ROAD SAFETY RESEARCH REPORT

CR 190
2000

Evaluating a Regulated Hours Regime On-Road and an Alternative Compliance Regime under Simulated Conditions

Ann Williamson
School of Psychology & NSW Injury Risk Management Research Centre,
University of New South Wales

Anne-Marie Feyer
New Zealand Environmental and Occupational Health Research Centre,
University of Otago, Dunedin, New Zealand

Rena Friswell
School of Psychology, University of New South Wales

Samantha Finlay-Brown
NSW Injury Risk Management Research Centre,
University of New South Wales

Department of Transport and Regional Services
Australian Transport Safety Bureau

Evaluating a regulated hours regime on-road and an alternative compliance regime under simulated conditions.

DEMONSTRATION PROJECT FOR FATIGUE MANAGEMENT PROGRAMS IN THE ROAD TRANSPORT INDUSTRY

Ann Williamson[1,2], Anne-Marie Feyer[3], Rena Friswell[1] and Samantha Finlay-Brown[2]

[1] School of Psychology, University of New South Wales, Sydney, Australia
[2] NSW Injury Risk Management Research Centre, University of New South Wales, Sydney, Australia.
[3] New Zealand Environmental and Occupational Health Research Centre, University of Otago, Dunedin, New Zealand

AUSTRALIAN TRANSPORT SAFETY BUREAU

DOCUMENT RETRIEVAL INFORMATION

Report No.	Date	Pages	ISBN	ISSN
CR 190	July 2000	102	0 642 25579 2	0810-770X

Title and Subtitle

Evaluating a regulated hours regime on-road and an alternative compliance regime under simulated conditions.

Authors

Williamson A; Feyer A-M; Finlay-Brown S; Friswell R.

Performing Organisation

School of Psychology and NSW Injury Risk Management Research Centre, University of New South Wales, Sydney, Australia.

New Zealand Occupational and Environmental Health Research Centre, PO Box 913, Dunedin, New Zealand

Sponsored by / Available from

Australian Transport Safety Bureau
GPO Box 967, CANBERRA ACT 2608

Project Officer: Margaret Smythe

Abstract

This is the second report in a series on the development of better work-rest schedules for the long distance road transport industry. This report describes the results of the evaluation of two work-rest schedules. The first is an on-road evaluation of the effectiveness of the current working hours regime for managing driver fatigue and the second was an evaluation of an alternative work-rest schedule that was conducted in a simulation mode because the trips were significantly longer than allowed under the working hours regulations. The on-road regulatory evaluation showed that fatigue increased and performance capacity deteriorated over the period of a trip, and over the period between long 24 hour breaks in the work-rest schedule. The extent of deterioration was not significantly large to be judged as likely to compromise safety on the basis of comparisons with alcohol-equivalent standards of performance decrement. Rather these changes were reflecting early warnings of fatigue and its consequences. In contrast, the alternative compliance simulation showed significant deterioration in fatigue and performance. By the middle to end of the second consecutive 16 hour work period, drivers were showing significantly increased fatigue and performance deterioration that was likely to compromise safety, as judged by a comparison with alcohol-equivalent performance standards. These results suggest that this alternative work-rest schedule would need to be modified by increasing the amount of rest between long work periods before it could be used on the road.

Keywords

FATIGUE MEASURES, ALCOHOL, PERFORMANCE

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

We would like to thank the company and drivers involved in these studies for their cooperation and support.

We would also like to thank Dr Jim Lemon at Bitwrit Software for his ongoing work on the computerised tests used in the studies.

EXECUTIVE SUMMARY

This project investigated the effects of two different work-rest schedules on fatigue and performance. The study addressed an important problem for Australian drivers and the long distance road transport industry in particular, that of driver fatigue. There is considerable evidence that for professional long distance drivers, fatigue is a major problem that is not being adequately addressed by current working hours regulations. Alternative approaches to fatigue management have been proposed and are in operation in Queensland, however these approaches have been hampered by a lack of specific evidence on what represents an effective work-rest schedule for managing fatigue.

In this project, so far, two work-rest schedules were chosen to examine their effects on fatigue and performance capacity. The first study looked at the influence of working the current working hours regulatory regime on fatigue and performance. This was designed to provide a baseline for comparison with alternative approaches to work-rest scheduling. The second study looked at an alternative work-rest schedule that involved drivers doing significantly longer trips than permitted under the regulatory regime. The first study was an on-road study where fatigue and performance were measured at intervals while drivers were working. In contrast, the second study was conducted as a simulation as the work-rest schedule being evaluated in this study contravened the working hours regulations. The simulation involved drivers going through all aspects of the trip, but not on the road.

The results of the first study showed that drivers working under the regulated hours regime did not report particularly high levels of fatigue over a work day nor over a work week. Over the first single shift these findings were mirrored by performance effects. There was no clear evidence of performance deterioration in this group of drivers whose trip lasted on average between 13 and 14 hours. The results suggest also that for a single trip of this length the effects on performance are the same for day and night work, provided drivers had a long rest immediately before the trip.

Over the work week, performance results were also much like the fatigue results since performance remained at much the same level as it had been at the beginning of the trip. There was, however, some performance deterioration across the week in the ability to pick up infrequent visual signals. The results of the Mackworth Vigilance test showed a significantly greater number of missed signals at the end of the week compared to the beginning. Compared to the performance standard developed against alcohol effects in the earlier laboratory studies (CR 189), this level of decrement was not sufficiently large to be of concern for safety. Nevertheless, it was concluded that this result is likely to be an early warning of more significant performance deteriorations that would occur in work-schedules with longer work periods and less rest over the work week.

Where the regulated regime study showed no changes of immediate concern, the results of the simulation study showed significant deterioration in fatigue and performance over the much longer trip. Furthermore, performance was significantly poorer than found for the laboratory-based alcohol standard for 0.05%BAC, thus leading to the conclusion that this trip would not be safe if it were conducted on the road. The long trips that were done in the simulation showed high levels of fatigue and clear effects on performance that would be very likely to affect driving, particularly affecting drivers' ability to pick up signals. The simulation trip also showed the consequence of not balancing rest with work demands in very long trips since after the second long day of work and with only around six hours overnight sleep, drivers did not recover and performance effects occurred very soon after starting the third day.

This conclusion was supported by the findings of the on-road study of consistent relationships between work, rest, fatigue and performance over a work shift and over the work week. Over a work shift, both fatigue and performance were positively related to the number of hours worked indicating that controlling the length of the work shift is appropriate for managing fatigue over a single shift. The results of this study do not allow us to be specific about the maximum number of hours that can be done with safety. The simulation study suggested that up to 16 hours of work may be done by rested drivers without producing significant adverse effects on performance, but that such a long shift cannot be sustained for longer than one shift. This finding needs to be replicated on the road as overall performance levels in this simulation were considerably poorer than expected.

From the on-road study, over a work week control of fatigue and safe performance were related to the steps drivers took to manage fatigue even at quite low levels of fatigue. Drivers who took more breaks and obtained more sleep had lower fatigue and showed better performance. The simulation revealed that recovery from fatigue is dependent on the accumulated level of fatigue and that recovery may not occur when too much fatigue had accumulated and insufficient recovery time was allowed. So it seems from these results that in the longer term, effective fatigue management should emphasise rest-taking rather than necessarily only limiting the length of the work shift.

These studies also provide a model for evaluating the effects of long hours on fatigue and performance. Using a set of performance tests that have demonstrated sensitivity for fatigue, based on laboratory results, allows interpretation of the meaning of the results for safety. Because the tests have been standardised, it is possible to conclude from the results of the working hours regime evaluation that the small performance change across the work week did not compromise safety, but from the extended trip simulation it must be concluded that performance during that trip would be highly likely to be unsafe.

A number of issues have not been resolved from these studies. While the results suggest that trips of around 14 hours and possibly even up to 16 hours do not compromise safe performance for rested drivers, it was not possible for this report to look at changes in performance within the trip. This issue needs to be addressed further. Similarly, the results of the on-road working hours regime study showed only very moderate effects on performance, but most drivers in this study did considerably shorter total weekly hours than are allowed by the working hours regulations. From these results it is not possible to suggest what the effects of longer weekly hours would be.

Clearly then, the results of these studies have shed light on the relative effects of a number of important aspects of the working hours regime on fatigue and safe performance. Not surprisingly, a number of issues need further clarification, but the results of these studies also show that we now have an effective method for evaluating the effects of work-rest scheduling on fatigue and performance.

INTRODUCTION

Driving for long periods is recognised to be one of the major road safety hazards due to the increasing likelihood of fatigue. Professional drivers are one of the groups most at risk for fatigue because of the long distances they are required to drive in the job and, for many drivers, the additional hours spent loading and unloading or simply waiting for loads. These activities are likely to increase the risk of drivers experiencing fatigue.

The traditional approach to managing the problem of driver fatigue for professional drivers has been to regulate the hours of work and the distribution of rests over each 24 hour period and over each seven day period. Unfortunately, however, this approach is acknowledged by many in the long distance road transport industry to be at best of limited success. The main problems with the approach are that it does not take into account individual differences between drivers in fatigue experiences, does not accommodate circadian variations in the likelihood of fatigue, and does not take operational demands for freight movements sufficiently into account.

Previous research on Australian truck drivers showed that many did extremely long distances and long hours of driving. Drivers worked 62.6 hours per week on average, with approximately 30% working more than 72 hours (Williamson, Feyer, Coumarelos, & Jenkins, 1992; see also Arnold, Hartley, Penna, Hochstadt, Corry, & Feyer, 1996). Given demands for long hours at the wheel for many drivers, the experience of fatigue was very common. Approximately one third of drivers rated fatigue as at least a substantial personal problem, and 85% experienced fatigue at least occasionally while driving. It is not surprising then, that recently there have been moves to develop alternative approaches to the regulation of work hours

Probably the most well-advanced alternative approach to managing driver fatigue has been the Fatigue Management Programme (FMP) developed by the Queensland Department of Transport. This programme recognised the need to moderate the number of hours worked by professional drivers, but proposed a different view about universal compliance to regulated working hours. Under the FMP, companies are permitted to implement a driver work-rest programme that is designed to minimise driver fatigue but which may contravene aspects of the working hours regulations. The benefit of this scheme for drivers is that fatigue is managed better and for companies the benefit is that work-rest schedules can be designed that will also fit better with operational needs.

The FMP has however encountered problems in defining what are the components of an effective fatigue management programme. For example, we do not have good evidence about how long an average driver can drive in a 24 hour period without experiencing adverse effects on performance, nor how much adjustment we need to make for night time driving compared to driving in the day time. From this viewpoint, the FMP has been a little ahead of our knowledge about fatigue management in the real world. What is needed is the development of models of work-rest schedules which have been demonstrated to manage driver fatigue in most drivers. These models would then be available to be used by companies that are interested in improving their approach to fatigue management for their drivers and would also expand our knowledge of how work-rest schedules should be formulated to manage fatigue.

One of the problems for the development of model schedules has been that relatively few techniques are available that could reliably demonstrate fatigue effects on performance (Dinges, Whitehouse, Orne & Orne, 1988; Pilcher & Huffcutt, 1996). Even where it is acknowledged that methods are likely to be sensitive to fatigue effects, there have been no standards for judging the relative importance of performance decrements. Recent work by Williamson, Feyer, Friswell and Finlay-Brown (in press, CR 189; see also Williamson, Feyer, Mattick, Friswell and Finlay-Brown, 2000) produced a set of performance tests which had demonstrated sensitivity to fatigue due to long hours of wakefulness and which had been standardised against the effects of varying doses of alcohol on performance.

The aim of this project was to use these performance tests to begin to develop some work-rest schedule models by investigating the effectiveness of some of the schedules proposed under the FMP project and to compare fatigue experienced working these schedules with fatigue experienced under the current work-rest regulations. In addition, the results of the standardisation study can be used to improve the interpretation of changes in performance over any trip by comparing these changes with performance levels found at blood alcohol concentrations (BACs) above the legal limits. This comparison should help us understand the real-world meaning of any performance changes.

This project addresses these aims in three parts. The first part was to examine the experience of fatigue in drivers who are working under the regulated working hours regime in order to develop a baseline for comparison of the alternative schedules. The second part was to evaluate driver fatigue during a long trip involving an alternative work-rest schedule that allowed longer periods of work than currently permitted in the regulated regime. This evaluation was conducted as a simulation because the trip contravened the current regulated hours, but the simulation involved drivers who would be doing these long trips if they were permitted. The third aspect had a methods focus and was an attempt to validate the results of performance tests conducted using a new palmtop computer which can be used for on-road testing by comparing them with the results of performance tests conducted off the road using a laptop computer which was the main method used in all of our previous studies.

1. Baseline study of the regulated regime

The aim of this part of the project was to evaluate the effects of the current working hours regulations on driver fatigue. Currently in NSW, drivers are permitted to work for 14 hours in any 24 hour period followed by a minimum of six continuous hours break. During the work period, drivers must take a break of at least 30 minutes within every 5.5 hours of driving. Drivers must take at least one 24 hour break in every seven days, so allowing them to work for six 24 hour periods consecutively provided that they work no more than 72 hours in each week. This regime was evaluated by measuring driver fatigue at the beginning and end of the first work shift after a 24 hour break, then again at the end of the work period before they had their next 24 hour break.

METHOD

Design

Long-distance truck drivers were studied during the course of a regular week of work, following a break of at least 24 hours. An attempt was made to study drivers whose weekly roster approximated the driving hours regulations. That is, they worked 12 to 14 hours in each 24 hour period and took one 24 hour break at the end of the week. However, because very few of the drivers worked such a regular and prolonged routine, the actual rosters sampled were quite variable.

Drivers were classified into two groups – those whose first shift of the week contained a majority of daytime (06:00 to 18:00) hours and those whose first shift comprised either a majority of night hours (18:00 to 06:00) or included most of the midnight to dawn hours (00:00 to 06:00).

Three cognitive psychomotor tasks, previously shown to be affected by sleep deficit, were administered at the start and end of the working week, and where possible, at the end of the first work "day". Subjective fatigue ratings were also recorded at these times. This testing regime allowed the effects of a single day's work to be assessed at the start of the work week when drivers were likely to be well rested, and to examine the cumulative effect on fatigue and performance of a week of work. Drivers who did not return to their place of origin at the end of their first "day" were not tested at this time. In addition to these test sessions, drivers were also asked to self-administer two of the tests, and to make fatigue ratings, at the start and end of every break from work which contained sleep during the week. Drivers were also asked to test themselves at the start of any non-sleep breaks spanning more than one hour, however compliance with this aspect of the procedure was not high (40.5% of subjects reported no such breaks). The self administered tests were presented on palmtop computers, and were conducted to assess their suitability for future field studies.

Drivers also provided background information about their health and lifestyle, and their recent work history.

Subjects

40 professional long distance drivers participated in the study. Three of the subjects had participated in an earlier study comparing fatigue and blood alcohol levels, and 2 others had been involved in the long shift simulation reported below. Eighteen drivers were regarded as starting their week with a day shift and 16 began with a night shift. The remaining subjects could not be unambiguously classified.

