in reaction time to a subsidiary task due to time-on-task did occur, there were no significant differences in reaction times due to time of day. Thus, these researchers concluded that factors other than biological rhythm must be responsible for accidents in the early morning hours. For instance, it might be that "drivers in the early morning often have too many continuous hours behind the wheel, with a consequent lack of sleep" (Lisper et al., 1979, p. 4).

Instructions to subjects. Since the success of using subsidiary task performance measures as indices of driving impairment depends in part on the subject's knowledge of which task is the primary one, the experimenter must be explicit in his or her instructions to the subject. The experimenter must emphasize the importance of the subject giving top priority to the primary task of driving, even at the expense of impaired subsidiary task performance.

Drugs. Davies and Parasuraman (1981) note that drugs can have various effects on the performance of vigilance tasks. For instance, amphetamines, nicotine, and caffeine have been shown to reduce the amount of performance decrement in vigilance tasks. On the other hand, depressants such as alcohol either exacerbate or show no effect on performance decrement.

Simulator Validity

The proponents of automobile simulators argue that both economy and maximum experimental precision can be achieved through the use of driving simulators. For instance, traffic, roadway, and meteorological variables which may affect the precision of on-the-road driving experiments may be successfully controlled in simulator studies. Additionally, the safety of both the experimenter and subject is assured (Sussman et al., 1971). Those who oppose the use of simulators, however, voice the concern that simulators may not provide an accurate or adequate representation of the real world. In fact, Boadle (1976) suggested that the subject's knowledge of the lack of danger associated with simulator driving may produce results which are not representative of on-the-road driving. Hence, she suggested that further research be conducted to provide clear evidence that the simulated and real tasks have common elements and similar effects on performance.

Just such an experiment was conducted by Leonard and Wierwille (1975). These investigators provided both experimental and theoretical verification of the validity of automobile simulators as research tools by comparing a full-scale vehicle with an automobile simulator. The results showed that, for each performance measure, at least one simulator condition produced corresponding valid results. Thus, Leonard and Wierwille concluded that it is possible to obtain performance validation with a full-scale system if the simulator is properly adjusted.

Nonetheless, as Sussman et al. (1971) point out, each research problem requires that the researcher make a decision concerning the relative efficacy of simulated versus on-the-road driving. In any event, on-the-road validation of the findings of simulator studies should be seriously considered.

Summary

A review of the literature shows that several types of variables have been employed by researchers as measures of driving impairment due to drowsiness or fatigue. These variables include driving performance measures, measures of physiological responses, subsidiary task performance measures, and subjective ratings of drowsiness and/or fatigue. This same literature also provides mixed reports concerning the validity or predictive capability of these variables. In general, the results of the driving performance literature indicates that, of all the driving performance variables studied, both lane position variability and reaction time to driving-related events are significantly affected by both sleep deprivation and time-on-task. Additionally, eyelid closures have been shown to be correlated with impaired driving performance and thus provide a reliable, predictive measure of drowsiness. Similarly, researchers have successfully employed subsidiary task performance as measures of reduced spare mental capacity due to driving-induced fatigue and/or drowsiness, and have shown that these measures correlate highly with other driving performance measures. Finally, the literature dealing with subjective ratings of fatigue and drowsiness suggests that these types of measures can provide a reliable method for assessing the subjective effects of sleep deprivation and time-on-task.

RESEARCH OBJECTIVES

The main thrust of the reported research was to determine if a task subsidiary to driving could be employed to maintain driver alertness or to predict driver impairment due to fatigue and/or drowsiness. To accomplish this, three active drowsiness countermeasures (subsidiary tasks) which varied only according to input modality were compared in terms of: 1) the degree to which each countermeasure aided the driver by maintaining or restoring alertness; and 2) the degree to which driver responses to these subsidiary tasks correlated with eye closures and other measures of drowsiness. Hence, the primary goal of this research was to evaluate three subsidiary tasks to determine which, if any, showed potential for use as both an active countermeasure and as a means of detecting drowsiness when it does occur.

Additionally, a subjective measure of driver impairment was assessed for its potential use as additional verification of driver drowsiness.

METHODOLOGY

Experimental Design

The experimental design was a 4 X 5 complete factorial within subject design. The two factors in the design were type of subsidiary task and time-on-task. The subsidiary task factor had four levels--no subsidiary task, auditory task, visual task, and tactual task. To control for order effects, the presentation order of the subsidiary tasks was varied using repeated latin squares. The second factor, time-on-task, had five levels with each level corresponding to one six-minute interval during each of four 30-minute driving segments. An example of the experimental design and subsidiary task presentation order for one subject is shown in Figure 1.

Subjects

Six male and six female volunteer subjects participated in all experimental levels and conditions in this study. The subjects were screened initially using a driving and resting habits questionnaire (Appendix A; Dingus, Hardee, and Wierwille, 1985) to eliminate drivers who were not prone

Time-on-Task

		1st 6-min	2nd 6-min	3rd 6-min	4th 6-min	5th 6-min
Type -of- Task	Auditory					
	Visual					
	No Task					
	Tactual					

|------------- 30 minutes -------------|

Figure 1: Example of experimental design for one subject.

to drowsiness or who exhibited pathological sleep disorders. Each was required to hold a valid driver's license and to have at least 20/30 far visual (minimum separable) acuity, uncorrected or corrected with contact lenses, as measured by the Landholt C-Ring vision test. Additionally, each subject passed a hearing test requiring that he or she be able to detect at least a 40 dB, 500 hz tone presented to each ear alone. These subjects received $4.00 per hour (including sleep deprivation time as described in the procedure section) for their participation.

At this point it should be noted that, while the data from only 12 subjects were used in the subsequently reported analyses, two additional subjects participated in the study. The experimental session for one of these subjects was aborted after only 20 minutes of data collection due to his inability to remain alert enough to perform the driving task. The data from a second subject was discarded due to his refusal to comply with the experimenter's instructions.

Apparatus

Automobile simulator. The automobile simulator used for the reported research is the computer-controlled, moving base, automobile driving simulator located in the Vehicle Simulation Laboratory at Virginia Polytechnic Institute and

State University. This simulator is programmed to produce a realistic nighttime highway driving scenario and to handle like a midsized, rearwheel drive, American sedan. The simulator has four major systems: a motion platform, a roadway imaging/display system, an audio system, and an analog/hybrid/digital system. These systems, which are discussed more fully below, produce four degrees of platform motion, six degrees of freedom image/display motion, and four audio cues. The interested reader is referred to Wierwille (1975) for a more complete description of the design and operation of this simulator.

The simulator has a simulated cab which consists of a seat, brake and accelerator pedals, a dashboard with speedometer and fuel gauge, a steering wheel, and a display system.

The motion platform system provides four degrees of vehicular motion--yaw, roll, lateral, and longitudinal translation--through the use of closed-loop, servo-controlled, hydraulic actuators with estimated time delays of 25 msec. Engine drive-train vibration is also simulated by means of a motor with an eccentric mass; the velocity of this motor is controlled by the simulated engine speed.

The roadway imaging/display system produces the image of a two-lane road with a dashed centerline, right and left lane boundaries, and additional horizontal lines to give the road the appearance of being embedded in the horizontal plane. The roadway image is produced on a monochrome monitor and is viewed through a fresnel lens which is located directly in front of the driver and behind the simulator's windshield. The roadway image as seen by the driver appears to be at a distance of approximately 33 feet; the image also occupies the driver's entire foveal, and most of his or her peripheral field-of-view when the eyes are focused on the roadway. The image may be varied in six degrees of display motion--forward velocity, lateral position, yaw, roll, pitch, and inverse radius of curvature.

The audio system enhances simulator realism and driver feedback cues by simulating four separate sounds: tire rolling resistance, velocity-dependent engine and drive-train noise, tire screech on severe braking, and tire squeal on severe cornering.

An analog/hybrid computational system simulates vehicle dynamics through the use of steering, braking, accelerator, wind gust, and curvature signals. Computer outputs, including vehicle velocity, lateral position, yaw, roll, and pitch, update the motion platform position continuously through control of electrohydraulic servos. Additional signals are also provided to the speedometer, the audio system, the image generator, and the motion servos.

Additional simulator features. The simulator also includes digital-to-analog and analog-to-digital interfaces which link the simulator to two TRS-80 Model III microcomputers. These interfaces increase the flexibility of the simulator system by providing a means by which additional stimuli may be introduced into the driving scenario. Additionally, these interfaces allow for on-line data analysis.

Since the reported research called for the monitoring of drivers' eye closures, the simulator was modified to include a low-light level, closed-circuit television camera. This camera allowed for continuous, real-time monitoring of drivers' eyes; recording of closures was accomplished via the use of a linear potentiometer and an FM recorder. Since the camera does not block the driver's view of the roadway image, and since it requires no additional lighting, the camera and eyelid recording were unobtrusive.

Finally, the simulator seat was modified to include pressure transducers in the upper and lower portions of the backrest cushions. This modification allowed for the unobtrusive and accurate measurement of driver restlessness or seat movement as defined by changes in pressure against the back of the seat.

Subsidiary task apparati. This research also required subjects to perform three subsidiary tasks during the driving sessions. The apparati which were used for the tactual, visual, and auditory tasks are described below.

The implementation of the tactual task required that the steering wheel be modified to include four "tab" actuators on the back side of the steering wheel, with two on each end of the spoke (see Figure 2). An extension of the tab approximately 1/16 in represented a stimulus input. To provide additional tactual cues, these tabs vibrated at a frequency of 50 Hz when they were extended. The tactual stimuli presentation lasted for five seconds with an additional 14 seconds between stimulus presentations; hence, stimulus presentation occurred at the rate of one per 19 seconds.

Additionally, since a manipulative response to all stimuli (tactual, visual, and auditory) was desired, the steering wheel was modified to include one thumb-actuated momentary pushbutton on each side of the steering wheel spoke. As can be seen in Figure 3, these pushbuttons were positioned so that the subjects could perceive the tactual input and perform the manipulative response with their hands in a comfortable position.

Figure 2: Location of the tabs on the steering wheel for the tactual subsidiary task.

Figure 3: Location of the pushbuttons on the steering wheel and sound sources for the auditory subsidiary task.

As can be seen in Figure 4, the visual input for the visual task was accomplished via four yellow LED's located in a rectangular pattern and projected onto the windscreen from the top of the dash. By projecting the stimuli onto the windscreen, driver distraction from the roadway scene was minimized. Once again, the visual stimuli were presented for five seconds with an additional 14 seconds between stimulus presentations.

The auditory task apparati included four sound sources located in a rectangular pattern along the posts on each side of the windscreen (see Figure 3). These sound sources allowed for the sequential presentation of four different tones (tonic, third, fifth, and octave) starting with the speaker located on the lower left. To provide the driver with additional information concerning the "starting point" of the sequence of tones, the tonic tone (located at the lower left) lasted for two seconds while the other three tones lasted for one second each. Thus, the auditory stimulus presentation also lasted for a total of five seconds with an additional 14 seconds between stimulus presentations.

A special circuit board was designed to provide the necessary stimuli. This circuit board was interfaced with an IBM-PC microcomputer containing a special interface card

Figure 4: Location of LED's for the visual subsidiary task as well as the lights themselves projected onto the windscreen.

(MetraByte Model PIO12). The microcomputer, interface card, and circuit board were used to control the order of stimulus presentation and to collect subsidiary task response data.

Heart rate monitor. An ear plethysmograph and a Hewlett-Packard 7807C heart rate monitor were used to measure subjects' heart rates.

Driving Scenario

The driving scenario was 120 minutes in length and was composed of four identical 30-minute segments. For data collection purposes, these 30-minute segments were further divided into 10 3-minute intervals. To provide a more realistic highway driving scenario, the 30-minute segments were divided into six 5-minute sections during which alternately straight and curved roads were presented to the subject. The vehicle dynamics, motion and sound systems, and the visual scene and its perspective created the illusion of late night driving on a deserted interstate highway.

Procedure

Subject selection. Initially, interested subjects were asked to complete a driver questionnaire (Appendix A) similar to the one employed by Tilley et al. (1973).

Subjects' responses to this questionnaire provided information concerning subjects' driving and resting habits as well as biographical data. This information was treated confidentially and was used to screen out those subjects who were not prone to drowsiness or who exhibited pathological sleep disorders.

Pretesting. Once subjects had been screened and selected, they were notified for participation in the study. Upon their arrival at the Simulation Laboratory, the subjects were administered a Landholt C-Ring vision test. Far visual (minimum separable) acuity of at least 20/30 (uncorrected or corrected with contact lenses) was required to assure that the vehicle instrumentation and roadway image could be properly interpreted. Additionally, subjects were required to pass an audiometry test. A rectangular-wave tone with a 500 Hz fundmental frequency was presented to each subject's ear through headphones, and the subject indicated when he or she heard the tone and in which ear. Three ascending trials were administered for each ear, and the subject was required to detect at least an average of 40 dB in each ear. This criterion was deemed to be an appropriate one since the tones used during the auditory subsidiary task were 47.5 dBA, and the ambient noise level was 40 dB.

Once the subject had met the above criteria, he or she was required to read the instructions and to complete the Informed Consent Form found in Appendix B. The experimenter then arranged a time for the subject to return and complete the experimental driving session.

Performance of the subsidiary tasks. As stated before, the three subsidiary tasks used in this study varied only in terms of stimulus input modality; hence, the subjects' responses to each of the subsidiary tasks' stimuli were similar. For instance, during the performance of the tactual task, the subject was required to perceive not only how many tabs were extended, but also the location of the extended tabs. If more tabs were extended on the right than on the left, the subject was to respond by depressing the button located on the right side of the steering wheel; the opposite response was required if more tabs were extended on the left. If one tab was extended on each side of the steering wheel, the subject was required to depress both buttons. However, if all four tabs were extended, the subject was to make no response.

Similar responses were required for the visual task. If the subject perceived that more LED's were lit to the left of center, he or she was to respond by depressing the left pushbutton. The opposite response was required if more

LED's were lit to the right of center. If one LED was lit on each side, the subject was to depress both buttons; however, if all LED's were lit, the subject was to make no response.

Finally, during the performance of the auditory task, if more tones originated to the left of the subject, the subject was to depress the left pushbutton. The opposite response was required if more tones originated to the right. If one tone was heard on each side, the subject was to respond by depressing both pushbuttons. However, if all four tones were presented, the subject was not to respond.

General procedure. The subject was required to report to the Vehicle Simulation Laboratory at approximately 6:00 p.m. on the evening of the experimental session. If necessary, transportation to the Lab was provided by the research team. A member of the research team then purchased dinner for the subject at a local restaurant. Once the subject finished dinner, he or she was returned to the Lab to participate in a familiarization session.

Prior to the start of the practice session, the subject received detailed instructions concerning the performance of the three subsidiary tasks (Appendix C). The subject then participated in 15 practice trials for each task before the simulator was set in motion. Next, the subject was asked to drive normally in the right lane while maintaining a speed of 55 mph. He or she was also reminded that the performance of the subsidiary tasks was of secondary importance to that of driving, and that the subsidiary tasks should be performed only when they did not interfere with driving. Once these instructions had been presented, the subject began a 20-minute practice session during which he or she drove both normally (without performing any subsidiary tasks) for five minutes and while performing the three subsidiary tasks during three different 5-minute time intervals.

After the familiarization session was completed, the subject began the sleep deprivation period. During this time, the subject was allowed to read, study, watch TV, or listen to music. To eliminate the potential effects of stimulants on driving performance, the subject was not allowed to eat, smoke, or drink coffee or soft drinks. A research assistant monitored the subject to insure that he or she remained awake and awakened him or her from any lapses of drowsiness.

Shortly after midnight two fresh researchers joined the team, and the subject began the experimental driving session. This session lasted for approximately 2 1/2 hours. The first 30 minutes of driving was conducted <u>without</u> the subject performing any subsidiary tasks. The purpose of this first segment of driving was twofold: 1) to further fatigue and sleep deprive the subject; and 2) to provide further familiarization with the simulator. No data were collected during this segment.

After the preliminary 30-minute segment, the subject continued to drive for two hours and performed the three subsidiary tasks during three different 30-minute time intervals. The remaining interval consisted of normal driving during which no subsidiary task was performed. As was stated before, the order of presentation of the three subsidiary tasks and the normal driving segment was varied.

Additionally, at the end of each 30-minute driving segment, the subject was asked to verbally rate his or her mean level of alertness during the preceding driving segment using the alertness scale found in Appendix D. This was accomplished by having the subject continue to drive, while the experimenter prompted the subject by asking him or her to rate how alert he or she had felt during the previous driving segment. The subject responded verbally, and the experimenter recorded the subject's response. No driving or subsidiary task data were recorded while the subject gave subjective ratings. Once the rating was recorded, however, the subject was given instructions concerning the subsidiary task which would be performed during the next driving segment, and data collection was recommenced. This process for acquiring subjective alertness ratings required approximately one minute. Five subjective ratings of drowsiness were obtained in this manner.

Once the 2 1/2-hour experimental run was completed, the simulator motion was stopped, and the subject exited the simulator. The subject was then debriefed, paid, and driven home and told to sleep.

Debriefing. Subjects received $4.00 per hour for their actual time involved in the study, including sleep deprivation time. Those two subjects whose data were not used were also paid for the time they actually spent in the study.

Data Collection and Handling Techniques

Driving performance, physiological/behavioral, and subsidiary task performance measures were collected during each 3-minute interval through the use of analog-to-digital interfaces and microcomputers. Additionally, subjective ratings of driver alertness were collected at the end of each 30-minute driving segment. The data collection and handling techniques for these various dependent measures are described below.

Driving performance and physiological/behavioral measures. Twenty driving and physiological/behavioral measures were collected and reduced on-line via the use of two Radio Shack TRS-80 Model III microcomputers. The mean values for each of these measures were computed during 3-minute intervals and were displayed on the microcomputer screens. After each 3-minute driving interval these measures were copied by an experimenter onto data collection sheets and were later entered from these sheets onto a mainframe computer for analysis. To guard against loss of measures due to equipment failure, however, all important simulator signals were also recorded on a Honeywell 5600E multichannel instrumentation recorder.

Definitions of each of the 20 driving and physiological/behavioral measures collected, along with a description of the data collection techniques, follow. For the sake of clarity, these 20 variables will be referred to as the 'driving' measures throughout the rest of the report.

1. Eyelid Closure. During the experimental session the subject's eyes were monitored via a closed-circuit, low-light level TV camera and were displayed on a video monitor. An experimenter manually tracked eyelid closures using a linear potentiometer. The resulting analog signal was sampled at 0.25 second

intervals by a microcomputer, and the following dependent measures were calculated:

a) EYEMEAN - The mean of the eyelid closure signal, calculated for a 3-minute interval.

b) EYEMEAS - The mean square of the eye closure signal calculated for a 3-minute interval. Since this measure was a mean square, as eye closures increased, EYEMEAS was weighted more heavily.

c) PERCLOS - The percentage of time that the eyes were 80-100% closed during a 3-minute interval. These percentages were chosen since it seems fairly obvious that driving impairment should result if the eyes are closed to this extent.

2. Heart Rate - Heart rate was collected through the use of the ear plethysmograph described earlier. The resulting signal was sampled at 0.25 second intervals by a microcomputer, and the following dependent measures were computed:

   a) HRTRTM - The mean heart rate calculated over a 3-minute interval.

   b) HRTRTV - The heart rate variance calculated over a 3-minute interval.

3. Seat Movements - Pressure transducers located in the upper and lower portions of the backrest cushions

provided a signal associated with driver movements. The measure derived from this signal, SEATMOV, was a movement count, tallied over a 3-minute interval. The signal amplitude necessary to trigger the seat movement counter was set high enough so that only driver movements, but not vehicle vibrations and movements, were recorded.

4. Lane Deviation Measures - The lane position was an analog signal from the simulator and was measured at 0.25 second intervals by a microcomputer. The center of the right lane was given a value of zero. Deviations to the left of center resulted in positively increasing values while deviations to the right resulted in negatively increasing values. This lane position signal was used to compute the following dependent measures:
   a) LANEX - The number of lane position samples during a 3-minute interval with values greater than those equivalent to the centerline of the simulated vehicle exceeding either the right or left lane boundary.
   b) LANDEVM - The mean of the lane position signal, calculated over a 3-minute interval.

c) LANDEVV - The variance of the lane position signal, calculated over a 3-minute interval.

d) LANDEVSQ - The mean square of the lane position signal, calculated over a 3-minute interval. This measure weighted larger deviations from center more heavily than small deviations.

e) LANDEV4 - The mean of the fourth power of the lane position signal, calculated over a 3-minute interval. Since this measure was a fourth-power calculation, large deviations from center were very heavily weighted.

