THE UNIVERSITY OF CALGARY

COGNITIVE ABILITIES AND TRAFFIC ACCIDENTS

by

Cheng Siew Lim

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL

FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF MASTER OF SCIENCE.

DEPARTMENT OF PSYCHOLOGY

CALGARY, ALBERTA

SEPTEMBER 1987

ISBN 0-315-38026-8

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read and recommend to the faculty of Graduate Studies for acceptance, a thesis entitled "Cognitive Abilities and Traffic Accidents", submitted by Cheng Siew Lim in partial fulfillment of the requirements for the Degree of Master of Science.

Dr R. E. Dewar (Supervisor)
Department of Psychology

Dr L. Radtke
Department of Psychology

Dr K. Shapiro
Department of Psychology

Dr L. Falkenberg
(External Examiner)
Faculty of Management

Date: September 8, 1987

ABSTRACT

The use of the information processing model (spare mental capacity (SMC), attention switching ability, reaction time and perceptual style) to predict traffic accidents has indicated that drivers who have lower SMC, poorer ability to switch attention, slower reaction times and who are field dependent are more likely to have traffic accidents.

A different approach, but using the same information processing model, is undertaken in the present study. Seventy-two bus drivers from the city of Calgary, half of whom had three or more traffic accidents in the previous five years and half of whom were accident-free, participated in a series of information processing/ cognitive tasks. These tasks consisted of tracking on a CRT display, two-choice reaction time to lights (RT), dichotic listening task (DLT), tracking and RT combined, tracking and DLT combined, and all three tasks combined. In contrast to earlier studies which typically measured these or similar abilities in isolation, the present approach has the advantage of assessing the simultaneous performance of three tasks, which simulated closely the demands of the actual driving task. Additional measures

were undertaken to assess drivers' RT performance under conditions of auditory input (listening) and verbal output (reporting of digits) while performing the three tasks simultaneously. Measures of field dependence and subjective mental workload were also taken. It was hypothesized that the accident group of drivers would have less SMC, poorer ability to switch attention, experience higher mental workload and be more field dependent as compared with drivers in the accident-free group.

The results indicated that the drivers in the accident group performed more poorly on three measures of SMC - tracking, DLT (total DLT and switching errors) and RT errors during verbal output. Furthermore, the accident group had poorer ability to switch attention. It was also found that the overall RT performance for both groups was slower during listening than when they were reporting the digits in the DLT. Field dependence did not discriminate between the two groups. However, in the case of the accident group, field dependence was related to specific skills (eg., RT and attention switching) important for driving. Subjective mental workload was not correlated with any of the other performance measures. Suggestions were made for the possible use of these measures in the selection of professional drivers.

ACKNOWLEDGEMENTS

Out of ignorance came knowledge. For this I owe to my supervisor, R. E. Dewar, for his understanding , care and wisdom in seeing me through the past two years putting my thesis together. His supervision of my work has enlarged my current knowledge of Applied Psychology in the area of Human Factors a thousand fold.

I wish to thank my examining committee, Dr. L. Falkenberg, Dr. K. Shapiro and Dr. L. Radtke for their contributions.

Thanks also to Dr. J. Ells for his contributions in the early stages of the formation of the thesis.

I also like to thank my fellow graduate students, too many to mention by names, for their contributions to my thesis. From discussions in statistics to the reading of my drafts, and from an occasional word of encouragement to simple friendship, to them I say "thank you".

To the support staff, who lent assistance from securing equipment to computer programming, I wish to acknowledge their contributions.

I am indebted to the staff, union and management of Calgary Transit who generously contributed their time and resources to participate in this study.

I am also indebted to Transport Canada who sponsored this study

Finally, my greatest gratitude goes to my dear wife, Lena, and delightful daughter, Hannah, without whom this intellectual persuit would be in vain. And of course, thanks to my parents, brothers and sisters for their support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PAGE

LIST OF APPENDICES

PAGE

GLOSSARY OF ABBREVIATIONS

Abbreviation	Meaning
DLT	DICHOTIC LISTENING TASK
DLTC	TOTAL ERROR IN DLT TASK C
DLTE	TOTAL ERROR IN DLT TASK E
DLTF	TOTAL ERROR IN DLT TASK F
DLCOM	OMISSION ERROR IN TASK C
DLEOM	OMISSION ERROR IN TASK E
DLFOM	OMISSION ERROR IN TASK F
DLCIN	INTRUSION ERROR IN TASK C
DLEIN	INTRUSION ERROR IN TASK E
DLFIN	INTRUSION ERROR IN TASK F
DLCSW	SWITCHING ERROR IN TASK C
DLESW	SWITCHING ERROR IN TASK E
DLFSW	SWITCHING ERROR IN TASK F
ERRTB	ERROR IN REACTION TIME TASK B
ERRTD	ERROR IN REACTION TIME TASK D
ERRTF	ERROR IN REACTION TIME TASK F
ERIN	ERROR IN REACTION TIME DURING DLT INPUT
EROUT	ERROR IN REACTION TIME DURING DLT OUTPUT
GEFT	GROUP EMBEDDED FIGURES TEST
GP	GROUP (ACCIDENT AND ACCIDENT-FREE)

MISRD	FAILED TO RESPOND TO LIGHTS IN RT TASK D
MISRF	FAILED TO RESPOND TO LIGHTS IN RT TASK F
MISRIN	FAILED TO RESPOND TO LIGHTS DURING DLT INPUT
MISROUT	FAILED TO RESPOND TO LIGHTS DURING DLT OUTPUT
RT	REACTION TIME
RTB	REACTION TIME IN TASK B
RTD	REACTION TIME IN TASK D
RTF	REACTION TIME IN TASK F
RTIN	REACTION TIME DURING DLT INPUT
RTOUT	REACTION TIME DURING DLT OUTPUT
SMC	SPARE MENTAL CAPACITY
SMC1	SPARE MENTAL CAPACITY UNDER TWO-TASK CONDITION
SMC2	SPARE MENTAL CAPACITY UNDER THREE-TASK CONDITION
SML	SUBJECT MENTAL WORKLOAD
SMLA	SUBJECTIVE WORKLOAD RATING FOR TASK A
SMLB	SUBJECTIVE WORKLOAD RATING FOR TASK B
SMLC	SUBJECTIVE WORKLOAD RATING FOR TASK C
SMLD	SUBJECTIVE WORKLOAD RATING FOR TASK D
SMLE	SUBJECTIVE WORKLOAD RATING FOR TASK E
SMLF	SUBJECTIVE WORKLOAD RATING FOR TASK F
TA	TRACKING ACCURACY IN TASK A
TD	TRACKING ACCURACY IN TASK D

TE	TRACKING ACCURACY IN TASK E
TF	TRACKING ACCURACY IN TASK F
RED	REACTION TIME TO RED LIGHT
GREEN	REACTION TIME TO GREEN LIGHT

INTRODUCTION

Motor vehicle accidents have plagued road users almost since the inception of the automobile (Eames, Lee, and Fell, 1970). There are myriads of causes that have been attributed to them. Official police investigations, insurance reports and annual accident statistics bear to that. As such, a wealth of data has been gathered through the years on accidents and on their supposed "causes" (Fell, 1976).

Operating a motor vehicle safely is not solely dependent on the driver. Granted that the driver bears the responsibility, other factors such as the design of roadways, vehicle design, and the roadway-vehicle environment all play major roles in contributing to the operation of the motor vehicle safely and efficiently. For example, highway design, traffic control devices, roadway conditions and environmental factors (eg., illumination, weather) can influence the driver's information acquisition and processing, decision making and reaction time (Shinar, 1978).

Not only has the driver to contend with information from the roadway, he/she has to process information that is within the vehicle. Poor instrument display design and layout can be a potential source of problems. Green and Pew (1978), for example, found that only 6 out of 19 symbols (used or under consideration for use in automobiles) evaluated were understood or met the acceptance criteria of a minimium 75% recognition and a maximium of 5% confusions set by Heard (1974). Such poor design of instrument displays may well contribute to traffic accidents.

In a study of driver expectancy and performance in locating automobile controls, McGrath (1975) reported that drivers of rented vehicles had difficulty finding and operating a simple and basic mechanism like the ignition/starter control. Perel (1983), in a review of the literature on vehicle familiarity and safety, noted that a disproportionately high number of accidents involved drivers who were unfamiliar with their vehicles (eg., rented, borrowed or recently purchased). He concluded that unfamiliar drivers are two to three times more likely to be involved in an accident than are familiar drivers. Thus, lack of familiarity with a motor vehicle may well be another source that contributes to traffic accidents.

Reports of vehicle defects, roadway conditions, and drug and alcohol abuse by drivers have all been frequently mentioned and discussed as being causes of traffic accidents in the highway safety literature (Hills, 1980; Little, 1966; MacFarland, Moore and Warren, 1955; Sussman, Bishop, Madnick, and Walter, 1985; Treat, Tumbas, MacDonald, Shinar, Hume, Mayer, Stansifer, and Castellan, 1977). Scores of studies relating to traffic accidents have been conducted, focusing mainly on the driver. These include studies of drivers' vision (Burg 1967, 1968), spare mental capacity (Brown, 1962; Brown and Poulton, 1961), and drivers' attention and information processing (Avolio, Galen Kroeck, and Penek, 1985; Kahneman, Ben-Ishai, and Lotan, 1973; Shinar, Zaidel, and Paarlberg, 1978; Sussman et al., 1985). The present study does not attempt to explain or explore the numerous causes that contribute to traffic accidents. However, the cognitive ability (information processing, spare mental capacity, attention switching ability, and perceptual style) of drivers will be examined in relation to traffic safety.

The Driving Task

In view of the interactions between roadway enviroment, vehicle design and the motor vehicle operator, the major component to successful operation of a man - machine system such as driving is the driver's attention and ability to process information quickly and accurately.

Great amounts of information are being processed continuously in the driving task. This is because inputs from both the vehicle and the environment concerning vehicle status are being received by the driver, and on the basis of this, certain control maneuvers are executed to operate the vehicle safely and efficiently. To the extent that the driver does not select and process the necessary information from both the environment and from within the vehicle, or does not respond in a timely manner, safety will be diminished (Dewar, 1986; Rumar, 1986; Sussman et al., 1985). This becomes evident when close to half of all traffic accidents have been attributed to human error associated with problems in attention, perception and information processing, as indicated in an in-depth study in the U.S. by Treat et al.(1977). They revealed that improper lookout (23%), inattention (15%), and internal distraction (9%) were three of the most common human causal factors in traffic accidents. Shinar (1978) concluded that problems with perception and information processing, rather than poor vehicle control capabilities, are causes of most accidents.

More recently, Sussman et al. (1985), in a report on driver inattention and highway safety, noted that close to 38% of drivers involved in automobile crashes in the U.S.

in 1982, as documented in the National Accident Sampling System (NASS), took no actions to avoid the collision. This suggests that attentional lapses are a major contributing factor in highway accidents. Driver inattention (or attentional lapses) was broadly defined as the attentional state where the "driver fails to respond to a critical situation" (p.42). In the analysis, the role of the vehicle was known to be either striking another vehicle or an object, or being struck by another vehicle or object while in motion (at least at .8 km/hr). Vehicles whose roles were unknown or were involved in chain reaction collisions were not included. There were 11,868 vehicles involved in the striking/struck accidents in the 1982 NASS file. It was found that 22.5% (or 2,665) were striking vehicles whose drivers took no avoidance action before the collision. Another 15.5% (or 1,838) were struck vehicles whose drivers took no avoidance action before the collision. These data indicate that in accidents where an avoidable maneuver might have been of value, a large proportion (38%) of the drivers involved took no action to avoid the collision. It is therefore, suspected that driver inattention played a major role in these accidents.

In a review of the demands of the driving task, Dewar (1986) indicated that drivers require a high level of perceptual and information processing skill to operate modern vehicles safely. This is because both vehicle and highway designs have improved over the past decades, resulting in the driving task becoming one involving greater cognitive ability and less physical strength and skill. This trend toward increasing cognitive demands on the driver has accelerated in recent years with the introduction of "high technology" into vehicles - largely due to the advances of electronics engineering which have changed vehicle display technology dramatically. As a result, there is available to the driver a plethora of information in a late model motor vehicle.

Furthermore, changes in the design of roadways, traffic control devices and vehicles are moving at a faster pace than are changes in driver abilities. This is evident when one observes that the efforts to promote driver education in the past decade have not been met with overwhelming success. As Rumar (1981) appropriately pointed out, human limitations in the driving task can best be reflected by the fact that the driver is an "outdated human with stone age characteristics and performance who is controlling a fast, heavy machine in an environment packed with unnatural, artificial signs and signals" (p.37). This is because driving, unlike most tasks, involves continuous and sometimes rapid changes in the information available to the driver. Moreover, the status of the system itself is continually changing in most situations. There is therefore a tremendous demand placed on the driver to take in and process information rapidly, to switch attention frequently, and to make decisions and execute maneuvers based on this information with a minimium of delay. Thus, it is essential to understand the cognitive abilities of the driver and the limitations placed on these abilities in the driving task.

Spare Mental Capacity

It is believed that the first attempt ever to measure "spare mental capacity" (SMC) while driving was conducted by Brown and Poulton (1961). SMC was measured in terms of performance on a secondary (auditory) task involving immediate memory while driving in a shopping area, as compared with driving in a residential area. The secondary task involved groups of eight digits being presented every four seconds, with each group being identical to the previous one except for one digit. The new digit (one to nine) and its position in the group were presented randomly. The subject's task was to detect the new digit each time and report it during a gap of two seconds before the next group was presented. Errors were in terms of omissions or incorrect responses. As the demands of the driving task increased, performance on the secondary task and/or the driving task deteriorated. SMC is measured by the driver's ability to carry out a secondary task while driving.

In another study, Brown (1962) reported that relatively small but significant differences in drivers' SMC were detected between driving in the two traffic conditions (shopping area and residential area). It was found that correct responses on the auditory task when performed alone were 90.6% and they decreased to 83.8% while driving in a residential area (light traffic), and to 79.5% while driving in a shopping area, where many more decisions had to be made. Brown noted that SMC may be related to traffic accidents, as the driver must have certain SMC to draw on to compensate for the increased task difficulty in an emergency situation. In yet another two studies, Brown (1966, 1968) found that trainee bus drivers' SMC is a predictor of their success or failure in the completion of a training course of limited duration (5 weeks). SMC was measured after only one week of training in the vehicle in question. Moreover, Brown (1968) also found that SMC is a sensitive measure of transfer of learning in the training of bus drivers. On a similar note, North and Gopher (1976), using a tracking and digit processing time-sharing task found that measures of SMC are reliable predictors of success in flight training performance among pilot trainees.

Attention Switching

Central to the concept of SMC is selective attention and divided attention. "Attention implies that when a person is attending to one thing, he cannot simultaneously attend to something else." (Keele, 1973, p.4). This suggests the serial model of attention, which assumes that switching of attention and perceiving of stimuli are carried out at separate times (Laberge, 1973). However, a study by Laberge (1973), employing two tasks (detection and discrimination) which were investigated in the auditory and visual modalities, found that perceptual processing can occur at the same time when attention is located elsewhere.

Using mathematical equations, the results demonstrated the parallel model of attention switching. That is, perceptual processing can occur at the same time that the switching operation is taking place. However, the author failed to mention that detection errors were small (5%) and discrimination (identification) errors were enormous (up to 30%) when the operation of attention switching was taking place. This indicates that the task of detecting stimuli while engaging in attention switching is more compatible than the task of identifying stimuli when attention switching is in progress. Thus, there may be different levels of ability to process information when the operating of attention switching is taking place.

For the driving task, the driver is required to divide attention among concurrent activities and among concurrent signals constantly in order to operate the motor vehicle safely. Of particular interest is the ability to switch attention, especially from the roadway to instruments in the vehicle and vice versa, as well as between instruments within the vehicle and between objects on the roadway. However, with experience, driving becomes fairly automatic and less attention is required. Nonetheless, failure in the ability to switch attention to the appropriate incoming stimulus cues may result in an inappropriate or delayed response sequence which can be fatal to the driver and other road users.

There are various ways of measuring selective attention and the ability to switch attention as is required in the driving task. One of the better methods is the dichotic listening task (DLT), where subjects must attend to auditory stimuli such as digits presented, typically in sets of three pairs. These digits are presented simultaneously, one to each ear, with each pair involving different digits. No digits are repeated within each set. The task is to report the digits coming to one ear while ignoring digits presented to the other ear. The ear to attend to changes randomly (indicated by a tone) from trial to trial, thus requiring good attention to the relevant input as well as good ability to switch attention.

The use of the DLT, termed " Auditory Selective Attention Task" (ASAT) by Gopher and Kahneman (1971), to measure selective attention was validated against criteria of flight proficiency in the Israeli Airforce. The DLT employed in this study consisted of a series of 48 dichotic messages in which different information was presented simultaneously to each ear. Each message consisted of two parts. A tone was presented at the begining of each part to indicate to the subject which ear to attend for the incoming message. A low tone (250 Hz) meant that the left ear was the relevant ear. A high tone (2500 Hz) indicated the right ear.

The message in Part I lasted for eight seconds, during which either two or four target digits were presented to the relevant ear, interspersed in a stream of words. The digits and unconnected words were presented at a rate of two items per second to each ear. The subject was to report immediately only digits heard in the relevant ear. The second tone was then presented (may be the same tone as Part I) to indicate which ear was relevant for Part II of the message. Thereafter three pairs of digits were presented to the two ears. Again, the subject was required to report the three digits from the relevant ear, as indicated by the tone.

On 50% of the occasions, the same ear was relevant on both parts. Thus, subjects were required to switch attention, to the relevant ear, from time to time as indicated by the tone . The task was to repeat immediately all digits in the message to the relevant ear.

The intent of Part I was to simulate the maintenance of selective attention to a relevant channel. Part II of the message was to get at the ability of attention switching. Errors were measured in terms of omissions (omission of items from the relevant ear in part I), intrusions (intrusions of items from the irrelevant ear in part I), and switching errors (all errors committed in part II of the message).

The results of the study indicated that errors in part II of the message (switching errors) provided a more valid score than the scores in part I (omission and intrusion errors) in the prediction of flight training performance. It was found that the selective attention test represents an independent contribution (from other cognitive and psychomotor tests) to the prediction of success in flight training for trainee pilots. Moreover, the test was able to discriminate between flyers of high performance (faster) aircraft and those flying lower performance (slower) aircraft, with pilots of faster planes performing better on the ASAT.

Kahneman, Ben-Ishai, and Lotan (1973), employing the same DLT as described earlier (Gopher and Kahneman, 1971), in the study of selective attention on professional bus drivers in Israel, found that ability to switch attention was related to traffic accidents. A sample of 117 male bus drivers were employed in this study. Accident records which included severity of the driver's error in each reported accident were kept by the bus company. Out of the 117 drivers, 39 had a total accident rating of 3.5 or more, indicating at least two moderately severe accidents. No details were provided regarding the rating measures, except that it involved both severity and frequencies of the accidents.

For each of these 39 drivers, two other drivers were obtained, matching him on the variables of age, number of years of driving experience, type of route (urban or inter-urban), marital status, and ethnic origin. One of the matched drivers had a zero-accident rating during the same period of time, and the other matched driver had an accident rating of .5 to 3.0. Combining the two categories of relatively safe drivers and comparing them to the unsafe drivers, the point-biserial correlation between accident ratings and the number of switching errors was .51. Moreover, differences in intelligence (a short intelligence test was administered to the bus drivers) did not affect the selective attention test. The results indicated that all three types of errors (omissions, intrusions and switching) were significantly related to driver's accident ratings, with switching errors being the most useful and valid score in predicting accidents. It was noted that some subjects showed extremely high frequencies of all types of errors, and practice did not improve the performance on the selective attention task.

In a study to devise and select predictors of traffic accident involvement, Mihal and Barrett (1976), using an information processing model (perceptual style, selective attention and perceptual-motor reaction time), found that the DLT (selective attention) was significantly related to traffic accidents. The DLT employed was essentially the same as that of Gopher and Kahneman (1971) and Kahneman et al. (1973). With 75 commercial drivers having at least five years' driving experience with a utility company in the U.S. as subjects, it was found that errors in the DLT correlated significantly with accident rates, with more attentive drivers (ie., fewer errors in the attention test) having fewer accidents.

