# A MULTIVARIATE INVESTIGATION OF DRIVER PERFORMANCE CHANGES DURING EXTENDED DRIVING PERIODS IN A COMPUTER-CONTROLLED DRIVING SIMULATOR

by

William H. Muto

Dissertation submitted to the Graduate 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, Chairman

E.S. Geller

A.M. Prestrude

D.L. Price

H.L. Snyder

July, 1981
Blacksburg, Virginia

# ACKNOWLEDGEMENTS

This research was sponsored jointly by the General Motors Corporation and Virginia Polytechnic Institute and State University.

A special note of gratitude is extended to Dr. Walter Wierwille who designed and made available the major apparatus for this research, and whose guidance and support made this dissertation possible. Special thanks are also extended to who developed the primary data reduction software, and to for his eager assistance during several stages of the dissertation research.

I would like to express my appreciation to each of my committee members, whose comments helped immeasureably in the accomplishment of this research; and also to those people not mentioned here, who gave me their technical and moral support.

Finally, an expression of thanks also goes to for her love and understanding throughout the years.

## TABLE OF CONTENTS

Page

Page

# LIST OF TABLES

## LIST OF TABLES

## LIST OF TABLES

# LIST OF FIGURES

## LIST OF FIGURES

## LIST OF FIGURES

## LIST OF FIGURES

# INTRODUCTION

Since the 1930s, driver fatigue has been a topic of great concern throughout the world. In Japan, for example, it has been suggested that up to 40% of the casualties due to traffic accidents have been related to driver fatigue (Yajima, Ikeda, Oshima, and Sugi 1976). In the United States, it has been estimated that ". . . 20 to 25 percent of all accidents on high-speed long distance roads, and 40 to 50 percent of the fatalities, are . . . [attributable] to driver fatigue" (Snook and Dolliver, 1976, p. 304). In a survey conducted in 1970, the Bureau of Motor Carrier Safety investigated 286 commercial vehicle accidents. Of these, 111 were attributed to the driver being either asleep at the wheel or inattentive, causing 142 deaths, 281 injuries, and over $4,785,000 in property damage (cited in Harris and Mackie, 1972).

In light of these statistics, it is understandable that substantial resources have been allocated to the problem of driver fatigue. Indeed, awareness of the relationship between accidents and fatigue has led to government regulations which control the on-duty time of commercial drivers. Present regulations state that (Harris and Mackie, 1972):

(1) No motor carrier shall permit or require any driver to drive more than 10 hours following 8 consecutive hours off duty.

(2) No motor carrier shall permit or require any driver to drive for any period after having been on duty 15 hours following 8 consecutive hours off duty.

(3) No motor carrier shall permit or require any driver to be on duty more than 60 hours in any 7 consecutive days except, if a carrier operates every day of the week, this may be extended to 70 hours in 8 consecutive days.

Although the intent of these regulations is commendable, the regulations on driving time seem rather arbitrary in view of the fact that, until quite recently, no objective studies had been conducted to determine the adequacy of these regulations.

## Definition and Measurement of Driver Fatigue

In conducting fatigue related research, the first problem that arises is one of definition. Although the concept of fatigue is well known to the majority of drivers through experience, it remains an elusive concept to define. Part of the problem, according to Crawford (1961), arises from the variety of the word's meaning in common usage and the "lack of direct research." One of many such definitions has been put forward by Bartlett (cited by Case, Hulbert, and Mellinger, 1970):

> Fatigue is a term used to cover all those determinable changes in the expression of an activity which can be traced to the continuing exercise of that activity under its normal operational conditions, and which can be shown to lead immediately or after delay, to deterioration in the expression of that activity, or more simply, to results within the activity that are not wanted (p. 4).

In other words, the performance of a task causes an eventual decrease in the performance of the task. Platt (1964) suggests that fatigue has at least three meanings:

> Subjective fatigue - defined as the feeling of being tired.
> Physiological fatigue - as determined from bodily changes.
> Objective fatigue - When performance of a task shows a progressive deterioration (p. 351).

Others have attempted to qualify further what is meant by adding a prefix to the term; consequently, references to "stress fatigue" (Heimstra, 1970), "generalized fatigue" (Crawford, 1961), etc., are often found in the literature.

Because fatigue cannot be measured directly, and because the definitions give researchers little guidance in this matter, researchers have had to use numerous methods in an attempt to define fatigue operationally. These operational definitions have generally fallen into three categories: performance measures, physiological measures, and subjective reports.

In addition to the determination of what variables are to be used, the investigator must also decide on other factors such as where, when, and how such measures are taken.

Driving environment. As suggested by Harris and Mackie (1972), the investigator has three basic choices of driving environments: (1) open-road--subjects drive actual vehicles on open (public) highways and roads; (2) closed course--subjects drive actual vehicles on a test track or course; and (3) driving simulator--subjects drive in controlled laboratory conditions.

Open-road driving would seem to be the most face-valid approach in that fatigue, presumably, could be observed as it occurs in the real world. Brown, Simmonds, and Tickner (1967), for example, conducted a study in which eight subjects were required to drive for 12 hours in city traffic. During the driving period, a number of subjects' responses were monitored. The measured variables included: brake, clutch, steering wheel movements, accelerator pedal movements, heart rate, oral temperature, and vigilance task performance. Results indicated no significant change in "car-control" skills, while, surprisingly, vigilance task performance improved with time. Brown et al. concluded ". . . a virtually continuous 12 hour period of driving during the normal working day need not affect either perceptual or motor skills adversely" (p. 655).

In a very extensive study conducted by Harris and Mackie (1972), performance and physiological responses of 195 bus and truck drivers were monitored during cross-country driving. Data were gathered during widely varied traffic and road conditions. Total observation time was approximately 1,550 hours, covering 62,000 miles of U.S. highways. With the use of on-board observers, Harris and Mackie were able to determine the effects of various independent variables: (a) type of operations, (b) on-duty time, (c) driving time, (d) rest breaks (frequency and time of occurrence), (e) cycle run (". . . the number of the run, following the drivers' last break of 24 hours or more, in which the driver was observed" (p. 140), and (f) time of day, and (g) various control variables which included terrain, road conditions, weather conditions, and traffic events. (The dependent measures employed will be discussed later.) The general results of the study revealed the following:

1. Significant increases in driver performance errors, and significant decreases in the level of psychophysiological arousal of the drivers occurred within the 10-hour limit on driving time set by the Department of Transportation (DOT).

2. Rest breaks varied in effectiveness on driver performance and psychophysiological arousal depending on the total trip time.

3. Drivers on the road for 40 to 100 hours ("sleeper drivers") appeared to recover less completely from rest breaks than did so-called relay drivers.

4. Several days on duty appeared to cause a cumulative detrimental effect on the drivers' lack of arousal.

5. Older drivers (over 45 years old) seem more susceptible to the adverse effects of prolonged driving than did younger drivers.

6. "There are marked diurnal variations in level of psychophysiological arousal with the lowest levels occurring, for most individuals, between about

2:00 a.m. and 7:00 a.m." (p. $x$).

One potentially serious deficiency of open-road driving environments is that many environmental variables (such as traffic and weather conditions) cannot be controlled or easily quantified, possibly resulting in experimental contamination. It is probable that driver behaviors (especially those related to fatigue) will be affected greatly by varied, uncontrolled environmental stimuli. In the above examples of open road driving, partly conflicting data may be due to the differences in the driving environments. For example, it is possible that the city traffic environment used by Brown et al. (1967) provided stimulation which nullified any possible fatigue effects (resulting in no observable changes in performance of physiological variables). Recognizing this possible confound, Brown et al. (1967) suggested that ". . . any conclusions drawn from these results are applicable only to the stimulating conditions under which driving was tested in city traffic and not to more monotonous conditions of motorway driving, where levels of arousal and performance could be depressed" (p. 670).

Fatigue studies employing closed-course driving usually require subjects to drive a vehicle around a test track for extended periods of time. One advantage of this methodology is that it allows control of many of the situational variables (such as traffic conditions) while maintaining much of the realism of open road driving.

Lisper, Dureman, Ericson, and Karlsson (1971) conducted a study in which six subjects were required to drive an instrumented vehicle around a 1.9 mile track. Two reaction time tasks were used to assess fatigue levels of the drivers. The first required subjects to respond to an auditory stimulus activated by the experimenter; the other task required subjects to respond to lights presented along the edge of the test track. During the four hour drive (between 10:00 p.m. and 2:00 a.m.), the reaction time to the auditory stimulus increased almost linearly with time. The mean response times to the light signals increased only slightly.

Although closed course driving has certain advantages, the methodology is not without disadvantages. Due to the repetitive nature of the task, it seems probable that experimental results can be contaminated by learning effects. For example, in the above study it is possible that the reaction time for the light detection task did not increase because the subjects merely learned to anticipate the onset of lights at certain locations around the track.

The third type of driving environment used in driving research is the driving simulator. The major advantage of driving simulators is that the experimenter is allowed almost total control of the environmental stimuli presented to the driver.

A study using a driving simulator was conducted by Sussman, Sugarman, and Knight (1971). In their investigation, 48 subjects were required to operate a driving simulator for four hours. The simulator employed a hydraulically actuated base, capable of movements in the yaw and roll axes, and a cathode ray tube displaying a roadway image. The movements of the platform and the apparent movements of the roadway were controlled and coordinated by an analog computer. During the course of the study, levels of acoustic noise, task complexity, and task duration were varied to determine their effects on several performance and physiological measures: steering wheel reversals of two degrees and greater; integrated absolute road position error during a simulated tire blowout; response latency to a light, integrated absolute velocity error; and the frequency of occurrence of alpha rhythms in the occipital EEG record. Analysis of the results indicated the following (pp. 30-31):

1. The driver's ability to maintain his vehicle on the road under low-event conditions decreases linearly with time over 4 hours ($p < 0.01$...).

2. The rate of steering wheel corrections made by the driver decreases linearly with time over four hours ($p < 0.01$...).

3. On a per subject basis, there is a significant negative correlation . . . between position error and steering wheel corrections ($p < 0.01$...). The investigators speculate that drivers either sample road position less frequently or they make fewer corrections as time on the road increases.

4. Measurements of position accuracy during a simulated emergency indicate that the driver is less likely to be able to control his vehicle accurately during an emergency after 4 hours of driving than after 1 hour of driving ($p < 0.01$...), and that this decrease in control during the emergency is most severe when the driver has been exposed to a high level of acoustic noise ($p < 0.05$...).

5. Analysis of occipital EEG recordings . . . reveals an overall increase in the frequency of occurrences of alpha bursts for all subjects ($p < 0.004$).

In a study by Snook and Dolliver (1976), a point light source simulator was used to assess (1) the effects of driving time on performance and (2) the effectiveness of other stimuli (recorded music, simulated news broadcasts, and lateral position feedback) as fatigue countermeasures. Ten subjects each drove the simulator for three hours. Lateral position error, speed variations, subjective fatigue, heart rate, and steering reversals were monitored during each trial. The results indicated that with increases in driving time, lateral position error, speed variations, and subjective fatigue increased significantly. Heart rate and steering wheel reversals were found to decrease significantly with driving time. It was further shown that certain lateral position error tolerances could be used as a measure of fatigue. The only "countermeasure" which was effective in reducing time related errors was lateral position feedback.

Heimstra (1970) conducted a study using a driving simulator to assess the effects of stress on driving performance. The simulator (actually a pursuit tracking device) required subjects to maintain a controlled element on the curving center line of a continuous rubber belt. Fifty-four subjects were assigned to three groups: (1) contingent shock group; (2) random shock group; (3) and a no shock (control) group. Subjects in the contingent shock group received shocks (pulsed 45 VDC through electrodes placed on the right forearm) when errors were made in one of the tasks. In the random shock group, subjects received shocks at random intervals, while subjects in the control group received no shocks. The measures obtained for each subject were: tracking error, speed maintenance, reaction time, and vigilance task performance. The task duration was six hours. The results were summarized as follows: (a) no significant difference existed between groups on tracking performance, (b) there was a significant increase in tracking error for the contingent group in the last two hours of the driving time, and (c) the contingent group showed an increase in reaction time with time on the task. The author suggested that "a heightened level of emotional arousal, as measured by changes in skin resistance, will affect performance on a number of tasks that are analogous to those required in the operation of a motor vehicle" (p. 216).

The inherent safety advantages of driving simulators may actually hinder the gathering of valid arousal and performance data. Because simulation removes the danger involved in driving, simulators are often criticized on the basis that they remove or reduce motivation which could result in changes in normal behavior. Some investigators have attempted to assess the changes in motivation, or have attempted to artificially maintain it at a high level. Such a study was conducted by Dureman and Boden (1972). A driving simulator (an "outside-in" simulator) consisted of a small round window through which the subject viewed a black "roadway" 0.3 cm wide, painted on a strip of paper. The subject used a steering wheel to keep a scale model automobile on the moving road during the four-hour driving runs. The independent variables were time on the task and "increased level of activation" determined by the presence of an electric shock. The dependent measures were: pulse rate, respiratory rate, skin resistance level, frequency of skin potential response, neck tension (measured by electromyogram), steering errors, and brake reaction time. The results showed that all subjects exhibited a progressive performance decrement over time. Pulse and respiratory rates were found to decline while skin resistance increased. Subjects expecting electrical shocks exhibited higher levels of arousal (as reflected by subjective and physiological measures), slower performance decrements over time, and lower between- and within-subject variability than subjects not expecting to receive shocks. The results seem to support the view that motivation could be an important factor in evaluating driving simulator data.

A second area of concern in the use of driving simulators is that they do not provide enough realism or "fidelity" to evoke responses which are found in actual driving. Harris and Mackie (1972) stated (in reference to certain types of driving simulators): ". . . by its very nature, the model terrain lacks full realism and the televised image is likely to have poor visual resolution as well as be relatively small in size. Similarly, computer driven simulators, such as the one used by Sussman, Sugarman, and Knight (1971) have . . . a major problem of realism . . . as well as the problem of resolution" (p. 40).

Although the lack-of-fidelity criticism should not be taken lightly, it should not be presumed that all simulators exhibit equal fidelity. The driving simulators used in research and training vary widely in their complexity, mode of implementation, and in their ultimate fidelity. For example, one type of driving simulator commonly used for driver training involves the projection of a motion picture of typical highway or street scenes. In motion picture simulators, the driver is usually required to respond to the projected roadway scenes through controls similar to those in actual automobiles. One problem in such simulators is that the driver's responses remain "open loop"; that is, his/her responses have no effect on the "vehicle's" direction, speed, etc. Because the driver receives little or no feedback of control inputs, the driver's responses in such a simulator may differ substantially from those which may be observed in an actual vehicle.

One of the most elaborate motion picture simulators has been developed at the University of California, Los Angeles (Kemmerer and Hulbert, 1975). The simulator consists of a conventional automobile around which is projected a 360 deg motion picture projection of the driving environment. The driver is given apparent control of the "vehicle" through the steering wheel (which varies the rotational position of the projectors) and accelerator and brake pedals (which control the motion picture frame speed through rotation of the rear wheels of the vehicle). This type of simulator, although capable of providing high-fidelity visual stimuli, cannot provide many cues which are normally available during on-road driving. In the UCLA simulator, partial feedback is provided for vehicle yaw and speed, while the remaining vehicular degrees of freedom remain essentially open loop. The presentation of some sensory cues in the absence of others could result in a conflict of cues which is the cause most often attibuted to "simulator sickness" (Casali and Wierwille, 1980).

Although they provide less visual realism than motion picture simulators, other types of simulators allow the operator to control the simulated vehicle in a manner more like real automobiles. In such devices, the simulated vehicle responds to the driver's steering wheel, accelerator, and brake pedal control inputs. An example of this type of simulator is the "point light source" simulator which has been used by several investigators (e.g., Ellingstad and Heimstra, 1970; Snook and Dolliver, 1976). Point light source simulators typically employ a direct link between the driver's controls and the controlled elements of the simulator. However, due to the "pure gain" characteristics of such a link, point light source simulators do not accurately reflect vehicular dynamics experienced in normal driving. The lack of accurate dynamics raises the question of the validity of data from studies using this type of simulator.

Another problem with driving simulators, alluded to by Harris and Mackie (1972), is that they often cause "a form of motion sickness" which could influence performance. Observations such as these cast doubt on the validity of the data gathered from the use of such simulators. (The term motion sickness is not always appropriate when applied to simulators, in that subjects have exhibited discomfort or sickness which has been shown to occur in simulators even when no actual physical motion is involved. To distinguish between motion sickness and sickness imparted by simulators, Barrett and Thornton (1968) have suggested the term "simulator sickness" to refer to the latter.)

It has been generally accepted that, although differences exist among individuals, motion/simulator sickness is caused by conflicts between the subject's visual cues and "body cues" from various kinesthetic and vestibular mechanisms (Alexander and Barrett, 1975; Barrett, Thornton, and Cabe, 1970). Therefore, it is plausible that at least some cases of simulator sickness are caused by the absence of motion cues when strong visual cues are present. It also seems plausible that the addition of appropriate motion cues to the simulation will eliminate visual/ body cue conflicts, and hence will reduce the possibility of simulator sickness. Wierwille (McLane and Wierwille, 1975) has developed a driving simulator which supplies the driver with apparently realistic motion cues. The simulator employs a computer-generated "roadway" display and a hydraulically activated motion platform which provides the driver with physical motion cues of roll, yaw, lateral translation, and longitudinal translation. The motion platform and visual display are controlled by analog and hybrid computers which allow, within certain limits, continuous adjustment of the simulated vehicle dynamics. In support of the notion that the presence of appropriate motion cues prevents simulator/motion sickness, McLane and Wierwille (1975, p. 492) observed that "the driver simulator . . . has never induced uneasiness or nausea to any observable extent in any subject."

As stated earlier, the question of fidelity of simulation can affect the validity of research results. Harris and Mackie state that, due to some of the problems mentioned previously, "it must be concluded that the relationship between performance observed in simulators and that on the open-road is tenuous at best" (p. 41). Again, although this argument is often justifiable, the claim of a "tenuous relationship" between open road driving performance and performance in a driving simulator has not been substantiated. For example, Leonard and Wierwille (1975) conducted a study comparing five measures of driver performance in "full-scale" (open road) driving with driver performance in a computerized moving base simulator. They found that among eight adjustment conditions of selected simulator parameters, at least one yielded driver performance data which were essentially equivalent to driver performance in open-road driving. These results led the investigators to conclude that valid driver performance data can be obtained ". . . if the simulator and experimental design satisfy the follow criteria:

1. The simulator must possess good fidelity in those aspects corresponding to the measures taken.
2. The simulator must have the capability of parameter adjustment.
3. A sufficient number of properly selected independent variables and corresponding settings must be employed.
4. Performance data must be obtainable for the standard full-scale vehicle and for each adjustment of the simulator and
5. Accepted methods of experimental design must be used to insure unbiased data and correct conclusions regarding validity (p. 451)."

In another example, Sussman, Sugarman, and Knight (1971), also using a computerized moving base simulator, investigated the effects of acoustic noise, task complexity, and driving duration on several physiological and driving performance measures. With four hours of driving, Sussman et al. found that driving performance declined while certain physiological variables indicated a reduction of arousal. In a subsequent study, Sugarman and Cozad (1972) conducted an open-road investigation of driver performance variables to determine the degree to which open-road data were equivalent to previous driving simulator data. Correlational analyses of the two studies concluded that the results of the open-road study generally supported those of the simulator study, thus helping to confirm the validity of the simulator data. The investigators commented:

> The few failures to confirm specific findings may be attributable to the lack of precise control over the on-road environment, in contrast to the control that can be achieved with a simulator. It is, therefore, our recommendation that future experimental investigations may be conducted with a simulator. On-road studies, however, should be conducted to verify critical results (p. 5-1).

It is clear that each of the driving environments discussed has advantages and disadvantages in dealing with the study of fatigue. Open road driving, having maximum realism, suffers from the difficulty of uncontrolled environmental stimuli. These uncontrolled events may cause changes in behavior which could either be inappropriately attributed to "driving fatigue" or serve to nullify otherwise observable fatigue effects. The end effect could be one of increasing the variability of dependent measures, causing a reduction in their reliability and validity.

Although closed-course driving restricts the quantity and variety of environmental variables more than open-road driving, the former may produce learning effects which could cause a systematic bias of the results.

Although driving simulators allow almost complete control of the environment, the lack of fidelity of (many) simulators casts doubt on conclusions drawn from studies using them. Because of these difficulties, it could be beneficial to use one or more of the above test environments and test methods in order to cross-validate results.

Pre-post test vs. periodic vs. continuous measures. Another choice to be made concerning experimental methodology is when measurements should be taken. The researcher can choose to measure (a) before and after the driving period ("pre-post" measures), (b) at regular intervals during the drive ("periodic" measures), or (c) continuously.

Pre-post measures are often employed when the selected dependent variables cannot be measured while the subject is operating the vehicle. Examples of such variables include skill test batteries (Herbert and Jaynes, 1963), critical flicker fusion frequency (Ohkubo, 1976), written subjective reports, blood sugar levels, etc. For many of these measures, it is most convenient to gather "data" before (the "pre-" test) and after (the "post-" test) a long driving session. The assumption is that changes between pre- and post- measurements reflect changes due to the independent variable in question. Pre-post measures, however, have at least two major drawbacks. First, according to Harris and Mackie (1972), pre-post tests have appeared not to be sensitive to changes presumed to be caused by fatigue. Secondly, pre-post measurements tell very little about what occurs during the driving period. Consequently, nonmonotonic trends may be undetectable using this technique.

Periodic measures fall into three general categories. The first is one in which the subject is required to respond to discrete events during the driving period. Examples of such periodic measures are brake reaction time tests and vigilance task performance.

A second type of periodic measure is one which constitutes a sampling strategy for variables which are normally continuously available. These techniques are used when experimental runs are very long, and it is infeasible or impractical to record data in a continuous record and subsequently analyze all of the data. Examples of this category are heart rate, EEG, and steering wheel reversals.

Some periodic measures fall into a third category. These are measures which are impractical to obtain while the subject is operating the vehicle. Examples of such measures are critical flicker frequency, tests of humoral factors requiring blood samples, blood pressure check, etc.

Continuous measures of drivers' physiological and performance responses have been used in various driving environments, and have been shown to be sensitive to changes occurring with increases in driving time. Continuous measurement variables are usually those which can be made without interfering with the driving task. Examples of these are measures related to driver performance, such as steering wheel reversals and lateral tracking error, or measures related to physiological effects, such as heart rate, EEG, and GSR.

## AN OVERVIEW OF PREVIOUS INVESTIGATIVE FINDINGS

As previously stated, driver fatigue measures fall into three general classes: (1) driver performance measures, (2) physiological measures, and (3) subjective reports. The following section discusses results of investigations using these measures.

### Performance Measures and Fatigue

It is generally assumed that after driving for a period of time on the highway, the driver undergoes certain mental and physiological changes which lead to impairment in driving ability. Although this notion is intuitively appealing, the measurement of impairment of driving performance (as a function of driving time or "fatigue") has not been straightforward. In addition to selecting the test environment, one of the first problems to arise is one of selecting valid performance variables -- those that are relevant to the driving changes over time. Typically, investigators have made various assumptions about fatigue mechanisms, and have hypothesized changes in selected performance measures. Platt (1964) for example, made the following assumptions about the fatigued driver.

1) Driver fatigue will have [an] effect on steering wheel reversal rates, speed change rates and average speed of the vehicle.

2) As the driver becomes fatigued he will accept wider tolerances of both vehicle tracking and speed control.

3) As a driver gets tired his speed may increase or decrease depending on whether his sensitivity to speed change or steering reversal rate is lost first.

4) The driver will usually take more risks as he becomes more fatigued. This will be indicated as an increase in tracking tolerance, and consequently, a decrease in steering reversals if speed is constant.

Degree of risk may also be indicated by an increase in the speed of the vehicle.

5) As the driver becomes tired his speed change rate increases but he usually makes some effort to keep it within balance by accelerator reversals or driving at a slower speed to accommodate it.

6) The most severe fatigue is encountered in driving when the speed change rate increases and accelerator reversal rate decreases. This indicates the driver has ceased to care about speed control (p. 356).

Although Platt's assumptions have not been experimentally verified, they have served as plausible hypotheses to be tested. In 1964, using the (then new) Greenshields Drivometer, Platt equipped a vehicle to measure "the fundamental actions of the driver in controlling direction and speed of his vehicle (p. 353)." The measures taken were steering wheel reversals, speed change rate, accelerator pedal reversals, and brake applications. Two subjects, alternating as observers every one and one half hours, drove a distance of 1200 miles on four lane divided highways. Unfortunately, due to the small sample size and relatively short driving periods, no statistically reliable results could be obtained from the study. Platt did, however, demonstrate that several of the measures used were sensitive to individual driver characteristics and to changes over time. Following Platt's investigation, many researchers have used the aforementioned variables to measure changes in driver performance.

In a study cited earlier, Harris and Mackie (1972) employed monitoring equipment in inter- and intra-state commercial vehicles to sense and automatically record a number of variables, including brake activations, speed, tie rod movements, steering wheel movements, and several physiological measures. On-board observers were used to record various environmental conditions such as traffic, weather, and terrain. The results indicated the following: a) brake activations were not a useful measure because no correspondence could be established between the speed or the magnitude of braking responses; b) speed variability was not related to time on the road; and c) steering reversal rate indicated a general downward trend with time on the road, with a sharp drop after the ninth hour of driving. It was also noted that steering reversals were sometimes inconsistently related to driving time.

In a study conducted by Riemersma, Biesta, and Wildervanck (1977), 12 subjects were required to drive an instrumented vehicle for eight hours during the night. The results revealed significant increases in lane position standard deviation over time. In contrast, steering wheel reversals and lane drift frequency exhibited no significant changes over time.

Safford and Rockwell (1967) conducted a study in which seven subjects were required to drive for 24 hours. One driver was also subjected to sleep deprivation and to sleep deprivation with a drug. The results indicated that measures of velocity, velocity variation, steering reversals, and accelerator reversals exhibited significant increases over the driving period. The velocity mean had a tendency to decrease significantly with time. When viewed on a per subject basis, however, it was found that some subjects exhibited high positive correlations between the variables and time, while other subjects exhibited low or even high negative (-.766) correlations. The results led the investigators to conclude, "that an increase or decrease in the value of one of the dependent variables cannot be taken as an indication of the onset of fatigue without knowing the specific characteristics of the driver that produced the increase or decrease" (p. 78).

Mast, Jones, and Heimstra (1966) conducted a study employing a driving simulator in which 60 male subjects drove for periods of four to six hours.

Measurements of tracking error, speed maintenance, and reaction time were obtained for each subject. Results indicated that subjects in the six-hour driving condition showed significantly greater performance degradation during the last hour of driving than they did during the first hour. Surprisingly, reaction times were found to be faster during the last hour than during the first hour.

Other driving performance measures. Some investigators have hypothesized that extended driving time can affect other unconventional driver performance measures. For example, Brown, Tickner, and Simmonds (1970) conducted a study to test the effects of 12-hours of driving on the number of "risky" overtaking maneuvers initiated by subjects. During the 12 hour drive, each of the six drivers was rated on the number of times unsafe passing maneuvers were initiated. After correcting for the number of passes attempted, the investigators found that risky maneuvers were attempted 50% more often during the last three hours of driving than during the first three hours.

In a similar study of driver "judgment", Lerner, Abbott, and Sleight (1964) investigated the effects of trip duration (in addition to speed, traffic, and illumination) on highway following distances. Fifteen subjects were required to drive for two hours at 33 and 55 mph, during which following distances were measured. The results indicated that following distances did not change with driving time.

Effects of fatigue on responses to unexpected events. Sussman, Sugarman, and Knight (1971) hypothesized that long distance driving could impair drivers' ability to respond to unexpected events. Using a moving-based simulator, simulated tire blowouts, simulated hills, and signal lights were presented to subjects at various times during the driving period. The results indicated that drivers would be less likely to respond accurately during an emergency after four hours of driving than after one hour of driving.

In a similar vein, Lisper, Dureman, Ericson, and Karlsson (1971) investigated the effects of sleep deprivation and prolonged driving on drivers' reaction times to an auditory signal. Eleven male subjects were instructed to drive a car for four hours during each of three experimental sessions. During the driving periods, subjects were required to respond to a 1000 Hz tone by pressing a foot pedal. Results showed that reaction time increased approximately 100 ms over the driving period. It was concluded that this increase in response time reflected drivers' diminished capacity to respond to unexpected traffic situations. In another study, Lisper, Laurell, and Stening (1973) tested reaction times of experienced and inexperienced drivers to an auditory stimulus over a three-hour driving period. Results indicated that inexperienced drivers exhibited an increase in reaction time over the driving period, while experienced drivers actually exhibited a decrease in reaction time. Using a similar approach, the effects of a number of variables have been tested, including car radio listening and personality (Fagerstrom and Lisper, 1977), driving and stationary conditions (Laurell and Lisper, 1976), and validity (comparison with reaction to simulated roadside obstacles, Laurell and Lisper, 1971).

Fatigue effects on information-handling performance. The human operator is often thought of as a "single channel" information-processing organism who processes information from multiple inputs by sampling ("time sharing") various inputs. It has been hypothesized that fatigue causes a reduction in the drivers' channel capacity (total information handling capacity), resulting in an increase of driver errors. This increase of driver errors could be due to the driver's attempt to compensate for his reduced channel capacity by time-sharing driving task inputs at a reduced rate, or by neglecting driving tasks perceived to be lower in priority. Referring to a study in which inconsistent performance results were observed, Hulbert (1972) stated that "This is exactly the type of result we would expect if the tired drivers were alternating between speed monitoring (gas pedal reversals and velocity variation) and path following (steering reversals)" (p. 290).

Several attempts have been made to measure the hypothesized reduction of information handling capacity of fatigued drivers. Most of these have employed subsidiary (secondary) task workload estimation procedures. In a study by Brown (1965), two secondary tasks involving responses to auditory stimuli were compared for their effectiveness in measuring fatigue. During experimental sessions, eight subjects were required to drive a 2.2 mile city traffic circuit while also attending to one of two subsidiary tasks. The first task (called "attention") required subjects to detect an "odd-even-odd" order from a sequence of numbers presented by a tape recorder. The second task (called "memory") required subjects to listen to a series of 10 letters (presented by tape recorder) and detect which letter had been repeated in the sequence. It was found that the attention task distinguished significantly between performance at the beginning and end of an 8-hour driving period and it correlated with driving speed. The results, contrary to Brown's expectations, indicated that the performance on the attention score was better at the end of the driving period than at the beginning. Brown suggested that this improvement was due to an increase of subjects' "reserve (information handling) capacity".

In an aforementioned study, Riemersma et al. (1977) conducted a study in which subjects drove for eight hours at night. During the drive, subjects were required to perform two subsidiary tasks. In the first, subjects were required to monitor driving distance on the vehicle's odometer and report driving distance increments of 20 km. In the second task, subjects were required to respond to the change of color in lights mounted on the dashboard. The results indicated that in the first task, the number of incorrect responses increased significantly during night driving. For the second subsidiary task, it was found that the number of missed signals and "mental blocks" (responses with reaction times more than twice the median) increased significantly from the first part of the driving period to the second part. Both of these results indicate an impairment in subsidiary task performance with driving time.

Brown et al. (1967) conducted a study in which eight subjects were required to drive while monitoring the onset of dim lights mounted in the side mirrors of the vehicle. Subjects' driving performance, as measured by brake, clutch, accelerator pedal, and steering wheel motions were recorded during the drive. The results indicated that vigilance performance actually improved during prolonged driving. In addition, it was found that none of the measures indicated any deterioration in driving performance during the driving period. Brown et al. suggested that with driving time there is an "automatization of control skills during prolonged work (which) leaves more time for perceptual skills" (p. 670).

A study to determine the effects of driving time on the performance of a subsidiary (vigilance) task was conducted by Dobbins, Tiedmann, and Skordahl (1963). Forty-two subjects drove heavily loaded trucks around an experimental highway for nine hours. Subjects were required to monitor concurrently a display containing a number of red and white colored lights. The task required subjects to detect the onset of red lights (presented about 30 times per hour) which were intermingled with the presentation of white lights (every 5 to 30 seconds). The results showed a surprisingly high performance on the subsidiary task, with only a small decrement over time. The authors suggested that the vigilance task "may have served as a stimulant to, rather than a measure of, vigilance performance" (p. 39). The notion that a subsidiary task can actually act as a stimulant to counteract fatigue effects may also be a plausible explanation for the lack of degradation in performance found by Brown _et_ _al_. (1967).

## Physiological Correlates of Fatigue

Researchers have suggested that arousal (or alertness) is maintained by impulses sent to the cortex via the brain stem reticular formation. It is theorized that these impulses cause depolarization of cortical cells so that they are more readily fired by sensory stimulation. Frequent sensory stimulation, which would tend to maintain a high level of arousal, is presumed to facilitate perceptual, psychomotor, and other higher brain mechanisms which, in turn, facilitate driving performance. It is hypothesized that driver fatigue occurs when mechanisms which facilitate driving performance are inhibited due to lack of stimulation.

In search of objective fatigue measures, researchers have employed a number of physiological variables (which presumably reflect changes in central nervous mechanisms), and have attemped to relate them to driving duration and/or measures of driving performance. Variables often measured are (a) heart rate, (b) heart rate variability, (c) brain wave frequency (using EEG), (d) electromyographic activity, (e) galvanic skin response/skin conductance, (f) blood sugar level, (g) blood eosinophil count, (h) respiration rate, (i) blood pressure, (j) oral temperature, (k) critical flicker frequency, (l) salivary pH, (m) eye movement frequency/amplitude, and (n) eye blink frequency.

Harris and Mackie (1972), for example, examined the effects of hours of driving on heart rate and heart rate variability (in addition to a number of other variables) of truck and bus drivers. It was found that of all the measures, heart rate, which was shown to decrease over time, was the most consistent indicator of the effects of time on the road. Heart rate was also found to be related to differences in the age of drivers, the error rates of different groups, the cumulative effects of successive runs, and time of day. In addition to the above, heart rate also showed: (a) low levels of arousal when extreme fatigue was exhibited, (b) positive correlations with driver performance, and (c) lower arousal with poor performance. It was concluded that heart rate is a useful indicator of the effects of time on the road. Heart rate variability was found to be an inconsistent indicator of time on the road in that it seemed highly sensitive to external stimuli.

In a study by Ohkubo (1976), four subjects were observed while driving instrumented vehicles on three different types of highways. Heart rate was measured continuously. Critical flicker frequency, response time, and subjective ratings of fatigue were recorded hourly. It was found that heart rate tended to decrease with time on the road, with unskilled drivers having earlier declines in heart rate than did skilled drivers. The critical flicker frequency was found to drop slightly after the beginning of a drive, but it did not change substantially after that.

Yajima, Ikeda, Oshima, and Sugi (1976) examined the effects of 10 to 24 hours of driving on a number of variables. Their results revealed the following: (a) critical flicker frequency--average CFF increased slightly and then decreased slightly after 4 to 5 hours; (b) blood pressure--showed no change with time; (c) salivary pH--a slight rise (not significant overall); (d) EOG--the saccadic movements decreased in frequency slowly with time, long duration eye movements increased with time; (e) GSR--GSR increased early, but then decreased gradually; (f) inspiratory volume and number--the values decreased slowly with driving time; (g) EEG--not measured reliably. The authors hypothesized that the "hypofunctions" of the cortex, exhibited by the above variables, apparently caused drowsiness, errors in recognition and judgment, and lack of control of the vehicle. The observed fatigue symptoms were: (1) drowsiness, (2) reduction or narrowing of attention, and (3) an increase of subjective fatigue.

## Subjective Ratings of Fatigue

The last category of measures of fatigue is subjective ratings, both those of drivers and of observers. The major problems with subjective ratings are mainly those of reliability and validity. One method used in an attempt to increase the reliability and validity of subjective ratings is the use of rating scales. Kashiwagi (1971), for example, used a factor analytic approach to construct a rating scale to allow a judgment of fatigue from the person's appearance. Twenty appearance items were selected and combined to form a tool for studying subjective aspects of fatigue.

Using a very simple self-rating (binary) technique, Snook and Dolliver (1976) asked subjects to indicate extreme fatigue when they felt they had "dozed off" by pressing a button in the center of the steering wheel. Snook and Dolliver found that the frequency of indicated fatigue increased significantly along with lateral position error and speed variables.

## Effects of Variables Other Than Driving Duration on Fatigue

The most frequently used independent variable in driving studies has been driving duration. Some researchers have attempted to manipulate other variables which are hypothesized to modify the effects of fatigue. Some of the variables studied include stress, circadian rhythm, lighting, noise, rest stops (type, time, and frequency), alcohol, sleep deprivation, driving speed, humoral factors, emergency events, motivation, drugs (including caffeine), food intake, traffic conditions, and fatigue countermeasures.

Konz and McDougal (1968), for example, studied the effects of background music on control activity of drivers. Steering movements, brake applications, and speed changes of 24 subjects were monitored while the drivers listened to three "levels" of music (fast music, slow music, and no music). It was found that certain control movements were affected by the type of music played. Accelerator movements were found to be less frequent with slow music than with no music; and accelerator movements with no music were less frequent than those with fast music.

Huntley and Centybear (1974) examined the effects of alcohol, sleep deprivation, and driving speed on performance. Twelve subjects were required to drive a vehicle through a short obstacle course after having ingested alcohol (or a placebo), and also after 29 hours of sleep deprivation (or normal sleep). Analysis of the results indicated that alcohol significantly increased the frequency of control usage, whereas sleep deprivation decreased frequency of control usage. In addition, a sleep deprivation-beverage interaction was found which implies that sleep deprivation tended to negate the control use symptoms of alcohol ingestion.