Seven of the subjects (2 day, 1 night, and 4 ungrouped subjects) terminated their participation at various times during the study week. Four of these cited work and time pressure as the factor responsible for their decision. In 3 of the withdrawals, technical problems with the on-road testing equipment contributed, and one subject's work roster changed so that it was no longer compatible with the requirements of the study.

Table 1 summarises the characteristics of the sample. Subjects were all men and had an average of 12 years professional driving experience. The minimum was 2.5 years.

The majority of the subjects (73%) were aged between 30 and 49 years and most (78.4%) were living in an ongoing relationship. Approximately two thirds of the group (73.0%) had ceased formal education before Year 11, while 10.8% had completed post-secondary qualifications. Most (83.8%) of the group had little or no prior computing experience. The majority of subjects drank alcohol (97.3%) but less than half (43.2%) currently smoked.

Materials and measures

i. Performance testing

Three of the tests from the Performance and Information Processing Systems (PIPS) Test Battery were administered on Compaq laptop computers. Subjects responded on peripheral Genovation keypads (Micropad 622) and standard serial mice. Each subject also used two custom-built "masks" that fitted over the keypads and concealed non-essential keys for the various performance tests. The masks were easily changed according to the instructions given at the start of each test.

TABLE 1: Demographic details of on-road study subjects

Demographic Factor		Percentage of subjects
AGE:		
• 20 – 29 years		16.2
• 30 – 39 years		51.4
• 40 - 49 years		21.6
• 50 - 59 years		10.8
EDUCATION LEVEL:		
• High school years 7 - 10		73.0
• High school years 11 - 12		16.2
• TAFE		8.1
• College/university		2.7
PC EXPERIENCE:		
• No previous experience		45.9
• A little experience		37.8
• Frequent PC user		16.2
SOCIAL DRUG HABITS:		
• Cigarettes	- Non-smoker	29.7
	- Ex-smoker	24.3
	- Current smoker	43.2
• Alcohol	- Non-drinker	2.7
	- Current drinker	97.3
RELATIONSHIP STATUS:		
• Married/Defacto		78.4
• Divorced/Widowed/Separated		13.5
• Single		8.1
DRIVING EXPERIENCE:		
• Years of driving experience	- ≤ 5 years	21.6
	- 5 - 10 years	24.3
	- 10 - 20 years	40.5
	- Over 20 years	13.5
** Least number of years driving experience - 2.5 years* *Greatest number of years driving experience - 36 years*		

The tests were selected on the basis that they showed a clear relationship between increasing time awake and performance deterioration, and comprised the Simple Reaction Time (RT) task, the Dual (RT and Tracking) task, and the Mackworth Clock Vigilance task. The tests took approximately 20 minutes to complete. Particulars of each test are presented below.

> *Simple Reaction Time (RT):* Subjects pressed a key as quickly as possible whenever a moving circle on the computer screen changed colour. Responses were made with the non-preferred hand to make them comparable with the later Dual Task. During the 2 minute task, 40 colour changes occurred with a minimum interstimulus interval (ISI) of 2sec. The maximum response time permitted was 1sec. Response speed and the number of correct responses were measured.
>
> *Dual Task:* In this task, subjects used the computer mouse in their preferred hand to pilot a small dot around the screen trying to keep it inside a moving circle. At the same time, they responded to colour changes in the circle with their non-preferred hand. As the subjects became more accurate at tracking the circle, the regularity or predictability of the circle's movement decreased, making the task more difficult. If a subject became less accurate, the movement of the circle became more regular or predictable, and therefore easier to track. The regularity of the circle's movement was updated every 5sec across the 3 minute task, and was used as a measure of tracking accuracy. In addition forty colour changes occurred with minimum ISIs of 2sec and maximum response times of 1sec. Speed and number of correct reaction time responses were measured as well as tracking accuracy.
>
> *Mackworth Clock Vigilance task:* Twenty four points lying on a circle flashed consecutively in a continuous circuit for 15 minutes. Subjects pressed a button on the keypad as quickly as possible whenever one of the points failed to flash. Omitted flashes occurred 15 times with a minimum ISI of 45sec. The maximum allowed response time following each omitted flash was 10sec. Flashes themselves lasted 500msecs. Reaction time, the number of missed responses and the number of false alarm responses were measured.

The tests were always presented in the same order – Simple Reaction Time, Dual Task, and finally the Mackworth Clock Vigilance task. At the beginning of the laptop tests, subjects were presented with three Visual Analogue Scales, anchored by the terms Fresh and Tired, Clear-headed and Muzzy-headed, and Very Alert and Very Drowsy, on which to rate their current fatigue.

The self-administered tests were run on small palmtop computers (Hewlett Packard 200LX) which travelled with the drivers. Each computer was packed in a standard lunchbox for ease of transport, together with a Genovation keypad (Serial Micropad 623) which plugged into the computer and served to increase the key size. To reduce the testing time required of subjects in the field, only the two tests demonstrating the clearest relationship between sleep deprivation and performance decrement were presented on the palmtops – the Simple Reaction Time task and a reduced (5 minute) version of the Mackworth Clock Vigilance task. Together these tests lasted 7 minutes. The palmtop tests were identical to the laptop versions, except that, in the Reaction Time task, the circle changed from a full line to a broken line, rather than changing colours and, in the Mackworth Clock Vigilance Task, only one third the number of flashes and omitted flashes occurred. The Simple Reaction Time task was always completed before the Mackworth Clock Vigilance task.

Visual Analogue Scale ratings were made in conjunction with the palmtop tests but were recorded manually in a "Work Diary", rather than electronically.

ii. Documentation

All subjects completed an informed consent form (Appendix 1) and a brief questionnaire addressing demographic characteristics, general health and lifestyle factors, workload in the previous week, recent sleep, and food and drug intake (Appendix 2).

Drivers were also provided with a "work diary" in which to record details of their work and rest schedule during the study, and to rate their fatigue before and after breaks (Appendix 3), and a set of written instructions for using the palmtops (Appendix 4).

Procedure

Each subject arrived at the testing room at the truck depot approximately 1 hour before they were due to depart on their first trip for the week. The nature of the study was explained and they were asked to give formal consent before participating. They then completed the background questionnaire and a practice session of the palmtop tests. Drivers who had obtained substantial palmtop test experience in earlier studies were not asked to do the practice session. All drivers then completed a baseline test session on the laptop computer, followed by a baseline palmtop session. This latter test session also allowed drivers to clarify the procedure for using the palmtops, so that they could confidently self-administer the tests when alone in the field. (Round the clock assistance for drivers on the road was obtainable by phone.) Finally, the use of the "work diary" was explained, and drivers were asked to complete some initial questions regarding any sleep, meals, and beverages they may have had immediately prior to the study.

Drivers returning to the depot at the end of their first work "day" and at the end of their final shift for the week reported to the test room just prior to leaving work, completed a palmtop test session and the appropriate diary entry, and a laptop testing session.

Statistical Analysis

In all analyses the three subjective rating scales were averaged to create a single rated fatigue measure.

Comparisons were made for fatigue and performance measures between the beginning and end of the first trip of the week and between the beginning of the first trip of the week and the end of last day of the working week as these were the common milestones across most drivers. For the first comparison, a two factor mixed model analysis of variance was used in which the first factor, test occasion (beginning and end of the first day), was repeated measures and the second factor was group (drivers whose first trip was primarily during the day or primarily at night). T-tests were used for the second comparison. Subject group was not included in these comparisons because drivers typically worked days and nights equally over the week.

The influence of the work and rest schedules before the trip on baseline fatigue and performance were analysed using linear regressions in which a range of work-rest factors were entered as predictors. The predictors used were work hours in the past seven days, number of night hours worked, number of hours between the end of the last work shift and the start of the study trip, total hours slept in this period, the hours since the end of the last sleep, and the rated quality of that sleep. A similar regression analysis was also used to estimate the influence of the work-rest experiences during the study week and the first day of the study on fatigue and performance measures.

Comparisons were also made between performance on each test during the trip and performance levels standardised against alcohol. This analysis should provide a yardstick for estimating the level of safety risk at intervals across the first trip and across the work week.

Missing data occurred throughout the study as a function of driver forgetfulness, driver tiredness, technical problems with the palmtops, work scheduling and unexpected changes to it, and the time pressures of other life commitments. This means that the analyses reported include varying numbers of people, depending on which measures were being examined and at which points in the work week they were recorded.

RESULTS

1. Comparison of fatigue and performance at the beginning of trip

Recent work history

In the seven days leading up to the beginning of this study drivers worked approximately 59 hours on average with about half of this time involving night work (see Table 2). Nearly three quarters of the drivers did more than 60 hours work in the past week, up to 90 hours in total. Night work was a common feature among almost all drivers with 78.4% doing more than 20 hours night work in the last week, up to 63 hours in total. Two drivers in the sample were on leave for the past seven days.

Over all drivers, the mean duration of last shift was between 15 and 16 hours, although for some drivers the last shift was up to 19 hours long. Consistent with the design of the study, however, all drivers had at least 25 hours break between the last shift and the start of the study and for most drivers the last shift occurred more than 48 hours before the study week started.

Table 3 describes the sleep patterns of drivers prior to the study. On average, drivers had around seven hours sleep before the study started which was rated as good quality by most drivers, and most drivers reported waking refreshed. A small proportion of drivers had considerably less sleep and, as a result, their ratings of refreshedness from sleep were also considerably lower.

Overall, therefore, the recent work and rest history of drivers in this study conformed with the study design since all drivers had at least 24 hours between the last work shift and the beginning of the study, ensuring that drivers were rested before they began the trip.

Comparing the day and night groups, there were no significant differences in the number of hours worked in the previous week ($t_{(32)}$=0.44,ns), the amount of night work in the past week ($t_{(32)}$=1.41,ns), or the number of trips done last week ($t_{(32)}$=1.15,ns).

TABLE 2: Distribution of work hours in the last 7 days

Hours worked	N	% of subjects
Total work hours:		
• On leave	2	5.4
• ≤ 60 hours	8	21.6
• 60-79 hours	23	62.2
• 80 + hours	4	10.8
Mean *SD* [1]	62.3 *11.5*	
Range [1]	40.0 – 90.0	
Night work hours:		
• On leave	2	5.4
• 0 hours	1	2.7
• ≤ 20 hours	5	13.5
• 20 - 40 hours	21	56.8
• 41 + hours	8	21.6
Mean *SD* [1]	31.8 *17.1*	
Range [1]	0.0 – 63.0	
Length of last shift:		
• ≤ 10 hours	4	10.8
• 10 – 15 hours	18	48.7
• 15 – 25 hours	10	27.0
• 25 - 48 hours	5	13.5
Mean *SD* [1]	16.3 *7.9*	
Range [1]	6.5 – 48.0	
Hours since end of last shift:		
• 24 - 48 hours	13	35.1
• 49 - 72 hours	18	48.7
• > 72 hours	6	16.2
Mean *SD* [1]	52.9 *15.2*	
Range [1]	25.0 – 78.5	
[1] *excludes 2 subjects who were on leave for past 7 days*		

Comparison of the group of drivers who began the study with a day trip with the group who started with a night shift showed that the day group tended to have a longer last trip than the night group ($t_{(19)}$=1.95,p<0.07) but there was no difference in the hours spent actually working in the last trip ($t_{(19)}$=1.32,ns.) nor the time since their last work shift (Tables 4 & 5). The day and night groups were similar in the total amount of sleep they had since their last work shift ($t_{(32)}$ =0.001,ns), in the quality of their last substantial sleep ($t_{(32)}$=0.001,ns) and in terms of how refreshed they felt after the substantial sleep ($t_{(32)}$=0.27,ns). There were differences however in the time between the last substantial sleep and the start of the study trip with day starters beginning work much sooner after sleep than night workers ($t_{(32)}$=4.19,p<0.001). The results showed, however, that half of the night starters made up for the long sleepless period before they started work by taking naps, usually a couple of hours before the trip began. Only night starters took naps and the results indicate that the quality of nap sleep was rated as high as drivers rated substantial sleep ($t_{(10)}$=1.42,ns) and their ratings of feeling refreshed were as high as for substantial sleep ($t_{(35)}$= 0.247,ns). When naps are figured into this analysis, there were no differences between the two groups in the time between sleep and starting work (t(21)=0.59,ns), although the length of the last sleep was, not surprisingly, statistically significantly lower for night workers ($t_{(26)}$=2.95,p<0.007).

TABLE 3: Details of rest since end of last shift

	Mean	Range	Median	SD
Hrs sleep since end of last shift - including leave	21:53	9:00 - 80:00	19:00	14:11
- without leave	18:48	9:00 - 34:00	19:00	5:30
Hrs length of last substantial sleep	7:34 hrs	3:0 – 12:00 hrs *8% - ≤3.5 hrs* *32% - ≤ 7.0 hrs* *68% > 7 hrs*	8:00	1:55
Rated quality of last substantial sleep(/100)	83.5	47.0 - 100.0	87.0	14.4
Rated refreshedness from last substantial sleep (/100)	76.0	8.0 - 100.0	77.0	22.2
Baseline fatigue ratings				
- laptop	22.5	0.0 – 68.3	20.0	18.7
- diary	26.1	0.0 – 63.8	18.3	20.7

TABLE 4: Distribution of work hours in last 7 days across day and night work groups on shift 1.

Hours worked	Mean	SD
Total work hours:		
• Majority Day Work	60.22	18.69
• Majority Night Work	57.31	19.67
Total no. hours at night:		
• Majority Day Work	26.11	13.79
• Majority Night Work	34.47	20.51
Total duration of last trip:		
• Majority Day Work	18.44 hrs	10.92
• Majority Night Work	13.25 hrs	2.74
Hours spent working last trip:		
• Majority Day Work	14.22 hrs	6.76
• Majority Night Work	12.03 hrs	1.80
No. trips last week:		
• Majority Day Work	3.83	1.38
• Majority Night Work	4.38	1.36
Decimal hrs between end of last shift & start of current one:		
• Majority Day Work	61.44	50.69
• Majority Night Work	66.74	47.87

Relationship between recent work-rest experience, fatigue and performance

The long rest break between the last trip and the beginning of this study was beneficial for drivers as their pre-study ratings of fatigue were very low for both day and night start groups (see Table 3). The relationship between fatigue at the beginning of the trip and prior work-rest schedules was analysed using regression analysis. The number of hours worked in the past seven days, the number of night hours, time since the last work shift, the number of hours sleep since the end of the last shift (including naps), quality of the last sleep and the time

`**TABLE 5:** Distribution of sleep across day and night work groups on shift 1.