In a recent study to replicate the work of Mihal and Barrett (1976), Avolio, Kroeck, and Panek (1985), also employing the DLT method, concluded that errors in attention switching were more useful and powerful than omission and intrusion errors, in predicting traffic accidents. A sample of 72 commercial drivers, 58 males and 14 females, from a utility company in the U.S. were employed as subjects in the study. Accident record data for the previous 10 years were collected. Drivers who had caused some damage to their vehicles during normal operation on the road (as cited in the company records) were classified into the accident group. The accidents cited in the company records included collisions with stationary objects as well as with other vehicles.

However, the company did not specify (reliably) who was at fault.

Preliminary analyses indicated that sex, age and driving experience were not significantly correlated with the number of motor vehicle accidents. The results indicated significant correlations of all three error scores (omission, intrusion and switching) with accident rate. They found a correlation of .43 (significant at a level of $p<.001$) between errors in attention switching and the accident criterion, and a correlation of .36 and .31 between omission errors and accidents, and intrusion errors and accidents respectively. It is interesting to note that, once again, ability to switch attention has the strongest (highest) correlation with accident involvement.

These studies have shown the importance of selective attention in relation to traffic accidents. Moreover, the findings concerning selective attention can be generalized from a sample of bus drivers in Israel to samples of commercial drivers in the U.S.. Not only is the DLT a good predictor of traffic accidents, the same task has also been found to be a good predictor of successful pilot training (Gopher, 1982).

In a recent study on cognitive abilities and road safety, McKeena, Duncan and Brown (1986), employing the same DLT as that of Gopher and Kahneman (1971) and Kahneman et al. (1973), failed to replicate earlier findings of the relationship between DLT performance and driving performance measures. A sample of 153 subjects (149 males) participated in this study. The DLT performance (and all other measures of cognitive abilities) were tested at the time when all the subjects were in a bus-driver training course conducted by a large city bus company in England.

A two-year follow-up of these subjects (those still in service) on their accident records was undertaken. Accidents were broadly defined and included all incidents such as collisions, passenger falls, attacks on staff and so on. Thus, whether the driver was at fault cannot be determined. However, incidents which had nothing to do with the drivers (eg. stones thrown through windows, or the bus being hit while stationary) and those incidents in which no report was available were excluded (not considered) as accidents. The final score (accidents) for each driver was the total number of non-excluded accidents occurring within two years of passing the Public Service Vehicle (PSV) driving test. The PSV test was usually undertaken by the bus trainees after several weeks of training. Of the 153 people tested, 111 (72.6%) passed the PSV test.

In using the DLT, it was reported that on several occasions subjects seemed to have understood the instructions but completed the task wrongly. During the first four weeks of DLT testing (54 subjects were tested), it was found that 18.5% of these subjects completed the test without fully following the instructions. From the fifth week of testing onwards, an objective learning criterion was introduced to minimize this problem. The DLT trials were repeated until, on at least two trials, the subject had reported digits correctly from the relevant ear in both part I and part II. After the learning criterion was introduced, only 3% of the 99 subjects failed to follow the DLT instructions. Subjects who failed to follow the instructions were excluded in the data analysis for the DLT. The results indicated that PSV test performance (passed vs failed) had non-significant correlations of -.08 and -.02 with errors in part I (omission and intrusion errors) and with part II (switching errors) respectively. Of the 111 people in the sample who passed the PSV test, 91 remained in service after two years. Analyses of accident data were based on these 91 cases, of which 86 cases were used for the DLT analysis, due to exclusion of subjects who failed to follow the instructions. The number of accidents range from 0 to 12 over the two years of service. The results indicated that the correlations between the number of accidents and DLT errors in both part I (.07) and part II (-.16) were not significant. The data in this study completely failed to replicate the results of Avolio et al. (1985), Kahneman et al. (1973) and Mihal and Barrett (1976) with respect to DLT performance and traffic accidents.

Perceptual Style

An additional area of research which may relate to driving ability and traffic accidents is the ability of the driver to pay attention to relevant stimuli in a complex environment. Studies have shown that there are stable individual differences in perceptual style along the field dependence - field independence continuum. An individual who is relatively field dependent in one situation is likely to be equally field dependent in others, with field dependent individuals being less proficient in disembedding relevant stimuli from irrelevant background stimuli (Witkin, Dyk, Faterson, Goodenough, and Karp, 1962).

There are a number of methods used in measuring field dependence. One of these is the Embedded Figures Test

(EFT). This method requires the individual to locate simple 2-dimensional geometric forms which are camouflaged by embedding them in complex whole figures. Field dependent individuals have difficulty in locating the simple forms, whereas field independent individuals are able to find the forms with ease.

Another method is the rod-and-frame test (RFT) which involves the perception of upright position in space under conditions of conflicting gravitational and visual cues. The RFT requires the subject to view a tilted rod within a square, tilted frame. The subject is to adjust the rod to the upright position within the tilted frame with all visual cues to the location of the upright position being eliminated.

Yet another measure of field dependence is the body-adjustment test (BAT) which also involves the perception of the upright position. This method is conducted in a tilting-room-tilting-chair apparatus in which the subject is seated in a rotatable chair within a rotatable room. Both the chair and the room are rotated about the horizontal axis. With the room being tilted, the subject's task is to guide the experimenter in tilting the chair to achieve an upright position.

In both the RFT and BAT, the field dependent subjects tend to be highly influenced by the tilted visual cues (field) in their judgement of the upright position. They do not make use of the vestibular-kinesthetic sensation embedded in the total complex stimulus environment. In the extremes, field dependent people locate the upright position by simply determining the orientation of the frame or the room. Field independent people on the other hand, are able to experience and use the gravitational cues without regard to the conflicting visual information in determining the upright position.

These three methods of measuring field dependence - independence have been widely studied, with only the EFT and RFT being widely used as measures of field dependence - independence in research on other characteristics or skills such as driving, sensation seeking and locus of control.

In relating perceptual style to traffic safety, particular attention has been paid to the effects of field dependence on driving. Mihal and Barrett (1976) used the portable RFT (Oltmant, 1968) and the first six figures of the EFT (Witkin, Oltman, Raskin, and Karp; 1971) to measure perceptual style on a sample of 75 commercial drivers. The results indicated that field dependent drivers had significantly more accidents than did the field independent drivers. The RFT and the EFT had correlations of .38 ($\underline{P}$,.001) and .24 ($\underline{P}$<.05) respectively with number of accidents. The RFT measures of perceptual style had a stronger relationship to the accident criterion than did the EFT.

In a review of individual differences in field dependence as a factor in automobile safety, Goodenough (1976) cited six studies (including Mihal and Barrett, 1976) that found significant correlations between number of accidents and/or traffic violations and measures of field dependence. He also reported that field independent drivers (measured by RFT) were more effective in their control of a skidding automobile than were field dependent drivers. Moreover, field dependent drivers failed to learn from repeated exposure to a simulated emergency (skidding) situation (Olson, 1974). Studies have also shown that field dependent drivers are slower in responding to embedded road signs (longer brake reaction times, slower deceleration) and do not quickly recognize developing hazards (Barrett and Thornton, 1968; Barrett, Thornton and Cabe , 1969; Loo, 1978) and also fail to drive defensively in high-speed traffic (Olson, 1974). These studies indicated that field dependent drivers tended to have more traffic accidents than did the field independent drivers.

Other studies are less conclusive about this relationship between field dependence and accidents. Avolio et al. (1985), for example, found only a marginal correlation between GEFT scores and accident involvement ($\underline{p}<.10$). McKeena et al. (1986), employing the individual version of the EFT (Witkin et al., 1971), found a very weak but significant correlation ($\underline{r}$ = .18) using the criterion of $\underline{p}$ < .05 with bus driver training success (based on the Public Service Vehicle test). The significance level of $\underline{p}$ < .05 was only approximate, as the PSV test success and the accident rate were not normally distributed. In this individual version of the EFT, subjects were required to find simple forms in complex figures with the scores being the time taken. In a two year follow-up on these drivers, they found a correlation of 0.19 with accident rate, which was not significant.

In another study, Clement and Jonah (1984), employing 285 undergraduate students, 130 males and 155 females, did not find any relationship between field dependence and subjects' self-report of the numbers of accidents or the various causes of accidents. In their study, the Perceived Accident Causes Questionaires with 21 causes of accidents (abstracted from the Indiana Tri-level Accident Study by Treat et al. (1977)) were filled out by the subjects to find out the attributed importance of the various possible causes of their last accident. All the subjects possessed a valid driver licence. The number of traffic accidents, the number of years they had been driving and the number of kilometers driven during the last year were obtained through self-reports. A 32-item hidden Figures test (HFT) developed by Jackson, Messick and Myers (1964) was used to obtain scores on field dependence - independence.

Out of the total of 285 subjects, 28.8% reported at least one accident during the previous three years. The number of accidents ranged from 0 to 6 with a mean of .36 and standard deviation of .68. The subjects declared having driven an average of 8841 kilometers during the previous year. Separate analyses for males and females partialling out age, annual driving distance travelled and number of years driving were performed. No significant correlation was found between field dependence and traffic accidents. Using a median split on the field dependence - independence scores on the subsample of individuals who declared having been involved in an accident, it was found that there was no significant difference between field dependent and field independent drivers in the rating of importance attributed to the 21 causes of accidents.

Although the findings of these studies produced conflicting results, it would seem logical that people who are field dependent might have more difficulty with the driving task, particularly in complex environments, or where there is a considerable demand on their mental capacity, as the ability to extract information from the environment is an important feature of successful driving under complex circumstances.

Mental Workload

Another aspect of cognitive ability relating to the driving task and to the design of modern motor vehicles is the operator's subjective mental workload. With increased complexity in instrumentation and layout, it is important to consider what demand the task imposes on the operator's limited resources (spare mental capacity). Therefore, the operator's subjective mental workload should be taken into consideration (Moray, 1982; Wickens, 1984). "The central concept of mental workload is the rate at which information is processed and the rate at which decisions are made and the difficulty at making the decision", (Moray, 1979, p. 13). From the standpoint of the actual system user (driver), the use of a subjective rating of task difficulty perhaps provides the most acceptable measure of workload. This is because the user would feel quite comfortable in simply stating or ranking the subjective feelings of effort or attention demands encountered in performing the task (Eggemeier, 1981; Moray, 1982; Reid, Shingledecker, and Eggemeier, 1981; as cited in Wickens, 1984).

Historically, confidence in this approach stems from the success of the Cooper-Harper scale in measuring the flyability or acceptability of a new aircraft on the basis of the subjective judgement of test pilots on the aircraft's flight handling characteristics. The original Cooper-Harper scale (Cooper and Harper, 1969), based on a 10-point scale decision-tree, is probably the oldest and best-validated subjective measure of workload in the rating of aircraft handling qualities. There is evidence that this rating scale is a reliable estimator of perceived workload (Moray, 1984). The Cooper-Harper ratings also provide a reliable, easy and valid measure within a relatively restricted domain of the tracking and manual-control task. For example, Jex and Clement (1979), in a study on measuring perceptual-motor workload in manual-control tasks, reported a correlation of .96 between the Cooper-Harper scale and a measure of spare capacity as assessed by the critical tracking task.

Essentially, the Cooper-Harper scale is used for workload rating on motor or psychomotor skills such as flying or tracking tasks. However, the original scale has been modified for broader applications for workload ratings of other activities such as perception, monitoring, evaluation, communications, and problem solving. The modified scale, which became known as the modified Cooper-Harper (MCH) scale, has all references to handling, pilot compensation, and controllability, being replaced by terms more appropriate to other human activities. The 10-point decision tree of the original scale was maintained. Terms such as task accomplishability, errors, difficulty, performance, and mental workload were used to broaden the range of applicability to situations commonly found in the modern systems (Wierwille and Casali, 1983).

Three experiments were conducted to evaluate the validity of the MCH scale on perceptual, mediational (cognitive), and communications abilities (respectively), not including motor or psychomotor activities. These experiments were carried out in a simulated aircraft environment on a GAT 1-B moving-base flight simulator.

Casali and Wierwille (1982) carried out the first experiment validating the MCH scale on perceptual activity. Six licensed pilot-subjects each flew three cross-country flights with a single load level (low, medium, or high) in each flight. The flight order was counterbalanced and the load was manipulated by varying the rate and number of "redline" danger conditions presented on oil pressure, oil temperature, cylinder head temperature, fuel tank gauges, and on a carburator ice warning indicator. The subject's task was to detect the presence of each danger condition and identify it by pressing a corresponding button on the simulator instrument panel. A correct response alleviated the danger condition and no diagnosis or compensation of danger conditions was required.

In the low load condition, only the danger conditon of carburator icing was used. Icing occurred at an average rate of once every 50 seconds. The medium load condition was limited to the fuel tank problems and the carburator icing, which occurred at an average rate of one failure every 10 seconds. In the high workload condition, danger occurred at an average rate of one failure every five seconds on all engine and fuel instruments in addition to carburator icing.

After each flight, subjects provided the rating on the MCH scale. The results indicated that the scores on the MCH scale differed significantly ($\underline{P}<.05$) for all load level comparisons (low vs medium, low vs high, and medium vs high) with the scores means increasing monotonically with load. Thus, the results indicated that the MCH scale is reliably sensitive to perceptual load.

The second experiment was conducted by Rahimi and Wierwille (1982) verifying the MCH scale on mediational activities. Another six licensed pilot-subjects participated by performing navigation tasks while flying in the simulator. Subjects were presented with a series of slides viewed through the simulator windscreen. The slides were presorted into low, medium, and high difficulty based on the number and complexity of the arithmetic and geometric operations required to solve them. The presentation of the mediational load was counterbalanced. The subjects maintained straight-and-level flight while performing the mental computations necessary for each navigational problem and verbalized the answers. Subjects were not required to execute the navigation solution in flight, thus eliminating the differential psychomotor load which can influence the navigation (mediational) tasks. Subjects provided the MCH rating immediately after each flight. The results indicated significant differences ($\underline{P}<.05$) on scores between low vs high load and between medium vs high load. Moreover, the MCH rating scale exhibited a monotonic increase with mediational load.

The third experiment (Casali and Wierwille, 1983) was designed to verify the MCH scale on verbal communications, including detection, recognition, comprehension and response. Another six licensed pilot-subjects flew three experimental flights in the simulator. A single communications load level was used in each flight, with the load level counterbalanced across flights for each subject. The task performed by the subjects consisted of aircraft control and communications. Aircraft control was invariant in difficulty across the three flights.

Subjects were instructed to maintain straight-and-level flight and carry out any commands given to them by the "tower". These commands included changes in altitude, headings, radio frequency and so on. Moreover, a series of abbreviated call signs were presented in between these commands from the tower. Each call sign consisted of two single-digit numbers and two phonetic letters combined in any order. Subjects were to transmit a "now" whenever they heard their own call sign and variations of their call sign with the correct leading digit appearing first.

This task represented a communications detection, comprehension and response task. Under the low load condition, call signs were presented at an average rate of one every 12 seconds with none of the extraneous (non-target) call signs being the presentation of the alphanumerics used in the target call signs. In the medium load condition, call signs were presented at an average rate of one every five seconds with 30% of the extraneous call signs being target permutations. For the high load condition, call signs were presented at an average rate of one every two seconds with 40% of non-target call signs being target permutations. Again, each subject provided a rating immediately after each flight. The results indicated significant differences ($\underline{P}<.05$) in MCH rating between low vs medium load and between low vs high load. Once again, the rating scale means demonstrated a monotonic increase with load level.

Taking the three experiments as a whole, the overall results indicated that the MCH scale ratings are a valid and statistically reliable indicator of overall mental workload. On the basis of these studies, Wierwille and Casali (1983) concluded that the MCH scale is a valid and reliable measure which may be used in experiments where overall mental workload is to be evaluated.

In a recent study to evaluate the decision-tree rating scale for mental workload estimation, Skipper, Rieger and Wierwille (1986), using the MCH scale and five other design variations of the scale on two independent aircraft simulator experiments, found that the MCH scale was generally more consistent than the other scales. The MCH scale was found to have a high degree of repeatability as well as high sensitivity to workload measurement. Thus, using the MCH scale, the subjective mental workload of the driver can be easily obtained.

Rationale For The Present Study

One source of contention of the earlier studies is the definition of an accident. In the Avolio et al. (1985) study, all subjects placed in the accident group were cited in the company record as having caused some damage to their vehicles during normal operations on the road, without verifying who was at fault. Clement and Jonah (1984) used subject self-report of accidents in their study. Accident data were not available to determine the nature of the accident and who was at fault. Moreover, the subjects were undergraduate students of limited age range and only 28.8% of the subjects reported having one or more accidents. As for McKeena et al. (1986), incidents such as collisions, passenger falls, attacks on staff and so on were recorded in the company's records as accidents. Again it was not possible to determine who was at fault as the 'accidents' were based on drivers' reports. Moreover, these accidents were based on two years of service with the company. It has been reported that accident reliabilities would likely be low if accident scores were obtained over only a couple of years (Mckeena et al., 1986). Therefore, accident data should be collected over a period of at least three years.

Another possible confounding factor was the inclusion of both male and female subjects in the analysis of the data in relating field dependence and traffic safety. It has been found that females tend to be more field dependent than their male counterparts (Shinar et al., 1978). With the exception of Clement and Jonah (1984), studies on field dependence had included both male and female subjects in their analyses on perceptual style. Avolio et al. (1985) employed 58 males and 14 females as subjects and McKeena et al. (1986) employed 149 males and 4 females in their study. By including both males and females in the analyses, the variability of the measure is increased. Therefore, it is recommended that data analyses on field dependence - independence be performed separately for males and females. Otherwise, the research should be conducted on either all male or all female subjects.

In the present study stricter criteria are employed. Subjects who had three or more accidents in the previous five years were classified into the accident group. Only accidents that were attributed to the drivers' errors are included. Only male drivers were employed as subjects even though the ability to generalize to the entire population is limited.

The use of the information processing model in traffic safety research is not uncommon (Avolio et al., 1985; McKeena et al., 1986; and Mihal and Barrett, 1976). However, these studies tested the subjects employing the DLT, together with a battery of tests, in isolation. The present study employed time-sharing tasks in which subjects were required to perform the DLT , response task and tracking task simultaneously. This is assumed to simulate closely the driving task in terms of information processing demands. Moreover, the DLT employed in the current study is different from those earlier studies, as it has been modified to suit the present experiment.

Although individual experiments have examined measures of attention and information processing separately as they relate to driving and traffic accidents, no previous work has attempted to gauge the relative contributions of spare mental capacity, attention switching, perceptual style, and subjective mental workload to traffic safety employing the time-sharing of two and three tasks. The present study examined drivers' SMC, ability to switch attention, subjective mental workload and perceptual style as these relate to their record of traffic accidents. Moreover, Pearson product moment correlations will be computed to explore for possible relationships between tasks performances between the two groups. A tracking (primary) task was used in replacement of the driving task and two secondary tasks (DLT and reaction time) were used in the measurement of SMC. Two levels of spare mental capacity were examined. The time-sharing of two tasks is classified as SMC1 (spare mental capacity 1) and time-sharing of three tasks is classified as SMC2 (spare mental capacity 2). The DLT also measured the drivers' ability to switch attention.

It is hypothesized that drivers who had three or more accidents in the previous five years, as compared with those who were accident-free during the same period, will have less SMC, poorer ability to switch attention, experience higher mental workload and will be more field dependent.

METHOD

Subjects

Seventy-two male Calgary Transit (CT) bus operators from the city of Calgary, participated in the study. The ages of the subjects ranged from 24 to 56 years, with a mean age of 35 years. All of them had met the medical criteria for a Class II driving licence required for the operation of buses. This involved minimium visual acuity of 6/9 (20/30) for the best eye and 6/15 (20/50) for the worst eye (with or without glasses) and "normal" hearing.