In the recent study, Pokorny, Opmeer, and Blom (1981) investigated the effects of shift (early, late, and "broken"), during driving (rural and urban driving), and age on compensatory tracking performance of 12 bus drivers. Oral temperature, presumed to be a measure of driver arousal, was also measured. The results of the study indicated that performance levels of subjects on the tracking task was better in late afternoon shifts than in early morning. The fact that oral temperatures also were higher in the late afternoon than in early morning led the investigators to conclude that tracking performance can be modified by the drivers level of arousal (measured by oral temperature). If these conclusions are valid, it appears that diurnal variations, could have an effect on levels of driver performance.

### Fatigue Countermeasures

One of the goals of driver fatigue research is to discover reliable fatigue "detectors," which could be used to predict and ultimately prevent fatigue-related automobile accidents.

Several devices have been marketed which purportedly will detect driver fatigue. Such devices, however, have not received widespread acceptance--mainly because they simply do not detect "fatigue symptoms" reliably. Designers of devices such as the "Electronic Transistor Safety Alarm" (produces an alarm when the driver's head nods) and the "Button Steering Wheel Alarm" (driver must continually press a button to prevent an alarm) (Hulbert, 1972) have grossly underestimated the complexity of the problem, and consequently have made some inappropriate assumptions concerning the nature of fatigue.

A number of other methods have been proposed to combat fatigue. Snook and Dolliver (1976), for example, have suggested vehicular lateral position feedback as a countermeasure for fatigue effects. (Unfortunately, even if a valid countermeasure, the difficulty of implementation of lateral position feedback on vehicles for highway use would probably prevent its widespread acceptance.) Others have suggested that amphetamines, caffeine, rest, and refreshments (Hulbert, 1972) may serve to counteract fatigue effects. Unfortunately, because there are no standardized methods to test the effects of fatigue on driving, there are likewise no standardized methods to test the effects of so-called countermeasures. Consequently, it is difficult to assess the validity of various methods proposed to counteract fatigue. Similar difficulties arise when attempting to resolve conflicting recommendations. For example, Yajima et al. (1976) suggest that driving time should not exceed two hours without a rest break (of 20-30 minutes). In contrast, Harris and Mackie (1972) believe that drivers should not be allowed to drive for more than three hours without a rest break.

## Vigilance Research

For a number of years, one of the most popular man-machine system research topics has been that of vigilance task performance. Mackworth (1968) describes the vigilance task as follows:

> The vigilance task usually requires a simple motor response to a signal which occurs at random time intervals against a background that is either unchanging, continuously changing, or presenting a regular series of events, most of which are "unwanted" or nonsignal events. The signal is a specified small change in one of these background events. The subject needs to pay continuous attention to the display in order to detect and respond as quickly as possible to these signals, whose occurrence he cannot predict (pp. 308-309).

From this description, it can be seen that vigilance research may have relevance for situations involving long duration, low frequency event driving. In fact, there is some question as to whether the effects of driver fatigue and the so-called "vigilance decrement" are distinctly different phemomena. For example, Harris (1977) states:

> The truck driving task is likely one of the most demanding vigilance tasks. The driver must maintain a continual vigil if he is to perform his task successfully. He must keep his truck positioned in a traffic lane, often scarcely wider than the truck itself, and he must constantly monitor the lane to make sure it is clear of threatening traffic or other obstacles (p. 134).

The lack of a clear distinction between underlying mechanisms of driver fatigue and vigilance has led to the inclusion of studies dealing with driver fatigue in a number of vigilance research publications (see Mackie, 1977).

Although caution should be used when attempting to generalize traditional vigilance research findings to driving performance situations (Smith and Lucaccini, 1969), it seems logical that subjective, physiological, and performance changes exhibited by operators in classical vigilance situations could lead to the understanding of the driver fatigue process. In a review of vigilance research, Davies and Tune (1969) describe the effects of signal characteristics, task variables, subject variables, and environmental variables on vigilance performance (and other correlated measures). Their publication continues as an excellent representation of the major issues in the vigilance research literature. Appendix I presents a brief outline of their findings. (See Mackie, 1977, for recent topics and references on vigilance research).

### General Summary

Because fatigue cannot be measured directly, it must be inferred from selected variables which reflect (a) changes in driver performance, (b) changes in physiological or psychophysiological mechanisms, or (c) subjective feelings of the driver (or judgments of the driver's state by an observer). Examples of these variables and changes found by various investigators are shown in Table 1.

Among driver performance measures, several have been shown to be promising indicators of the effects of prolonged driving. Unfortunately, only a few variables show universal consistency among investigators. For example, there appears to be no disagreement among investigators as to the nature of vehicle lateral position variability over time. In contrast, steering wheel reversals, one of the most widely used performance variables, has been found to exhibit highly divergent trends among various investigations, with some showing increasing trends over time and others showing decreasing or no reliable trends over driving time. Other measures such as vehicle velocity variation, brake activations, accelerator pedal movements, and tie rod movements have been used to measure fatigue with mixed success. Many of these appear to suffer from larger inter- and intra-subject variability. Other performance measures such as those based on driver judgments such as following distance, passing criteria, speeding frequency, tailgating, etc. have had similar mixed success in detecting fatigue. A few investigators who have studied the effects of fatigue on drivers' responses to unanticipated events (simulated emergencies or other auditory or visual cues) have shown that the driver's ability to respond to such events diminishes over time.

A number of physiological variables (and psychophysiological variables) have been used to measure driver fatigue. Among measures which have been shown to be sensitive to changes over time are heart rate, heat rate variability, frequency of change of galvanic skin response, and eye movements. Others which have

**TABLE 1.** Examples of Previous Investigative Findings

Driver Performance Measures

| Variable | Result | Investigation | |
|---|---|---|---|
| Vehicle lateral deviation | Increase position variability over time | Snook and Dolliver | (1970) |
| | | Sussman and Harris | (1970) |
| | | Riemersma et al. | (1977) |
| Steering wheel reversals | Frequency decreases over time | Platt | (1964) |
| | | Sugarmann and Cozad | (1972) |
| | | Sussman and Morris | (1970) |
| | Frequency increases over time | Safford and Rockwell | (1967) |
| | No change or highly variable over time | Brown et al. | (1967) |
| | | Harris and Mackie | (1972) |
| | | Riemersma et al. | (1977) |
| Vehicle velocity variation | Increase with time | Heimstra | (1970) |
| | No change | Harris and Mackie | (1972) |
| Response to unexpected events | Reduced ability to respond to the event after driving for extended periods. | Sussman et al. | (1971) |
| | | Lisper et al. | (1971) |
| Subsidiary (secondary) task performance | No changes over time | Brown | (1965) |
| | | Brown et al. | (1967) |

Physiological Variables

| Variable | Result | Investigation | |
|---|---|---|---|
| Heart rate | Decreases over time | Dureman and Boden | (1972) |
| | | Harris and Mackie | (1972) |
| | | Ohkubo | (1976) |
| | Increase or decrease over time depends on individual | Yajima et al. | (1976) |

**TABLE 1 (cont)**

| | | | |
|---|---|---|---|
| Heart rate variability | Increase over time | Harris and Mackie | (1972) |
| EEG-alpha activity (brain waves, 8-13 Hz) | Increased occurrence over time | Sussman and Morris<br>Sussman *et al.* | (1970)<br>(1971) |
| Changes in galvanic skin response | Decrease over time | Yajima *et al.* | (1976) |
| Respiration rate | Decrease overtime | Dureman and Boden | (1972) |
| Eye movements | Frequency decreases overtime | Kuroki *et al.* | (1975) |
| Critical flicker frequency (CFF) | Decrease over time | Ohkubo | (1976) |
| | Increased over time | Yajima *et al.* | (1976) |
| Subjective Measures | | | |
| Ratings on appearance items | | Kashiwagi | (1971) |
| Button press when driver felt he "dozed off" | Increased frequency | Snook and Dolliver | (1976) |

shown some sensitivity to time include critical flicker frequency, blood pressure, and salivary pH.

Subjective reports by drivers, observers, or both have been used in fatigue studies. One of the major values of subjective measures is that they can be used as a means to validate data from other measures. However, these measures should be used cautiously because of their susceptibility to drivers' biases and lack of reliability. Additional problems lie in the scaling of fatigue level judgments.

"Symptoms" exhibited by operators in vigilance studies are similar to those observed in drivers during long distance, relatively uneventful driving (See Appendix I). The major difference seems to be that changes in vigilance tasks occur more rapidly than in driving situations. The most obvious explanation for the difference in the apparent rate of change of arousal is that the driving task imposes a heavier workload on the operator than does the vigilance task environment. Such a difference in workload could affect the rate at which arousal changes over time on the task.

# STATEMENT OF THE PROBLEM

## Deficiencies of Past Investigations

As is evident from the previous discussion, driver fatigue studies have yielded widely varied, often conflicting results concerning the nature of performance, physiological correlates, and subjective rating changes over prolonged driving periods. A possible source for some of these inconsistencies lies in methodological differences and deficiencies of past investigations. An important goal of the present investigation was to identify major methodological considerations which could affect the ultimate results. The following discussion identifies four such methodological considerations and discusses means to eliminate, minimize, or specify their effects.

Inappropriate statistical methods. As discussed, a single reliable and valid measure of driver fatigue has not yet been identified. Consequently, investigators have typically employed multiple dependent variables, where each dependent variable is presumed to measure independent but parallel behaviors or symptoms of fatigue. For the most part, dependent variables in these studies have been treated as if obtained during independent experiments, using conventional univariate analyses, such as analysis of variance (ANOVA). Independent analyses of multiple dependent variables could result in erroneous experimental conclusions caused by: (a) inflation of per-experiment alpha, (b) reduction of experimental power, and (c) the loss of information regarding the interdependence of dependent variables (Finkelman, Wolf, and Friend, 1977).

Successive univariate analyses conducted on multiple dependent variables in the same experiment causes inflation of per-experiment alpha (i.e., the probability of Type I errors increases). If the dependent variables are independent (i.e., uncorrelated), alpha error is estimated by the equation: $\alpha = 1-(1-\alpha)^k$, where k is the number of tests performed. Where dependent variables are correlated, alpha error can be as high as (but cannot exceed) $\alpha k$. Because multiple dependent variables, especially those typically used in driver studies, are likely to be correlated (e.g., heart rate and heart rate variability, steering movements, and vehicle yaw excursions), the possibility of alpha error inflation cannot prudently be ignored.

The use of separate univariate ANOVAs can also result in reduction of experimental power (i.e., an increase in the probability of Type II error -- the probability that an existing experimental effect will not be detected). This can occur when an experimental effect exists, but is "spread out" over several dependent variables. As a consequence, it is possible that separate univariate tests will not detect statistical differences when an effect does, in fact, exist (Finkelman et al., 1977; Kerlinger and Pedhazur, 1973). In such cases, it is possible that multivariate analyses would detect effects where equivalent univariate analyses would not.

Because separate univariate analyses cannot assess the correlation among dependent variables, the result can be a loss of information regarding the interdependence of dependent variables. The loss of information, according to Finkelman et al., can result in serious errors in interpretation.

The above discussion indicates that the use of univariate statistics on multiple dependent measures could result in a peculiar situation wherein the probabilities of Type I and Type II errors are inflated simultaneously. The investigator is left with a potentially serious (but most often ignored) interpretive dilemma. Specifically, if the investigation results in a rejection of the null hypothesis (on one or more of the dependent variables), the result may be due to a true experimental effect, or may be a spurious effect of multiple testing and the resultant alpha error. On the other hand, if all tests yielded non-significance, it may be that no experimental effect existed, or that an effect, distributed over several dependent variables, was not (and could not be) detected. Until recently, the use of multivariate statistical analyses has been limited due to the large computational requirements demanded by the analyses and to the lack of adequate statistical tables for determining critical levels. The use of high speed computers and statistical software packages has made multivariate statistics both feasible and practical.

The present investigation employed multivariate analysis of variance (MANOVA) to analyze continuously monitored dependent variables obtained in extended duration driving runs. The analyses made extensive use of the Statistical Analysis System, a sophisticated statistical software system developed by Barr, Goodnight, Sall, and Helwig (1976).

Simple reaction time (RT) as a measure of fatigue decrement. Previous investigations which have attempted to assess the effects of fatigue on emergency performance have often employed "simple reaction time" measurements (e.g., Laurell and Lisper, 1976; Lisper, Laurell, and Stening, 1973; Mast, Jones and Heimstra, 1966). Such measures may not be generalizable to driving performance for a number of reasons. For one thing, emergency responses made during driving emergencies most often require the driver to attend to multiple inputs and choose from a large repetoire of possible responses. In contrast, simple reaction tasks involve a single response to the presence or absence of a single stimulus. The disparity in the complexity of these tasks introduces some doubt as to the validity of simple RT as an instrument for measuring fatigue performance decrement. A related argument against the use of simple RT measures is based on the theoretical nature of reaction time. Donders (1969), Sternberg (1969), and others have theorized that differences in response latency to so-called "simple", "select", and "choice reaction time" tasks are due to the involvement of differing cognitive stages or subprocesses in the decision making process. Given that such stages do exist, it is likely that prolonged driving does not have an equal effect on all cognitive stages. In fact, based on current arousal theories, driving is likely to have a greater detrimental effect on the decision making process than on motor processes. If so, it would be expected that tasks assumed to involve relatively few decisions, as in simple reaction time tasks, would be affected less by prolonged driving than would more complex tasks. This supposition is supported by the findings of some investigators who have found that simple reaction time measures often show little or no change as a function of prolonged driving (e.g. Lisper _et_ _al._, 1971; Lisper _et_ _al._, 1973).

The present investigation employed three types of "emergency" stimuli in three separate experiments. The first two experiments employed simulated emergencies to assess the effects of prolonged driving on performance in a realistic emergency driving scenario. Experiment I involved the simulation of an emergency in which the driver was required to respond to a lead vehicle stopping suddenly on the roadway. Experiment II required drivers to maneuver to maintain road position during a simulated sudden, high speed wind gust. Both experiments were presumed to involve complex stimuli and/or contingent responses typical of those in actual driving. It was hypothesized that performance of subjects in each of these experimental situations would deteriorate after prolonged driving periods. Experiment III employed a choice reaction time task similar to those found in reaction time studies (e.g., Sternberg, 1969; Theios, 1973). It was hypothesized that choice reaction time task (CRT) performance, which presumably requires complex cognitive as well as motor components, would exhibit degradation (increase in reaction time) with prolonged driving.

*The use of repeated response trials in the assessment of emergency performance.* Studies which have attempted to assess the effects of fatigue on reactions to unexpected driving situations typically require the subject to respond to an emergency stimulus which is presented repeatedly throughout the driving period. Lisper *et al.* (1973), for example, conducted a study which required subjects to respond to an emergency stimulus which occurred on the average of every 50 s during a three hour driving period (over 200 total stimulus presentations).

Although Laurell and Lisper (1976) suggested that successive trials can (presumably due to muscular fatigue) cause an increase in reaction time; other possibilities have not been examined. Two factors which could have important effects on investigations employing successive response trials, especially in those dealing with driver fatigue, are learning and arousal. In terms of learning processes, it seems likely that the stimulus-response pairing of a trial presents an opportunity for the driver to improve his or her performance of the task with each successive trial. Such an effect could negate any decrements (i.e., reduce reaction time) caused by prolonged driving (presumed to be fatigue) effects. A second possible effect of repeated trials is that each successive trial serves to stimulate or increase the arousal state of the driver. The resulting increase in arousal could inhibit or counteract fatigue effects, which could cause a reduction in reaction time relative to what might normally occur. If, indeed, learning and arousal are significant factors, the results of investigations using successive trials may not be valid indications of actual driving conditions. The exclusion of such considerations may explain why a number of investigators have shown minimal decrements or even improvements in emergency reaction time over extended driving periods (e.g., Ellingstad and Heimstra, 1970; Laurell and Lisper, 1976; Mast, et al., 1966). The present experiments also employed repeated response trials to assess the effects of extended duration driving on emergency performance. Unlike most other studies, however, response trials were presented after (rather than during) uninterrupted driving sessions. This technique allowed the opportunity to assess the effects of both driving time and repeated trials on performance, without interfering with time related fatigue processes. It was hypothesized that the performance level over repeated trials would improve with each succeeding trial with the first trial (which presumably reflects the greatest amount of fatigue) indicating the lowest level of performance.

Steering wheel movements as a measure of prolonged driving effects. As discussed earlier, frequency characteristics of steering wheel movements have been commonly measured in driver fatigue studies. One category of methods used for analyzing such characteristics basically attempts to identify discrete steering movements made by the driver, and then determines the frequency of such movements for a specified unit of time. The most common of these frequency measures is "steering wheel reversals." The measure has been widely used in driver research because it is relatively easy to measure, and poses no problem in terms of interference with the driving task. Unfortunately, as discussed earlier, steering wheel reversals have not always been found to be reliable measures of driver state changes over extended driving periods.

Although several factors could account for inconsistencies found among various studies, the biggest obstacle to reliable and ultimately valid steering reversal measures appears to be the lack of agreement on how to measure steering wheel reversals. According to McLean and Hoffman (1975), steering wheel reversals have most often been defined as ". . . the number of times the steering wheel crosses the zero angle position (or a small arc encompassing that position), or, alternatively, the number of times the direction of steering wheel movement is reversed through a definite angle" (p. 248). The zero crossing method, as it is called, has at least two major shortcomings: first, the method does not measure movements which do not involve zero crossings; and secondly, the measure may contain a considerable amount of "noise", due mainly to the fact that the method counts small movements (about the center position) which may have little or no relevance to the driver's state or the driving task.

A more frequently used method which will be referred to as the "fixed gap method," is one which uses instrumentation to count the number of times the steering wheel is turned past selected limits on either side of the center steering wheel position (0 deg). With the appropriate gap limits, the noise problem encountered using the zero crossing method may be reduced or eliminated. The method, however, is not without disadvantages. The biggest (and most often ignored) problem with fixed gap methods is one of selecting an "optimal" gap size, which usually involves the placement of limit switches along the steering wheel arc or setting of a threshold detector/counter on the analog steering signal. According to McLean and Hoffman (1975), most investigators have depended on arbitrary judgments or trial-and-error testing to determine "optimal" gap sizes. The lack of a standard steering reversal criterion is reflected by the wide disparity of gap sizes used in the fatigue literature. Gaps range from 2 deg (e.g., Sussman and Morris, 1970) to 12 deg (e.g., Huntley and Centybear, 1974). This lack of consensus on an optimal gap could affect the ultimate reliability and/or validity of the measure. If, for example, a gap size is chosen which is too large, small steering wheel movements which may be of interest will not be detected. Similarly, if gap limits are made too small, the measure cannot distinguish small steering movements from larger ones.

McLean and Hoffman (1975) developed a technique which overcomes many of the problems described above. Essentially the method employs computational hardware and software which classifies all steering wheel movements within a broad range, rather than those of restricted magnitude or frequency. This was accomplished by digitizing steering wheel data, and detecting changes in direction of steering movements. Steering wheel reversals were then classified in a range from 0.1 to 10.0 deg. In a study of "sight-distance" and lane width conditions, McLean and Hoffman found that the steering reversal rates were most reliable as measures of driver performance (using test-retest reliability correlation) at "gap" sizes of 0.5 to 0.7 deg.

The present investigation employed techniques similar to those used by McLean and Hoffman to examine the effects of prolonged driving on steering wheel movements. It was hypothesized that as the driving period is prolonged, the driver tolerates larger vehicle excursions than when he is alert, and that accompanying the driver's increased tolerance for lane drift would be associated changes in the driver's overall steering behavior.

## Primary Goal and Rationale

Problems of interpreting fatigue studies. Inconsistencies notwithstanding, the concensus of investigations presented thus far has shown that certain performance, physiological, and subjective rating measures can show changes in drivers' apparent states of arousal over driving time. Even where statistically reliable changes of such variables are found, however, a number of questions arise concerning how such data should be interpreted.

Lisper, Laurell, and Stening (1973) have suggested that such data should be viewed with caution in view of the apparent flaws in the methods used in selecting variables and interpreting associated data. According to Lisper et al., the general approach of investigations has been to base "the decision to accept or reject a specific measure . . . on the trend over driving time. If there is deterioration, the measure is accepted as valid and if there is none the measure is rejected" (p. 501). Using such an approach, Lisper et al. have asked, "How is it possible . . . to avoid reaching any conclusion other than that driving is fatiguing?" (p. 501).

A similar problem exists in the use of definitions and hypotheses typically used in fatigue investigations. For example, a frequently stated hypothesis proposes that "fatigue" is a process involving changes in central nervous system mechanisms which cause changes to occur in performance and/or physiological variables. The argument becomes circular, however, when it is concluded that a driver exhibiting certain performance and/or physiological changes is said to be "fatigued."

Aside from these problems, other questions arise concerning the interpretation of physiological, performance, or subjective data. For example, how much change in a variable or set of variables constitutes a "fatigued" driver state? If observed changes reflect changes in arousal mechanisms, to what extent is the "normal" driving ability of the driver impaired when such symptoms are exhibited? If changes in certain variables can be associated with the impairment of performance, how much change constitutes a dangerous level of performance degradation in the driver? From the above discussion it appears that the most significant problem of driver fatigue research has continued to be the lack of valid criteria for the definition of fatigue. In view of these difficulties, it appears that alternative approaches should be examined.

Primary goal. The primary goal of the present investigation was to develop a means to anticipate and ultimately predict degradations in driving performance associated with prolonged driving.

The means for predicting these "fatigue decrements" are based on the validity of the following hypotheses:

(1) During prolonged driving periods, the ability of the driver to respond to unanticipated driving events degrades over time;

(2) During prolonged driving periods, certain variables which reflect the driver's state of alertness show significant changes over time; and

(3) During prolonged driving periods, changes in the driver's emergency performance capability (1) are related to changes in variables reflecting drivers' states of alertness (2). Such a relationship could be used to predict decrements in emergency performance from changes in alertness variables.

A number of driver fatigue studies discussed earlier have demonstrated that certain variables (e.g., steering wheel reversals, lateral deviation, heart rate variability) measured continuously or periodically throughout the driving period exhibit reliable changes over time. These changes are most often attributed to changes in arousal mechanisms of drivers' central nervous systems. Similarly, investigations have shown that the driver's ability to respond to sudden events diminishes over extended driving periods (e.g., Lisper et al., 1971; Lisper et al., 1973; Sussman et al., 1971). As with continuously monitored variables, the performance degradation in emergencies is often attributed to changes in the driver's central nervous system. If, in fact, both classes of behaviors are modified by common central nervous system mechanisms, it is reasonable to hypothesize that changes in these variables are correlated. Assuming that driving time is the major, controllable determinant of fatigue, it was hypothesized that changes in selected continuously measured variables would be correlated with changes in emergency performance over extended driving periods.

The present investigation employed three experiments to determine the effects of prolonged, monotonous driving periods (of 30, 60, or 150 min) on a wide variety of driving behaviors which include drivers' responses to simulated emergencies. The three experiments employed were similar in terms of experimental protocol, but differed in the nature of the emergency stimuli presented to subjects and in the responses required of them.

Experiments I and II employed emergencies which could be encountered in normal driving - differing mainly in the major sensory modalities through which emergencies are perceived. In particular, Experiment I employed a visually presented emergency, while Experiment II employed an emergency with visual, vestibular, and kinesthetic cues. Experiment III employed an "emergency" involving a choice reaction task. Such a task was employed to determine the effects of prolonged driving on certain classes of decision making processes.

A secondary goal of the present investigation was to determine the interactive effects of steering gain, emergency performance, and extended duration driving. In 1975, 89.9% of all domestic passenger vehicles and 76.6% of all small trucks (10 000 GVW and less) sold in the United States were equipped with power steering systems (Motor Vehicle Manufacturers Association, 1976). The popularity of power steering has been a primary reason for continued efforts in determining optimal design parameters for such systems. Of these parameters, one of obvious importance is steering gain (or steering ratio). Although a number of investigators have studied the effects of steering gain on various aspects of driver performance (e.g., Hoffmann and Joubert, 1966; Olson, 1970; Repa, Alexandridis, Howell, and Wierwille, 1978), none appear to have addressed the possible interactive effects of steering gain and driving induced fatigue.

The second experiment (Experiment II) of the present investigation, in addition to the stated primary goal, was designed to determine the nature and extent of interactions between steering gain (normal and high) and driving duration on drivers' responses to a vehicular emergency (simulated, large amplitude, sudden onset wind disturbances). It was hypothesized that alert drivers in the high steering gain group would perform better than alert drivers in the low gain group. Such a difference was anticipated due to the advantage in steering response "rise-time" provided by the high steering gain over the lower steering gain. It was further expected that the performance advantage of high steering gain over low steering gain would reverse after prolonged driving, so that after prolonged driving, drivers (presumed to be "fatigued") in the low steering gain group would perform better than drivers (presumed to be equally fatigued) in the high steering gain group. The expectation for a performance reversal was based on the following premises: (a) the latency to respond to an emergency increases equally with prolonged driving in normal and high steering groups; (b) the driver responds with large amplitude, initially "open-loop" steering response (i.e., response is made without vehicular response feedback); and (c) the high steering gain level is higher than is optimal for emergency steering inputs - resulting in steering overcompensation.

Table 2 presents a summary of the primary and secondary goals of the present investigation.

**TABLE 2.** Summary of Investigative Goals and Associated Hypotheses

I. Primary Goal: Develop a method to predict degradations in driver emergency performance from changes in other variables.

Associated Hypotheses

1. During prolonged driving periods, the ability of drivers to respond to unanticipated driving events degrades over time.
2. During prolonged driving periods, certain variables (which presumably reflect the state of the driver alertness) exhibit significant changes over time.
3. During prolonged driving periods, degradations in the driver's ability to respond to unanticipated driving events can be predicted from changes in alertness variables.

II. Secondary Goals

A. Identification and avoidance of possible experimental contaminants:

1. Inappropriate statistical methods
2. Simple reaction time measures
3. Repetitive response trials
4. Inadequate methods for measuring steering movements

B. Determine the interactive effects of steering gain, emergency performance, and extended driving durations.

# GENERAL METHOD

## Subjects

The subjects were paid volunteers ranging in age from 18 to 31. All were required to have a valid drivers license and at least 10,000 miles or two years of driving experience. All subjects had 20/25 visual acuity or better (with correction where needed) and normal color vision, measured with a laboratory vision tester (Bausch and Lomb Ortho-Rater). Subjects were paid by the hour, and received a small bonus payment for completing all trials.

## Driving Simulator

The primary apparatus used in the present investigation was the driving simulator facility located in the Human Factors Laboratory at Virginia Polytechnic Institute and State University. Descriptions of the simulator have been given by Wierwille (1975) and McLane and Wierwille (1975). The following describes the simulator's four major systems: the roadway imaging/display system, the roadway object imaging system, the motion platform, and the audio system.

Roadway imaging/display system. The roadway imaging and display system located on the system's "video bench" (Figure 1) was responsible for the generation and display of the roadway image on which the operator "drives" the vehicle. The image was generated electronically by the "roadway image generator" and projected onto a cathode ray tube display (Tektronix 604). A closed-circuit television (CCTV) camera was used to convert the cathode ray tube image to a video signal which was routed to the driver's platform display and experimental

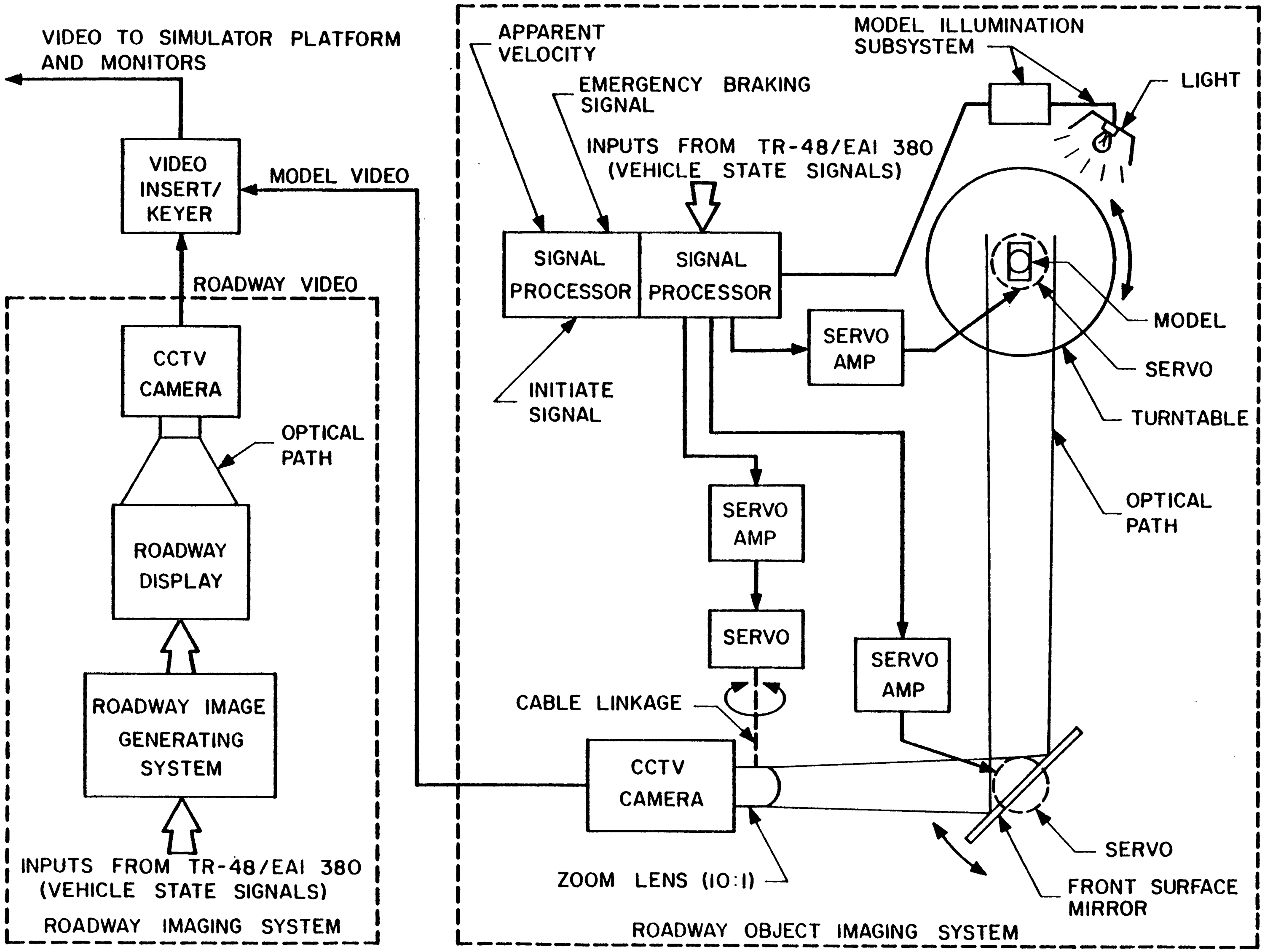

Figure 1. Roadway object/imaging system.

monitors. The driver's roadway image was displayed by a television display mounted behind a Fresnel lens. Figure 2 shows the driver's platform display and the roadway image.

Roadway object imaging system (video bench). The roadway object imaging system (Figure 1) of the simulator's video bench was used to superimpose an image of an automobile onto the roadway image. To accomplish this, a CCTV camera (Dage RGS-50) was used to produce a video image of an automobile scale model (scale 1:72). The video image was combined electronically with the roadway image so that the roadway image was blanked by a high-speed video insert keying device where the automobile image appeared.

To produce realistic orientation and distance cues, the model image was controlled in three dimensions by the following system elements: (a) apparent distance of the model image--controlled by a servo actuated zoom lens (10:1 zoom ratio) attached to the CCTV camera; (b) lateral position--controlled by a servo actuated mirror which reflected the optical image approximately 90 deg to the camera lens; and (c) horizontal rotation orientation--controlled by a servo actuated turntable on which the scale model was positioned. All three servo systems were controlled simultaneously with two analog signal processors (Comdyna Systems, model 808) through associated servo amplifiers. The signal processors received signals (representing vehicular motion) from the simulator's analog/hybrid computer (EAI 380) so that proper "on-the-road" orientation was maintained with respect to the driver's vantage point, despite the driver's control actions.

In the present investigation, the imaging system was used to simulate a car following scenario, in which the rear of an automobile scale model was continu-

Figure 2. Driver's platform display/dashboard and roadway image.

ously displayed on the right lane of the roadway. The lead vehicle's apparent distance, determined by the camera's zoom lens position, was dependent on the apparent speed of the simulator with respect to a preset velocity threshold. When the simulated vehicle velocity was greater than 89 km/h, the apparent distance of the model image decreased; when less than 89 km/h, the distance increased. At maximum apparent distance, the car image subtended approximately 1.6 deg of horizontal arc (from the driver's vantage point), which was equivalent to a distance of approximately 70 m. At minimum apparent distance, the image subtended approximately 15.8 deg of horizontal arc, equivalent to a distance of approximately 7 m.

Motion platform. A second major system of the driving simulator was the motion platform, which simulated vehicular motions of yaw, roll, lateral translation, and longitudinal translation. The motion platform was actuated by hydraulic servos which were controlled by signals from an analog computer (EAI TR-48). Hydraulic power to the servos was supplied by a high volume hydraulic power supply which operated at a pressure of $1.03 \times 10^7$ Pa.

Vehicular dynamics were simulated in the following manner. Driver control signals (representing steering wheel, accelerator, and brake positions) were inputted to the system's main analog computer. The computer, programmed to simulate vehicular equations of motion, supplied control voltages to the roadway imaging/display system, the platform servo valves, the simulator's speedometer, the video bench, and the audio system. (Feedback to the computer was supplied by linear potentiometers in various systems of the simulator.) The "normal" vehicle parameters used in Experiments I, II (normal steering gain group), and III were representative of a late model intermediate American-made sedan. (Refer to Appendix II for relevant simulator parameters.)

Audio system. To enhance simulator realism and driver feedback cues (McLane and Wierwille, 1975), four channels of sound were used to simulate driving sounds: wind/road noise, velocity-dependent engine and drive-train noise, tire screech with severe braking, and tire squeal with large steering inputs at highway speeds. Wind and road noises were recorded on audio tape and played back through speakers affixed to the motion platform. A velocity-dependent engine/drive-train sound was produced by rotating a permanent magnet proximate to a magnetic sensor. The sensor output was amplified and played through a platform loudspeaker. The rotation of the magnet was accomplished by a DC motor, whose rate of rotation was controlled by a computer output signal representing velocity.

### Body Sensors

Two devices were used in conjunction with the simulator to measure the driver's upper body movements and heart pulse.

Upper body movement detector. Movements of the driver's upper torso were sensed by linear potentiometers placed in the back cushion of the driver's seat. The lower potentiometer was positioned in the center of the seat, approximately 15 cm above the bottom of the seat back (approximately at the driver's lumbar region). The upper potentiometer was positioned approximately 46 cm above the bottom of the seat back. The inputs to the potentiometers were reference voltages of 10 VDC. The outputs of both potentiometers were filtered and routed to a comparator referenced to a selected threshold. Compression or extension of either potentiometer (through the seat back) resulted in a short duration square wave pulse which was recorded by the FM recorder.

Pulse monitor. The driver's pulse was monitored by an ear plethysmograph connected to a heart monitor (Hewlett-Packard 7807C). The plethysmograph lead was affixed to a headband which acted as a strain relief while providing sufficient lead slack for the subject to move freely in the seat without restriction.

## Procedure

The three experiments of the present investigation were scheduled with driving trials conducted on two consecutive days (see Figures 3 and 4). Experiment III was conducted in a manner similar to Experiments I and II except that two additional runs were designated for Day 1 of the experiment (Figure 4).

Day 1. After reading a set of instructions (Appendix IIIA-IIID), and signing a consent form (Appendix V), the subject was seated in the simulator and fitted with an ear plethysmograph. The subject was then given a briefing during which the simulator's safety features were described, and any remaining questions were answered.

The first run was a practice run which allowed the subject to become accustomed to the simulator's control and handling characteristics. At the beginning of the practice run, the subject was instructed to bring the simulator up to speed and maintain a normal position in the right lane of the roadway. At various times during the practice run, the subject was instructed to perform various maneuvers which included steering the vehicle into the opposite lane of the roadway applying the brakes quickly and re-accelerating to normal speed. When the subject's proficiency was judged to be adequate, the practice run was ended. Practice runs had an average duration of approximately three minutes and

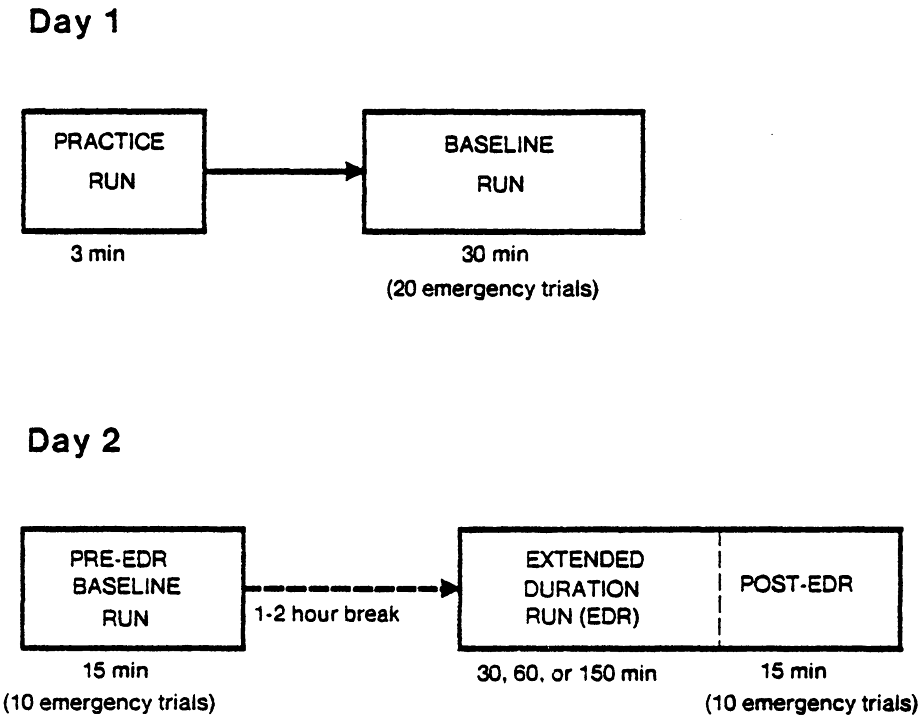

Figure 3. Timeline for Experiments I and II.