5. Yaw Deviation - Yaw deviation was measured as an analog signal of the difference between the simulated vehicle's heading and the roadway tangent. This value was sampled at 0.15 second intervals by one of the microcomputers, and the following dependent measures were derived from this signal:

   a) YAWMEAN - Mean yaw deviation, calculated over a 3-minute interval.

   b) YAWVAR - Yaw deviation variance calculated over a 3-minute interval.

   c) HIPASSYAW - The mean square of the output of a high pass filter with yaw deviation as the input to the filter. The filter's corner frequency was set at 0.1 radians/second.

6. Simulated Accelerometer - The lane position signal was used to create a measure simulating a high pass lateral position indicator based on a lateral accelerometer. This measure was sensitive to faster changes in lateral position and slowly returned to zero through the use of a recursion relation after a period of 60 seconds. This incoming signal was used to calculate the variable LATPOSMS, which was the high-pass mean square calculated over a 3-minute interval. Since this measure was a mean square calculation, rapid and large changes in lateral position were heavily weighted.
7. Steering Velocity - Steering wheel velocity was measured as an analog signal from the simulator and was sampled at 0.15 second intervals by a microcomputer. The following dependent measures were computed from this signal:
    a) STEXEED - The number of times steering velocity exceeded the equivalent of 150 degrees/second during a 3-minute interval.
    b) STVELM - Mean absolute steering velocity calculated over a 3-minute interval.
    c) STVELV - Steering velocity variance calculated over a 3-minute interval.

8. Steering Reversals - Steering wheel position was measured as an analog signal and was sampled at 0.15 seconds by a digital microcomputer. The following dependent measures were derived from this signal:
   a) LGREV - The number of times the magnitude of steering movement exceeded 5 degrees or more after steering wheel velocity passed through zero, tallied during a 3-minute interval.
   b) MDREV - The number of times the magnitude of steering wheel movement exceeded two degrees or more after steering wheel velocity passed through zero, tallied during a 3-minute interval.

Pre-analysis driving performance data reduction. Prior to data analysis, the driving data were reduced from 10 3-minute intervals to 5 6-minute intervals. This was accomplished by collapsing the data for each subject so that the 3-minute interval mean values for each variable in intervals one and two were combined to form a new, 6-minute mean. The same was done for intervals three and four, five and six, seven and eight, and nine and ten. In all subsequent reports of the data analysis, the data set used, 3- or 6-minute, is specified.

Subsidiary task performance measures. As was mentioned before, the subsidiary task stimuli were presented at the rate of one per 19 seconds. Therefore, during each 3-minute driving interval, the subject responded to nine task stimuli, resulting in a total of 90 stimulus presentations per 30-minute driving segment. Four subsidiary task response measures for each stimulus presentation were collected via an IBM-PC with its internal interface card and an external circuit board. These measures were printed out on a Radio Shack TRP-100 printer during the experimental session and were later entered on a mainframe computer for analysis. The TRP-100 was used because of its low noise output while printing. The four subsidiary task measures were:

1. The response time for the left pushbutton - For the visual and tactual tasks, this was computed as the time in seconds between stimulus onset and the subject's response. For the auditory task, however, response time was calculated as the time between the beginning of the fourth tone and the subject's response. If the fourth tone was not presented, the response time was computed as the time between the beginning of the fourth tone if it would have been presented and the subject's response.
2. The response time for the right pushbutton - This was computed in the same manner as the left pushbutton response time.

3. The response time for both pushbuttons together - In those instances where the subject responded by pressing both pushbuttons simultaneously, this single measure of response time was used. In those cases where the subject responded slightly more quickly with one hand however, the longer response time of the two pushbuttons was used as a measure of response time for both hands.
4. The subject's response to the stimulus (i.e., right, left, or both) was also recorded.

Pre-analysis subsidiary task data reduction. Prior to data analysis it was necessary to reduce the subsidiary task data so that it could be used in subsequently reported correlation analyses. To accomplish this, five new variables were created. These variables, which were used in all subsequent analyses, are listed and described below.

1. RTLEFT - The mean response time in seconds for the left pushbutton, calculated over each 3-minute interval. Any data corresponding to an incorrect response by the subject was disregarded.
2. RTRIGHT - The mean response time in seconds for the right pushbutton, calculated over each 3-minute interval. Once again, incorrect responses were not included.

3. RTBOTH - The mean response time in seconds for both pushbuttons, calculated over each 3-minute interval. Incorrect responses were not included.
4. RTMEAN - The overall mean response time per 3-minute interval computed as: (RTLEFT + RTRIGHT + RTBOTH)/3.
5. NUMHIT - The number of correct responses per 3-minute interval.

Additionally, these same new variables were computed for 6-minute intervals as described in the previous section.

<u>Subjective alertness ratings</u>. Finally, subjects' verbal responses to a subjective alertness rating scale were recorded by the experimenter at the end of each 30-minute driving segment.

RESEARCH HYPOTHESES

Given the previously described research objectives and experimental methodology, several hypotheses were developed. First, it was anticipated that subjects would demonstrate significantly greater driving impairment during the "no task" driving segment than during those segments when they performed the subsidiary tasks. Prior to the experiment, however, it was not possible to predict which subsidiary task would provide increased driver alertness and less drowsiness-induced impairment. It was hypothesized, however, that, since driving itself requires the processing of visual information, the visual subsidiary task may not be associated with a reduction in driving impairment but may actually be associated with an increase in impairment. This result could be expected if the performance of the visual subsidiary task interfered with that of driving.

To determine the value of using subsidiary task performance measures as indicators of driving impairment, it was necessary to demonstrate relatively high correlations between subjects' reaction times to the subsidiary tasks and other measures (such as eyelid closure, yaw deviation, and lane standard deviation) which have been shown to be sensitive to drowsiness-induced impairment (Skipper et al., 1984). Since the reviewed subsidiary task literature does not include variables such as yaw deviation and lane standard deviation, it was difficult to make a prediction concerning the outcome of this analysis. However, given that other researchers have found significant correlational relationships between subsidiary task performance and other types of driving performance variables, it seemed probable that similar results would be obtained in this instance also.

Finally, it was hypothesized that subjects would rate themselves as being less alert during intervals in which no subsidiary task was presented. Additionally, it was expected that they would rate themselves as being less alert as the duration of the driving task increased. These latter results, if obtained, would be similar to those found by Muto (1981).

RESULTS

It should be recalled that the major objectives of this research were twofold: 1) to compare the three subsidiary tasks to determine which, if any, aided the driver by helping to maintain alertness; and 2) to determine if driver responses to the subsidiary tasks could be used to detect drowsiness and/or driving impairment. Along with these two major research questions was the issue involving the use of subjective rating scales to verify driver drowsiness. Given these objectives, it was necessary to conduct the data analyses in three stages to answer these research questions.

First, the value of using subsidiary tasks as drowsiness countermeasures was assessed through the use of multivariate analyses of variance (MANOVAs). The results of these MANOVAs were further explored by conducting univariate analyses of variance (ANOVAs) and Newman-Keuls Sequential Range Tests.

Second, to determine the predictive ability of subsidiary task performance measures, correlation analyses were first conducted to show which driving measures were sensitive to drowsiness as indicated by their correlations with eye measures. The subsidiary task measures were then correlated with both the eye and driving measures to determine which subsidiary task measures were also sensitive to drowsiness. Next, several stepwise and linear discriminant analyses were performed in an attempt to build a predictive model which could be used to classify driving performance as being either impaired or not impaired by drowsiness.

Third, the subjective ratings of alertness were analyzed via the use of two nonparametric Kruskal-Wallis tests to determine if these ratings were affected by the performance of the subsidiary tasks or by time-on-task.

Finally, the driving data were also subjected to a post hoc analysis aimed at determining the effect, if any, of driver experience on driving behavior.

To facilitate understanding, the results of the three stages of data analysis are presented along with the associated research question in the following sections. The results of the post hoc analyses are also included.

Subsidiary Tasks as Drowsiness Countermeasures

MANOVAs. Of primary interest to this research was the effect of subsidiary task performance on driver alertness and driving. In an attempt to assess this effect, three 4 x

5 MANOVAs were conducted using type of task and 6-minute driving interval as the independent variables and the 20 6-minute driving variables as the dependent variables in these analyses.

Since, to conduct a MANOVA it is necessary to include no more than n-1 (or in this case, 12-1=11) dependent variables, and since statistical power varies inversely with the number of dependent measures relative to the number of subjects, it was decided to divide the 20 driving variables into three groups, thus taking advantage of increased statistical power associated with the lowered ratio of number of subjects to number of dependent measures. The following three groups of variables were thus analyzed in three separate MANOVAs:

1. The physiological/behavioral variables - PERCLOS, EYEMEAN, EYEMEAS, SEATMOV, HRTRTM, and HRTRTV;
2. The lane/yaw variables - LANEX, LANDEVM, LANDEVV, LANDEVSQ, LANDEV4, YAWMEAN, YAWVAR, LATPOSMS, and HIPASSYAW;
3. The steering variables - STEXEED, STEVLM, STVELV, LGREV, and MDREV.

The results of these three MANOVAs (shown in Tables 1, 2, and 3) indicate a significant interval effect for each of the three categories of variables which far exceeds the a

TABLE 1

MANOVA Summary for Physiological/Behavioral Variables.

Source	dv	df_H	df_E	U	F	p
Between Subject						
Subj						
Within Subject						
Interval	6	24	137	0.0841	5.91	0.0001
Interval x Subj						
Task	6	18	79	0.5007	1.23	0.2617
Task x Subj						
Interval x Task	6	72	696	0.6147	0.91	0.6957
Interval x Task x Subj						

where: dv = number of dependent measures

df_H = degrees of freedom for treatment effect

df_E = degrees of freedom for error effect

U = Wilks' likelihood ratio statistic

$$\frac{|E|}{|E + H|}$$, where:

$|E|$ = determinant of sum of squares and cross-products matrix for error

$|E + H|$ = determinant of the sum of the sum of squares and cross-products matrix for error, and the sum of squares and cross-products matrix for treatment

TABLE 2

MANOVA Summary for Lane/Yaw Variables.

Source	dv	df_H	df_E	*U*	*F*	*p*
Between Subject						
Subj						
Within Subject						
Interval	9	36	136	0.0835	3.57	0.0001
Interval x Subj						
Task	9	27	73	0.3634	1.13	0.3319
Task x Subj						
Interval x Task	9	108	915	0.4459	0.99	0.5010
Interval x Task x Subj						

where: dv = number of dependent measures

df_H = degrees of freedom for treatment effect

df_E = degrees of freedom for error effect

U = Wilks' likelihood ratio statistic

$$\frac{|E|}{|E + H|}$$, where:

$|E|$ = determinant of sum of squares and cross-products matrix for error

$|E + H|$ = determinant of the sum of the sum of squares and cross-products matrix for error, and the sum of squares and cross-products matrix for treatment

TABLE 3

MANOVA Summary for Steering Variables.

Source	dv	df_H	df_E	U	F	p
Between Subject						
Subj						
Within Subject						
Interval	5	20	133	0.0044	27.61	0.0001
Interval x Subj						
Task	5	15	80	0.6205	1.01	0.4519
Task x Subj						
Interval x Task	5	60	603	0.6359	1.02	0.4377
Interval x Task x Subj						

where: dv = number of dependent measures

df_H = degrees of freedom for treatment effect

df_E = degrees of freedom for error effect

U = Wilks' likelihood ratio statistic

$$\frac{|E|}{|E + H|}$$, where:

$|E|$ = determinant of sum of squares and cross-products matrix for error

$|E + H|$ = determinant of the sum of the sum of squares and cross-products matrix for error, and the sum of squares and cross-products matrix for treatment

priori criterion for significance of $p=0.05$ ($F=5.91$, $p=0.0001$ for the physiological/behavioral variables; $F=3.57$, $p=0.0001$ for the lane/yaw variables; and $F=27.61$, $p=0.0001$ for the steering variables).[1] No task or interaction effects were indicated, however. This absence of a task effect was a disappointing one since it indicated that the hypothesis that performance of subsidiary tasks would affect driving performance was not supported.

One-Way ANOVAs. Given this significant interval effect for each of the MANOVAs, 20 one-way ANOVAs were next conducted to assess which of the dependent variables made significant contributions to this main effect of interval. The results of these ANOVAs are grouped according to category of dependent measure (physiological/behavioral, lane/yaw, steering) and are presented in Tables 4, 5, and 6, respectively.

These one-way ANOVAs indicate that all 20 dependent measures contributed significantly to the interval main effect at $p=0.05$ or less. More specifically, the variables EYEMEAN, EYEMEAS, PERCLOS, YAWMEAN, HIPASSYAW, STVELM, STVELV, LGREV, and MDREV showed p-values less than $p=0.0002$

---

[1] It should be noted at this point that the reported F-values for these and all subsequent MANOVAs are approximate F's which are based on Wilks' Liklihood Ratio statistic (U). The interested reader is referred to Tabachnick and Fidell (1983) for a detailed discussion of these multivariate test statistics.

TABLE 4

Summary of One-Way ANOVAs for Effects of Interval on Physiological/Behavioral Variables.

Variable	df_F	SS_{source}^{*}	SS_{error}^{*}	F	p
EYEMEAN	4,44	18 252.215	9107.078	22.05	0.0001
EYEMEAS	4,44	1.24×10^9	8.36×10^8	16.26	0.0001
PERCLOS	4,44	1138.765	1479.827	8.46	0.0001
HRTRTM	4,44	45.758	141.699	3.55	0.0135
HRTRTV	4,44	4867.048	9751.626	5.49	0.0011
SEATMOV	4,44	3696.100	7827.675	5.19	0.0016

* Source of variation for all ANOVAs is Interval.

** Error term for all ANOVAs is Subject x Interval.

TABLE 5

Summary of One-Way ANOVAs for Effects of Interval on Lane/Yaw Variables.

Variable	df_F	SS_{source}*	SS_{error}**	*F*	*p*
LANEX	4,44	9644.6292	27 513.0208	3.86	0.0090
LANDEVM	4,44	112.4453	298.4825	4.14	0.0062
LANDEVV	4,44	5.56×10^5	1.82×10^6	3.36	0.0175
LANDEVSQ	4,44	6.21×10^5	1.96×10^6	3.49	0.0147
LANDEV4	4,44	5.60×10^{13}	2.27×10^{14}	2.71	0.0422
YAWMEAN	4,44	0.5097	0.7466	7.51	0.0001
YAWVAR	4,44	1.03×10^5	1.77×10^5	6.38	0.0004
LATPOSMS	4,44	4.01×10^5	1.31×10^6	3.38	0.0171
HIPASSYAW	4,44	2.66×10^5	4.18×10^5	7.00	0.0002

* Source of variation for all ANOVAs is Interval.

** Error term for all ANOVAs is Subject x Interval.

TABLE 6

Summary of One-Way ANOVAs for Effects of Interval on Steering Variables.

Variable	df_F	SS_{source}*	SS_{error}**	*F*	*p*
STEXEED	4,44	113.69	210.56	5.94	0.0007
STVELM	4,44	1.11	0.05	240.93	0.0001
STVELV	4,44	72 172.98	46 607.46	17.03	0.0001
LGREV	4,44	3263.42	1008.80	35.56	0.0001
MDREV	4,44	37 756.58	4377.37	94.88	0.0001

* Source of variation for all ANOVAs is Interval.

** Error term for all ANOVAs is Subject x Interval.

and were thus greatly affected by driving interval. All other variables, except for LANDEV4, also demonstrated p-values well below 0.05. LANDEV4 was least affected by interval, since it was associated with a p-value of 0.0422.

Newman-Keuls Sequential Range Tests. While the previously described statistical analyses did provide information indicating that driving interval reliably affected all performance measures, these analyses did not determine the exact loci of this significance. To further investigate significant differences among pairs of treatment means, Newman-Keuls Sequential Range Tests were performed. This multiple comparison procedure was chosen because it provides a sensitive test between means while still maintaining the protection level at alpha for each set of ordered comparisons (Williges, 1982).

The mean values associated with each level of driving interval are presented in Tables 7, 8, and 9 for each of the dependent measures, along with the results of the associated Newman-Keuls test. As before, the results are presented according to category of dependent measure. For each comparison, a criterion of significance of $p<0.05$ was used. These results indicate that for all variables except LANDEV4, at least one pair of mean values differ significantly with probability $p=0.05$. The lack of

TABLE 7

Newman-Keuls Results for Driving Interval Using Physiological/Behavioral Measures.*

EYEMEAN

Interval 5	Interval 4	Interval 3	Interval 2	Interval 1
133.93	128.53	126.05	124.30	107.97
A	A			C
	B	B	B	

EYEMEAS

Interval 5	Interval 4	Interval 3	Interval 2	Interval 1
20 217	18 474	17 981	17 224	13 363
A	A			C
	B	B	B	

PERCLOS

Interval 5	Interval 4	Interval 2	Interval 3	Interval 1
11.47	8.62	8.33	7.68	4.65
A	B	B	B	C

---

*Mean values designated by different letters are significantly different with $p < 0.05$.

TABLE 7

Newman-Keuls Results for Driving Interval Using Physiological/Behavioral Measures (cont.).*

HRTRIM

Interval				
4	3	5	2	1
61.84	61.80	61.71	61.44	60.66
A	A	A	A	B

HRTRTV

Interval				
3	2	4	5	1
36.11	35.17	34.98	33.49	23.38
A	A	A	A	B

SEATMOV

Interval				
5	4	3	2	1
19.98	18.46	16.99	14.77	8.73
A	A	A	A	B

---

*Mean values designated by different letters are significantly different with $p < 0.05$.

TABLE 8

Newman-Keuls Results for Driving Interval Using Lane/Yaw Variables.*

LANEX

Interval				
4	2	5	3	1
24.77	20.14	17.89	16.09	5.64
A	A	A B	A B	B

LANDEVM

Interval				
3	5	2	4	1
8.94	8.91	8.19	7.41	7.39
A	A	A B	B	B

LANDEVV

Interval				
4	2	5	3	1
270.16	222.42	170.86	163.52	134.64
A	A B	A B	A B	B

*Mean values designated by different letters are significantly different with $p < 0.05$.

TABLE 8

Newman-Keuls Results for Driving Interval Using Lane/Yaw Variables (cont.).*

LANDEVSQ

Interval				
4	2	5	3	1
352.27	325.64	278.71	266.24	205.20
A	A	A	A	
		B	B	B

LANDEV4

Interval				
4	2	5	3	1
1 419 456	1 282 434	872 669	463 784	138 625
A	A	A	A	A

YAWMEAN

Interval				
3	4	1	5	2
0.657	0.605	0.591	0.569	0.516
A	A		C	C
	B	B	B	

YAWVAR

Interval				
4	2	3	5	1
119.09	83.38	80.13	76.07	55.00
A	B	B	B	B

---

*Mean values designated by different letters are significantly different with $p < 0.05$.

TABLE 8

Newman-Keuls Results for Driving Interval Using Lane/Yaw Variables (cont.).*

LATPOSMS

Interval

4	2	5	3	1
229.47	195.22	157.10	134.77	116.75
A	A	A		
	B	B	B	B

HIPASSYAW

Interval

4	3	2	5	1
193.87	131.50	130.50	125.05	90.91
A	B	B	B	B

---

*Mean values designated by different letters are significantly different with $p < 0.05$.

TABLE 9

Newman-Keuls Results for Driving Interval Using Steering Variables.*

STEXEED

Interval				
4	3	5	2	1
2.104	1.802	0.740	0.604	0.396
A	A	B	B	B

STVELM

Interval				
3	1	4	5	2
1.49	1.48	1.46	1.39	1.31
A	A	B	C	D

STVELV

Interval				
4	3	1	2	5
100.46	86.35	66.98	61.07	53.14
A	B	C	C	C

*Mean values designated by different letters are significantly different with $p < 0.05$.

TABLE 9

Newman-Keuls Results for Driving Interval Using Steering Variables (cont.).*

LGREV

Interval				
4	3	1	2	5
17.20	14.32	12.21	9.60	6.54
A	B	C	D	E

MDREV

Interval				
4	1	3	2	5
82.16	75.03	69.62	57.52	47.04
A	B	C	D	E

*Mean values designated by different letters are significantly different with $p < 0.05$.

significant differences between mean values for the variable LANDEV4 may be due in part to the relatively large (i.e., relatively large compared to the corresponding p-values associated with the remaining variables) p-value associated with the interval mean effect shown in the ANOVA ($p=0.0422$). Alternatively, since the Newman-Keuls test is not the most sensitive of the tests for paired comparisons (Williges, 1982), it may be that the use of another test such as the Least Significant Difference test would have demonstrated a significant difference between at least one pair of treatment conditions.