The subjects were classified into two groups with respect to their on-the-job traffic accident records. Those who had three or more 'avoidable' accidents in the previous five years were classified into one group and those who were accident-free during the same period of time formed the other group. 'Avoidable' accidents were defined by CT as "when the operator failed to take some action that he reasonably could have taken to avoid accident involvement. Calgary Transit operators are expected to drive defensively by recognizing the hazards, understanding the defence, and acting in time to avoid accidents, despite adverse road, weather and traffic conditions, or error of other drivers or pedestrians" (Vehicle Accident Appeal

Board: Transportation Department, Calgary Transit Operation, 1985, p.3). In the event that an operator disagrees with the decision of the Safety Officer with respect to the classification of an accident, the operator may appeal to the Accident Appeal Board set-up by CT.

The Accident Appeal Board was chaired by the supervisor of Safety and consisted of a senior training officer and an operator (with a minimium of 10 years safe driving experience with CT) chosen and agreed to by managment and the union. The board reviewed all accidents submitted and the affected operator was given the opportunity to appear before the board, answering questions pertaining to the accident. In making the decision regarding the classification of the accident, the appeal board took into consideration factors such as knowledge, alertness, foresight, judgement, and skill of the operator. Based on the above criteria, the final decision in classifying the accident as "avoidable' or otherwise depended on whether the driver did everything reasonably possible to avoid the accident. It should be noted that accidents involving alcohol and/or drugs were not included, as the ability for information processing and attention are affected by these substances.

These two groups of subjects were matched for age and driving experience (years as a full-time driver with CT) where possible. All subjects were recruited through telephone calls by the experimenter. A notice (see Appendix A) regarding the nature of the present study was posted on the CT bulletin board two weeks prior to contacting the bus operators. The same notice was again posted mid-way through the study, requesting for volunteers. Each subject participated voluntarily and received $5.00 for participating.

Apparatus

An Amiga personal computer (model 1000) and a 13-inch full colour monitor (model 1080), together with a mouse connected was used. It has the ability to incorporate and synchronize the auditory and visual inputs, as well as to record reaction times.

A foot pedal, similar to the automobile gas pedal, was connected together with the mouse to the right side of the main unit of the computer. The pedal was pivoted in the middle and was spring-loaded to maintain a neutral position. Pressing the pedal with the toes, or with the heel, extinguished the stimuli appearing on the screen.

A stereo head-phone model SP40A, manufactured by Superior Electronics Ltd of Canada, was connected to a SA-150 Realistic integrated stereo amplifier. The amplifier was in turn connected to the audio connector of the computer main unit. A volume control on the amplifier enabled the subject to adjust the volume of the tone and digits generated by the computer. A balance control was available to cater to the need to counter any ear dominance that the subject would have in the DLT. A number of earlier studies had reported right-ear dominance in dichotic listening (Bryden, 1969; Gopher and Kahneman, 1971; Kimura, 1967; Treisman and Geffen, 1968).

The Group Embedded Figures Tests Booklet developed by Oltman, Raskin and Witkin (1971) was employed to test for perceptual style (field dependence). Subjective mental workload was measured using the MCH scale (Skipper, Rieger and Wierwille, 1986). The wording of the MCH scale was modified to suit the present experiment - see Appendix B.

TASKS

Task A - Tracking

Task A was a pursuit tracking task with a dot and a circle of diameter 6mm and 12 mm respectively (see Figure 1). The dot (white) was generated by the computer to move continuously at a constant rate (7.7 mm per sec) and the circle (black) was moved by the mouse under the control of the subject. The background of the computer screen was blue. Subjects were told to maintain the circle around the dot at least 95% of the time. When tracking fell below 95% accuracy, the entire computer screen flashed an orange colour momentarily (200ms), providing feedback to subjects reminding them that their tracking performance was below the required 95% accuracy. The 5% error tolerance level was to ensure similar high performance by all subjects in the tracking task. The time taken for this task was 2 minutes 15 sec.

Task B - Reaction Time (RT)

Task B was a reaction time task in which subjects were required to respond by pressing the foot pedal as quickly as possible to one of two stimuli (red or green). The stimuli were presented randomly (with a limitation that the maximium number of same stimuli occurring consecutively be three) at specified intervals (between three and nine seconds) on the screen.

The diameter of the stimulus was 4.5 mm and appeared at approximately 13 mm away (centre of the dot to centre of

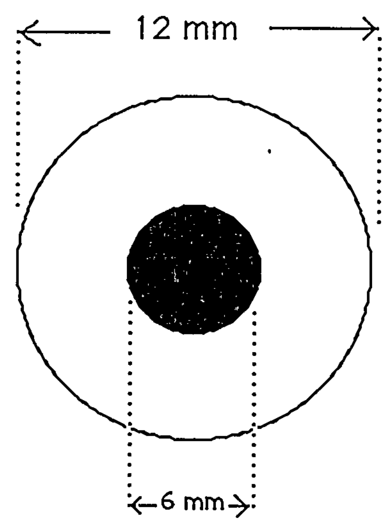

Figure 1. Dot and Circle

the stimulus) from the moving dot at one of the four possible locations (up, down, left or right) in relation to the moving dot (see Figure 2). These stimuli came on at an average of once every six seconds. The red and green stimuli were extinguished by pressing the pedal with the heel and toes respectively, which constituted correct responses. When a subject did not respond after three seconds, the stimulus was automatically extinguished and it was considered a 'missed response'. Altogether 20 stimuli were presented, with the number of red and green stimuli occurring equally (10 red and 10 green) over a period of 2 minutes 15 seconds. The first reaction time trial commenced after the target had been moving for approximately ten seconds. The dot also continued to move for approximately five seconds after the last stimulus came on. Measurements were in terms of the reaction time in milliseconds, number of wrong responses committed in responding to the stimuli, and the number of missed responses (failure to respond).

Task C - Dichotic Listening Task (DLT)

The objective of Task C was to measure the subject's ability to switch attention. A dichotic listening task (DLT) was employed using a pure tone sound to signify the relevant ear to attend. A 250 Hz tone meant the subject

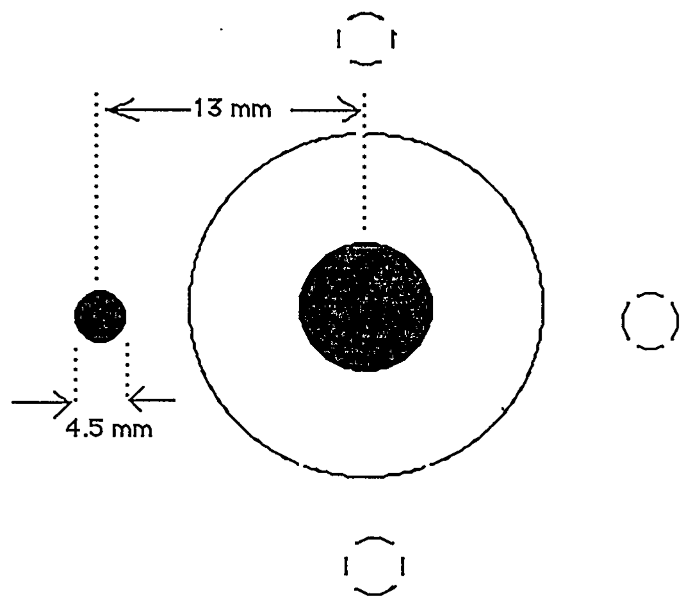

Figure 2 : Location of RT stimuli with respect to moving dot.

was to respond to digits coming into the left ear and a 2500 Hz tone, to respond to digits coming into the right ear. Following the tone, three pairs of digits were presented simultaneously, one to each ear. The subject's task was to report the three digits from the ear indicated by the tone. The computer generated the pure tone sound and digits required for this task.

Figure 3 shows the sequence of events for the DLT. The pure tone was presented for 200 ms to both ears with the volume adjusted by the subject, using the amplifier, to suit the individual. After a lapse of 1300 ms following the pure tone sound (or 1500 ms from the onset of the pure tone), pairs of digits were presented simultaneously, one to each ear. Each of the three pairs of digits was presented at an interval of 500 ms and the utterance of each digit (generated by the computer) was 500 ms. The digits (one to ten) were adapted from Procter, Ponton, and Jamieson (1986) and digitized by the Amiga computer. Subjects were to respond by reporting verbally the digits presented to the relevant ear during the 3000 ms lapse which immediately followed each set. The reported digits were recorded by the experimenter. Due to the inherent limitations of the computer, a systematic error of 50 ms existed. That is, the onset of the tone (and the rest of

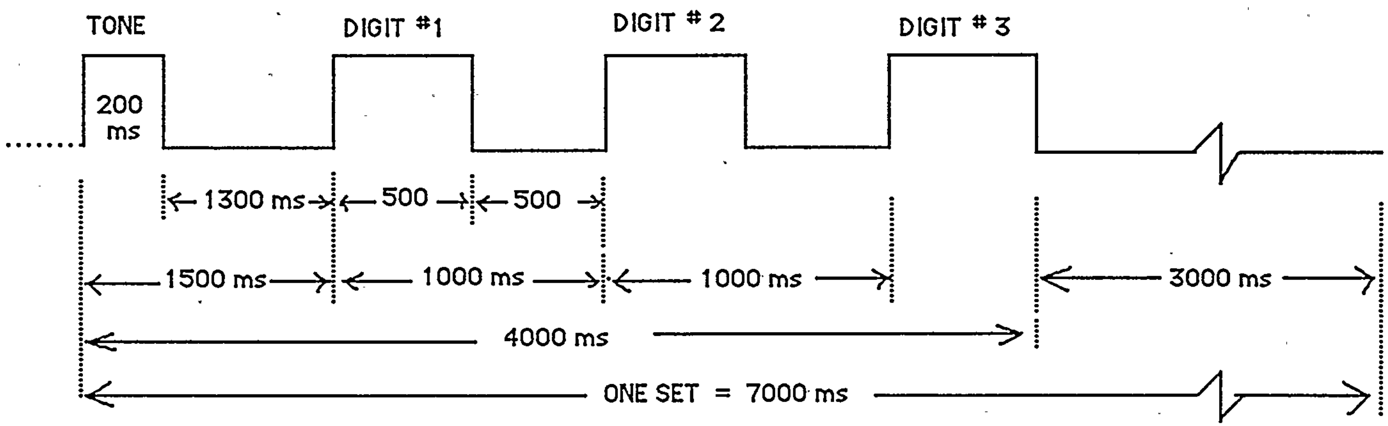

Figure 3. Sequence of Events in the Dichotic Listening Task.

the digits) may be brought forward or set back by 50 ms. Thus, the time lapse during which the subjects were to report the digits fluctuated between 2950 - 3050 ms. This however, made no difference to subjects' perception of the 3000 ms time lapse between each set of digits.

Altogether there were 18 sets of digits. Each set consisted of three pairs of digits (ranging from one to ten). Within each set, no digits were repeated. Limitation was also placed on the sequence of which ear to attend to (the pure tone sound). The same ear to attend to did not occur more than three times consecutively. The 18 tones (as there were 18 sets of digits) were arranged such that there were ten switches required of the subject for this task (see Figure 4 for an example). Measurement was in terms of omissions, intrusions, and switching errors as defined by Gopher (1982).

Task D - Tracking and RT Task

Task D was the combination of the tracking task and the RT task (Task A and Task B). Subjects were required to perform both tasks simultaneously while maintaining 95% accuracy in the tracking task. The first reaction time stimulus appeared after approximately ten seconds of tracking. This was to allow subjects to adapt to the

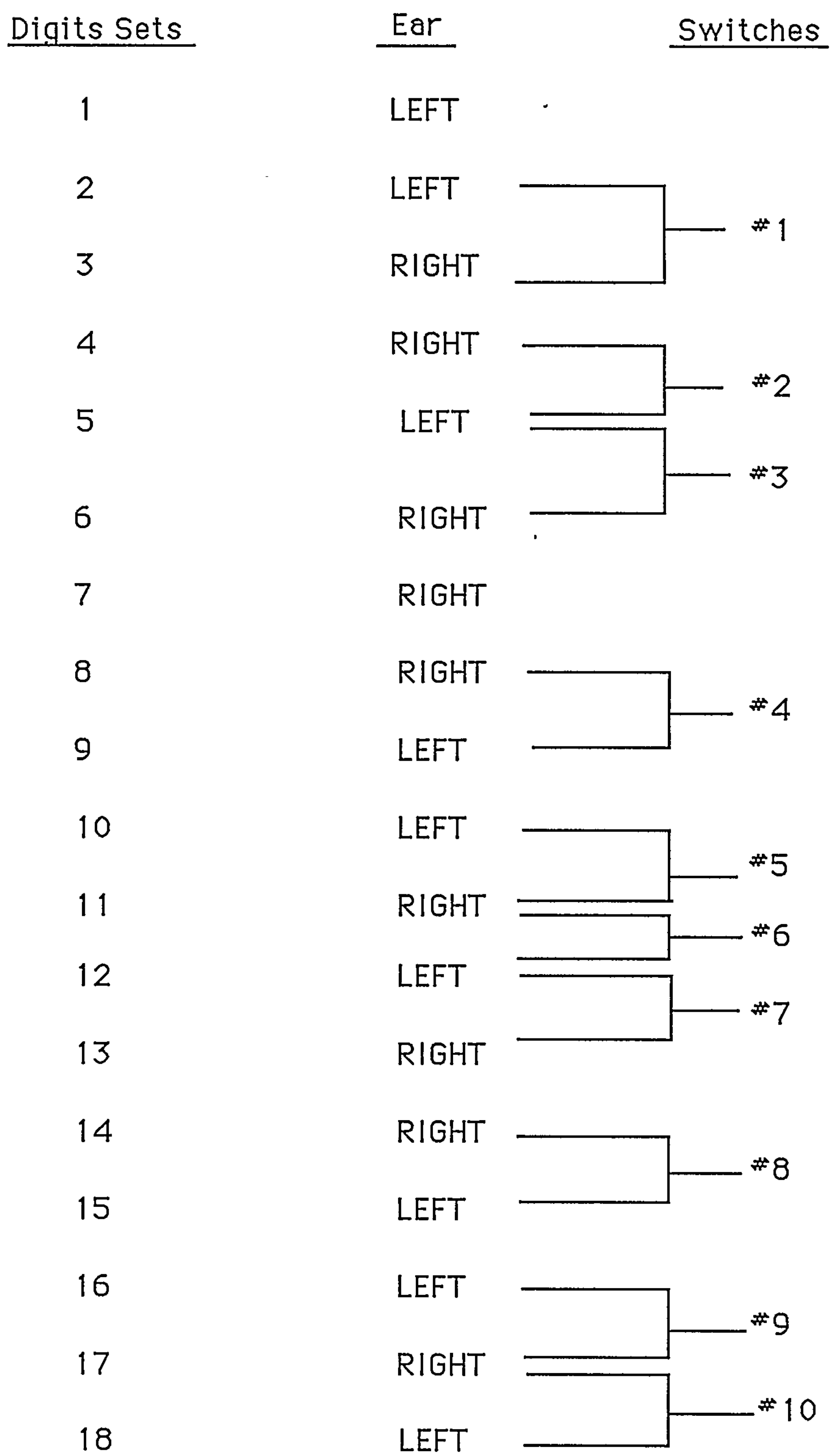

Figure 4. Pattern of Switches In Dichotic Listening Task.

tracking task. Again, there were in total 20 responses, with the number of red and green stimuli occurring equally, over the period of 2 minutes and 15 seconds. The same colour stimulus, however, was not allowed to occur more than three times consecutively. After the last stimulus came on, tracking continued for another five seconds. Measurements were in terms of tracking accuracy, reaction time, number of missed responses and number of errors committed in response to the green and red stimuli.

<u>Task E - Tracking and DLT</u>

Task E was a combination of the tracking task and the DLT (Task A and Task C). Subjects were required to perform them simultaneously while maintaining 95% accuracy in the tracking task. The onset of the first set of digits in the DLT occurred at approximately eight seconds into the tracking task. The final set of digits in the dichotic listening task was terminated on the 134th second (2 minutes 14 seconds) of the task. Measurements were in terms of tracking accuracy, omissions, intrusions, and switching errors.

<u>Task F - Combination (Tracking, RT and DLT)</u>

Task F was the combination of the three individual tasks described above. Subjects were required to perform the tracking task, the RT task and the DLT (Task A, Task B and Task C) simultaneously while maintaining 95% accuracy in the tracking. This task combination was assumed to simulate closely the demands of the actual driving situation where visual and auditory stimuli are continuously being processed. The DLT was essentially the same as described above. However, presentation of stimuli for the RT task occurred 50% of the time during DLT auditory input and 50% of the time during DLT verbal output. Measurements were in terms of tracking accuracy, input reaction time (reaction time responses while the subject was listening), output reaction time (reaction time responses during DLT output, i.e. while the subject was reporting the digits), RT errors in input (listening) and output (reporting of digits) responses, missed RT responses during input and output, and omissions, intrusions, and switching errors in the dichotic listening task. The task took 2 minutes 15 seconds.

There were altogether five different tracking patterns (generated by the computer) used for the tracking tasks (practice trial, Task A, Task D, Task E and Task F). This was to ensure that subjects could not anticipate the tracking pattern for the actual task. All subjects had the same tracking pattern for each task. Likewise, the sequences for the switching of attention (indicated by the tone) for the DLT (practice trial, Task C, Task E, Task F) and the sequences for the red and green stimuli for the RT task (practice trial, Task B, Task D, Task F) were all different. In total there were four different sequences of attention switching with four different blocks of 18 sets of digits used for the DLT, (one for all the practice trials and three for the remaining tasks - C, E and F). For each of the tasks, all subjects went through the same attention switching sequence (eg. see Figure 4) with the same 18 sets of digits, but each task had a different set of 18 digits. For the RT task, there were also four different sequences used for the red and green stimuli. One sequence was used for all the practice trials and three were used for the other tasks (B, D, and F). All subjects again had the same sequence for each task.

Procedure

The study was conducted in the presence of the experimenter and was carried out at the three main Calgary Transit garages at which the subjects reported for work. At each garage, the inspectors' room (approximately 5 meters by 7 meters) or a small training room (7 meters by 10 meters) was used to conduct the study. In the three rooms, all natural light was blocked out and normal office white diffuse lighting was used. Noise was minimal (not distracting) in all three rooms.

Subjects participated in this study during their off-day, just before they began their work, or in between their shifts (ie., they had driven for approximately four hours prior to participating in the study). None of the subjects took the test after a full day's work.

The study was broken down into two sections with section one always preceeding section two. The GEFT was administered in section one, and section two consisted of the six tasks (Task A to Task F). Appropriate instructions in written form were given prior to each task (see appendix C for all the instructions). Before the commencement of section two, subjects were asked whether they played video games in the past two years. It was felt that experience in playing video games might have effects on the present study.

A within-subjects repeated measures design was employed in section two, consisting of two phases. Phase I consisted of performance of all the single tasks (Task A, Task B, and Task C) with the order of presentation for Tasks B and C being randomized for all the subjects, and Task A always being performed first.

Phase II, which followed immediately after Phase I, consisted of the performance of the time-sharing tasks (Task D, Task E and Task F). These three tasks were presented in random order with a restriction placed on Task F, which followed D and E. The purpose was to examine the influence of these time-sharing tasks, while maintaining 95% accuracy on the primary (tracking) task so as to establish the amount of spare mental capacity available to drivers under these conditions.

For each of the six tasks, there was a practice trial to familiarize the subjects with the required task, prior to the performance of the actual task. If subjects could not maintain 95% accuracy in the first practice trial, more practice was given to bring the subjects' tracking performances up to 95% accuracy level. And if subjects failed to follow the instructions for the DLT in the first practice trial, more practice was provided until subjects could perform the DLT task. Practice trials in Phase I (single tasks) were one minute each and practice trials for tasks in Phase II were 2 minutes 15 seconds each. There was a one-minute rest period between the practice trials and the actual tasks.

Subjective mental workload was measured immediately after each task, using the MCH rating scale. Another rest period of one minute was provided before the practice trial for the next task.

The total time taken for each subject in the study was between an hour and an hour 15 minutes, depending on the number of practice trials required by the subject.