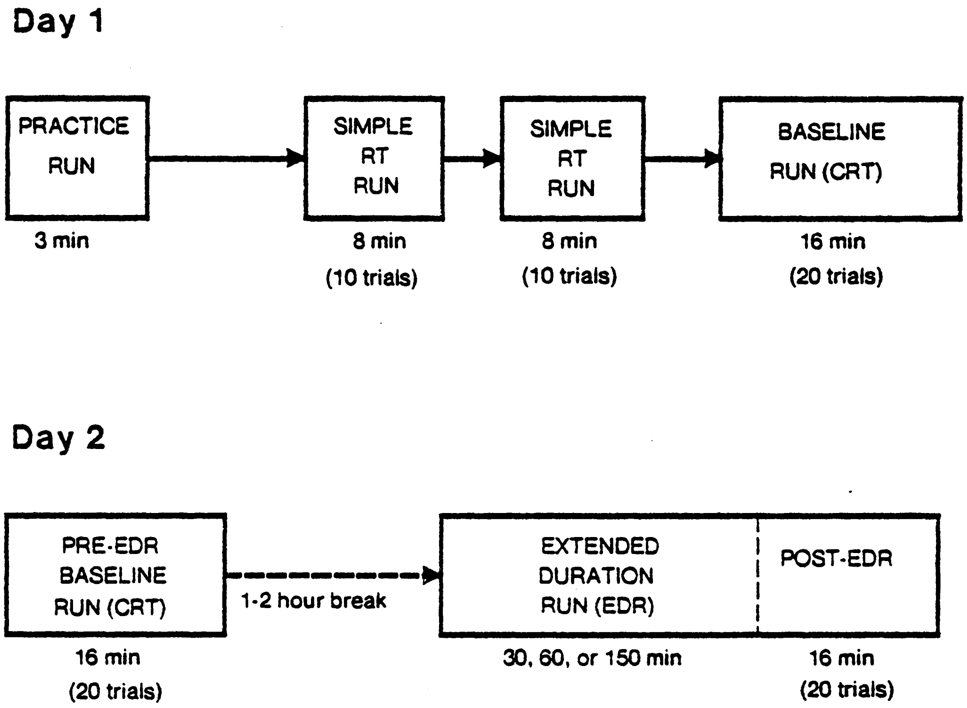

**Figure 4. Timeline for Experiment III.**

were followed by a three minute rest break. No data were recorded during the practice run.

The practice run was followed by a "baseline" run. This run was used to establish initial baseline data for emergency response performance. During this segment, the subject was required to operate the simulator and respond to "emergency" trials which were presented at random intervals throughout the run. The number and types of emergencies differed for the three experiments, as will be described later.

Upon leaving the facility, subjects were given instructions requesting that they receive a normal amount of sleep and refrain from taking stimulants (including drinks containing caffeine) before the Day 2 runs.

Day 2. The second day events occurred in the following order (Figures 3 and 4): (a) pre-extended duration run (pre-EDR), (b) rest period (away from the experimental facility), (c) extended duration run (EDR), and (d) post-EDR run.

Upon arrival for the pre-EDR run, the subject was first briefed, seated in the simulator, and fitted with an ear plethysmograph. After the room lights were turned off, the simulator was activated, and the run was conducted in a manner identical to the Day 1 baseline run.

After the run, the subject was asked to leave the experimental facility and to return approximately two hours later. Subjects were reminded not to take stimulants of any kind during the interim.

For extended duration runs (EDR), subjects were again briefed, seated in the simulator, and fitted with the ear plethysmograph. As before, each subject was instructed to drive in the right lane and remain alert to possible emergency situations which may occur at anytime during the driving session.

During the EDR, each subject was required to drive continuously for a period of 30, 60, or 150 minutes. Driving time was determined by random assignment of the subject to a given duration condition. No emergencies were presented during this period. To minimize the possibility of anticipatory responses, each subject, regardless of the assigned driving duration, was informed only that the driving period could last as long as three hours. In addition, any time cues, including subjects' watches and facility time bells were removed or disabled before the run was started. During the EDR, no communication was allowed between the subject and the experimenter. Subjects were observed continually during the run for obvious signs of inattention or falling asleep.

During post-EDRs (which followed EDR without interruption), subjects were required to operate the simulator and respond to emergencies presented at random intervals throughout the run. During the transition from EDR to the post-EDR, the simulator was kept in continuous operation, and drivers were not alerted to any change. At the end of the EDR, the experimenter reset the run timer and recorded event markers on the FM data record (indicating the end of the EDR and the beginning of the post-EDR). These markers clearly separated the runs for subsequent data analysis.

At the end of the experiment, the subject was escorted from the simulator and asked to fill out a questionnaire (Appendix IV). Upon completion of the questionnaire, the experimenter paid the subject and answered any questions the subject had concerning the experiment or the apparatus.

## Data Collection and Initial Data Reduction

Data were recorded on 11 channels of an analog FM tape recorder

(Honeywell, Model 5600E). Analog signals (ranging from -2.5 VDC to +2.5 VDC) were continuously recorded during all baseline, extended duration, and emergency trial runs. The recorded signals represented the following variables or events:

1. Simulated vehicle speed,
2. Steering wheel position (angular displacement),
3. Vehicle yaw angle,
4. Vehicle lateral position,
5. Accelerator pedal position,
6. Car following distance,
7. Heart beat (square pulse output of heart monitor),
8. Upper torso movements,
9. Emergency initiate,
10. Brake activation, and
11. Driving period marker.

Simulated vehicle speed, accelerator pedal position, and car following distance were used to assess subjects' general adherence to instructions (i.e., speed limits and following distance), and were not formally employed as dependent variables in any analyses.

The analog recordings were played back into a digital computer (Digital Equipment Corporation PDP 11/55) through a Laboratory Peripheral System (DEC, LPS-11) which accomplished analog-to-digital conversion and real-time clocking. The data were analyzed by means of several software packages developed to perform real-time analyses of the data at a data sampling and analysis rate of 100 Hz.

The separation of driving periods was determined by the "driving period marker" recorded on one analog tape channel. The presence or absence of the signal was used to direct data sampling by the computer. In addition, a tape footage counter and voice markers were used to ensure that the recorder's recorded tape segment corresponded to a given subject during a given driving run.

For most variables, raw analog data were initially reduced by the laboratory computer which executed the following steps: (a) data sampling (during presence of a driving period marker); (b) scaling of digitized data in appropriate units (voltage to equivalent quantity in deg, cm, etc.); (c) performance of intermediate calculations (e.g., sums, sums of squares); (d) transfer of intermediate calculations to disk storage at the end of a driving period; and (e) final calculation and printout of reduced data (accomplished after termination of data sampling). The dependent measures computed from analog recordings were separated into two categories, emergency run variables and EDR variables. The emergency run variables, which will be described later, were used to assess driver performance in emergencies of the three experiments. EDR variables were used for the continuous assessment of driver behavior during EDRs. These variables, common to all three experiments, are described below.

Means/standard deviations. Means and standard deviations were calculated for steering wheel position, vehicle lateral position, vehicle yaw angle, simulated vehicle speed, and accelerator pedal position. These data (scaled in appropriate units) were outputted for every minute and every 10 minutes of driving time.

Steering wheel movements. Displacement of the steering wheel from the straight ahead (0 deg) position was expressed in terms of (a) steering wheel position mean and standard deviation, and (b) number of steering reversals by magnitude for every minute and 10 minutes of driving time.

Steering reversals were defined as the movement between steering inflection points (i.e., points where the steering motion changed direction). Computer and related software were used to detect and classify all steering wheel movements occurring during the driving period (in real-time) by executing the following steps: (a) calculation of the first derivative for each digitized sampling point; (b) detection of a change in sign of the derivative (indicating a reversal in direction of steering motion); (c) calculation of the magnitude difference between the current reversal point and the previous one (calculates steering reversal magnitude); (d) storage of the computed reversal magnitude; (e) categorization of each reversal according to magnitude classification; and (f) computation of reversal means and standard deviations.

The algorithm used for computing the first derivative for each sampling point was based on the "convolution integer" method described by Evans and Gutmann (1978). (The technique described by Evans and Gutmann was originally designed for the analysis of eye movement data.) The algorithm estimates a derivative (while minimizing noise effects) by performing a least squares fit of a polynomial to a sliding window of data points.

Steering wheel reversal data were expressed in terms of mean steering reversal magnitude and frequency distribution of reversals by magnitude (in degrees). Reversal means were outputted for every minute and 10 minutes of continuous data. Frequency distributions were outputted for every five minutes of continuous data. Dependent variables derived from these analyses were: (a) mean steering reversal amplitude, (b) number of steering reversals in the range of 0.5 deg to 2.0 deg (inclusive), and (c) number of steering reversals greater than 2.0 deg.

Yaw angle reversals. Vehicle yaw angle reversals were calculated and outputted in a manner identical (except for scaling factors) to steering reversal data. The dependent variable derived from the output was the number of yaw reversals greater than 2.0 deg.

Upper torso movements. The leading edge of each square pulse (which signified a change in upper body position) was detected and counted for every minute and 10 minutes of driving time.

Heart beat. Heart beats were analyzed by computer to obtain heart beat interval standard deviation, a measure of heart arrhythmia. An analysis routine was developed to analyze recorded square wave pulses from the heart monitor. (These square waves marked the presence of arterial pressure pulses caused by cardiac ventricular contractions. The interval between the leading edge of these square waves corresponds to the heart inter-beat interval, Venables and Martin, 1971). The computer algorithm accomplished the following: a) computed the interval between leading edges of consecutive square pulses; b) computed and stored the sum and sum of squares of the beat-to-beat interval; c) computed interval means and standard deviations; and d) printed reduced data for every minute and 10 minutes of driving time.

Summary of EDR variables. Nine dependent variables were selected as EDR variables, variables employed for assessment of driver behavior during extended duration runs. The final choice of variables was based on the results of a pilot study and general driver fatigue literature.

1. BDYMOVES -- number of upper torso movements per 10 min.
2. HRTINTSD -- heart beat interval standard deviation.
3. LATDEVSD -- vehicle lateral position standard deviation (cm).

4. STEERSD -- steering wheel standard deviation (deg).
5. STREVLRG -- number of large ( >2.0 deg) steering reversals per 10 min.
6. STREVSML -- number of small (0.5 deg to 2.0 deg) steering reversals per 10 min.
7. STRREVMN -- mean steering reversal amplitude (deg).
8. YAWDEVSD -- vehicle yaw angle standard deviation (deg).
9. YAWREVII -- number of yaw reversals > 2 deg.

## EXPERIMENT-SPECIFIC METHODS

### Experiment I

Subjects. Twelve subjects, six males and six females, were randomly assigned to one of three duration conditions. The number of males and females per duration condition was equal. One subject was replaced when he showed obvious signs of falling asleep during the EDR.

Apparatus. Experiment I used a simulated "car following emergency" in which the simulated lead vehicle appeared to stop suddenly in the roadway. The major apparatus for the simulation of this emergency scenario was the video bench described earlier. The experimenter initiated an emergency by closing a switch which resulted in: (a) the recording of an event marker on one channel of the FM recording and (b) the coupling of a constantly increasing (ramp) voltage to the zoom servo controller of the roadway object imaging system (Figure 1). The ramp voltage caused the zoom servo to rotate the zoom lens of the camera so that the apparent distance of the lead vehicle decreased rapidly.

Procedure for Experiment I emergency runs. At the beginning of emergency trial runs (Day 1 baseline, pre-EDR, and post-EDR), the subject was instructed to accelerate to, and maintain, a nominal speed of 55 miles per hour (89 km/h). As per written instructions (Appendix IIIB), subjects were instructed to maintain vehicle position in the right lane of the roadway at a safe distance from the vehicle ahead.

During the pre- and post-EDRs, which were each 15 minutes in duration, 10 car-following emergencies were initiated at random intervals ranging from 20 to

180 s. Upon initiation of the emergency, the apparent distance between the lead vehicle and the subject's vehicle appeared to diminish rapidly. The subject was required to step on the brake pedal as soon as the emergency state was detected. After braking, the lead vehicle distance increased to normal, at which time the subject was required to resume normal speed and following distance.

### Experiment II

Subjects. Twenty-four subjects, 12 males and 12 females, were assigned randomly to either of two steering gain groups ("normal" and "high" gain). Within each group, males and females were assigned in equal numbers to the three duration conditions. One male subject was replaced after exhibiting obvious signs of falling asleep.

Apparatus. Emergencies for Experiment II were simulated, high velocity cross-wind gusts. These gusts were simulated by inserting a DC voltage (switched by the experimenter) into the gust input circuit of the simulator dynamics. Actuating the wind gust switch caused the simultaneous displacement of the yaw, lateral, and roll actuators of the motion platform and the roadway display. The resulting actions were similar to those which can be experienced in a standard intermediate size U.S. sedan with cross-wind gust forces equal to approximately 890N (200 lb.). The duration of each wind gust was approximately 3 s (automatically controlled by a preset timing circuit).

Two levels of steering gain were used in Experiment II. These steering gains resulted in yaw rate gains of .15 deg/s per degree displacement of the steering wheel for "normal" steering gain, and .35 deg/s per degree displacement of the steering wheel for the "high" gain group. (Normal steering gain settings were also used in Experiments I and III.) The normal steering gain would correspond to an American intermediate sedan, while the high gain would be higher than typically found in production automobiles.

Procedure for Experiment II emergency runs. During emergency runs (Day 1 baseline, pre- and post-EDRs), subjects were required (per instructions, see Appendix IIIc) to "drive" the simulator in the right lane of the displayed roadway at a speed of 55 mph (89 km/h). During the run, which lasted 15 min, wind gusts were initiated at random inter-trial intervals ranging from 20 to 180 s. Upon detection of the wind gust, the subject was required to steer in the direction opposite the wind gust to maintain the vehicle's position in the right lane.

### Experiment III

Subjects. Twelve subjects, six males and six females were randomly assigned, in equal numbers, to one of three duration conditions.

Apparatus. Two brackets (3.4 x 6.9 cm) were mounted on the dashboard of the simulator. Vertically aligned in each of the brackets was a red (top) and green (bottom) light. The fixtures were placed so that the left and right brackets were approximately 9 deg to either side of the driver's straight-ahead visual axis. The selection and activation of the lights were accomplished with toggle switches located at the experimenter's station. Actuation of switches also placed event markers on FM tape which signified stimulus onset.

Procedure for Experiment III emergency runs. During emergency trials, subjects were required to monitor the activation of dashboard mounted lights, and respond to the lights as instructed (Appendix IIID). Subjects were required to respond to given light combinations with the responses shown in Table 3.

**Table 3.** Signal Light Configurations and Required Reponses for Experiment III

| Left Fixture | Right Fixture | Required Response |
|---|---|---|
| Green | Red | Steer into left lane |
| Red | Green | Steer into right lane |
| Red | Red | Step on brake pedal |

Day 1 runs consisted of a practice run, two simple reaction time runs, and a baseline choice reaction time run. Day 2 runs consisted of an EDR, pre- and post-EDRs (Figure 4).

For simple reaction time runs (Day 1 only), subjects were informed by verbal instructions that only one type of response was required during that set of trials. During the trials (although appropriate light combinations were presented), subjects were instructed to respond to the onset of lights rather than to the particular light combination. For example, during simple brake reaction time trials, subjects were informed that the only response required was to step on the brake pedal at the onset of the lights (only red lights were presented during these trials). Each of the simple reaction time runs, which were 8 min in duration, contained 10 reaction time trials, presented at random intervals ranging from 20 to 90 s.

Choice reaction time runs were conducted during Day 1 trials following simple reaction time trials. All subsequent emergency trials for Experiment III employed CRT trials (See Figure 4). During CRT emergency trials, subjects were required to respond according to the light combinations with required responses shown above. Subjects were informed before the beginning of the run that either lane change or brake responses could be required during any given trial. During the 16 min CRT emergency runs, 10 braking and 10 lane change trials were presented. Brake and lane change trials were randomly interspersed throughout the run at random intervals ranging from 20 to 90 s.

## Emergency Run Dependent Variables

Experiment I - Brake response time. The emergency run dependent variable for Experiment I was brake response time. Brake response time was defined as the time elapsed between the activation of the emergency initiate switch (switch activating the zoom servo) and the depression of the brake pedal (signalled by a brake actuated switch).

Experiment II - Mean lateral deviation error. Performance of the driver's recovery from a sudden cross-wind gust was measured by the average deviation from the mean lateral position immediately prior to the wind gust.

During emergency run analyses, successive lateral position data points were summed continously on a sliding window of 6 s. Upon detection of the wind gust (indicated by the presence of an event marker), the mean lateral position was computed for the 6-s window by dividing the sum of the lateral position data points by the number of sampling points (600) in the period. A mean lateral error score was computed for the 2-s period immediately following initiation of the wind gust by (1) summing the absolute value of the deviations from the just computed mean lateral position score, and (2) dividing the sum of the absolute deviations by the number of sampling points in the period (200).

Experiment III - Choice response times (braking and steering responses). Two types of responses were measured for Experiment III, brake response time and lane change response time. Brake response time was defined as the time elapsed between the onset of the "command" to brake (two red lights) and a brake pedal response. Lane change response time was defined as the elapsed time between the onset of lane change lights (one light red, the other green) and a steering wheel movement greater than or equal to 10 deg.

# RESULTS

## EDR Variables

Multivariate analyses. Because extended duration runs (EDRs) for the three experiments (except Experiment II, high gain group) were identical in terms of run formats, schedules, dependent variables, etc., it was anticipated that the EDR data would not differ significantly among the three experiments. It was further anticipated that the type of emergency presented to each subject would cause no differential effects on EDR data among experimental groups. To test for experimental group differences for effects due to time, the EDR data for Experiments I, II (normal steering gain), and III were combined (with corresponding dependent variables), and analyzed in three two-way multivariate analysis of variance (MANOVA) tests (one for each duration condition). Each analysis included the following factors: experiment (I, II (normal steering gain), and III), time (successive 10-min observations), and subjects, nested within experiment. Tests of significance for time and experiment x time were conducted using Roy's criterion (Harris, 1975). Tests of significance for experiment main effects were conducted using Wilk's criterion rather than Roy's criterion due to the inadequacy of available tables for the specified degrees of freedom for the latter. All multivariate analyses were calculated by the Statistical Analysis System (Barr, Goodnight, Sall, and Helwig, 1976).

The MANOVA for the short duration group, shown in Table 4, revealed no significant main effects or interactions ($p > 0.05$). MANOVAs for the medium (Table 5) and long duration groups (Table 6) revealed significant effects due to

**TABLE 4.** MANOVA Summary for EDR Variables Across Experiments I, II (Normal gain), and III - Short Duration

(Wilk's Criterion)

| SOURCE | p | $\nu_H$ | $\nu_E$ | U-statistic [a] |
|---|---|---|---|---|
| Experiment (E) | 9 | 2 | 9 | .270459 |
| Subjects/E | (Error Term for E) | | | |

(Roy's Union Intersection Test)

| SOURCE | s | m | n | $ch_{max}(H\,E^{-1})$ | $\theta^b$ |
|---|---|---|---|---|---|
| Time (T) | 2 | 3.0 | 4.0 | 1.9393 | .6598 |
| E x T | 4 | 2.0 | 4.0 | 1.6940 | .6288 |
| T x S/E | (Error term for T, E x T) | | | | |

Notes a and b are presented following Table 6.

**TABLE 5.** MANOVA Summary for EDR Variables Across Experiments I, II (Normal gain), and III - Medium Duration.

(Wilk's Criterion)

| SOURCE | p | $\nu_H$ | $\nu_E$ | U-Statistic[a] |
|---|---|---|---|---|
| Experiment (E) | 9 | 2 | 9 | .000201 |
| Subjects/E (S/E) | (Error Term for E) | | | |

(Roy's Union Intersection Test)

| SOURCE | s | m | n | $ch_{max}(H E^{-1})$ | $\theta^b$ |
|---|---|---|---|---|---|
| Time (T) | 5 | 1.5 | 17.5 | 3.4490 | .7752** |
| E x T | 9 | 0.0 | 17.5 | 1.4507 | .5920 |
| T x S/E | (Error Term for T, E x T) | | | | |

**$p$ < .01

Notes a and b are presented following Table 6.

**TABLE 6.** MANOVA Summary for EDR Variables Across Experiments I, II (Normal gain), and III - Long Duration

(Wilk's Criterion)

| SOURCE | p | $\nu_H$ | $\nu_E$ | U-Statistic[a] |
|---|---|---|---|---|
| Experiment (E) | 9 | 2 | 9 | .000017 |
| Subjects/E (S/E) | (Error term for E) | | | |

(Roy's Union Intersection Test)

| SOURCE | s | m | n | $ch_{max}(H E^{-1})$ | $\theta$[b] |
|---|---|---|---|---|---|
| Time (T) | 9 | 2.0 | 58.0 | 2.3134 | .6982** |
| E x T | 9 | 9.0 | 58.0 | .6561 | .3962 |
| T x S/E | (Error term for T, E x T) | | | | |

** $p < .01$

Notes a and b are presented on the following page.

Notes for Tables 4, 5, and 6

a. $p$ = number of dependent variables

$\nu_H$ = degrees of freedom for treatment

$\nu_E$ = degrees of freedom for error

U - statistic (likehood ratio) = $\dfrac{\underline{E}}{\underline{E} + \underline{H}}$

Where H = Sum of squares and cross-products matrix for treatment

E = Sum of squares and cross-products matrix for error

b. $s = \min(\nu_H, p)$

$$m = \frac{\nu_H - p - 1}{2}$$

$$n = \frac{\nu_E - p - 1}{2}$$

$p$, $\nu_H$, and $\nu_E$ are defined as above.

$$\theta = \frac{ch_{max}(\underline{H}\,\underline{E}^{-1})}{1 + ch_{max}(\underline{H}\,\underline{E}^{-1})}$$

where $ch_{max}(\underline{H}\,\underline{E}^{-1})$ = largest charactistic root of the matrix time ($p < 0.01$), and no effect due to experiment or the experiment x time interaction ($p > 0.05$).

The significant MANOVA main effects due to time indicate that among the array of dependent variable means, there was at least one dependent variable (or combination of dependent variables) which exhibited significant differences among time period observations.

Univariate analyses. Employing methods proposed by Finkelman *et al.* (1977), separate ANOVAs were conducted on each dependent variable to determine the impact of the significant independent variables (time) on each of the dependent variables. Each ANOVA included the same sources of variance used in the overall MANOVAs. Summary results for ANOVAs on effects due to time (the only significant effect detected by the MANOVAs), for medium and long duration groups are shown in Table 7 (medium duration) and Table 8 (long duration).

As indicated in Table 7, all EDR variables showed significant changes due to time in the medium duration group, except for YAWREVII (number of yaw reversals > 2 deg). Table 8 shows that for the long duration group, time was significant for all nine EDR variables. Means and 90% confidence limits are shown for significant EDR variables in Figures 5-12 for medium duration conditions and in Figures 13-21 for long duration conditions.

To determine the nature of time trends (i.e., the trend among successive 10-min observations), two regression analyses were conducted on each of the EDR variables significantly affected by time, for medium and long duration conditions. The first regression analyses conducted were (simple) linear regressions on group mean observations for successive time periods; the second were linear regressions based on individual driver data for successive time periods. Tables 9 and 10 show

**TABLE 7.** Summary of Univarate ANOVAs for Effects of Time on EDR Variables Across Experiments I, II (Normal gain), and III - Medium Duration

| VARIABLE | df | MS | F | p |
|---|---|---|---|---|
| BDYMOVES | 5,45 | 10.3667 | 4.04 | $4x10^{-3}$ |
| HRTINTSD | 5,45 | 1.0889 | 11.17 | $5x10^{-7}$ |
| LATDEVSD | 5,45 | 184.8099 | 6.60 | $1x10^{-4}$ |
| STEERSD | 5,45 | .0009 | 11.43 | $4x10^{-7}$ |
| STREVLRG | 5,45 | 8794.1250 | 5.68 | $4x10^{-4}$ |
| STREVSML | 5,45 | 6581.8250 | 4.35 | $3x10^{-3}$ |
| STRREMN | 5,45 | 1.0722 | 9.53 | $3x10^{-6}$ |
| YAWDEVSD | 5,45 | .0149 | 4.28 | $3x10^{-3}$ |
| YAWREVII | 5,45 | 95.3917 | 1.91 | $> 1x10^{-1}$ |

**TABLE 8.** Summary of Univarate ANOVAs for Effects of Time on EDR Variables Across Experiments I, II (Normal gain) and III - Long Duration

| VARIABLE | df | MS | F | p |
|---|---|---|---|---|
| BDYMOVES | 14,126 | 32.5341 | 1.99 | $2x10^{-2}$ |
| HRTINTSD | 14,126 | .0020 | 6.20 | $3x10^{-9}$ |
| LATDEVSD | 14,126 | 156.5832 | 8.83 | $3x10^{-13}$ |
| STEERSD | 14,126 | .9486 | 6.96 | $2x10^{-10}$ |
| STREVLRG | 14,126 | 9423.4698 | 4.93 | $3x10^{-7}$ |
| STREVSML | 14,126 | 15473.1667 | 9.17 | $1x10^{-13}$ |
| STRREVMN | 14,126 | 1.4320 | 6.07 | $5x10^{-9}$ |
| YAWDEVSD | 14,126 | .0206 | 7.75 | $1x10^{-11}$ |
| YAWREVII | 14,126 | 157.2460 | 4.37 | $3x10^{-6}$ |

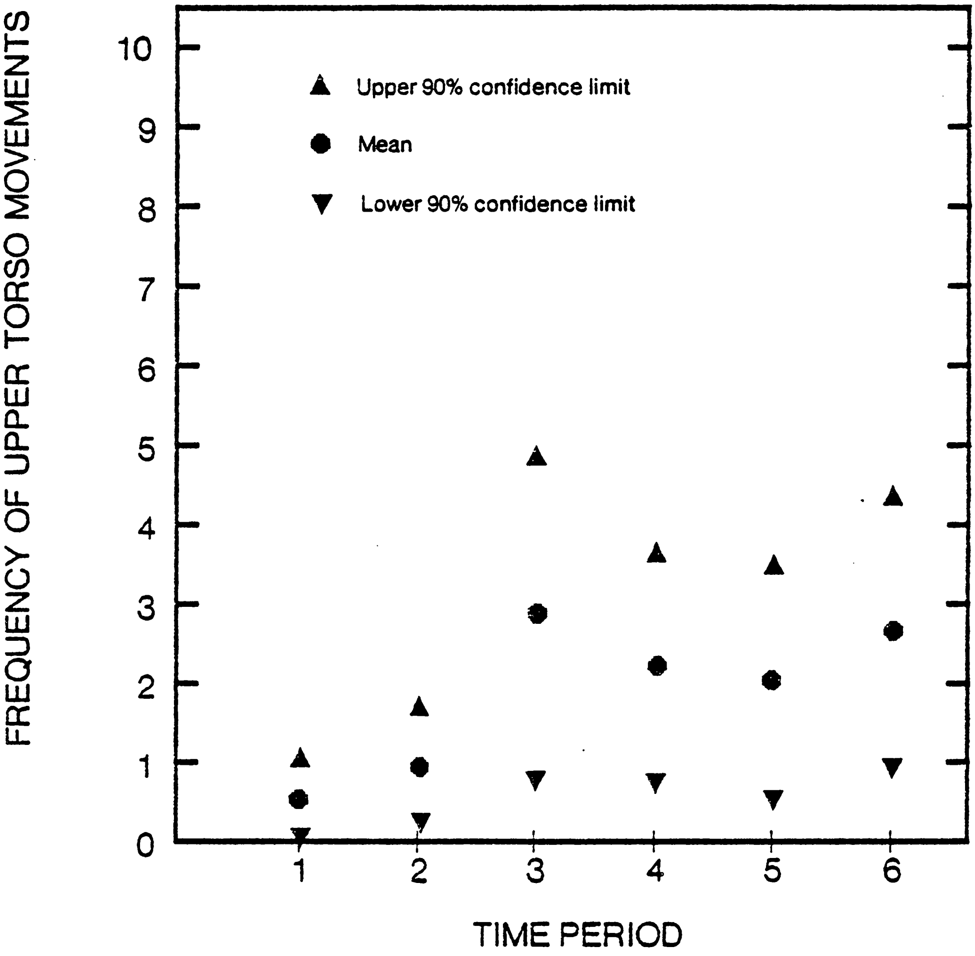

Figure 5. Mean and 90% confidence limits for frequency of upper torso movements (BDYMOVES) for each 10 min period- Medium duration.

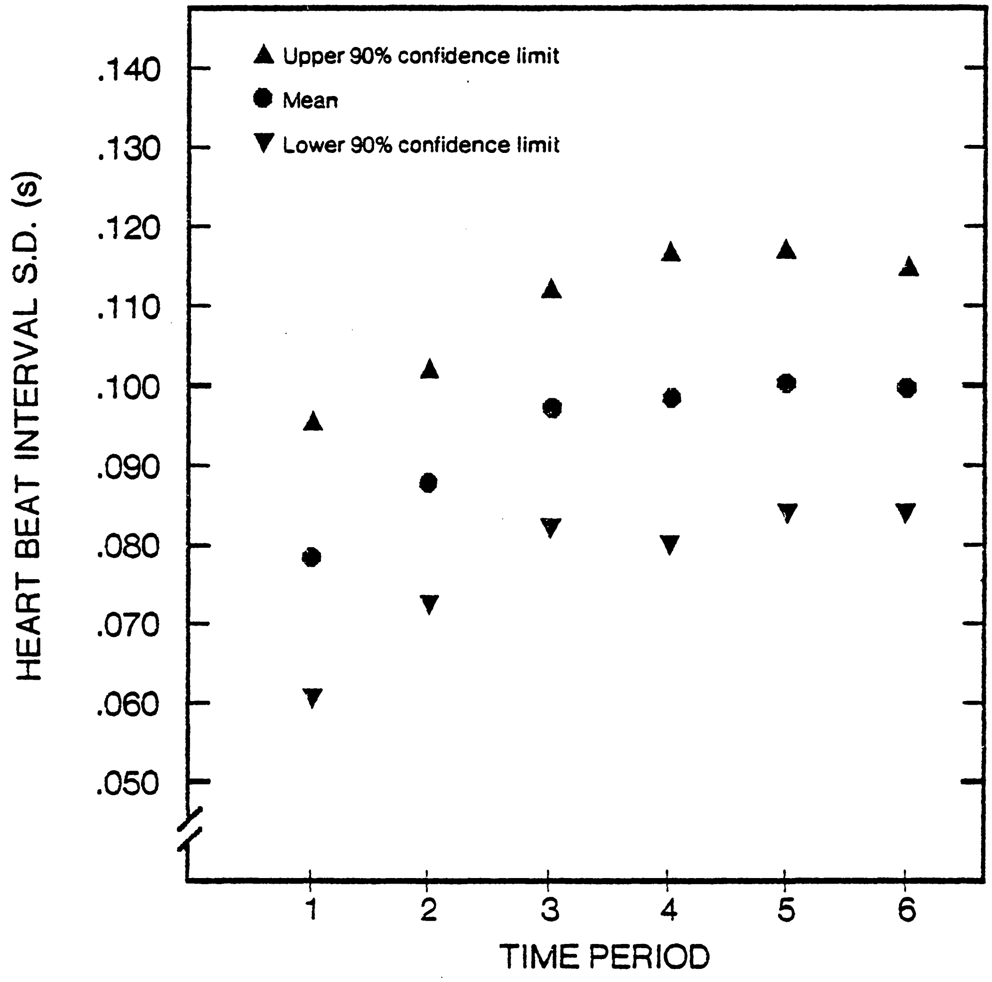

Figure 6. Mean and 90% confidence limits for heart beat interval s.d. (HRTINTSD) for each 10 min period-Medium duration.

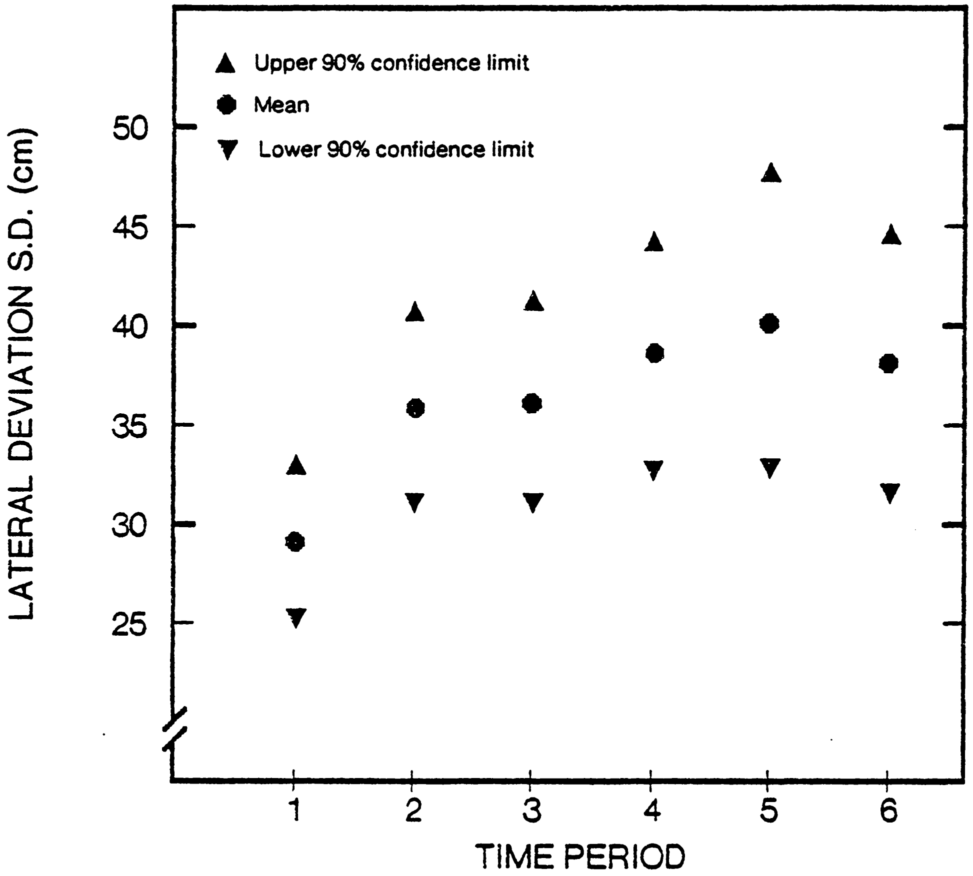

Figure 7. Mean and 90% confidence limits for lateral deviation s.d. (LATDEVSD) for each 10 min period· Medium duration.

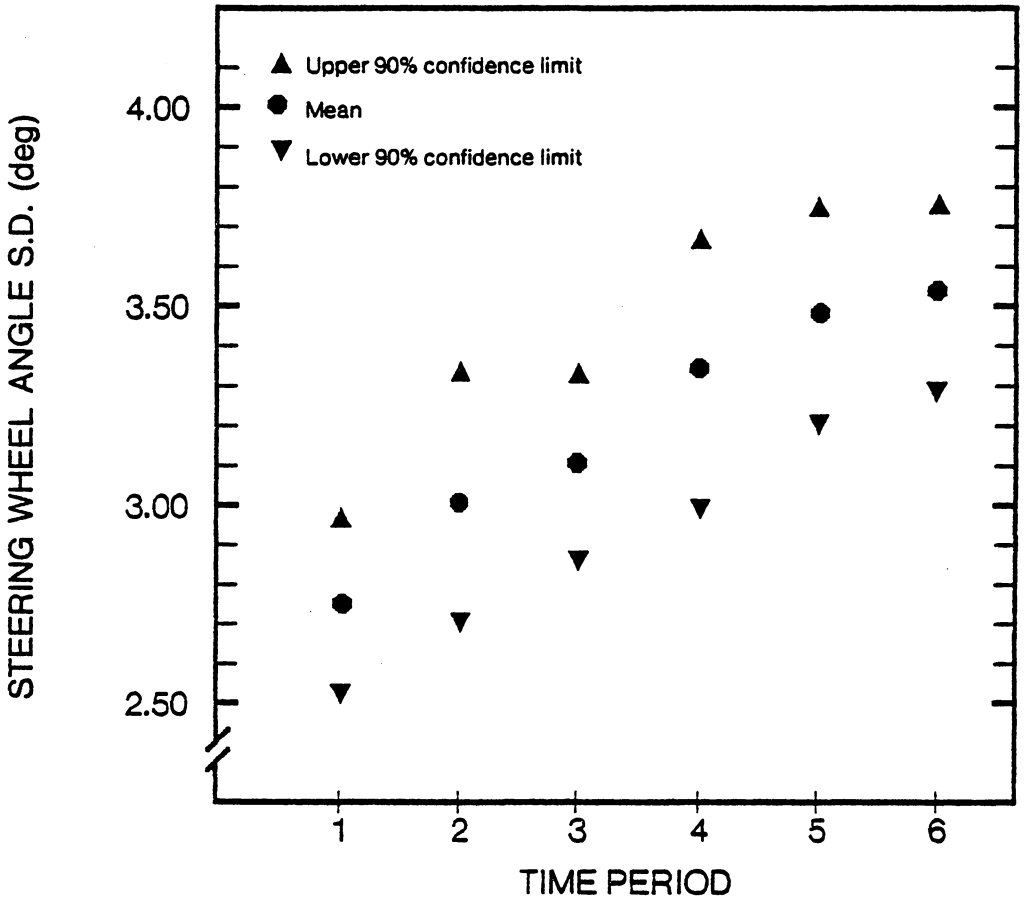

Figure 8. Mean and 90% confidence limits for steering angle s.d. (STEERSD) for each 10 min period-Medium duration.