The three eye closure measure means demonstrate a general increase in the number of closures as the driving segment progressed, along with a significant increase in the number of closures in interval 5 compared to interval 1. The three lane variables LANEX, LANDEVV, and LANDEVSQ also exhibit a significant increase in the number of lane deviations at the end of the driving segment compared to interval 1.

Additionally, HRTRTM, HRTRTV, and SEATMOV show their smallest mean values for interval 1 with a significant increase in means associated with interval 2. For all three of these variables, the mean values do not change significantly during intervals 2 to 5.

Finally, the interval means associated with the remaining 10 variables tend to demonstrate a cyclic pattern, with the smallest mean value usually occurring in interval 1 or 2. This pattern is most clearly represented by the mean values of the variables LGREV and MDREV. These variables were the most sensitive to time-on-task effects, since the number of steering reversals changed significantly from interval to interval for both of these variables. The mean values for LANDEVM, YAWMEAN, YAWVAR, LATPOSMS, HIPASSYAW, STEXEED, STVELM, and STVELV also generally alternately increase and decrease across the five intervals indicating a rise and fall in subject performance during the 30-minute driving segment. This phenomenon of cyclic peaks and valleys in the performance of sleep-deprived subjects is similar to the one described by Davies and Parasuraman (1981).

Summary. The results of the MANOVAs indicate that the subsidiary tasks did not significantly affect driving performance as measured by the dependent measures. Hence, the hypothesis that the tasks would act as drowsiness countermeasures was not supported. However, it was found that driving interval accounted for a significant amount of variance in driving performance. The 20 one-way ANOVAs showed that all 20 dependent measures made significant contributions to this main effect of interval. Newman-Keuls multiple comparison tests further revealed that, for 10 of the driving variables the mean interval values demonstrated a cyclic effect of increasing and decreasing alertness which is consistent with other research findings (Davies and Parasuraman, 1981). The mean interval values of EYEMEAN, EYEMEAS, and PERCLOS also significantly increased from interval 1 to interval 5.

Subsidiary Task Performance Measures as Predictors of Impairment

A second series of analyses was undertaken to determine if driver responses to the subsidiary tasks could be used to predict drowsiness and/or driving impairment. To accomplish this, several correlation and discriminant analyses were performed. The first of the correlation analyses was conducted to determine which of the 17 driving measures correlated most highly with the eye closure measures and thus may be predictive of drowsiness-induced impairment. Next, the subsidiary task performance measures were also correlated with both the eye measures and the driving performance measures to determine if subsidiary task performance was associated with eye closures and/or other driving variables which were found to be sensitive to drowsiness.

Finally, several discriminant analyses were conducted using combinations of driving and subsidiary task performance measures which appeared to be most sensitive to drowsiness. The purpose of these analyses was to determine if a model using combinations of the driving and subsidiary task performance measures could be used to classify subjects as being either impaired by drowsiness or not impaired.

A report of the results of these correlation and discriminant analyses follows. It should be noted that all correlation analyses were conducted using both 3-minute and 6-minute interval data. This was done to determine if the length of data collection time substantially affected the size of the correlations and hence, the predictive ability of the variables. Also, the correlations were computed by subsidiary task; in other words, each reported correlation matrix includes data from only one of the three subsidiary tasks. This was done to assess the relative predictive ability of the measures associated with each of the tasks.

Three-minute correlations between eye and other driving measures for the auditory task. The first step in the correlation analysis was to correlate the eye measures with other driving measures for each task to determine which driving measures were most sensitive to drowsiness. This was accomplished using 3- and 6-minute data for each subsidiary task.

The 3-minute correlations between the eye and other driving measures for the auditory task are presented in Table 10. As shown, the highest correlations exist between the PERCLOS measure and the lane and yaw variables, with correlations in excess of $r=0.50$ for LANEX, LANDEVV, LANDEVSQ, LANDEV4, YAWVAR, LATPOSMS, and HIPASSYAW. Likewise, the correlations between EYEMEAS and these same variables, with the exception of LANDEV4, are also greater than $r=0.50$. The correlations between these variables and EYEMEAN are somewhat lower ($r>0.40$), however. Additionally, relatively high correlations exist between HRTRTV and EYEMEAN and EYEMEAS ($r=-0.496$ and $r=-0.430$, respectively.) These correlations are very encouraging since, given the degree of inter-subject variability associated with driving performance (e.g., Safford and Rockwell, 1967; Yajima et al., 1976), correlations in excess of 0.40 can be considered to be high. In the case of the current research, all correlations greater than $r=0.40$ are significant with $p<0.009$. For this reason, correlations in excess of $r=0.40$ are highlighted in these and all subsequent analyses.

Six-minute correlations between eye and other driving measures for the auditory task. Table 11 reveals the same pattern of correlations between the eye and other driving measures as that exhibited by the 3-minute data. However, a

TABLE 10

Three-Minute Correlations Between Eye and Driving Measures for Auditory Task.

	EYEMEAN	EYEMEAS	PERCLOS
HRTRTM	-0.230	-0.276	-0.331
HRTRTV	-0.496	-0.430	-0.218
SEATMOV	0.120	0.099	0.049
LANEX	0.549	0.619	0.725
LANDEVM	0.196	0.218	0.255
LANDEVV	0.468	0.532	0.647
LANDEVSQ	0.530	0.597	0.714
LANDEV4	0.377	0.435	0.542
YAWMEAN	0.035	-0.011	-0.139
YAWVAR	0.477	0.535	0.641
LATPOSMS	0.491	0.555	0.667
HIPASSYAW	0.496	0.554	0.655
STEXEED	0.205	0.203	0.203
STVELM	0.203	-0.224	-0.259
STVELV	0.213	0.205	0.187
LGREV	0.238	0.232	0.218
MDREV	0.029	0.033	0.052

Correlations greater than 0.200 are significant with $p \leq 0.05$.

TABLE 11

Six-Minute Correlations Between Eye and Driving Measures for Auditory Task.

	EYEMEAN	EYEMEAS	PERCLOS
HRTRTM	-0.226	-0.274	-0.336
HRTRTV	-0.525	-0.450	-0.211
SEATMOV	0.172	0.151	0.102
LANEX	0.579	0.658	0.802
LANDEVM	0.256	0.286	0.352
LANDEVV	0.515	0.588	0.729
LANDEVSQ	0.586	0.664	0.816
LANDEV4	0.482	0.555	0.693
YAWMEAN	0.100	0.029	-0.187
YAWVAR	0.545	0.611	0.731
LATPOSMS	0.535	0.607	0.744
HIPASSYAW	0.555	0.621	0.740
STEXEED	0.277	0.283	0.294
STVELM	-0.215	-0.253	-0.327
STVELV	0.284	0.288	0.292
LGREV	0.288	0.292	0.300
MDREV	0.103	0.119	0.161

Correlations greater than 0.253 are significant with $p \leq 0.05$.

comparison of the 3- and 6-minute data shows that the 6-minute correlations tend to be generally higher. In fact, two measures, LANEX and LANDEVSQ, show correlations with PERCLOS in excess of 0.80. Also, according to Tables 10 and 11, the variables HRTRTV, LANEX, LANDEVV, LANDEVSQ, LANDEV4, LATPOSMS, and HIPASSYAW tend to correlate most highly with eye closure when the auditory task was performed, and may thus be predictive of driving impairment.

Three-minute correlations between eye and other driving measures for the tactual task. Table 12 indicates much lower correlations between the eye and other driving measures than those for the auditory task. However, several different driving variables exhibit correlations with the eye measures in excess of 0.40. For instance, YAWVAR correlates highly with both EYEMEAS and PERCLOS (*r*=0.456 and *r*=0.504, respectively). Additionally, both STVELV and LGREV correlate highly with EYEMEAS and PERCLOS (*r*=0.460 and *r*=-0.469, respectively for STVELV; *r*=0.538 and *r*=0.580, respectively for LGREV). Once again, HIPASSYAW also correlates highly with both EYEMEAS and PERCLOS.

Six-minute correlations between eye and other driving measures for the tactual task. The 6-minute correlations for the tactual task are presented in Table 13. A comparison of Tables 12 and 13 indicates similar patterns of

TABLE 12

Three-Minute Correlations Between Eye and Driving Measures for Tactual Task.

	EYEMEAN	EYEMEAS	PERCLOS
HRTRTM	-0.103	-0.147	-0.219
HRTRTV	-0.159	-0.087	0.056
SEATMOV	0.029	0.098	0.224
LANEX	-0.090	-0.024	0.099
LANDEVM	-0.306	-0.225	-0.044
LANDEVV	0.032	0.067	0.123
LANDEVSQ	-0.151	-0.078	0.063
LANDEV4	0.038	0.054	0.106
YAWMEAN	-0.091	-0.124	-0.237
YAWVAR	0.365	0.456	0.504
LATPOSMS	0.084	0.129	0.198
HIPASSYAW	0.357	0.452	0.499
STEXEED	0.219	0.381	0.306
STVELM	-0.043	-0.005	-0.067
STVELV	0.313	0.460	0.469
LGREV	0.429	0.538	0.580
MDREV	0.210	0.291	0.383

Correlations greater than 0.200 are significant with $p \leq 0.05$.

TABLE 13

Six-Minute Correlations Between Eye and Driving Measures for Tactual Task.

	EYEMEAN	EYEMEAS	PERCLOS
HRTRTM	-0.102	-0.148	-0.217
HRTRTV	-0.168	-0.091	0.070
SEATMOV	0.030	0.122	0.265
LANEX	-0.091	-0.026	0.132
LANDEVM	-0.316	-0.243	-0.036
LANDEVV	0.058	0.092	0.162
LANDEVSQ	-0.155	-0.088	0.089
LANDEV4	0.099	0.114	0.176
YAWMEAN	-0.048	-0.115	-0.278
YAWVAR	0.423	0.498	0.583
LATPOSMS	0.113	0.155	0.242
HIPASSYAW	0.414	0.492	0.577
STEXEED	0.274	0.409	0.371
STVELM	0.032	0.046	-0.036
STVELV	0.360	0.494	0.545
LGREV	0.501	0.605	0.686
MDREV	0.296	0.390	0.523

Correlations greater than 0.253 are significant with $p \leq 0.05$.

correlation. Once again, the 6-minute correlations are generally higher than the 3-minute correlations. Also, the 6-minute correlations show relatively high correlations between STEXEED and EYEMEAS (*r*=0.409) and between MDREV and PERCLOS (*r*=0.523). In summary, the variables YAWVAR, HIPASSYAW, STVELV, LGREV, MDREV, and perhaps STEXEED correlate most highly with the eye closure measures for the tactual task.

Three-minute correlations between eye and other driving measures for the visual task. As can be seen in Table 14, the correlations between HRTRTM, LANEX, LANDEVSQ, YAWVAR, LATPOSMS, HIPASSYAW, STVELV, and LGREV and all three eye measures are in excess of *r*=0.40. In addition, LANDEVV correlates highly with both EYEMEAS and PERCLOS (*r*=0.417 and *r*=0.419, respectively), while MDREV correlates highly with EYEMEAN (r=0.429).

Six minute correlations between eye and other driving measures for the visual task. Once again, a comparison of Table 14 with the 6-minute correlations in Table 15 indicates a general increase in the magnitude of the correlations. While a similar pattern of correlations exists for both the 3- and 6-minute correlations, several other variables demonstrate correlations with all three eye measures in excess of 0.40. These variables are SEATMOV,

TABLE 14

Three-Minute Correlations Between Eye and Driving Measures for Visual Task.

	EYEMEAN	EYEMEAS	PERCLOS
HRTRTM	-0.454	-0.458	-0.488
HRTRTV	-0.350	-0.270	-0.136
SEATMOV	0.399	0.340	0.298
LANEX	0.417	0.444	0.440
LANDEVM	0.110	0.097	0.087
LANDEVV	0.392	0.417	0.419
LANDEVSQ	0.415	0.437	0.431
LANDEV4	0.313	0.337	0.343
YAWMEAN	-0.148	-0.170	-0.227
YAWVAR	0.511	0.525	0.524
LATPOSMS	0.412	0.438	0.440
HIPASSYAW	0.534	0.548	0.545
STEXEED	0.385	0.335	0.345
STVELM	-0.146	-0.134	-0.151
STVELV	0.525	0.469	0.481
LGREV	0.619	0.552	0.591
MDREV	0.429	0.385	0.385

Correlations greater than 0.200 are significant with $p \leq 0.05$.

TABLE 15

Six-Minute Correlations Between Eye and Driving Measures for Visual Task.

	EYEMEAN	EYEMEAS	PERCLOS
HRTRTM	-0.462	-0.480	-0.498
HRTRTV	-0.363	-0.294	-0.135
SEATMOV	0.501	0.430	0.420
LANEX	0.462	0.501	0.502
LANDEVM	0.132	0.119	0.116
LANDEVV	0.438	0.477	0.485
LANDEVSQ	0.469	0.504	0.507
LANDEV4	0.369	0.406	0.423
YAWMEAN	-0.166	-0.214	-0.274
YAWVAR	0.589	0.618	0.622
LATPOSMS	0.459	0.497	0.504
HIPASSYAW	0.611	0.642	0.643
STEXEED	0.526	0.509	0.499
STVELM	-0.114	-0.097	-0.125
STVELV	0.625	0.594	0.606
LGREV	0.728	0.707	0.731
MDREV	0.548	0.523	0.535

Correlations greater than 0.253 are significant with $p \leq 0.05$.

LANDEVV, STEXEED, and MDREV. Additionally, LANDEV4 correlates highly with both EYEMEAS and PERCLOS. Hence, for the visual task, the variables HRTRTM, SEATMOV, LANEX, LANDEVV, LANDEVSQ, LANDEV4, YAWVAR, LATPOSMS, HIPASSYAW, STEXEED, STVELV, LGREV, and MDREV indicate potential as drowsiness predictors.

Summary of correlation analyses of eye and other driving variables. The results of the previously described correlation analyses are presented in Table 16. This table provides a list of those driving measures whose correlations with the eye closure measures (especially with EYEMEAS and PERCLOS) for each task exceed 0.40 (r=0.40, p<0.001 for the 3-minute data; r=0.40, p<0.009 for the 6-minute data), and which therefore show potential as predictors of drowsiness. According to these analyses, the variables HRTRTV, LANEX, LANDEVV, LANDEVSQ, LANDEV4, LATPOSMS, and HIPASSYAW may be predictive of impairment for the auditory task; YAWVAR, HIPASSYAW, STVELV, STEXEED, LGREV, and MDREV show potential for the tactual task; and SEATMOV, LANEX, LANDEVV, LANDEVSQ, LANDEV4, YAWVAR, LATPOSMS, HIPASSYAW, STEXEED, STVELV, LGREV, and MDREV may be used as predictors of drowsiness for the visual task. Additionally, these analyses suggest that the 6-minute data produce slightly larger correlations which may provide slightly greater predictive power.

TABLE 16

Summary of Variables Which Show Potential as Predictors of Drowsiness for Each Task.

Auditory	Tactual	Visual
HRTRTV	YAWVAR	HRTRTM
LANEX	HIPASSYAW	SEATMOV
LANDEVV	STVELV	LANEX
LANDEVSQ	STEXEED	LANDEVV
LANDEV4	LGREV	LANDEVSQ
LATPOSMS	MDREV	LANDEV4
HIPASSYAW		YAWVAR
		LATPOSMS
		HIPASSYAW
		STEXEED
		STVELV
		LGREV
		MDREV

Three-minute correlations between subsidiary task measures and all driving measures for the auditory task. The next step in the correlation analysis was to correlate the subsidiary task measures with all driving measures to determine if they correlated highly with both eye closures and with driving measures which were found to be sensitive to drowsiness. Once again, both 3- and 6-minute correlations are presented for each subsidiary task. The first of these correlation matrices is presented in Table 17. To facilitate interpretation, only those correlations greater than r=0.20 are included. As before, however, those correlations with magnitudes greater than 0.40 are underlined. The correlations in Table 17 indicate that the response time measures do not correlate highly with either the eye or other driving measures. However, the variable NUMHIT does correlate highly with both EYEMEAS and PERCLOS (r=-0.419 and r=-0.471, respectively) and very highly with LANEX, LANDEVV, LANDEVSQ, LANDEV4, YAWVAR, LATPOSMS, and HIPASSYAW (r>0.50).

Six-minute correlations between subsidiary task measures and all driving measures for the auditory task. Table 18 shows that, for the 6-minute data, correlations are generally higher than for the 3-minute data. NUMHIT still shows high correlations with some of the lane and yaw

TABLE 17

Three-Minute Correlations Between Subsidiary Task Measures and Eye and Driving Measures for Auditory Task.

	RTLEFT	RTRIGHT	RTBOTH	RTMEAN	NUMHIT
EYEMEAN	--	--	--	--	-0.373
EYEMEAS	--	--	--	--	-0.419
PERCLOS	--	--	--	--	-0.471
HRTRTM	--	--	--	--	-0.251
HRTRTV	--	--	-0.260	--	--
SEATMOV	--	0.279	--	--	--
LANEX	--	--	--	--	-0.604
LANDEVM	--	--	--	--	--
LANDEVV	--	--	--	--	-0.629
LANDEVSQ	--	--	--	--	-0.619
LANDEV4	--	--	--	--	-0.611
YAWMEAN	--	--	--	--	0.242
YAWVAR	--	--	--	--	-0.545
LATPOSMS	--	--	--	--	-0.645
HIPASSYAW	--	--	--	--	-0.568
STEXEED	0.304	0.335	--	0.271	--
STVELM	--	--	--	--	--
STVELV	0.326	0.330	--	0.322	--
LGREV	0.297	--	--	0.244	--
MDREV	0.343	0.227	0.215	0.334	--

Correlations greater than 0.181 are significant with $p \leq 0.05$.

TABLE 18

Six-Minute Correlations Between Subsidiary Task Measures and Eye and Driving Measures for Auditory Task.

	RTLEFT	RTRIGHT	RTBOTH	RTMEAN	NUMHIT
EYEMEAN	--	--	-0.247	--	-0.416
EYEMEAS	--	--	--	--	-0.471
PERCLOS	--	--	--	--	-0.550
HRTRTM	--	--	-0.340	-0.228	0.277
HRTRTV	--	--	--	--	--
SEATMOV	--	0.325	--	--	--
LANEX	--	--	--	--	-0.667
LANDEVM	-0.226	--	--	--	--
LANDEVV	0.216	--	--	--	-0.669
LANDEVSQ	--	--	--	--	-0.655
LANDEV4	--	--	--	--	-0.632
YAWMEAN	--	--	--	--	--
YAWVAR	0.278	--	--	--	-0.588
LATPOSMS	0.231	--	--	--	-0.684
HIPASSYAW	0.315	--	--	--	-0.623
STEXEED	0.426	0.384	--	0.341	--
STVELM	0.245	--	--	0.215	--
STVELV	0.472	0.447	0.223	0.436	--
LGREV	0.410	0.234	--	0.316	-0.238
MDREV	0.482	0.364	0.303	0.464	--

Correlations greater than 0.258 are significant with $p \leq 0.05$.

variables which were shown to have potential as drowsiness predictors from the previous analysis and which are listed in Table 16; however, some of the response time variables also show correlations in excess of *r*=0.40 with other performance measures. For instance, RTLEFT correlates highly with STEXEED, STVELV, LGREV, and MDREV. Additionally, both RTRIGHT and RTMEAN correlate highly with STVELV. RTMEAN also shows a high correlation with MDREV. In summary, however, only NUMHIT correlates highly with those performance variables which were shown to have potential as predictors of drowsiness in the previously reported correlation analyses; hence, the variable NUMHIT may itself be used to predict drowsiness.

Three-minute correlations between subsidiary task measures and all driving measures for the tactual task. As can be seen in Table 19, variables RTLEFT, RTRIGHT, RTBOTH, and RTMEAN show correlations with HRTRTM in excess of *r*=-0.40.

Six-minute correlations between subsidiary task measures and all driving measures for the tactual task. As before, a comparison of Table 19 with the 6-minute correlations in Table 20 indicates similar patterns of correlations, with the 6-minute correlations being slightly larger than the 3-minute correlations. Unfortunately,

TABLE 19

Three-Minute Correlations Between Subsidiary Task Measures and Eye and Driving Measures for Tactual Task.