RESULTS

The abbreviations defined in the glossary on page xv will be used in the results section for ease of reference.

Subjects in the accident group had an average of 3.53 on-the-job accidents in the previous five years of their driving with Calgary Transit. The number of accidents ranged from three to six. Table 1 shows the overall, accident group and accident-free group means and standard deviations of the variables measured.

Practice trials were given to all subjects to bring their tracking performance up to the 95% accuracy level for Tasks A, D, E and F. The 95% accuracy criterion had to be met for the practice trials to ensure a high level of performance. However, this criterion was not intended for the actual trials. This was partly because the tracking patterns for each of the four actual tracking tasks were different from those in the practice. We would therefore expect tracking performance might fall below the 95% level in the actual task.

As anticipated, some subjects (especially those in the accident group) did not manage to maintain the 95% accuracy level when performing the actual tasks for the reason stated above. However, all subjects' data were included in

Table 1

Means and Standard Deviations for all Variables

	OVERALL (N=72)		ACCIDENT (n=36)		ACCIDENT-FREE (n=36)	
VARIABLE	MEAN	STD DEV	MEAN	STD DEV	MEAN	STD DEV
ACCIDENTS	-	-	3.53	.84	-	-
AGE	35.04	7.53	35.86	7.90	34.22	7.15
DX	7.11	2.57	7.00	2.61	7.22	2.55
GEFT	10.56	4.84	10.05	4.95	11.08	4.74
TA	97.51	1.52	97.17	1.83	97.85	1.04
TD	95.66	3.09	95.00	3.59	96.32	2.36
TE	96.53	3.05	95.84	3.62	97.21	2.18
TF	94.71	4.29	93.70	5.29	95.73	2.68
RTB	599.79	96.28	599.47	90.74	600.11	102.80
RTD	725.55	121.82	713.13	93.15	737.97	145.30
RTF	909.34	196.77	920.83	220.82	897.86	171.79
ERRTB	1.50	1.30	1.52	1.13	1.47	1.46
ERRTD	1.93	1.58	2.13	1.69	1.72	1.46
ERRTF	2.45	2.21	2.91	2.79	2.00	1.30
RTIN	981.86	240.14	1009.00	253.74	954.72	226.01
RTOUT	826.04	166.07	811.91	171.67	840.16	161.45
ERIN	.84	1.15	.91	1.42	.77	.83
EROUT	1.62	1.40	2.02	1.66	1.22	.95
DLTC	2.80	2.57	3.25	2.30	2.36	2.77
DLTE	2.98	2.88	3.67	2.88	2.30	2.75
DLTF	5.22	3.67	6.30	3.98	4.13	3.01
DLCOM	.23	.51	.33	.63	.13	.35
DLEOM	.19	.74	.30	.98	.08	.36
DLFOM	.29	.82	.44	1.10	.13	.35
DLCIN	1.18	1.37	1.25	1.25	1.11	1.50
DLEIN	1.09	1.17	1.30	1.23	.88	1.08
DLFIN	2.02	1.89	2.41	2.07	1.63	1.64
DLCSW	1.36	1.45	1.66	1.54	1.05	1.30
DLESW	1.65	1.74	2.05	1.77	1.25	1.64
DLFSW	2.90	2.09	3.44	2.32	2.36	1.69
SMLA	2.75	1.08	2.86	1.01	2.63	1.15
SMLB	2.58	1.31	2.77	1.09	2.38	1.49
SMLC	4.34	1.88	4.41	1.77	4.27	2.00
SMLD	4.38	1.43	4.30	1.32	4.47	1.55
SMLE	4.77	1.62	4.77	1.53	4.77	1.72
SMLF	6.12	1.78	6.00	1.65	6.25	1.91
VI	1.48	.50	-	-	-	-

Table 1 (continued)

Means and Standard Deviations for all Variables

	OVERALL (N=72)		ACCIDENT (n=36)		ACCIDENT-FREE (n=36)	
VARIABLE	MEAN	STD DEV	MEAN	STD DEV	MEAN	STD DEV
RT RED	726.12	149.72	-	-	-	-
RT GREEN	735.43	152.83	-	-	-	-
MISRD	.01	.11	0.0	0.0	.02	.16
MISRF	.40	1.41	.55	1.87	.25	.69
MISRIN	.29	.86	.36	1.04	.22	.63
MISROUT	.11	.61	.19	.85	.02	.16

the analyses, as all of them were able to maintain the 95% accuracy level during practice. Moreover, the average tracking performance for Task F (the most difficult task) was close to the 95% accuracy level (94.7%).

Preliminary analysis indicated that there were no significant differences between the accident and accident-free groups of drivers in age and driving experience. Employing t-Tests, no significant difference was found in the GEFT scores between the two groups ($t(70) = -.90$, $\underline{P}=.371$). It was also found that playing of video games was not related to performance in any of the tasks employed in the present study.

A repeated measures ANOVA using the BMDP2V programme (Dixon, Brown, Engleman, Frane, Hill, Jennrich and Toporek; 1985) and Tukey's method of pairwise comparisons, (Glass and Hopkin, 1984), were employed, unless otherwise stated.

Tracking Tasks

The mean tracking accuracy for all the subjects in Tasks A, D, E and F was 97.5%, 95.7%, 96.5% and 94.7% respectively. Table 2 shows the summary of the repeated measures ANOVA of the tracking tasks between the two groups. It was found that the accident-free group displayed significantly superior tracking performance over

Table 2

Group By Task ANOVA Summary Table (Tracking Performance)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	131.62	1	131.62	4.54*
ERROR	2028.08	70	28.97	
TRACKING TASK (T)	308.46	3	102.82	33.91***
T by G	16.23	3	5.41	1.78
ERROR	636.82	210	3.03	

* P<.05
*** P<.001

the accident group in the tracking tasks ($F(1,70) = 4.54$ $p<.05$). Moreover, the result also indicated significant main effects for taskload ($F(3,210) = 33.91$, $p<.001$), that is, tracking accuracy between Task A, Task D, Task E and Task F. Post-hoc pairwise comparisons found all subjects' mean tracking performance differed significantly from one task to the other ($p<.05$), with tracking performance deteriorating as taskload increased. Tracking in Task F was poorer than tracking in Task D and Task E; tracking in Task D and task E were poorer than tracking in Task A, as shown in Figure 5. This indicated that spare mental capacity (SMC) decreased for subjects as taskload increased. However, tracking for Task E (DLT and tracking) was found to be significantly superior compared with Task D (RT task and tracking) for the time-sharing of two tasks ($p<.05$).

Dichotic Listening Task

Measures of omission, intrusion, switching and overall errors were analyzed to find which of these discriminate between the accident and accident-free groups. Tables 3 and 4 show group differences in the overall DLT ability and attention switching ability, respectively. No interaction effect was found between the number of overall errors in the DLT and the two groups (Table 3). However, it was

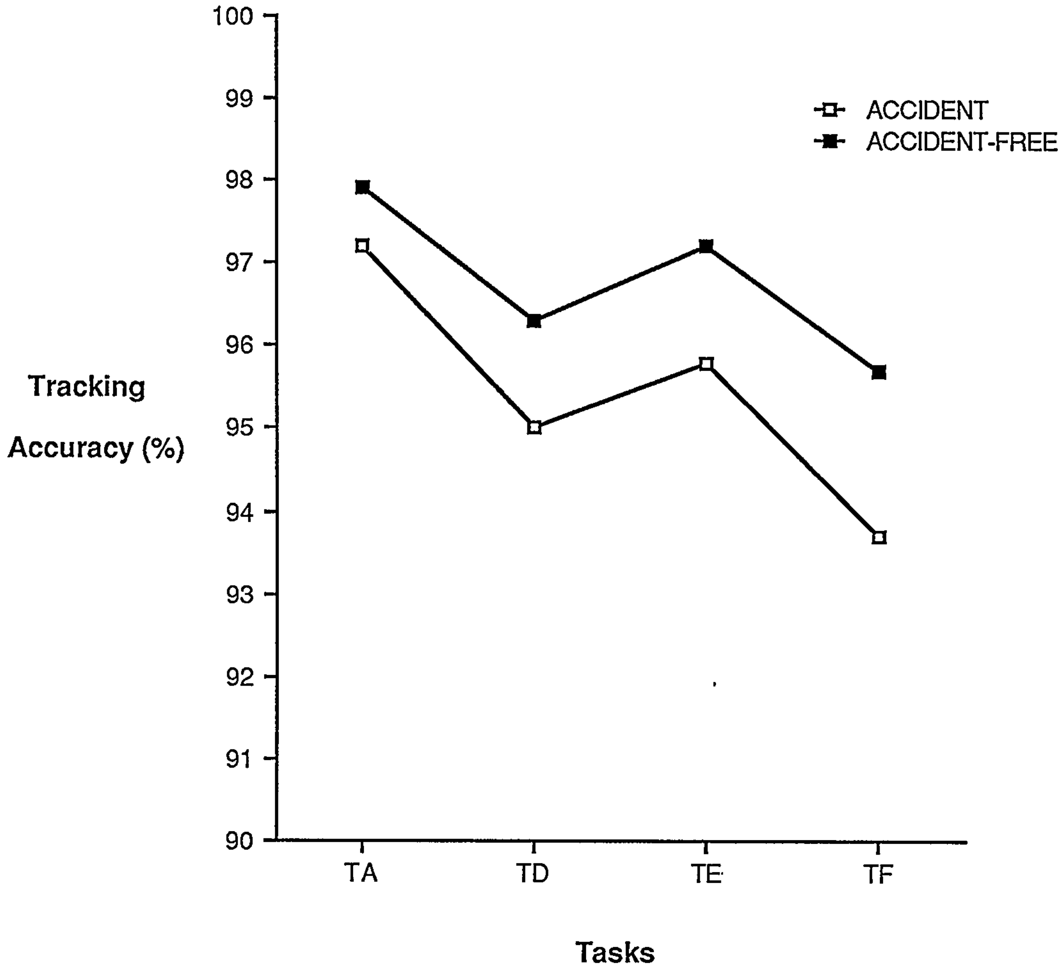

Figure 5. Tracking Accuracy for each Group as a function of Task

Table 3

Group By DLT ANOVA Summary Table (Overall Errors)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	117.04	1	117.04	5.89*
ERROR	1391.95	70	19.88	
DLT	260.95	2	130.47	36.78***
DLT by G	15.02	2	7.51	2.12
ERROR	496.68	140	3.54	

Table 4

Group By DLT ANOVA Summary Table (Switching Errors)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	37.50	1	37.50	6.47*
ERROR	405.66	70	5.79	
DLT	96.58	2	48.29	29.13***
DLT by G	2.02	2	1.01	0.61
ERROR	232.05	140	1.65	

* P<.05
*** P<.001

found that the accident group made significantly more overall errors in the DLT than did the accident-free group ($\underline{F}(1,70) = 5.89$, $\underline{P}<.05$) in all DLT taskload conditions (Task C (3.25 versus 2.36), Task E (3.67 versus 2.30), and Task F (6.30 versus 4.16)) as shown in Figure 6. The result also indicated a significant main effect of DLT ($\underline{F}(2,140) = 36.78$, $\underline{P}<.001$). A post-hoc pairwise comparison found significant differences in the number of overall errors in DLT between Task C (mean = 2.80) and Task F (mean = 5.22) ($\underline{P}<.001$), and between Task E (mean = 2.98) and Task F (mean = 5.22) ($\underline{P}<.001$), but not between Task C and Task E. This indicated that SMC was reduced as taskload increased from SMC1 (involving two tasks) to SMC2 (involving three tasks) and from single task to time-sharing of three tasks for all subjects, but not between the single task and time-sharing of two tasks.

When the overall DLT errors were broken down into omission, intrusion, and switching errors, analyses employing repeated measure ANOVA between the two groups and each of the three categories of errors indicated no significant interaction effects. It was also found that there were no significant differences between the two groups in omission and intrusion errors. However, in terms

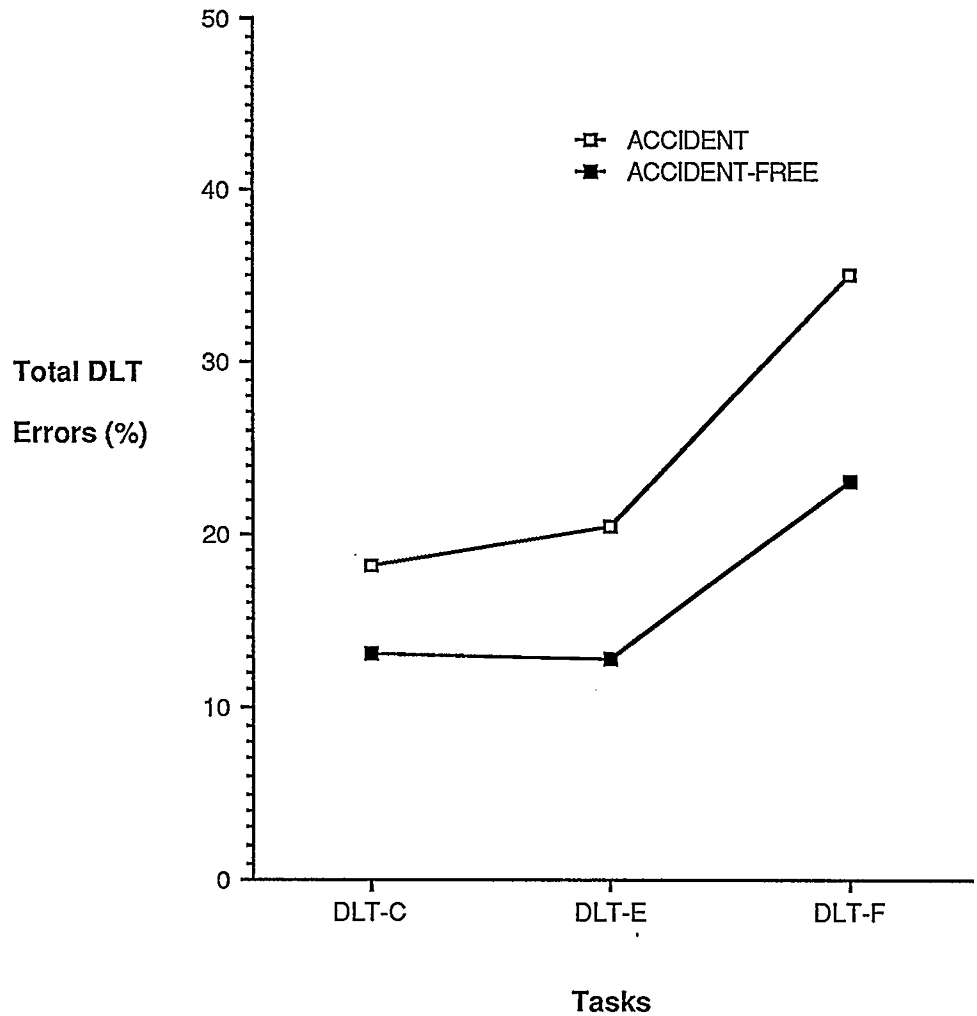

Figure 6. Total DLT Errors for each Group as a function of Task

of switching errors, the accident group made significantly more switching errors than did the accident-free group ($F(1,70) = 6.47$, $P<.05$) in all taskload conditions (Task C (1.66 versus 1.05), Task E (2.05 versus 1.25), and Task F (3.44 versus 2.36)) as shown in Figure 7. Thus, switching error discriminated between the two groups of drivers. The results also showed a significant main effect of DLT task ($F(2,140) = 29.13$, $P<.001$). A post-hoc pairwise comparison found significant differences in the number of switching errors between Task C (mean = 1.36) and task F (mean = 2.90) ($P<.001$), and between Task E (mean = 1.65) and Task F (mean = 2.90) ($P<.001$), but not between Task C and Task E. This indicated that switching error is a good measure of SMC, with SMC decreasing as taskload increased from a single task to SMC2 and from SMC1 to SMC2, but no significant decrement in SMC from a single task to a dual task condition.

For the DLT omission errors, there was no significant difference between the two groups and no significant main effect. As for intrusion errors, Table 5 indicated only a significant main effect for taskload ($F(2,140) = 16.76$, $P=<.001$). A post-hoc pairwise comparison found the same significant results as those in the DLT overall errors and the DLT switching errors. That is, significantly different

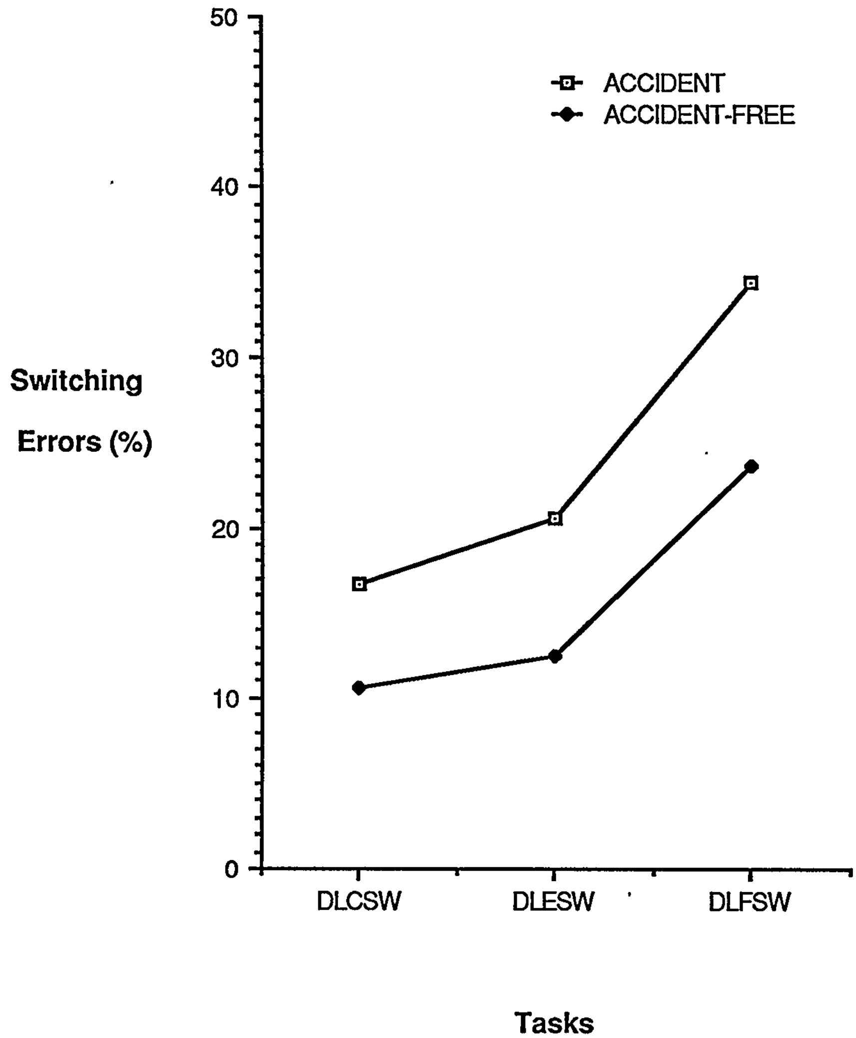

Figure 7. Switching Errors for each Group as a function of Task

Table 5

Group By DLT ANOVA Summary Table (Intrusion Errors)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	10.66	1	10.66	2.37
ERROR	315.09	70	4.50	
DLT	38.17	2	19.08	16.76***
DLT by G	3.69	2	1.84	1.62
ERROR	159.46	140	1.13	

*** P<.001

numbers of intrusion errors were made between Task C (mean = 1.18) and Task F (mean = 2.02) ($\underline{P}$=<.001), and between Task E (mean = 1.09 and Task F (mean = 2.02) ($\underline{P}$=<.001), but not between Task C and Task E. This again indicated that SMC decreased as taskload increased from a single task to SMC2, and from SMC1 to SMC2, but there was no significant decrement of SMC from a single task condition to a dual task condition.

Reaction Time Task

For the response task, only reaction time (RT) of correct responses were used in the analyses.