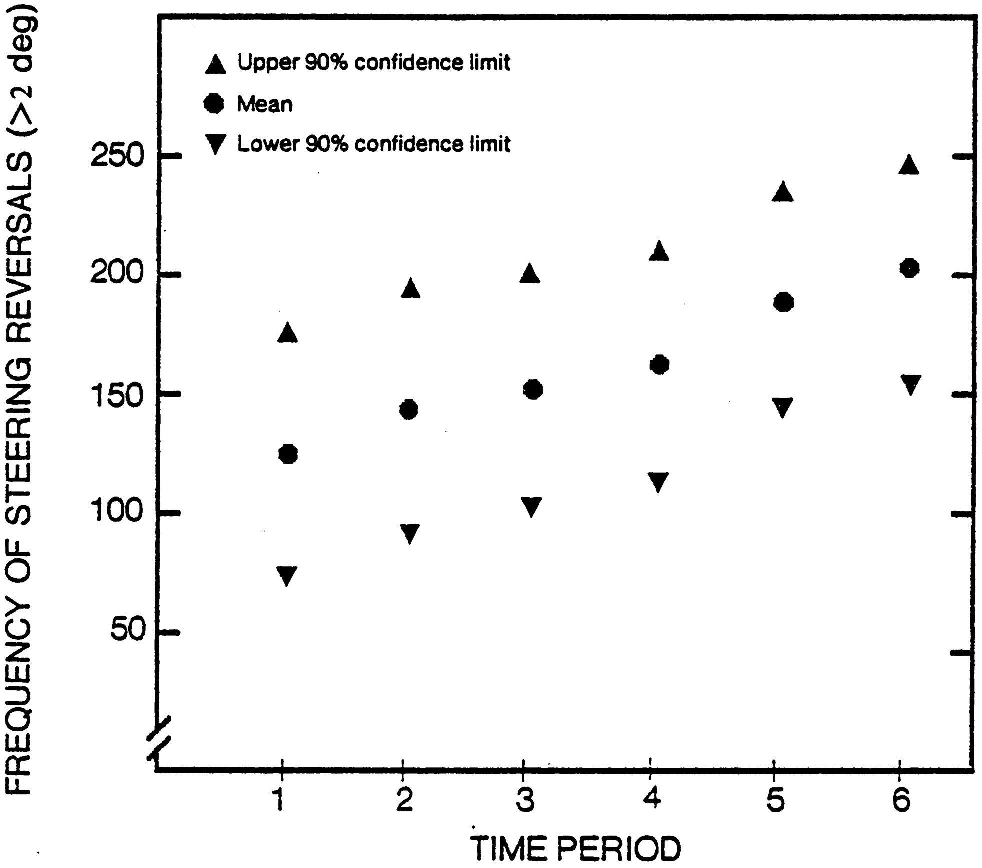

Figure 9. Mean and 90% confidence limits for frequency of large (> 2 deg) steering reversals (STREVLRG) for each 10 min period- Medium duration.

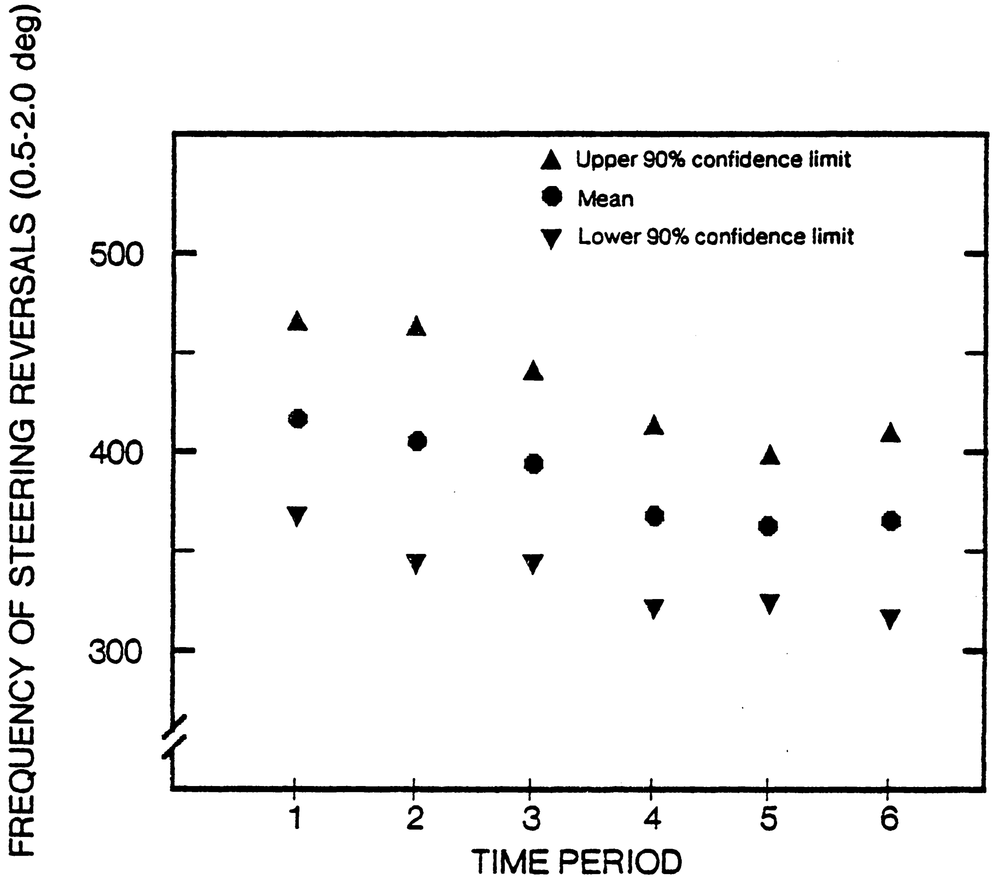

Figure 10. Mean and 90% confidence limits for frequency of small (0.5-2.0 deg) steering reversals (STREVSML) for each 10 min period- Medium duration.

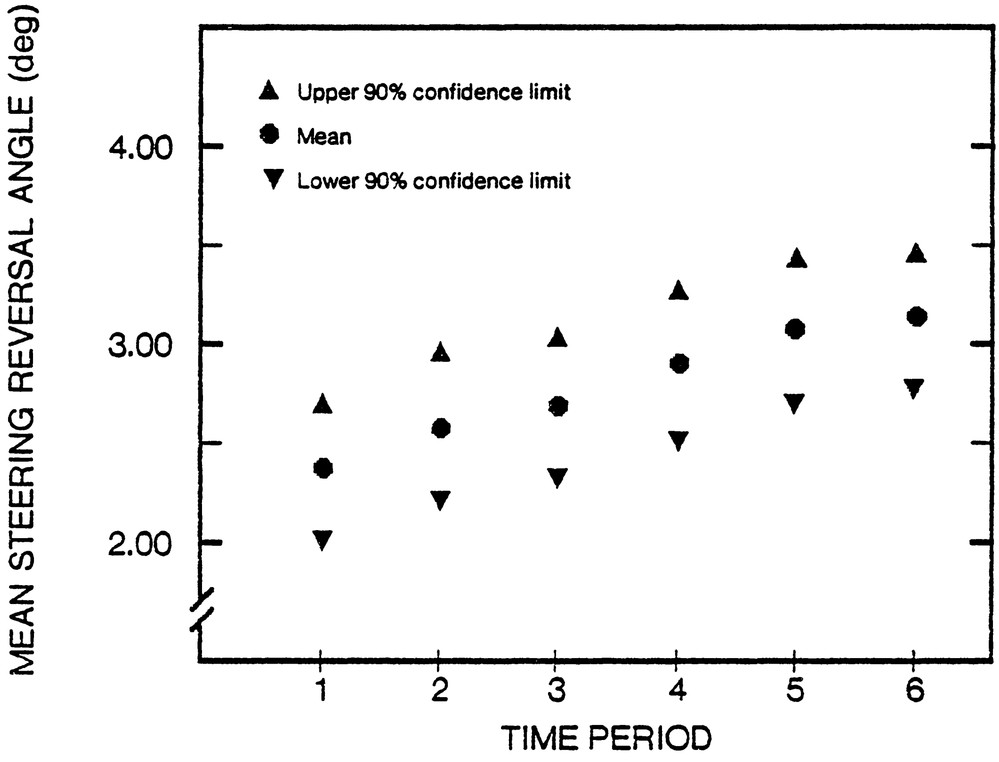

Figure 11. Mean and 90% confidence limits for steering reversal means (STRREVMN) for each 10 min period-Medium duration.

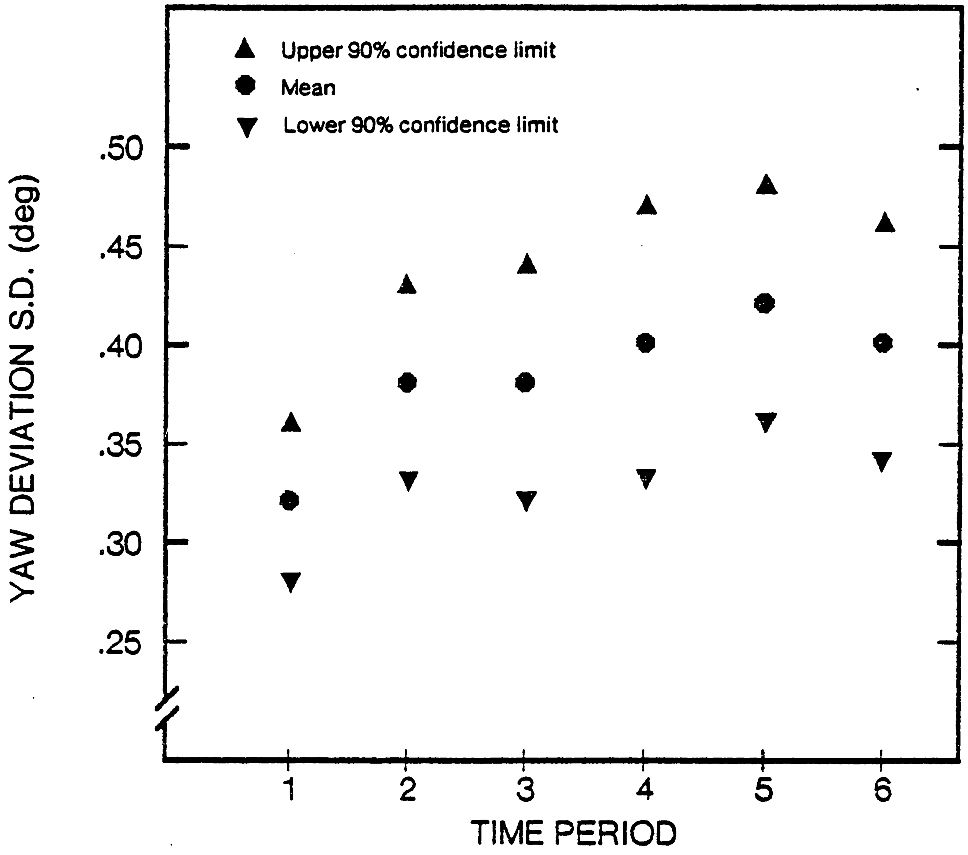

Figure 12. Mean and 90% confidence limits for yaw deviation s.d. (YAWDEVSD) for each 10 min period-Medium duration.

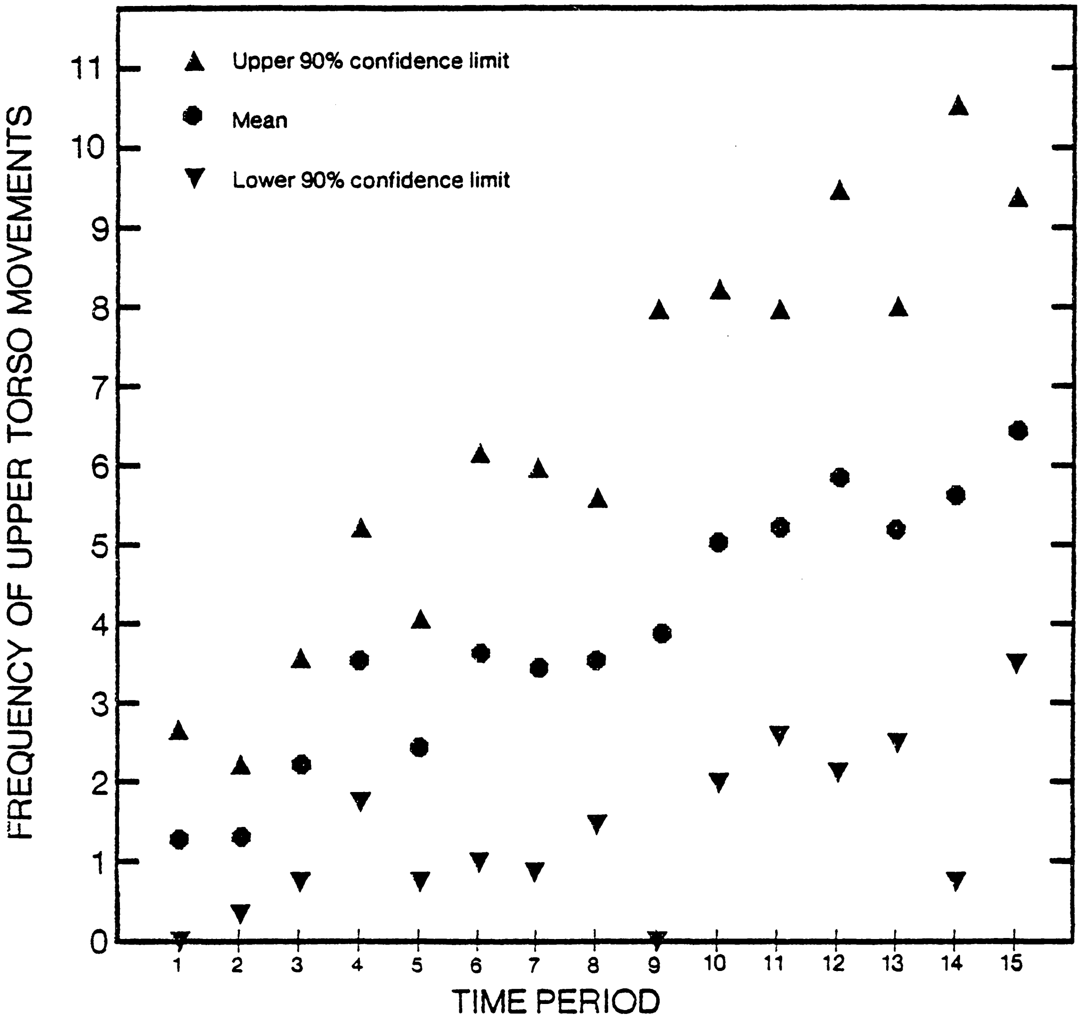

Figure 13. Mean and 90% confidence limits for frequency of upper torso movements (BDYMOVES) for each 10 min period- Long duration.

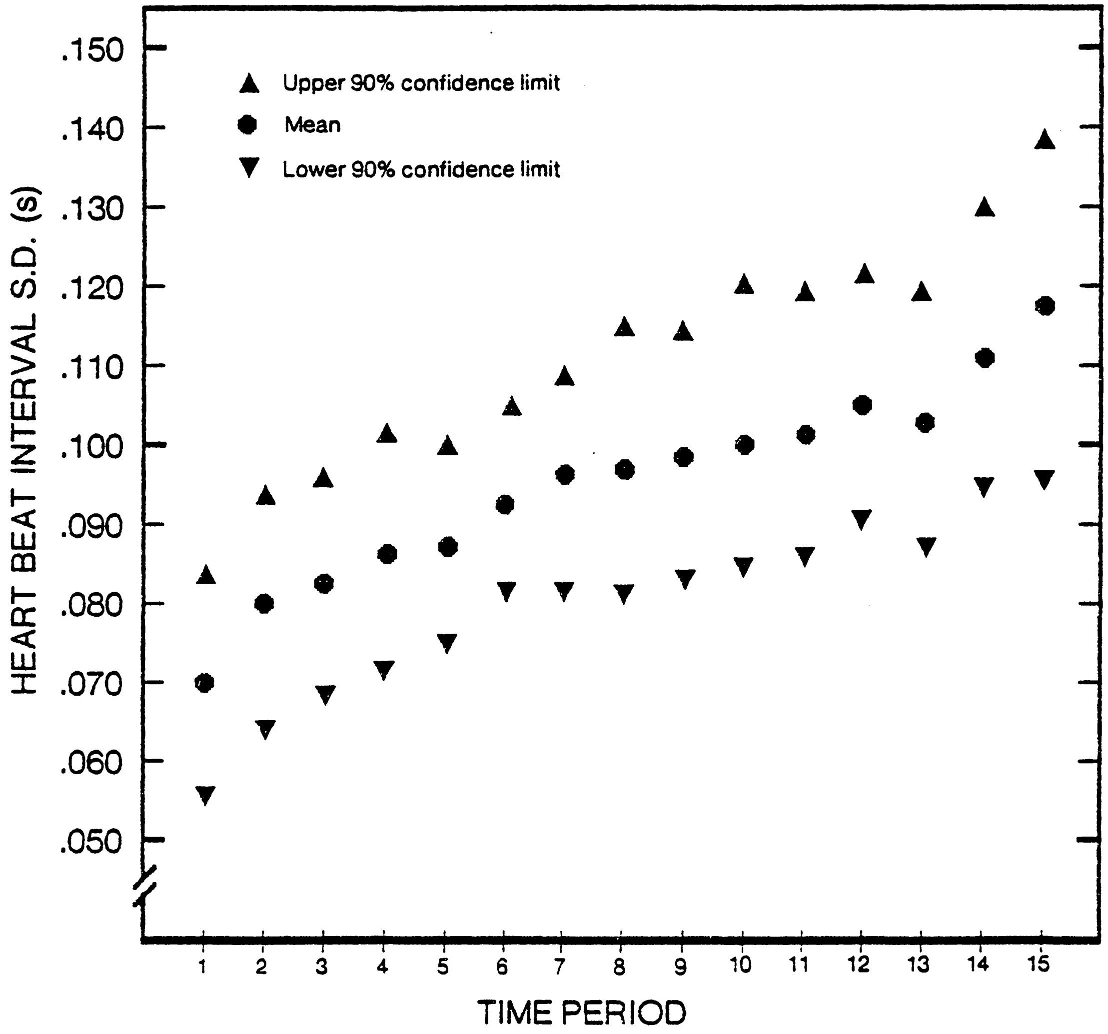

Figure 14. Mean and 90% confidence limits for heart beat interval s.d. (HRTINTSD) for each 10 min period-Long duration.

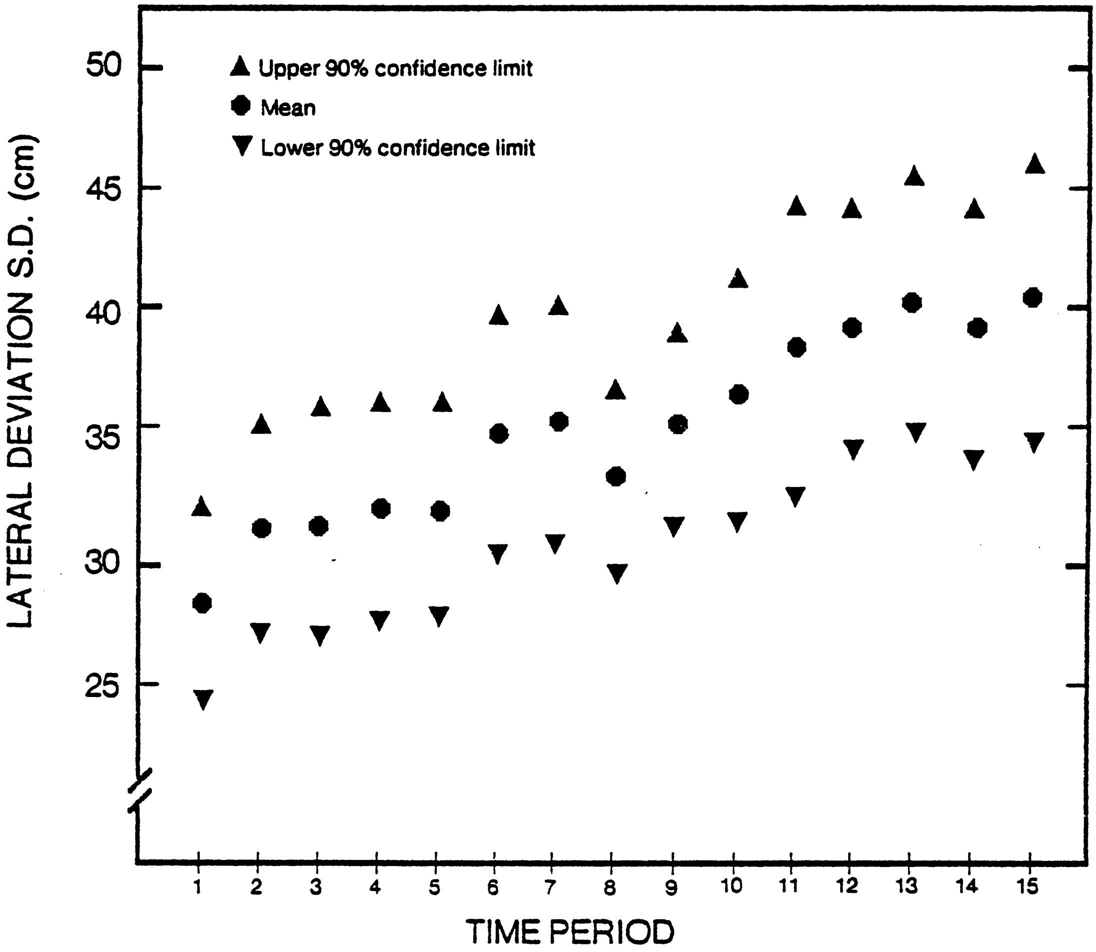

**Figure 15. Mean and 90% confidence limits for lateral deviation s.d. (LATDEVSD) for each 10 min period-Long duration.**

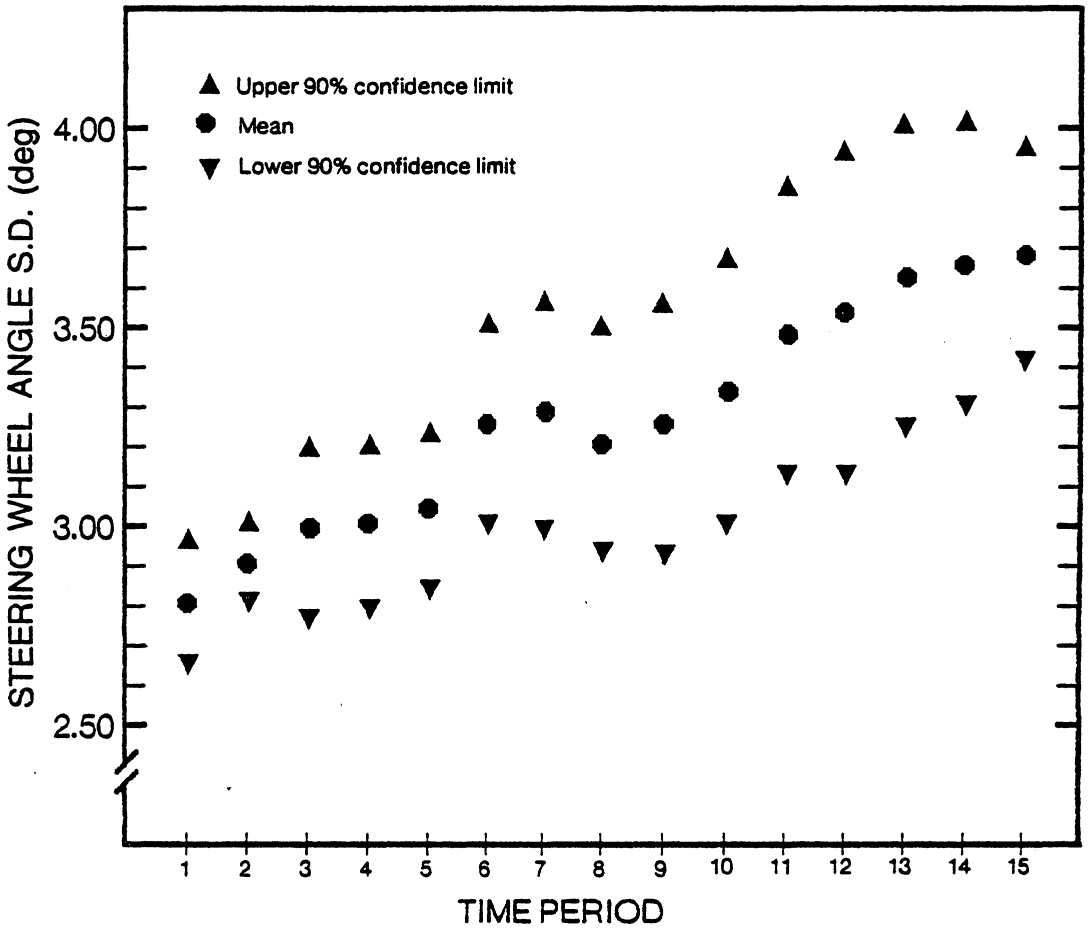

Figure 16. Mean and 90% confidence limits for steering angle s.d. (STEERSD) for each 10 min period- Long duration.

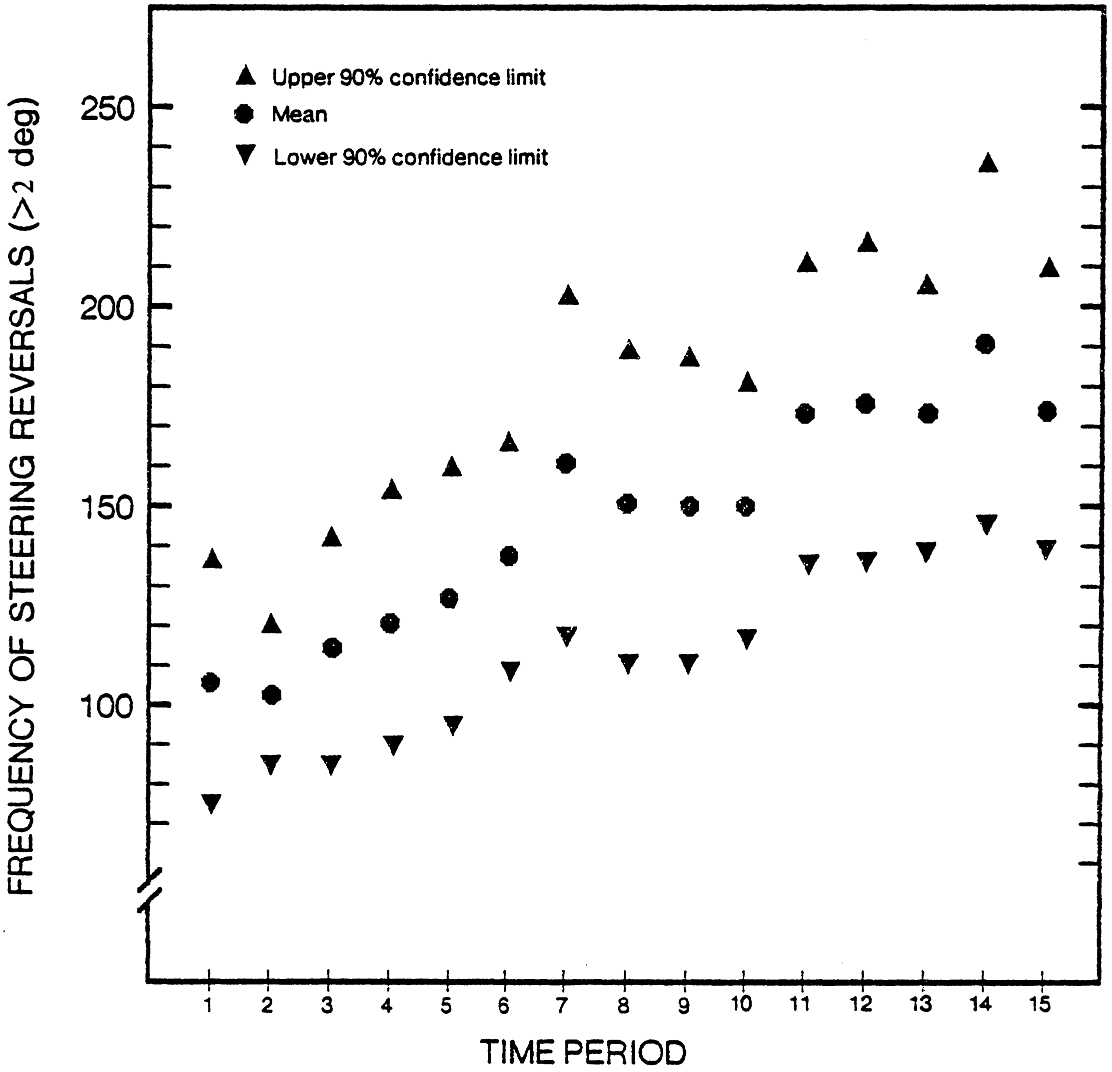

Figure 17. Mean and 90% confidence limits for frequency of large (> 2 deg) steering reversals (STREVLRG) for each 10 min period- Long duration.

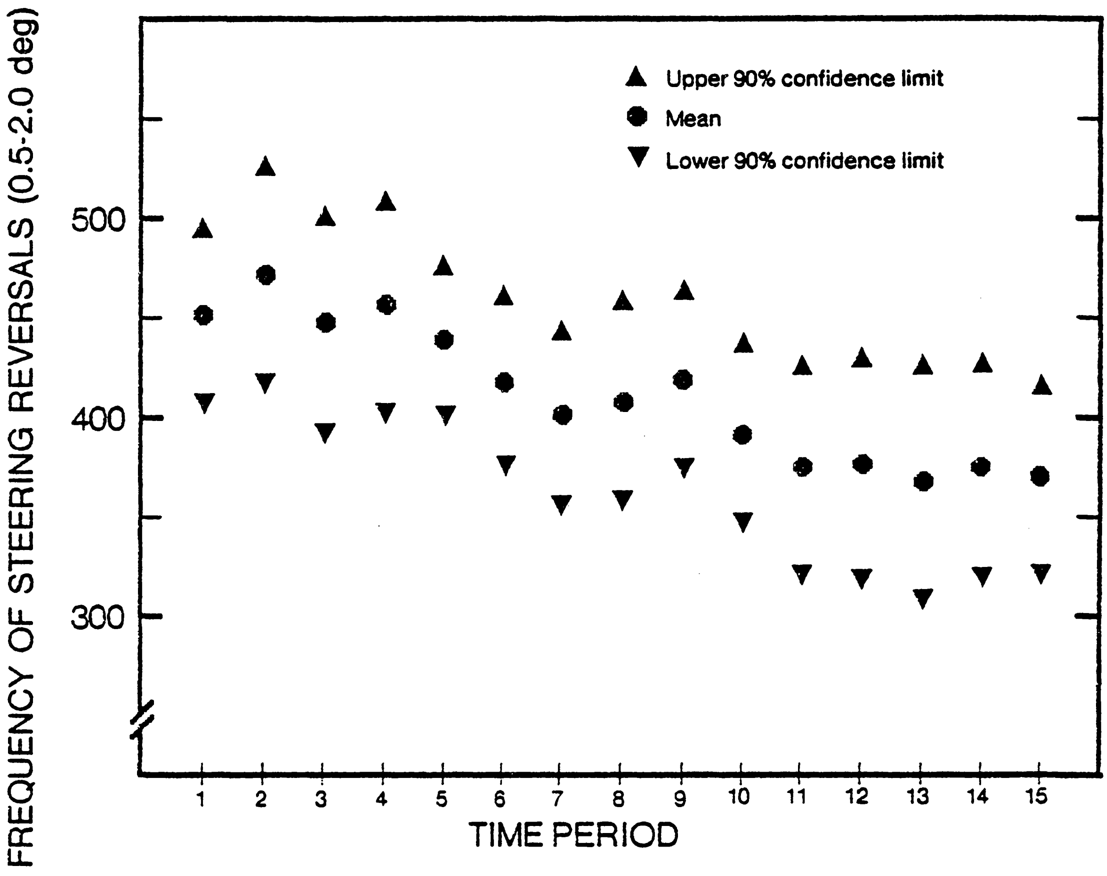

Figure 18. Mean and 90% confidence limits for frequency of small (0.5-2.0 deg) steering reversals (STREVSML) for each 10 min period- Long duration.

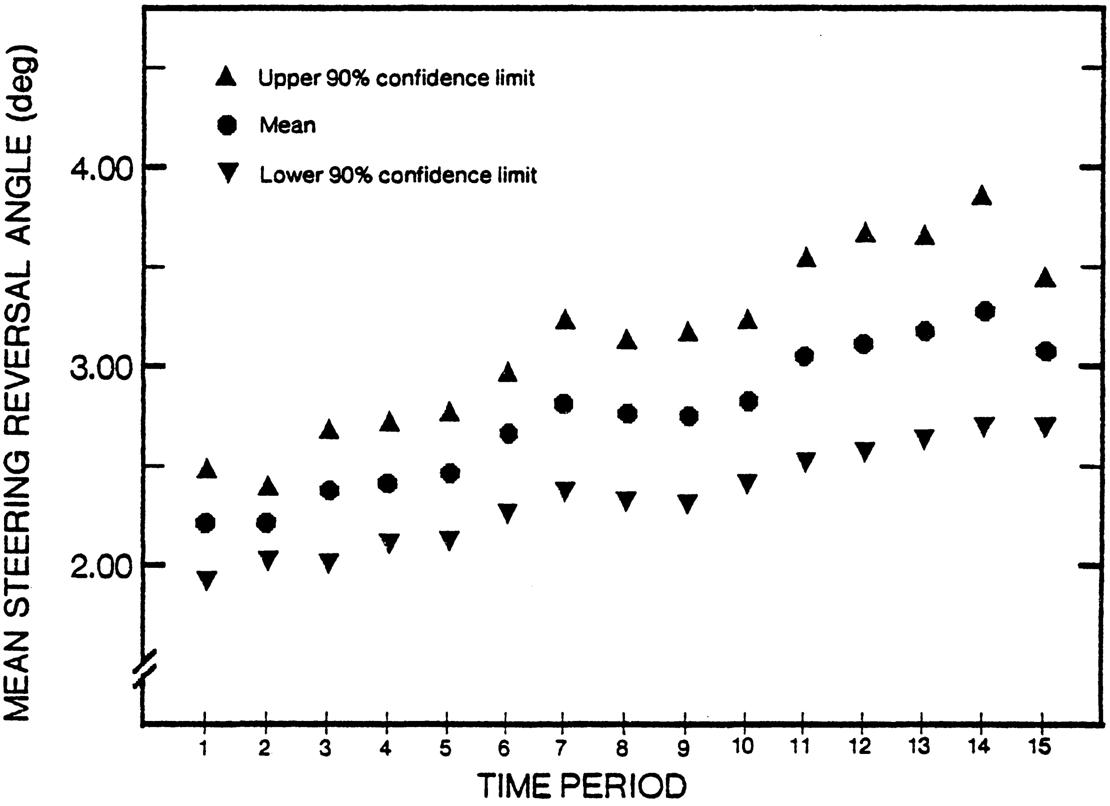

Figure 19. Mean and 90% confidence limits for steering reversal means (STREVMN) for each 10 min period-Long duration.

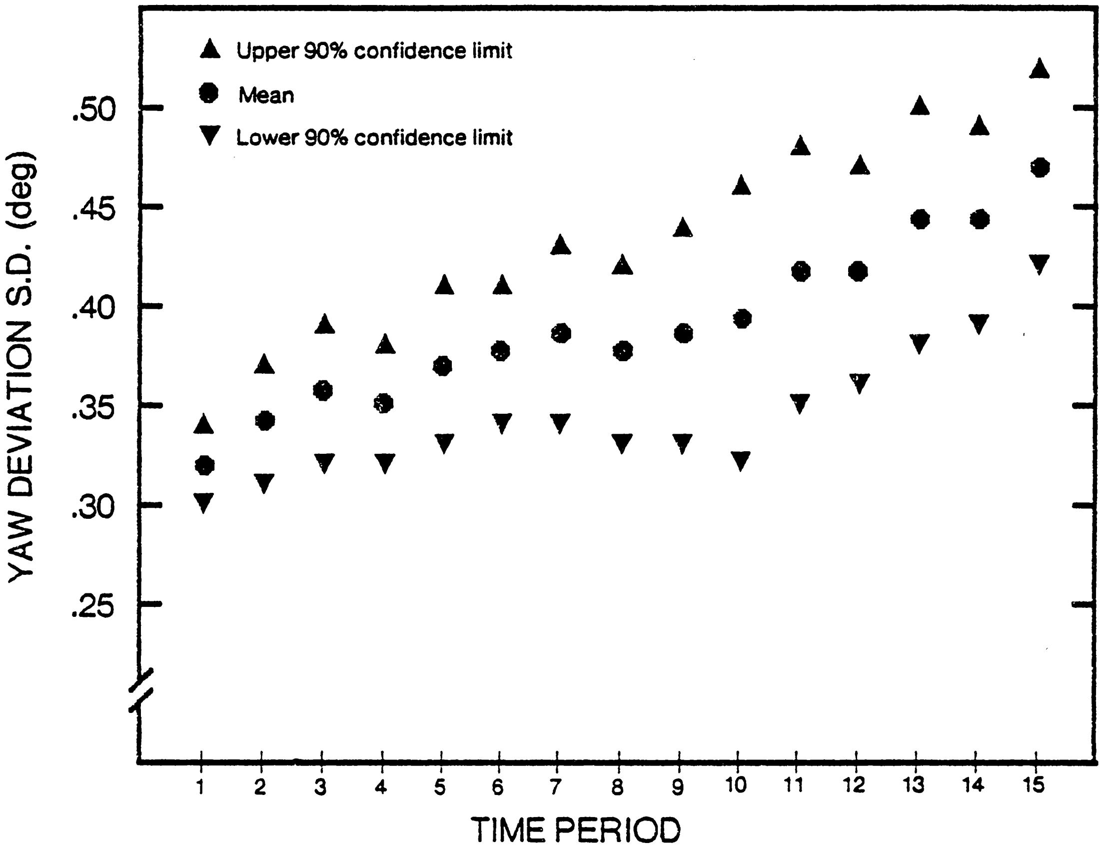

Figure 20. Mean and 90% confidence limits for yaw angle s.d. (YAWDEVSD) for each 10 min period-Long duration.