Sleep	Mean	SD
Hours length of last substantial sleep: • Majority Day Work • Majority Night Work	7.76 7.27	2.17 1.98
Hrs sleep since end of last shift: • Majority Day Work • Majority Night Work	22.06 22.06	15.42 14.25
Decimal hrs between end of last substantial sleep & start of current shift: • Majority Day Work • Majority Night Work	4.91 hrs 11.33 hrs	3.06 5.64
Quality rating of last sleep: • Majority Day Work • Majority Night Work	82.44 82.44	14.64 14.89
Refreshed rating at end of last sleep: • Majority Day Work • Majority Night Work	74.47 76.56	25.10 19.02
Decimal hrs between end of nap & start of current shift: • Majority Day Work • Majority Night Work	n/a 2.04 hrs	n/a 1.53
Quality rating of last nap: • Majority Day Work • Majority Night Work	n/a 74.44	n/a 22.74
Refreshed rating after last nap: • Majority Day Work • Majority Night Work	n/a 72.88	n/a 24.45
Hours length of last substantial sleep or nap: • Majority Day Work • Majority Night Work	7.76 4.97	2.17 3.18

since that sleep were all entered as predictors of fatigue ratings. The regression model using these predictors was not significant using either diary ratings of fatigue or laptop ratings of fatigue (for diary ratings, $F_{(6,30)}$=1.07,ns.; r^2=0.011; for laptop ratings, $F_{(6,29)}$=1.23,ns.; r^2=0.037). It seems that any variation in the already low levels of fatigue at baseline was not related to variations in the drivers' recent work and rest experiences.

Regression analysis using the same predictor variables for baseline performance on the laptop version of the Simple Reaction Time test also showed no overall relationships for Simple Reaction Time ($F_{(6,30)}$=0.7,ns, r^2=0.053), the variation in reaction speed ($F_{(6,30)}$=0.7,ns, r^2=0.050) or the number of missed responses ($F_{(6,30)}$=0.8,ns, r^2=0.031). Inspection of the correlation matrices, however, showed significant correlations between the number of night hours done in the previous week and reaction speed ($r_{(37)}$=0.3,$p<0.04$) and variation in reaction speed at baseline ($r_{(37)}$=0.29,$p<0.04$). Reduced speed and increased speed variability were associated with more night work. Regression analysis of the palmtop version of this test showed similar nonsignificant results for each of the measures (for reaction speed, $F_{(6,30)}$=0.16,ns, r^2=0.16; variability of reaction speed, $F_{(6,30)}$=0.78,ns, r^2=0.038; missed responses, $F_{(6,30)}$=0.66,ns, r^2=0.06). Inspection of the correlations showed that the time since the last sleep or nap was significantly associated with both the variability in reaction time and the number of missed responses in the palmtop test (for variability of reaction speed; $r_{(37)}$=0.29,$p<0.04$; for missed responses, $r_{(37)}$=0.28,$p<0.04$), such that people with longer periods since the last sleep tended to miss more responses and have more variable reaction times. At the beginning of the study, therefore, simple reaction performance was not strongly predicted by work-rest experiences in the past week. The results do suggest, however, that the number of hours worked at night in the last week and the time since the last sleep influence task performance.

Similar regression analysis was performed using the same predictor variables for all measures on the laptop and palmtop versions of the Mackworth Vigilance test. For the laptop version, the reaction time and variability in reaction time measures were predicted significantly by the work-rest variables entered (for reaction time, $F_{(6,30)}$=3.78,$p<0.006$, r^2=0.32; for variability in reaction time, $F_{(6,30)}$=4.93,$p<0.001$, r^2=0.4) and as shown in Table 6 this was mainly due to the influence of the time since the last sleep or nap for both measures. Neither of the other two measures of the laptop version of the Mackworth test were significantly predicted by the prior work-rest variables (for false alarms, $F_{(6,30)}$=1.27,ns., r^2=0.043; for missed responses, $F_{(6,30)}$=1.26,ns., r^2=0.04). Examination of the correlation matrices showed that the

TABLE 6: Regression statistics for laptop Mackworth Clock vigilance measures at baseline

	REGRESSION STATISTICS			
	Beta	Standard error	Standardised Beta	Test result
PREDICTORS OF RT				
• Constant	844.14	154.23	-	t=5.47, p<0.001
• Hours worked in last 7 days	-1.01	1.52	-0.16	t=-0.66, p=0.51
• Night hours worked in last 7 days	0.95	1.03	0.14	t=0.93, p=0.36
• Hours off work before study	0.80	1.34	0.32	t=0.60, p=0.56
• Hours sleep since last shift	-0.93	3.98	-0.11	t=-0.23, p=0.82
• Hours between last sleep and start of study	11.50	3.80	0.48	t=3.03, p=0.005
• Rated quality of last sleep	-0.26	1.06	-0.04	t=-0.24, p=0.81
PREDICTORS OF VARIABILITY IN RT				
• Constant	137.46	214.87	-	t=0.64, p=0.53
• Hours worked in last 7 days	-2.67	2.12	-0.28	t=-1.26, p=0.22
• Night hours worked in last 7 days	1.02	1.43	0.11	t=0.71, p=0.48
• Hours off work before study	1.37	1.86	0.37	t=0.73, p=0.47
• Hours sleep since last shift	-1.44	5.54	-0.12	t=-0.26, p=0.80
• Hours between last sleep and start of study	14.06	5.29	0.40	t=2.66, p=0.01
• Rated quality of last sleep	-0.07	1.47	-0.01	t=-0.05, p=0.96

number of false alarms was significantly correlated with the number of night hours worked last week ($r_{(37)}$=0.36,p<0.013) and there was a trend for missed responses to be correlated with the number of night hours last week and the time since the last sleep or nap ($r_{(37)}$=0.24,p<0.08; $r_{(37)}$=0.23,p<0.09, respectively).

For the palmtop version of the Mackworth test, none of the measures showed significant prediction by the work-rest experience variables (for reaction time, $F_{(6,30)}$=0.76,ns., r^2=0.04;

for variability in reaction time, $F_{(6,30)}$=0.65,ns., r^2=0.06; for false alarms, $F_{(6,30)}$=1.96, p<0.1, r^2=0.14; for missed responses, $F_{(6,30)}$=0.72,ns., r^2=0.05) although again, inspection of the correlation matrices showed that reaction time was significantly associated with time since the last sleep or nap ($r_{(37)}$=0.27,p<0.05) and all other measures were significantly correlated with the number of night hours worked in the past week (for variability in reaction time, $r_{(37)}$=0.23,p<0.08; for false alarms, $r_{(37)}$=0.41,p<0.006; for missed responses, $r_{(37)}$=0.3,p<0.04).

Overall, the results for both versions of the Mackworth test indicate that the time since the drivers last slept is an important determinant of their ability to respond quickly to infrequent events. Similarly, the number of night hours worked is likely to be important for the accuracy with which drivers can respond to infrequent signals.

Regression analysis for the Dual Task showed no significant prior work-rest predictors for any of the measures (for reaction time, $F_{(6,30)}$=1.6,ns., r^2=0.09; for variability in reaction time, $F_{(6,30)}$=2.08,p<0.085., r^2=0.15; for missed responses, $F_{(6,30)}$=0.52,ns., r^2=0.09; wander, $F_{(6,30)}$=0.8,ns., r^2=0.035), but again, the correlation matrix revealed that the number of night hours worked in the past week was significantly related to reaction speed and variability and the degree of difficulty achieved (wander) in the tracking task component (for reaction time, $r_{(37)}$=0.36,p<0.01; for variability in reaction time, $r_{(37)}$=0.41,p<0.006; wander, $r_{(37)}$=0.41,p<0.006;). In addition, the amount of variability in the reaction time measure was also correlated with the number of hours since last sleep ($r_{(37)}$=-0.32,p<0.03). The Dual Task, like the other two tasks, showed some relationships between performance and the exposure to night work and sleep in the past week, however the nature of these relationships varied across measures. For the reaction time measures performance became worse as night work increased and the amount of sleep decreased. In contrast, tracking performance, as indexed by the level of difficulty achieved, was better in drivers who had done more night work. As the tracking task was continuous, it is likely to have assumed a primary focus in the Dual Task, so that reaction performance became poorer as more attention was paid to the tracking element. In tired drivers it might be expected that a decline in reaction time occurred as a consequence of more attention being paid to the tracking component.

The relationship between rated fatigue and performance was also investigated using correlation analysis. The range of fatigue scores was very small at baseline so it would not be expected that fatigue and performance would exhibit relationships. The only tests that did show a significant relationship with fatigue were the palmtop version of the Simple Reaction Time test in which reaction speed was slower in drivers who were more fatigued ($r_{(36)}$=0.33,p<0.05) and the level of difficulty achieved in the tracking component of the Dual

Task in which drivers with higher fatigue showed better tracking performance ($r_{(37)}=0.38, p<0.02$).

Comparisons across first day of the study week

As planned, night starters did significantly more hours at night on day 1 of the study than the day starters ($t_{(17)}=3.94, p<0.001$) and spent proportionately more of their work hours driving at night than day workers ($t_{(17)}=4.44, p<0.001$) however there was no difference between the two groups in the overall number of hours worked on the first day ($t_{(17)}=0.87$,ns). There was also no difference between the two groups in the number of breaks reported over the first day ($t_{(17)}=0.63$, ns.). These results indicate that the two groups differed only in the amount of night work done over the first day, as intended (Table 7).

Experience of fatigue over the first day of the study week

Changes in subjective fatigue ratings for each subgroup are shown in Table 8. Analysis of variance results showed a significant overall multivariate effect of test occasion for laptop measures of subjective fatigue ($F_{(1,19)}=22.35, p<0.0001$) but the group effect and occasion by group interaction were not statistically significant (group effect $F_{(1,19)}=0.01$,n.s.; interaction effect $F_{(1,19)}=0.21$,n.s.). These results show that fatigue increased significantly across the first day for both groups. Analysis of diary fatigue ratings associated with the palmtop tests showed the same finding with significant increases in fatigue across the day ($F_{(1,34)}=29.30, p<0.0001$) but no group differences ($F_{(1,34)}=0.025$,ns.) or interaction effects ($F_{(1,34)}=0.15$,ns.).

TABLE 7: Work experiences on the first day of the trip for all drivers and for day starter and night starter groups separately showing means *(SD)* and ranges of scores

	Day starters (n=8)	Night starters (n=11)	All drivers (n=19)
Hrs work on day 1	13.07	14.26	13.76
	(3.75)	*(2.18)*	*(2.91)*
	7.67 – 19.00	12.00 – 18.50	7.67 – 19.00
Night hours on day 1	4.8	8.92	7.18
	(2.53)	*(2.03)*	*(3.02)*
	0.00 – 7.25	*6.25 – 11.50*	*0.00 – 11.50*
% working hours done at night	34.87	62.37	50.8
	(16.8)	*(10.3)*	*(19.0)*
	0.00 – 50.00	*46.02 – 79.17*	*0.00 – 79.17*
Mean number of breaks on day 1	0.5	0.73	0.63
	(0.79)	*(0.79)*	*(0.76)*
	0.00 – 2.00	*0.00 – 2.00*	*0.00 – 2.00*

TABLE 8: Mean *(SD)* ratings of subjective fatigue at the beginning and end of the first day of the trip and at the end of the working week

	Group		
Fatigue Ratings	Day	Night	Total
Laptop	n=10	n=11	n=21
• Beginning day 1	24.50 *(22.42)*	22.88 *(13.81)*	23.65 *(17.95)*
• End day 1	47.17 *(23.32)*	46.97 *(23.21)*	47.06 *(22.67)*
Diary/Palmtop	n=20	n=16	n=36
• Beginning day 1	25.62 *(20.03)*	28.03 *(21.85)*	26.69 *(20.59)*
• End day 1	48.60 *(19.27)*	47.96 *(21.66)*	48.31 *(20.07)*

Changes in performance over the first day of the study week.

Simple Reaction Time performance across day 1 of the study is shown in Table 9. Analysis of the laptop and palmtop versions of the Simple Reaction Time test was by two factor analysis of variance with one repeated factor. For the laptop version, the results showed no significant effect of test occasion, group or interaction effect for Simple Reaction Time (for occasion, $F_{(1,17)}$=0.11,ns.; for group, $F_{(1,17)}$=0.005,ns.; for

TABLE 9: Mean performance on Simple Reaction Time tests at the start and end of the first shift of the week, with predicted performance equivalents for 0.05% BAC derived from an earlier laboratory study.

Performance measure	*Group (n)*	Start of week Mean	SD	End of first shift Mean	SD	BAC 0.05% Mean	SD
LAPTOP TESTS						*(n=38)*	
• RT	*Day starters (8)*	534.50	53.98	525.88	39.30		
	Night starters (11)	524.09	50.00	539.64	66.92		
	Total (19)	528.47	50.50	533.84	56.01	534.0	58.90
• RT variability	*Day starters (8)*	77.63	33.52	101.63	40.51		
	Night starters (11)	85.82	26.64	97.45	42.46		
	Total (19)	82.37	29.13	99.21	40.55	94.79	52.05
• # missed	*Day starters (8)*	1.13	0.99	0.63	0.74		
	Night starters (11)	0.36	0.67	0.45	0.82		
	Total (19)	0.68	0.89	0.53	0.77	1.17	238
PALMTOP TESTS							
• RT	*Day starters (16)*	607.63	65.45	630.00	76.62		
	Night starters (15)	619.13	72.36	610.40	68.73		
	Total (31)	613.19	67.97	620.52	72.38	na	na
• RT variability	*Day starters (16)*	90.94	24.46	100.94	28.62		
	Night starters (15)	95.23	32.83	98.80	32.15		
	Total (31)	93.03	28.40	99.90	29.88	na	na
• # missed	*Day starters (16)*	1.81	2.66	2.00	2.92		
	Night starters (15)	2.20	3.03	0.93	1.22		
	Total (31)	2.00	2.80	1.48	2.29	na	na

interaction, $F_{(1,17)}$=1.36,ns.). For the variability measure of reaction speed, there was a trend for variability to be higher at the end of the work day ($F_{(1,17)}$=3.86,p<0.066), but again no group or interaction effects ($F_{(1,17)}$=0.02,ns.; $F_{(1,17)}$=0.47,ns. respectively). There was no effect of test occasion on the number of missed responses in this test ($F_{(1,17)}$=1.48,ns.), nor of group ($F_{(1,17)}$=1.95,ns.) but the interaction effect indicated a trend which was due to day starters improving performance by missing fewer responses over the work day and night starters remaining at much the same level of performance ($F_{(1,17)}$=3.09,p<0.097).

The palmtop version of the Simple Reaction Time test showed no significant effects of test occasion, group or their interaction for any measure indicating that performance on this version of the test did not change across day one (for reaction speed, occasion, $F_{(1,29)}$=0.39,ns.; group, $F_{(1,29)}$=0.031,ns.; interaction, $F_{(1,29)}$=2.0,ns.; for variability of reaction speed, occasion, $F_{(1,29)}$=1.64,ns.; group, $F_{(1,29)}$=0.014,ns.; interaction, $F_{(1,29)}$=0.38,ns.; for missed responses, occasion, $F_{(1,29)}$=0.70,ns.; group, $F_{(1,29)}$=0.26,ns.; interaction, $F_{(1,29)}$=1.27,ns.).