Tables 6 and 7 show the summary of the repeated measures ANOVA for the task and the errors committed in the RT task respectively for the two groups. It was found that there were no significant differences between the two groups in their RT or the number of errors in their responses. There was no interaction between taskload and groups. However, significant main effects were found for the RT ($\underline{F}(2,140) = 136.21$, $\underline{P}<.001$) and for errors in responses ($\underline{F}(2,140) = 7.67$, $\underline{P}<.001$). Post-hoc pairwise comparisons found that all RT's differed significantly one from the other ($\underline{P}<.001$) with Task B having the fastest RT and Task F, the slowest RT.

Table 6

Group By Response Task ANOVA Summary Table (Reaction Time)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	37.50	1	37.50	0.00
ERROR	2646942.92	70	37813.47	
RESPONSE TASK (RT)	3490093.78	2	1745046.89	136.21***
RT by G	20569.36	2	10284.68	0.80
ERROR	1793578.18	140	12811.27	

Table 7

Group By Response Task ANOVA Summary Table
(Reaction Time Errors)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	11.57	1	11.57	2.49
ERROR	325.46	70	4.64	
RESPONSE TASK (RT)	33.17	2	16.58	7.67***
RT by G	6.73	2	3.36	1.56
ERROR	302.75	140	2.16	

*** P<.001

In terms of errors committed in the responses, post-hoc pairwise comparisons found only significant differences between Tasks B (mean = 1.50) and F (mean = 2.45) ($\underline{P}<.001$) and no significant differences between Tasks B and D, or between Tasks D and F. Once again, the phenomenon of SMC was displayed. As taskload increased from a single task condition to the time-sharing of three tasks, RT increased significantly and errors in responses were significantly more in the time-sharing of three tasks condition as compared to the single task condition.

Although 3000 ms were allotted for responding to each of the stimuli in the RT task, one subject did not respond to one stimulus (5%) in Task D. In Task F, thirteen subjects (18%) failed to respond at an average of 11.15 % (2.23) of the stimuli. Under the Task B condition all the subjects responded to all the stimuli. A two by two repeated measures ANOVA was carried out between groups and tasks (D and F). Task B was not included in the analysis, as all the subjects had responded to all the stimuli. No significant group differences or interaction effects were found (Table 8). However, a significant difference was found between Task D (mean = .01) and Task F (mean = .40) ($\underline{F}$ (1,70) = 5.40, $\underline{P}<.05$) with Task F having more missed

Table 8

Group By Response Task ANOVA Summary Table
(Missed Responses)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	0.69	1	0.69	0.69
ERROR	70.05	70	1.00	
RESPONSE TASK (RT)	5.44	1	5.44	5.40*
RT by G	1.00	1	1.00	0.99
ERROR	70.55	70	1.00	

* P<.05

responses (subjects failed to respond). It therefore indicated that as taskload increased, SMC decreased.

Reaction Time Task Conditions: Input vs Output

Table 9 shows the analysis of RT during DLT input (RT while subject was listening) and during DLT output (RT when subject was reporting the digits) by groups. A significant interaction effect (Figure 8) was found ($F(1,70) = 4.48$, $P<.05$). Therefore interpretation of main effects is generally not meaningful (Pedhazur, 1982; Kirk, 1968). An analysis of simple main effects (Kirk, 1968) was conducted and it was found that there was a significant difference in RT between Input (listening to the digits) (mean =981.86 ms) and Output (reporting the digits) (mean = 826.04 ms), with RT being significantly slower during Input for both groups ($P<.001$). However, there were no significant group differences.

An interaction effect ($F(1,70) = 4.73$, $P<.05$) was also found in the number of RT errors during input and output (Table 10 and Figure 9). An analysis of simple main effects shown in Table 11 indicates significant differences in RT errors during output between the two groups ($P<.01$) with the accident group having significantly more errors (mean = 2.02) than the accident-free group (mean = 1.22).

Table 9

Group By Response Task Condition ANOVA Summary Table
(Input-Output RT)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	6097.00	1	6097.00	0.08
ERROR	5026647.15	70	71809.24	
RESPONSE TASK CONDITION (RTC)	874069.17	1	874069.17	63.82***
RTC by G	61297.50	1	61297.50	4.48*
ERROR	958743.81	70	13696.34	

*   P<.05
***   P<.001

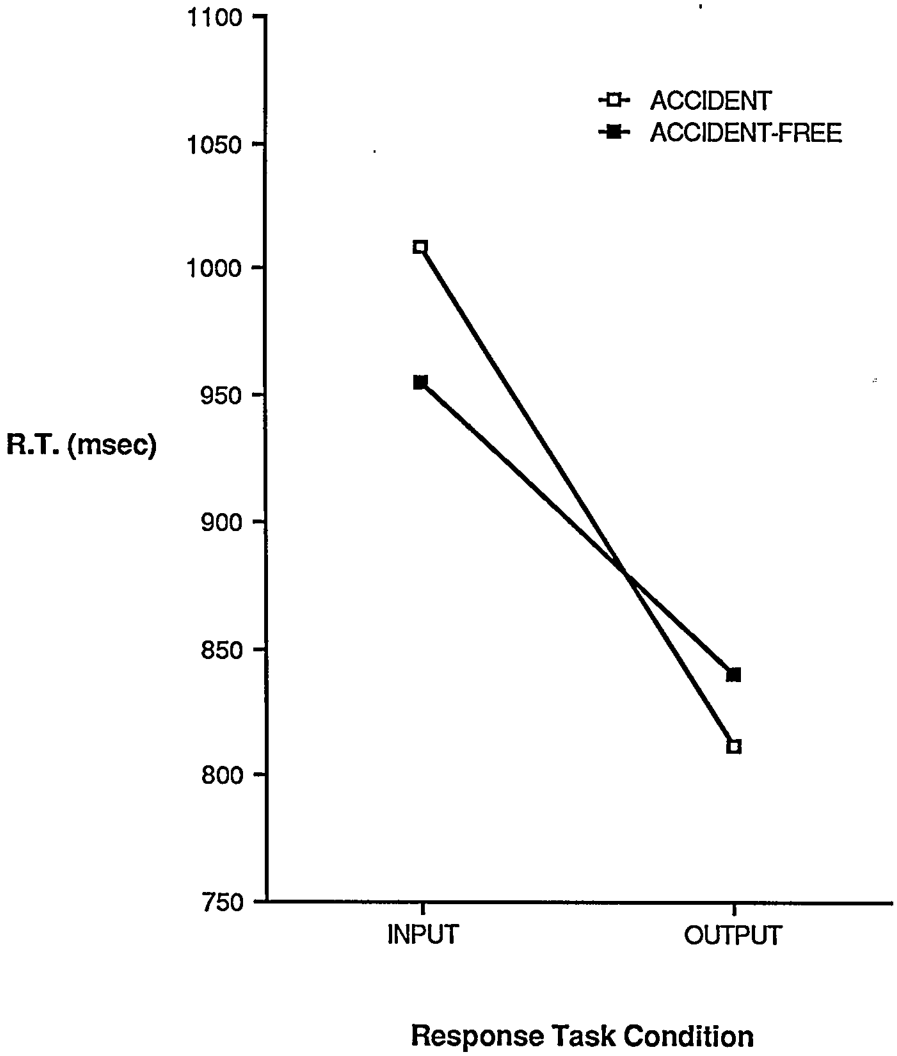

Figure 8. RT Performance for each Group as a function of Response Task Condition

Table 10

Group By Response Task Condition ANOVA Summary Table
(Input-Output RT Errors)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	8.02	1	8.02	3.41
ERROR	164.94	70	2.35	
RESPONSE TASK CONDITION (RTC)	21.77	1	21.77	25.74***
RTC by G	4.00	1	4.00	4.73*
ERROR	59.22	70	0.84	

Table 11

Simple Main Effect of Group by Input-Output (RT Errors)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
ACCID GROUP AT INPUT-OUTPUT	22.222	1	22.22	26.27 ***
ACCID-FREE GROUP AT INPUT-OUTPUT	3.56	1	3.56	4.20
ERROR (CELL)	59.22	70	.84	
INPUT AT ACCID/ACCID-FREE	.347	1	.34	<1.00
OUTPUT AT ACCID/ACCID-FREE	11.68	1	11.68	7.30 **
ERROR (POOLED)	222.17	140	1.60	

* P<.05
** P<.01
*** P<.001

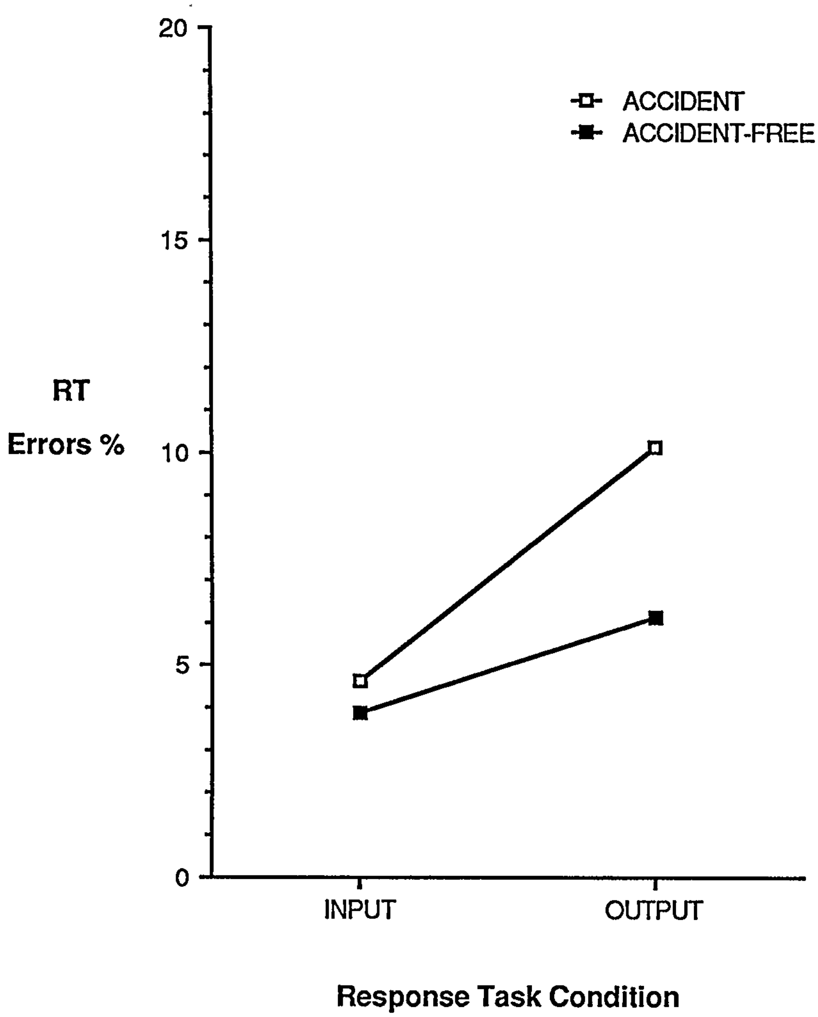

Figure 9. Input-Output RT Errors for each Group as a function of Response Task Condition

The groups did not differ significantly in their RT errors during input. Moreover, there were significantly more output errors (mean = 2.02) than input errors (mean = .91) for the accident group ($\underline{P}$<.001) but not for the accident-free group. Table 12 showed no significant negative correlation between output RT (RTOUT) and output errors (EROUT), which indicated that speed-accuracy trade-off did not occur. This was important because RTOUT was faster than RTIN, with EROUT having significantly more errors than input RT (ERIN).

Table 13 shows the analysis of groups by missed reponses during input (MISRIN) and output (MISROUT) during the RT task. Only a significant main effect in the response conditions was found ($\underline{F}$(1,70) = 8.82, $\underline{P}$<.01). The result indicated that more responses were missed during DLT input (2.9%) than during DLT output (1.1%). Again no trade-off between speed and missing reponses was found, as indicated by the positive correlation between RTIN and MISRIN, and between RTOUT and MISROUT (Table 12).

<u>Perceptual Style by Groups</u>

Although T-tests did not show any significant differences between the two groups in the GEFT scores, an attempt was made to explore the relationship of field

Table 12

Correlation Matrix of Input-Output (RT Task)

	RTIN	RTOUT	ERIN	EROUT	MISRIN	MISROUT
RTIN						
RTOUT	.708**					
ERIN	-.082	.025				
EROUT	-.001	-.167	.473**			
MISRIN	.616**	.399**	.017	.056		
MISROUT	.458**	.218	-.035	.145	.810**	

* P<.01
** P<.001

Table 13

Group By Response Task Condition ANOVA Summary Table
(Input-Output Missing Responses)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	0.84	1	0.84	0.84
ERROR	69.81	70	0.99	
RESPONSE TASK CONDITION (RTC)	1.17	1	1.17	8.82**
RTC by G	0.01	1	0.01	0.05
ERROR	9.31	70	0.13	

** P<.001

dependence - field indpendence with respect to the two groups in their taskload performances. Analyses of GEFT (low, medium and high scores) by tasks were performed for each group separately.

The overall GEFT scores had a mean of 10.56 and standard deviation of 4.84. The scores were divided into three categories with one standard deviation below the mean scores being categorized as 'Low', one standard deviation above the mean scores as 'High', and with those falling in between as 'Medium'. Although dividing the GEFT scores into quartiles has been reported (Loo, 1978; Witkin et al., 1971), the present study used one standard deviation above and below the sample mean to categorize field dependence - independence. There is no standard or norm for classifying field dependence - independence, as the mean and standard deviation of the GEFT scores vary from sample to sample (Witkin et al., 1971). In the present study, dividing the GEFT scores into quartiles shows clustering of scores around the first and second quartiles and also around the third and fourth quartiles. However, dividing the scores into three categories using one standard deviation shows fairly clear demarcations of scores with less clustering around the dividing points. Furthermore, it is the low and high GEFT scores that are of interest here. Thus, it was considered statistically prudent to categorize the GEFT scores into these three categories using one standard deviation above and below the sample mean to differentiate between field dependent and field independent subjects.

The maximum possible score for the GEFT by any subject was 18. Scores between 0 to 6 were classified as 'Low' or field dependent, between 7 to 15 as 'Medium', and scores between 16 to 18 as 'High' or field independent. Table 14 shows the number of subjects within each category of the two groups with their mean GEFT scores.

Among all the taskloads analysed, a significant difference was found only for GEFT by group by tracking task (Table 15). A significant two-way interaction was found between GEFT and group ($\underline{F}(2,66) = 3.43$, $\underline{P}<.05$) as indicated in Figure 10. An analysis of simple main effects followed up by post-hoc pairwise comparisons found significant differences between the accident group and the accident-free group ($\underline{P}<.01$) for those who had low GEFT scores (field dependent subjects) and not for those who had medium and high GEFT scores. That is, field dependent subjects in the accident group performed poorer in all tracking tasks than did those in the accident-free group. It was also found that GEFT scores were able to discriminate tracking performance within the accident

Table 14. Number of Subjects in each Group as a function of GEFT.

	GEFT		
	LOW	MEDIUM	HIGH
ACCIDENT GROUP	$\underline{n}$= 10 ($\bar{x}$= 3.9)	$\underline{n}$= 20 ($\bar{x}$= 11.1)	$\underline{n}$= 6 ($\bar{x}$=17.3)
ACCIDENT-FREE GROUP	$\underline{n}$= 6 ($\bar{x}$= 3.0)	$\underline{n}$= 24 ($\bar{x}$= 11.6)	$\underline{n}$= 6 ($\bar{x}$= 16.8)

Table 15

Group By GEFT By Tracking Task ANOVA Summary Table
(Tracking Performance)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	118.58	1	118.58	5.11*
GEFT	248.88	2	124.44	5.36**
G by GEFT	159.33	2	79.66	3.43*
ERROR	1532.75	66	23.22	
TRACKING TASK (T)	239.11	3	79.70	27.10***
T by G	18.21	3	6.07	2.06
T by GEFT	31.36	6	5.22	1.78
T by G by GEFT	14.96	6	2.49	0.85
ERROR	582.41	198	2.94	

* P<.05
** P<.01
*** P<.001

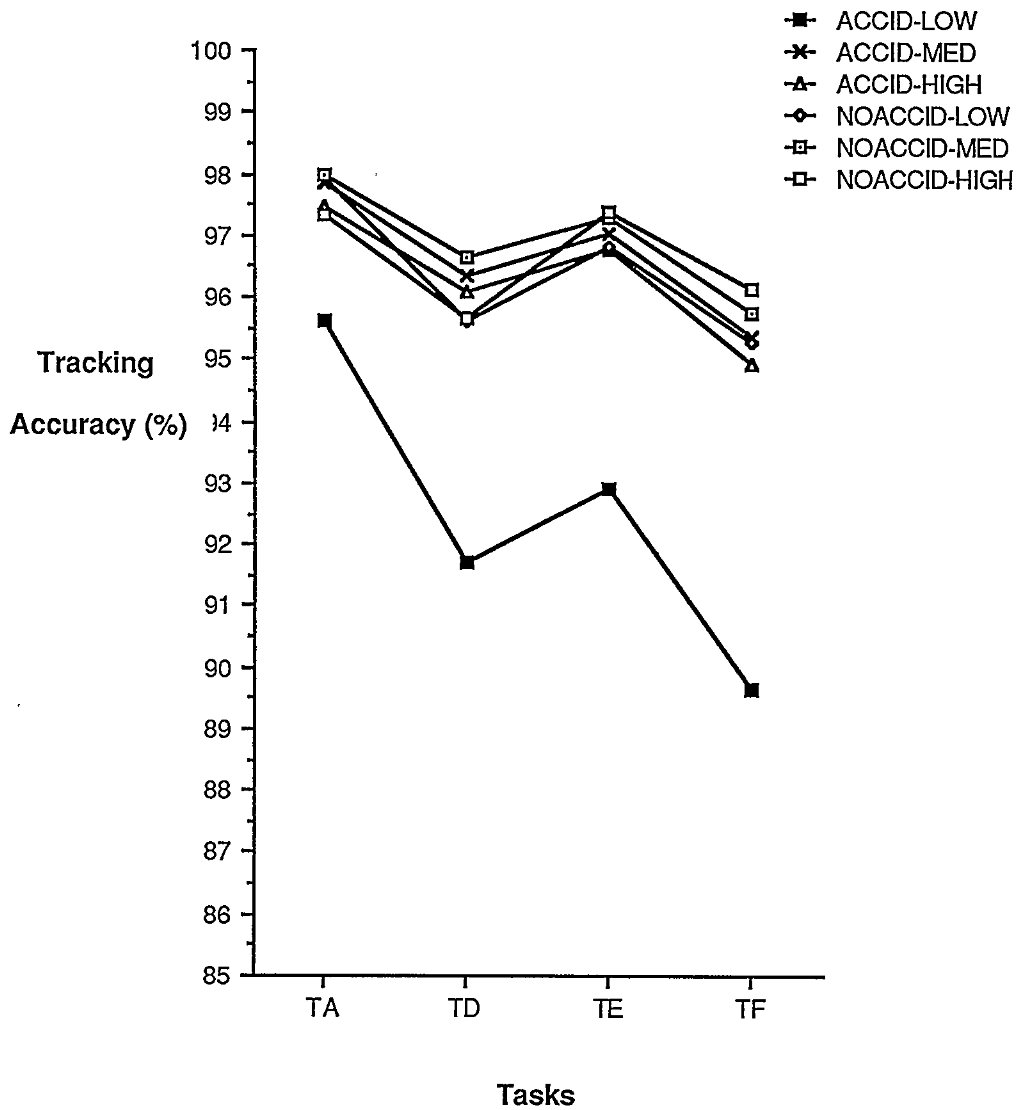

Figure 10. Tracking Accuracy for GEFT within each Group as a function of Task

group. Significant differences in tracking performance were found between those who had low and medium GEFT scores ($p<.01$), and between the low and high GEFT scores ($p<.05$). No significant difference was found between the medium and high scorers in their tracking performances. However, GEFT scores did not discriminate tracking performances of those in the accident-free Group.

An analysis comparing GEFT and the number of accidents within the accident group was also conducted. It was found that the mean number of accidents for the low, medium and high GEFT scores were 3.70, 3.45 and 3.50 respectively. A univariate ANCOVA with age and driving experience as covariates indicated no significant differences between the field dependent and field independent subjects in the number of accidents committed.