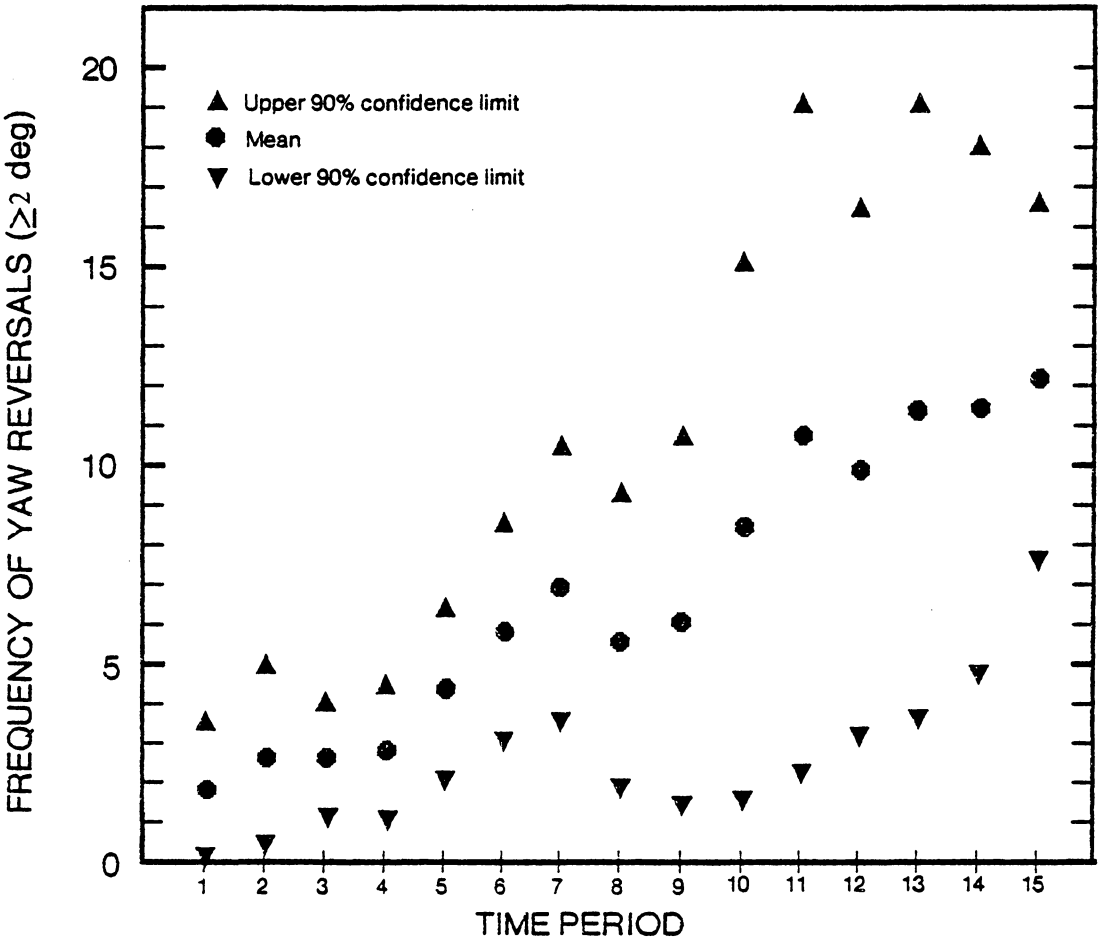

Figure 21. Mean and 90% confidence limits for frequency of yaw reversals (YAWREVII) for each 10 min period- Long duration.

equations and $R^2$ (proportion of the variance accounted for) values derived from regression analyses for significant EDR variables of the medium and long duration groups. As shown in Table 9 (also illustrated in Figures 5-12), among the eight EDR variables which exhibited significant time effects, all except STREVSML exhibited positive (increasing) linear trends over successive time periods. STREVSML exhibited a linearly decreasing trend over successive time periods. Regressions on group mean data for the medium duration condition revealed $R^2$ values ranging from .56 to .98, whereas $R^2$ values for regressions using individual data yielded values of .050 to .218.

For the long duration condition, Table 10 (and Figures 13-21) indicates that among the nine EDR variables which exhibited significant time effects, eight variables exhibited increasing linear trends over time. As in the medium duration condition, STREVSML exhibited decreasing linear trends over time. $R^2$ values for regressions on group mean data ranged from .90 to .97; regressions on individual data yielded $R^2$ values ranging from .080 to .187.

Test for serial dependence. To ensure that successive time observations were uncorrelated (i.e., independent - an assumption of the above parametric analyses), a Durbin-Watson analysis (Durbin and Watson, 1951) was conducted on all dependent variables (combined across duration conditions). The results indicated that all EDR variables were serially uncorrelated ($p < 0.05$) except for STREVLRG of Experiment II, normal gain (for which the test was inconclusive).

## EDR Steering Variables - Experiment II

Significant differences were anticipated among EDR variables of the two steering gain groups in Experiment II. These differences were anticipated

**TABLE 9.** Regression Analysis Summary on EDR Variable Trends Over Time - Medium Duration

| EDR Variable | Equation - Linear Estimate | $R^2$ (Group Means) | $R^2$ (Individual Data) |
|---|---|---|---|
| BDYMOVES | $Y = .53 + .37 X$ | .56 | .050 |
| HRTINTSD | $Y = .078 + .004 X$ | .80 | .052 |
| LATDEVSD | $Y = 30.09 + 1.74 X$ | .69 | .071 |
| STEERSD | $Y = 2.65 + .16 X$ | .96 | .218 |
| STREVLRG | $Y = 110.1 + 14.3 X$ | .98 | .075 |
| STREVSML | $Y = 425.1 - 11.9 X$ | .91 | .049 |
| STRREVMN | $Y = 2.22 + .16 X$ | .98 | .141 |
| YAWDEVSD | $Y = .324 + .016 X$ | .74 | .066 |

**TABLE 10.** Regression Analysis Summary on EDR Variable Trends Over Time - Long Duration

| EDR Variable | Equation - Linear Estimate | $R^2$ (Group Means) | $R^2$ (Individual Data) |
|---|---|---|---|
| BDYMOVES | $Y = 1.04 + .35 X$ | .92 | .080 |
| HRTINTSD | $Y = .073 + .003 X$ | .96 | .148 |
| LATDEVSD | $Y = 29.1 + .78 X$ | .93 | .134 |
| STEERSD | $Y = 2.77 + .06 X$ | .97 | .187 |
| STREVLRG | $Y = 98.5 + 6.0 X$ | .91 | .135 |
| STREVSML | $Y = 470.2 - 7.6 X$ | .90 | .110 |
| STRREVMN | $Y = 2.12 + .075 X$ | .94 | .153 |
| YAWDEVSD | $Y = .316 + .009 X$ | .94 | .165 |
| YAWREVII | $Y = .47 + .79 X$ | .95 | .123 |

primarily in variables involving steering wheel movements, but not necessarily on vehicle control, heart beat, and body movement variables. To test this hypothesis, MANOVAs were conducted separately on steering wheel variables (i.e., STEERSD, STRREVMN, STREVSML, and STREVLRG), and on the remaining vehicle control and body variables (BDYMOVES, HRTINTSD, LATDEVSD, YAWDEVSD, and YAWREVII) for each of the three duration conditions.

Steering control variables, analyzed with two-way MANOVAs (steering gain x time) are shown for each duration group in Tables 11-13. As in the previous overall EDR analyses, time was significant for medium ($p < 0.01$) and long ($p < 0.01$) duration conditions. Followup univariate ANOVAs for the time effect (summarized in Table 14) indicated significant time effects in the medium duration condition for STEERSD, STRREVMN, and STREVLRG, but not for STREVSML. In the long duration condition, only STEERSD and STRREVMN exhibited significant time effects. As shown in Figures 22-24 for medium duration, and Figures 25 and 26 for long duration groups (by steering gain) these variables are essentially consistent with previously data showing a positive trend toward larger steering movements over time.

The MANOVAs on EDR steering variables (Tables 11-13) also indicated steering gain was significant for short ($p < 0.05$), medium ($p < 0.05$) and long ($p < 0.05$) duration conditions. Followup univariate ANOVAs performed on steering control variables (summarized in Table 15) revealed that steering gain had a significant effect on STEERSD, STRREVMN and STREVLRG, across all duration conditions. STREVSML failed to exhibit any effect due to time for any of the duration conditions.

Table 16, which presents steering variable means by steering gain and

**TABLE 11.** MANOVA Summary for EDR Steering Variables of Experiment II - Short Duration

| SOURCE | p | $\nu$ H | $\nu$ E | U - Statistic[a] |
|---|---|---|---|---|
| Steering Gain (G) | 4 | 1 | 6 | .0275* |
| Subjects/G (S/G) | (Error term for G) | | | |
| Time (T) | 4 | 2 | 12 | .3853 |
| G x T | 4 | 2 | 12 | .2535 |
| T x S/G | (Error term for T, G x T) | | | |

* $\underline{p} < .05$

a. p = number of dependent variables

$\nu_H$ = degrees of freedom for treatment

$\nu_E$ = degrees of freedom for error

U - statistic (likehood ratio) = $\dfrac{\underline{E}}{\underline{E} + \underline{H}}$

Where $\underline{H}$ = Sum of squares and cross products matrix for treatment

$\underline{E}$ = sum of squares and cross-products matrix for error

**TABLE 12.** MANOVA Summary for EDR Steering Variables of Experiment II - Medium Duration

| SOURCE | p | $\nu_H$ | $\nu_E$ | U - Statistic[a] |
|---|---|---|---|---|
| Steering Gain (G) | 4 | 1 | 6 | .0743* |
| Subjects/G (S/G) | | (Error term for G) | | |
| Time (T) | 4 | 5 | 30 | .2598** |
| G x T | 4 | 5 | 30 | .4961 |
| T x S/G | | (Error term for T, G x T) | | |

* $\underline{p}$ <.05
** $\underline{p}$ <.01

a. p = number of dependent variables

$\nu_H$ = degrees of freedom for treatment

$\nu_E$ = degrees of freedom for error

U - statistic (likehood ratio) = $\dfrac{\underline{E}}{\underline{E} + \underline{H}}$

Where $\underline{H}$ = Sum of squares and cross products matrix for treatment

$\underline{E}$ = Sum of squares and cross products matrix for error

**TABLE 13.** MANOVA Summary for EDR Steering Variables of Experiment II - Long Duration

| SOURCE | p | $\nu_H$ | $\nu_E$ | U - Statistic[a] |
|---|---|---|---|---|
| Steering Gain (G) | 4 | 1 | 6 | .0164** |
| Subjects/G (S/G) | (Error term for G) | | | |
| Time (T) | 4 | 14 | 84 | .3338** |
| G x T | 4 | 14 | 84 | .5895 |
| T x S/G | (Error term for T, G x T) | | | |

** $\underline{p}$ <.01

a. p = number of dependent variables

$\nu_H$ = degrees of freedom for treatment

$\nu_E$ = degrees of freedom for error

U - statistic (likehood ratio) = $\dfrac{\underline{E}}{\underline{E} + \underline{H}}$

Where $\underline{H}$ = Sum of squares and cross products matrix for treatment

$\underline{E}$ = Sum of squares and cross-products matrix for error

**TABLE 14.** Summary of Univariate ANOVAs for Effects of Time on EDR Steering Variables of Experiment II

Short Duration

Time was not significant

Medium Duration

| Variables | df | MS | F | p |
|---|---|---|---|---|
| STEERSD | 5,30 | .1818 | 30.48 | .0015 |
| STRREVMN | 5,30 | .3166 | 4.87 | .0022 |
| STREVSML | 5,30 | 776.67 | .34 | >.10 |
| STREVLRG | 5,30 | 2728.04 | 3.04 | .025 |

Long Duration

| Variable | df | MS | F | p |
|---|---|---|---|---|
| STEERSD | 14,84 | .177 | 141.90 | .00002 |
| STREVMN | 14,84 | .249 | 53.74 | .0003 |
| STREVSML | 14,84 | 8022.73 | 2.28 | >.10 |
| STREVLRG | 14,84 | 1532.07 | 1.60 | .097 |

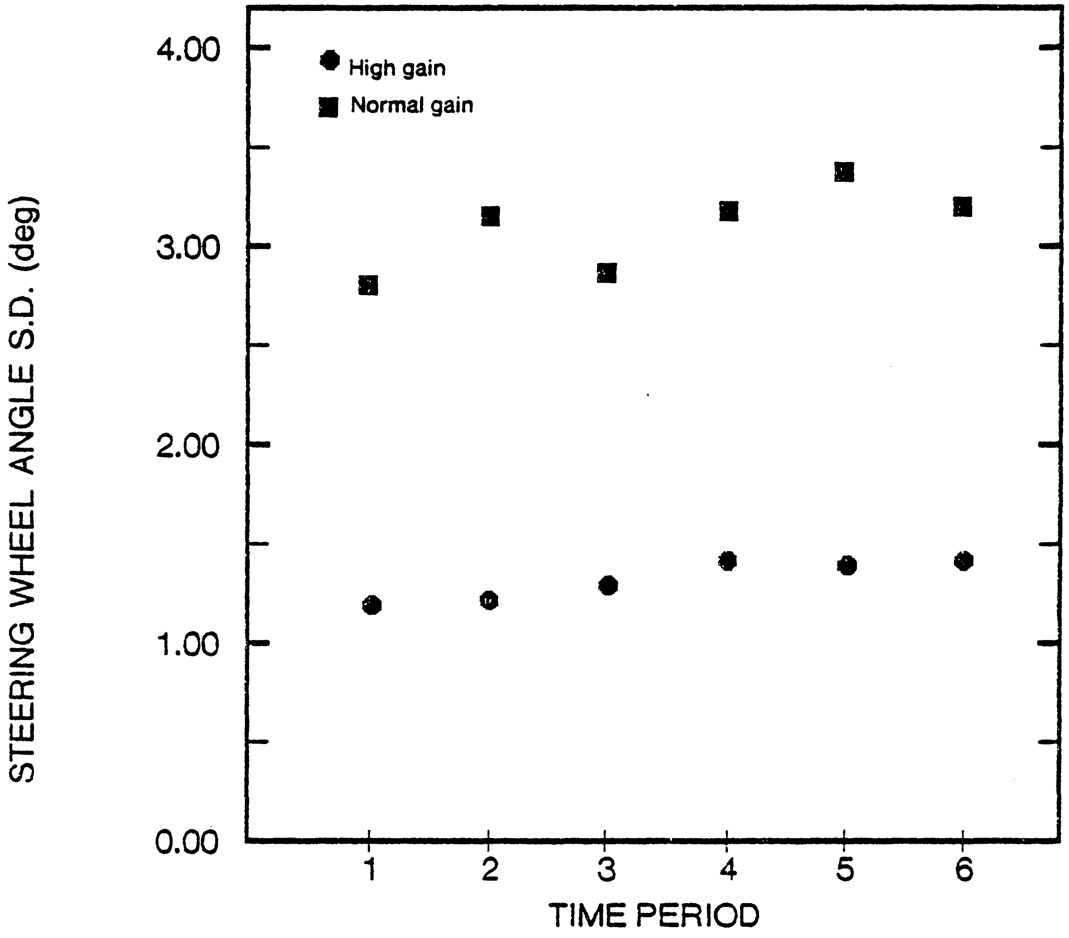

Figure 22. Mean steering angle s.d. (STEERSD), by time period, for normal and high steering gain groups of Experiment II- Medium duration.

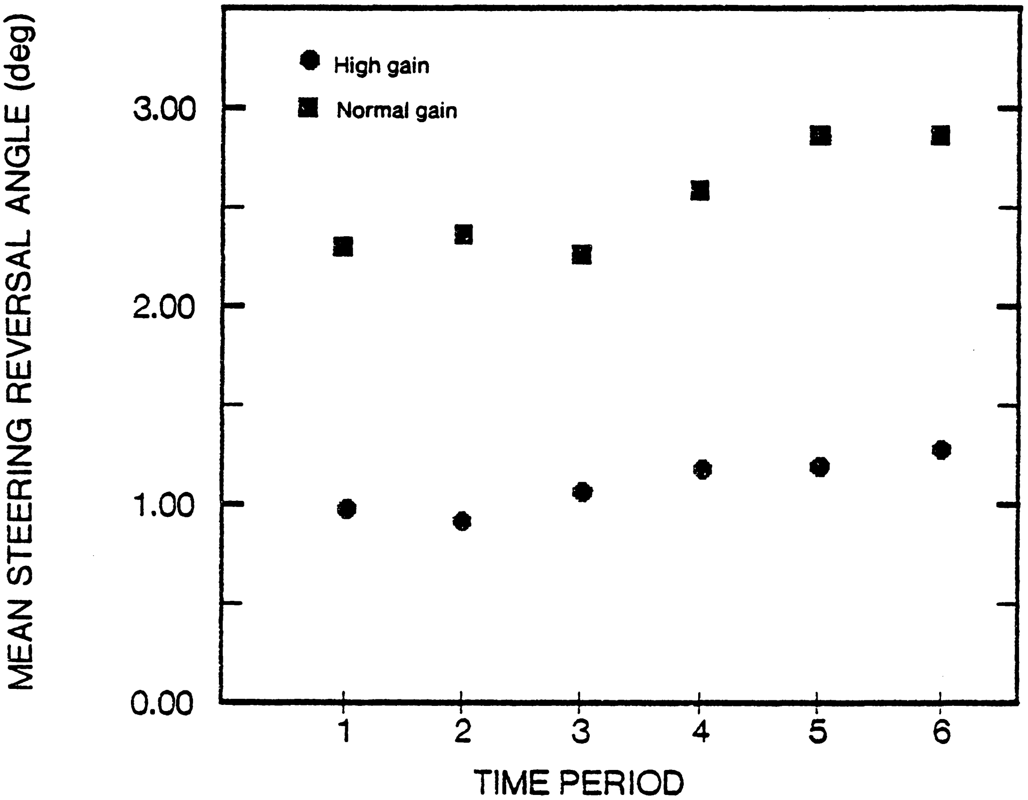

Figure 23. Mean steering reversal means (STREVMN), by time period, for normal and high steering gain groups of Experiment II- Medium duration.

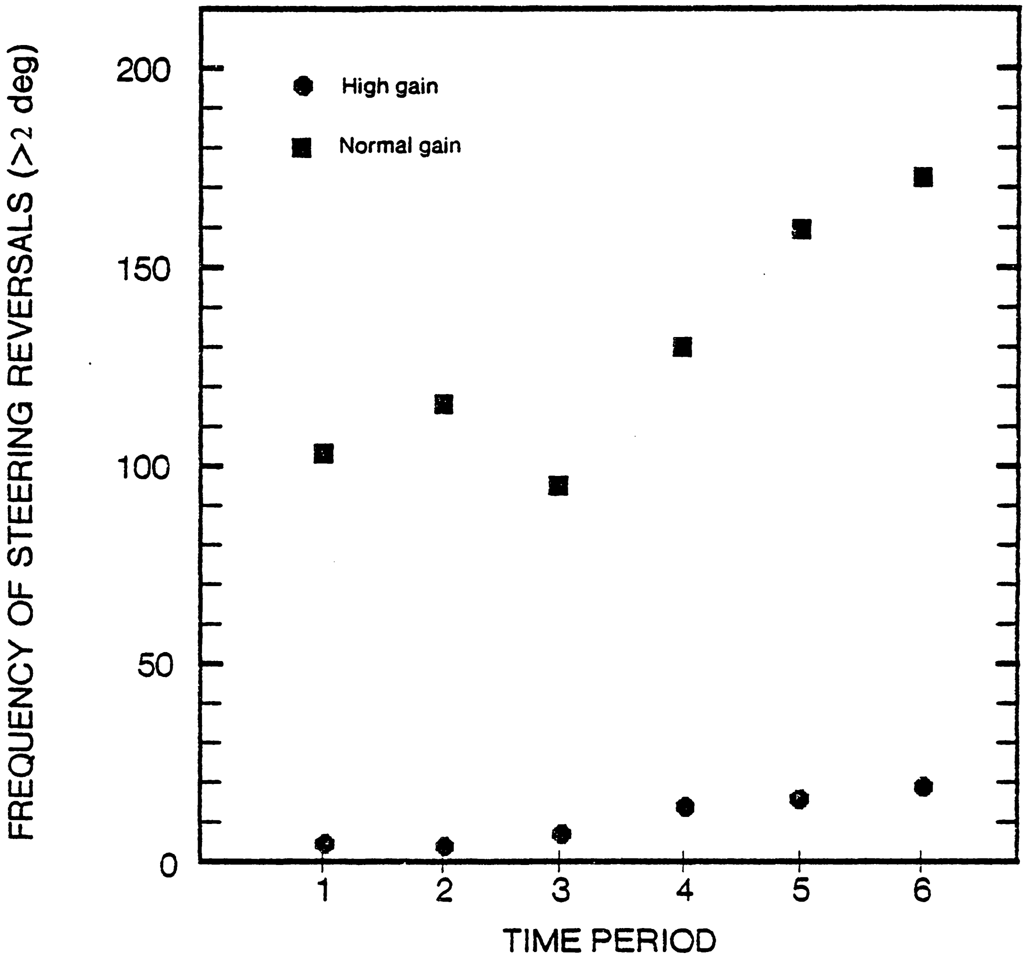

Figure 24. Mean frequency of large (> 2 deg) steering reversals (STREVLRG), by time period, for normal and high steering gain groups of Experiment II- Medium duration.

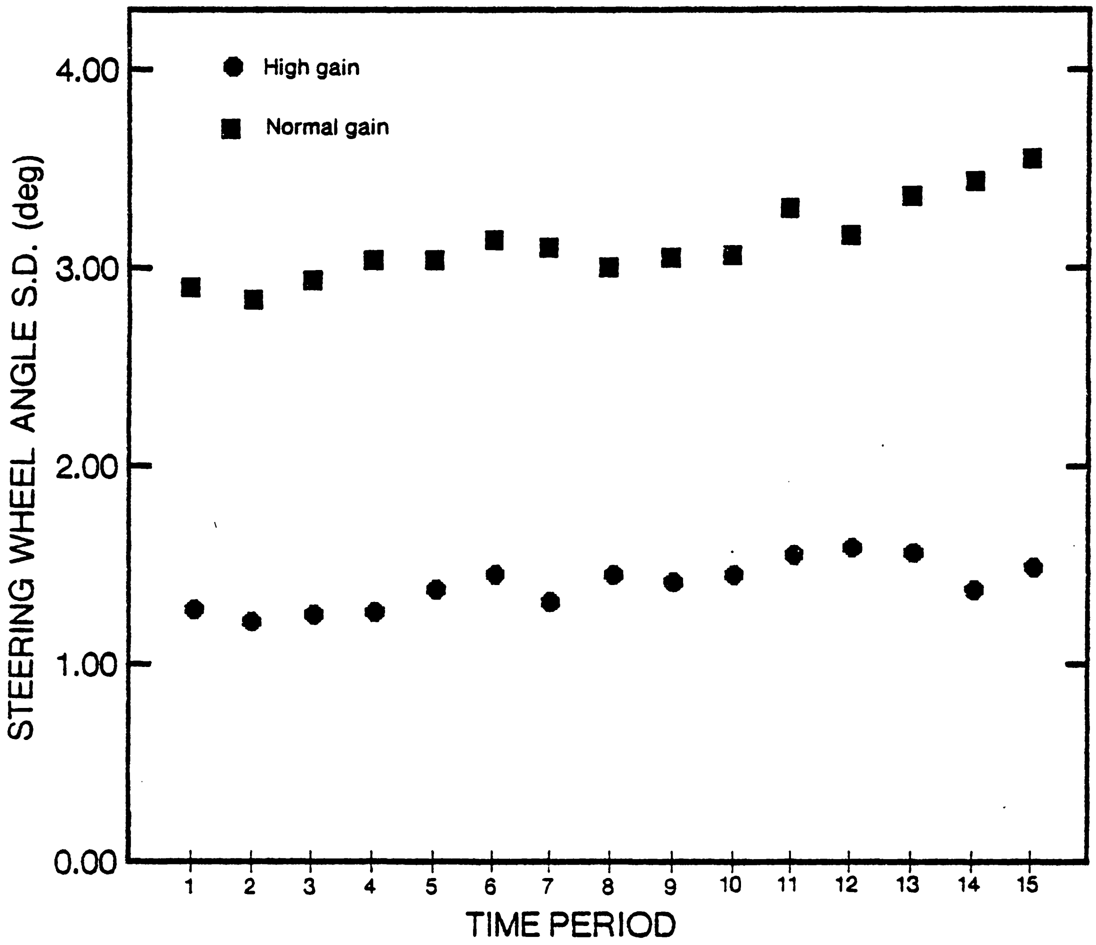

Figure 25. Mean steering angle s.d. (STEERSD), by time period, for normal and high steering gain groups of Experiment II- Long duration.

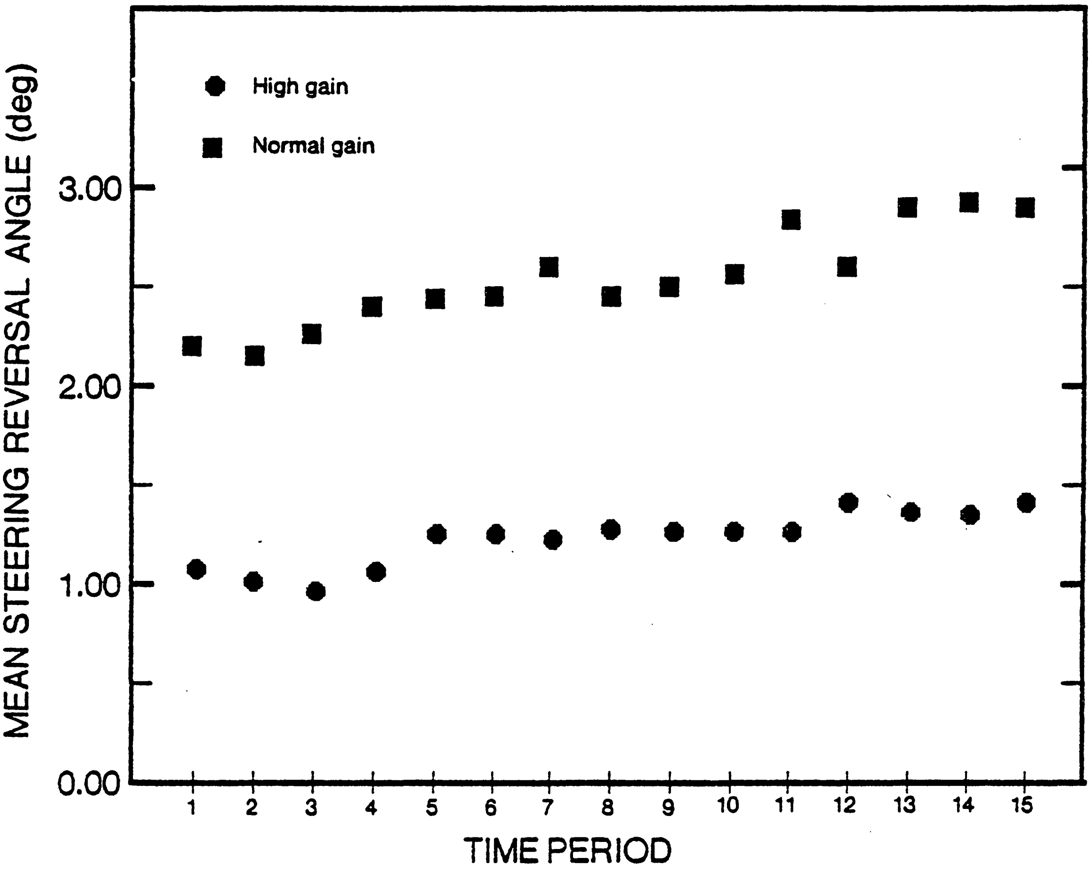

Figure 26. Mean steering reversal means (STREVMN), by time period, for normal and high steering gain groups of Experiment II- Long duration.

**TABLE 15.** Summary of Univariate ANOVAs for Effects of Steering Gain on EDR Steering Variables of Experiment II

Short Duration

| VARIABLE | df | MS | F | p |
|---|---|---|---|---|
| STEERSD | 1,6 | 13.50 | 94.92 | .00007 |
| STRREVMN | 1,6 | 6.70 | 27.08 | .002 |
| STREVSML | 1,6 | 532.0 | .04 | >.10 |
| STREVLRG | 1,6 | 52266.7 | 18.79 | .005 |

Medium Duration

| VARIABLE | df | MS | F | p |
|---|---|---|---|---|
| STEERSD | 1,6 | 38.16 | 30.48 | .00015 |
| STRREVMN | 1,6 | 25.95 | 18.78 | .005 |
| STREVSML | 1,6 | 13567.7 | .43 | >.10 |
| STREVLRG | 1,6 | 170289.2 | 10.69 | .017 |

Long Duration

| VARIABLE | df | MS | F | p |
|---|---|---|---|---|
| STEERSD | 1,6 | 91.00 | 141.90 | .00002 |
| STRREVMN | 1,6 | 53.61 | 53.74 | .0003 |
| STREVSML | 1,6 | 174498.1 | 2.28 | > .10 |
| STREVLRG | 1,6 | 382279.4 | 47.03 | .0005 |

**TABLE 16.** EDR Steering Variable Means, by Steering Gain and Duration - Experiment II

| EDR Steering Variable | Duration | Normal Gain | High Gain | p |
|---|---|---|---|---|
| STEERSD (deg) | Short | 2.87 | 1.37 | .00007 |
| | Medium | 3.09 | 1.30 | .00200 |
| | Long | 3.12 | 1.38 | .00002 |
| STRREVMN (deg) | Short | 2.31 | 1.26 | .0020 |
| | Medium | 2.55 | 1.08 | .0050 |
| | Long | 2.54 | 1.20 | .0003 |
| STREVLRG<br>(number per 10 min) | Short | 110.6 | 17.3 | .0050 |
| | Medium | 129.0 | 9.9 | .0170 |
| | Long | 131.3 | 18.4 | .0005 |
| STREVSML<br>(number per 10 min) | Short | 481.8 | 472.4 | >.10 |
| | Medium | 431.0 | 464.7 | >.10 |
| | Long | 408.3 | 484.6 | >.10 |

duration condition, shows that mean values for STEERSD, STRREVMN and STREVLRG for the normal steering gain group were greater than means of the high gain group, indicating that drivers in the normal steering gain group employed larger steering movements than those in the high gain group.

The MANOVAs, shown in Tables 11-13, failed to reveal significant steering gain x time interactions for any duration condition. These results indicate that among the steering control variables, the effects of time were not differentially affected by steering gain (indicating that the time trends for the two steering gain groups are parallel).

### Vehicle Control and Body Variables - Experiment II

The MANOVAs conducted on non-steering variables of Experiment II (i.e., BDYMOVES, HRTINTSD, LATDEVSD, YAWDEVSD, and YAWREVII), shown in Tables 17-19, indicate that the only significant effect was due to time in the long duration condition ($p < 0.05$). All other effects across duration conditions were not significant ($p > 0.05$). Univariate ANOVAs for significant time effects (summarized in Table 20) revealed that all of the vehicle control and body variables in the long duration group condition were significantly affected by time. The nature of these time trends, shown in Figures 27-31, are consistent with corresponding EDR data of the overall analyses.

### Experiment I - Emergency Response Data

Brake response time data were analyzed in a 3 x 2 x 10 analysis of variance. The factors included in the analysis were: duration (30, 60, and 150 min), run (pre- and post-EDR), and trials (10 emergency trials per run). The results of the ANOVA, shown in Table 21, revealed significant effects of run, $F(1,9) = 13.96$, $p =$

**TABLE 17.** MANOVA Summary for EDR Vehicle Control and Body Variables of Experiment II - Short Duration

| SOURCE | p | $\nu_H$ | $\nu_E$ | U - Statistic[a] |
|---|---|---|---|---|
| Steering Gain (G) | 5 | 1 | 6 | .2704 |
| Subjects/G (S/G) | | (Error term for G) | | |
| Time (T) | 5 | 2 | 12 | .3425 |
| G x T | 5 | 2 | 12 | .3916 |
| T x S/G | | (Error term for T, G x T) | | |

a. p = number of dependent variables

$\nu_H$ = degrees of freedom for treatment

$\nu_E$ = degrees of freedom for error

U - statistic (likehood ratio) = $\dfrac{\underline{E}}{\underline{E} + \underline{H}}$

Where $\underline{H}$ = Sum of squares and cross-products matrix for treatment

$\underline{E}$ = sum of squares and cross-products matrix for error

**TABLE 18.** MANOVA Summary for EDR Vehicle Control and Body Variables of Experiment II - Medium Duration

| SOURCE | p | $\nu_H$ | $\nu_E$ | U - Statistic[a] |
|---|---|---|---|---|
| Steering Gain (G) | 5 | 1 | 6 | .0223 |
| Subjects/G (S/G) | | (Error term for G) | | |
| Time (T) | 5 | 5 | 30 | .4018 |
| G x T | 5 | 5 | 30 | .5155 |
| T x S/G | | (Error term for T, G x T) | | |

a. p = number of dependent variables

$\nu_H$ = degrees of freedom for treatment

$\nu_E$ = degrees of freedom for error

U - statistic (likehood ratio) = $\frac{\underline{E}}{\underline{E} + \underline{H}}$

Where $\underline{H}$ = Sum of squares and cross-products matrix for treatment

$\underline{E}$ = sum of squares and cross-products matrix for error

**TABLE 19.** MANOVA Summary for EDR Vehicle Control and Body Variables of Experiment II - Long Duration

| SOURCE | p | $\nu$ H | $\nu$ E | U - Statistic[a] |
|---|---|---|---|---|
| Steering Gain (G) | 5 | 1 | 6 | .0886 |
| Subjects/G (S/G) | (Error term for G) | | | |
| Time (T) | 5 | 14 | 84 | .2051** |
| G x T | 5 | 14 | 84 | .4122 |
| T x S/G | (Error term for T, G x T) | | | |

** $p$ <.01

a. p = number of dependent variables

$\nu_H$ = degrees of freedom for treatment

$\nu_E$ = degrees of freedom for error

U - statistic (likehood ratio) = $\frac{\underline{E}}{\underline{E} + \underline{H}}$

Where $\underline{H}$ = Sum of squares and cross-products matrix for treatment

$\underline{E}$ = sum of squares and cross-products matrix for error

**TABLE 20.** Summary of Univariate ANOVAs for Effects of Time on EDR Vehicle Control and Body Variables Experiment II - Long Duration

| VARIABLE | df | MS | F | p |
|---|---|---|---|---|
| BDYMOVES | 14,84 | 15.4643 | 2.19 | .025 |
| HRTINTSD | 14,84 | .0018 | 7.38 | .001 |
| LATDEVSD | 14,84 | 123.0577 | 4.94 | .001 |
| YAWDEVSD | 14,84 | .0016 | 3.02 | .005 |
| YAWREVII | 14,84 | 70.0929 | 2.40 | .025 |

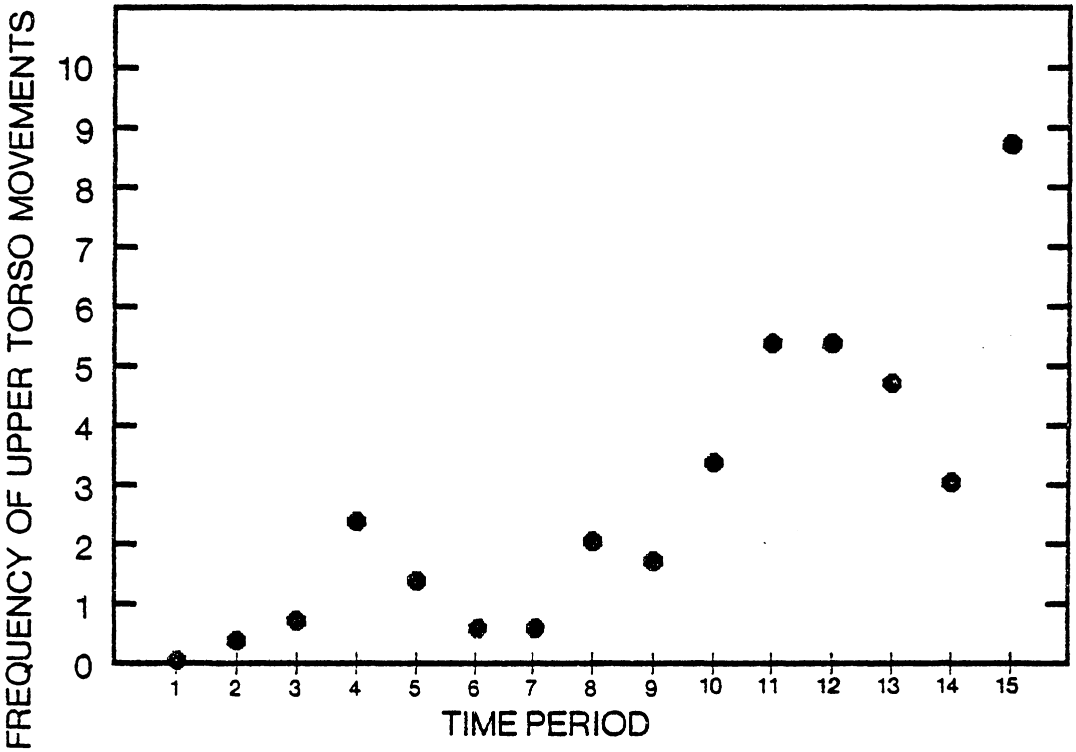

Figure 27. Mean frequency of upper torso movements (BDYMOVES) for each 10 min period of Experiment II-Long duration.