	RTLEFT	RTRIGHT	RTBOTH	RTMEAN	NUMHIT
EYEMEAN	--	--	0.218	--	--
EYEMEAS	--	-0.225	0.216	0.213	-0.200
PERCLOS	0.200	-0.296	0.214	0.264	--
HRTRTM	-0.489	-0.479	-0.563	-0.577	0.200
HRTRTV	--	--	-0.268	-0.217	--
SEATMOV	--	--	--	--	--
LANEX	--	--	--	--	--
LANDEVM	-0.202	--	--	--	--
LANDEVV	--	--	--	--	--
LANDEVSQ	--	--	--	--	--
LANDEV4	--	--	--	--	--
YAWMEAN	--	--	--	--	--
YAWVAR	--	--	--	--	--
LATPOSMS	--	--	--	--	--
HIPASSYAW	--	--	--	--	--
STEXEED	--	--	--	--	--
STVELM	0.349	0.282	0.250	0.328	--
STVELV	0.201	0.287	--	--	--
LGREV	0.246	0.336	--	0.235	--
MDREV	--	0.239	--	--	--

Correlations greater than 0.181 are significant with $p \leq 0.05$.

TABLE 20

Six-Minute Correlations Between Subsidiary Task Measures and Eye and Driving Measures for Tactual Task.

	RTLEFT	RTRIGHT	RTBOTH	RTMEAN	NUMHIT
EYEMEAN	--	0.224	0.242	0.226	-0.283
EYEMEAS	--	0.258	0.242	0.250	-0.270
PERCLOS	0.248	0.345	0.238	0.299	--
HRTRTM	-0.554	-0.524	-0.624	-0.620	0.219
HRTRTV	--	--	-0.306	-0.234	--
SEATMOV	--	--	--	--	--
LANEX	--	--	--	--	--
LANDEVM	0.200	--	--	--	--
LANDEVV	--	--	--	--	--
LANDEVSQ	0.209	--	--	--	--
LANDEV4	--	--	--	--	--
YAWMEAN	--	--	--	--	0.237
YAWVAR	--	--	--	--	--
LATPOSMS	--	--	--	--	--
HIPASSYAW	--	--	--	--	--
STEXEED	--	--	--	--	--
STVELM	0.300	0.385	0.332	0.368	--
STVELV	0.238	0.316	--	0.220	--
LGREV	0.284	0.385	--	0.292	--
MDREV	--	0.325	--	0.227	--

Correlations greater than 0.258 are significant with $p \leq 0.05$.

however, none of the subsidiary task variables in either the 3- or 6-minute analyses correlates highly with the eye measures or with other driving variables which were shown to correlate highly with eye measures for the tactual task. Hence, the tactual task performance measures show little promise as being predictive of drowsiness-impaired driving.

Three-minute correlations between subsidiary task measures and all driving measures for the visual task. Table 21 shows that, while none of the visual subsidiary task measures correlates highly with the eye closure measures, some do correlate highly with other driving variables which were shown in previous analyses to be sensitive to eye closures. For example, RTLEFT, RTRIGHT, RTMEAN, and NUMHIT all correlate highly with LANEX. RTRIGHT, RTMEAN, and NUMHIT also correlate highly with LANDEVV, LANDEVSQ, YAWVAR, LATPOSMS, and HIPASSYAW. Additionally, RTMEAN correlates highly with both LGREV and MDREV, while RTLEFT correlates highly with LANEX.

Six-minute correlations between subsidiary task measures and all driving measures for the visual task. The 6-minute data in Table 22 exhibit correlations between RTLEFT, RTMEAN, and the three eye closure measures in excess of *r*=0.40. Additionally, all five subsidiary task measures show high correlations with LANEX, LANDEVV, LANDEVSQ,

TABLE 21

Three-Minute Correlations Between Subsidiary Task Measures and Eye and Driving Measures for Visual Task.

	RTLEFT	RTRIGHT	RTBOTH	RTMEAN	NUMHIT
EYEMEAN	0.352	0.284	0.325	0.389	-0.288
EYEMEAS	0.323	0.260	0.287	0.352	-0.296
PERCLOS	0.369	0.266	0.320	0.385	-0.259
HRTRTM	-0.328	-0.288	-0.318	-0.379	0.218
HRTRTV	--	--	--	--	--
SEATMOV	--	0.264	--	0.206	--
LANEX	0.445	0.567	0.371	0.565	-0.587
LANDEVM	--	0.267	--	0.236	--
LANDEVV	0.322	0.422	0.280	0.419	-0.547
LANDEVSQ	0.376	0.522	0.351	0.512	-0.584
LANDEV4	0.215	0.347	--	0.310	-0.535
YAWMEAN	--	--	--	--	0.294
YAWVAR	0.335	0.418	0.292	0.427	-0.565
LATPOSMS	0.321	0.410	0.283	0.414	-0.560
HIPASSYAW	0.367	0.453	0.309	0.461	-0.585
STEXEED	0.273	0.280	0.254	0.328	-0.296
STVELM	--	--	--	--	0.259
STVELV	0.323	0.320	0.331	0.396	--
LGREV	0.382	0.321	0.335	0.420	--
MDREV	0.365	0.331	0.397	0.444	--

Correlations greater than 0.181 are significant with $p \leq 0.05$.

TABLE 22

Six-Minute Correlations Between Subsidiary Task Measures and Eye and Driving Measures for Visual Task.

	RTLEFT	RTRIGHT	RTBOTH	RTMEAN	NUMHIT
EYEMEAN	0.502	0.327	0.347	0.423	-0.376
EYEMEAS	0.525	0.315	0.347	0.426	-0.392
PERCLOS	0.525	0.289	0.336	0.412	-0.332
HRTRTM	-0.417	-0.343	-0.370	-0.409	0.294
HRTRTV	--	--	--	--	0.200
SEATMOV	--	0.349	0.287	0.302	--
LANEX	0.712	0.598	0.528	0.665	-0.741
LANDEVM	--	0.339	0.222	0.258	--
LANDEVV	0.566	0.403	0.404	0.495	-0.696
LANDEVSQ	0.656	0.549	0.516	0.623	-0.735
LANDEV4	0.491	0.329	0.340	0.417	-0.680
YAWMEAN	-0.387	--	--	-0.212	0.352
YAWVAR	0.633	0.465	0.457	0.561	-0.724
LATPOSMS	0.582	0.402	0.414	0.503	-0.703
HIPASSYAW	0.644	0.482	0.472	0.576	-0.739
STEXEED	0.563	0.481	0.466	0.547	-0.423
STVELM	--	--	--	--	0.316
STVELV	0.469	0.454	0.478	0.510	-0.217
LGREV	0.472	0.419	0.437	0.482	-0.218
MDREV	0.449	0.462	0.524	0.523	--

Correlations greater than 0.258 are significant with $p \le 0.05$.

YAWVAR, LATPOSMS, HIPASSYAW, and STEXEED. RTLEFT, RTRIGHT, RTBOTH, and RTMEAN correlate highly with STVELV, LGREV, and MDREV. Finally, RTLEFT and RTMEAN correlate highly with LANDEV4. Hence, it appears that, since all of the visual subsidiary task measures correlate highly with several of the driving measures which were also shown to be sensitive to drowsiness, all of these task measures may be potentially predictive of drowsiness-induced impairment.

Summary of correlation analyses of driving and subsidiary task measures. The results of the previously reported correlation analyses can be found in Table 23. The subsidiary task variables listed are limited to those whose correlations with the driving measures listed in Table 16 and/or with the eye measures exceed 0.40 (r=0.40, p=0.001 for the 3-minute data; r=0.40, p=0.009 for the 6-minute data). According to these results, the variable NUMHIT was shown to correlate highly with the eye and driving measures for both the auditory and visual tasks. Additionally, the variables RTLEFT, RTRIGHT, RTBOTH, and RTMEAN were shown to have potential as predictors of driving impairment for the visual task data.

These subsidiary task variables, along with those driving measures listed in Table 16, were next employed in a series of stepwise and linear discriminant analyses in an

TABLE 23

Summary of Subsidiary Task Performance Variables Which Correlate Most Highly with Driving Variables for Each Task.

Auditory	Tactual	Visual
NUMHIT	--	RTLEFT RTRIGHT RTBOTH RTMEAN NUMHIT

attempt to build a model, using subsidiary task and driving measures, which is predictive of driving impairment due to drowsiness.

Stepwise discriminant analyses. The next step in the subsidiary task data analysis was to conduct stepwise discriminant analyses for each task using both 3- and 6-minute interval data. The stepwise approach was chosen since it considers combinations of dependent measures and determines which of these variables contribute useful, independent information to a linear discriminant model. Hence, redundant information from one or more variables is not included in the resulting model.

To accomplish this, however, three parameters must first be specified:

1. A criterion for discrimination (i.e., in this case, some rule which can be used to classify an observation as being impaired or not impaired by drowsiness);
2. A threshold value for entry of a variable into the model being built; and
3. A threshold value for removing a variable from the model.

The selection of each of these parameters is discussed below.

First, it was necessary to determine how to classify an observation as being impaired or not impaired. Since other researchers have found that eye closures are reliable detectors of drowsiness impairment (Dingus et al., 1985; Erwin et al., 1976; Skipper et al., 1984), and since many of the lane and other driving performance measures employed in the current study demonstrated significant correlations with PERCLOS and EYEMEAS, the following impairment criterion based on these measures was used:

> If EYEMEAS > 7000 or PERCLOS > 2 then classify the observation as impaired.
>
> Else, classify the observation as not impaired.

The value of 2 for PERCLOS was chosen since, as Dingus et al. (1985) argued, a PERCLOS value greater than 2% in one 3-minute interval represents eye closures of 3-4 seconds duration. (It should be remembered that PERCLOS is the percentage of time the eyes are closed 80-100%.) A closure of this duration could certainly result in an accident.

The choice of 7000 as the criterion value for EYEMEAS was based on the finding by Dingus et al.(1985) that EYEMEAS values of 5000 or less indicated an unimpaired, rested driver while values of 9000 or more indicated impairment due to drowsiness. The value of 7000 was thus chosen in an attempt to minimize misclassification in the reported discriminant analyses such that neither misses nor false alarms were exceptionally high.

Next, the criteria for entry and removal of a variable from the model were selected. The statistical software package used for this analysis (SAS, 1982) defaults to p=0.15 for both of these parameters; however, a less stringent value of p=0.20 was chosen for use. This statistical probability level was believed to be appropriate since it provided 80% confidence that the variables in the model were useful without severely limiting the number of variables entered into the model.

Once these parameters were specified, the variables listed in Tables 16 and 23 were entered in four stepwise discriminant models--two for the auditory task (one using the 3-minute data and another using the 6-minute data) and two for the visual task. The goal of this analysis was to identify driving performance and subsidiary task measures which provide independent, useful information that is predictive of drowsiness impairment. Also, since no subsidiary task measures associated with the tactual task were found to correlate highly with eye or other performance measures, no stepwise analyses were performed using the tactual task data. The results of the stepwise analyses are summarized below.

Stepwise model for the auditory task using 3- and 6-minute data. The previously reported correlation analyses indicated that the variables NUMHIT, HRTRTV, LANEX, LANDEVV, LANDEVSQ, LANDEV4, LATPOSMS, and HIPASSYAW showed potential as predictors of impairment when the auditory task was performed. These variables were thus entered into two stepwise discriminant analyses using both the 3- and 6-minute data associated with the auditory task. Through a process of successive elimination, the following variables were designated as providing a significant amount of independent information:

<u>3-minute data</u>	<u>6-minute data</u>
HRTRTV	HRTRTV

<u>Stepwise model for the visual task using the 3- and 6-minute data</u>. Given the results of the correlation analyses, the following variables showed potential as drowsiness predictors when the visual task was performed: RTLEFT, RTRIGHT, RTBOTH, RTMEAN, NUMHIT, HRTRTM, SEATMOV, LANEX, LANDEVV, LANDEVSQ, LANDEV4, YAWVAR, LATPOSMS, HIPASSYAW, STEXEED, STVELV, LGREV, and MDREV. Since only 11 degrees of freedom (number of subjects minus one) were available for the multivariate calculations, these variables were divided into two groups and entered into two stepwise discriminant analyses using both the 3- and 6-minute data. Through a process of successive elimination, the following variables were shown to contribute a significant amount of independent information:

3-minute data		6-minute data
SEATMOV	LANEX	NUMHIT
LANDEVSQ	LGREV	LGREV
LANDEV4		

It is interesting to note at this point that the 3- and 6-minute data yielded quite different results. Furthermore, the 6-minute interval visual task model is the only one of the four which includes both subsidiary task *and* driving measures.

Linear discriminant analyses--Stage 1. Initially, five linear discriminant analyses were conducted to determine if observations could be optimally classified as being impaired or not impaired based upon the linear function derived from the information contained in each model's dependent measures. These analyses and their resulting models are described below.

Since the purpose of the previously discussed stepwise analyses was to select groups of subsidiary task and driving performance measures which provide independent information and which may be used to predict driving impairment, the first discriminant analysis was conducted using the 6-minute visual data and the variables NUMHIT and LGREV. For this analysis, the same criterion for classification was used as that employed in the stepwise discriminant analyses, namely:

If EYEMEAS > 7000 or PERCLOS > 2 then classify the observation as impaired.

Else, classify the observation as not impaired.

The discriminant analysis uses this type of classification rule to determine which observations are actually impaired or not impaired. It also uses a derived linear function to predict, based on the values of the predictor variables, whether an observation is impaired or not impaired. The results of the discriminant analysis are presented in a confusion matrix such as the one below which shows actual versus predicted group membership.

		Predicted	
		Impaired	Not impaired
Actual	Impaired	HIT (correct)	MISS (incorrect)
	Not Impaired	FALSE ALARM (incorrect)	CORRECT REJECTION (correct)

The effectiveness of the model can thus be judged by the number of misclassifications which result. One index of the success of the discriminant function is the Apparent Error Rate (APER). The APER is calculated by dividing the sum of the misclassifications by the total number of observations. In general, an APER of 20% or less is considered to provide good discrimination. However, when the number of observations in the two classified groups are greatly disproportionate (as is the case with these data), the APER tends to provide an optimistically low estimate of the actual error rate. In instances such as this, the percentages of the misses and false alarms provide a better estimate of model success (Johnson and Wichern, 1982).

The results of the first discriminant analysis, shown in Table 24, reveal 24 misses but no false alarms, yielding an APER=40.0%. Additionally, the standardized canonical coefficients indicate that LGREV provided only a slightly larger portion of separation information than did NUMHIT. These findings are somewhat surprising since, based upon the results of the stepwise discriminant analysis, one would expect that these variables would provide a significant amount of independent information which would lead to good discrimination.

Given the disappointing result associated with the previously reported discriminant analysis, the decision was made to conduct four other discriminant analyses using a slightly different criterion for impairment classification and a different set of predictor variables. For these analyses, the criterion was:

> If EYEMEAS > 7000 then classify the observation as impaired.
>
> Else, classify the observation as not impaired.

TABLE 24

Discriminant Analysis Results Using the Visual Task Model Suggested by the Stepwise Analyses.

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	32	24 (46.67%)	56
	Not Impaired	0 (0.00%)	4	4
		32	28	60

APER = 40.0%

Model Variables:	Standardized Canonical Coefficients:
NUMHIT	0.6060
LGREV	-0.7035

Criterion: If EYEMEAS > 7000 or PERCLOS > 2 then classify the observation as impaired.

Else, classify the observation as not impaired.

This simpler classification rule was used because of the findings by previous investigators (Dingus et al., 1985; Skipper, et al., 1984), which are also consistent with the current reported findings, that several driving performance measures correlate highly with EYEMEAS. These findings suggest that EYEMEAS is a stable and reliable index of drowsiness.

Additionally, based upon the results of both the stepwise and correlation analyses, two driving performance measures (LGREV and HRTRTV) and one subsidiary task measure (NUMHIT) were chosen for inclusion in the model. The variable LGREV was chosen due to its high correlation with EYEMEAS ($\underline{r}$=0.707 for the 6-minute visual task data) and due to the stepwise results for the visual task which indicate that this variable contributes a significant amount of independent information in a discriminant model. The variable HRTRTV was chosen based upon the auditory task's stepwise results. Also, it was thought that the inclusion of a physiological variable along with a performance variable may provide important, nonredundant information which may aid in the detection of impairment. Finally, the subsidiary task measure NUMHIT was chosen since it was the only subsidiary task measure which was found by the stepwise analyses to provide independent information.

The 6-minute data for the auditory and visual tasks were then entered into four discriminant analyses using two different combinations of the three predictor variables. More specifically, two discriminant models were derived for the auditory task data using the variables NUMHIT and HRTRTV in one analysis and the variables NUMHIT and LGREV in another. The same was done using the visual task data. Once again, the tactual task data were not used since the subsidiary task measures associated with this task did not correlate highly with the eye or other driving measures. Also, only the 6-minute data were used since these data were shown to result in higher correlations among measures. The results of these four discriminant analyses are reported next.

The first of these four discriminant analyses was conducted using the auditory task data and the variables NUMHIT and HRTRTV. The results of this analysis, presented in Table 25, reveal only 5 misclassified observations, thus yielding an APER of 8.3%. Additionally, the standardized canonical coefficients indicate that HRTRTV provided a larger portion of separation information than did NUMHIT. While this model appears to provide a reliable function for predicting drowsiness, a closer look at the misclassified observations indicates that 40.00% of the unimpaired and

TABLE 25

Discriminant Analysis Results Using the Six-Minute Auditory Task Data and the Variables NUMHIT and HRTRTV.

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	52	3 (5.45%)	55
	Not Impaired	2 (40.00%)	3	5
		54	6	60

APER = 8.3%

Model Variables:	Standardized Canonical Coefficients:
NUMHIT	0.0908
HRTRTV	1.1372

Criterion: If EYEMEAS > 7000 then classify the observation as impaired.

Else, classify the observation as not impaired.

5.45% of the impaired observations were misclassified. This may thus represent an undesirably high false alarm rate.

The auditory task data were employed in a discriminant analysis once again along with the second set of predictors, NUMHIT and LGREV. The results for this discriminant analysis are summarized in Table 26. This model indicates that 27 of the 60 observations were misclassified; hence, the APER=45.0%. Also, LGREV provided only a slightly larger portion of separation information than did NUMHIT. This model, thus, is a poor predictor of drowsiness.

A discriminant analysis was next conducted using the visual task data and the predictors NUMHIT and HRTRTV. The results of this model, presented in Table 27, indicate only 2 misclassifications and a resulting APER of 3.3%. Furthermore, this model yielded a 1.82% miss rate and a 20.00% false alarm rate. These rates are much more acceptable than those associated with the corresponding auditory task model in Table 25. Once again, the standardized canonical coefficients associated with the predictor variables show that HRTRTV provided slightly greater separation information than NUMHIT. Hence, these variables allow for both the reliable prediction of drowsiness and more acceptable hit and false alarm rates when the visual task, as opposed to the auditory task, is performed.

TABLE 26

Discriminant Analysis Results Using the Six-Minute Auditory Task Data and the Variables NUMHIT and LGREV.

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	29	26 (47.27%)	55
	Not Impaired	1 (20.00%)	4	5
		30	30	60

APER = 45.0%

Model Variables:	Standardized Canonical Coefficients:
NUMHIT	0.7393
LGREV	-0.8226

Criterion: If EYEMEAS > 7000 then classify the observation as impaired.

Else, classify the observation as not impaired.

TABLE 27

Discriminant Analysis Results Using the Six-Minute Visual Task Data and the Variables NUMHIT and HRTRTV.

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	54	1 (1.82%)	55
	Not Impaired	1 (20.00%)	4	5
		55	5	60

APER = 3.3%

Model Variables:	Standardized Canonical Coefficients:
NUMHIT	0.0829
HRTRTV	1.2523

Criterion: If EYEMEAS > 7000 then classify the observation as impaired.

Else, classify the observation as not impaired.

Finally, a discriminant analysis was performed using the visual data and the predictors NUMHIT and LGREV. The results of this analysis, presented in Table 28, show that 25 observations were misclassified (APER=41.6%) and that LGREV contributed greater separation information than NUMHIT.

It is interesting to note that the discriminant model summarized in Table 24 differs from the one presented in Table 28 only in regard to the classification criterion used. A comparison of these two models thus indicates that, when the simplified criterion was used, only 1 more misclassification resulted. Hence, it may be concluded that the use of the simplified classification criterion produced results similar to those obtained when the more complex criterion was used.

Unresolved issues. The results of the previously reported discriminant analyses raised two important issues which should be addressed:

1. The disparity in the number of impaired vs. unimpaired observations.
2. The poor results of the model suggested by the stepwise discriminant analyses.