Mean performance on the Mackworth Clock Vigilance task across day 1 of the study is shown in Table 10. Analysis of the laptop version of the test revealed that the reaction time measure showed no significant effect of test occasion ($F_{(1,17)}$=1.06,ns), group ($F_{(1,17)}$=1.13,ns) or interaction ($F_{(1,17)}$=1.14,ns). The variability of reaction speed showed similar non-significant effects (for test occasion, $F_{(1,17)}$=1.88,ns.; group, $F_{(1,17)}$=1.88,ns.; interaction, $F_{(1,17)}$=1.56,ns), as did the number of false alarms (for occasion, $F_{(1,17)}$=0.69,ns.; group, $F_{(1,17)}$=0.088,ns.; interaction, $F_{(1,17)}$=0.095,ns) and the number of missed signals (for test occasion, $F_{(1,17)}$=1.03, n.s.; for group, $F_{(1,17)}$=0.14,n.s.; $F_{(1,17)}$=0.026, n.s.). For the palmtop version, reaction speed results also showed no significant effect of test occasion ($F_{(1,27)}$=0.06,ns.), group or the interaction ($F_{(1,27)}$=1.34,ns, $F_{(1,27)}$=0.074,ns. respectively) and variability of reaction speed also showed no significant effects (for test occasion, $F_{(1,27)}$=1.46,ns; for group effect, $F_{(1,27)}$=0.8,ns; $F_{(1,27)}$=0.12,ns). In addition, the false alarm measure showed no significant effects for this version of the test (for test occasion, $F_{(1,27)}$=1.14,ns; group, $F_{(1,27)}$=0.61,ns, interaction effect, $F_{(1,27)}$=0.83,ns). In contrast, the palmtop version showed significant effects of test occasion in the number of missed responses ($F_{(1,27)}$=6.19,p<0.019) indicating that missed signals decreased significantly across the shift for both groups. The group and interaction effects for the missed signals measure were not significant (for group effect, $F_{(1,27)}$=0.91,ns; for interaction effects,

TABLE 10: Mean performance on Mackworth Clock Vigilance tests at the start and end of the first shift of the week, and predicted performance equivalents for 0.05% BAC from an earlier laboratory study

Performance measure	*Group (n)*	Start of week Mean	Start of week SD	End of first shift Mean	End of first shift SD	BAC 0.05% Mean	BAC 0.05% SD
LAPTOP TESTS						*(n=38)*	
• RT	*Day starters (8)*	862.88	84.70	860.75	91.29		
	Night starters (11)	883.27	84.31	1008.36	368.93		
	Total (19)	874.68	82.74	946.21	290.63	1094.	312
• RT variability	*Day starters (8)*	83.75	33.94	100.50	61.96		
	Night starters (11)	118.36	53.21	488.91	826.91		
	Total (19)	103.79	48.26	325.37	648.22	304	443
• # missed	*Day starters (8)*	1.50	0.93	1.88	1.25		
	Night starters (11)	1.81	1.54	2.09	2.43		
	Total (19)	1.68	1.29	2.00	1.97	4.09	3.48
• # false alarms	*Day starters (8)*	1.50	1.41	1.13	1.36		
	Night starters (11)	2.00	3.52	1.18	2.64		
	Total (19)	1.79	2.78	1.16	2.14	1.63	2.12
PALMTOP TESTS							
• RT	*Day starters (15)*	882.80	102.31	858.93	131.69		
	Night starters (14)	911.71	143.85	954.93	287.49		
	Total (29)	896.76	122.71	905.28	222.33	na	na
• RT variability	*Day starters (15)*	62.20	49.00	103.53	110.42		
	Night starters (14)	91.07	69.99	165.71	356.55		
	Total (29)	76.14	60.75	133.55	257.14	na	na
• # missed	*Day starters (15)*	0.67	0.82	0.40	0.51		
	Night starters (14)	1.07	1.00	0.43	0.76		
	Total (29)	0.86	0.92	0.41	0.63	na	na
• # false alarms	*Day starters (15)*	0.27	0.46	10.13	37.04		
	Night starters (14)	0.93	1.69	1.71	4.27		
	Total (29)	0.59	1.24	6.07	26.70	na	na

$F_{(1,27)}$=1.06,ns). Overall, therefore, only the missed signals measure showed any significant changes in this analysis, although this was only shown for the palmtop version of the test. Both groups of drivers showed improvements in detecting signals across the work day.

The Dual Task measures on the laptop (Table 11) showed that again there were no significant effects for reaction speed (for test occasion, $F_{(1,17)}$=1.9,ns; for group, $F_{(1,17)}$=0.76,ns; for the interaction, $F_{(1,17)}$=0.14,ns) although the variability of reaction speed showed a trend to be lower at the end of the day than at the start (for test occasion, $F_{(1,17)}$=3.45,p<0.08; for group, $F_{(1,17)}$=1.22,ns; for the interaction, $F_{(1,17)}$=0.61,ns). For the number of misses, no analysis could be conducted as no errors were made. Lastly, the level of difficulty achieved in the tracking task showed no effect of test occasion or for the interaction effect (for test occasion, $F_{(1,17)}$=0.35,ns; for the interaction, $F_{(1,17)}$=0.78,ns), but there was a significant group effect, ($F_{(1,17)}$=7.71,p=0.013), indicating that night drivers achieved lower levels of difficulty on both test occasions.

TABLE 11: Mean performance on Dual Task tests at the start and end of the first shift of the week, and predicted performance equivalents for 0.05% BAC derived from an earlier laboratory study

Performance		Start of week		End of first shift		BAC 0.05%	
measure	*Group (n)*	Mean	SD	Mean	SD	Mean	SD
• RT	*Day starters (8)*	660.88	72.64	633.75	81.17	*(n=38)*	
	Night starters (11)	719.00	182.90	671.55	132.60		
	Total (19)	694.53	146.65	655.63	112.69	725	175
• RT variability	*Day starters (8)*	187.13	115.58	143.50	87.25		
	Night starters (11)	289.09	228.82	181.73	149.24		
	Total (19)	246.16	192.25	165.63	125.34	253	200
• # missed	*Day starters (8)*	0.00	0.00	0.00	0.00		
	Night starters (11)	0.00	0.00	0.00	0.00		
	Total (19)	0.00	0.00	0.00	0.00	0.04*	0.22
• Task difficulty	*Day starters (8)*	47.35	20.80	49.28	27.07		
	Night starters (11)	30.30	26.86	20.57	14.89		
	Total (19)	37.48	25.37	32.66	24.90	45.43	26.6

* No significant relationship was demonstrated between this variable and sleep deprivation in the laboratory study, so the BAC equivalents tabled should be interpreted with caution.

Across the first work period, therefore, there was little evidence of performance decrements due to fatigue. This conclusion is reinforced by the finding that there were no significant correlations between fatigue and any of the performance measures for any performance tests. The only changes over the first study day were relatively minor. Both day and night drivers showed a trend to more variable reaction speed at the end of the work period in the laptop version of the Simple Reaction Time test. However, day drivers significantly improved in accuracy in detecting signals in both laptop Simple Reaction Time test and the palmtop version of the Mackworth Vigilance test whereas, night drivers only showed an improvement in accuracy in the palmtop Mackworth test. Both groups became less variable across the work period for reaction speed in the Dual Task.

The influence of work-rest experiences over the first day on fatigue and performance at the end of the day were investigated by regression analysis. For this analysis, the work-rest variables included were the number of hours worked over the first day, the percentage of night hours worked and the number of breaks recorded over the day. When entered as a block, none of the work-rest variables were significant predictors for either the laptop or diary measures of fatigue (for laptop measures $F_{(3,17)}=0.86$,ns, $r^2=-0.02$; for diary measures, , $F_{(3,16)}=1.31$,ns, , $r^2=-0.05$). Inspection of the correlation matrix showed that the number of hours worked in the day was significantly correlated with fatigue ratings from both laptop and diary measures (for laptop, $r_{(21)}=0.34, p<0.065$; for diary, $r_{(20)}=0.44, p<0.027$) indicating that drivers who worked longest over the day rated themselves as most tired.

For the Simple Reaction Time test, the work-rest variables did not significantly predict differences in performance across the work day for any measures for either laptop and palmtop versions (laptop version, reaction time, $F_{(3,17)}=1.93$,ns, , $r^2=0.12$; variability in reaction speed, $F_{(3,17)}=0.44$,ns, , $r^2=-0.09$; missed signals, $F_{(3,17)}=0.13$,ns, , $r^2=-0.15$; palmtop version, , reaction time, $F_{(3,17)}=1.33$,ns, , $r^2=-0.05$; variability in reaction speed, $F_{(3,17)}=1.72$,ns, , $r^2=0.10$; missed signals, $F_{(3,17)}=0.26$,ns, , $r^2=-0.12$). The bivariate correlations revealed that hours worked over the first day was significantly related to reaction time for both laptop and palmtop versions of this test (laptop, $r_{(21)}=0.44, p<0.02$; palmtop, $r_{(21)}=0.38, p<0.04$). Drivers who worked the longest hours showed the slowest reaction times.

Regression analysis for the Mackworth Vigilance test showed no significant predictors from the set of work-rest variables for any of the measures for either the laptop or palmtop versions of this test (laptop version, reaction time, $F_{(3,17)}= 1.89$,ns, , $r^2=0.12$; variability in reaction speed, $F_{(3,17)}=2.40$,ns., $r^2=0.17$; missed signals, $F_{(3,17)}=2.17$,ns, , $r^2=0.15$; false alarms, $F_{(3,17)}=0.78$,ns, $r^2=-0.04$; palmtop version, , reaction time, $F_{(3,15)}=0.59$,ns, $r^2=-0.07$; variability in reaction speed, $F_{(3,15)}$=0.64,ns, , r^2=-0.07; false alarms, $F_{(3,15)}$=0.58,ns, r^2=-0.08; missed signals, $F_{(3,15)}$=0.29,ns, r^2=-0.13). For the laptop version of this test, bivariate correlations showed that drivers who worked longer hours over the first day missed a significantly greater number of signals ($r_{(20)}$=0.46,p<0.018). The palmtop version did not show any significant bivariate correlations between work-rest variables and performance.

The relationship between Dual Task performance at the end of the work day and work-rest variables was also investigated using regression analysis. The work-rest variables did not predict reaction time or variability of reaction speed in this test (reaction time, $F_{(3,17)}$= 1.30,ns, , r^2=0.04; variability in reaction speed, $F_{(3,17)}$=0.81,ns., r^2=-0.03), but the level of difficulty achieved in the tracking component of the test was significantly predicted by the work-rest variables ($F_{(3,17)}$= 4.38,p<0.02, , r^2=0.34), in particular the percentage of night hours worked (See Table 12). The only significant bivariate correlations were for tracking performance with higher night hours being associated with poorer tracking performance ($r_{(21)}$)=-0.60,p<0.002) and more breaks also being associated with poorer tracking performance ($r_{(21)}$)=-0.38,p<0.05).

TABLE 12: Regression statistics for laptop Dual Task measures at end of first shift

	REGRESSION STATISTICS			
	Beta	Standard error	Standardised Beta	Test result
PREDICTORS OF TASK DIFFICULTY				
• Constant	87.33	22.10	-	t=3.95, p=0.001
• Hours worked in first shift	-1.27	1.57	-0.15	t=-0.80, p=0.43
• Percent of work hours at night in first shift	-0.63	0.25	-0.49	t=-2.52, p=0.02
• Number of breaks reported in first shift	-8.00	6.05	-0.25	t=-1.32, p=0.20

Comparing performance across the first study day with the performance standards set in the laboratory study (Williamson et al., 2000) for 0.05%BAC equivalence, revealed that for all measures and all tests, performance levels were better than the standard at study baseline and at end of the study day. The only exceptions were the degree of variability in reaction time in the Mackworth task which increased by the end of the first shift to slightly higher levels than found at 0.05%BAC and with a considerably larger standard deviation. Also, the level of difficulty achieved in the tracking component of the Dual Task was lower and more poorly performed than drivers had been achieving when their blood alcohol level was at 0.05%.

Description of the study week

There was considerable variation in work experiences over the study week. On average drivers did around 65 hours work in the study week, but this ranged from lower than 50 hours in the week for 11.5% of drivers to more than 72 hours for nearly one-quarter of drivers (Table 13). Around half of study week hours were done at night for all drivers, but again there was a wide range, with one-quarter of drivers doing less than 24 hours of night driving over the study week and an equivalent proportion doing more than 37 hours. The percentage of work hours done at night

TABLE 13: Work experiences in the study week for all drivers showing means, standard deviations and range of scores

	All drivers		
	Mean	SD	Range
Work			
• Hrs work in week	65.4	12.6	42.2 – 93.1
• Night hours in week	32.3	9.2	18.3 – 51.7
• % working hours done at night	50.1	13.0	30.9 – 89.2
Rest			
• Mean number of breaks in week	6.9	2.4	4.0 – 12.0
• Total sleep in week (hrs)	27.2	6.5	16.0 – 38.0
• Median length of sleeps during breaks	6.1	1.4	3.0 – 10.0
• Number of sleeps taken in breaks across the week	4.8	1.4	3.0 – 8.0
• % breaks with sleep	75.4	24.2	37.5 – 100.0

was also very variable. For some drivers (around 10 percent) less than one-third of work hours were done at night, whereas a significant percentage of drivers (19.2%) did twice that proportion. No drivers did no night driving or all night driving.

Most drivers had between 6 and 7 substantial breaks in the study week, with all having a least 3 breaks. On average drivers had around 27 hours of sleep in the study week, although for a few drivers this ranged to below 18 hours dependent, in part, on the number of shifts comprising the work week. For most drivers the majority of breaks included sleep, with the median length of sleep being around 6 hours, consistent with the minimum break length required by regulation.

Changes in fatigue ratings across the study week.

Analysis by T-test of the change in self-rated fatigue over the study week showed significant increases for both laptop and palmtop versions of this measure (for laptop ratings, $t_{(30)}=-3.99, p<0.0001$; for palmtop ratings, $t_{(31)}=-3.99, p<0.0001$). The change in ratings of fatigue increased by a very similar amount for both forms of the rating measure.

Regression analysis was performed to examine the role of work-rest factors across the study week on the experience of fatigue. The predictor variables were the total known hours spent working, the number of night hours worked, the number of hours spent sleeping and the total number of breaks taken over the study week. For the laptop ratings of fatigue, none of the predictor variables was revealed as significant for predicting changes in fatigue ratings ($F_{(4,26)}=0.93$,ns., $r^2=0.009$, ns.). Inspection of the correlation matrix showed a trend for fatigue to be related to the number of breaks taken over the past week ($r_{(31)}=-0.25, p<0.085$) suggesting that fatigue tended to be higher in drivers who took fewer breaks. Interestingly there was no relationship between the number of total hours worked or the number of night hours worked in the week and the number of breaks taken (for hours worked, $r_{(31)}=0.003$, ns.; for night hours worked, $r_{(31)}=0.05$,ns.)