Subjective Mental Workload

Table 16 and Figure 11 show the results of the SML ratings of the accident and accident-free groups. No significant differences were found in the ratings between the two groups. However, a significant main effect of task was found for SML ratings ($F(5,350) = 106.37$, $p<.001$). SML ratings increased monotonically as taskload increased, which indicated that subjects rated mental workload as

Table 16

Group By SML ANOVA Summary Table

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
GROUP (G)	0.33	1	0.33	0.04
ERROR	595.99	70	8.51	
SML	633.90	5	126.78	106.37***
SML by GP	5.25	5	1.05	0.88
ERROR	417.17	350	1.19	

*** P<.001

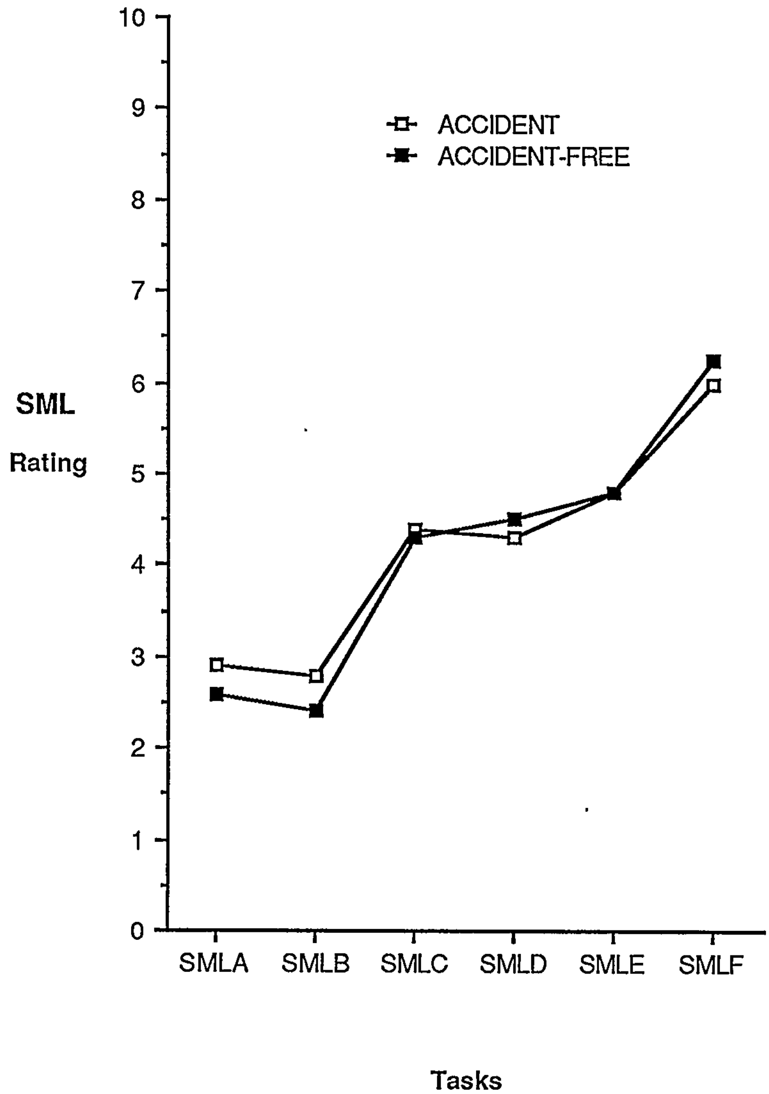

Figure 11. SML Rating For each Group as a function of Task

increasingly heavier as taskload increased from a single task situation to the time-sharing of three tasks. Post-hoc pairwise comparisons found no significant differences between Tasks A and B, and between Tasks C, D and E. However, significant differences ($\underline{P}<.001$) in mental workload ratings were found between the single task conditions A (mean = 2.75) and B (mean = 2.58) (but not Task C (mean =4.34)) and the time-sharing tasks conditions (Tasks D, E and F with means of 4.38, 4.77 and 6.12, respectively). The SML rating of Task C was found to be significantly different ($\underline{P}<.001$) from Tasks A and B, and from Task F. It was also found the the rating differed significantly ($\underline{P}<.001$) between Task F and those of Tasks D and E.

Task Correlations

Pearson product moment correlations were computed to test the relationship among task performance measures for the time-sharing tasks, with both groups combined, as well as for each group separately for Task F. Correlations among variables within each time-sharing tasks were explored separately. Correlations for Tasks A, B and C (all single tasks) which served as base-line and practice for the time-sharing tasks of Tasks D, E and F were not explored.

Table 17,18 and 19 show correlations of variables in Task D, Task E and Task F respectively. A probability level of less than .01 ($\underline{P}<.01$) instead of $P<.05$ was chosen as the significant level because of the large number of variables in the correlation matrix (41 by 41). Table 17 shows the overall correlations of Task D. Significant negative correlations were found between tracking and RT ($\underline{P}<.001$). This indicated that speed - accuracy trade-off did not occur. That is, subjects with better tracking performance were also faster in their RT. Moreover, lack of significant negative correlations between RT and RT errors and between tracking and RT errors also indicated that there was no speed-accuracy trade-off. It should be noted that RT for green (toes) and red (heel) stimuli did not correlate significantly with RT errors. This was an important feature which indicated that reacting to the stimuli using toe movement was not superior to heel movement as anticipated. A T-test was performed between RT for green and red stimuli. No significant difference was found between their RT's. It further indicated that heel movement was not a confounding factor in this study.

Table 18 shows the overall correlations for Task E. It indicates that tracking performance (TE) correlated significantly with all DLT measures. Of interest here was

Table 17

Overall Correlations For Task D

	TD	RTD	ERRTD	RED	GREEN	SMLD
TD						
RTD	-.411**					
ERRTD	-.207	-.010				
RED	-.238	.866**	-.041			
GREEN	-.432**	.824**	.034	.468**		
SMLD	-.108	.201	.079	.094	.193	

Table 18

Overall Correlations for Task E

	TE	DLTE	DLEOM	DLEIN	DLESW	SMLE
TE						
DLTE	-.493**					
DLEOM	-.352*	.382**				
DLEIN	-.369*	.817**	.074			
DLESW	-.418**	.910**	.139	.640**		
SMLE	-.093	.351*	.258	.218	.315*	

* P<.01
** P<.001

Table 19

Overall Correlations for Task F

	TF	RTF	ERRTF	RTIN	RTOUT	ERIN	EROUT
TF							
RTF	-.545**						
ERRTF	-.372*	-.079					
RTIN	-.487**	.955**	-.065				
RTOUT	-.465**	.867**	-.103	.708**			
ERIN	-.315*	-.057	.829**	-.082	.025		
EROUT	-.347*	-.041	.882**	-.001	-.167	.473**	
MISRF	-.409**	.595**	.012	.577**	.340*	-.004	.098
MISRIN	-.361*	.623**	.002	.616**	.399**	.017	.056
MISROUT	-.421**	.487**	.024	.458**	.218	-.035	.145
DLTF	-.415**	.380**	.250	.342*	.397**	.223	.215
DLFOM	-.214	.064	.056	.072	.069	.120	-.013
DLFIN	-.308*	.313*	.278	.280	.317*	.174	.298
DLFSW	-.365*	.358*	.165	.319*	.382**	.185	.111
SMLF	.046	-.081	-.082	-.117	.058	-.052	-.110

	MISRF	MISRIN	MISROUT	DLTF	DLFOM	DLFIN	DLFSW	SMLF
MISRF								
MISRIN	.966**							
MISROUT	.933**	.810**						
DLTF	.213	.223	.175					
DLFOM	-.005	.017	-.0366	.264				
DLFIN	.190	.175	.1894	.812**	-.148			
DLFSW	.204	.226	.1502	.914**	.203	.579**		
SMLF	-.266	-.271	-.2300	.144	.070	.127	.109	

* P<.01
** P<.001

that speed-accuracy trade-off did not occur. This was indicated by the significant negative correlations between TE and DLTE ($P<.001$). That is, subjects who performed well in the tracking task also performed well (with fewer errors) in the DLT. It was also found that SML correlated (positively) significantly with DLTE ($P<.01$), particularly with switching error (DLESW).

The overall correlations of Task F (Table 19) were analysed to examine for possible overall speed-accuracy trade-off. It should be noted that the total sum of errors in RT (ERRTF) was the combination of RT errors during input (ERIN) and during output (EROUT). Therefore, correlations among these measures would be expected. Likewise, it was the same for RTF and MISRF, which were the sum total of RTIN and RTOUT, and MISRIN and MISROUT, respectively.

SML ratings and DLT omission errors were not significantly correlated with any of the measures employed in Task F. However, it was found that tracking correlated (negatively) significantly ($P<.01$) with all other measures, which indicated that speed-accuracy trade-off did not take place. In terms of the Response Task, RT was found to be correlated significantly (positively) with MISRF ($P<.001$), MISRIN ($P<.01$), MISROUT ($P<.001$) and overall DLT ($P<.001$), particularly with intrusion errors ($P<.01$) and with switching errors ($\underline{P}<.01$). These indicated that subjects with faster RT's were missing fewer RT responses and having fewer DLT errors, particularly intrusion and switching errors. Significant positive correlations were also found between RTIN and RTOUT ($\underline{P}<.001$), between RTIN and MISRIN ($\underline{P}<.001$), between RTIN and DLFSW ($\underline{P}<.01$), between RTOUT and DLFIN ($\underline{P}<.01$) ,and between RTOUT and switching errors ($\underline{P}<.001$). There were no significant correlations between RT errors (ERRTF) and input-output RT, between ERRTF and DLTF, and between RTOUT and EROUT. Finally, significant positive correlations were found between ERIN and EROUT ($\underline{P}<.001$) and between MISRIN and MISROUT ($\underline{P}<.001$), which indicated that subjects tended to perform poorly in both the input and output conditions at the same time. Despite the complexity of Task F and the demands required to perform the tasks, no speed-accuracy trade-off occurred, as indicated by the correlations in Table 19.

Tables 20 and 21 show the correlations of Task F for the accident group and accident-free groups respectively. A comparison was made between these correlations. Only differences in correlations between the two groups will be mentioned here, as these differences are of main interest in this study. These correlations are underlined in Tables 20 and 21.

Table 20

Accident Group Correlations for Task F

	TF	RTF	ERRTF	RTIN	RTOUT	ERIN	EROUT
TF							
RTF	-.601**						
ERRTF	-.340	-.165					
RTIN	-.529**	.976**	-.182				
RTOUT	-.597**	.897**	-.156	.811**			
ERIN	-.300	-.156	.883**	-.182	-.086		
EROUT	-.343	-.088	.910**	-.096	-.160	.616**	
MISRF	-.428*	.650**	-.029	.623**	.386	-.046	.086
MISRIN	-.426*	.672**	-.038	.637**	.439*	-.036	.059
MISROUT	-.415	.602**	-.016	.586**	.308	-.056	.116
DLTF	-.430*	.409	.195	.355	.530**	.221	.140
DLFOM	-.168	.006	.012	-.024	.093	.115	-.084
DLFIN	-.343	.327	.292	.296	.373	.196	.327
DLFSW	-.350	.406	.067	.356	.531**	.149	-.010
SMLF	.167	-.169	-.173	-.174	.002	-.084	-.259

	MISRF	MISRIN	MISROUT	DLTF	DLFOM	DLFIN	DLFSW	SMLF
MISRF								
MISRIN	.987**							
MISROUT	.981**	.940**						
DLTF	.164	.205	.107					
DLFOM	-.067	-.043	-.093	.240				
DLFIN	.166	.165	.162	.761**	-.281			
DLFSW	.164	.226	.084	.920**	.187	.547**		
SMLF	-.377	-.346	-.403	.194	.093	.166	.141	

* P<.01
** P<.001

Table 21

Accident-Free Group Correlations for Task F

	TF	RTF	ERRTF	RTIN	RTOUT	ERIN	EROUT
TF							
RTF	-.449*						
ERRTF	-.333	.092					
RTIN	-.414	.933**	.110				
RTOUT	-.366	.859**	.053	.624**			
ERIN	-.347	.147	.682**	.083	.246		
EROUT	-.154	-.001	.773**	.079	-.140	.063	0
MISRF	-.245	.524*	.094	.588**	.374	.148	0.0
MISRIN	-.144	.528**	.068	.596**	.382	.149	-.0363
MISROUT	-.462*	.156	.130	.161	.093	.045	.1390
DLTF	-.240	.333	.238	.284	.328	.206	.1470
DLFOM	-.257	.269	.062	.351	.104	.108	-.0094
DLFIN	-.114	.277	.146	.218	.305	.106	.1068
DLFSW	-.264	.269	.270	.221	.267	.241	.1603
SMLF	-.196	.022	.102	-.049	.099	0.0	.1397

	MISRF	MISRIN	MISROUT	DLTF	DLFOM	DLFIN	DLFSW	SMLF
MISRF								
MISRIN	.971**							
MISROUT	.433*	.209						
DLTF	.325	.221	.503*					
DLFOM	.323	.241	.420	.224				
DLFIN	.232	.160	.351	.876**	.089			
DLFSW	.286	.188	.469*	.885**	.105	.572**		
SMLF	-.113	-.187	.245	.156	.116	.129	.129	

* P<.01
** P<.001

Tracking was found to be significantly (negative) correlated with MISRF ($\underline{P}<.01$), MISRIN ($\underline{P}<.01$), and DLTF ($\underline{P}<.01$) for the accident group, but not for the accident-free group. However, a significant negative correlation (r = -.463) was found between tracking and MISROUT ($\underline{P}<.01$) for the accident-free group and not for the accident group. However, this difference is likely to be marginal, as the accident group had a correlation of -.415, which was close to the significant level. It does appear that for the accident group, subjects who performed poorly in the tracking task also missed more responses in the response task, particularly during auditory input. Likewise, subjects who performed poorly in the tracking task also had more errors in the DLT in the accident group. On the contrary, a significant (negative) correlation was found between tracking and MISROUT ($\underline{P}<.01$) with no significant correlations between tracking and MISRF, MISRIN, and DLT in the accident-free group. Thus, subjects in the accident-free group who performed poorly in the tracking task had more missed responses in their response task when they had to report the digits at the same time. This phenomenon seemed to occur for the accident group too, as a correlation of -.415 (close to significant level) was reported between tracking and MISROUT. It was also found that overall RT (RTF) was significantly (positive) correlated with MISROUT ($\underline{P}$<.001) for the accident group and not for the accident-free group. Thus, subjects in the accident group who were slower in their RT also failed to respond to stimuli when they had to report the digits at the same time. This however, was not the case for the accident-free group.

It was found that RTOUT was significantly (positive) correlated with DLTF and DLFSW ($\underline{P}$<.001) for the accident group and not for the accident-free group. This again indicated that subjects in the accident group who were slower in their RT while reporting digits were also making more errors in reporting digits in the overall DLT particularly when attention switching was required. This however, was not found in the accident-free group. ERIN was also found to be significantly (positive) correlated with EROUT for the accident group and not for the accident-free group. Thus, it appeared that subjects in the accident group were making RT errors both during DLT input and output, whereas subjects in the accident-free group were not doing so systematically. Likewise , it appeared to be the same for MISRIN and MISROUT, as indicated by the significant (positive) correlation between MISRIN and MISROUT ($\underline{P}$<.001) for the accident group and not for the accident-free group.

With respect to MISROUT, a significant (positive) correlation was found between MISROUT and DLTF ($\underline{P}<.01$) particularly for switching error (DLFSW) at $\underline{P}<.01$ for the accident-free group and not for the accident group. Thus, subjects' ability to switch attention in the accident-free group was tied to their failure to respond to stimuli when they have to report the digits at the same time.

Although T-tests did not find any significant differences between the two groups on field dependence, looking at the differences in correlations of the GEFT scores with taskload between the two groups provided further insights into group differences. Table 22 shows significant correlations between GEFT and taskloads for the accident group, whereas no significant correlations were found between GEFT and taskloads for the accident-free group. It was found that GEFT was significantly correlated (positive) with tracking performances under all tasks conditions (TA, TD, TE, TF) at $\underline{P}<.01$ (that is, from single task to the combination of three tasks) and negatively correlated with RTF ($\underline{P}<.001$), RTIN ($\underline{P}<.001$), and RTOUT ($\underline{P}<.01$). Thus, it indicated that field dependent subjects performed poorer in the tracking tasks and were also slower in their overall RT than were field independent subjects within the accident group. However, GEFT was not

Table 22

Correlations between GEFT and Taskload

For Tracking and RT Tasks

	GEFT		
	Overall (N=72)	Accident (n=36)	Accident-Free (n=36)
TA	.2373	.4712*	-.2233
TD	.3020*	.4934*	-.0280
TE	.3563*	.5136*	.0708
TF	.3192*	.4772*	-.0106
RTF	-.3153*	-.5853**	-.2726
RTIN	-.2922	-.6107**	.0997
RTOUT	-.2322	-.4910*	.0309

* P<.01

** P<.001

significantly correlated with such performances within the accident-free group. These correlations confirmed the ANOVA results obtained above.

Contribution of the Various Measures

Finally, a Dicriminant Function Analysis using SPSSX (1983) was carried out to determine a function that best discriminated between the accident and accident-free groups on the variables in which the two groups showed differences. All the tracking tasks (TA, TD, TE and TF) with all the dichotic listening tasks (DLTC, DLTE and DLTF), all switching errors (DLCSW, DLESW and DLFSW) and RT errors during output(EROUT) were included in the analysis. A preliminary analysis using the DIRECT method (forced entry) indicated 66.67% of the cases were classified correctly. Tables 23 and 24 show the standardized and unstandardized canonical discriminant function coefficients and the classification results (respectively) using the DIRECT method. A follow-up to the analysis employing the method of minimization of Wilks' Lamda (stepwise procedure) showed that the overall errors in DLTF and EROUT were the variables chosen in the discriminant function, with DLTF (standardized coefficient = .68) being a more useful predictor than EROUT (standardized coefficient = .65). Tables 25 and 26 show the standardized and unstandardized

Table 23

Standardized and Unstandardized Canonical Discriminant Function Coefficients (Direct Method)

	Standardized	Unstandardized
TA	0.12929	0.8656365E-01
TD	-0.12110	-0.3980104E-01
TE	-0.20234	-0.6755279E-01
TF	-0.02785	-0.6636063E-02
EROUT	0.60516	.4454496
DLTC	-0.55275	-.2163184
DLTE	0.02368	0.8389081E-02
DLTF	0.44163	.1249880
DLCSW	0.46419	.3237670
DLESW	0.22343	.1306771
DLFSW	0.02867	0.1410336E-01
(CONSTANT)		1.023536

Table 24

Classification Results of Group Membership (Direct Method)

ACTUAL GROUP		NO. OF CASES	PREDICTED GROUP MEMBERSHIP 1	2
Accident Group	(1)	36	21 58.3%	15 41.7%
Accident-free Group	(2)	36	9 25.0%	27 75.0%

Percent of "grouped" cases correctly classified: 66.67%

Table 25

Standardized and Unstandardized Canonical Discriminant Function Coefficients (Stepwise Selection)

	STANDARDIZED COEFFICIENT	UNSTANDARDIZED COEFFICIENT
EROUT	0.64691	.4761790
DLTF	0.67644	.1914422
(CONSTANT)		-1.7735440

Table 26

Classification Results of Group Membership (Stepwise Selection)

ACTUAL GROUP	NO. OF CASES	PREDICTED GROUP MEMBERSHIP (1)	(2)
ACCIDENT GROUP (1)	36	23 63.9%	13 36.1%
ACCIDENT-FREE GROUP (2)	36	9 25.0%	27 75.0%

Percent of "grouped" cases correctly classified: 69.44%

canonical discriminant function coefficients and the classification results (respectively) using the stepwise procedure. These two predictors with the discriminant function :

$$D = -1.77 + 1.91\ (DLTF) + .48\ (EROUT)$$

were able to classify 69.44% of the cases correctly (Table 26) which was slightly better than the classification results obtained from the DIRECT method.