These two issues are discussed in turn below along with their implications for further exploration of the data through the use of discriminant analyses.

TABLE 28

Discriminant Analysis Results Using the Six-Minute Visual Task Data and the Variables NUMHIT and LGREV.

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	30	25 (45.45%)	55
	Not Impaired	0 (0.00%)	5	5
		30	30	60

APER = 41.6%

Model Variables:	Standardized Canonical Coefficients:
NUMHIT	-0.4603
LGREV	0.8225

Criterion: If EYEMEAS > 7000 then classify the observation as impaired.

Else, classify the observation as not impaired.

First, given the criterion for impairment used in these discriminant analyses (i.e., if EYEMEAS > 7000 then the observation is impaired), the results indicate that very few of the observations employed in the analyses were unimpaired. Hence, the resulting models demonstrate a large disparity between the number of observations in the impaired vs. not impaired groups. This disparity is statistically undesirable since the reported APERs associated with these models are optimistically low as a result. Additionally, this result suggests that the subjects in this experiment may have been more impaired than those in previous studies (e.g., Dingus et al., 1985); therefore, a comparison of the eye closure data from the current study with that reported by Dingus et al. may be desirable to further refine the criterion for impairment.

Since the experimental design employed by Dingus et al. (1985) included both rested and sleep deprived driving conditions, a comparison of mean eye closures per 30 minute driving segment for each of these two conditions was possible. These results, shown in Figure 5, indicate that rested subjects exhibited mean EYEMEAS values below 6000 while sleep deprived subjects exhibited mean EYEMEAS values above 9000. A similar plot for the current study demonstrates that all mean EYEMEAS values are in excess of

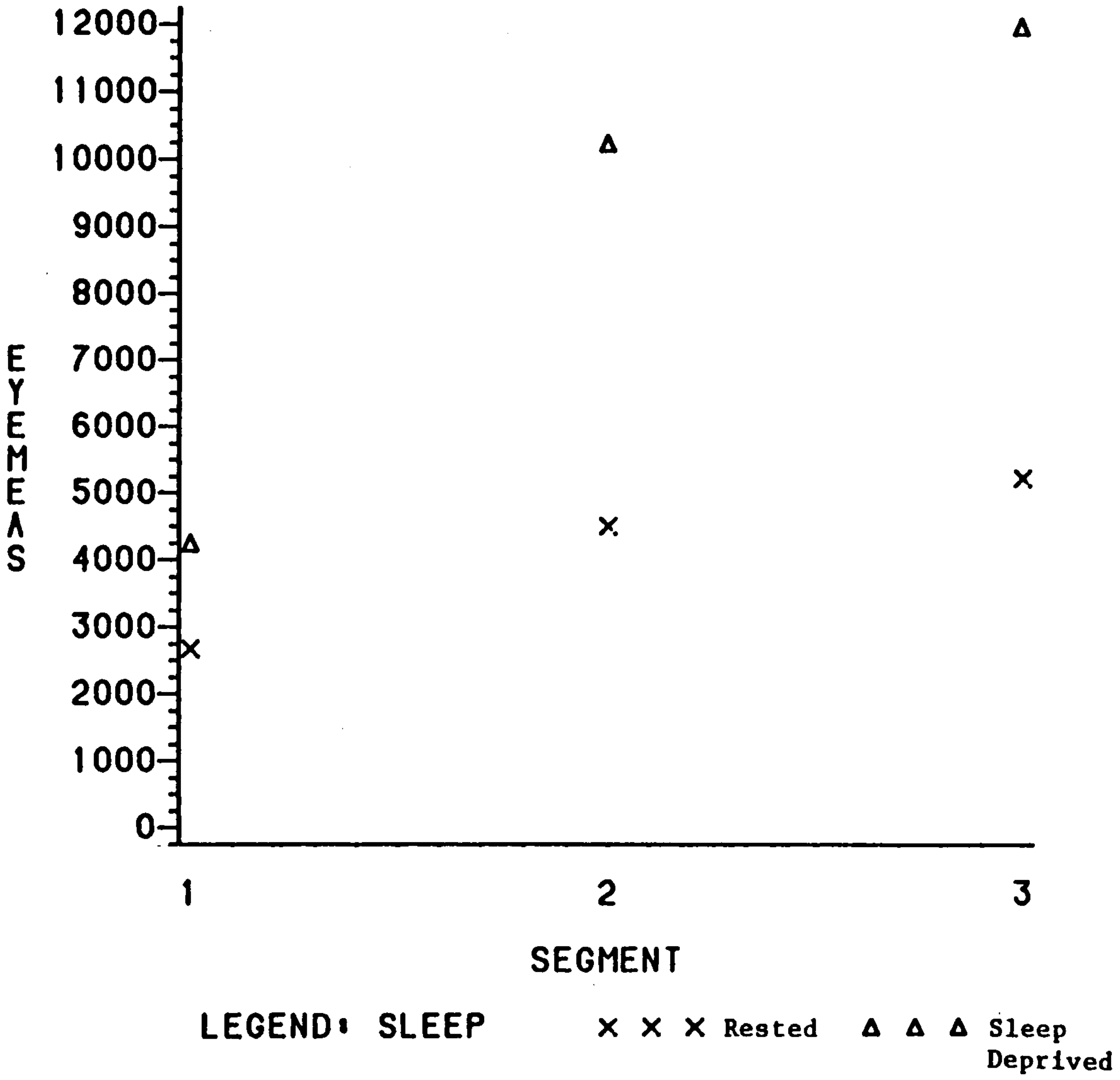

Figure 5: EYEMEAS means vs. driving segment plotted by sleep level (taken from Dingus et al., 1985).

10 000 (see Figure 6). In fact, the smallest EYEMEAS mean in the current study (12 887 for segment 1) is larger than the largest mean reported in the Dingus et al. study (12 000 for segment 3). Hence, these results provide fairly convincing evidence that the subjects in the current study were more impaired than those in the Dingus et al. study.

Additionally, the plot in Figure 6 suggests that, during segments 1 and 2 impairment continued to increase but then began to level off during segments 3 and 4. These observations were supported by the results of a Newman-Keuls multiple comparison of the EYEMEAS means (see Table 29). According to these results, the means for segments 1 and 2 are significantly different from each other and from the means for segments 3 and 4. The means for segments 3 and 4 do not differ significantly, however.

Given these results, a decision was made to modify the criterion for impairment and to conduct another series of discriminant analyses employing two new criteria. These new criteria were:

1. If EYEMEAS > 9000 then classify the observation as impaired.

   Else, classify the observation as not impaired.

2. If EYEMEAS > 17 500 then classify the observation as severely impaired.

   Else, classify the observation as impaired.

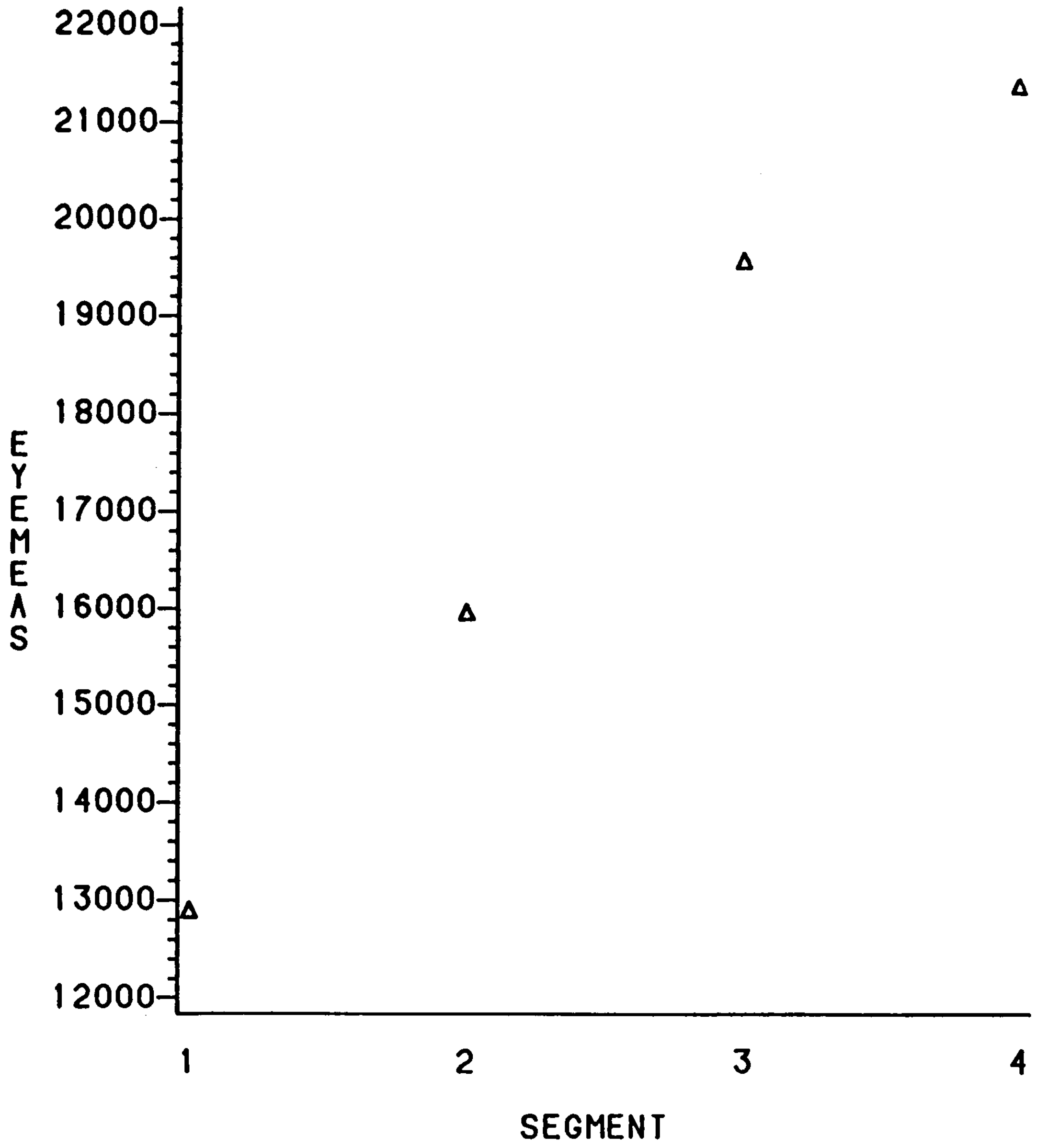

Figure 6: EYEMEAS means vs. driving segment.

TABLE 29

Newman-Keuls Results for Effect of Driving Segment on EYEMEAS.*

1	2	3	4
12 887 A	15 958 B	19 575 C	21 386 C

* Mean values designated by different letters are significantly different with $p < 0.05$.

The first criterion provides for a higher cutoff for EYEMEAS values which was deemed to be more consistent with the EYEMEAS data of the current study; however, this selected value for EYEMEAS is also consistent with the finding by Dingus et al. (1985) that EYEMEAS values for impaired observations were greater than 9000.

The second criterion was selected so that a comparison of the detectability of severely impaired vs. impaired observations could be made. The value of 17 500 for EYEMEAS was chosen since it represents the point approximately midway between the EYEMEAS means for segments 2 and 3 (15 958 and 19 575, respectively) where eye closures appear to be leveling off.

The second issue raised by the previously reported discriminant analyses involves the poor results obtained from the discriminant model which employed the combination of variables chosen by the stepwise analyses. As the reader will recall, the first discriminant model reported included the variables NUMHIT and LGREV and employed the 6-minute visual data (see Table 24). This model was the only model which included only those combinations of variables which were designated by the stepwise procedure as providing independent information. Yet, the ability of this model to discriminate impaired vs. not impaired observations was much poorer than that of two other models which included combinations of variables which were chosen by the experimenters (Tables 25 and 27). Hence, it would seem logical that more extensive exploration of the data is warranted to determine if other combinations of variables, combinations not suggested by the stepwise analyses, can be used to predict driver impairment.

Linear discriminant analyses--Stage 2. Based upon the two issues raised by the first series of discriminant analyses, the decision was made to conduct another series of analyses which would address these issues. To this end, 4 groups of 17 discriminant analyses were conducted. Each of these 4 groups employed the same 17 combinations of subsidiary task and performance measures. More specifically, the variable NUMHIT was paired with each of the following performance variables, resulting in 17 combinations of variables: HRTRTM, HRTRTV, SEATMOV, LANEX, LANDEVM, LANDEVV, LANDEVSQ, LANDEV4, YAWMEAN, YAWVAR, LATPOSMS, HIPASSYAW, STEXEED, STVELM, STVELV, LGREV, and MDREV. It should be noted that the variable NUMHIT was chosen since it was previously shown to correlate highly with EYEMEAS and with other performance measures which are sensitive to drowsiness.

The four groups of analyses are characterized below in terms of their criterion for impairment and the type of subsidiary task used. Once again, only the 6-minute data and the auditory and visual task data were used in these analyses for the reasons discussed previously.

Group	Criterion	Task
1	If EYEMEAS > 9000 then impaired. Else, not impaired	Auditory
2	If EYEMEAS > 9000 then impaired. Else, not impaired	Visual
3	If EYEMEAS > 17 500 then severely impaired. Else, not impaired	Auditory
4	If EYEMEAS > 17 500 then severely impaired. Else, not impaired	Visual

All of the resulting 68 discriminant analyses are not reported here. Instead, only those which show the most promise as predictors of impairment (i.e., those with APERs less than or equal to 20.0%) have been reproduced in the text which follows. The interested reader is thus referred to Appendix E for a complete summary of these results.

As the previous discussion indicates, the discriminant analyses for Groups 1 and 2 utilized the auditory and visual data, respectively, along with a revised criterion for impairment which raised the cutoff for the EYEMEAS values. The results of these analyses are discussed next. Special emphasis is given to comparisons of the auditory and visual task results, as well as to comparisons of results obtained with the old vs. the revised impairment criteria.

The 17 discriminant analyses conducted for Group 1 yielded only one promising model. This model, which can be seen in Table 30, employed the variables NUMHIT and HRTRTV and was built using the auditory task data. The APER of 6.7% is well below the "rule of thumb" value of 20.0%; however, the model also demonstrates a false alarm rate of 37.50% since three of the eight unimpaired observations were misclassified. This rate is, unfortunately, undesirably high. Finally, the standardized canonical coefficients associated with the model's variables indicate that HRTRTV contributes more separation information than does NUMHIT.

It is interesting to compare the discriminant model in Table 30 with the one summarized in Table 25 since these models differ only in regard to their criteria for impairment. The associated APERs indicate that, when the higher cutoff value for impairment is employed (as in Table 30), discrimination is somewhat improved (APER=6.7% vs. APER=8.3%). Additionally, the miss rate associated with the

TABLE 30

Auditory Discriminant Model Employing the Revised Criterion and Variables NUMHIT and HRTRTV.

		Predicted: Impaired	Predicted: Not Impaired	
Actual	Impaired	51	1 (1.92)	52
	Not Impaired	3 (37.50)	5	8
		54	6	60

APER = 6.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.0748
HRTRTV	1.3103

Criterion: If EYEMEAS > 9000 then classify the observation as impaired.

Else, classify the observation as not impaired.

revised model is more satisfactory (1.92% vs. 5.45%). Hence, it may be concluded that, in the case of the auditory task, revising the criterion for impairment resulted in an improved model.

The 17 discriminant analyses associated with Group 2 also indicated the visual data model employing the variables NUMHIT and HRTRTV as the sole promising predictive model. This model, summarized in Table 31, yielded an APER=13.3%. Both the miss and false alarm rates are fairly high, however (8.16% and 36.36%, respectively). The standardized canonical coefficients reveal that HRTRTV contributed more separation information than did NUMHIT.

A comparison of the visual model in Table 31 with the previously discussed auditory model in Table 30 suggests that, when the revised criterion for impairment is used, the auditory task allows for somewhat better prediction of impairment than does the visual task (APER=6.7% for the auditory data and APER=13.7% for the visual data). However, by comparing the visual model in Table 31 with its corresponding visual model in Table 27, one might conclude that discriminability suffers when the revised criterion is employed (APER=3.3% for the visual model with the old criterion vs. APER=13.3% for the visual model with the revised criterion). This conclusion is also supported by a

TABLE 31

Visual Discriminant Model Employing the Revised Criterion and Variables NUMHIT and HRTRTV.

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	45	4 (8.16)	49
	Not Impaired	4 (36.36)	7	11
		49	11	60

APER = 13.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.3853
HRTRTV	1.1152

Criterion: If EYEMEAS > 9000 then classify the observation as impaired.

Else, classify the observation as not impaired.

comparison of the miss and false alarm rates associated with the two models (miss rates=1.82% and 8.16%; false alarm rates=20.20% and 36.36% for the models employing the old and revised criteria, respectively).

The analyses of Groups 3 and 4 employed, as the reader will recall, yet another criterion for impairment. This criterion classified the observations as either <u>severely</u> impaired or impaired. These analyses were undertaken to determine if discriminability improves as impairment increases. The results of these analyses are described below.

The discriminant analyses conducted for Group 3 yielded two promising models for the auditory task data. The model employing the variables NUMHIT and SEATMOV is summarized in Table 32. According to these results, the APER=18.3%; however, both the miss and false alarm rates are fairly high (26.92% and 11.76%, respectively). Thus, while this model's APER suggests that it provides reasonably good prediction of severe impairment, the miss and false alarm rates are unsatisfactory. Finally, the variable NUMHIT contributed more separation information than did SEATMOV.

The second promising model, shown in Table 33, employed the variables NUMHIT and STVELV and yielded an APER of 20.0%. While the false alarm rate for this model is fairly

TABLE 32

Auditory Discriminant Model Employing the Criterion for Severe Impairment and Variables NUMHIT and SEATMOV.

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	19	7 (26.92)	26
	Impaired	4 (11.76)	30	34
		23	37	60

APER = 18.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.1516
SEATMOV	-0.2591

Criterion: If EYEMEAS > 17 500 then classify the observation as severely impaired.

Else, classify the observation as impaired.

TABLE 33

Auditory Discriminant Model Employing the Criterion for Severe Impairment and Variables NUMHIT and STVELV.

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	17	9 (34.62)	26
	Impaired	3 (8.82)	31	34
		20	40	60

APER = 20.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0308
STVELV	-0.4903

Criterion: If EYEMEAS > 17 500 then classify the observation as severely impaired.

Else, classify the observation as impaired.

low (8.82%), the miss rate is unacceptably high (34.62%). A comparison of the standardized canonical coefficients associated with the two variables reveals that NUMHIT provides more separation information than does STVELV.

The discriminant analyses associated with the fourth and final group resulted in five predictive models for the visual task data. The results of the first of these models can be found in Table 34. According to these results, this model, which includes the variables NUMHIT and SEATMOV, indicates an APER of 18.3% with a miss rate of 19.23% and a false alarm rate of 17.65%. Additionally, the variable NUMHIT is shown to provide only slightly more separation information than does SEATMOV.

A comparison of this model with the one summarized in Table 32 reveals that the variables NUMHIT and SEATMOV allow for the comparable discriminability of impaired vs. not impaired drivers and severely impaired vs. impaired drivers in terms of APERs. However, the auditory model has a higher miss rate (26.92% compared to 19.23%) and a lower false alarm rate (11.76% compared to 17.65%) than the visual model.

The second model, summarized in Table 35 was built using the variables NUMHIT and YAWVAR and exhibits an APER of 18.3%. The false alarm rate is very low (2.94%);

TABLE 34

Visual Discriminant Model Employing the Criterion for Severe Impairment and Variables NUMHIT and SEATMOV.

		Predicted Severely Impaired	 Impaired	
Actual	Severely Impaired	21	5 (19.23)	26
	Impaired	6 (17.65)	28	34
		27	33	60

APER = 18.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.9829
SEATMOV	-0.7012

Criterion: If EYEMEAS > 17 500 then classify the observation as severely impaired.

Else, classify the observation as impaired.

TABLE 35

Visual Discriminant Model Employing the Criterion for Severe Impairment and Variables NUMHIT and YAWVAR.

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	16	10 (38.46)	26
	Impaired	1 (2.94)	33	34
		17	43	60

APER = 18.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.1629
YAWVAR	-1.1225

Criterion: If EYEMEAS > 17 500 then classify the observation as severely impaired.

Else, classify the observation as impaired.

however, the miss rate is high (38.46%). Thus, this model may not allow for an acceptable level of prediction of severely impaired drivers.

Another model using the variables NUMHIT and HIPASSYAW also yielded results which are worthy of discussion. Table 36 shows that this model demonstrates both a low APER (15.0%) and false alarm rate (2.94%). The miss rate of 30.77% is fairly high, however. Finally, the standardized canonical coefficients reveal that HIPASSYAW provides more separation information than does NUMHIT.