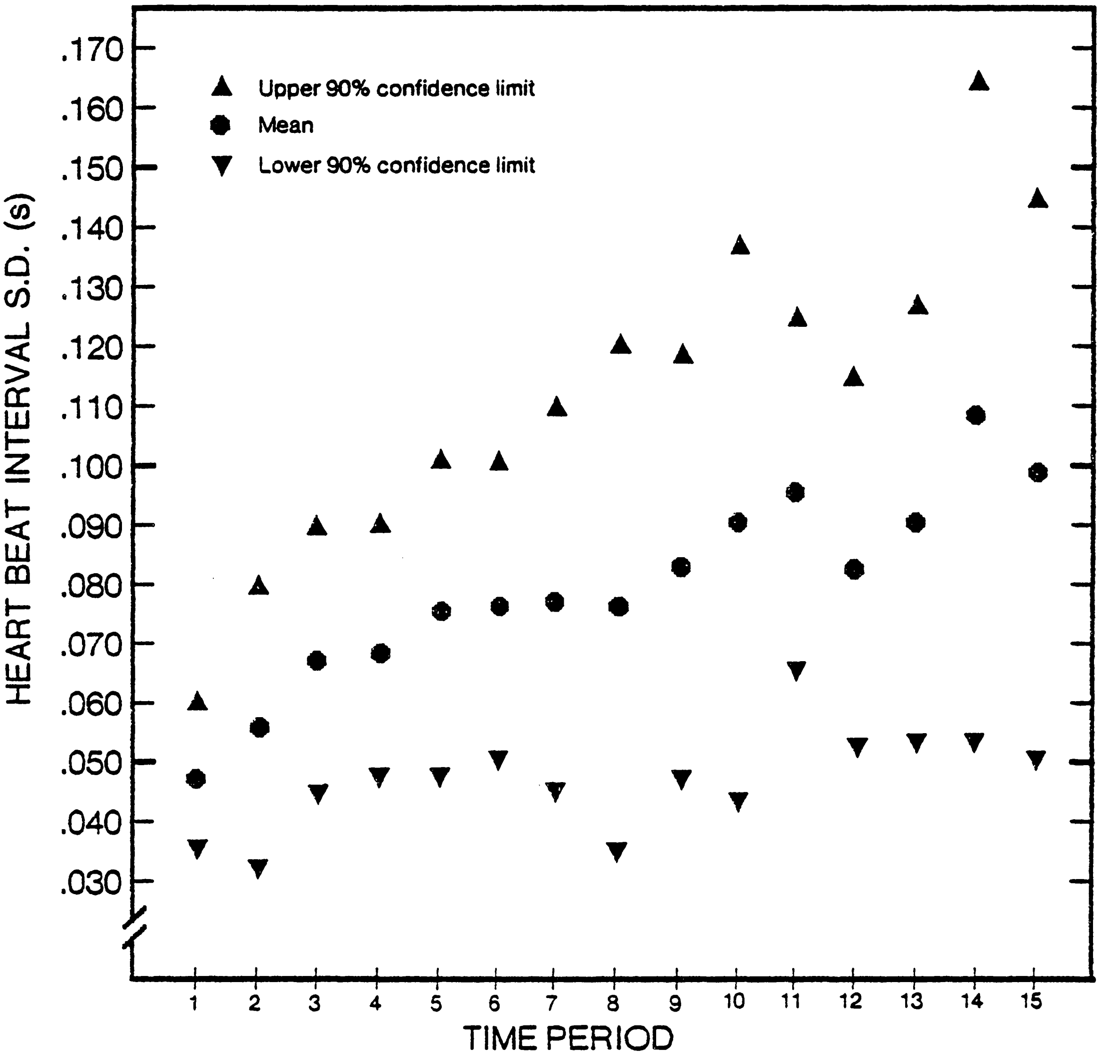

Figure 28. Mean and 90% confidence limits for heart beat interval s.d. (HRTINTSD), for each 10 min period of Experiment II-Long duration.

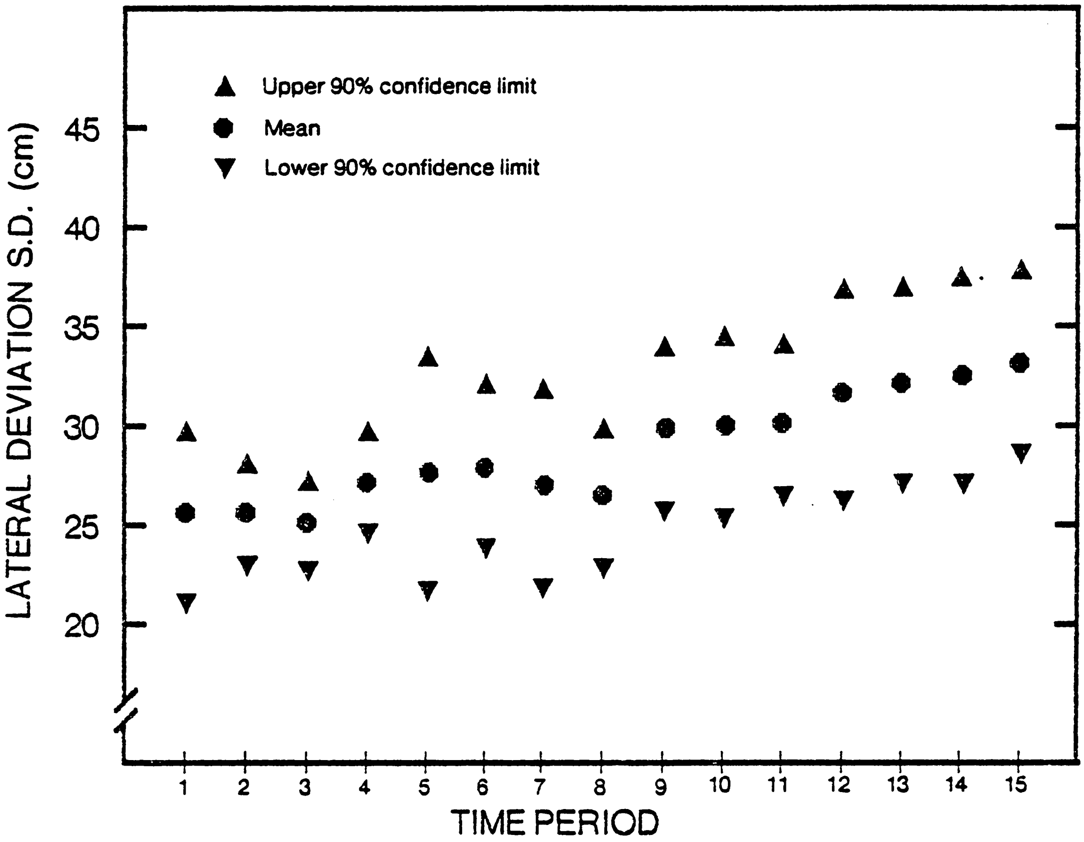

Figure 29. Mean and 90% confidence limits for lateral deviation s.d. (LATDEVSD), for each 10 min period of Experiment II- Long duration.

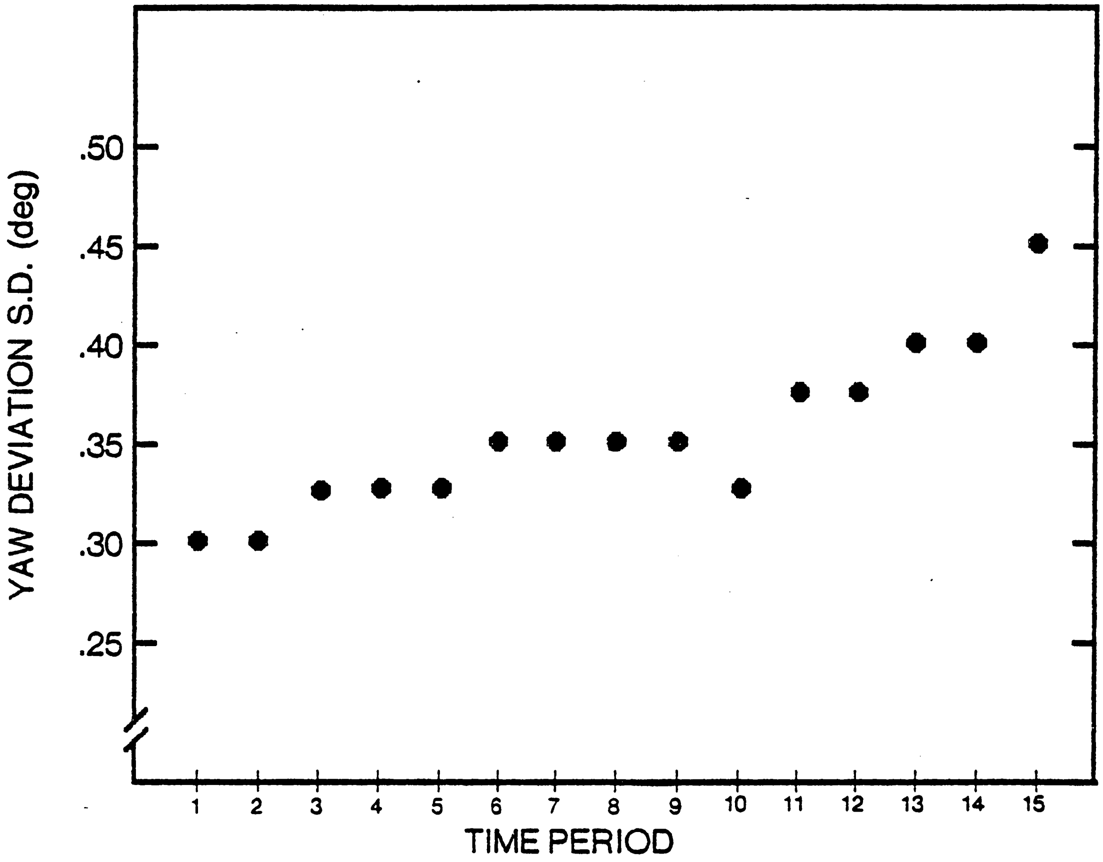

Figure 30. Mean yaw deviation s.d. (YAWDEVSD) for each 10 min period of Experiment II- Long duration.

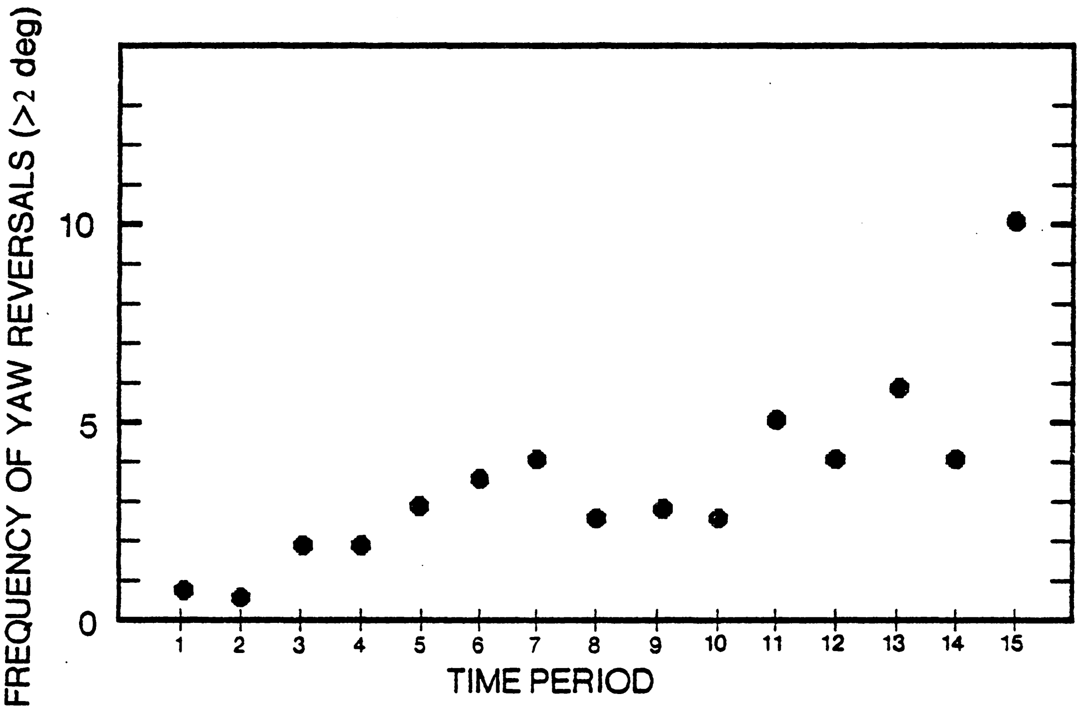

Figure 31. Mean frequency of yaw reversals (YAWREVII) for each 10 min period of Experiment II- Long duration.

**TABLE 21.** Analysis of Variance Summary for Brake Response Times of Experiment I

| SOURCE | df | MS | F | p |
|---|---|---|---|---|
| BETWEEN SUBJECTS | | | | |
| Duration (D) | 2 | .3079 | 1.94 | >.10 |
| Subjects/D (S/D) | 9 | .1589 | | |
| WITHIN SUBJECTS | | | | |
| Run (R) | 1 | .4611 | 13.96 | .0046 |
| Trials (T) | 9 | .0768 | 1.90 | >.05 |
| D x R | 2 | .0046 | .14 | >.10 |
| D x T | 18 | .0311 | .77 | >.10 |
| R x T | 9 | .1167 | 3.76 | .0006 |
| R x S/D | 9 | .0330 | | |
| T x S/D | 81 | .0403 | | |
| D x R x T | 18 | .0225 | .72 | >.10 |
| R x T x S/D | 81 | .0310 | | |
| Total | 239 | | | |

0.0046, and run x trials, $F(9,81) = 3.76$, $p = 0.0006$. A Newman-Keuls analysis of the run x trials interaction indicated that the mean reaction time of drivers for the first emergency trial of the post-EDR ($\bar{X} = 1.64$ s) was significantly longer than all trials ($p<0.05$) except trial 4 of the pre-EDRs and trials 2, 3, 4, 5, 6, and 8 of the post-EDRs ($p>0.05$). The mean response time of drivers for trial 3 of the post-EDR runs ($\bar{X} = 1.74$ s) was significantly longer than all other response times ($p<0.05$) except trial 1 of post-EDRs ($p>0.05$). All other differences were not significant ($p>0.05$). These results, which are illustrated in Figure 32, indicate that the predominant source of the run main effect and the run x trials interaction was the longer reaction times observed among early post-EDR trials (specifically trials 1 and 3) relative to correponding pre-EDR trials.

## Experiment II - Emergency Response Data

Wind gust recovery performance data (mean lateral deviation error scores) were analyzed in a 2 x 2 x 3 x 10 analysis of variance. Factors included in the analysis were run (pre- and post-EDRs), steering gain (normal and high gain), duration (short, medium, and long), and trials (1-10). The ANOVA, shown in Table 22, revealed significant effects due to run, $F(1,18) = 10.85$, $p = .0040$, and duration x trial interaction, $F(18,162) = 1.83$, $p = 0.0262$. All other main effects and interactions were not significant ($p > 0.05$).

The main effect due to run indicates that the mean lateral deviation error for the post-EDR trials ($\bar{X} = 65.7$ cm) was significantly larger than that of the pre-EDR trials ($\bar{X} = 57.0$ cm).

The significance of the duration x trial interaction, illustrated in Figure 33, (in the absence of other interactions) indicates that driving duration group

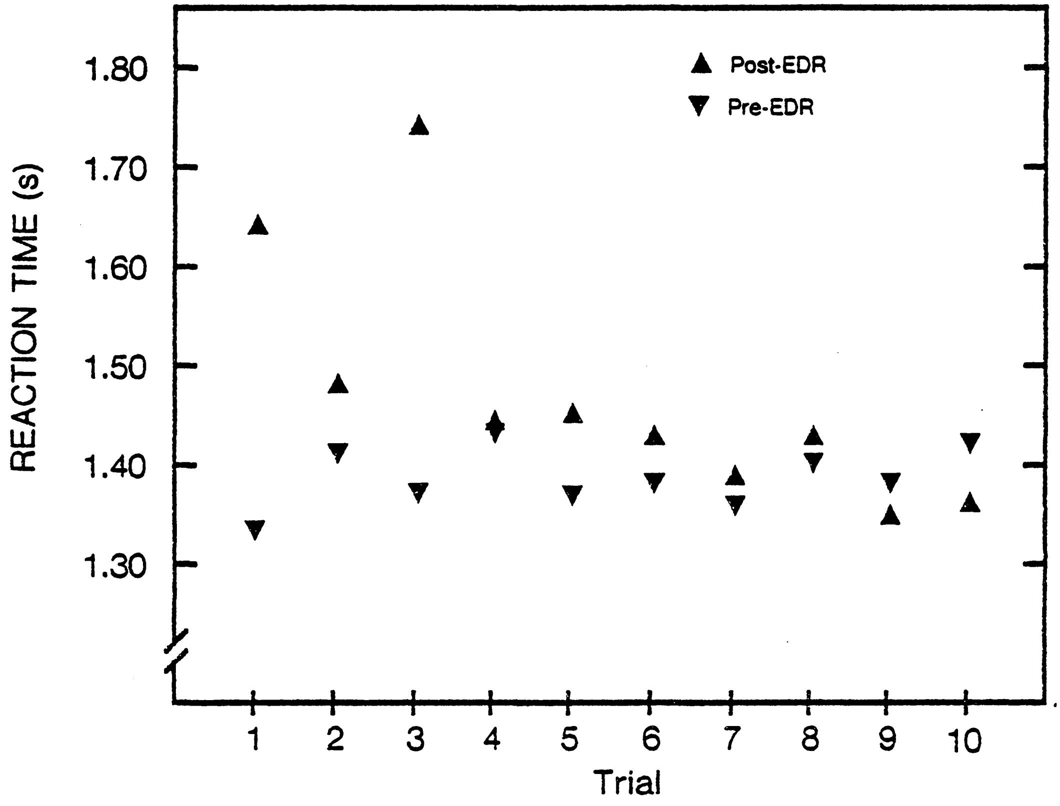

Figure 32. Mean brake response times by run and trial-Experiment I.

**TABLE 22.** Analysis of Variance Summary for Mean Lateral Deviation Error Scores of Experiment II

| SOURCE | df | MS | F | p |
|---|---|---|---|---|
| BETWEEN SUBJECTS | | | | |
| Steering Gain (G) | 1 | 253.898 | .11 | >.10 |
| Duration (D) | 2 | 4967.748 | 2.10 | >.10 |
| G x D | 2 | 1391.226 | .59 | >.10 |
| Subjects/GD (S/GD) | 18 | 2369.586 | | |
| WITHIN SUBJECTS | | | | |
| Run (R) | 1 | 9115.019 | 10.85 | .004 |
| Trial (T) | 9 | 918.333 | 1.84 | >.05 |
| G x R | 1 | 262.109 | .31 | >.10 |
| D x R | 2 | 2213.085 | 2.63 | >.05 |
| G x T | 9 | 800.336 | 1.60 | >.10 |
| D x T | 18 | 912.839 | 1.83 | .026 |
| R x T | 9 | 346.346 | .76 | >.10 |
| G x D x R | 2 | 174.395 | .21 | >.10 |
| G x D x T | 18 | 595.701 | 1.19 | >.10 |
| G x R x T | 9 | 471.995 | 1.04 | >.10 |
| D x R x T | 18 | 378.190 | .83 | >.10 |
| R x S/GD | 18 | 840.225 | | |
| T x S/GD | 162 | 500.125 | | |
| R x T x G x D | 18 | 560.297 | 1.23 | >.10 |
| R x T x S/GD | 162 | 455.814 | | |
| Total | 479 | | | |

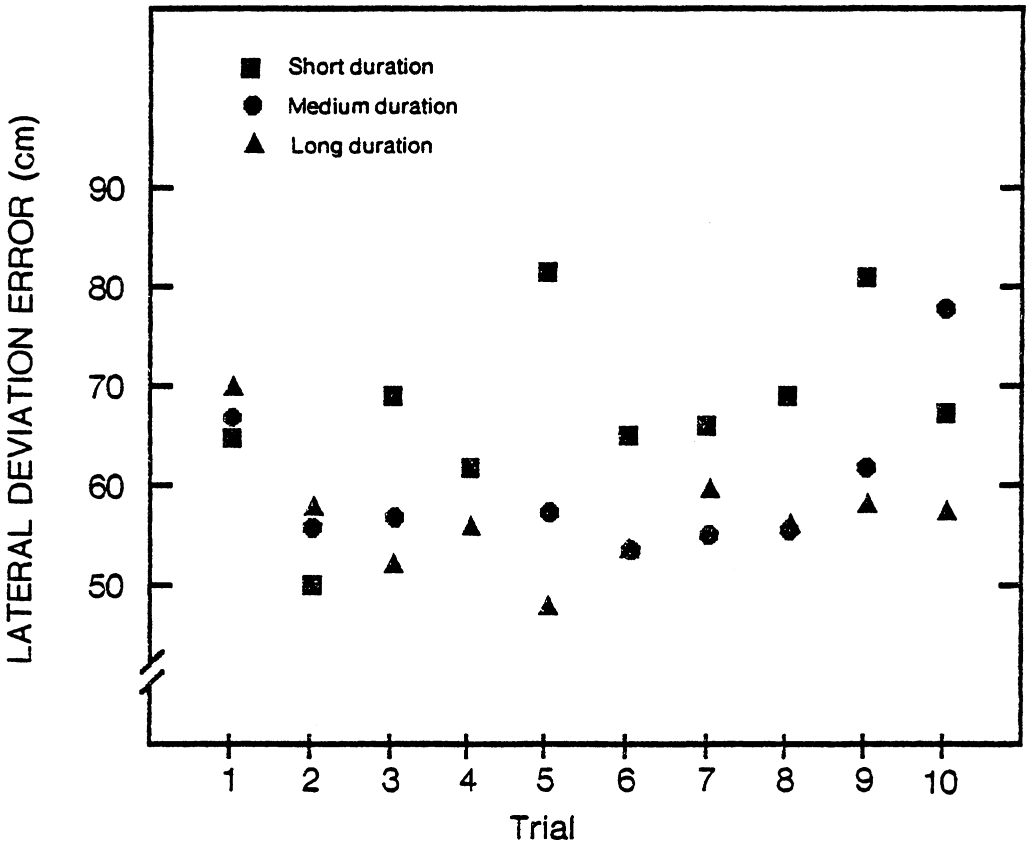

Figure 33. Mean lateral deviation error by duration and trial- Experiment II.

performances were affected differentially by trials across both pre- and post-EDR runs. A simple effects analysis (analysis of variance) indicated that significant group mean differences existed for trials 5, 9, and 10 ($p < 0.05$).

## Experiment III - Emergency Response Data

Missing observations. A number of trials during choice reaction time runs of Experiment III were discarded when an inappropriate response was made (e.g., brake application when a steering response should have been made), or when the time to respond exceeded 2.00 s. In the analysis of the data for each trial, cell means were substituted for discarded observations, and the degrees of freedom for the particular factor was reduced for each substitution. The number (and approximate percentage) of missing observations for brake and steering responses by run are shown in Table 23.

Comparison of simple and choice reaction times. The reaction times for simple and choice reaction times were analyzed in separate three-way ANOVAs (duration x runs x trials) for brake and steering response times.

The analysis of variance for brake reaction times, shown in Table 24, revealed a significant effect due to run, $F(1,9) = 68.77$, $p = .0001$, and trials, $F(9,81) = 3.84$, $p = .0005$. All other effects were not significant ($p > 0.05$). The brake reaction time run main effect implies that the mean for CRT baseline runs ($\bar{X} = 0.892$ s) was significantly longer than the mean for simple reaction time runs ($\bar{X} = 0.704$ s). A Newman-Keuls analysis of the significant trials effect revealed that trial 1 differed significantly from all subsequent trials ($p < 0.01$). No other trial mean differences were significant ($p > 0.05$).

The ANOVA for steering responses, shown in Table 25, also indicated

**TABLE 23.** Number (and Percentage) of Missing Observations for Brake and Steering Responses by Run - Experiment III

| Run | Brake Responses | | Steering Responses | |
|---|---|---|---|---|
| Simple Reaction | 0 | | 1 | (1%) |
| Day 1 Baseline (Choice reaction time - CRT) | 9 | (8%) | 7 | (6%) |
| Pre-EDR CRT | 13 | (11%) | 3 | (3%) |
| Post-EDR CRT | 10 | (8%) | 2 | (2%) |

**TABLE 24.** Analysis of Variance Summary for Brake Response Times of Simple and Day 1 CRT Baseline Runs of Experiment III

| SOURCE | df | MS | F | p |
|---|---|---|---|---|
| BETWEEN SUBJECTS | | | | |
| Duration (D) | 2 | .0169 | .22 | >.10 |
| Subjects/D (S/D) | 9 | .0780 | | |
| WITHIN SUBJECTS | | | | |
| Run (R) | 1 | 2.1414 | 68.77 | .0001 |
| Trials (T) | 9 | .3670 | 3.84 | .0005 |
| D x R | 2 | .0017 | .05 | >.10 |
| D x T | 18 | .0133 | 1.40 | >.10 |
| R x T | 9 | .0207 | 1.75 | |
| R x S/D | 9 | .0311 | | |
| T x S/D | 81 | .0096 | | |
| D x R x T | 18 | .0145 | 1.22 | >.10 |
| R x T x S/D | 72[a] | .0118 | | |
| Total | 230 | | | |

a. Degrees of freedom were adjusted for missing observations.

**TABLE 25.** Analysis of Variance Summary for Steering Response Times of Simple and Day 1 CRT Baseline Runs of Experiment III

| SOURCE | df | MS | F | p |
|---|---|---|---|---|
| BETWEEN SUBJECTS | | | | |
| Duration (D) | 2 | .1091 | .59 | >.10 |
| Subjects/D (S/D) | 9 | .0780 | | |
| WITHIN SUBJECTS | | | | |
| Run (R) | 1 | 2.6021 | 46.13 | .0001 |
| Trials (T) | 9 | .0373 | 2.11 | .0373 |
| D x R | 2 | .0036 | .06 | >.10 |
| D x T | 18 | .0140 | .79 | >.10 |
| R x T | 9 | .0245 | 1.37 | >.10 |
| R x S/D | 9 | .0564 | | |
| T x S/D | 81 | .0176 | | |
| D x R x T | 18 | .0113 | .64 | >.10 |
| R x T x S/D | 73[a] | .0179 | | |
| Total | 231 | | | |

[a] Degrees of freedom adjusted for missing observations.

significant effects due to run, F(1,9) = 46.13, p = .0001, and trials, F(9,81) = 2.11, p = .0373. All other effects were not significant ($p > 0.05$). The main effect due to run implies that the mean reaction time for the CRT Day 1 baseline run ($\bar{X}$ = .804 s) was significantly longer than the mean reaction time for simple reaction time runs ($\bar{X}$ = 0.596 s). A Newman-Keuls analysis on the trials effect failed to detect any significant differences among trials $p > 0.05$).

*Comparison of pre- and post-EDR runs.* The reaction times for pre- and post-EDR runs were analyzed in separate three-way ANOVAs (duration x run x trials) for brake and steering response times. Table 26 shows that the ANOVA for brake reaction times revealed significant effects due to run, F(1,9) = 7.38, p = .024, and trials, F(9,81) = 3.97, p = .0003. No other effects were significant ($p > 0.05$). The main effect due to run implies that the mean brake response time was significantly longer in post-EDR runs ($\bar{X}$ = 0.944 s) than in pre-EDR runs ($\bar{X}$ = 0.891 s). A Newman-Keuls comparison of trial means for brake reaction times (shown in Figure 34), revealed that the time for trial 2 was significantly longer than for all other trials except trial 1 ($p < 0.01$). No other trial mean differences were significant.

The ANOVA for steering reaction times, shown in Table 27, revealed significant effects due to run, F(1,9) = 6.05, p = .0362, and trials, F(9,81) = 3.93, p = 0.0004). The significant run effect indicates that the mean steering response time for the post-EDR run ($\bar{X}$ = 0.839 s) was significantly longer than that of the pre-EDR run ($\bar{X}$ = 0.778 s). A Newman-Keuls test of the trials effect indicated that the mean reaction time for trial 1 was longer than for trials 4, 8, 9, and 10 ($p < 0.05$); and also that the mean reaction time for trial 7 was significantly longer than for trial 8 ($p < 0.05$). No other mean differences were significant ($p > .05$).

**TABLE 26.** Analysis of Variance Summary for Brake Reaction Times of Pre- and Post-EDR Runs of Experiment III

| SOURCE | df | MS | F | p |
|---|---|---|---|---|
| BETWEEN SUBJECTS | | | | |
| Duration (D) | 2 | .0045 | .07 | > .10 |
| Subjects/D (S/D) | 9 | .0645 | | |
| WITHIN SUBJECTS | | | | |
| Run (R) | 1 | .1739 | 7.38 | .0237 |
| Trials (T) | 9 | .0477 | 3.97 | .0003 |
| D x R | 2 | .0229 | .97 | > .10 |
| D x T | 18 | .0163 | 1.35 | > .10 |
| R x T | 9 | .0232 | 1.37 | > .10 |
| R x S/D | 9 | .0236 | | |
| T x S/D | 81 | .0120 | | |
| D x R x T | 18 | .0237 | 1.40 | > .10 |
| R x T x S/D | 58[a] | .0237 | | |
| Total | 216 | | | |

[a] Degrees of freedom were adjusted for missing observations.

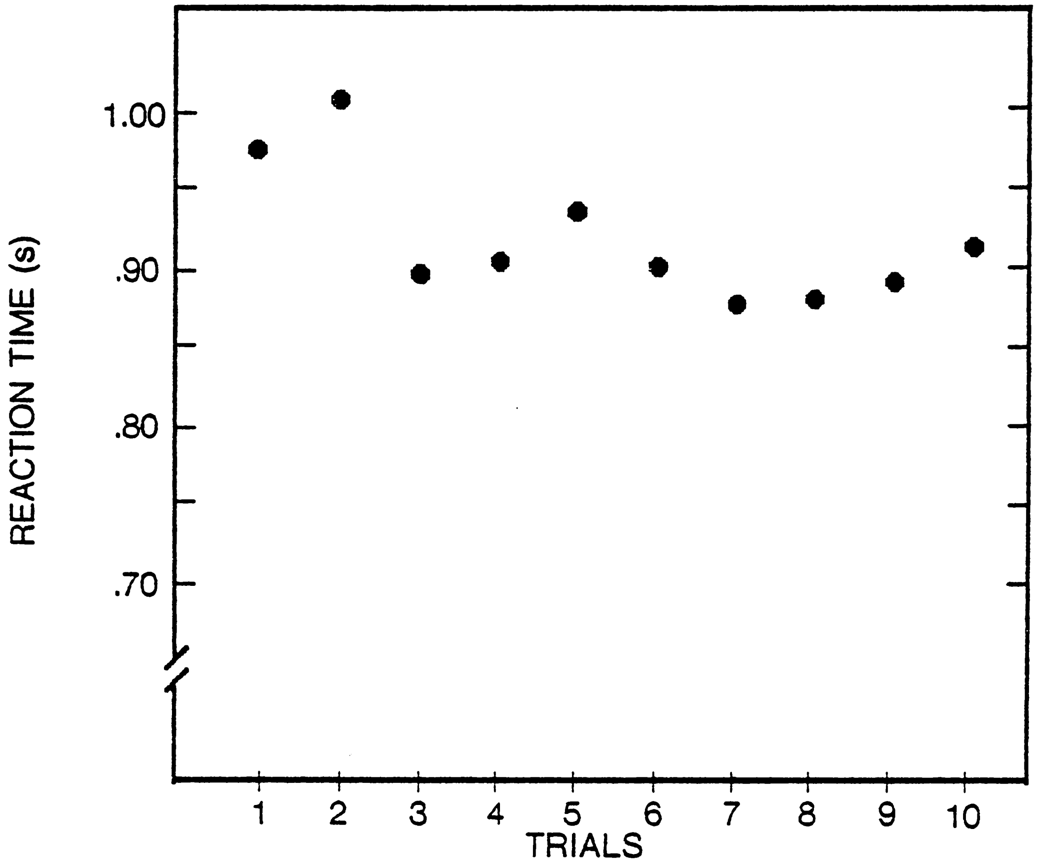

Figure 34. Mean brake response times by trial-Experiment III.

**TABLE 27.** Analysis of Variance Summary for Steering Reaction Times of Pre-and Post-EDR Runs of Experiment III

| SOURCE | df | MS | F | p |
|---|---|---|---|---|
| BETWEEN SUBJECTS | | | | |
| Duration (D) | 2 | .1812 | .69 | >.10 |
| Subjects/D (S/D) | 9 | .2615 | | |
| WITHIN SUBJECTS | | | | |
| Run (R) | 1 | .2190 | 6.05 | .0362 |
| Trials (T) | 9 | .0781 | 3.93 | .0004 |
| D x R | 2 | .0393 | 1.09 | >.10 |
| D x T | 18 | .0138 | .70 | >.10 |
| R x T | 9 | .0311 | 1.44 | >.10 |
| R x S/D | 9 | .0362 | | |
| T x S/D | 81 | .0199 | | |
| D x R x T | 18 | .0199 | .92 | >.10 |
| R x T x S/D | 76[a] | .0216 | | |
| Total | 234 | | | |

a. Degrees of freedom were adjusted for missing observations.

Mean steering reaction times by trial are shown in Figure 35.

## Correlation Between EDR Variables and Emergency Performance Data

To determine the relationship between EDR variable changes and emergency performance decrements, EDR variable difference scores were correlated with emergency performance difference scores. EDR variables difference scores were calculated for Experiments I, II (normal gain), and III by subtracting the score of the first 10-min period from the score of the last 10-min period for each subject. Emergency performance difference scores for subjects in Experiments I and II were determined by subtracting the performance score of the first trial of post-EDR emergency trials from the mean of the pre-EDR emergency trials for each subject. For Experiment III, the first emergency trials (whose presentation orders were counterbalanced among subjects) were subtracted from their respective pre-EDR means.

The difference scores were then correlated with a Pearson product moment correlation. The computed correlation coefficients, shown in Table 28, ranged in magnitude from 0.01 to 0.58 (absolute value). STREVLRG was the only variable yielding a correlation significantly greater than zero ($p < 0.05$).

## Subjective Measures

Subjective ratings of fatigue for all experiments were combined in a two-way ANOVA with factors "duration" and "test" (post-hoc ratings of subjective fatigue assessment for before and after the extended duration runs). As shown in Table 29, significant effects were found due to test, $F(1,45)= 163.64$, $p= 0.0001$, and the duration x test interaction, $F(2,45)= 8.36$, $p= 0.0008$. A Newman-Keuls analysis of the duration x test interaction (Figure 36) revealed that: (a) mean

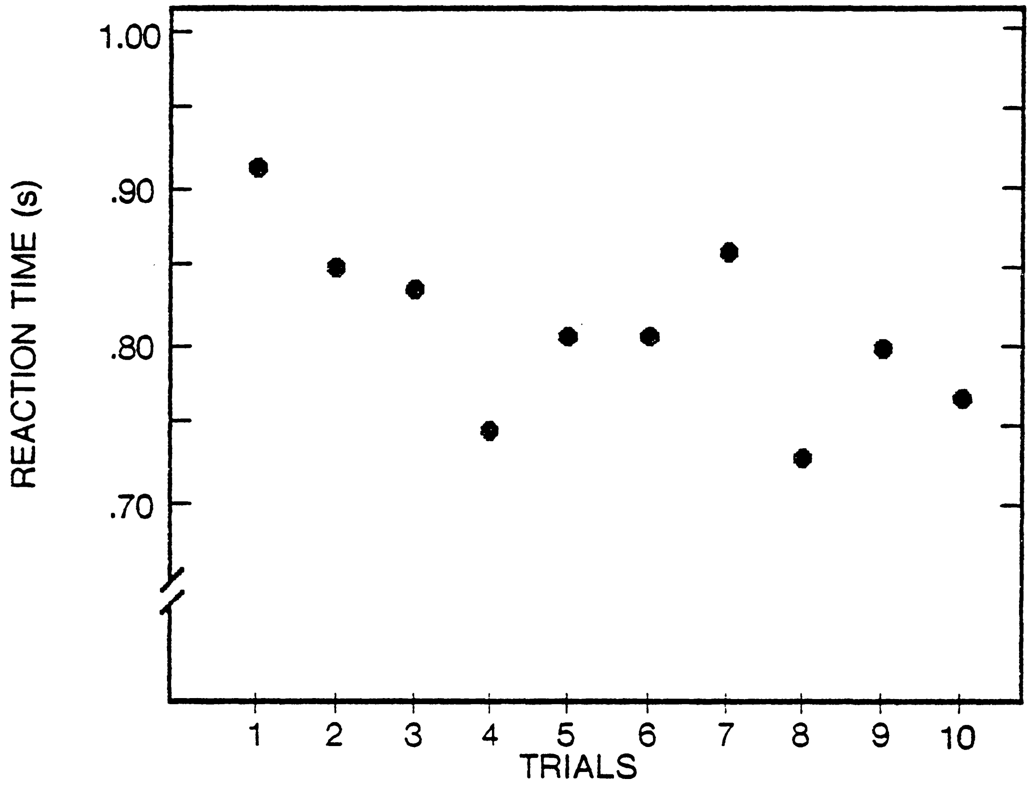

Figure 35. Mean steering response times by trial-Experiment III.