Performance changes over the study week

Comparison of performance on laptop and palmtop tests at the beginning and end of the study week was by matched T-test. Means are presented in Table 14. For both the laptop and palmtop versions of the Simple Reaction Time, there was no significant change in for any of the measures, reaction speed, variability of reaction speed or the number of missed signals across the study week (laptop tests - reaction speed, $t_{(30)}=0.6$,ns.; variability, $t_{(30)}=0.47$,ns.; missed signals, $t_{(30)}=1.3$,ns.; palmtop tests, reaction speed, $t_{(33)}=0.36$,ns.; variability, $t_{(33)}=1.35$,ns.; missed signals, $t_{(33)}=0.63$,ns.).

For the Mackworth Vigilance test, in the laptop test mode, only the number of missed signals showed a significant increase across the study week ($t_{(31)}$=2.54,p<0.016) while the other measures showed no significant change across the week (for reaction speed, $t_{(31)}$=1.04,ns.; for variability of reaction speed, $t_{(31)}$=0.69,ns.; for false alarms, $t_{(31)}$=1.63,ns.). The palmtop version of the Mackworth Vigilance test showed a significant increase in variability of reaction speed across the study week ($t_{(31)}$=2.17,p<0.04), but not for any of the other measures of this test (for simple reaction speed, $t_{(31)}$=0.35,ns.; for false alarms, $t_{(31)}$=1.06,ns.; for number of missed signals, $t_{(30)}$=1.49,ns).

The Dual Task also showed no significant change in performance for any measure between the beginning and end of the study week (for level of difficulty, $t_{(30)}$=0.23,ns.; for Simple Reaction Time, $t_{(30)}$=0.19,ns.; for variability of reaction speed, $t_{(30)}$=0.15,ns.; for number of missed signals, $t_{(30)}$=0.89,ns).

These results indicate that performance in general stayed at a similar level for most tests and measures for the duration of the study week. The only exceptions were in the Mackworth test, in which the laptop version showed a decreased ability to detect signals at the end of the week, compared to the beginning, and the palmtop test showed an increased variability in the speed of reacting to signals.

Comparison of performance across the study week, with the 0.05%BAC equivalence performance standards set in the laboratory study (Williamson et al., in press, CR189), again revealed that performance on all measures and all tests was still better than the standard even at the end of the study week. The only exception, again, was performance on the Dual Task where performance on the tracking component did not change over the study week and remained low in comparison to the laboratory standard and the number of misses was slightly higher at then end of the trip than the 0.05%BAC standard.

TABLE 14: Mean performance at the start and end of the study week

	Measure	Start of week Mean	Start of week SD	End of week Mean	End of week SD
LAPTOP TESTS					
• Simple Reaction Time (n=31)	*RT*	515.03	44.01	520.16	57.20
	RT variability	80.29	25.94	83.65	29.13
	# missed	0.42	0.76	0.71	1.22
• Mackworth Vigilance (n=31)	*RT*	875.32	120.43	916.35	208.02
	RT variability	134.23	185.89	195.84	456.42
	# missed	1.16	1.24	2.35	2.63
	# false alarms	1.71	1.99	0.97	1.72
• Dual Task (n=31)	*RT*	689.61	153.69	698.84	254.32
	RT variability	251.03	207.70	243.81	222.43
	# missed	0.10	0.30	0.23	0.80
	Task difficulty	34.29	24.92	35.21	24.34
PALMTOP TESTS					
• Simple Reaction Time (n=33)	*RT*	603.55	69.85	599.45	75.80
	RT variability	92.30	28.31	102.06	35.84
	# missed	1.94	2.81	1.64	2.50
• Mackworth Vigilance (n=32)	*RT*	884.25	118.20	891.81	122.86
	RT variability	75.53	50.46	119.19	110.06
	# missed	0.75	0.88	0.53	0.95
	# false alarms	0.81	1.80	0.50	0.67
FATIGUE RATINGS					
• Laptop (n=31)		21.45	18.64	42.42	26.60
• Diary (n=32)		26.17	20.72	48.66	26.47

The relationship between performance at the end of the study week and work-rest experiences throughout the week was investigated by linear regressions. As for the previous regressions, the work-rest variables entered as predictors were the total hours worked in the study week, the number of night hours worked, the hours of sleep obtained and the number of breaks during the week. The results for the Simple Reaction Time test showed that for both laptop and palmtop versions, none of the work-rest variables were significant predictors of reaction speed, the variability of reaction speed, nor the number of targets missed during this test (for the laptop version, simple reaction time, $F_{(4,25)}=2.32, p<0.085, r^2=0.15$; variability of reaction speed, $F_{(4,25)}=1.03$,ns., $r^2=0.004$; missed signals, $F_{(4,25)}=1.03$,ns., $r^2=0.004$; for the palmtop version, simple reaction time, $F_{(4,27)}=1.11$,n.s., $r^2=0.014$; variability of reaction speed $F_{(4,27)}=1.73$,ns., $r^2=0.086$; and missed signals, $F_{(4,25)}=1.93$,ns., $r^2=0.11$, ns.). Inspection of the correlation matrix showed similar results for both versions of the test in that reaction speed was significantly correlated with hours spent sleeping (for laptop, $r_{(30)}=-0.44, p<0.008$; for palmtop, $r_{(32)}=-0.35, p<0.025$) and the number of breaks taken (for laptop, $r_{(30)}=-0.33, p<0.036$; for palmtop, $r_{(32)}=-0.25, p<0.08$). Drivers who had more sleep and breaks showed faster reaction time. The variability of reaction speed measure was also significantly correlated with the number of breaks taken in the last week for both versions of the test (for laptop, $r_{(30)}=-0.27, p<0.07$; for palmtop, $r_{(32)}=-0.38, p<0.016$) indicating that drivers who had more breaks were less variable in speed of reaction. Similarly for the number of misses, both versions showed a significant correlation with the number of breaks (for the laptop version, $r_{(30)}=-0.31, p<0.05$); for palmtop, $r_{(32)}=-0.35, p<0.027$) and for the palmtop version, also with the amount of sleep ($r_{(32)}=-0.37, p<0.019$). Again, drivers who took more breaks and had more sleep missed fewer signals.

For the Mackworth Vigilance test, regression analysis using the same work-rest predictor variables showed no significant prediction by the work-rest variables for any of the Mackworth Vigilance test measures; Simple Reaction Time, variability of reaction time, false alarms or the number of misses signal. This was found for both the laptop and palmtop versions of the test (for the laptop version Simple Reaction Time, $F_{(4,25)}=0.36$,ns., $r^2=0.1$; variability of reaction speed, for the palmtop version, $F_{(4,25)}=0.3$,ns. , $r^2=0.11$; false alarms, $F_{(4,25)}=0.42$,ns., missed signals, $F_{(4,25)}=0.94$,ns., $r^2=0.008$; for the palmtop version, Simple Reaction Time, $F_{(4,26)}=1.06$,ns., $r^2=0.007$; variability of reaction speed $F_{(4,26)}=0.59$,ns., $r^2=0.058$; false alarms, $F_{(4,26)}=0.72$,ns., $r^2=0.039$; and missed signals, $F_{(4,25)}=1.51$,ns., $r^2=0.064$, ns.). Examination of the correlations between work-rest and performance measures on this test confirmed the regression results for the laptop version of this test since there were no significant correlations. For the palmtop version, however, the correlations indicated that the hours spent working in the study week were significantly associated with simple reaction speed such that longer hours were correlated with faster reaction speed ($r_{(31)}=-0.34, p<0.03$). Both the number of false alarms and missed signals were significantly higher in drivers who had less sleep in the past week (for false alarms $r_{(31)}=-0.31, p<0.04$; misses $t_{(31)}=-0.34, p<0.029$). Misses were also higher for drivers who had fewer breaks over the study week ($r_{(31)}=-0.3, p<0.05$).

Regression analysis was also performed for measures of the Dual Task using the same set of predictor variables. None of the measures showed significant regression results (for reaction time, $F_{(4,25)}=0.34$,ns., $r^2=0.10$; for variability of reaction time, $F_{(4,25)}=0.46$,ns., $r^2=0.08$;for the number of misses, $F_{(4,25)}=0.73$,ns., $r^2=0.04$;for level of difficulty, $F_{(4,25)}=1.35$,ns., $r^2= 0.05$). In addition, there were no significant correlations with the work-rest variables for reaction time, variability of reaction time or the number of misses. Only the level of difficulty in tracking showed significant correlations with the number of night hours worked and the number of breaks in the study week (with night hours worked, $r_{(30)}=-0.3, p<0.05$; with the number of breaks, $r_{(30)}=-0.31, p<0.05$). Drivers who had worked more night hours or had more breaks in the past week showed poorer tracking ability.

Investigation of the correlations between rated fatigue and performance at the end of the study week showed that fatigue was significantly associated with reaction time in all tests which had a reaction time component, such that drivers with higher ratings of fatigue had slower reaction time (for the Simple Reaction Time test, $r_{(31)}=0.46, p<0.009$; for Mackworth test, $r_{(31)}=0.42, p<0.017$; for Dual Task, $r_{(31)}=0.4, p<0.026$). In addition, fatigue was significantly associated with variability of reaction speed in the Simple Reaction Time and Mackworth tests (for the Simple Reaction Time test, $r_{(31)}=0.43, p<0.015$; for Mackworth test, $r_{(31)}=0.37, p<0.038$) and for the missed signals measure of the Mackworth and Dual Tasks (for the Mackworth test, $r_{(31)}=0.49, p<0.009$; for the Dual Task, $r_{(31)}=0.38, p<0.037$).

DISCUSSION

The results of this study indicate that on average professional drivers do not experience particularly high levels of subjective fatigue over a trip that was intended to comply with regulated hours. Reported fatigue increased over the trip, but only to quite moderate levels. Performance measures also suggested that the trip did not have any notable effects on performance capacity. Day and night trips were similar in that rated fatigue was no different at the end of each type of trip and they were also similar in performance for most tests at the beginning and end of the both trips. In fact the only significant changes in performance over the first day were improvements in the ability to detect irregular signals in both the Simple Reaction Time and Mackworth Vigilance tests. This change is suggestive of the effects of increased arousal or practice rather than the decreased arousal associated with fatigue.

These findings suggest that overall the 14 hour trip did not constitute a problem for drivers, even those doing their trips at night. Performance did not deteriorate significantly and compared to the laboratory-generated performance standard, over the 14 hour trip drivers were almost always below the 0.05%BAC level. The only notable exception was the increase in the average variability of reaction time and the greater range of responses in this measures seen in drivers at the end of the first shift compared to alcohol standards. This finding seems to reflect early deterioration in the capacity to perform this test for some but not all drivers.

It must be acknowledged, however, that this analysis was only able to look at the gross milestones of the beginning and end of the work day. The analysis reveals that at the end of the trip, drivers' performance was as good as it had been when they were rested at the beginning of the trip. In contrast, this analysis is not able to reveal whether there were periods within the trip during which performance was significantly impaired. Based on the laboratory study (Williamson et al., in press, CR189) and other studies (e.g., Monk, 1994), it might be expected, for example, that performance would be poorer between the hours of 2am and 6am. In this study performance was only measured within the trip when drivers stopped driving to take a significant break, therefore it is not possible to obtain estimates of changes in performance during the trip. It is possible that the changes in reaction time variability in the Mackworth test may reflect effects for some drivers more than others after the first shift.

Over the work week, the change in rated fatigue was similar to that reported across the first day, suggesting that most drivers were not experiencing an accumulation of fatigue across the week. Drivers did fairly long trips during the study week, but on average the amount of work was less than that allowed by regulations. On the other hand, around half of the working hours were done at night for most drivers which might be expected to increase fatigue. The results show that drivers took steps to manage fatigue however, by actively balancing their work demands with as much sleep as possible. Drivers took rest regularly, most breaks of an hour or longer involved sleep and most rest time was spent sleeping. The results show that most drivers were quite successful in achieving this balance as they were able to report relatively low levels of fatigue even at the end of the study week.

There was evidence of some deterioration in performance capacity over the study week. Drivers showed a significant drop in performance across the study week in the Mackworth Vigilance test. The number of missed signals and the degree of variability in speed of reaction increased across the study week. This suggests that experiences during the study week put pressures on drivers' performance capacity, particularly their ability to sustain attention over long periods. A small proportion of drivers reported considerably higher fatigue at the end of the study week. This suggestion is reinforced by the finding in every test, that at the end of the study week, higher self-reported fatigue was associated with significantly poorer reaction time performance and significantly poorer ability to detect infrequent signals. These results signal the importance of managing the build up of fatigue over the work week since there is a clear association between increased fatigue and poorer performance. It should be noted, however, that compared to the performance standards developed for the known effects of alcohol, on-road performance levels on all measurement occasions in this study for almost all tests were better than the 0.05% BAC standard. The only exception was an indication in the Dual Task which showed that missed signals increased over the work week to become slightly more likely than would be expected based on the alcohol standard.

Despite the fact that overall, drivers tended to be working within the regulatory hours, there was a fairly broad range of work-rest experiences for drivers before the study trip and during the study trip. This provided an opportunity to look at the influence of a range of work-rest variables on fatigue and on performance. The results showed predictable relationships between work-rest experiences and fatigue. Over the first day of the study week, fatigue was associated with more hours worked, and over the study week, it was associated with the number of rest breaks taken in the study week. At baseline none of the work-rest variables predicted fatigue, probably because of the relatively narrow range in fatigue ratings and because all drivers were "standardised" by having at least 25 hours rest before the study started.

The results also showed similar lawful relationships between work-rest experiences and performance. At baseline, all reaction time measures, all variability of reaction speed measures and the accuracy measures for the Simple Reaction Time and Mackworth Vigilance tests were associated with the amount of night work done in the past week, or the time since the driver's last sleep, or both. Drivers had taken a relatively long break between their last trip and the beginning of the study week. This should have allowed all drivers to overcome fatigue from the previous week's work. The results suggest that for a group of drivers, namely those who do most night work, the break was not long enough to overcome the chronic effects of fatigue. It is also likely that the performance of night drivers is poorer at the beginning of the study since they had been awake for significantly longer than the day drivers. This is likely to have increased their fatigue levels again, so beginning to counteract the effects of the long 25 hour rest break.

Across the first day of work, the only significant work-rest variable was the number of hours worked. Contrary to expectations, over this single work period, night work was not shown to be a significant influence on performance levels. The length of the trip influenced reaction speed in both versions of the Simple Reaction Time test and the number of missed signals in the Mackworth task. Although it was intended that drivers work strictly to the working hours regulations of a maximum of 14 hours work in a trip, operational needs meant that over the first day, a few drivers did slightly longer hours. These results suggest that the ability to react quickly and to detect infrequent signals is likely to be compromised when working hours are longer. From this study, it is not possible to estimate the exact length of working hours at which performance begins to be compromised. This is an important issue for setting safety limits on working hours and therefore needs further follow-up. It should be recognised, however, that the performance levels achieved across the first work shift in this study were within safe limits compared to the laboratory-based standards. These results therefore signal the *beginnings* of adverse effects on performance rather than actual unsafe performance. Further work is needed to determine how rapidly and at which point longer hours begin to produce unsafe performance.