DISCUSSION

A variety of factors could influence performance on tracking and RT tasks such as those used in this experiment. One of these is the effect of playing video games on the task performance measures employed in this study. We would generally expect people who played video games to have superior performance on eye-hand co-ordination tasks. However, no significant difference was found between those who reported that they played video games and those who did not play, on task performance in the present study. The tasks employed in the present study (except Task A) required more than eye-hand co-ordination. Task B required eye-foot co-ordination, while Task C involved listening and reporting digits and Tasks D, E and F required much more. For example, Task D required eye-hand-foot co-ordination, while Task E involved eye-hand co-ordination and at the same time, listening and reporting digits. Task F, which was the combination of all the tasks, involved eye-hand-foot co-ordination together with listening and reporting of digits at the same time.

Another concern was the forward (toes) and backward (heel) motions used in responding to the green and red stimuli (respectively), which may be a confounding factor. During the course of this study, some subjects reported that they felt that their responses were faster for the red stimuli than for the green stimuli. This was due to the pedal being slanted at an angle and the positioning of the pedal in relation to their seating positions. Employing a T-test, no significant difference was found between the RT for the green and red stimuli. Moreover, no significant correlation (Table 17) was found between the RT to red stimuli and RT errors, and between RT to green stimuli and RT errors, which indicated that the experiment was not being confounded in using a single pedal employing both toe and heel movements.

In section II of the study, where subjects performed the six tasks, it was observed that some subjects (particularly those from the accident group) were upset with their performance. Generally they seemed to have difficulty in the performance of the time-sharing tasks, especially in maintaining 95% accuracy in their tracking. They complained that the time-sharing tasks, especially Task F, were very difficult to perform, stating that these tasks were totally unrelated to driving a bus. Some subjects came close to giving up participating in the study. Moreover, these distressed subjects were given more practice trials than the others before they performed the actual tasks. Some subjects reported feeling more confident after having more practices. However, they still showed a high frequency of errors (tracking and DLT) in the performance of the actual tasks. Thus, we may conclude that these subjects failed to improve with practice, similar to the findings of an earlier study (Kahneman et al., 1973).

Spare Mental Capacity

Generally, it was found (as expected) that as taskload increased, SMC decreased for both groups. For example, tracking performance was poorest for Task F and the best for task Task A (Table 2 and Figure 5). Likewise, overall DLT errors, switching errors and intrusion errors (Tables 3, 4 and 5), were fewer under the single task condition (Task C) than under the dual task and 3-task conditions (Task E and F), with most errors under the 3-task condition. It was the same for RT, with fastest RT in Task B and the slowest RT in Task F. More errors in RT were also committed in Task F than in Tasks D and B, with Task B having the fewest RT errors. Furthermore, more responses were missed in Task F than Task D, and no responses were missed in Task B. Thus, the tasks employed in the present study were able to measure SMC.

Figure 5 shows an interesting feature with respect to SMC. Although SMC decreased as taskload increased from single task to 3-task conditions, it was also found that tracking performance in Task D (RT and tracking) differed significantly from that in Task E (DLT and tracking) under the dual task conditions. Tracking in Task E was found to be significantly superior to tracking in Task D by a small margin of 1%. Although the difference was small, it does indicate that tracking under eye-hand co-ordination together with listening-reporting combinations, was better than the task requiring eye-hand-foot co-ordination. The small margin of 1% under some critical circumstances may be a deciding factor between safety and disaster (for example, tracking aircraft movements by air traffic controllers or monitoring critical instruments in a nuclear power plant). This has implications for vehicle design, or more generally for designing of equipment that requires human-task compatibility. For example, when eye-hand co-ordination is employed, and an additional task is required, using an auditory task rather than using the foot should be the choice of the additional taskload where possible.

The difficulty of performing Task F is undeniable. Subjects were required to perform the task involving eye-hand-foot co-ordination together with listening and reporting of digits. A major concern here was speed-accuracy trade-off. However, as indicated in Table 19, trade-off did not take place. Under such a situation, it was interesting to note that RT was faster when subjects were reporting (RTOUT) than when they were listening (RTIN), as indicated in Figure 8. Moreover, more RT responses were missed during input (MISRIN) than during output (MISROUT) (Table 13). The likely explanation for these phenomena is that, when subjects were listening, they were processing the information at the same time. This is the phenomenon of limited capacity in processing of information, as there were many pieces of information to process at almost the same time. Therefore, RT during input was slower and/or they failed to respond. On the other hand, when subjects were reporting, the information had already been processed. Therefore, they were able to respond faster and missed fewer responses.

The percentages of missed responses were small (2.9% during input and 1.1% during output). But the increase in missed responses from output to input condition was 163%. Under this experimental condition, it showed that 'forced' listening slowed down RT and/or caused subjects to fail to respond. Thus, in the driving world where bus operators are sometimes 'forced' to listen to messages coming into the bus intercoms, as well as to passengers (when drivers are not separated from passengers), they need to be made aware that responses important to driving will slow down and/or they may fail to make appropriate reponses when required under such circumstances.

RT errors were found to be significantly more frequent during output (EROUT) than during input (ERIN) (Table 10 and Figure 9). However, there is no speed-accuracy trade-off, as indicated in Table 12. One possible explanation for this phenomenon is that while the subject was in the "output mode" after processing the information (i.e., reporting the digits), the presentation of a stimulus that required the subject to switch to the "input mode" (i.e., subject was required to detect the stimulus) was apparently incompatible. Subjects were not able to process the input stimuli accurately while at the same time reporting the digits. Therefore, it accounts for more errors being committed in responding to the light stimuli. Once again, if we are able to generalize to the driving world, there is a good possibility of making errors in the foot movement when drivers are talking at the same time. An error of such nature has great consequences, especially when the gas pedal is stepped on instead of the brake pedal in an emergency situation.

In terms of the attention switching model, the present data seem to support the parallel model of attention switching at a superficial level. This was indicated by the low level of missed responses during output (misrout). Subjects were able to detect the presence of the stimuli (red or green) while they were reporting the digits. This phenomenon required the subject to switch to the "input" mode from the "output" mode. However, at another level, the results of the present study actually supported the serial model of attention switching. This is because subjects were making more identification errors in their RT responses (EROUT), although they were able to detect these stimuli. They were not processing the information accurately. The present study confirmed Laberge's findings (1973) with respect to both detection and identification errors (which he failed to mention). That is, when the operation of switching of attention is taking place, the process of detecting stimuli is easier than identifying the stimuli. This was indicated by the proportion of errors in both the detection and identification errors in both studies.

Perceptual Style

The present study confirmed the finding of some of the recent studies (McKeena et al., 1986; Avolio et al., 1985;

Clement and Jonah, 1984) regarding the relationship between field dependence and traffic accidents. The measure of perceptual style (field dependence) using GEFT did not differentiate between drivers who had traffic accidents and those who were accident-free. Thus, the hypothesis that drivers who had accidents would be field dependent was not supported.

Although GEFT did not discriminate between 'good' and 'bad' drivers, Table 22 indicates differences in perceptual style, separately for the accident group and the accident-free group. For the former group, those who were field dependent were slower in their RT and also performed poorer in all the tracking tasks. Table 15 and Figure 10 show clearly that subjects in the accident group who were field dependent were inferior in the tracking task to field dependent subjects in the accident-free group and also inferior to other subjects who were not field dependent. There seemed to be two separate categories of field dependence or perceptual style. One category was those in the accident group who did poorly in tracking (eye-hand co-ordination), whether the tracking task was being performed alone (TA) or in combination with other tasks (TD, TE, and TF). The other category of field dependence performed equally well together with the less field dependent subjects within both the accident and accident-free groups. If we can in some way differentiate between these two categories of field dependence, we may then have greater confidence in associating perceptual style (field dependence) with driving behaviors.

Overall, field dependence as it relates to driving does not differentiate between the accident and accident-free groups of drivers in this study. However, field dependence may be related to some aspects of the driving task for the accident group , as indicated above. Field dependent subjects in the accident group were slower in their RT which confirmed the findings of some of the earlier studies (Barrett and Thornton, 1968; Barrett et al., 1969). This same group of subjects also had poorer ability in tracking (eye-hand co-ordination) tasks.

Subjective Mental Load

Although ratings of mental workload increased as taskload increased, no group differences were found for these ratings. Because Task A was always administered first and Task F administered last, there could be an order effect. Subjects could have rated task A as easy and Task F as most difficult, with Task B, C, D, and E in between due to fatigue or some other order effects. Thus, the SML rating may be confounded. However, it is interesting to note that the accident group of drivers did not 'perceive' or rate the tasks as more difficult than the accident-free group of drivers, even though their performance (tracking and DLT in the time-sharing task conditions) was poorer than the accident-free group. In fact, under the time-sharing conditions (Tasks D, E, and F), the accident group rated their workload as lighter (though not significantly) than did the accident-free group. It is therefore possible that the accident group have a less realistic perception of the tasks' difficulty than the accident-free group of drivers.

The MCH rating scale was indeed a valid and reliable scale in measuring workload, as indicated by the monotonic increase in the rating scale as workload increased. However, it was not a useful measure in predicting or differentiating between the safe and unsafe drivers in the present study.

Group Differences

Of all the various tasks employed, it was found that tracking (Figure 5), output RT errors (Table 10 and Figure 9), and DLT (particularly switching errors - Figures 6 and

7) were able to discriminate between the accident and accident-free groups.

It should be noted that, although the DLT employed in the present study was different from those in the earlier studies (Avolio et al., 1985; Gopher, 1982; Gopher and Kahneman, 1971; Kahneman et al., 1973; McKeena et al., 1986; and Mihal and Barrett, 1976), the present study was able to replicate the earlier findings of Avolio et al. (1985), Kahneman et al. (1973), and Mihal and Barrett (1976), with switching errors discriminating between the accident and accident-free groups of drivers.

As indicated in Figure 6, the accident group of drivers as a whole committed more errors in the overall DLT than did the accident-free group of drivers. Of the three categories of errors, switching error was found to discriminate between the two groups (Table 4). The accident group had significantly more switching errors than did the accident-free group (Figure 7). Therefore, ability to switch attention is a good predictor of safe driving behaviour. This same ability was also found to be a good predictor of successful pilot training (Gopher, 1982).

In a Discriminant Function Analysis (stepwise procedure), DLTF and EROUT were the only two significant predictors of group differences. As switching errors were subsumed under the overall DLT errors, it therefore did not come through as a significant predictor of group differences in the analysis. Looking at the structure matrix of pooled within-groups correlations between discriminating variables and the canonical discriminant function (Table 27), switching errors in Task F (DLFSW) has the third highest correlation ($\underline{r}$=.64) following immediately after DLTF ($\underline{r}$=.77) and EROUT ($\underline{r}$=.74).

Nonetheless, these two predictors with the discriminant function :

$$D = -1.77 + 1.91\ (\mathrm{DLTF}) + .48\ (\mathrm{EROUT})$$

were able to classify 69.44% of the cases correctly (Table 26). Within the accident group, 63.9% of the cases were classified correctly, which was slightly above chance level (50%). This does indicate that there were many other causes that contributed to traffic accidents which neither the measures of DLTF nor the EROUT were able to discriminate. However, for the accident-free group, 75% of the cases were classified correctly. It appeared that, although these two measures (DLTF and EROUT) had a

Table 27

Correlations between Discriminating Variables
and Canonical Discriminant Functions (Stepwise selection)

	FUNCTION
DLTF	0.76805
EROUT	0.74270
DLFSW	0.63963
DLTC	0.50263
DLTE	0.47255
DLCSW	0.46136
TF	-0.44560
TA	-0.44390
DLESW	0.42052
TD	-0.38222
TE	-0.35321

misclassification rate of 36.1% of the accident group, it classified fairly well for this sample of accident-free drivers. As for the rest of the predictors, tracking measures were not potentially good predictors, and RT did not show group differences based on earlier analysis.

Implications for Bus Driving

Under some circumstances for the bus operators in the One Man Operator (O-M-O) system, a great deal of information is required to be processed at almost the same time. For example, driving during rush (busy) hours with heavy volumes of traffic and large numbers of passengers wanting transfer tickets, as well as having to listen to messages coming in from the intercom, the workload (information processing) is tremendous. Thus, it is essential that during training and in the training procedure, bus operators be made aware that their RT will slow down and/or they may fail to make appropriate responses when they are attending to the information relevant to these tasks. At the same time, they must also be made aware that while they are responding (talking) to the passengers and/or the intercom, they may make errors in their responses in the driving task. Therefore, the Calgary Transit policy that requires drivers to bring their vehicles to a halt while interacting with passengers is appropriate.

Besides making the bus operators aware of the possibilities of such situations arising, an alternate solution is to employ a two-operator system during busy periods. Thus, the driver is separated from the passengers (perhaps in an enclosed area) and concentrates solely on the driving task and listening to the intercom, while the "conductor" handles the passengers. This will lessen workload and improves safety but will incur extra cost.

Another aspect that relates to safety and operator workload is the use of the intercom. In order not to further increase the workload of drivers, the intercom should be of good quality. With poor quality intercoms, drivers may be distracted or extra effort may be required to decipher the incoming messages. Thus, a good quality, properly maintained communication system is essential.

One aspect of training is building of confidence. Some subjects reported that they had more confidence in performing the tasks in the experiment after a number of practice trials. The use of time-sharing of three tasks, which gets at the information processing ability, can be a useful tool for training purposes. Using a simulator, a mock-up of the driving task, with bus operators training to perform two or three tasks at one time before they go on the road, can be very useful. In this manner, bus operators will have greater confidence in handling situations when workload (information processing) demand suddenly becomes greater.

For driver selection purposes, the ability to transfer training from one vehicle to another is important. Brown (1968) found that spare mental capacity is a sensitive measure of the transfer of learning. With SMC being measured only one week after training, he found that those with greater SMC were more capable of transferring their learning from one vehicle to another. As the tracking task and the DLT were able to differentiate between the accident and accident-free groups of drivers in their SMC in the present study, these measures (tracking and DLT) can be used for selecting potential drivers who would be more capable of transferring what they have learned in training to the real world situation. As indicated earlier, some subjects were not able to improve their performance in these two tasks with practice. This perhaps indicates the limited mental capacity available to these subjects. Thus, the time-sharing tasks may be employed for selection purposes. As poor ability to perform the tracking task and

DLT is an indication of lesser SMC, it may be concluded that the people doing poorly on these tasks will have poor transfer of learning. Furthermore, as the aging population is on the increase, a large proportion of the driving population will be in the older age groups. Thus, employing the time-sharing tasks (particularly time-sharing of three tasks) may be a useful device in selecting older drivers, who are already with the company, to remain in their jobs or to request their early retirement. Thus, the functional age of the older drivers can be determined.

Finally, the Discriminant Function (Stepwise Selection) was able to classify 69.44% of the overall cases correctly. Using the tasks employed in the present study for selection of drivers, and looking at potential drivers who did well in the DLT in Task F and had fewer errors in RT during output, we may be able to have 70% of potentially safe drivers. Thus, the use of time-sharing of three tasks can be a useful tool for selection of potential bus drivers.

Cautionary Notes

There are a number limitations that should be pointed out in the present study. Firstly, although stricter criteria were used in defining accidents, these accidents were generally minor in nature. For example, turning and hitting a mailbox, pulling into a bus stop and hitting the side of a parked vehicle, or failing to halt in time to avoid a rear-end collision, resulting in no serious injury were typical occurances. Secondly, the tasks employed in this experiment cannot replace the actual driving task. At best, these tasks were only able to simulate the information processing required in driving. Furthermore, in an experimental setting such as this study, subjects were "forced" to perform uncommon tasks in an unnatural environment (tracking and listening to and reporting digits in a room) which they will never encounter in the course of their driving careers.

A further limitation is that the tasks in the present study were performed over a period of slightly more than an hour. If subjects were to perform the tasks for eight hours, for example, or the study was conducted after subjects had worked for a full day, the results would most likely be very different from what were obtained. This may represent a more accurate state of the subject's SMC, as fatique and motivation will certainly play a major role, as in the case of driving for a full day on the job. Performing the tasks for eight hours also has the advantage of measuring the SMC with greater validity, in that automaticity in multi-task performance will set in. This will affect the availability of spare capacity (Schneider, Dumais, and Shiffrin, 1984; Shiffrin and Schneider, 1977). Thus, the results obtained in a study of greater duration may be quite different.

Finally, the subjects were all male bus drivers from the City of Calgary. (Sex differences have been found with respect to RT (Testin and Dewar, 1981) and perceptual style (Shinar et al., 1978), with male subjects having faster RT and being less field dependent than the females.) Moreover, this sample of drivers participated voluntarily and therefore may not be a representation of the actual bus driver population. Thus, interpreting and generalizing the findings of the present study should be done with caution.

Future Research

The use of the time-sharing of three tasks employed in the present study showed potential as a tool for selection purposes for skilled jobs requiring eye-hand-foot co-ordination while listening and talking. In performing such a task, the ability to process information in parallel is essential. The ability to track, to respond with the foot, and to switch attention at almost the same instant, amount to processing of information in parallel. The majority of the accident group of drivers were less capable of processing the information in parallel, as shown by their poor performance in the tracking task and DLT. It is more likely that they were processing the information in serial or sequentially. However, it is premature to draw such a conclusion. Further studies using a different sample of subjects (eg., truck drivers) are needed to verify the present data and the utility of time-sharing of three tasks in discriminating between safe and unsafe drivers. It is recommended that stricter criteria should be applied to define traffic accidents as it had been in the present study. In this manner, we can have greater confidence that traffic accidents will be due to drivers' errors instead of some extenuating circumstances. Moreover, more serious traffic accidents should be included, as they may be more indicative of safe driving behaviour of commercial drivers.

As sex differences have been found for field dependence (Shinar et al., 1978) and RT (Testin and Dewar, 1981), future studies should include both males and females.

Age is another factor which has been found to correlate with driving behaviour. Barrett, Mihal, Panek, Sterns and Alexander (1977) found that older drivers were slower in their RT (choice and complex RT) and made more errors in the DLT (total and switching errors) than did younger drivers. Their results indicated that, with increasing age, drivers become less efficient in processing information from the environment. Younger drivers were found to be more efficient on information processing tasks, with faster RT and fewer accidents. Therefore, future research should include both males and females as well as non-commercial drivers of a wider age range (especially drivers above age 55) for broader application and greater generalization.

Further research is recommended to explore the complexity of processing information with respect to auditory input versus verbal output while performing other manual tasks that require eye, hand and foot co-ordination. This has immense implications for other industries besides transportation.

Finally, time was the main consideration in using the GEFT to measure perceptual style in the present study. If time permits, the use of the Rod-and-Frame test would be the better choice. As suggested by some researchers (Avolio et al, 1985 Barrett et al., 1969; Mihal and Barret, 1976;), the Rod-and-Frame test may be a more effective measure of perceptual style than the GEFT. Moreover, it was reported that the Rod-and-Frame test has a stronger relationship with accident records.

CONCLUSIONS

On the whole, the results confirmed the hypothesis that the accident group of drivers had poorer ability to switch attention than did the accident-free group. With respect to SMC, not all taskloads employed in this study discriminated between the two groups in their SMC. Only tracking and DLT were able to discriminate the SMC of the two group of drivers, with the accident group having lesser SMC.

The hypotheses that the accident group of drivers were more field dependent and that they rated their mental workload as higher than the accident-free group were not confirmed in the present study.

REFERENCES

Avolio, B. J., Galen, K. G., and Panek, P. E. (1985). Individual differences in information-processing ability as a predictor of motor vehicle accidents. Human Factors, 27, 577-587.

Barrett, G. V., and Thornton, C. L. (1968). Relationship between perceptual style and driver reaction to an emergency situation. Journal of Applied Psychology, 52, 169-176.

Barrett, G. V., Thornton, C. L., and Cabe, P. A. (1969). Relation between embedded figures test performance and simulator behavior. Journal of Applied Psychology, 53, 253-254.

Barrett, G. V., Mihal, W. L., Panek, P. E., Sterns, H. L., and Alexander, R. A. (1977). Information - processing skills predictive of accident involvement for the younger and older commercial drivers. Industrial Gerontology, 4, 173-182.