The discriminant model reported in Table 37 includes the variables NUMHIT and STVELV and is identical to the corresponding auditory data model presented in Table 33 in terms of its associated APER (20.0%), miss rate (34.62%), and false alarm rate (8.82%). The sole difference between these two models is the weightings of the variables as indicated by the standardized canonical coefficients. The visual model shows that NUMHIT and STVELV contribute virtually equal amounts of separation information. Thus, the variables NUMHIT and STVELV allow for comparable discriminability when either the auditory or visual task is employed.

The final model associated with the Group 4 discriminant analyses is summarized in Table 38. This

TABLE 36

Visual Discriminant Model Employing the Criterion for Severe Impairment and Variables NUMHIT and HIPASSYAW.

		Predicted Severely Impaired	Predicted Impaired	
Actual	Severely Impaired	18	8 (30.77)	26
	Impaired	1 (2.94)	33	34
		19	41	60

APER = 15.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.0549
HIPASSYAW	-1.2295

Criterion: If EYEMEAS > 17 500 then classify the observation as severely impaired.

Else, classify the observation as impaired.

TABLE 37

Visual Discriminant Model Employing the Criterion for Severe Impairment and Variables NUMHIT and STVELV.

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	17	9 (34.62)	26
	Impaired	3 (8.82)	31	34
		20	40	60

APER = 20.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.8181
STVELV	-0.7618

Criterion: If EYEMEAS > 17 500 then classify the observation as severely impaired.

Else, classify the observation as impaired.

TABLE 38

Visual Discriminant Model Employing the Criterion for Severe Impairment and Variables NUMHIT and LGREV.

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	18	8 (30.77)	26
	Impaired	3 (8.82)	31	34
		21	39	60

APER = 18.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.7476
LGREV	-0.9278

Criterion: If EYEMEAS > 17 500 then classify the observation as severely impaired.

Else, classify the obsevation as impaired.

model, which employed the variables NUMHIT and LGREV, demonstrates an APER of 18.3% and a fairly low false alarm rate (8.82%). The miss rate, however, is fairly high (30.77%). The standardized canonical coefficients associated with the variables show that LGREV contributes slightly more separation information than does NUMHIT.

Summary of the discriminant analyses. Initially, several stepwise discriminant analyses were conducted using the driving and subsidiary task variables which were indicated by the correlation analyses to have potential for predicting driving impairment due to drowsiness. The results of these analyses yielded only one set of driving and subsidiary task measures which provided significant, independent information. These variables, NUMHIT and LGREV, were associated with the 6-minute visual data.

Five different linear discriminant analyses were conducted next using the criterion for impairment: If EYEMEAS > 7000 then the observation is impaired; else the observation is not impaired. Of the five resulting discriminant models, two exhibited promise as good predictors of drowsiness. These two models, one using the 6-minute auditory task data and the other using the 6-minute visual task data, included the variables NUMHIT and HRTRTV and yielded APER's of 8.3% and 3.3%, respectively.

Surprisingly, however, the models including the variables which were shown by the stepwise results to provide independent information (NUMHIT and LGREV) yielded poor results. Given this finding, and the finding concerning the large disparity in number of observations in the impaired vs. nonimpaired groups, the decision was made to revise the criterion for impairment and to conduct another series of discriminant analyses using the 17 combinations of variables produced when NUMHIT was paired with 17 of the driving measures. These results indicated that, when a revised criterion for impairment was employed (If EYEMEAS > 9000 then the observation is impaired; else the observation is not impaired), once again the auditory and visual task data models including the variables NUMHIT and HRTRTV were superior (APER=6.7% for the auditory task model and 13.3% for the visual task model).

A second criterion which allowed for a discrimination between severely impaired vs. impaired observations was also employed (If EYEMEAS > 17 500 then the observation is severely impaired; else the observation is not impaired) along with the 17 pairs of predictor variables described earlier. The results of these 34 analyses revealed that, for the auditory task data, the models including NUMHIT and SEATMOV (APER=18.3%) and NUMHIT and STVELV (APER=20.0%)

showed promise as predictors of severe impairment. In regard to the visual task data, the following five models yielded promising results: NUMHIT and SEATMOV (APER=18.3%); NUMHIT and YAWVAR (APER=18.3%); NUMHIT and HIPASSYAW (APER=15.0%); NUMHIT and STVELV (APER=20.0%); NUMHIT and LGREV (APER=18.3%).

These results taken together indicate that, as hypothesized, it is possible to employ groups of subsidiary task and driving measures to provide a reliable means by which driver impairment due to drowsiness can be predicted. Furthermore, it appears that the discrimination between impaired vs. not impaired drivers is superior to that of severely impaired vs. impaired drivers.

Subjective Alertness Ratings as a Measure of Driver Drowsiness

One other analysis was undertaken to determine whether or not subjective ratings by subjects were influenced by the subsidiary task or by total time on the driving task. The answers to these questions were explored through the use of two Kruskal-Wallis tests. The Kruskal-Wallis test was chosen for use since it is the nonparametric equivalent of a one-way ANOVA and is an appropriate statistical test for use with ordinal data. The results of these analyses follow.

Effect of type of subsidiary task on subjective alertness ratings. At the end of each 30-minute driving segment subjects were asked to verbally rate, using the scale found in Appendix D, how alert they had felt during the previous driving segment. These ratings were subsequently analyzed through the use of a Kruskal-Wallis test using type of subsidiary task as the independent measure. The results indicate that these ratings were not significantly influenced by the type of task performed ($H'=7.159$, $p<0.070$). However, the relatively marginal nonsignificance of this statistical test may suggest that any significant differences which were present in the subjective data may not have been identified due to a lack of power. Perhaps an increase in the number of observations would have resulted in statistical significance.

Effect of time-on-task on subjective alertness ratings. A second Kruskal-Wallis test was conducted using driving segment as the independent measure. This measure, in this case, had five levels since a subjective rating was obtained for the first driving segment when no other driving or subsidiary task measures were taken, as well as for the four driving segments when experimental data were collected. According to these results, the independent variable, driving segment, did not significantly influence the subjective ratings of alertness ($H'=9.265$, $p<0.060$).

Given that this test statistic is marginally nonsignificant, however, the mean rank sums used in the statistical test were inspected to determine if any trends in the subjective data were present. This inspection revealed that the mean rank sums increased slightly from segment 1 to segment 5, thus suggesting that subjects tended to rate themselves as less alert with increasing time-on-task. Thus, it may be that the lack of significance associated with this analysis was due to a lack of power. With a larger number of observations, significance may have been obtained.

Summary of subjective rating analyses. Contrary to expectation, subjective ratings of alertness were not significantly affected by either type of task or time-on-task.

Post hoc analyses for effect of driver experience on driving performance. One variable which may affect driving performance is the experience level of the driver. For this reason, the decision was made to analyze the driving performance data to determine if driver experience had an effect on driving performance in the current experiment

Before performing these analyses, it was necessary to divide the subjects into two groups representing experienced and inexperienced drivers. This was accomplished by reviewing the subjects' responses to a question included in the driving habits questionnaire which asked how many miles they drove per year. Based on this information, the six subjects (three males and three females) who indicated that they drove more than 10 000 miles per year were assigned to the experienced driver group while the remaining three males and three females were assigned to the inexperienced (less than 10 000 miles per year) driver group.

To analyze the performance data for effects of driver experience, three one-way MANOVAs using the variable driving experience as the independent measure were conducted. As previously discussed, since the total number of dependent measures exceeded the number of degrees of freedom available for statistical computation, it was necessary to conduct more than one MANOVA. The dependent variables were thus divided into three groups as before (i.e., the behavioral/physiological variables, the lane/yaw variables, and the steering variables). The results of these three MANOVAs are shown in Table 39 and indicate that driver experience did not account for a significant portion of variance in driving performance ($F=4.32$, $p=0.0650$ for the behavioral/ physiological variables; $F=4.66$, $p=0.1890$ for the lane/yaw variables; $F=0.55$, $p=0.7372$ for the steering variables).

TABLE 39

MANOVA Summary for Test of Effect of Driver Experience on Driving Variables.*

Dependent Measures	dv	df_H	df_E	U	F	p
Physiological/Behavioral						
Source Experience	6	6	5	0.1619	4.32	0.0650
Lane/Yaw						
Source Experience	9	9	2	0.0455	4.66	0.1890
Steering						
Source Experience	5	5	6	0.6866	0.55	0.7372

*Error term for all MANOVAs is Subject(Experience).

where: dv = number of dependent measures

df_H = degrees of freedom for treatment effect

df_E = degrees of freedom for error effect

U = Wilks' likelihood ratio statistic

$$\frac{|E|}{|E + H|}$$, where:

$|E|$ = determinant of sum of squares and cross-products matrix for error

$|E + H|$ = determinant of the sum of the sum of squares and cross-products matrix for error, and the sum of squares and cross-products matrix for treatment

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

Subsidiary Tasks as Drowsiness Countermeasures

As was discussed in the Results Section of this report, the three subsidiary tasks were not found to be effective as drowsiness countermeasures. More specifically, no reliable decrease in eye closure measures or improvement in driving performance measures was obtained during driving segments when subjects performed the subsidiary tasks as compared with segments when no task was performed. This result appears to be somewhat counterintuitive since one would expect that performing an additional task should cognitively load the driver, or at least load the driver at the perceptual level, thus helping him or her to maintain alertness.

Additional support for this hypothesis was also provided by the subjects themselves. When questioned after the experiment, five of the 12 subjects expressed the opinion that the tasks had helped them to maintain alertness; of these five subjects, one chose the auditory task as the one which helped best to maintain alertness, one chose the visual task, and three chose the tactual task.

However, despite these claims by the subjects themselves, the objective driving data did not provide support for the hypothesis.

One possible explanation for these findings could be that the subjects were too exhausted for any type of stimulation to provide an effective alerting mechanism. The combination of the sleep deprivation and the length of the driving session (2 1/2 hours) may have resulted in subjects who were simply too exhausted to perform such a low event driving task. That this was the case is suggested by both subjective and objective comparisons of the subjects in the current and previous driving studies. For instance, a subjective comparison of subject driving and eye closure behavior during this study and the two previous driving studies performed in the Vehicle Simulation Laboratory at Virginia Tech (Dingus, et al., 1985; Skipper, et al., 1984) suggests that the subjects in the current study appeared to be more likely to actually fall asleep during the driving task than those who drove for only 1 1/2 hours in the previous two studies. As a matter of fact, during the current study, two subjects exceeded the lane boundaries of the simulated highway so badly during bouts of sleep that it was necessary to awaken the subjects and reset the simulator before continuing with the driving session. This type of incident did not occur during previous sleep deprivation studies. That the subjects were too exhausted and drowsy to perform the task was also exemplified by the one subject who was withdrawn from the study due to his inability to remain awake. Finally, a more objective comparison of the mean EYEMEAS values per driving segment for the subjects in the Dingus et al. (1985) study and the subjects in the current study (which was previously reported in detail in the Results Section) also strongly suggests that the subjects in this study were simply too drowsy for any type of countermeasure to provide an alerting effect.

Another, alternative explanation for the ineffectiveness of the subsidiary tasks as drowsiness countermeasures may be that the tasks were simply not stimulating enough. The tasks used in this study were repetitive, simple to perform, and did not require a tremendous amount of cognitive processing. Hence, the subjects may not have been sufficiently stimulated by these tasks to remain alert. It may be that more meaningful, intrinsically interesting tasks would provide a greater alerting effect for the drowsy driver. However, the possibility does exist that such tasks, while providing the driver with an effective means for counteracting drowsiness, may in fact be so distracting as to present a hazard to the driver. This potential for distraction must thus be kept in mind as the search for countermeasures continues.

Subsidiary Task Performance Measures as Predictors of Impairment

The first series of correlation analyses was encouraging in that it pointed to several driving measures which, due to their high correlations with eye measures in this and other studies (Dingus, et al., 1985; Skipper, et al., 1984), have been shown to be predictive of drowsiness-induced impairment. Additionally, the visual task measures appeared to correlate highly with other driving measures which appear to be sensitive to drowsiness.

The results of a series of linear discriminant analyses also proved to be encouraging since a single combination of driving and subsidiary task measures (NUMHIT and HRTRTV) was identified as being highly predictive of eye closures during both the auditory and visual tasks. These results are also especially promising since either of these subsidiary tasks along with a detector for the subsidiary task predictor variable could be implemented in a vehicle; however, it may be necessary to first develop an unobtrusive and reliable methodology for monitoring heart rate before this physiological variable can be utilized in a drowsiness detection model. Nonetheless, these results appear to provide support for the hypothesis that subsidiary task measures may be used in conjunction with driving measures to predict drowsiness-induced impairment.

One unexplained result of these analyses, however, relates to the finding that the models which were shown to provide superior discrimination included variables other than those designated as optimal by the stepwise discriminant analyses (i.e., NUMHIT and LGREV). Regardless of the criterion employed for classifying observations as being impaired or not impaired, the variables NUMHIT and HRTRTV were shown to provide for the superior prediction of impairment as compared with the variables NUMHIT and LGREV, or as compared with any other combination of variables.

In an attempt to explain these results, a series of 20 one-way ANOVAs was conducted which employed type of subsidiary task as the independent measure. While the complete results of these analyses are not presented here, Table 40 includes the ANOVA summary table for HRTRTV, which was the only one of the 20 driving measures shown to be significantly affected by type of subsidiary task ($F=3.17$, $p=0.0372$). Of course, it is entirely possible that this result represents a Type I error since one would expect that, given an a priori criterion for significance of $p=0.05$, one out of 20 F-tests would demonstrate significance

TABLE 40

ANOVA Summary of Type of Task Effect on HRTRTV.

Source	df	MS	*F*	*p*
Task	3	1524.66	3.17	0.0372
Task x Subj	33	481.29		

due to chance alone. However, it may well be that these results indicate that the variable HRTRTV is in fact sensitive to the subsidiary tasks and thus provides important information which can be used in conjunction with subsidiary task performance measures to predict impairment due to drowsiness. Certainly further investigation and validation of this finding is warranted before any firm conclusions can be drawn.

As for the detection of severe impairment, the results of the discriminant analyses suggest that several driving measures can be used in conjunction with NUMHIT to detect severe impairment; however, the promising models associated with these analyses do not discriminate as well as those which predict impairment vs. no impairment. This may be partially due to the fact that the number of observations in the severely impaired and impaired groups are nearly equal; whereas, the analyses employing the criterion for impairment vs. no impairment demonstrate a great disparity in the number of observations in each category. Since such a disparity tends to deflate the APER, it may be that the results associated with the prediction of impairment appear too optimistic. For this reason, these results should be interpreted with caution until they are further validated.

Subjective Alertness Ratings as a Measure of Driver Drowsiness

The analysis of the subjective ratings indicated that subjects did not rate themselves as being significantly more alert during those 30-minute driving segments when they performed the subsidiary tasks compared to the segment when no task was performed. These results are consistent with the analysis of the driving performance measures; neither the subjective ratings nor the driving measures changed significantly due to the presence or absence of the subsidiary tasks.

The subjective ratings were also not significantly affected by total time-on-task. While this finding is somewhat counterintuitive and is inconsistent with other results obtained by experimenters who employed the same rating scale (Wierwille and Muto, 1981), these results may lend further support to the theory that the subjects who participated in this study were extremely exhausted. In fact, a subjective inspection of the alertness rating data tends to indicate that subjects rated themselves as being fairly drowsy early on in the experiment. Hence, it may be concluded that these subjective data provide further indication of the extreme impairment experienced by these subjects.

Validation of Current Research Findings

The sleep deprivation/driving paradigm used for this research was similar to those used in two previous drowsy driving studies by Dingus et al. (1985) and Skipper et al. (1984). The similarity between these two studies and the current one thus allows a comparison of results.

Table 41 presents a list of those variables which have been shown in three studies at Virginia Tech to be potentially reliable predictors of drowsiness-induced driving impairment due to their high correlations with eye closures (especially EYEMEAS and PERCLOS). It should be noted that the measures listed for the current study represent those variables which were shown to correlate highly with eye closure measures, regardless of type of subsidiary task performed. Additionally, those variables (i.e., LANEX, LANDEVV, LANDEVSQ, YAWVAR, LATPOSMS, HIPASSYAW, STVELV, and LGREV) highlighted in Tables 42 and 43 which demonstrate high correlations with eye closures when no task was performed are also represented in Table 41. Only the variables YAWVAR and HIPASSYAW were found to correlate highly with the eye measures under all subsidiary task conditions (i.e., auditory, tactual, visual, and no task).

TABLE 41

Potentially Reliable Predictors of Drowsiness-Induced Impairment Identified During Three Studies at Virginia Tech.

Skipper et al. (1984)	Dingus et al. (1985)	Current Research
LANDEVV*	HRTRTM	HRTRTM
LANDEVSQ*	HRTRTV	HRTRTV
STVELM	SEATMOV	SEATMOV
	YAWVAR	LANEX
	STEXEED	LANDEVV
	STVELV	LANDEVSQ
	LGREV	LANDEV4
		YAWVAR**
		LATPOSMS
		HIPASSYAW**
		STEXEED
		STVELV
		LGREV
		MDREV

* These variable names were not the names used by Skipper et al. (1984) in their report; however, the variables named by Skipper et al. are equivalent to the variables LANDEVV and LANDEVSQ reported by Dingus et al. (1985) and by the current author.

**These variables correlated highly with eye closures for all subsidiary tasks _and_ when no task was performed.

TABLE 42

Three-Minute Correlations Between Eye and Driving Measures When No Task was Performed.

	EYEMEAN	EYEMEAS	PERCLOS
HRTRTM	-0.338	-0.355	-0.351
HRTRTV	-0.287	-0.193	0.042
SEATMOV	-0.101	-0.033	0.145
LANEX	0.248	0.323	0.506
LANDEVM	0.205	0.189	0.217
LANDEVV	0.143	0.224	0.415
LANDEVSQ	0.208	0.282	0.471
LANDEV4	0.072	0.127	0.289
YAWMEAN	0.147	0.140	0.069
YAWVAR	0.210	0.272	0.433
LATPOSMS	0.183	0.265	0.461
HIPASSYAW	0.231	0.297	0.460
STEXEED	0.088	0.116	0.188
STVELM	-0.155	-0.167	-0.175
STVELV	0.138	0.172	0.269
LGREV	0.248	0.263	0.322
MDREV	0.054	0.070	0.141

Correlations greater than 0.200 are significant with $p \leq 0.05$.

TABLE 43

Six-Minute Correlations Between Eye and Driving Measures When No Task was Performed.

	EYEMEAN	EYEMEAS	PERCLOS
HRTRTM	-0.342	-0.362	-0.372
HRTRTV	-0.320	-0.219	0.041
SEATMOV	-0.121	-0.048	0.139
LANEX	0.293	0.385	0.648
LANDEVM	0.249	0.236	0.297
LANDEVV	0.151	0.248	0.512
LANDEVSQ	0.234	0.323	0.588
LANDEV4	0.063	0.129	0.339
YAWMEAN	0.286	0.279	0.180
YAWVAR	0.261	0.342	0.588
LATPOSMS	0.194	0.293	0.560
HIPASSYAW	0.274	0.357	0.603
STEXEED	0.153	0.195	0.333
STVELM	-0.179	-0.205	-0.283
STVELV	0.205	0.252	0.419
LGREV	0.305	0.326	0.429
MDREV	0.106	0.133	0.254

Correlations greater than 0.253 are significant with $p \leq 0.05$.

Table 41 indicates that two variables, LANDEVV and LANDEVSQ, have been shown to be predictors of drowsiness by both Skipper et al. (1984) and the current research. Additionally, both the research by Dingus et al. (1985) and that currently reported identified six other variables which show potential as predictors of driving impairment due to drowsiness: HRTRTM, SEATMOV, YAWVAR, STEXEED, STVELV, and LGREV. Therefore, the results of the reported research are consistent with past experimental findings, lending validity to the conclusion that driving measures can be used to predict drowsiness.

Suggestions for Future Research

While the evaluation of the three subsidiary tasks as drowsiness countermeasures rendered disappointing results, it should be recalled that this represents only the second reported attempt to employ subsidiary tasks in such a manner. Hence, this line of research should not be totally discarded until it is more fully explored and investigated. Therefore, the current author recommends that future research be conducted along the following lines:

1. First, while the three subsidiary tasks used in the recent study were found to be ineffective as drowsiness countermeasures, one should not conclude

that all subsidiary tasks will also be ineffective. Future research should emphasize the search for other countermeasures which can capture and maintain a driver's alertness over substantial periods of time. Of course, while it is probably impossible to develop a countermeasure which is effective without being somewhat distracting, the search for countermeasures which provide optimal driver arousal and minimal distraction should be pursued. In particular, an evaluation of tasks which are intrinsically interesting and perhaps which require more active interaction between the driver and the task may prove to be fruitful.