Across the study week, there were consistent relationships between two work-rest variables, hours since last break and amount of sleep obtained, and fatigue and performance. Fatigue levels were significantly positively associated with the number of breaks taken across the study week. Higher levels of fatigue were also associated with slower reaction time in all tests, greater variability in reaction speed in the Simple Reaction Time and Mackworth Vigilance tests and a greater number of missed signals in the Mackworth Vigilance and Dual Tasks. Furthermore, the number of breaks taken and the amount of sleep obtained during the week were associated with performance. Drivers who had fewer breaks and less sleep during the week, by the end of the week showed slower and more variable reaction speed, and missed more signals in both reaction time and Mackworth Vigilance tests. From these results, a uniform picture has emerged that links activity aimed at reducing the build up of fatigue, namely taking of breaks and sleep, with subsequent fatigue and even more importantly from the viewpoint of safety, with performance.

Both baseline and end of week results confirm that controlling the accumulation of fatigue through adequate breaks and sleep is the key to fatigue management and safe performance. From this data, it is not possible to specify what patterns of rest are likely to be sufficient to maintain fatigue and performance at safe levels. Investigation of alternative arrangements for rest will be necessary to define the appropriate limits of work and rest for effective fatigue management.

Overall, these results indicate that the drivers' performance did not constitute a risk to safety at the measurement occasions across the study. The adverse changes in fatigue and performance that were seen across the week, however, were warnings of more significant change if drivers did not get adequate rest and sleep to keep their fatigue levels low.

2. Study of an alternative schedule for a simulated long trip

The aim of this part of the project was to evaluate the effects on driver fatigue of a long return trip which contravened current working hours regulations. The main work section of this trip involved a journey from Wagga Wagga to Brisbane and return which is a trip of approximately 3,000 km in total and was scheduled to take around 48 hours to complete. The schedule designed for this long trip also included a six hour continuous break immediately before the trip started and immediately after it ended.

This trip was chosen as a model as it was fairly typical of the usual demands on drivers in terms of the distance they needed to travel. Previous surveys of long distance drivers and fatigue showed that drivers averaged trips of 1260kms (Williamson et al, 1992). A useful model for industry therefore would be a work-rest schedule which allowed a long trip to be done within operational constraints but which was demonstrated to minimise fatigue.

METHOD

Design

A single group of long-distance drivers simulated the pattern of work and rest that would be required of them under a proposed 60 hour shift cycle incorporating a return trip from Wagga Wagga in southern NSW to Brisbane in Queensland. Although the proposed shift began and ended with 6 hour rest breaks, the current study focussed only on the intervening 48 hours. Under the proposed shift cycle, drivers were required to take a 6 hour break sometime within these 48 hours. Brief (15 minute) breaks were also required after every 3 continuous hours of work. Within these constraints, drivers tried to pattern their work and rest periods in a way that might realistically allow them to complete the return trip to Brisbane, but at the same time, did not involve working at fatigue levels beyond those at which they would feel comfortable driving. This meant that the work and rest patterns for individual drivers might vary widely. Figure 1 compares one interpretation of this shift regime with that allowed under current driving regulations.

At 2 hourly intervals, during "work" periods, performance on a battery of six cognitive and psychomotor tasks was tested and drivers rated their subjective fatigue. The performance tests had previously been examined over a 28 hour period of sustained wakefulness and were known to be systematically affected by sleep deficit. At 4 hourly intervals, and before and after every substantial (>1 hour) break from "work", two of the tests and the fatigue ratings were presented on palmtop computers. Comparison of palmtop and the PC versions of the tests was designed to further evaluate the usefulness of the handheld tests.

Subjects

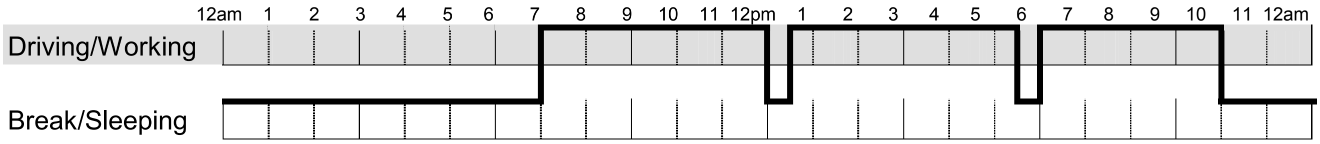
Thirteen professional long-haul drivers and three ex-drivers employed by a NSW transport company were recruited as subjects. Table 15 summarises the characteristics of the group. Subjects were all men and had an average of 12.9 years professional driving experience. All had been driving for at least three years. Twelve of the subjects had participated in an earlier study during which test performance when fatigued was calibrated against performance at known blood alcohol levels.

The subjects were almost all over 30 years of age. The majority of the subjects (75%) were aged between 30 and 49 years and almost one-fifth were aged between 50 and 59. Most (87.5%) were living in an ongoing relationship. Approximately two thirds of the group had ceased formal education before Year 11, with the rest completing post-secondary qualifications. Most (87.5%) of the group had little or no prior computing experience (except that obtained in the earlier calibration study). The majority of subjects drank alcohol (93.8%) but only a third (31.3%) currently smoked.

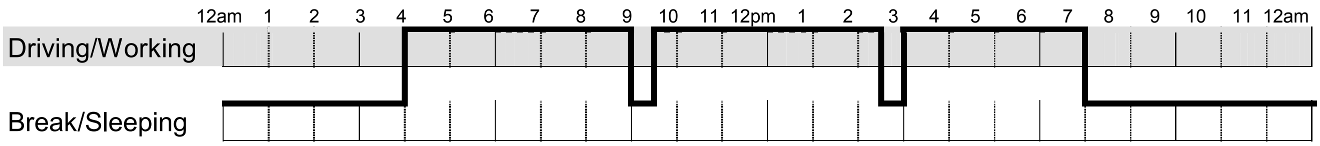
FIGURE 1: Comparison of simulated and regulated work-rest regimes.

Working Hours Regulations Guidelines:

DAY 1

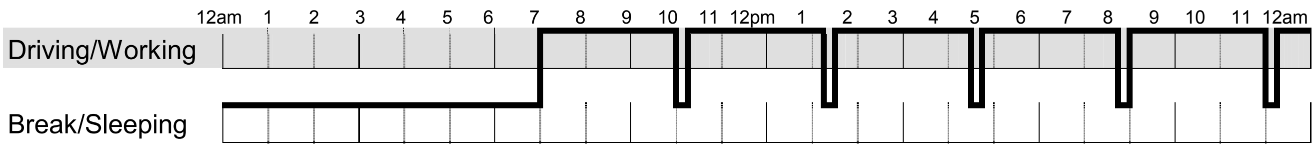


DAY 2

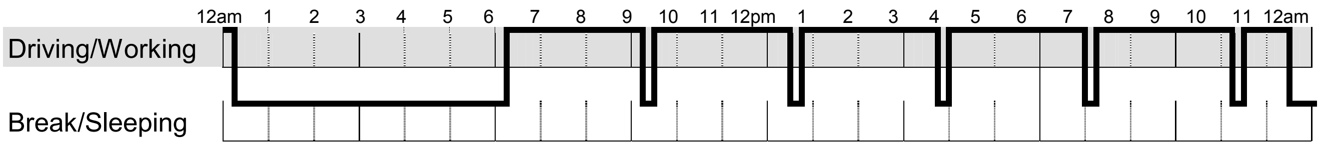


Driver Simulation FMP Guidelines:

DAY 1



DAY 2



Note: These timelines represent only one of many possible interpretations of each set of guidelines. For example, under FMP guidelines drivers must break for a MINIMUM only of 15 minutes every 3 hours but if they feel the need to break for longer they are able to do so. Also, drivers are not required to drive 16 hours each day. The guidelines stipulate that drivers may drive a maximum of 32 hours in any 48 hour period and for ease of presentation we choose to break this into 16 hours per day for the 2 days. The Working Hours Regulations also offer more flexibility than is demonstrated above, eg Drivers are required to break for 30 minutes in every 5 ½ hour block but it may be taken as one 30 minute break (as shown) or two 15 minute breaks taken within that 5 ½ hours.

TABLE 15: Demographic details of simulation study subjects.

Demographic Factor		Percentage of subjects
AGE:		
• 20 – 29 years		6.25
• 30 - 39 years		37.5
• 40 -49 years		37.5
• 50 -59 years		18.75
EDUCATION LEVEL:		
• High school years 7 - 10		62.5
• TAFE		25.0
• College/university		12.5
PC EXPERIENCE:		
• No previous experience		50.0
• A little experience		37.5
• Frequent PC user		12.5
SOCIAL DRUG HABITS:		
• Cigarettes	- Non-smoker	56.25
	- Ex-smoker	12.5
	- Current smoker	31.25
• Alcohol	- Non-drinker	6.25
	- Current drinker	93.75
RELATIONSHIP STATUS:		
• Married/Defacto		87.5
• Single		12.5
DRIVING EXPERIENCE:		
• Currently driving	- Yes	81.25
	- No	18.75
• Years of driving experience	- ≤ 6 years	30.75
(current drivers only)	- 6 - 13 years	30.75
	- Over 13 years	38.5
* *Least number of years driving experience - 3 years* *Greatest number of years driving experience - 30 years*		

Materials and measures

i. Performance testing

Six psychomotor and cognitive tests were selected from the Performance and Information Processing Systems (PIPS) Test Battery and were administered using Compaq laptop computers (4 in all). Subjects responded on peripheral Genovation keypads (Micropad 622) and standard serial mice. Each subject also used two custom-built "masks" that fitted over the keypads and concealed non-essential keys for the various performance tests. The masks were easily changed according to the instructions given at the start of each test.

The tests were selected on the basis that they had previously shown performance decrements as a function of increasing time awake. They included a Simple Reaction Time (RT) task, a Manual Tracking task, a Dual (RT and Tracking) Task, the Mackworth Clock Vigilance task, the Symbol Digit task and a Sequential Spatial Memory task. The tests took approximately 30 minutes to complete. Particulars of each test are presented below.

Simple Reaction Time (RT), Dual Task, and Mackworth Clock Vigilance Task: These tasks were identical to those used in the on-road study.

Manual Tracking: Subjects used a computer mouse to pilot a small dot around the screen in an effort to keep it inside a moving circle. As the subjects became more accurate, the regularity or predictability of the circle's movement decreased, making the task more difficult. If a subject became less accurate, the movement of the circle became more regular or predictable, and therefore easier to track. The regularity of the circle's movement was updated every 5 seconds across the 3 minute task, and was used as a measure of tracking accuracy.

Symbol Digit task: At the top of the screen, ten nonsense symbols were paired with the digits from 0 to 9. At the bottom of the screen, the nonsense symbols were presented individually and subjects were asked to type the paired digit, on the keypad, as quickly as possible. The task lasted 90sec. The number of symbols completed, reaction time, and the percentage of correct responses were measured.

Sequential Spatial Memory task: Squares in a 3 x 3 grid flashed for 250msecs each in a random sequence on the screen. Subjects then reproduced the sequence by clicking the computer mouse within each square in the sequence in turn. If the subject correctly reproduced the sequence, it was repeated with an additional square added to the end. If the subject made a reproduction error, a new sequence (trial), containing a single square, was initiated. Three such trials were completed. The length of the sequence correctly recalled on each trial was measured.

Nine different orders of the six tests were devised such that each test occurred equally often in the first, middle and last thirds of the orders. A unique schedule of test orders was then prepared for each subject to cover the entire study, assuming that they would complete a maximum of 18 test sessions. All nine test orders were included twice in each subject's schedule.

Testing also required four Hewlett Packard 200LX palmtop computers each attached to a Genovation keypad (Serial Micropad 623). Only the Simple Reaction Time task and the 5 minute version of the Mackworth Clock Vigilance task were used, as in the on-road study.

At the beginning and end of the laptop tests, subjects were presented with three Visual Analogue Scales, anchored by the terms Fresh and Tired, Clear-headed and Muzzy-headed, and Very Alert and Very Drowsy, on which to rate their current fatigue. On the palmtops, these scales were only presented before the tests, and were accompanied by questions about the timing of the last sleep and by Visual Analogue Scales of sleep quality and refreshingness.

ii. Documentation

All subjects completed an informed consent form (Appendix 5) and a brief questionnaire requesting details about their workload in the previous week, and their recent food and drug intake (Appendix 6). Basic demographic details had been collected from most of the subjects during their participation in an earlier study. Information on driving experience, age, marital and educational status, and social drug taking was collected verbally from those subjects who had not participated previously.

iii. "Work" activities

During their simulated "worktime", subjects were asked to engage in activities requiring a degree of concentration. To this end they were provided with a variety of computer games and puzzles, with particular focus on driving tasks. Drinks and snacks were provided as requested during "work" time.

Procedure

Subjects were tested in groups of four. All testing was conducted at a local motel. Subjects were asked to stay at the motel for the duration of the study (2 days and 2 nights) to ensure that commuting, socialising and family duties did not disrupt the work/rest pattern of their simulated trip. A function room was dedicated to "work" and testing and subjects spent the majority of their time there, only retiring to separate rooms to sleep.

Subjects arrived at the testing room at approximately 7:00 am on day 1 of the study. The nature of the study was explained and they were asked to give formal consent before participating. Those who had not participated in the earlier calibration study were asked to arrive somewhat earlier (6:30 am) on the first day of the study to complete 2 practice sessions of the laptop tests immediately prior to completing their baseline test. The other subjects, who had completed the tests 23 times as part of the calibration study, were not required to do additional practice and began their baseline test session as soon as practical after arriving. At the end of their baseline laptop test, all subjects completed a baseline palmtop testing session. As part of this initial palmtop test session, brief questions about the amount, timing and quality of sleep obtained in the previous night were also recorded.

Once the baseline testing was complete, subjects were asked to complete a brief questionnaire about their work and rest patterns in the 7 days preceding the study, and about their food and drug consumption in the hours preceding the study. Those subjects who had not participated in the earlier calibration study were also asked to provide some background demographic information.

Subjects were then asked to formulate a provisional plan for their simulated trip, bearing in mind that under the FMP proposal the following breaks from work were mandatory:

i. 15 minutes after not more than 3 hours continuous work, and
ii. 6 continuous hours at least once during their 48 hours of participation.

Other sleep and meal breaks were included in the schedules according to the drivers' personal preferences. Drivers were also encouraged to include a low but reasonable estimate of loading and unloading time at midtrip. This was typically set at 4 hours, but some subjects scheduled their simulated trip in an "ideal" manner, which included only enough time at midtrip to change trailers, but not to load or unload, and not to make local load deliveries. It was emphasised that the schedules were only preliminary, and that if subjects became too tired to drive, they should take a break regardless of whether it had been included in their original schedule, and regardless of whether they would then have sufficient driving time to complete the trip. The schedule would then be updated to take the extra break time into account.

Once the schedules had been prepared, drivers began "working". Work activities comprised driving simulation games and other computer-based and paper and pencil games requiring concentration and some motor skill.

Starting from the 7:00am baseline, the 30 minute test sessions were administered on the odd hours up to and including 7:00am on the 3rd test day. Tests were only administered if drivers were "at work". Starting from the 7:00am baseline, the short palmtop tests were administered every 4 hours, when drivers were "at work" and also immediately before and after breaks of 1 hour or more. All tests were considered to occupy "work" time. A running tally of work and rest time for each subject was kept by the researchers.