Brown, I.D. (1962). Measuring the 'Spare Mental Capacity' of car drivers by a subsidiary auditory task. Ergonomics, 5, 247-250.

Brown, I. D. (1966). Subjective and objective comparisons of successful and unsuccessful trainee drivers. Ergonomics, 9, 49-56.

Brown, I. D. (1968). Some alternative methods of predicting performance among professional drivers in training. Ergonomics, 11, 13-21.

Brown, I. D., and Poulton, E. C. (1961). Measuring the 'Spare Mental Capacity' of car drivers by a subsidary auditory task. Ergonomics, 4, 35-40.

Bryden, M. P. (1969). Binaural competition and division of attention as determinants of the laterality effect in dichotic listening. Canadian Journal of Psychology, 23, 101-113.

Burg, A. (1967). Relationship between vision test scores and driving record : General findings (Report no. 67-24). Department of Engineering, University of California, Los Angeles.

Burg, A. (1968). Vision test scores and driving record : Additional findings (Report no. 68-27). Department of Engineering, University of California, Los Angeles.

Calgary Transit: Transportation Department (1985). Vehicle accident appeal board.

Casali, J. G., Wierwille, W. W. (1982). A sensitivity/intrusion comparison of mental workload estimation techniques using a flight task emphasizing perceptual piloting activities. Proceedings of the 1982 IEEE International Conference of Cybernetics and Society, (pp. 598-602). Seattle, Washington.

Casali, J. G., and Wierwille, W. W. (1983). Communications - imposed pilot workload: a comparison of sixteen estimation techniques. Proceedings, Second Symposium on Aviation Psychology, Ohio State University, Columbus, Ohio.

Clement, R., and Jonah, B. A. (1984). Field dependence, sensation seeking, and driving behavior. Personality and Individual Differences, 5, 87-93.

Cooper, G. E., and Harper, R. P. (1969). The use of pilot ratings in the evaluation of aircraft handling qualities (NASA Ames Technical Report TN-D-5153). Moffett Field, CA: NASA Ames Research Center.

Dewar, R. E. (1986). In-vehicle information and driver overload. Proceedings of the 3rd International Association for Vehicle Design Meeting on Vehicle Design and Components (pp. D43-D54). Geneva, Switzerland, Mar. 3-5.

Dixon, W. J., Brown, M. B., Engleman, L., Frane, J. W., Hill, M. A., Jennrich, R. I., and Toporek, J. D. (1985). BMDP statistical software. Berkely: University of California Press.

Eggemeier, T. F. (1981). Current issues in subjective assessment of workload. Proceedings of the 25th annual meeting of the Human Factors Society Santa Monica, CA.

Emaes, W. G., Lee, S. N., and Fell, J. C. (1970). State-of-the-art : Motor vehicle accident studies. Society of Automotive Engineers Report no. 700437, International Automotive Safety Conference, Detroit, Michigan.

Fell, J. C. (1976), A motor vehicle accident causal system : The human element. Human Factors, 18, 85-94.

Glass, G. V., and Hopkins, K. D. (1984). Statistical methods in education and psychology (2nd ed.). Englewood Cliffs, New Jersey: Prentice-Hall.

Goodenough, D. R. (1976). A review of individual differences in field dependence as a factor in auto safety. Human Factors, 18, 53-62.

Gopher, D. (1982). A selective attention test as a predictor of success in flight training. _Human Factors_, _24_, 173-183.

Gopher, D. and Kahneman, D. (1971). Individual differences in attention and prediction of flight criteria. _Perceptual and Motor Skills_, _33_, 1335-1342.

Green, P., and Pew, R.W. (1978). Evaluating pictographic symbols: An automotive application. _Human Factors_, _20_, 103-114.

Heard, E. A. (1974). _Symbol studies - 1972_. Society of Automotive Engineers, paper no. 740304. Detriot, Michigan.

Hills, B. L. (1980). Vision, visibility and perception in driving. _Perception_, _9_, 183-216.

Jackson, D. N., Messick, S., and Myers, C. T. (1964). Evaluation of group and individual forms of embedded figures measurement of field-independence. _Educational Psychological Measurement_, _24_, 177-192.

Jex, H. R., and Clement, W. F. (1979). Defining and measuring perceptual-motor workload in manual control tasks. In N. Moray (Ed.), _Mental workload: Its theory and measurement_. New York: Plenum Press.

MacFarland, R. A., Moore, R. C., and Warren, A. B. (1955). Human variables in motor vehicle accidents : A review of the literature. Harvard School of Public Health, Boston.

McGrath, J. J. (1975). Driver expectancy and performance in allocating automotive controls. Vehicle Research Institute Report, No. 14.1

McKenna, F. P., Duncan, J., and Brown, I. D. (1986). Cognitive abilities and safety on the road : a re-examination of individual differences in dichotic listening task and search for embedded figures. Ergonomics, 29, 649-663.

Mihal, W.L., and Barrett, G.V., (1976). Individual differences in perceptual information processing and their relationship to automobile accident involvement. Journal of Applied Psychology , 61, 229-233.

Moray, N. (1979). Models and measures of mental workload. In N. Moray (Ed.), Mental workload: Its theory and measurement. New York : Plenum Press.

Moray, N. (1982). Subjective mental workload. Human Factors, 24, 25-40.

7

Kahneman, B., Ben-Ishai, R., and Lotan, M. (1973). Relation of a test of attention to road accidents. Journal of Applied Psychology, 58, 113-115.

Keele, S. W. (1973). Attention and human performance. Pacific Palisades, California: Goodyear Publishing Co. Inc.

Kimura, D. (1967). Functional asymmetry of the brain in dichotic listening. Cortex, 3, 163-178.

Kirk, R. E. (1968). Experimental design : procedures for the behavioral sciences. Belmont, California: Brooks/Cole.

Laberge, D. (1973). Identification of two components of the time to switch attention: A test of a serial and a parallel models of attention. In S. Kornblum (Ed.), Performance and performance IV (pp.71-86). New York: Academic Press.

Little, A. D. (1966). The state-of-the-art of traffic safety. Summary report for the Automobile Manufacturers Association Inc. Cambridge, Mass.

Loo, R. (1978). Individual differences and the perception of traffic signs. Human Factors, 20, 65-74.

Moray, N. (1984). Mental workload. Proceedings of the 1984 International Conference on Occupational Ergonomics, Vol.2, 41-46. Toronto.

North, R. A., and Gopher, D. (1976). Measures of attention as predictors of flight performance. Human Factors, 18, 1-14.

Olson, P. L. (1974). Aspects of driving performance as a function of field dependence. Journal of Applied Psychology, 59, 192-196.

Oltman, P. K. (1968). A portable rod-and-frame apparatus. Perceptual and Motor Skills, 26, 503-506.

Oltman, P. K., Raskin, E., and Witkin, H. A. (1971). Group embedded figures test. Palo Alto: Consulting Psychologists Press.

Pedhazur, E. J. (1982). Multiple regression in behavioral research (2nd ed.). New York: Holt, Rinehart and Winston.

Perel, M. (1983). Vehicle familiarity and safety. NHTSA Technical Note, DOT HS-806-509, U.S. Dept. of Transportation. Washington, DC.

Procter, M. A., Ponton, C. W., and Jamieson, D. G. (1986). Programmes to produce high quality dichotic tapes for central auditory testing. Computers and Biomedical Research, 19, 508-519.

Rahimi, M., and Wierwille, W. W. (1982). Evaluation of the sensitivity and intrusion of workload estimation techniques in piloting tasks emphasizing mediating activity. Proceedings of the 1982 IEEE International Conference on Cybernatics and Society, (pp. 593-597). Seattle, Washington.

Reid, G. B., Shingledecker, C., and Eggemeier, T. (1981). Application of conjoint measurement of workload scale development. In R. Sugarman (Ed.), Proceedings of the 25th Annual Meeting of the Human Factor Society, Santa Monica, CA.

Rumar, K. (1981). Impacts on road design of the human factor and information systems. Proceedings of the 9th International Road Federation Meeting, (pp.31-49). Stockholm.

Rumar, K. (1986). In Vehicle information systems. _Proceedings of the 3rd International Association for Vehicle Design, meeting on Vehicle Design and Components_ (pp. D33-D42). Geneva, Switzerland, Mar. 3-5.

Schneider, W., Dumais, S. T., and Shiffrin, R. M. (1984). Automatic and control processing and attention. In R. Parasuraman, and R. Davis (Eds.), _Varieties of attention_ (pp. 1-28). New York: Academic Press.

Shiffrin, R. M., and Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attenting, and a general theory. _Psychological Review_, _84_, 127-190.

Shinar, D. (1978). _Psychology on the road_: _The human factor in traffic safety_. New York: Wiley.

Shinar, D., Zaidel, D. M., and Paarlberg, W. T. (1978). _Driver performance and individual differences in attention and information processing Volume II_ : _Field dependence and highway safety_. NHTSA DOT HS-8-01819, U.S. Department of Transportation. Washington, DC.

Skipper, J. H., Rieger, C. A., and Wierwille, W. W. (1986). Evaluation of decision-tree rating scales for mental workload estimation. Ergonomics, 29, 585-599.

Statistical Package for the Social Science, Inc. (1983) SPSSX User's Guide. New York: McGraw-Hill.

Sussman, E. D., Bishop, H., Madnick, B., and Walter, R. (1985). Driver inattention and highway safety. Transportation Research Record no. 1047.

Testin, F. J., and Dewar, R. E. (1981). Divided attention in a reaction time index of traffic sign perception. Ergonomics, 24, 111-124.

Treat, J.R., Tumbas, N.S., McDonald, S.T., Shinar, D., Hume, R.D., Mayer, R. E., Stansifer, R. L., and Castellan, N. J. (1977). Tri-level study of the cause of traffic accidents. Report No. DOT-HS-034-3-535-77 (TAC), Indiana University.

Treisman, A. M., and Geffen, G. (1968). Selective attention and cerebral dominance in perceiving and responding to speech messages. Quarterly Journal of Experimental Psychology, 20, 139-151.

Wickens, C. (1984). _Engineering psychology and human performance_. Merrill, Columbus.

Wierwille, W. W., and Casali, J. G. (1983). A validated rating scale for global mental workload measurement applications. _Proceedings of the Human Factors Society 27th Annual Meeting_ (pp.129-133). Norfolk, Virginia.

Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R., and Karp, S. A. (1962). _Psychological differentiation_ : _Studies of development_. New York : Wiley.

Witkin, H. A., Oltman, P. K., Raskin, E., and Karp, S. A. (1971). _Manual for embedded figures test, children's embedded figures test and group embedded figures test_. Palo Alto, California: Counselling Psychologists.

Appendix A

#

1986 DECEMBER 18

SUBJECT: STUDYING THE DRIVING TASK

We are asking for your help in a study on driving behavior.

Operating a motor vehicle safely often requires the performance of two or more tasks simultaneously while driving. The purpose is to find out whether the task of driving a motor vehicle is related to the ability to process information (eg., detecting and reacting to traffic lights). The results of the study will help us to gain better understanding of some of the factors which contribute to good driving. This study is funded by Transport Canada and conducted by the University of Calgary.

The tasks that you are required to do will be interesting and challenging. The main task is a tracking task on a computer screen, a bit like the 'Pac-Man' game. Two other tasks are also required, which you are to perform simultaneously with the tracking task. One of them is to see how quickly you respond to lights and the other requires you to pay close attention to numbers presented to you through a headphone.

We are looking for approximately 100 male volunteers of all ages who have at least three years of driving experience. In order to use both male and female Operators we would need to have equal numbers of males and females. As only a small percentage of Calgary Transit Operators are females, we would not be able to conduct a statistically valid study for the female Operator group. We therefore have asked for only male volunteers. This study will spread over a period of time from 1987 January to 1987 March.

Should you agree to participate, <u>all of the information you provide will be treated as strictly confidential.</u> Our records will identify you by a number <u>not</u> by your name, and your employer will not know your individual performance in these tasks. You also have the right to withdraw from the study at any time should you wish to do so.

If you agree to participate, we will be asking you to take part in an individual session lasting about one hour. The study will be conducted at your place of work and all appointments will be made to suit your convenience. If you are interested, a summary of the results will be sent to you upon completion of the study. Each participant will receive a small honorarium of $5.00.

Over the next three months, some of you will be approached (contacted by phone) by us and we hope you will help us by volunteering to take part in the study.

We thank you in anticipation of your help.

J. Gendron
Superintendent of Operations
Calgary Transit

Dr. R. E. Dewar
Professor of Psychology

Cheng S. Lim
Graduate Student

jgv

Appendix B

SUBJECTIVE WORKLOAD RATING SCALE

PLEASE INDICATE YOUR RATING OF THE TASK BY A X

		DIFFICULTY LEVEL	OPERATOR MENTAL DEMAND	RATING
EASY	1	VERY EASY, HIGHLY DESIRABLE	MENTAL EFFORT MINIMAL AND DESIRED TASK IS EASILY PERFORMED	
	2	EASY, DESIRABLE	MENTAL EFFORT IS LOW AND DESIRED TASK CAN BE PERFORMED	
	3	FAIRLY EASY	ACCEPTABLE MENTAL EFFORT IS REQUIRED TO PERFORM THE TASK	
DIFFICULT	4	MILD DIFFICULTY	MODERATELY HIGH MENTAL EFFORT IS REQUIRED TO PERFORM THE TASK	
	5	MODERATELY DIFFICULT	HIGH MENTAL EFFORT IS REQUIRED TO PERFORM THE TASK	
	6	FAIRLY DIFFICULT	MAXIMIUM MENTAL EFFORT IS REQUIRED TO PERFORM THE TASK	
VERY DIFFICULT	7	MAJOR DIFFICULTY	MAXIMIUM MENTAL EFFORT IS REQUIRED TO BRING ERRORS TO MODERATE LEVEL	
	8	VERY HIGH LEVEL OF DIFFICULTY	MAXIMIUM MENTAL EFFORT IS REQUIRED TO AVOID LARGE AND NUMEROUS ERRORS	
	9	EXTREME DIFFICULTY	INTENSE MENTAL EFFORT IS REQUIRED TO ACCOMPLISH TASK BUT WITH NUMEROUS ERRORS	
IMPOSSIBLE	10	IMPOSSIBLE	THE TASK CANNOT BE ACCOMPLISHED RELIABLY	

TASK A __; B __; C __; D __
E __; F __

SUBJECT NO: ______

Appendix C

Part I of instructions

Introduction

The purpose of this study is to find out whether the ability to perform a number of tasks at the same time is related to the driving task. This study requires you to perform a number of tasks which are broken down into two sections.

Section I is an exercise on finding hidden figures and in section II you are asked to perform 6 different tasks (Task A to Task F). Task A is a simple tracking task (where you follow a target on the screen); Task B is a response task which requires you to press a foot-pedal when a light comes on; Task C is a listening task; Task D consists of the tracking task and the response task; Task E is a combination of the tracking and listening tasks; and Task F requires you to perform the tracking, response and listening tasks all at the same time.

In addition, you are asked in section II to rate each task individually in terms of the level of difficulty.

Appendix C (continued)

Practice trials will be given for all the tasks that you are asked to perform. This is to help you to be familiar with the tasks.

During the course of this experiment, you have the right to withdraw from it at any time if you wish to do so. All information will be kept confidential.

Section I

The task in the first section requires you to trace out figures on a booklet which will be shown to you shortly.

(<u>GEFT booklet given to subject at this point</u>).

Appendix C (continued)

<u>Part II of Instructions</u>

Section II

In this section, you are asked to perform 6 different tasks (Task A to F) and also a rating of the level of difficulty of each task.

Task A

Task A is a simple tracking task which requires you to keep a dot inside a circle. The dot will move continuously and the circle is moved by a mouse which you will control. You are asked to keep the dot inside the circle 95% of the time (ie., 95% accuracy). In the course of your tracking, the screen will turn orange in colour momentarily whenever your tracking is below 95% accuracy. This is to make you aware that you should be more careful to track the dot accurately. It is very important that you maintain at least 95% accuracy in your tracking.

<u>NOTE:</u> Mistakes are common for the tasks in section II, as some of them are rather difficult and require a lot

Appendix C (continued)

of your attention. However, your maintenance of 95% accuracy in the TRACKING TASK is very important.

WORKLOAD RATING

In this rating, the scale ranges from 1 to 10 and falls into 4 categories. These categories are "EASY", "DIFFICULT", "VERY DIFFICULT", and "IMPOSSIBLE". Ratings of 1, 2 and 3 mean that the task is easy. Within this category, there are levels of how easy the task is. And a rating of 4,5 and 6 means that the task is difficult; 7, 8, and 9 mean that the task is very difficult; and a rating of 10 means that the task is impossible. You are asked to mark an "X" on the scale to indicate what you think the level of difficulty of the task is.

(<u>Subject practices and performs Task A</u>)

Appendix C (continued)

Part III of Instructions

Task B

Task B is a response task which requires you to press a foot-pedal as quickly as possible to lights appearing on the screen. When a GREEN light comes on, you press the foot-pedal with your TOES. If it is a RED light, press the pedal with your HEEL. Each time when a light comes on, it will appear close to the moving dot.

REMINDER: Green - press the pedal with your toes.
Red - press the pedal with your heel.

(*Subject practices and performs Task B*)

Appendix C (continued)

Part IV of Instructions

Task C

This is a listening task. Different numbers will be presented to each ear at the same time. Before the presentation of the numbers, a sound tone will indicate which ear you must pay close attention to. If a LOW TONE is presented, it means that you pay close attention to the LEFT ear. If a HIGH TONE is presented, it means that you pay close attention to numbers coming into the RIGHT ear. The tone will be presented to both ears. Your task is to report all the numbers coming into the ear indicated by the tone while ignoring the numbers from the other ear. Report only after 3 pairs of numbers have been presented when there is a period of silence over your ear-phones. You may report the numbers in any order.

REMINDER: Low tone - pay close attention to numbers in LEFT ear

High tone - pay close attention to numbers in RIGHT ear

(Subject practices and performs Task C)

Appendix C (continued)

Part V of Instructions

Task D

Task D consists of the tracking task in addition to the response task. The response task requires you to press a foot-pedal as quickly as possible to lights appearing on the screen close to the moving dot. When a GREEN light comes on, you press the pedal with your TOES. If it is a RED light, press the pedal with you HEEL. In this task, it is also very important that you maintain at least 95% accuracy in tracking the moving dot. Again, the screen will turn orange momentarily whenever your tracking is below 95% accuracy.

REMINDER: Green - press the pedal with your toes.
Red - press the pedal with your heel.

(Subject practices and performs Task D)

Appendix C (continued)

Part VI of Instructions

Task E

Task E consists of the tracking task in addition to the listening task. If a LOW TONE is presented, it means that you must pay close attention to numbers coming into you LEFT ear. If it is a HIGH TONE, it means that you must pay close attention to numbers coming into the RIGHT ear. Now your task is to report all numbers coming into the ear indicated by the tone while ignoring numbers from the other ear. Only report the numbers after 3 pairs of numbers have been presented when there is a period of silence over the ear-phones. You may report the numbers in any order. At the same time, you are asked to maintain 95% accuracy in your tracking task and again the screen will turn orange momentarily when your performance of the tracking task is below 95% accuracy.

REMINDER: Low tone - pay close attention to numbers in LEFT ear

High tone - pay close attention to numbers in RIGHT ear

(Subject practices and performs Task E)

Appendix C (continued)

<u>Part VII of Instructions</u>

Task F

This is a combination of Tasks A, B, and C. In this task, you are asked to perform the tracking task again maintaining at least 95% accuracy (very important to do that). At the same time, you are to press the pedal as quickly as you can when the lights (RED or GREEN) come on, and also to report numbers coming into your ear indicated by the tone, as in the listening task.

(any Questions?)

<u>NOTE:</u> Making mistakes is very common for such tasks, as a lot of your attention is required. However, it is important that you maintain 95% accuracy in the tracking.

REMINDER: GREEN light - use toes
RED light - use heel
Low tone - left ear
High tone - right ear

(<u>Subject practices and performs Task F</u>).