**TABLE 28.** Pearson Product Moment Correlations Between EDR Variable and Emergency Performance Changes Over Time

| EDR VARIABLE | Exp. I | Correlation Coefficients (r)<br>Exp. II (Normal gain) | Exp. III |
|---|---|---|---|
| BDYMOVES | .46 | .26 | .06 |
| HRTINTSD | .27 | -.16 | -.18 |
| LATDEVSD | .05 | -.27 | .42 |
| STEERSD | .39 | .37 | .18 |
| STREVLRG | .37 | .58* | -.46 |
| STREVSML | -.12 | -.14 | -.16 |
| STRREVMN | .48 | .50 | -.03 |
| YAWDEVSD | .21 | .01 | .57 |
| YAWREVII | .38 | .21 | .24 |

* p <.05

**TABLE 29.** Analysis of Variance Summary for Subjective Fatigue Ratings of All Experimental Groups

| SOURCE | df | MS | F | p |
|---|---|---|---|---|
| Duration (D) | 2 | .0495 | .03 | >.10 |
| Test (T) | 1 | 119.2604 | 163.64 | .0001 |
| Subject/D (S/D) | 45 | 1.9087 | | |
| D x T | 2 | 6.0964 | 8.36 | .0008 |
| T x S/D | 45 | .7288 | | |
| Total | 95 | | | |

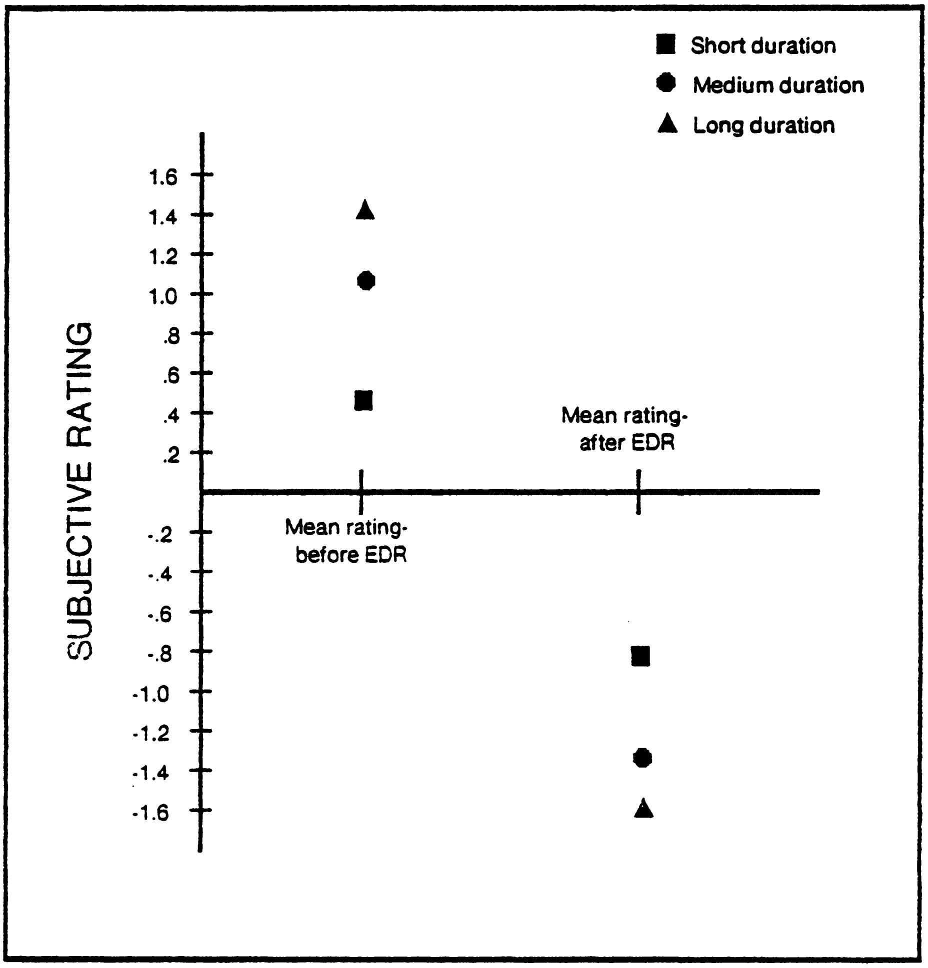

Figure 36. Mean subjective (post-hoc) ratings of "before" and "after" extended driving runs by duration for all experiments.

subjective rating scores for all duration conditions were significantly lower (less "alert") for after-EDR ratings than for before-EDR ratings ($p < 0.01$); (b) mean before-EDR rating scores were significantly higher (more "alert") for the medium and long duration conditions than for the short duration condition ($p < 0.01$); (c) mean after-EDR rating scores were significantly lower for medium and long duration conditions than for the short duration condition ($p < 0.01$); and (d) no significant differences were found among corresponding before- and after-EDR ratings of medium and long duration groups.

These results indicate first, that drivers, regardless of how long a period they had driven, rated themselves to be significantly less alert after the extended duration run than before the run began. Secondly, drivers in medium and long duration groups indicated themselves to be equally alert (equally tired) in corresponding ratings for before- and after-EDR periods. Lastly, drivers in the short duration group rated themselves to be less alert before, and more alert after the drive than drivers in the medium and long duration groups.

# DISCUSSION

## Effects of Driving Time on EDR Variables

Vehicle control. Significant changes among lane-keeping performance variables over time indicate that vehicle control becomes increasingly erratic as the driving period continues. This tendency was observed in drivers required to drive for periods of 60 min or longer (as in medium and long duration groups), as indicated by significant linear increases in lateral position standard deviation and yaw angle standard deviation (LATDEVSD and YAWDEVSD) over successive time periods. Similar changes were observed in yaw reversals greater than 2 deg (YAWREVII), which indicated significant increases among successive time periods in the long duration driving groups. The tendency for increased lateral position variability over time is consistent with the findings of Heimstra (1970), Snook and Dolliver (1976), Sussman et al. (1971), and others. Increases in yaw angle variability and number of large yaw reversals over time appear to be unprecedented in the driving fatigue literature.

Significant changes in lane keeping performance were evident in the long driving duration group of Experiment II, but not in the medium or short duration groups. The fact that vehicle lane-keeping performance variables did exhibit significant changes over time in the medium duration groups of the combined EDR analysis and in the steering variables of Experiment II (medium duration) indicates that these vehicle control variables are perhaps less sensitive to time changes than are other variables (such as steering wheel movements).

Steering behavior. The analysis of performance variables indicated that the increasing tendency to drift within the driving lane was accompanied by a general tendency of drivers to employ larger, more variable steering inputs after 60 minutes or more of continuous driving. These changes in steering behaviors are comprised of an apparent shift from frequent, small amplitude steering inputs (early in the driving period) to larger, less frequent steering inputs as the driving period progresses. This shift in steering strategy was reflected by linear increases in steering wheel angle standard deviation and steering reversal angle means, and more specifically by significant linear increases in the frequency of large steering reversals and concurrent decreases in the frequency of small steering reversals over driving periods of 60 min and longer.

Reconciling the present data with the results of previous investigations is problematic. The first problem is that the results of previous investigations often do not agree with each other. For example, while several investigators have shown that steering reversals tend to decrease in frequency over time (e.g., Snook and Dolliver, 1976; Sussman and Knight, 1971; Sussman and Morris, 1970), others have found that steering reversals tend to increase in frequency over time (e.g., Safford and Rockwell, 1967). Still others have shown that steering reversal frequencies can either increase or decrease over time (Brown et al., 1967), or are too highly variable to be reliable (Harris and Mackie, 1972). Further complicating any comparisons is the result of the present investigation, which indicated that small and large reversal frequencies tend to exhibit significant, opposing trends over time.

Part of the conflict of results found among various investigations may be due to the inadequacy of steering reversal measures used by most investigators to distinguish different aspects of steering behaviors (such as found in the present investigation). How measurement reliability and validity could be affected will be discussed later.

Another source of conflict of data could lie in the differences of vehicle steering gain used in present and previous investigations. As will be shown later, steering gain has a significant effect on steering behavior of drivers, the overall effect being that lower steering gain causes drivers to exhibit larger and more variable steering than does higher steering gain. At least part of this overall effect could be due to a shift in distribution caused by an increase in the magnitude of (and drivers' use and perception of) "small" steering reversals. If a proportion of these small reversals was sufficiently large to exceed the detection threshold employed, and these small reversals follow similar trends of the small reversals in the present investigation (i.e., decreasing frequencies over time), it is likely that the negative trend of these reversals would completely mask (or reverse) the increasing trends of large steering reversals detected.

Conclusions on steering behavior and measurement of driving time effects. The present results reveal that drivers exhibit increasingly variable steering movements during extended driving periods of 60 min and longer. The increase in variability, measured by increases in steering wheel angle standard deviation and mean steering reversal angle, reflect dynamic changes in steering behavior characterized by a concomitant increase in the frequency of large steering inputs and a decrease in the frequency of small steering inputs.

The steering reversal measurement technique employed in the present investigation (which detected and categorized all steering wheel movements) and the follow-up dichotomous categorization provided information on the dynamic changes in steering behavior which have not been previously reported. Although results derived from the dichotomous classification scheme are illuminating, the technique may lack the flexibility needed to measure more dynamic changes or to allow data to be generalized to other environment-vehicle-driver combinations. In view of the present data, it appears that future investigators would be prudent to detect and capture all steering inputs, and assess changes in the continuous frequency distribution, or in discrete multiple categories (e.g., histograms), of the continuous distribution. More parsimonious measures, such as mean steering reversal angle, could be equally effective indicators of time related changes; however, care should be taken in attempting to generalize such results because of the inability of such measures to distinguish among certain types of changes. For example, an increase in mean steering reversal angle could be the result of an increase in the frequency of large movements, a decrease in the frequency of small movements, or a combination of both of these.

Steering reversal threshold measures used in previous investigations have several deficiencies which could jeopardize the reliability and validity of the outcome of such measures. Among possible outcomes are the non-detection of steering movements smaller than the selected threshold, mutual opposition or masking of trends of differing classes of steering movements, or increases in variability (or shifts in the distribution) caused by differences in steering gain or other vehicular characteristics.

Heart beat and body movements. The analysis of the combined EDR data indicated significant changes of the heart beat variability and gross body movements, two variables which have been purported to be indicators of changes in a driver's state of alertness.

In drivers required to drive for 60 min or longer, heart beat data revealed that drivers' heart rhythms exhibited significant linear increases in beat-to-beat interval variability over time. These results agree with the findings of a number of investigators (e.g., Riemersma, Sanders, Wildervanck, and Gaillard, 1977; Sugarman and Cozad, 1972), and seem to support the notion that increases in heart beat variability reflect time changes presumed to be caused by mechanisms responsible for reductions in alertness.

The frequency of drivers' upper torso movements were also found to increase (linearly) in drivers who drove continuously for 60 min and longer. These changes are consistent with observations made by O'Hanlon and Kelley (1977), but appear not to be as dramatic as changes observed in O'Hanlon and Kelley's subjects (who apparently exhibited sudden, significant changes in speed control, lane drift performance, and steering behavior when "gross, whole body movements" were observed). Whether these movements are signs of the subjects' attempt to "shake off feelings of fatigue", as the authors speculated, or merely reflect drivers' attempts to restore location circulation, or are due to certain involuntary mechanisms is not immediately apparent.

Heart beat variability and gross body movement frequency in Experiment II exhibited increases in the long duration condition only. The fact that these variables indicated significant changes over time in the medium duration condition of the combined EDR analysis, but not in the medium duration condition of Experiment II (which employed fewer degees of freedom) implies that heart beat variability and gross body movement frequency data are somewhat less sensitive to changes due to driving time than are other variables.

Trends of EDR variables over time. EDR variables indicated that drivers exhibited highly significant changes over time. The $R^2$ values obtained for regressions on group means, which averaged .83 (range, .56 to .98) for the medium duration group and .94 (range, .90 to .97) for the long duration group, indicate that group mean trends can be accurately modeled by linear trends over time. In contrast, $R^2$ values obtained for regressions employing individual data indicate that the terms in the regression equations account for relatively small portions of the variances. For example, the highest $R^2$ value obtained among EDR variables in medium or long duration groups was .218 (for STEERSD, medium duration). Most others, however, were much lower, averaging .090 for the medium duration group and .137 for the long duration group.

From these results it is evident that although the regression equations of Tables 6 and 7 are good models of group behavior, they are relatively poor models of individual behavior. Consequently, it would appear that attempts to predict individual behavior from mean trends would be unreliable, and hence would have questionable utility. The major source of the problem appears to exist in the high variablity exhibited among subjects. Figures 37, 38 and 39, which present examples of EDR data for HRTINTSD, LATDEVSD, and STEERSD of three subjects, illustrate that considerable differences do exist both in terms of (initial) baseline levels, and in the slopes of the EDR variables over time.

Individual differences in susceptibility to driving time effects. Individual differences in the slope of EDR variable changes over time indicate that individual differences exist in the susceptibility of drivers to the effects of prolonged driving time. In the present study, for example, two drivers were removed from the study after exhibiting obvious signs of falling asleep (in one case, the subject fell asleep in less than 30 minutes of driving time). In contrast, other investigators (Safford and Rockwell, 1967) have observed that drivers are

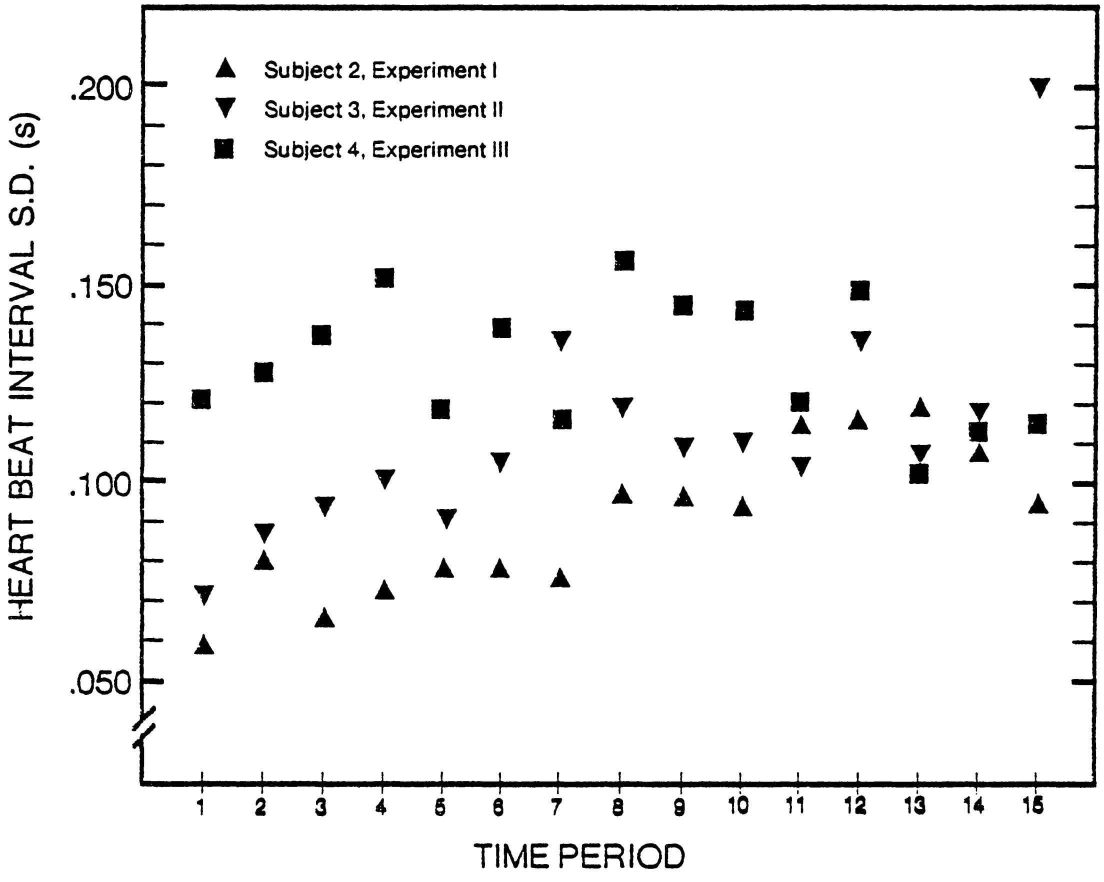

Figure 37. Mean heart beat interval s.d. (HRTINTSD), by time period, for three subjects- Long duration.

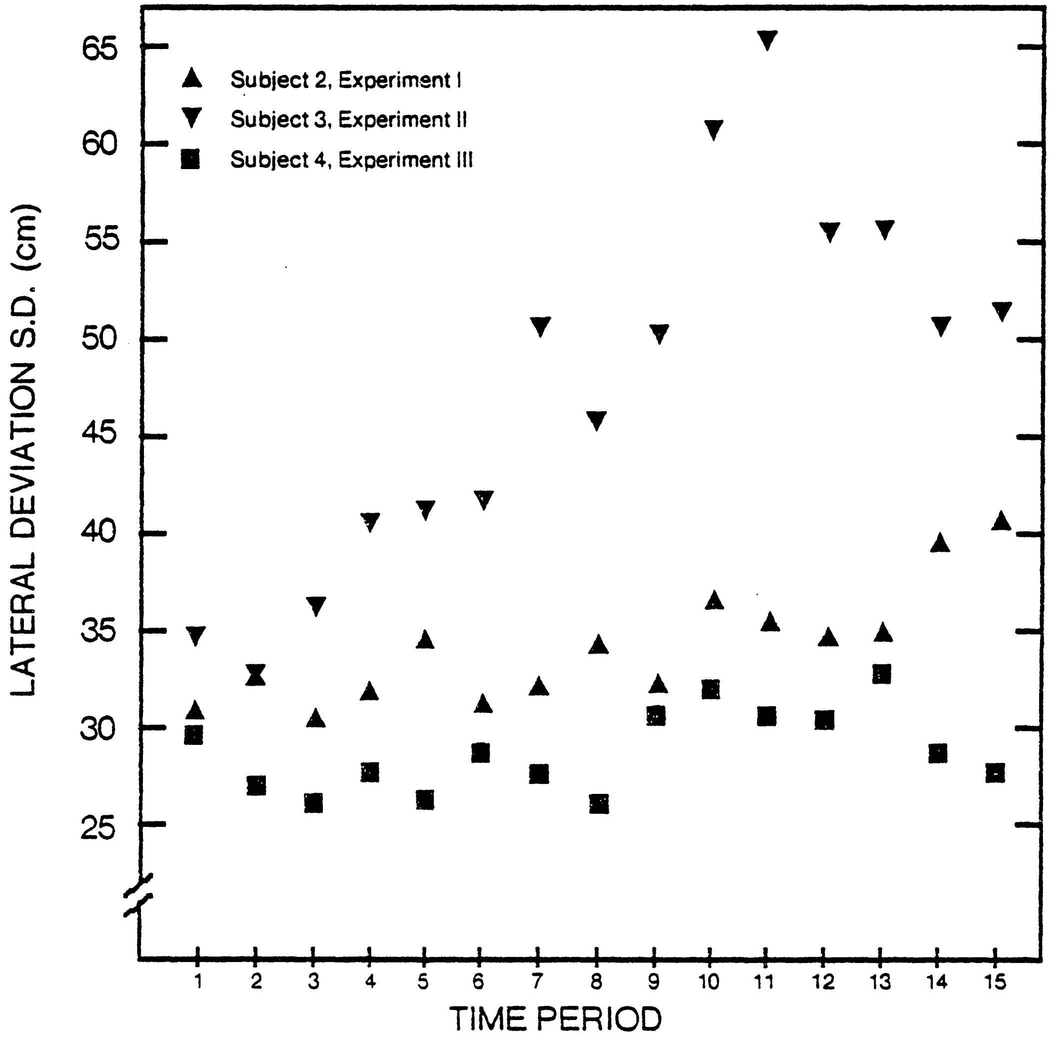

Figure 38. Mean lateral deviation s.d. (LATDEVSD), by time period, for three subjects- Long duration.

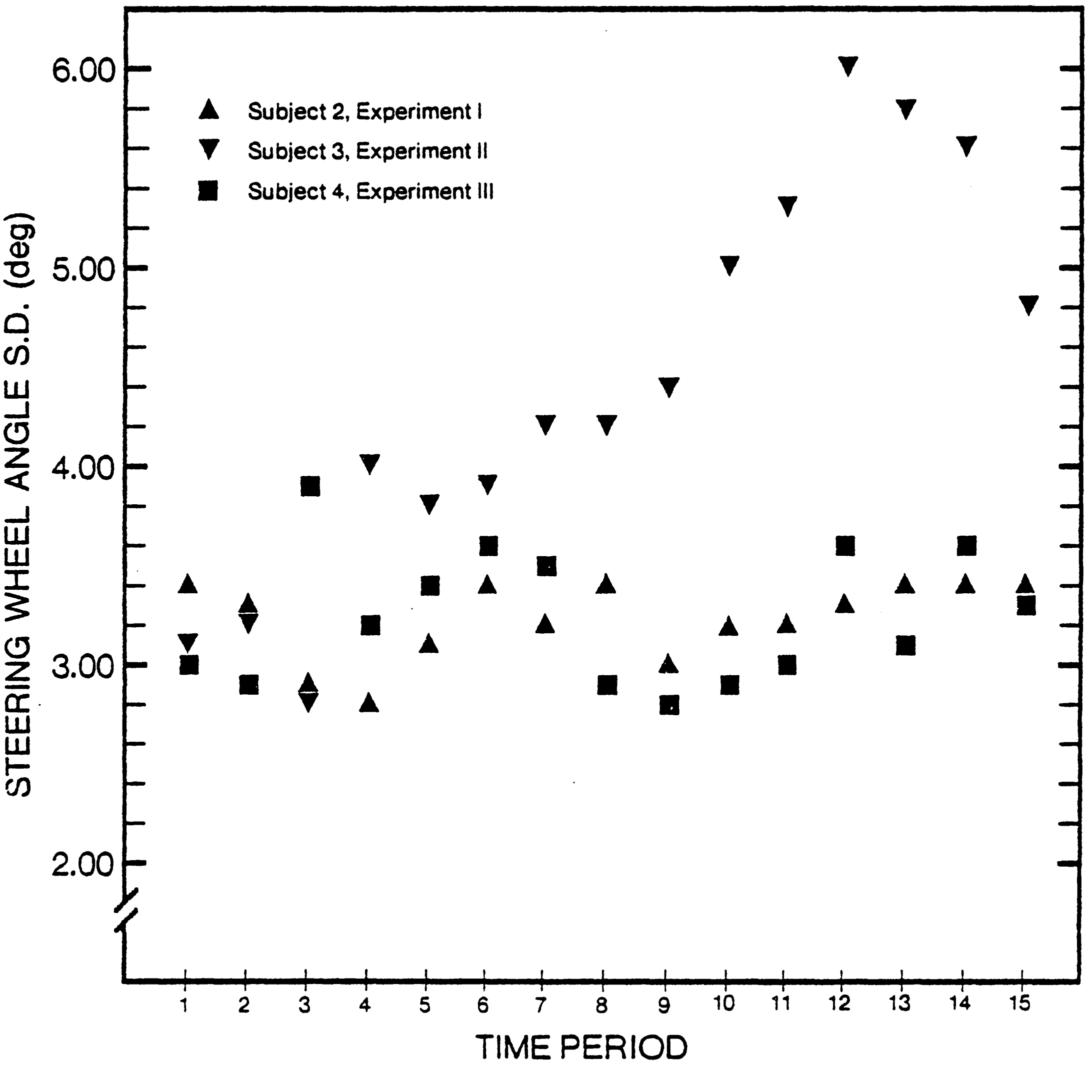

Figure 39. Mean steering angle s.d. (STEERSD), by time period, for three subjects- Long duration.

able to operate a vehicle effectively for periods up to 24 hours (with some rest breaks).

The factors differentially affecting drivers' susceptibility to prolonged driving effects could include motivation, basal metabolic rate, amount of sleep deprivation, amount of drug intake, etc. In addition, certain environmental factors could affect the drivers' attentional workload, which could in turn change the drivers' level of arousal. External variables which could affect the drivers' alertness could include terrain, characteristics of roadway construction, weather, traffic conditions, temperature, roadway noise, rhythm of background music, etc. The present investigation employed a carefully controlled driving environment in which environmental stimuli were, to the extent possible, held constant and minimal. Even with careful control of the environment, however, it is apparent that driver behavior over time can differ substantially among individuals. Consequently, it appears that differences observed in the susceptibility to driving time effects are due to the action of a set of, as yet, unspecified individual factors.

From the present results, it would have to be concluded, as others have (e.g., Harris and Mackie, 1972; Safford and Rockwell, 1967), that a measure of group performance changes over time cannot be used reliably as a universal measure of fatigue. An attempt to use such a group measure as a criterion for "fatigue" would almost certainly result in non-detection of changes for a portion of the population and a large number of "false alarms" for another portion. It is likely, however, that factors could be identified that would predict individual performance levels (both in terms of baseline levels and change rates). As a consequence, the ability of regression equations to account for significant proportions of the population variances would be greatly enhanced.

An alternative approach. Another approach would acknowledge the fact that drivers exhibit individualistic (perhaps unique) patterns of change in EDR-type variables over time, and would use "time change profiles" as a basis for fatigue assessment. These time change profiles would represent "normal" (non-fatigued) changes in continuously monitored variables that are known to vary over over time. A profile would be established for a given variable by monitoring that variable during several "calibration" runs, and forming a composite score of runs (e.g. means) over time. This profile would serve as the basis of comparison for detecting out-of-tolerance states for each variable employed. Such a system would be analagous to "speaker-dependent" speech recognition systems in which incoming word patterns are compared with a person's word pattern "template" (See Moshier, Osborn, Baker, and Baker, 1980). However, unlike speech recognition devices which detect template matches, the fatigue detection system would detect mismatches.

Out-of-tolerance states for several variables could be combined logically, or differentially weighted to form a composite criterion. With such a criterion, it would seem reasonable that accurate and reliable fatigue detection could be achieved.

## Effects on Steering Gain on EDR Variables (Experiment II)

Steering gain and vehicle control. The results of Experiment II EDR variables indicate that steering gain has no effect on lane-keeping performance (measured by lateral deviation standard deviation, yaw angle standard deviation, and yaw reversal frequency). In addition, the time trends of lane-keeping performance appear not to be affected by steering gain differences (indicated by the lack of significance of the steering gain x time interaction).

Effects of steering gain on continuous steering behavior. The results of Experiment II indicate that steering gain has a direct and significant effect on drivers' steering behaviors (except in the frequency of small steering reversals). The general tendency was for drivers in the high steering gain group to employ smaller steering inputs than drivers in the normal group throughout extended duration runs. This tendency was evident in the significantly lower values in the high steering gain group (relative to the normal steering gain group) for steering angle standard deviation, mean steering wheel reversal angle, and large steering reversals frequency. The non-significance of the steering gain x time interaction for all steering behavior measures indicates that the magnitudes of changes due to time (i.e., in the changes among successive time periods) were not differentially affected by steering gain.

The non-significance of steering gain effects on the frequency of small (0.5-2.0 deg) steering reversals indicates that there exists a class of steering behaviors which remain relatively invariant among drivers (and probably within individual drivers), even with large differences in vehicle steering characteristics. One possible reason for this is that such reversals are partly the result of small involuntary muscular tremors resulting from normal physiological functions. Another possibility is that such actions reflect a sampling strategy of drivers, wherein drivers continually assess the responsiveness of the driver-vehicle system and employ the information for ongoing vehicle control actions.

Implications of steering gain effects on EDR variables over time. The fact that steering gain had significant effects on all but one of the steering variables, and not on any of the vehicle control variables implies that drivers adopt equivalent "lane-keeping" performance levels, even though the steering responses required to maintain equivalent lane tracking performance levels differed significantly. In the long duration condition of Experiment II, the significance of the time effect and the non-significance of the steering gain x time interaction implies that the equivalency of lane tracking performance among drivers in the two groups was maintained even though the actual performance levels of both groups changed significantly over successive time periods.

Although previous studies have shown concomitant changes in vehicle control performance and steering behavior similiar to those shown in the present investigation, the causal relationships of such changes has remained unclear. In viewing concomitant vehicle control/steering behavior changes over time, two general hypothetical modes of change seem possible. The first is that extended driving time results in physiological changes in drivers which, in turn, affect motor responses comprising vehicle control (particulary steering wheel) inputs. Increasingly variable motor responses (exhibited as increasingly variable steering inputs) would result in corresponding changes in vehicle lateral and yaw deviations. A second hypothesis is that driving time results in changes in the driver's perception (or judgment) of acceptable criterion levels for lane-keeping performance (lateral position deviation and yaw angle deviation). The less stringent lane-keeping criterion level would allow the driver to make corresponding changes in his steering behavior. If the former hypothesis were true, changes in steering behavior associated with driving time (being essentially "open-loop") would likely be equivalent among drivers in the two steering gain groups. Although steering inputs among steering gain groups would be equivalent, the resultant vehicle responses in yaw and lateral deviation would differ in proportion to the difference in levels of steering gain. Contrary to the latter expectations, Experiment II results indicate that the drivers in the normal and high steering gain groups maintained equivalent lane-keeping performance levels, but significantly different steering behaviors. This relationship (equivalent lane-keeping performance with differential steering behaviors) indicates that drivers adopt a level of lane-keeping performance which is subject to changes over time but is not directly affected by changes in steering gain. The parallel changes in steering behavior among the steering gain groups over time (indicated by significant main effects due to time, steering gain, and no steering gain x time interaction) indicates that drivers modify their steering inputs to maintain their adopted lane-keeping performance levels.

From the above discussion, it appears that driving time results in a modification in the amount of lateral and yaw deviations tolerated by the driver. This could either be due to a change in the driver's perception of signs of vehicle drift, or simply a relaxation of the driver's criterion for acceptable lane tracking performance levels. Either alternative would allow the driver to devote less attention to steering inputs and serve as a mechanism through which the driver could adapt to reductions in workload capacity presumed to be caused by time-related fatigue mechanisms.

### Subjective Ratings of Alertness

The results of drivers' post-hoc ratings of alertness revealed that drivers among groups were able to reliably judge changes in alertness associated with extended driving periods. The mean ratings for "before" and "after" EDR ratings of short, medium, and long duration conditions indicate that subjects judged themselves to be less "alert" after having driven for a prolonged period than before the period began. Subjects who had driven for periods of 60 and 150 min (medium and long duration conditions) judged themselves to be significantly less alert than subjects who had driven for 30 minutes (short duration condition). The lack of significant differences of EDR ratings among medium and long duration conditions indicates an asymtotic driving time effect on subjective ratings. This asymtotic effect, which was exhibited after 60 min of continuous driving, indicates either that (a) drivers are maximally "fatigued" after 60 min of continuous driving, (b) the ability of drivers to judge any further changes in alertness reliably is diminished over driving time, or (c) subjective changes are sufficiently non-linear so that much longer driving times are needed to elicit "fatigued" ratings.

The significance of the duration x time interaction reflects an apparent tendency of subjects in the medium and long duration conditions to rate themselves to be more "alert" for before-EDR ratings and increasingly "tired" for after-EDR ratings than subjects in the short duration condition. Although the after-EDR differences are attributable to alertness decrements brought on by increments in driving time, the differences in before-EDR ratings are not obvious. Assuming that no inherent biases were present in subject sampling or in experimental conditions, the most plausible explanation for the differences among groups in before-EDR ratings is that subjects tend to anchor their initial ratings (the "before-EDR" ratings) according to their perceived magnitude of change in alertness over the driving period. That is, subjects in the longer driving duration conditions (60 min and longer) judge themselves to be more alert than subjects judged themselves to be in the short duration condition in an effort to increase the dynamic range of alertness level changes. Although the present results indicate that subjective ratings can discriminate among levels of alertness, they also reinforce the fact that subjective judgements can be unreliable or not clearly interpretable. It is apparent that care should be taken when attempting to assess relative or absolute levels of alertness based on such post-hoc judgments.

## Effects of Driving Time on Emergency Performance

Experiment I - Car following emergency. The significantly longer reaction times of early trials of the post-EDR over trials of the pre-EDR (indicated by the significant run x trial interaction) supports the hypothesis that extended driving causes degradations in drivers' abilities to respond to emergency road situations over time.

The significant differences existing among early trials of the post-EDR run and corresponding pre-EDR trials, and the non-significance of differences among later trials indicate that the effects of prolonged driving on reaction time tend to diminish with successive response trials. The mean differences in post-EDR reaction times of trials 1 and 3 from corresponding pre-EDR reaction times were .306 s and .372 s. At normal highway speeds, these changes in response latency could easily constitute the difference between involvement in or avoidance of an accident.

The most likely cause for the apparent neutralization of the performance decrement with successive response trials would seem to be an alerting effect of the emergency stimulus presentation and the associated driver's response. Such an effect would counteract the driving induced decrements in performance and tend to cause a return in performance to "non-fatigued" levels. In this experiment, after the third trial mean, reaction times improved (decreased) to pre-EDR response time levels.

Among the post-EDR trials, the shortness of the mean reaction time for trial 2, relative to those of trial 1 and 3, is somewhat puzzling. One possibility is that the first emergency trial after 30, 60, or 150 min of continuous, uneventful driving constituted a novel, and hence highly alerting, stimulus. Such a stimulus could have resulted in a temporary but significant elevation in the alertness level of drivers, which would, in turn, facilitate quicker responses to the second trial stimulus.

The fact that increases in driving duration had no incremental effects on reaction time was somewhat unexpected. These results, indicating no differences among reaction times of drivers who had driven continuously for periods of 30, 60, or 150 min, implies that (a) the effects of continuous driving time on reaction time is truly asymptotic after 30 min of driving, or (b) the effect of continuous driving time on reaction time is non-linear, and that longer driving periods are required to elicit any further decrements in reaction time performance.

Experiment II - Wind gust emergency. The difference in mean lateral deviation error found between pre- and post-EDR runs indicates that drivers were less able to maintain vehicle control after having driven continously for periods greater than 30 min than they were able to before having driven for that period. The actual magnitude of the statistically significant difference, however, was quite small, with the post-EDR mean lateral deviation error exceeding that of the pre-EDR mean by only 5.8 cm. Although it is doubtful that such a difference would constitute a practical difference in actual emergency driving performance, the result does imply that certain psychomotor response mechanisms undergo reliable changes over driving time. (In terms of accident avoidance, a more important measure might be maximum excursions from the mean.)

A possible reason why larger performance decrements (due to extended driving) were not observed is that the stimulus cues for wind gusts are multimodal and redundant. That is, because cross wind gusts cause simultaneous changes in roll, yaw, and lateral movements of the vehicle, the driver receives parallel cues through visual, vestibular, and kinesthetic sensory modalities. The presence of such multidimensional cues could result in a reduction in the uncertainty of onset of an event, which would likely result in shorter response latencies to emergencies relative to emergencies presented through single stimulus modalities. (According to McCormick, 1970, the use of multidimensional coding tends to facilitate information transfer.)

The non-significance of the duration effect indicates the effects of extended driving remained constant across the three duration conditions. This absence of an incremental driving time effect beyond 30 min indicates, as in Experiment I, that (a) the driving time effect is non-linear and the driving times employed were not sufficiently long to elicit any changes in performance; or (b) the driving time effect is asymptotic after 30 min of continuous driving.

Finally, the absence of a significant steering gain effect indicates that widely disparate ("normal" and "high") levels of steering gain do not affect wind gust recovery performance. The fact that drivers performed equally well in the wind gust emergency despite widely disparate vehicle steering gain characteristics indicates either that (a) drivers employ vehicle ("closed-loop") feedback to maintain vehicle control during such emergencies, or (b) that drivers modify their "open-loop" steering inputs according to learned vehicle response characteristics.

Experiment III - Simple and choice reaction times. A major reason for employing a choice reaction task was to provide a means to ensure that the decision making processes in the emergency task of Experiment III more closely resembled those of real world emergencies. The significant differences in (mean) simple reaction times from Day 1 CRT baseline means for braking and steering responses indicate that even though the same stimuli and motor tasks were involved, the addition of alternate choices to simple reaction tasks results in a significant increase in the time to respond to the stimulus.

As discussed earlier, Donders (1969) and others have presented evidence which indicates that the time required to respond to a choice reaction stimulus depends on the time required to execute a series of successive stages or processes, specifically (a) stimulus input time, (b) stimulus indentification time, (c) response selection time, and (d) response output time. If simple reaction time (which presumably consists only of stages (a) and (c)) is subtracted from the estimate of choice reaction time, the remainder is presumed to be the total decision time consisting of the time to identify the stimulus and to select the action required. Assuming the validity of Donder's subtraction method, simple reaction times (for brake and steering response tasks) subtracted from the choice reaction times yields "decision time" latencies of .188 s for brake response trials and .208 s for steering response trials. The significantly longer reaction times for the choice reaction time runs, relative to the simple reaction time runs, indicate that the increase in "processing requirements" of the choice reaction task does constitute an increase in task complexity over simple reaction tasks (which are often employed in driver fatigue studies).

The fact that the mean reaction times for the first trial of simple and CRT baseline runs (note that the run x trials interactions were not significant) were longer than most of the mean reaction times for subsequent trials seems to indicate that drivers were better able to anticipate stimulus onset after the first response trial. This apparent change in the ability of drivers to anticipate stimulus onset appears to be the reason why a number of investigators have employed signals to warn subjects that a stimulus is about to be presented (e.g., Donders, 1969; John, 1969; Sternberg, 1967).

Experiment III - Pre- and post-EDRs (choice reaction time). Analysis of steering and brake (choice) reaction times indicates that reaction times for trials presented immediately after continuous driving (in post-EDR trials) were significantly longer than trials that did not follow extended driving (pre-EDR run). Although statistically significant, the differences between pre-and post-EDR runs were surprisingly small. For brake responses, the mean run difference was .053 s; and for steering responses, the difference was .061 s. It is doubtful that such a difference could account for "fatigue induced" accidents. The question is, is this a valid measure of the overall response decrement?