During breaks, drivers were encouraged to leave the testing room.

Statistical Analysis

For these analyses, the three subjective rating scales were averaged to create a single fatigue ratings measure.

Comparisons were made on fatigue and performance measures at a series of five test milestones across the simulated trip; at the start and end of day 1, the start and end of day 2 and the start of day 3 which corresponded to approximately 1, 16.5, 24.8, 40.4 and 48.9 hours after the beginning of the study respectively. A repeated measures analysis of variance on one factor (test milestones) was performed for each measure. Multivariate F-test statistics are reported unless they differ from the univariate F-test statistic, in which case both are reported. Post hoc contrasts were done between each of the five tests and the test immediately preceding it. Three additional T-tests were conducted post hoc to compare the start of day 1 with the starts of day 2 and 3 and between the end of day 1 and the end of day 2.

Comparisons were also made between performance on each test during the simulation and performance levels standardised against alcohol. This analysis should provide a yardstick for estimating the level of safety risk at intervals across the simulated trip.

A comparison of laptop and palmtop data is presented separately in Appendix 7.

RESULTS

Recent work and rest history

Three subjects had been on leave from work in the 7 days leading up to the study (see Tables 16 and 17). The remaining subjects had worked an average of 56.13 hours (SD = 20.4). Most had worked 40 hours or more, up to 90 hours for one driver. On average, almost half of the work hours (mean = 21.8, SD = 12.6) were done at night (6pm to 6am). The majority of drivers had done at least some night work ranging up to 40 hours.

TABLE 16: Work in the last 7 days

	Mean	Range	Median	SD
Hrs since end of last shift				
- including leave *	71:28 hrs	4:30 – 240:00 *(10 days)*	59:00	41:3
- without leave *	46:04 hrs	4:30 – 142:00 *(~6 days)*	38:00	77:15
**(3 subjects on leave for past 7 days)*				
Work in past week *(7 days)*	56.13 hrs	15 –90 hrs	50.0	20.4
No. night hours in past week	21.79 hrs	0 – 40 hrs	50.0	12.6
Length of last trip	12:22 hrs	3:30 – 21:00 hrs	12:0	6:22

The length of the last work "day" averaged 12.4 hours (SD = 6.4), but just over one-third of drivers worked longer than 12 hours in the last shift up to 21 hours for one driver. A substantial number of the subjects (57.1%) had been off work for 48 hours or more at the start of the study, however 3 (18.8%) had had less than 12 hours off, and one of these had only had 4.5 hours off, in contravention of the FMP proposal being tested.

On the evening before the commencement of the study, subjects averaged 6.1 hours sleep (SD = 1.6), however the minimum was only 2.5 hours (Table 18). Because drivers are only assured a 6 hour break prior to their trip under the FMP proposal, many of the subjects got more sleep prior to the simulation than they might under real operational conditions. Indeed 44% of the sample slept for longer than 6 hours. Visual analogue ratings of sleep quality and refreshedness upon waking were both moderate, averaging slightly over 50.

On the morning of the test most drivers had consumed at least one drink caffeine containing (62.5%) and most had consumed breakfast (68.8%). It seems that in the past 12 to 24 hours most drivers were quite well-prepared for the long trip simulation since they had a reasonable amount of sleep and had eaten before they started. A proportion of drivers however had only a limited amount of sleep in the last 24 hours and a larger proportion had worked long hours over the past week. It is possible that some drivers may have started the simulated trip with significant amounts of fatigue so the relationship between recent work-rest experience and fatigue was examined. In addition, as a check, the relationship between work-rest history and performance at the start of the trip was also investigated for the two most reliable fatigue tests, Simple Reaction Time and Mackworth Clock Vigilance.

Relationship between recent work/rest history and fatigue and performance at the beginning of the trip (baseline)

Correlational analysis showed that there were statistically significant relationships between work-rest experience leading up to the trip and ratings of subjective fatigue.

TABLE 17: Distribution of work hours in the last 7 days.

Hours worked	N	% of subjects
Total work hours:		
• On leave	3	18.8
• ≤ 40 hours	3	18.9
• 41-60 hours	5	31.5
• 61 + hours	5	31.5
Night work hours:		
• On leave	3	18.8
• 0 hours	2	12.6
• ≤ 20 hours	5	31.5
• 21 + hours	6	37.8
Length of last shift:		
• ≤ 6 hours	3	18.9
• 7 – 14 hours	5	31.5
• 15 - 24 hours	6	37.8
Hours since end of last shift:		
• < 6 hours	1	6.3
• 6 – 12 hours	1	6.3
• 13 - 18 hours	2	12.6
• 19 - 24 hours	2	12.6
• 24 – 48 hours *(2 days)*	1	6.3
• > 2 days	9	56.7

TABLE 18: Details of rest the night before the start of the study

	Mean	Range	Median	SD
Waking time on Day 1 of the Study	05:14	04:15 – 06:00	05:30	0:31
Hrs of sleep night before start of study	6:04 hrs	2:30 – 8:00 hrs	6:07	1:35
Rating of quality of last sleep(/100)	54.67	10.0 - 90.0	55.0	25.94
Rating of refreshedness from sleep (/100)	52.67	10.0 - 85.0	50.0	20.86

As shown in Tables 19 and 21, fatigue ratings were significantly higher for drivers who had worker longer hours in the past week and for those who had worked more night hours. While there was no relationship between fatigue ratings and the number of hours since the last work shift finished, fatigue ratings were higher for drivers who had woken later. This apparently inconsistent finding occurred because in most cases drivers who had not finished their last shift until late the night before the simulated trip started, tended to sleep later. This is reinforced by the findings that the length of the last sleep period and the rated sleep quality before the study were quite strongly negatively correlated with fatigue ratings. Drivers who got less sleep and who felt it was poor quality tended to report more fatigue. Similarly, there was a strong negative correlation between rated freshness at waking and subjective fatigue.

TABLE 19: Relationships between rated fatigue, Simple Reaction Time performance and Mackworth Vigilance performance at baseline and pre-study sleep (Laptop data)

	Pre study waking time (N=14)	Length of pre study sleep (N=13)	Rated sleep quality[1] (N=14)	Rated freshness at waking[1] (N=14)
Simple Reaction Time				
RT	r=-0.37, p=0.19	r=0.25, p=0.41	r=0.60, p=0.02	r=0.20, p=0.49
RT variability	r=0.02, p=0.96	r=0.12, p=0.67	r=0.61, p=0.02	r=0.23, p=0.43
# missed	r=0.26, p=0.38	r=-0.01, p=0.97	r<0.01, p>0.99	r=0.12, p=0.68
Mackworth Vigilance				
RT	r=-0.07, p=0.83	r=0.44, p=0.13	r=0.62, p=0.02	r=0.57, p=0.04
RT variability	r=-0.11, p=0.71	r=0.39, p=0.18	r=0.56, p=0.04	r=0.36, p=0.20
# missed	r=-0.23, p=0.43	r=0.18, p=0.56	r=0.40, p=0.15	r=0.51, p=0.06
# false alarms	r=0.27, p=0.35	r=0.32, p=0.29	r=0.23, p=0.42	r=0.17, p=0.55
	(N=15)	(N=14)	(N=15)	(N=15)
Averaged fatigue ratings	r=0.55, p=0.03	r=-0.62, p=0.02	r=-0.64, p=0.01	r=-0.77, p<0.01

[1] Spearman's rho nonparametric correlations

Recent work-rest experiences therefore appear to be important influences of drivers' fatigue state before the trip started. As might be expected, drivers who had done more work and had poorer recovery sleep before the trip were more tired before they had even started the long trip simulation.

The results showed little influence of immediate past work-rest history on performance. Table 20 shows Simple Reaction Time test results for drivers who had finished their last shift a short (24 hours or less) or longer time ago (48 hours or more). Analysis by T-test demonstrated that drivers who had been off work for shorter periods tended to be faster on this test at baseline ($t_{(11)}$=2.12, p<0.058), but their performance was not more variable ($t_{(11)}$=0.066,n.s.) and they did not differ on the number of missed signals ($t_{(11)}$=0.35, n.s.) compared to drivers who had been off work for longer periods. Similarly, in the Mackworth Clock test, none of the measures, reaction time, variability in reaction time, number of missed signals or the number of false alarms, differed between drivers with more or less time off in the past week ($t_{(11)}$=0.29, n.s.; $t_{(11)}$=0.75, n.s.; $t_{(11)}$=0.098, n.s.; $t_{(11)}$=0.86, n.s.).

TABLE 20: Simple Reaction Time (RT) test results for drivers who had a short vs long time since last shift

Time since end of last shift	N	Mean	SD
Overall Mean RT -			
• 24 hours or less	6	501.7	40.0
• 48 hours or more	7	561.9	58.8
Overall SD of RT -			
• 24 hours or less	6	77.7	15.2
• 48 hours or more	7	78.6	10.6

There was some evidence of an association between fatigue and performance, but it was not in the expected direction (Table 21). Correlational analysis showed that reaction time was faster for drivers reporting greater fatigue in both Simple Reaction Time and Mackworth Vigilance tests, but there was no relationship between variability of responding and fatigue. For the Mackworth Vigilance test only, drivers reporting higher fatigue were also less likely to miss signals. These results suggest that at baseline prior work-rest experiences did influence the amount of fatigue drivers experienced, but higher reported fatigue did not have an adverse effect on performance.

TABLE 21: Relationships between rated fatigue, Simple Reaction Time performance and Mackworth Vigilance performance at baseline and pre-study work (Laptop data)

	Hours since last shift (N=14)	Hours worked in last week (N=15)	Hours worked at night (N=15)	Averaged fatigue ratings[1] (N=15)
Simple Reaction Time				
RT	r=0.23, p=0.42	r=-0.14, p=0.61	r=-0.29, p=0.29	r=-0.59, p=0.02
RT variability	r=-0.01, p=0.96	r=0.06, p=0.83	r=0.20, p=0.48	r=-0.40, p=0.14
# missed	r=-0.19, p=0.51	r=0.20, p=0.48	r=0.19, p=0.50	r=-0.35, p=0.20
Mackworth Vigilance				
RT	r=0.16, p=0.60	r<0.01, p>0.99	r=0.05, p=0.85	r=-0.68, p<0.01
RT variability	r=-0.11, p=0.71	r=0.23, p=0.41	r=0.27, p=0.34	r=-0.43, p=0.11
# missed	r=-0.04, p=0.88	r=0.16, p=0.58	r=0.08, p=0.77	r=-0.53, p=0.04
# false alarms	r=0.06, p=0.85	r=0.28, p=0.32	r=0.28, p=0.32	r=-0.11, p=0.70
	(N=15)	(N=16)	(N=16)	
Averaged fatigue ratings	r=-0.36, p=0.19	r=0.51, p=0.04	r=0.50, p=0.046	

[1] Spearman's rho nonparametric correlations

Description of simulated trip

The simulated trip involved 30.4 hours of work on average (range 28.2 – 33.2) spread over the 48 hours scheduled for the trip. Breaks took around one-third of trip time with an average of 18.04 hours being spent resting and sleeping (range 15.3 – 20.5). With long laptop tests every two hours during work periods, most drivers had between 17 and 19 tests over the study period.

The distribution of breaks over the trip are shown in Table 22. Most drivers had 11 breaks in total with most breaks being of around 15 minutes duration. All drivers had two long breaks of around six hours as planned in the schedule and for almost all drivers the first occurred during the fifth or sixth break and the second at the 11th break. Most drivers also had a break of around one hour usually for a meal in the fourth break and in the eighth or ninth break.

TABLE 22: Details of breaks on the trip

Break Number	Hours after start of simulation	% of subjects from total	Duration of break *(hrs)*	% of subjects taking break
1	3:00 – 3:45	93.75	0:15	100
2	5:15 – 6:20	100	0:15 0:25 0:30	25 25 50
3	8:15 – 9:10	93.75	0:15 1:20	93.75 6.25
4	10:20 – 12:00	93.75	0:15 1:05-2:05	6.25 93.75
5	12:30 – 14:30	81.1	0:15 ≤ 0:30 6:00 – 9:00	56.3 62.5 37.5
6	14:45 – 17:45	87.5	≤ 1:00 6:00 – 8:00	37.5 62.5
7	16:25 – 23:15	100	0:15 1:00	75 25
8	19:35 – 25:00	100	0:15 0:30 1:00 – 1:20	56.3 12.5 31.3
9	21:45 – 28:00	100	0:15 0:30 – 1:00 2:15 >6:00	25 63.0 6.3 6.3
10	23:55 – 30:00	93.75 *(1 subject stopped at 9 breaks)*	0:15 0:30 – 1:30 >6:00	37.5 44.1 12.5
11	27:05 – 32:45	87.5 *(1 subject stopped at 10 breaks)*	<0:30 4:25 – 7:25	12.5 87.5
12	28:00 - 29:00	14.3 *(only 2 subjects had a 12th break)*	0:30 6:05	50 50

Changes in subjective fatigue levels over the simulated trip

As can be seen in Table 23, subjective fatigue ratings varied considerably across the simulated trip. Repeated measures ANOVA showed an significant overall effect of time ($F_{(4,12)}$=6.90, p<0.004). Post hoc contrasts showed significant increases in ratings of fatigue between the beginning and end of each of the two days on the trip ($F_{(1,15)}$=14.7, p<0.002 for day 1; ($F_{(1,15)}$=20.23, p<0.001 for day 2).