2. If a drowsiness countermeasure is developed, the conditions under which it maintains its effectiveness should be identified. It is almost certainly true that a countermeasure will work only if the driver is not past a certain point of sleep deprivation and/or exhaustion. Therefore, the limits within which the countermeasure is effective should be determined to insure that the incorporation of a drowsiness countermeasure in a vehicle is worth the cost involved.

The reported research was more successful in identifying combinations of subsidiary task and driving measures which can be used to predict drowsiness-induced driving impairment. However, these results do not represent an end to this line of research. The following suggestions are offered in regard to future research dealing with subsidiary task measures as predictors of drowsiness:

1. The search for other subsidiary task measures which are predictive of drowsiness-induced impairment should be continued. While the current research provided evidence that the response time variables were not predictive of drowsiness, the variable NUMHIT was demonstrated to be sensitive to drowsiness. It is possible, however, that other variables associated with different tasks may provide even greater predictive ability. Research thus needs to be conducted to identify these variables.
2. The discriminant models which were found to be predictive of drowsiness-induced impairment should next be validated using a data sample other than that from which the models were derived. Additionally, to more adequately validate these models, a sample using both sleep deprived _and_ rested subjects should be used. This should eliminate the statistical problems

associated with discriminant analyses results based upon groups with greatly disproportionate sizes. This type of validation is necessary to estimate the potential success of using these measures in an on-board drowsiness detection system.

3. The search for an unobtrusive method of obtaining heart rate data should be continued so that, if the previously reported detection models are validated, heart rate data can be feasibly and reliably obtained through the use of an on-board monitor.
4. Additional analyses should be conducted to determine the improvement in drowsiness detection capability that is afforded by using one of the subsidiary tasks. While the results of the present study suggest that a measure of subsidiary task response correctness provides an effective means by which drivers can be classified as drowsy or alert, they do not identify the exact improvement in detection capability provided by this subsidiary task measure. Hence, this question needs to be answered.
5. Another possible use of the subsidiary task which should be investigated involves employing the subsidiary task as a second level of detection. In this type of scenario, other driving performance

measures could be used to predict potential drowsiness. Once drowsiness occurs, the driver could be requested to perform the subsidiary task. Subsequently, when drowsiness is detected through the use of combinations of subsidiary task and driving performance measures, the driver could be warned of this hazard. This dual level approach to drowsiness detection may well result in very low detection error rates.

REFERENCES

Attwood, D. A. and Scott, P. L. (1981). The on-line use of performance measures to predict driver fatigue. In _Proceedings of the 25th Annual Meeting of the Human Factors Society_ (pp. 726-730). Human Factors Society.

Boadle, J. (1976). Vigilance and simulated night driving. _Ergonomics_, _19_, 217-225.

Brown, I. D. (1962). Measuring the 'spare mental capacity' of car drivers by a subsidiary auditory task. _Ergonomics_, _5_, 247-250.

Brown, I. D. (1965). A comparison of two subsidiary tasks used to measure fatigue in car drivers. _Ergonomics_, _8_, 467-473.

Brown, I. D. (1966). Effects of prolonged driving upon driving skill and performance of a subsidiary task. _Industrial Medicine and Surgery_, _35_, 760-765.

Brown, I. D. (1967). Decrement in skill observed after seven hours of car driving. _Psychonomic Science_, _7_, 131-132.

Brown, I. D. and Poulton, E. C. (1961). Measuring the spare 'mental capacity' of car drivers by a subsidiary task. _Ergonomics_, _4_, 35-40.

Brown, I. D., Simmonds, D. C. V., and Tickner, A. H. (1967). Measurement of control skills, vigilance and performance of a subsidiary task during 12 hours of car driving. _Ergonomics_, _10_, 665-673.

Buck, L. (1973). Developments in measuring fatigue impairment. _Society of Automotive Engineers_.

Crawford, A. (1961). Fatigue and driving. _Ergonomics_, _4_, 143-154.

Davies, D. R. and Parasuraman, R. (1981). _The psychology of vigilance_. London: Academic Press.

Dingus, T. A., Hardee L., and Wierwille, W. W. (July, 1985). Development of impaired driver detection measures. IEOR Department Report #8504. Vehicle Simulation Laboratory, Human Factors Group. Virginia Polytechnic Institute and State University, Blacksburg, VA.

Drory, A. (1985). Effects of rest and secondary task on simulated truck-driving task performance. Human Factors, 27, 201-207.

Dureman, E. I. and Boden, C. (1972). Fatigue in simulated car driving. Ergonomics, 15, 299-308.

Erwin, C. W. (Ed). (1976). Studies of drowsiness--Final report. Durham, North Carolina: The National Driving Center.

Erwin, C. W., Hartwell, J. W., Volow, M. R., and Alberti, G. S. (1976). Electrodermal change as a predictor of sleep. In C. W. Erwin (Ed.) (1976), Studies of drowsiness--Final report. Durham, North Carolina: The National Driving Center.

Erwin, C. W., Volow, M. R., and Gray, B. (1973). Psychophysiologic indices of drowsiness. Society of Automotive Engineers, International Automotive Engineering Congress, Detroit, Michigan, January 8-12, Report No. 730124.

Fagerstrom, K. O. and Lisper, H. O. (1977). Effects of listening to car radio, experience, and personality of the driver on subsidiary reaction time and heart rate in long-term driving task. In R. R. Mackie (Ed.), Vigilance: Theory, operational performance and physiological correlates. (pp. 73-85). New York: Plenum Press.

Fagerstrom, K. O. and Lisper, H. O. (1978). Sleepy drivers: Analysis and therapy of seven cases. Accident Analysis and Prevention, 10, 241-250.

Forbes, T. W. (Ed.). (1972). Human factors in highway traffic safety research. New York: Wiley and Sons.

Fredrik, W. S. (1966). Alert-O-Matic, Liberty Mutual Research Center, Hopkinton, Massachussetts. In T. W. Forbes (Ed.) (1972), Human factors in highway traffic safety research. New York: Wiley and Sons.

Harris W. V. (1967). The effect of ammonia and ammonium carbonate in the reduction of drowsiness in the human operator, Masters Thesis, Department of Systems Engineering, University of Arizona. In T. W. Forbes (Ed.) (1972), Human factors in highway traffic safety research. New York: Wiley and Sons.

Harris, W. and Mackie, R. R. (1972). A study of the relationship among fatigue, hours of service and safety of operations of truck and bus drivers. Human Factors Research Inc. Technical Report 1727-2, Goleta, CA.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., and Dement, W. C. (1973). Quantifications of sleepiness: A new approach. Psychophysiology, 10, 431-436.

Hoffman, H., Mayer, H., Grundei, M., and Meier, W. (1971-72). Verkehrsmedizinische untersuchunger zur frage des kreislaufrerhaltens, der korperlichen und seelischen belastung von kraftfahrzeugfuhrern in verschiedenen fahrzeugtypen und auf fernfahrten. Zentralblatt Verkehrs-Medizin, 17, 193-224 and 18, 1-28. Cited in Lisper, H. O., Eriksson, B., Fagerstrom, K. O., and Lindholm, J. (1979). Diurnal variation in subsidiary reaction time in a long-term driving task. Accident Analysis and Prevention, 11, 1-5.

Hulbert, S. (1972). Effects of driver fatigue. In T. W. Forbes (Ed.), Human factors in highway traffic safety research. New York: Wiley and Sons.

Huntley, M. S. and Centybear, T. M. (1974). Alcohol, sleep deprivation, and driving speed effects upon control use during driving. Human Factors, 16, 19-28.

Johnson, R. A. and Wichern, D. W. (1982). Applied multivariate statistics. New Jersey: Prentice-Hall, Inc.

Jones, T. O., Kelly, A. H., and Johnson, D. R. (1978). Half a century and a billion kilometers safely. Society of Automotive Engineers, Inc. Technical Paper Series 780621, Passenger Car Meeting, Troy, MI, June 5-9.

Kaluger, N. A., and Smith, G. L., Jr. (1970). Driver eye-movement patterns under conditions of prolonged driving and sleep deprivation. Night Visibility and Driver Behavior, Highway Research Record, 336. Cited in T. W. Forbes (Ed.) (1972), Human factors in highway traffic safety research. New York: Wiley-Interscience.

Kearney, P. W. (1966). Highway homicide. New York: Cromwell.

Kinsman, K. A. and Weiser, P. C. (1976). Subjective symptomatology during work and fatigue. In E. Simonson and P. C. Weiser (Eds.) Psychological aspects and physiological correlates of work and fatigue. Springfield, Illinois: Charles C. Thomas.

Kinsman, K. A., Weiser, P. C., and Stamper, D. A. (1973). Multidimensional analysis of subjective symptomatology during prolonged strenuous exercise. Ergonomics, 16, 211.

Kousmens, A. J. R., Tursky, B., and Solomon, P. (1968). Electrodermal levels and fluctuations during normal sleep. Psychophysiology, 5, 300-306.

Laurell, H. and Lisper, H. O. (1978). A validation of subsidiary reaction time against detection of roadside obstacles during prolonged driving. Ergonomics, 21, 81-88.

Lemke, M. (1982). Correlation between EEG and driver's actions during prolonged driving under monotonous conditions. Accident Analysis and Prevention, 14, 7-17.

Leonard, J. and Wierwille, W. W. (1975). Human performance validation of simulators: Theory and experimental verification. In Proceedings of the 19th Annual Meeting of the Human Factors Society, (pp. 446-456). Dallas, Texas: Human Factors Society.

Lisper, H. O. (1966). Trotthet hos bilforare. En provning av reaktionstid som prestationsmatt. Unpublished Phil. lic. thesis. Department of psychology, University of Uppsata. Cited in Lisper, H. O., Dureman, I., Ericsson, S., and Karlsson, N. G. (1971) Effects of sleep deprivation and prolonged driving on a subsidiary auditory reaction time. Accident Analysis and Prevention, 2, 335-341.

Lisper, H. O., Dureman, I., Ericsson, S., and Karlsson, N. G. (1968). Effects of prolonged driving upon a subsidiary serial reaction time. 1. Comparison with brake reaction time to stimuli on the road. Department of Psychology, University of Uppsala, Report 52. Cited in Lisper, H. O., Dureman, I., Ericsson, S., and Karlsson, N. G. (1971) Effects of sleep deprivation and prolonged driving on a subsidiary auditory reaction time. Accident Analysis and Prevention, 2, 335-341.

Lisper, H. O., Dureman, I., Ericsson, S., and Karlsson, N. G. (1971). Effects of sleep deprivation and prolonged driving on a subsidiary auditory reaction time. Accident Analysis and Prevention, 2, 335-341.

Lisper, H. O., Eriksson, B., Fagerstrom, K. O., and Lindholm, J. (1979). Diurnal variation in subsidiary reaction time in a long-term driving task. Accident Analysis and Prevention, 11, 1-5.

Lisper, H. O., Laurell, H., and Stening, G. (1973). Effects of experience of the driver on heart-rate, respiration rate, and subsidiary reaction time in a three hours continuous driving task. Ergonomics, 16, 501-506.

Mackie, R. R. (Ed.). (1977). Vigilance: Theory, operational performance and physiological correlates. New York: Plenum Press.

Mast, T. M., Jones, H. W., and Heimstra, N. W. (1966). Effects of fatigue on performance in a driving device. Highway Research Record, 122, 93.

McDonald, N. (1981). Safety and regulations restricting the hours of driving of goods vehicle drivers. Ergonomics, 24, 475-485.

McLean, J. R. and Hoffman, E. R. (1975). Steering reversals as a measure of driver performance and steering task difficulty. Human Factors, 17, 248-256.

Muto, W. H. (1981). A multivariate investigation of driver performance changes during extended driving periods in a computer-controlled driving simulator. Unpublished doctoral dissertation. Virginia Polytechnic Institute and State University, Blacksburg, VA.

Muto, W. H. and Wierwille, W. W. (1982). The effects of repeated emergency response trials on performance during extended-duration simulated driving. Human Factors, 24, 693-698.

Nelson, T. M., Laden, C. J., and Carlson, D. (1979). Perceptions of fatigue as related to alcohol ingestion. Waking and Sleeping, 3, 115-135. Cited in Ryder, J. M., Malin, S. A., and Kinsley, C. H. (1981), The effects of fatigue and alcohol on highway safety. National Highway Traffic Safety Administration Report No. DOT-HS-805-854.

Nunney, D. N. (1963). Fatigue, impairment, and psycho-motor learning. Perceptual and Motor Skills, 16, 369-375.

O'Hanlon, J. F. and Kelley, G. R. (1977). Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation. In R. R. Mackie (Ed.) (1977), Vigilance: Theory, operational performance and physiological correlates (pp. 87-109). New York: Plenum.

Onboard electronic advisor warns of erratic steering behavior. (1984). Automotive Engineering, 92, 101-102.

Pearson, R. G. and Byars, G. E., Jr. (1956). The development and validation of a checklist for measuring subjective fatigue. United States Air Force School of Aviation Medicine Report, No. 56-115. Cited in Ryder, J. M., Malin, S. A., and Kinsley, C. H. (1981). The effects of fatigue and alcohol on highway safety. National Highway Traffic Safety Administration Report No. DOT-HS-805-854.

Pierson, W. R. (1963). Fatigue, work decrement, and endurance in a simple repetitive task. British Journal of Medical Psychology, 36, 279.

Riemersma, J. B. J., Sanders, A. F., Wildervanck, C., and Gaillard, A. W. (1977). Performance decrement during prolonged night driving. In R. R. Mackie (Ed.), Vigilance: Theory, operational performance and physiological correlates. (pp. 41-58). New York: Plenum Press.

Ryder, J. M., Malin, S. A., Kinsley, C. H. (1981). The effects of fatigue and alcohol on highway safety. National Highway Traffic Safety Administration Report No. DOT-HS-805-854.

Safford, R. and Rockwell, T. H. (1967). Performance decrements in twenty-four hour driving. Highway Research Record, 163, 68-79.

SAS Institute Inc. (1982). SAS user's guide: Statistics, 1982 edition. Cary, NC: SAS Institute Inc.

Skipper, J. H., Wierwille, W. W., and Hardee, L. (1984, September). An investigation of low-level stimulus-induced measures of driver drowsiness. IEOR Department Report #8402. Vehicle Simulation Laboratory, Human Factors Group. Virginia Polytechinc Institute and State University, Blacksburg, VA.

Sussman, E. D., Sugarman, R. C., and Knight, J. R. (1971). Use of simulation in a study investigating alertness during long-distance, low-event driving. Highway Research Record, 364, 27-32.

Tabachnick, B. G. and Fidell, L. S. (1983). Using multivariate statistics. New York: Harper and Row, Publishers.

Technology Associates, Inc. (1978). Description and operating instructions for eye closure detector. Technology Associates, Inc., W. Lafayette, Indiana.

Tilley, D. H., Erwin, C. W., and Gianturco, D. T. (1973). Drowsiness and driving: Preliminary report of a population survey. Society of Automotive Engineers, International Automotive Engineering Congress, Detroit, Michigan, January 8-12, Report No. 730121.

Travis, R. C. and Kennedy, J. L. (1947). Prediction and automatic control of alertness. I. Control of lookout alertness. Journal of Comparative and Physiological Psychology, 40, 457-461.

Virginia traffic crash facts. (1983). Department of State Police, Richmond, VA.

Volow, M. R. and Erwin, C. W. (1973). The heart rate variability correlates of spontaneous drowsiness onset. Society of Automotive Engineers, International Automotive Engineering Congress, Detroit, Michigan, January 8-12, Report No. 730124.

Volow, M. R., Erwin, C. W., Cipolat, A. L., and Hartwell, J. W. (1976). Operant conditioning of skin potential level. In C. W. Erwin (Ed). (1976), Studies of drowsiness--Final report. Durham, North Carolina: The National Driving Center.

Wetherell, A. (1981). The efficacy of some auditory-vocal subsidiary tasks as measures of the mental load on male and female drivers. Ergonomics, 24, 197-214.

Wierwille, W. W. and Muto W. H. (1981, May). Significant changes in driver-vehicle response measures for extended simulated driving tasks. Paper presented at the First European Annual Conference on Human Decision Making and Manual Control, Delft, Netherlands, 298-314.

Wierwille, W. W. (1975). Driving simulator design for realistic handling. In H. K. Sachs (Ed.), Proceedings of the Third International Conference on Vehicle System Dynamics. Amsterdam, Netherlands: Swets and Zeitlinger, 186-199.

Wilkinson, R. T. (1965). Sleep deprivation. In O. G. Edholm and A. L. Bacharach (Eds.), The physiology of human survival. (pp. 399-430). New York: Academic Press.

Williams, H. B. (1966). Personal communications with the author, Alertmaster Company, Houston, Texas. In T. W. Forbes (Ed.) (1972), Human factors in highway traffic safety research. New York: Wiley and Sons.

Williams, H. L., Lubin, A., and Goodnow, J. J. (1959). Impaired performance with acute sleep loss. Psychological Monographs, 73, (whole no. 484).

Williges, R. C. (1982). Class notes.

Wolf, G. (1967). Construct validation of measures of 3 kinds of experimental fatigue. Perceptual and Motor Skills, 24, 1067-1076.

Yajima, K., Ikeda, K., Oshima, M., and Sugi, T. (1976). Fatigue in automobile drivers due to long time driving. Society of Automotive Engineers, Report No. 760050.

APPENDICES

Appendix A: Driving Questionnaire

To be filled in by experimenter:

Subject Number: __________

The following questionnaire is designed to investigate your driving and sleeping habits for participation in the automobile simulator study. All answers will be confidential and will be treated anonymously.

Complete the following items:

1. Year of Birth ________________________

2. Check one: Male ______ Female ______

3. Occupation: _________________________

Circle the correct response to the following:

4. Marital Status: Married Single Divorced Separated Widowed

5. Do you smoke? Yes No

6. If you are a smoker, do you smoke:

Lightly Moderately Heavily

7. Do you have a valid drivers license? Yes No

9. What are your normal sleeping hours?

Retire________ Awake________

10. On the average, what is your depth of sleep?

Restless Light Moderate Deep Very Deep

11. Do you have any pathological (serious) sleep disorders?

Yes No

Sleeping Habits

Check the blank that is most appropriate.

	Never	Almost Never	Occasionally	Moderately Often	Very Often
1. Do you normally fall asleep during the day?	_____	_____	_____	_____	_____
2. How often do you fall asleep watching TV?	_____	_____	_____	_____	_____
3. How often do you fall asleep reading?	_____	_____	_____	_____	_____
4. Do you fall asleep after lunch?	_____	_____	_____	_____	_____
5. Do you fall asleep after dinner?	_____	_____	_____	_____	_____
6. How often do you take daily naps?	_____	_____	_____	_____	_____

Driving Habits

Check the blank that is most appropriate.

	Never	Almost Never	Occasionally	Moderately Often	Very Often
7. How often have you experienced drowsiness while driving?	_____	_____	_____	_____	_____
8. How often have you driven for more than three hours at a time?	_____	_____	_____	_____	_____
9. How often have you had trouble staying awake in situations other than driving?	_____	_____	_____	_____	_____

10. Estimate mileage driven yearly: (Check one)

_____a. 0 - 5,000 miles.
_____b. 5,000 - 10,000 miles.
_____c. 10,000 - 20,000 miles.
_____d. greater than 20,000 miles.

11. On long trips: (Check one)

_____a. I do all of the driving.
_____b. I do most of the driving.
_____c. I share the driving equally.
_____d. I do little of the driving.
_____e. I never drive on long trips.

12. When I drive when I am tired: (Check one)

______a. My driving is the same as when I am rested.
______b. My driving is not as good as when I am rested.
______c. I occasionally doze or nod off.
______d. I often doze or nod off.

Appendix B: Instructions to Subjects

Introduction to the Automobile Simulator Study

The purpose of this study is to investigate driving performance after a period of partial sleep deprivation. The study is being conducted at the Vehicle Simulation Laboratory, Department of Industrial Engineering and Operations Research (IEOR), Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061, telephone number 961-7962. The research team consists of Lenora Hardee, Thomas Dingus, and Jonathan Antin who are graduate students in Industrial Engineering and Operations Research, and Dr. Walter W. Wierwille, (Principal Investigator), Professor of Industrial Engineering and Operations Research.