From the results of Experiment I and other studies, it was anticipated that reaction times for the early trials of the post-EDR would be longer than corresponding trials of the pre-EDR. The absence of a run x trial interaction indicates that this was not the case. However, to protect against systematic biases of one set of data, the first trial presentations for brake and steering signals were counterbalanced so that half of the subjects received the signal to brake on the first trial, the other half received the signal to make a steering response. It is possible that the counterbalancing of these trials acted to dilute the "first trial" effect due to the fact that, in half the cases, the "first" response trial was preceded by a trial which may have alerted the driver and facilitated a faster response to the signal that followed. If such contamination did occur, then the true effect of extended driving would likely be longer than the effect due to the run would indicate.

The main effect of trials indicated that the mean reaction times for trial 1 (for steering responses) and trial 2 (for brake responses) were longer than for subsequent trials. This result, combined with absence of a run x trials effect implies that both pre- and post-EDR runs exhibited similar changes over trials. This finding, as with the simple and Day 1 baseline data, reflect an apparent change in the anticipatory responses of drivers. This seems to indicate that a portion of what might have been referred to as the "fatigue decrement" is due to the "surprise" effect of the first trial stimulus.

As in Experiments I and II, effects due to driving duration were not significant. The significant decrement in driving performance (indicated by significant differences between pre- and post-EDR runs) apparently did not increase (which would have resulted in a significant run x duration interaction) even though driving time of the long duration condition was five times longer than the short duration condition. These results indicate that (a) the effects of continuous driving time is asymptotic after 30 min of driving or (b) the effect of continuous driving time on reaction time is non-linear, and longer driving periods are required to elicit further decrements in performance.

### Relationship of Emergency Response Decrements and EDR Variables

The results of the correlational analysis indicate that, in general, the relationships of performance degradations in emergency situations to changes in continuously monitored (EDR) are rather weak. Of these correlations (Table 28), few of them appear to be high enough to account for more than a minor proportion of the variance. Furthermore, none of the EDR variables correlated highly with emergency performance measures across all three experiments. From these results, it appears that none of the EDR variables could be used to predict reliably driving-time-induced emergency performance decrements. Several reasons could explain why a greater number of significant correlations were not attained.

Emergency performance measures. Perhaps the most significant factor in the above relationships is that EDR variables for medium and long duration conditions continued to change over time, while the emergency performance measures did not indicate any further decrements in driving periods longer than 30 min. This "ceiling" effect (of time) on emergency performance would obviously limit the magnitude of time related correlations. Several reasons could account for the asymptotic effect of time on emergency performance. First, it is possible that time does, in fact, have an asymptotic effect on the emergency performance measures used in present investigations. This notion is not unprecedented in view of the findings of several investigators (noted earlier), who have been unable to demonstrate that drivers exhibit any emergency response performance degradations over time. Secondly, it is possible that drivers do not exhibit degradations in performance in the particular emergencies employed -- as they might in other types of emergencies. For example, it has been suggested that fatigue and other variables (such as narcotic or alcohol usage) causes "tunnel vision" in drivers, a phenomenon in which the effective visual field of the driver is narrowed. If true, then perhaps the most significant emergency performance degradations would be observed in emergencies where drivers are required to respond to stimuli normally occurring in the peripheral field of vision. A third possibility is that driving times longer than 150 min (the longest time employed in the present investigation) might have been required to elicit further degradations in emergency performance. Although the use of the driving simulator in the present investigation was expected to yield more rapid degradations over driving time than actual roadway driving (because of its presumed "stimulus improverished" environment relative to actual roadway driving), this assumption may have been unfounded.

Another reason why no differences in emergency performance were detected among driving duration groups concerns the statistical sampling problem, which appears to be unique to fatigue studies. In particular, the significant run x trial interactions of Experiment I (for example) indicates that the emergency performance decrement associated with driving time (indicated by performance differences between pre-and post-EDR measures) tends to diminish with repeated emergency response trials. This change is presumed to be due to the alerting effect of the response trial itself. If so, it may be that the only valid indication of the driving related decrement is the first (or some limited number of early) trial. This result indicates that the number of useable data points per driver is severely limited when dealing in emergency response measures. With small sample sizes (as in the present investigation), low degrees of freedom associated with first emergency response trials (only) in addition to high subject variability could severely limit the sensitivity of measures to driving duration effects. This could be one of the reasons why pre-post measures (such as the emergency performance measures of the present experiments) often are not sensitive to changes over time.

<u>EDR variables.</u> The correlational analyses of the present investigation employed EDR "change scores" which were calculated by subtracting the EDR variable score of the first 10-min period from the score of the last 10-min period. One possible problem is that the 10-min averaging period for the EDR data may have "smoothed" transient changes in the variable which would have otherwise better reflected the current state of alertness of the driver. A possible way to combat such a problem would be to shorten the time period over which the continuous data are averaged, taking care not to reduce the time to where the measures are statistically "noisy".

<u>Concluding statements on the ability of EDR variables to predict decrements in emergency response performance.</u> The present results indicate that changes in EDR variables over driving time do not correlate highly with changes in emergency performance. These results, which were contrary to expectations, indicate that the EDR variables employed would not be able to predict accurately decrements in drivers' abilities to respond to emergencies of the type used in the present investigation. Before abandoning the approach completely, future investigations should continue to explore (a) the effects of longer driving periods on emergency performance, (b) driving time effects on different emergency tasks, (c) the effects of differing averaging periods for EDR variables, (d) the use of various other EDR variables, and (e) the effectiveness of multiple regression models as predictors by using several EDR variables to predict changes in emergency response decrements.

### Methodological Considerations

One of the goals of the present investigation was to identify methodological considerations that might have affected the results (and resulted in experimental biases of past investigations), and to identify methods for minimizing or controlling their effects.

Statistical methods. As discussed earlier, the majority of previous driver fatigue studies have employed multiple dependent variables to measure fatigue effects, most of which have been analyzed using univariate statistical methods. The use of univariate statistical methods could result in interpretive errors in the final analysis. In contrast to the aforementioned approach, the present investigation employed multivariate analysis of variance with followup univariate analyses. As indicated by Tables 4 and 5, the univariate ANOVAs for all EDR variables (except YAWREVII of the medium duration condition) yielded highly significant effects due to time. If all dependent variables were perfectly correlated, the overall $\alpha$ level for k comparisons could not exceed $\alpha k$ (for a given fixed $\alpha$ level, Finkelman et al., 1977), the sum of individual alpha levels (where individual levels are specified for each comparison). In the present results, even assuming this most conservative case, the maximum (combined) alpha levels would be 0.01 for the medium duration condition (excluding YAWREVII) and 0.02 for all variables of the long duration condition. From these highly significant results, it would seem that the use of multivariate analysis was unnecessary. This, however, may not be true in all situations. For example, the MANOVA conducted on the combined EDR variables of the short duration condition indicated no significant effects due to time, but univariate ANOVAs of the EDR data would have indicated that two of these variables, HRTINTSD and STREVMN, exhibited significant effects due to time ($p < .05$). It would appear that the multivariate analysis served to "screen out" what would have otherwise been essentially spurious results. From such findings, it is reasonable to assert that the use of multivariate analysis is an appropriate and necessary "screening" technique for analyzing multiple dependent measures which interact to an unspecified degree.

Emergency response data and repetitive trials. From the results of the three sets of emergency response data, it appears that care need be taken in future studies concerned with measuring emergency response decrements. In Experiment I, the mean response times for the post-EDR run indicated a significant performance decrement for early trials, but this decrement appeared to decrease (tending toward non-"fatigued", pre-EDR levels) with subsequent trials. This result indicates that the measurement of the performance itself apparently negates the effect of the mechanism that is being measured. Certain steps could be taken to conteract such effects, such as increasing the "recovery interval" between trials (presumably to allow alerting effects of a given response trial to "settle"). However, the problem then remains in determining the duration of the transient alerting effects.

From the results of Experiment I, it appears that the "fragility" of fatigue induced by extended driving casts doubt as to the reliability and validity of results of previous investigations which have employed repeated response trials during an extended driving period. In a study discussed earlier, Lisper et al. (1973) required subjects to respond to 200 response trials during a three-hour driving period. Although an overall "fatigue" decrement was observed, it is quite likely that the repeated trials could have counteracted "normal" fatigue-related mechanisms. Consequently, based on the present results, it is highly likely that the performance decrements shown by Lisper et al. do not reflect the "true" response decrements exhibited by drivers in actual roadway driving.

Experiment III emergency response trials also revealed improvements in response performance with repetitive trials. The trials effect, however, was present in both pre- and post-EDR runs, indicating a general trend toward lowered response time with successive trials. This improvement in performance with successive trials is apparently due to a learning effect gained from repeated exposure to the stimulus/response combinations. The implication of this result is similar to that of Experiment I in that, if successive response trials are employed, but the investigator fails to measure or control the trials effect, it is quite likely that the effects of the independent variables of interest will be less reliable and valid than thought otherwise.

Experiment II, unlike the other experiments, showed no significant effects due to trials. As discussed earlier, it seems plausible that the redundancy of wind gust cues causes performance to be less susceptible to driving time effects. If so, then the performance levels among successive trials would tend to be homogeneous.

Measurement of steering behavior. As discussed earlier, investigators have employed various techniques for measuring steering reversal movements. The majority of these techniques operate on a principle of detecting and counting discrete steering movements which are larger in magnitude than a preset threshold angle. A major key to success of these methods lies in the selection of optimal steering reversal thresholds. Such "optimal" thresholds would serve to detect "important" steering movements while excluding inconsequential movements, thus maximizing the "signal-to-noise" ratio. Threshold angles used in previous investigations, often determined arbitrarily or by trial-and-error methods, fall within a range starting from about 1 to 2 deg (e.g., Sussman et al.,

1971) to 12 deg (Sussman and Morris, 1970). Although these methods are attractive in terms of their relative simplicity, the present results (which indicate that steering reversals fall into dichotomous classes with opposing time trends) cast suspicion on the validity of any unitary steering reversal threshold measure. If, for example, a high amplitude steering threshold is employed, only reversals large enough to exceed the threshold would be counted.

On the other hand, if a low amplitude reversal threshold is employed, reversals which exceed the threshold value would include large and small steering reversals. Unless special steps were taken, small steering reversals would be indistinguishable from large ones.

Based on the present data, it is likely that the trends of small steering reversals (which exhibited declines in occurrences over time) would tend to mask or completely dominate the trends of the large steering reversals. This would especially be true in view of the present findings, which indicated that the frequency of small steering reversals was generally much higher than those of large steering reversals. Therefore, any change in the frequency of steering reversals would likely exceed a proportional change in the frequency of reversals. Such might have been the case in several previous investigations. For example, Sussman and Morris (1970) found that road position error increased, while fine steering reversals (>2 deg) decreased during four hours of continuous driving. Finding the data somewhat paradoxical, because decreases in steering frequency are usually associated with increased tracking accuracy, the investigators speculated that the decrease in steering reversal frequency was due to a decrease in the frequency of "perceptual sampling" of the environment by the driver (p. 40). In view of the present results, an equally plausible explanation is that the increase in road position variability was caused by the increase in frequency of large steering inputs, the presence of which was masked by decreasing trends of the frequency of small steering reversals which were included in the distribution.

A contradiction with past investigations is evident in the fact that several studies indicating declines in steering reversal frequencies over time have employed measures which supposedly detect steering reversals of 2 deg and larger (the definition for "large" reversals of the present investigation). There are several possible reasons for the conflict. First, because the dichotomous "break point" between small and large steering reversals in the present investigation was chosen somewhat arbitrarily, the actual break point could vary substantially from the selected angle of 2 deg. Secondly, it is possible that past studies have suffered from measurment errors due to calibration, electronic drift, and hysteresis errors in measuring small steering reversals. The present investigation employed state-of-the-art measurement instruments and techniques which serve to minimize such errors. These included: a high resolution, drift-compensated analog tape recorder, high speed digital analysis, and software and hardware noise reduction techniques. Such may not have been available to previous investigators.

## SUMMARY AND CONCLUSIONS

### EDR Variable Changes Over Time

On a group basis, it is apparent that drivers exhibit significant changes in steering behavior, vehicle lane-keeping performance, and certain physiological variables over extended driving periods. Of the three driving duration groups (short, 30 min; medium, 60 min; and long, 150 min), only the short duration group failed to exhibit significant EDR variable changes over time, implying that drivers begin to exhibit changes after 30 to 60 min of continuous driving.

EDR variable means exhibited highly significant, predominately linear, trends over successive time periods (among drivers in the medium and long duration groups). The regression analyses on group mean trends indicate that major proportions of variances can be accounted for by linear models of time trends. In contrast, the regression analyses on individual driver behavior indicated that only small proportions of the variances could be accounted for by such models. It is apparent that inter-subject variability in terms of EDR variable baselines, slopes (reflecting susceptibility to time effects), and linearity (sudden deviation from means) would limit the ability of any group model to reliably detect individual changes in driver behavior.

### Emergency Response Performance Changes Over Time

Drivers exhibited a decline in the ability to respond to emergencies after having driven continuously for 30 min or longer. The absence of any duration condition differences in the three emergency performance task groups indicates that drivers apparently reached an asymptotic level of performance degradation after 30 min of driving, and that driving time increments of up to 2 hours did not produce any further degradations in emergency performance. This was the case in spite of the fact that subjective fatigue ratings of drivers in the medium and long duration groups indicated that they were significantly "more tired" than drivers in the short duration group. Although a number of reasons could account for why no time-related differences were found, the biggest obstacle to detecting significant changes may have been the alerting effect of sequential response trials. Because response trials apparently increase the alertness of the driver, the most valid measures of the response decrement might exist in the first trial (or two) after the driving period. If true, larger sample sizes would undoubtedly be required to detect any significant differences with respect to driving time.

### Predictability of Emergency Performance Decrements by EDR Variables

Correlations between emergency performance decrements and EDR variable changes indicate that the relationship between the variables is, at best, tenuous. Based on these results, it would appear that changes in a set of EDR variables could not reliably be used to predict an individual's ability to respond to emergencies of the type in the present experiment. Future studies could address some of the methodological problems of both types of variables described. In addition, the use of multiple regression and other techniques which would make better use of combinations of variables could prove to be reliable predictors of behavior.

### Methodological Considerations for Future Studies

It is apparent that certain methodological considerations should be addressed when attempting to measure fatigue effects. These considerations should include (a) the use of multivariate analyses when multiple dependent measures are employed, (b) assessment of the "trials" effect in emergency response performance measures, (c) examination of the nature of the emergency task in terms of validity, and (d) evaluation of methods for measuring complex operator behaviors such as steering wheel movements.

# REFERENCES

Alexander, R. A. and Barrett, G. V. Relationship between perceptual style and responses to visual motion under active and passive viewing conditions. *Journal of Applied Psychology*, 1975, *60(4)*, 507-512.

Barr, A. J., Goodnight, J. H., Sall, J. P., and Helwig, J. T. *A User's Guide to SAS 76*. Raleigh: Sparks Press, 1976.

Barrett, G. V. and Thornton, C. L., Relationship between perceptual style and simulator sickness. *Journal of Applied Psychology*, 1968, *52(4)*, 304-308.

Barrett, G. V., Thornton, C. L. and Cabe, P.A. Cue conflict related to perceptual style. *Journal of Applied Psychology*, 1970, *54(3)*, 258-264.

Brown, I.D. A comparison of two subsidiary tasks used to measure fatigue in car drivers. *Ergonomics*, 1965, *8(4)*, 467-473.

Brown, I. D. Car driving and fatigue. *Triangle*, 1967, *8*, 131-137.

Brown, I. D., Simmonds, D. C. V., and Tickner, A. H. Measurement of control skills, vigilance, and performance on a subsidiary task during 12 hours of car driving. *Ergonomics*, 1967, *10(6)*, 665-673.

Brown, I. D., Tickner, A. H., and Simmonds, D. C. V. Effects of prolonged driving on overtaking criteria. *Ergonomics*, 1970, *13(2)*, 239-242.

Casali, J. G. and Wierwille, W. W. The effects of various design alternatives on moving-base driving simulator discomfort. *Human Factors*, 1980, *22(6)*, 741-756.

Case, H. W., Hulbert, S., and Mellinger, R. L. Effects of fatigue on skills related to driving. UCLA Rept. 7060. Los Angeles: University of California, Department of Engineering, 1970.

Crawford, A. Fatigue and driving. *Ergonomics*, 1961, *4(2)*, 143-154.

Davies, D. R. and Tune, G. S. *Human vigilance performance*. New York: American Elsevier Publishing, 1969.

Dobbins, D. A., Tiedemann, J. G. and Skordahl, D. M. Vigilance under highway driving conditions. *Perceptual and Motor Skills*, 1963, *16*, 38.

Donders, F. C. On the speed of mental processes. *Acta Psychologica, Attention and Performance II*, 1969, *30*, 412-431.

Durbin, J. and Watson, G. S. Testing for serial correlation in least squares regression II. Biometrika, 1951, 38, 159-178.

Dureman, E. E. and Boden, C. H. Fatigue in simulated car driving. Ergonomics, 1972, 15(3), 299-308.

Ellingstad, V. S. and Heimstra, N. W. Performance changes during the sustained operation of a complex psychomotor task. Ergonomics, 1970, 13(6), 693-705.

Evans, J. E. and Gutmann, J. C. Minicomputer processing of dual Purkinje image eye-tracker data. Behavior Research Methods and Instrumentation, 1978, 10(5), 159-178.

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

Finkelman, J. M., Wolf, E. H. and Friend, M. A. Modified discriminant analysis as multivariate post-comparison extension of MANOVA for interpretation of simultaneous multimodality measures. Human Factors, 1977, 3, 253-262.

Harris, R. J. A primer of multivariate statistics. New York: Academic Press, 1975.

Harris, W. Fatigue, circadian rhythm, and truck accidents. In R. R. Mackie (Ed.) Vigilance: theory, operational performance and physiological correlates. New York: Plenum Press, 1977.

Harris, W. and Mackie, R. R. A study of the relationships among fatigue, hours of service, and safety of operations of truck and bus drivers. NTIS No. PB-213 963, 1972.

Heimstra, N. W. The effects of "stress fatigue" on performance in a simulated driving situation. Ergonomics, 1970, 13(2), 209-218.

Herbert, M. J. and Jaynes, W. E. Performance decrement in vehicle driving. NTIS No. Ad-430 226, 1963.

Hoffman, E. R. and Joubert, P. N. The effect of changes in some vehicle handling variables on driver steering performance. Human Factors, 1966, 8(3), 245-263.

Huntley, M. S. and Centybear, T. M. Alcohol, sleep deprivation and driving speed effects upon control use during driving. Human Factors, 1974, 16(1), 19-28.

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

John, I. D. Mediating processes in choice reaction times. Acta Psychologica, Attention Performance II, 1969, 30, 58-64.

Kashiwagi, S. Psychological ratings of human fatigue. Ergonomics, 1971, 14(1), 17-21.

Kemmerer, R. H. and Hulbert, S. F. UCLA driving simulation laboratory: with a 360 degree scene around a full size car. Society of Photo Optical Instrumentation Engineers, 1975, 59, 158-170.

Kerlinger, F. N. and Pedhazur, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart, and Winston, Inc., 1973

Konz, S. and McDougal, D. The effect of background music on the control activity of an automobile drive. Human Factors, 1968, 10(3), 233-244.

Kuroki, Y., Aso, T., Hori, H. and Matsuno, M. Variation of driver's arousal level during car driving in Japan. Society of Automative Engineers, SAE paper 760081, 1975.

Laurell, H. and Lisper, H.-O. Changes in subsidiary reaction time and heart-rate during car driving, passenger travel and stationary conditions. Ergonomics, 1976, 19(2), 149-156.

Laurell, H. and Lisper, H.-O. A validation of subsidiary reaction time against detection of roadside obstacles during prolonged driving. Ergonomics, 1978, 21(2), 81-88.

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

Lerner, H. D., Abbot, H. E. Jr., and Sleight, R. B. Following distance on the highway related to speed, trip duration, traffic, and illumination. Human Factors, 1964, 6(4), 343-350.

Lisper, H.-O., Laurell, H., and Stening, G. Effects of experience of the driver on heartrate, respiration-rate, and subsidiary reaction time in a three hours continuous driving task. Ergonomics, 1973, 16(4), 501-506

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

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

Mackworth, J. F. Vigilance, arousal, and habituation. Psychological Review, 1968, 75(4), 308-322.

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

McComick, E. J. Human factors engineering. New York: McGraw-Hill Book Co., 1970.

McLane, R. C. and Wierwille, W. W. Influence of motion and audio cues of driver performance in an automobile simulator. Human Factors, 1975, 17(5), 488-501.

McLean, J. R. and Hoffman, E. R. Steering reversals as a measure of driving performance and steering task difficulty. Human Factors, 1975, 17(3), 248-256.

Moshier, S. L., Osborn, R. R., Baker, J. M., and Baker, J. K. Dialog Systems automatic speech recognition capabilities - present and future. Proceedings of a Symposium: Voice Interactive Systems: Applications and Payoffs, Dallas, Texas, 1980, 163-187.

Motor Vehicle Manufacturers Association. Motor vehicle facts & figures '76, Detroit: Motor Vehicle Manufacturers Association, 1976.

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

Ohkubo, T. Psycho-physiological reactions of vehicle drivers under long distance driving. Proceedings of the 6th Congress of the International Ergonomics Association, College Park, Maryland, 1976, 299-303.

Olson, P. L. The effects of variable-ratio steering gears on driver performance. Human Factors, 1970, 12(6), 553-558.

Platt, F. New method of evaluating the effects of fatigue on driver performance. Human Factors, 1964, 6, 351-358.

Pokorny, M. L. I., Opmeer, C. H. J. M., and Blom, D. H. J. Effect of learning, age and type of shift on the performance on a critical instability task by bus drivers before their work. A preliminary report. Paper presented at the First European Annual Conference on Human Decision and Manual Control. Delft, Netherlands, May 25-27, 1981.

Repa, B. S., Alexandridis, A. A., Howell, L. J., and Wierwille, W. W. Study of vehicle steering and response characteristics in simulated and actual driving. Society of Automotive Engineers, SAE paper 780011, 1978.

Riemersma, J. B. J., Biesta, P. W., and Wildervanck, C. Fatigue and stress due to prolonged driving and changing task demands. Society of Automotive Engineers, SAE paper 770134, 1977.

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

Safford, R. and Rockwell, T. H. Performance decrement in 24 hour driving. Highway Research Record, 1967, 163, 68-79.

Smith, R. L. and Lucaccini, L. F. Vigilance research: its application to industrial problems. Human Factors, 1969, 11(2), 149-156.

Snook, S. H. and Dolliver, J. J. Driver Fatigue: a study of two types of countermeasures. Proceedings of the 6th Congress of the International Ergonomics Association, College Park, Maryland, 1976, 304-311.

Sternberg, S. Retrieval of contextual information from memory. Psychonomic Science, 1967, 8(2), 55-56.

Sternberg, S. The discovery of processing stages: extensions of Donders' method. Acta Psychologica, Attention and Performance II, 1969, 30, 276-315.

Sugarman, R. C. and Cozad, C. P. Road tests of alertness variables. Washington, D. C.: National Highway Traffic Safety Administration, U.S. Department of Transportation, NTIS No. PB-215 450/8, 1972.

Sussman, E. D. and Morris, D. F. An investigation of factors affecting driver alertness. Washington, D. C.: National Highway Safety Bureau, U.S. Department of Transportation, NTIS No. PB-195, 1970.

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

Theios, J. Reaction time measurements in the study of memory processes: theory and data. The Psychology of Learning and Motivation, 1973, 7, 43-85.

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

Venables, P. H. and Martin, I. (Ed.) A manual of psychophysiological methods. New York: American Elsevier Publishing Co., 1967.

Yajima, K., Ikeda, K., Oshima, M. and Sugi, T. Fatigue in automobile drivers due to long time driving. Society of Automotive Engineers, SAE paper 760050, 1976.

# APPENDIX I

## VIGILANCE RESEARCH SUMMARY

(Davies and Tune, 1969)

## VIGILANCE RESEARCH

The following is a brief outline of variables which have been found to have (or are hypothesized to have) an effect on vigilance performance. (Davies and Tune, 1969).

I. **Signal Characteristics**

A. **Signal duration.** Studies have generally found that the longer a signal is presented, the greater the probability of its being detected, and less the decrement observed with time on task.

B. **Signal intensity.** High amplitude signals are more likely to be detected than those with lower amplitudes. Signals with low amplitudes increase the probability of a rapid decrement over time.

C. **Signal frequency.** Generally, the more often a signal is presented, the higher the signal detection probability. Some investigators have hypothesized that performance follows an inverted U function with respect to the stimulus frequency; therefore, very high and very low frequencies result in low performance levels, while some point between the extremes results in "optimum" performance levels.

D. **Signal probability.** It appears that signal probability (in addition to signal frequency) is important in signal detection performance.

E. **Intersignal interval.** As signal regularity increases, so does the probability of detection. Specifically, in low variability conditions, response time appears to be a decreasing function of inter-stimulus interval; is invariant in medium conditions of variability; and is an increasing function of inter-stimulus interval in high variability conditions.

F. **Spatial characteristics of the signal.**

1. When fixation is required, peripheral signals tend to be overlooked.

2. Performance improves when signals occur in one area of a display more often than others.

G. **Artificial signals.** Some investigations have shown that artificial signals improve vigilance task performance, while others have shown only temporary effects. Still others have found that such signals produce no effect at all.

II. **Task Variables**

A. **Duration of the task.**

1. The optimum length is about 30 minutes, i. e., "Most of the decrement which will take place occurs during this initial period."

2. Subjects' knowledge of the length seems to have the following effects:
   a. Subjects expecting long duration runs exhibit rapid performance decrements almost immediately after beginning the task. Those expecting short duration runs show less rapid decrements.

   b. Subjects aware of the end of the task exhibit an end effect-where performance tends to improve as the end of the run approaches.

B. **Rest pauses.** Rest breaks from continuous performance of 5-10 minutes appear to permit some recovery from performance decrements.

C. **Multiple operators.** Multiman systems seem to maintain higher performance levels than single man systems. The evidence, however, is equivocal.

D. **Time sharing and bimodal tasks.**

1. The effect of performing a vigilance task in conjunction with a secondary task is not clear.

2. Where redundant stimuli are presented (through different sensory modality), the overall performance is better than in single modality conditions.

E. **Incentives.**
1. Instructions which emphasize the importance of the task (with a supervisory figure present) help to raise performance levels, but have no effect on the rate of decrement.

2. Monetary rewards have been shown to maintain performance at high levels.

F. **Knowledge-of-results.**

1. Complete feedback seems to reduce performance decrements to a minimum (even when the knowledge of results is false).

2. Cueing signals result in higher detection rates and reduction of commissive errors.

G. **Practice.** Repeated exposure to the task reduces errors from session to session, but performance decrements still occur.

III. **Subject variables**

A. **Personality variables.** Some investigators have attempted to find relationships between certain personality characteristics and vigilance task performance. For example, it is generally found that extraverts make fewer commissive errors than intraverts, but extraverts show greater decrements with time on task.

B. **Intelligence.** There appears to be no relation between intelligence and vigilance task performance.

C. **Sex.** There appears to be no relation between gender and vigilance task performance.

D. **Age.** Age differences are sometimes found to cause differences in vigilance performance (presumably due to factors such as, deficits in short term memory, excessive risk taking, etc.); but the area has not been thoroughly explored.

## IV. Environmental variables

A. **Auditory noise.**

1. In general, noise has been found to improve detection rates, but appears to increase the time to detect a signal.

2. Noise appears to increase commissive errors.

3. The effect of noise on decrement rates has been found to be variable. In some cases it has been shown to accelerate decrements, in others retard it.

B. **Reductions in sensory stimulation.** Generally, during sensory deprivation, detection latencies and detection rates improve. However, after sensory deprivation, performance is usually impaired.

C. **Fatigue/sleep deprivation.** "Mental fatigue" and sleep deprivation has been found to impair vigilance task performance.

D. **Temperature.** Both heat and cold seem to impair vigilance task performance.

E. **Time of day.** Vigilance task performance has been shown to follow diurnal body temperature variations. Generally performance is better in the afternoon.

# APPENDIX II

## SIMULATOR PARAMETERS

## SIMULATOR PARAMETERS

STEERING GAIN $\frac{\dot{\psi}}{\delta_{sw}}$ STEADY STATE

Normal: $\frac{0.15\ \text{deg/s}}{{}^{\circ}_{sw}}$ (.72g/100 deg sw)

High (Experiment II): $\frac{0.35\ \text{deg/s}}{{}^{\circ}_{sw}}$ (1.67g/100 deg sw)

(sw denotes angular displacement of the steering wheel)

YAW RATE

Overshoot: 26 %

Rise Time (0-100%): .20 s

LATERAL ACCELERATION

Rise Time (0-100%): .57 s

UNDERSTEER: 5 deg/g

ROLL RESPONSE

Roll Compliance: 6 deg/g

Rise Time (0-100 %): .67 s

NOTE: Steering input and wind gust input transient response shapes assumed identical for all axes noted except roll.

WIND GUST (Experiment II):

Step sideforce equivalence of 890 N.

# APPENDIX III A

## GENERAL INSTRUCTIONS

(Instructions given to all subjects)

## INTRODUCTION

You are about to participate in an experiment designed to evaluate how drivers respond to certain "emergency" road situations. This will be accomplished with the aid of a driving simulator.

After you have been seated in the simulator, you will receive instructions on the simulator's operation and safety features. Basically, you will "drive" the simulator as you would an actual automobile, using standard controls (i.e., steering wheel, brake and accelerator pedals).

Before beginning the actual experiment, you will be given a chance to practice driving the simulator.

Your primary task during this experiment is to drive the automobile simulator without stopping. During the drive, you are asked to:

1. Maintain the car's position in the center of the right hand lane.
2. Keep both hands on the steering wheel at all times.
3. Maintain a speed of 55 miles per hour (as indicated by the speedometer).
4. Maintain a safe distance between your car and the one in front.

In addition, please observe the following:

1. Do not talk, smoke, or chew gum during the experiment.
2. Continue driving unless you are experiencing urgent personal discomfort. If for any reason you should stop before being instructed to do so, this will signify the end of the experiment, at which time you would be paid for

your time spent in the simulator (with no bonus).

3. But ... If you feel that you must stop for any reason - tell the experimenter immediately!

If you have any questions, please ask them now.

## EXPLANATION OF PAYMENT

1. The experiment in which you are about to participate will be conducted in several sessions. These sessions may be distributed over several days.

2. Your "base" pay for participating in this experiment will be $2.25 per hour. However, if you complete all trial sessions, you will receive a bonus payment of $0.75 per hour. If for some reason you cannot finish all sessions completely, you will be paid for your time at "base" rate.

3. The success of this experiment depends largely on your completing all sessions, so we ask that you make every effort in this regard.

4. If you have any questions, please ask the experimenter.

# APPENDIX III B

## INSTRUCTIONS FOR EXPERIMENT I

# INSTRUCTIONS

In this experiment, your task will be to drive the simulator as you would a conventional automobile. For the most part, this simply involves keeping the vehicle centered in the right hand lane at a speed of 55 miles per hour (as indicated by the speedometer). During the experiment your vehicle will appear to be following another car on the roadway. You are asked to maintain a constant, safe following distance between your vehicle and the one in front.

Several times during the experiment, the vehicle in front will appear to stop. When this occurs, immediately depress the brake pedal (press hard). After braking, the car in front will resume a normal speed. At this time resume a speed of 55 miles per hour with your vehicle centered in the right hand lane.

Before beginning the experiment you will be given a few minutes to get used to driving the simulator.

If you have any questions, please ask the experimenter.

# APPENDIX III C

## INSTRUCTIONS FOR EXPERIMENT II

# INSTRUCTIONS

In this experiment, your primary task will be to drive the simulator as you would a conventional automobile. For the most part, this simply involves keeping the vehicle centered in the right hand lane at a speed of 55 miles per hour (as indicated by the speedometer). In addition, your vehicle will appear to be following another vehicle on the roadway. You are asked to maintain a constant, safe following distance between your vehicle and the one in front.

At various times during the experiment, you will experience several very strong simulate cross wind gusts, which will tend to push your vehicle off the roadway. When this occurs, your job will be to maintain the vehicle's lane position and speed as well as possible (this will require that you steer in the opposite direction of the wind gust). After the wind gust stops, check to make sure that your speed is 55 miles per hour and that your lane position is correct.

If you have any questions, please ask the experimenter.

# APPENDIX III D

## INSTRUCTIONS FOR EXPERIMENT III

## INSTRUCTIONS

In this experiment, your task will be to drive the simulator as you would a conventional automobile. For the most part, this simply involves keeping the vehicle centered in the lane at a speed of 55 miles per hour. During the experiment, your vehicle will appear to be following another car on the roadway. You are asked to maintain a constant, safe following distance between your vehicle and the one in front.

As you will see, there are two small fixtures mounted on top of the dashboard of the simulator. Each of these fixtures contain two lights. When lit, the top lights are red, and the bottom ones are green. The lights will signify the status of two lanes of the roadway ahead; with the right fixture indicating the status of the right lane, and the left fixture indicating the status of the left lane. For example, a red light in the right fixture and a green light in the left fixture indicates that the right lane is closed and the left lane is open.

During most of the experiment, the lights will be turned off. At various times during the run, the lights will be turned on, and you will be expected to respond as follows (to the lights indicated):

1. **Left-red/Right-red:** step on the brake pedal immediately (press hard on the brake pedal). Resume 55 mph after the lights are turned off.
2. **Left-green/Right-red:** Quickly steer into the left lane. Remain in the lane until another signal indicates otherwise.
3. **Left-red/Right-green:** Quickly steer into the right lane. Remain in the lane until another signal indicates otherwise.

During early sessions, you will be told to expect only one signal combination during a trial. In later sessions, any of the light signal combinations may occur.

Please try to respond to these signals as if they were on-the-road emergencies, so when you see the signals - respond as quickly as you can! (As if, for example, there were obstacles immediately ahead of you on the roadway.)

# APPENDIX V

## PARTICIPANT'S CONSENT FORM

## PARTICIPANT'S CONSENT

As a participant in the driving simulator experiment, you have certain rights. The purpose of this note is to make you aware of these rights and to obtain your consent to participate.

1. You may stop the experiment in which you are participating at any time if you feel that it is not agreeable to you. Should you terminate the experiment, you will receive pay only for the proportion of time you participated.
2. You may see your data and you may withdraw it from the experiment if you feel that you should. In general, data are processed after all runs are completed. In this experiment, we can provide you with some qualitative information immediately after the experiment. Subsequently, all data are treated with anonymity. Therefore, if you wish to withdraw your data, you must do so immediately after your participation is completed.
3. You have the right to be informed on the results of the overall experiment. If you wish to receive information on the results, please include your address three months hence with your signature below. A summary will be sent to you. If you would then like further information, please contact the Human Factors Laboratory and a full report will be made available to you.

We hope you will find the experiment a pleasant and interesting experience. The faculty and graduate students involved greatly appreciate your help as a participant. If you have any questions about the experiment or your rights as a participant, please do not hesitate to ask. We will do our best to answer them, subject only to the constraint that we do not want to pre-bias the experimental results.

Your signature below indicates that you have read your above stated rights as a participant and that you consent to participation. If you include your printed name and address below, a summary of experimental results will be sent to you.

SIGNATURE: ____________________
DATE: ____________________
PRINTED NAME: ____________________
PRINTED ADDRESS: ____________________
(if you want
a summary of ____________________
results ____________________

# APPENDIX IV

## QUESTIONNAIRE

Subject # ______________ Group ______________ (No Name)

Sex _________ Age _________ Years Driving _________

1. How much sleep did you get last night? _________ hrs.

2. How much sleep do you normally get? _________ hrs.

3. Rate how you felt at the beginning of today's (tonight's) experiment. (Mark with an "x" on top of the line.)

| -3 | -2 | -1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|
| Extremely tired | | | | | | Extremely alert |

4. Rate how you felt at the end of today's (tonight's) experiment. (Mark with an "x" on top of the line.)

| -3 | -2 | -1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|
| Extremely tired | | | | | | Extremely alert |

5. Are you prone to drowsiness at the wheel? Yes No (Circle one)

6. Have you ever fallen asleep at the wheel? Yes No (Circle one)

7. Have you ever had an accident? Yes No (Circle one)

   Briefly describe the nature of the accident

8. Did you have coffee, tea or any other type of stimulant during the three hour period prior to your participation in today's experiment?

   Yes No (Circle one)

# A MULTIVARIATE INVESTIGATION OF DRIVER PERFORMANCE CHANGES DURING EXTENDED DRIVING PERIODS IN A COMPUTER-CONTROLLED DRIVING SIMULATOR

by

William H. Muto

## (ABSTRACT)

An investigation was conducted to determine the effects of driving time on nine continuous measures of driver performance and alertness, and to determine the extent to which such variables could be used to predict driver performance degradations in driving emergencies.

Subjects were required to respond to one of three types of driving emergencies presented after each had driven continuously for periods of 30, 60 or 150 minutes (determined by random group assignment).

The general findings of continuous measures indicated that in driving periods of 60 minutes and longer: (a) drivers exhibited linear increases in a tendency to cause the vehicle to drift laterally (within the lane boundaries) and yaw over time; (b) drivers exhibited significant linear increases in the number of large steering reversals, and significant decreases in the number of small steering reversals over time; and (c) drivers exhibited significant increases in heart beat interval standard deviation and in body movement frequency over time.

In all driving duration conditions, drivers exhibited significant degradations in their ability to respond to emergencies relative to their baseline driving performance levels. However, the amount of degradation did not differ among duration conditions.

Correlational relationships of continuous and emergency performance data, possible methodological problems of previous investigations (in light of present findings), effects of steering gain on continuous and emergency performance, and suggestions of possible future investigations were discussed.