Your task for this study is to drive an automobile simulator in a simulated highway driving scenario. While you are driving, you will also be asked to perform three different subsidiary tasks. You should remember, however, that these tasks are of secondary importance to the task of driving. You should therefore not let the performance of these subsidiary tasks interfere with driving. Please drive in the right lane and maintain a speed of 55 mph. For your safety, you will be required to wear the lap seat belt at all times during the driving session.

You have the right to discontinue driving at any time if you feel it is necessary. However, for your own safety you must not disconnect the lap belt or attempt to leave the simulator before the motion platform has been stopped. If you feel you must discontinue the driving session, you may do so in one of the two following ways:

1. a. Inform the experimenter of your wish to stop.
    b. Remain seated while the experimenter stops the simulator.
    c. Upon instruction from the experimenter, disconnect your lap belt and wait for the experimenter to assist you while you exit the simulator.

Or alternatively,

2. a. Inform the experimenter of your wish to stop.
    b. Press the white emergency stop button located on the dash of the simulator.
    c. Remain seated with the lap belt attached until the experimenter can assist you in exiting the simulator.

While it is your right to terminate the driving session at any time, the research team would prefer for you to finish the experiment if you can.

The study consists of one short (approximately 30 minutes) practice session and one experimental driving session. The purpose of the practice session is to familiarize you with the simulator. During the practice session you will also be given complete instructions concerning the performance of three subsidiary tasks. The experimental session will take place shortly after midnight and will last for approximately three hours. A member of the research team will arrange a time for you to complete the experiment.

For the experimental session, it is necessary that you drive after you have been deprived of sleep. To accomplish this, a member of the research team will pick you up at your residence at approximately 6:00 p.m. and will take you to dinner at a fast food restaurant. You will then be taken to the building in which the Vehicle Simulation Laboratory is located where you will participate in a familiarization session. Once you have completed the familiarization session you will be required to wait until it is time for your driving session to begin. During this time you may read, study, watch TV, or listen to music. You will not be allowed to eat, smoke, or drink coffee or soft drinks, however, since your ingestion of these stimulants may affect the outcome of the experiment. You will remain awake throughout the evening. A member of the research team will remain with you during this time and will awaken you if you fall asleep.

Shortly after midnight you will begin the driving session. This session will take approximately three hours. It is possible that you will fall asleep during this session. Should this happen, the experimenter will awaken you. You will be in no danger if the simulator runs off the road; however, you should try to remain awake during the driving task. It is important that you try to keep driving even though you may be tired and sleepy.

During the driving session, data will be recorded, and will be later analyzed and treated with anonymity. While you may be concerned that you are not able to perform the driving task well due to drowsiness, you should not worry. This experiment is not designed to test your driving skills, but rather to obtain information on the influence of sleep deprivation on driving. Therefore, all the research team asks is that you make an effort to drive normally.

Once you have completed the experimental driving session, you will be paid at the rate of $4.00 per hour for all the time you spent in performing the experiment, including the evening hours prior to the driving session. If for some reason you are unable to complete the experiment, you will be paid only for the amount of time you actually spent in the study. You will then be driven home by a member of the research team and told to go to sleep. Under no circumstances will you be allowed to drive yourself or walk home.

If you have any questions about the experiment or your rights as a participant, please feel free to ask. We will answer your questions as honestly as possible; however, answers to some of your questions may be delayed until all experimental sessions are completed to avoid affecting the outcome of the study. Please do not discuss the experiment with anyone until after data for all participants have been collected. Since data collection is expected to be completed by May 1, 1985, you may discuss the study with anyone you wish after that date.

There are certain potential risks associated with this experiment of which you should be aware. First, a risk of injury to you exists if you attempt to leave the driving simulator before its motion has stopped. Second, participation in this experiment may result in some interference in your activities (due to possible fatigue or drowsiness) on the day following the driving session.

It will be necessary for you to provide us with informed consent before you may participate in the experiment. The Informed Consent Form is on the next two pages.

The faculty and graduate students involved in the study greatly appreciate your help as a participant.

Participant's Informed Consent

The purpose of this document is to obtain your consent to participate in this experiment and to inform you of your rights as a participant.

Rights of Participation

(1) You have the right to stop participating in the experiment at any time. If you choose to terminate the experiment, you will receive pay only for the time you participated in the experiment.

(2) You have the right to be informed of the overall results of the experiment, but anonymity will be preserved. If, after participation, you wish to receive summary information please include your address (six months hence) with your signature below. If more detailed information is desired after receiving the results summary, please contact the Vehicle Simulation Laboratory, and a full report will be made available to you.

Inherent Risks

(1) The risk of injury to you exists if you attempt to exit the driving simulator before motion has stopped.

(2) You may experience some interference in your activities on the day following the driving session.

Obligations

(1) You will not be allowed to eat, smoke, or drink coffee or soft drinks after 6:00 p.m. and before the experimental session.

(2) After the driving session, you will be driven to your residence by a member of the research team. Under no circumstances will you be allowed to drive yourself or walk home.

Your signature below indicates you have read the above stated rights and you consent to participate in the study described. If you include your printed name and address below, a summary of the experimental results will be sent to you.

______________________ Signature	______________________
______________________ Date	______________________
	______________________ Printed name and address
______________________ Witness Date	Vehicle Simulation Laboratory IEOR Department Virginia Tech Blacksburg, Virginia 24061 961-7962

This is to verify that this experiment has been explained to the subject and he/she has had an opportunity to have all aspects of the research explained.

Appendix C: Subsidiary Task Instructions

During the driving session you will be required to also perform three different subsidiary tasks while you drive. These tasks will provide you with tactual, visual, and auditory stimuli to which you will respond with a manipulative output (i. e., you will respond by depressing one or both of the pushbuttons located on the right and left sides of the steering wheel). All three tasks are very similar and require similar responses. The major difference among the three tasks is the sensory modality of the stimulus. A description of the three types of stimuli and the responses which you must make to them is given below.

Visual task. During the performance of the visual task you will be required to monitor the four yellow lights which can be seen on the roadway display of the simulator. If one or more are lit, you should respond in one of the following ways:

1. If more lights are lit to the left of center than to the right, depress the left pushbutton.
2. If more lights are lit to the right of center than to the left, depress the right pushbutton.

3. If one light is lit on both sides, depress both pushbuttons.
4. If all four lights are lit, do not make any response.

The lights will remain on for five seconds. You therefore have five seconds to see the lights and to decide how to respond.

If you feel that you understand how to perform the visual task, you may now practice for a few minutes.

<u>Tactual task</u>. The tactual task is very similar to the visual task. To perform this task, you will be required to "feel" for the extension of four tabs on the back of the steering wheel. (There are two tabs on each spoke of the steering wheel.) If one or more of these tabs extend and vibrate, this represents a stimulus input. You must respond to this input in one of the following ways:

1. If more tabs are extended on the left side than on the right, depress the left pushbutton.
2. If more tabs are extended on the right side than on the left, depress the right pushbutton.
3. If one tab is extended on each side, depress both pushbuttons.
4. If all four tabs are extended, do not make any response.

While you are performing this task, you should try to remember that, if you press too tightly against the tabs on the back of the steering wheel, you may prevent them from extending. Therefore, you should try to find a comfortable hand position which will allow you to feel the tabs but which will allow the tabs to extend. Once you feel them extend, however, if you press lightly against the tabs, one at a time, you will be better able to tell which ones are vibrating. Also, once the tabs have extended, they will remain out for five seconds; you therefore have five seconds to "feel" the tabs and to decide how to respond.

Auditory task. The auditory task is also similar to the other two tasks. To perform this task, you must listen for four different tones. These tones will originate from four different sources which are located at the top and bottom of the posts on each side of the windscreen. During the task, the tones will be presented to you in sequence, one tone at a time. The tone presentation will begin with the bottom left speaker and will continue in a clockwise direction. The first tone will last for two seconds, and the other three tones will last for one second each. Therefore, it will take five seconds for all four tones to be presented. You thus have five seconds to listen for all the tones and to decide how to respond. It is important that you wait until after the fourth tone is presented (or would be presented) before you make your response. If you respond before this time, you may make an incorrect response. However, if you do respond too soon and make an incorrect response, just depress the button(s) again and make the correct response.

Once the task begins, you will be required to respond in one of the following ways:

1. If more tones originate to your left than to your right, depress the left pushbutton.
2. If more tones originate to your right than to your left, depress the right pushbutton.
3. If one tone is heard on each side, depress both pushbuttons.
4. If all tones are presented, do not make any response.

Now that you understand how to perform the auditory task, you may practice for a few minutes.

Some things to remember. During the experimental driving session, you will be asked to drive under four different conditions: (1) while performing the tactual task; (2) while performing the visual task; (3) while performing the auditory task; and (4) while performing no other tasks. The order in which you perform each of the three tasks during the experimental session may be different from the order in which you performed them during the practice session. You will be told when to perform each of the tasks, however.

Finally, you should always remember that your primary task is to drive the simulator. You should therefore only perform the three subsidiary tasks when they do not interfere with driving.

If you have any questions at this point about the three tasks or about the experiment in general, please feel free to ask. Your participation and cooperation is greatly appreciated.

Appendix D: Subjective Rating Scale

Please rate how alert you were during the last driving segment using the following scale:

-3______-2______-1______0_______1_______2_______3

nearly asleep   extremely alert

Appendix E: Complete Discriminant Analyses Results

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	31	21 (40.38)	52
	Not Impaired	2 (25.00)	6	8
		33	27	60

APER = 38.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0434
HRTRTM	-0.2996

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	51	1 (1.92)	52
	Not Impaired	3 (37.50)	5	8
		54	6	60

APER = 6.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.0748
HRTRTV	1.3103

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	28	24 (46.15)	52
	Not Impaired	1 (12.50)	7	8
		29	31	60

APER = 41.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	-0.7598
SEATMOV	0.8161

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	21	31 (59.62)	52
	Not Impaired	0 (0.00)	8	8
		21	39	60

APER = 51.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.1211
LANEX	1.0973

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	27	25 (48.08)	52
	Not Impaired	2 (25.00)	6	8
		29	31	60

APER = 45.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.7537
LANDEVM	-0.5949

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted: Impaired	Predicted: Not Impaired	
Actual	Impaired	21	31 (59.62)	52
	Not Impaired	0 (0.00)	8	8
		21	39	60

APER = 51.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.0097
LANDEVV	1.0213

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	17	35 (67.31)	52
	Not Impaired	0 (0.00)	8	8
		17	43	60

APER = 58.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.2884
LANDEV4	-0.8006

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	26	26 (50.00)	52
	Not Impaired	1 (12.50)	7	8
		27	33	60

APER = 45.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.1681
LANDEVSQ	1.1276

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	39	13 (25.00)	52
	Not Impaired	2 (25.00)	6	8
		41	19	60

APER = 25.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	-0.6494
YAWMEAN	-0.9471

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	26	26 (50.00)	52
	Not Impaired	0 (0.00)	8	8
		26	34	60

APER = 43.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.0089
YAWVAR	1.0275

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	21	31 (59.62)	52
	Not Impaired	0 (0.00)	8	8
		21	39	60

APER = 51.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.1010
LATPOSMS	1.0840

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	28	24 (46.15)	52
	Not Impaired	0 (0.00)	8	8
		28	32	60

APER = 40.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.0991
HIPASSYAW	1.0834

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	24	28 (53.85)	52
	Not Impaired	0 (0.00)	8	8
		24	36	60

APER = 46.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	-0.5822
STEXEED	0.8135

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	21	31 (59.62)	52
	Not Impaired	2 (25.00)	6	8
		23	37	60

APER = 55.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0131
STVELM	-0.1249

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	29	23 (44.23)	52
	Not Impaired	2 (25.00)	6	8
		31	29	60

APER = 41.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	-0.5536
STVELV	0.8276

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	30	22 (42.31)	52
	Not Impaired	1 (12.50)	7	8
		31	29	60

APER = 38.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	-0.3673
LGREV	0.8733

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	23	29 (55.77)	52
	Not Impaired	3 (37.50)	5	8
		26	34	60

APER = 53.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.9374
MDREV	-0.2945

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	32	17 (34.69)	49
	Not Impaired	2 (18.18)	9	11
		34	26	60

APER = 31.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.7612
HRTRIM	0.5369

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	45	4 (8.16)	49
	Not Impaired	4 (36.36)	7	11
		49	11	60

APER = 13.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.3853
HRTRTV	1.1152

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	33	16 (32.65)	49
	Not Impaired	0 (0.00)	11	11
		33	27	60

APER = 26.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	-0.6892
SEATMOV	0.8732

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	25	24 (48.98)	49
	Not Impaired	1 (9.09)	10	11
		26	34	60

APER = 41.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.8562
LANEX	-0.2300

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	26	23 (46.94)	49
	Not Impaired	1 (9.09)	10	11
		27	33	60

APER = 40.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.8808
LANDEVM	-0.5263

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	25	24 (48.98)	49
	Not Impaired	1 (9.09)	10	11
		26	34	60

APER = 41.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.9454
LANDEVV	-0.1268

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	26	23 (46.94)	49
	Not Impaired	1 (9.09)	10	11
		27	33	60

APER = 40.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.6572
LANDEVSQ	-0.4590

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	25	24 (48.98)	49
	Not Impaired	1 (9.09)	10	11
		26	34	60

APER = 41.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.1298
LANDEV4	0.1434

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	24	25 (51.02)	49
	Not Impaired	1 (9.09)	10	11
		25	35	60

APER = 43.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.1059
YAWMEAN	-0.6025

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	24	25 (51.02)	49
	Not Impaired	0 (0.00)	11	11
		24	36	60

APER = 41.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.5641
YAWVAR	-0.5614

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	25	24 (48.98)	49
	Not Impaired	1 (9.09)	10	11
		26	34	60

APER = 41.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.8802
LATPOSMS	-0.2096

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	23	26 (53.06)	49
	Not Impaired	0 (0.00)	11	11
		23	37	60

APER = 43.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.4878
HIPASSYAW	-0.6336

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	29	20 (40.82)	49
	Not Impaired	1 (9.09)	10	11
		30	30	60

APER = 35.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.7417
STEXEED	-0.4904

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	25	24 (48.98)	49
	Not Impaired	1 (9.09)	10	11
		26	34	60

APER = 41.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0318
STVELM	0.0166

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	26	23 (46.94)	49
	Not Impaired	0 (0.00)	11	11
		26	34	60

APER = 38.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.7039
STVELV	-0.6634

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	30	19 (38.78)	49
	Not Impaired	0 (0.00)	11	11
		30	30	60

APER = 31.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.6379
LGREV	-0.7433

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Visual

		Predicted		
		Impaired	Not Impaired	
Actual	Impaired	29	20 (40.82)	49
	Not Impaired	1 (9.09)	10	11
		30	30	60

APER = 35.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.8307
MDREV	-0.5342

Criterion: If EYEMEAS > 9000 then the observation is classified as impaired.

Else, the observation is classified as not impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	18	8 (30.77)	26
	Impaired	5 (14.71)	29	34
		23	37	60

APER = 21.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.9180
HRTRTM	0.5082

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	16	10 (38.46)	26
	Impaired	4 (11.76)	30	34
		20	40	60

APER = 23.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0007
HRTRTV	0.4452

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	19	7 (26.92)	26
	Impaired	4 (11.76)	30	34
		23	37	60

APER = 18.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.1516
SEATMOV	-0.2591

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	13	13 (50.00)	26
	Impaired	3 (8.82)	31	34
		16	44	60

APER = 26.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.5890
LANEX	-0.6935

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted: Severely Impaired	Predicted: Impaired	
Actual	Severely Impaired	18	8 (30.77)	26
	Impaired	7 (20.59)	27	34
		25	35	60

APER = 25.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.9674
LANDEVM	-0.6002

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	13	13 (50.00)	26
	Impaired	3 (8.82)	31	34
		16	44	60

APER = 26.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.8281
LANDEVV	-0.4075

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	15	11 (42.31)	26
	Impaired	4 (11.76)	30	34
		19	41	60

APER = 25.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.5914
LANDEVSQ	-0.6978

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	11	15 (57.69)	26
	Impaired	1 (2.94)	33	34
		12	48	60

APER = 26.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.7974
LANDEV4	-0.4630

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	18	8 (30.77)	26
	Impaired	5 (14.71)	29	34
		23	37	60

APER = 21.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.1172
YAWMEAN	0.0573

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	13	13 (50.00)	26
	Impaired	2 (5.88)	32	34
		15	45	60

APER = 25.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.6653
YAWVAR	-0.6521

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	13	13 (50.00)	26
	Impaired	2 (5.88)	32	34
		15	45	60

APER = 25.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.7450
LATPOSMS	-0.5020

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	13	13 (50.00)	26
	Impaired	3 (8.82)	31	34
		16	44	60

APER = 26.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.6259
HIPASSYAW	-0.6775

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	15	11 (42.31)	26
	Impaired	5 (14.71)	29	34
		20	40	60

APER = 26.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0670
STEXEED	-0.3896

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	16	10 (38.46)	26
	Impaired	5 (14.71)	29	34
		21	39	60

APER = 25.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0729
STVELM	0.2472

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	17	9 (34.62)	26
	Impaired	3 (8.82)	31	34
		20	40	60

APER = 20.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0308
STVELV	-0.4903

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	16	10 (38.46)	26
	Impaired	3 (8.82)	31	34
		19	41	60

APER = 21.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.9728
LGREV	-0.4336

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Auditory

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	16	10 (38.46)	26
	Impaired	4 (11.76)	30	34
		20	40	60

APER = 23.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0723
MDREV	-0.3294

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	16	10 (38.46)	26
	Impaired	7 (20.59)	27	34
		23	37	60

APER = 28.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.7901
HRTRTM	0.7140

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	16	10 (38.46)	26
	Impaired	7 (20.59)	27	34
		23	37	60

APER = 28.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0320
HRTRTV	0.3233

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	21	5 (19.23)	26
	Impaired	6 (17.65)	28	34
		27	33	60

APER = 18.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.9829
SEATMOV	-0.7012

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	15	11 (42.31)	26
	Impaired	2 (5.88)	32	34
		17	43	60

APER = 21.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.4444
LANEX	-0.8047

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	15	11 (42.31)	26
	Impaired	3 (8.82)	31	34
		18	42	60

APER = 23.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0757
LANDEVM	-0.3161

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	15	11 (42.31)	26
	Impaired	3 (8.82)	31	34
		18	42	60

APER = 23.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.5813
LANDEVV	-0.6810

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	15	11 (42.31)	26
	Impaired	2 (5.88)	32	34
		17	43	60

APER = 21.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.5478
LANDEVSQ	-0.7164

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	12	14 (53.85)	26
	Impaired	3 (8.82)	31	34
		15	44	60

APER = 28.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.8387
LANDEV4	-0.3863

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	16	10 (38.46)	26
	Impaired	7 (20.59)	27	34
		23	37	60

APER = 28.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.1094
YAWMEAN	0.0458

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	16	10 (38.46)	26
	Impaired	1 (2.94)	33	34
		17	43	60

APER = 18.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.1629
YAWVAR	-1.1225

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	15	11 (42.31)	26
	Impaired	3 (8.82)	31	34
		18	42	60

APER = 23.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.5498
LATPOSMS	-0.7122

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	18	8 (30.77)	26
	Impaired	1 (2.94)	33	34
		18	42	60

APER = 15.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.0549
HIPASSYAW	-1.2295

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	16	10 (38.46)	26
	Impaired	3 (8.82)	31	34
		19	41	60

APER = 21.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.7122
STEXEED	-0.7081

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	16	10 (38.46)	26
	Impaired	6 (17.65)	28	34
		22	38	60

APER = 26.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	1.0325
STVELM	0.2434

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	17	9 (34.62)	26
	Impaired	3 (8.82)	31	34
		20	40	60

APER = 20.0%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.8181
STVELV	-0.7618

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	18	8 (30.77)	26
	Impaired	3 (8.82)	31	34
		21	39	60

APER = 18.3%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.7476
LGREV	-0.9278

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

Task: Visual

		Predicted		
		Severely Impaired	Impaired	
Actual	Severely Impaired	18	8 (30.77)	26
	Impaired	5 (14.7)	29	34
		23	37	60

APER = 21.7%

Variables:	Standardized Canonical Coefficients:
NUMHIT	0.8786
MDREV	-0.7285

Criterion: If EYEMEAS > 17,500 then the observation is classified as severely impaired.

Else, the observation is classified as impaired.

The two page vita has been removed from the scanned document. Page 1 of 2

The two page vita has been removed from the scanned document. Page 2 of 2