TABLE 23: Mean (*SD*) fatigue and performance at the beginnings and ends of study days (Laptop data)

	Start day 1	End day 1	Start day 2	End day 2	Start day 3	Multivariate test result
Time of day						
	7:33 *0:27*	21:37 *5:31*	7:27 *0:49*	21:29 *5:26*	6:38 *1:04*	
Hours since start of study						
	1.00 *0.42*	16.51 *0.51*	24.82 *0.81*	40.35 *0.74*	48.91 *2.76*	
Averaged fatigue ratings (/100)						
	32.34 *19.85*	48.33 *17.42*	29.84 *17.89*	47.45 *18.60*	44.32 *20.22*	F(4,12)=6.90, p=0.004
Mackworth Vigilance						
RT	920.27 *202.03*	963.13 *186.51*	941.20 *156.80*	1115.93 *496.61*	960.53 *186.48*	F(4,11)=0.45 p=0.77
RT variability	187.07 *323.84*	186.13 *162.69*	166.93 *108.25*	312.13 *509.92*	157.97 *110.14*	F(4,11)=0.40 p=0.81
# missed	1.93 *2.28*	3.00 *2.48*	2.80 *2.04*	4.87 *3.62*	4.87 *3.29*	F(4,11)=10.61 p=0.001
# false alarms	1.73 *2.12*	2.20 *2.11*	1.87 *2.53*	3.13 *4.93*	2.73 *4.74*	F(4,11)=0.63 p=0.65
Simple Reaction Time						
RT	536.40 *55.03*	536.87 *34.89*	542.67 *55.15*	546.33 *49.64*	569.60 *75.95*	F(4,11)=1.60 p=0.24
RT variability	79.27 *29.78*	82.73 *21.59*	84.07 *35.85*	90.73 *41.03*	101.33 *33.74*	F(4,11)=1.07 p=0.42
# missed	1.19 *2.10*	0.69 *1.54*	0.81 *1.52*	2.44 *3.42*	2.50 *4.03*	F(4,12)=7.20 p=0.003
Dual Task						
RT	702.92 *175.27*	709.77 *124.85*	693.31 *137.71*	685.54 *95.94*	762.08 *173.89*	F(4,9)=1.41 p=0.31
RT variability	240.62 231.25	253.77 203.64	260.00 203.87	251.62 159.34	308.46 243.08	F(4,9)=0.20 p=0.93
# missed	4.67 *11.53*	0.13 *0.35*	0.00 *0.00*	0.13 *0.52*	0.01 *0.26*	F(4,11)=1.45 p=0.28
task difficulty	39.27 *21.84*	45.55 *22.62*	47.94 *20.20*	48.77 *24.78*	41.15 *21.53*	F(4,11)=1.71 p=0.22
Tracking						
task difficulty	38.61 *24.84*	39.75 *24.56*	37.16 *23.98*	35.08 *22.78*	36.33 *23.22*	F(4,11)=0.08 p=0.99

TABLE 23: continued

	Start day 1	End day 1	Start day 2	End day 2	Start day 3	Multivariate test result
Symbol Digit task						
RT	3104.33 *984.74*	2450.80 *429.61*	2567.53 *689.71*	2386.73 *391.91*	2372.47 *429.37*	F(4,11)=2.45 p=0.11 [1]
SD of RT	1522.80 *1326.90*	811.20 *359.16*	868.27 *574.47*	786.40 *252.96*	896.07 *415.03*	F(4,11)=1.49 p=0.27
# presented	29.80 *7.49*	35.60 *6.16*	34.87 *7.13*	36.20 *6.53*	36.60 *6.10*	F(4,11)=6.73 p=0.005
% correct	97.73 *3.01*	98.13 *2.45*	97.93 *2.89*	97.13 *3.60*	97.87 *3.02*	F(4,11)=0.21 p=0.93
Sequential Spatial Memory						
recall length	4.33 *2.17*	5.07 *2.35*	5.45 *2.00*	4.17 *2.35*	4.64 *1.89*	F(4,10)=1.99 p=0.17

[1] Univariate test of the occasion factor was significant F(2.1,29.9)=6.76, p<0.005

Between the end of day 1 and the beginning of day 2 there was a significant drop in fatigue ratings ($F_{(1,15)}$=29.4, p<0.001) indicating recovery following the rest period between day 1 and day 2. In contrast, there was no recovery of fatigue between the end of day 2 and the beginning of day 3 ($F_{(1,15)}$=0.45,n.s.). Repeated measures T-tests also showed that fatigue levels were about the same at the beginning and end of each of the two days of the trip ($t_{(15)}$=0.68, n.s.; $t_{(15)}$=0.43, n.s. respectively) but that fatigue increased significantly between the first and last tests of the trip ($t_{(15)}$=0.2.27, p<0.038).

Changes in test performance over the simulated trip

Simple Reaction Time test

Simple Reaction Time data at the end of study days are shown in Table 23. Analysis of the results for the Simple Reaction Time test across the trip showed a trend for the effect of time of test although this was not statistically significant (multivariate $F_{(4,11)}$= 1.60,n.s.; univariate $F_{(4,56)}$=2.39, p=0.06). There was no significant change in reaction speed between the

beginning and end of day 1 ($F_{(1,15)}$=0.002, n.s.) or day 2 ($F_{(1,15)}$=0.11, n.s.), between the end of day 1 and the start of day 2 ($F_{(1,15)}$=0.23, n.s.)
or the end of day 2 and the start of day 3 ($F_{(1,15)}$=1.99, n.s.). Post hoc t-tests however supported the overall trend for time of test by revealing a significant slowing in reaction time between the beginning of the trip and the end of the trip on day 3 ($t_{(14)}$=2.46, p=0.03).

Analysis of the variability of reaction time across the trip showed no main effect of time of test ($F_{(4,11)}$=0.4, n.s.) nor were any post hoc contrasts significant (beginning and end of day 1 ($F_{(1,14)}$=0.28,n.s.); beginning and end of day 2 ($F_{(1,14)}$=0.0.36,n.s.); between day 1 and day 2 ($F_{(1,14)}$=0.16,n.s.); and between day 2 and day 3 ($F_{(1,14)}$=0.38,n.s.).

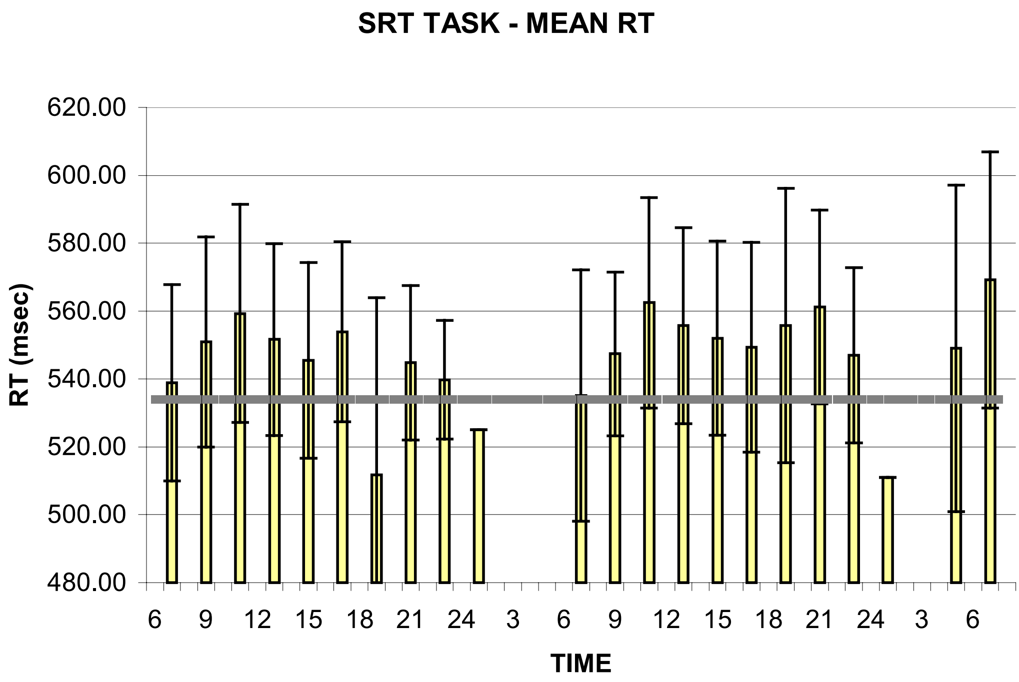
The third measure, number of missed signals, also showed no significant overall effect of time of test ($F_{(4,11)}$=1.07,n.s.), but post hoc contrasts revealed a significant increase in the number of missed signals across day 2 of the trip ($F_{(1,15)}$=7.75, $p<0.01$); which remained high between the end of day 2 and the beginning of day 3 ($F_{(1,15)}$=0.003,n.s.). There were no significant differences across the first day and between the end of day 1 and the beginning of day 2 ($F_{(1,15)}$=0.75,n.s.; $F_{(1,15)}$=0.07,n.s.). These results show that the likelihood of missing signals increased as the trip progressed.

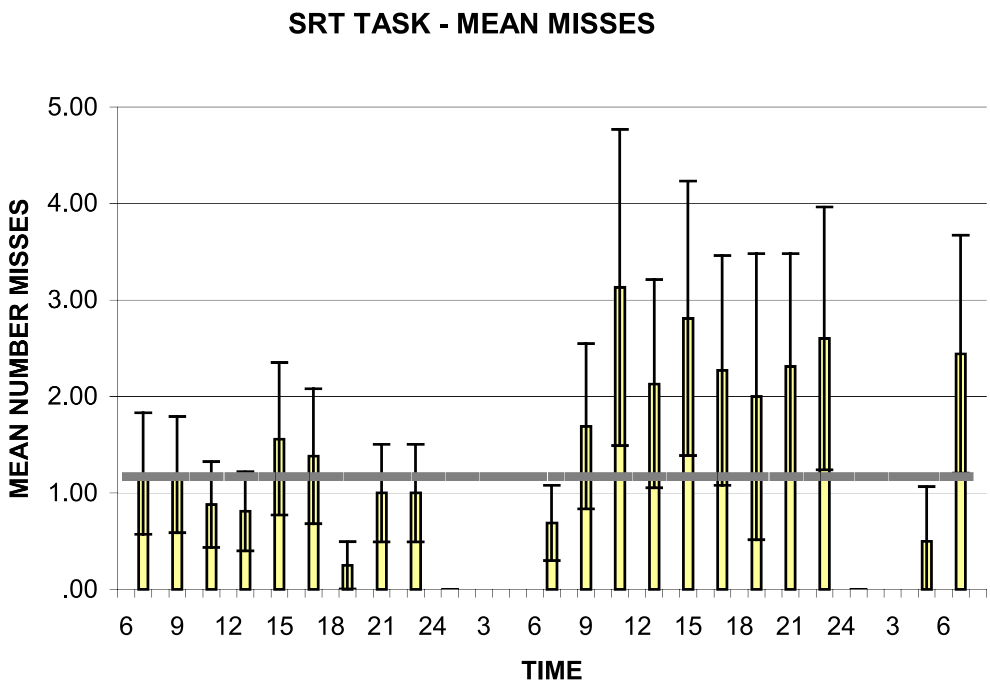
Comparison of Simple Reaction Time performance over the simulated trip and predicted performance at 0.05% BAC (see Figure 2) showed that performance in the simulation was poorer than the 0.05% BAC level for most of the trip. Even at the beginning of the trip, average reaction speed in the simulation was slower than predicted 0.05% level although the number of misses was not.

Mackworth Clock test

Reaction speed results across the trip for the Mackworth Clock test are shown in Table 23. Analysis showed no significant overall effect of time of test ($F_{(4,11)}$=0.45,n.s.) and post hoc contrasts also showed no differences in reaction time between any of the test occasions (beginning and end of day 1 ($F_{(1,14)}$=0.35,n.s.) beginning and end of day 2 ($F_{(1,14)}$=1.93,n.s.), between days 1 and 2 ($F_{(1,14)}$=0.47,n.s.) and between days 2 and 3 ($F_{(1,14)}$=1.45,n.s.).

FIGURE 2: Simple Reaction Time measures with 95% confidence intervals. Dotted lines indicate performance at BAC = 0.05%.





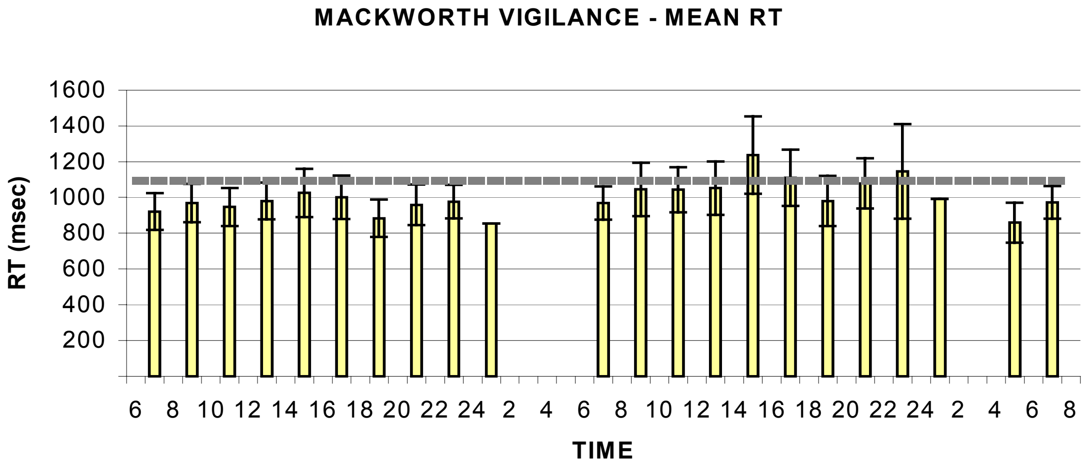
Analysis of the variability in reaction speed across the trip also showed no significant overall trend ($F_{(4,11)}$=0.40,n.s.) or post hoc contrasts (beginning and end of day 1 ($F_{(1,14)}$=0.00,n.s.) beginning and end of day 2 ($F_{(1,14)}$=1.20,n.s.), between days 1 and 2 ($F_{(1,14)}$=0.26,n.s.) and between days 2 and 3 ($F_{(1,14)}$=1.43,n.s.).

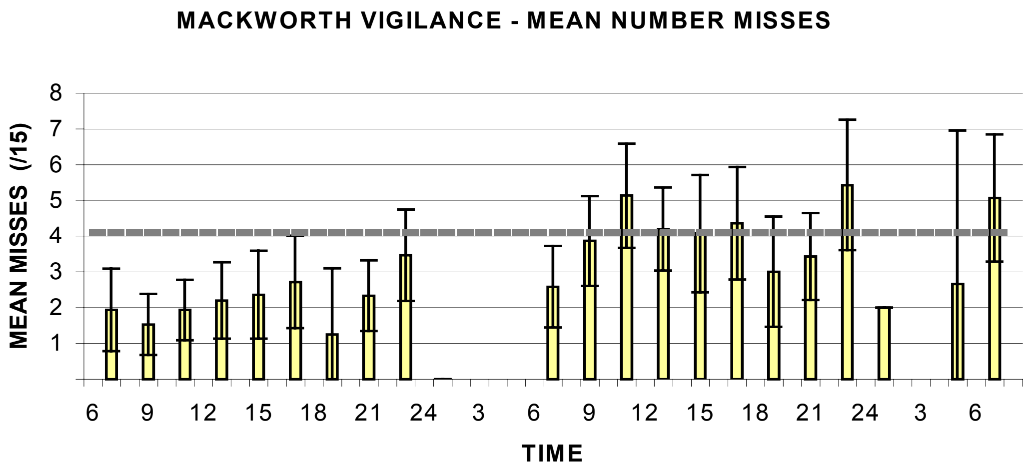
The number of missed signals increased significantly across the trip (see Table 23) as shown by the significant repeated measures main effect for time of test ($F_{(4,11)}$=10.61,p<0.001). Within the trip, post hoc contrasts showed no significant change in the number of misses up to the beginning of the second day of the trip (across day 1, $F_{(1,14)}$=1.60,n.s.; between day 1 and 2, $F_{(1,14)}$=0.09,n.s.), but a large and significant increase in the number of missed signals occurred across day 2 ($F_{(1,14)}$=5.13,p=0.04) which remained high until the end of the trip at the beginning of day 3 ($F_{(1,14)}$=0.001,n.s.). Further T-tests also confirmed a significant increase in missed signals between the beginning and end of the trip ($t_{(14)}$=3.97, p<0.001).

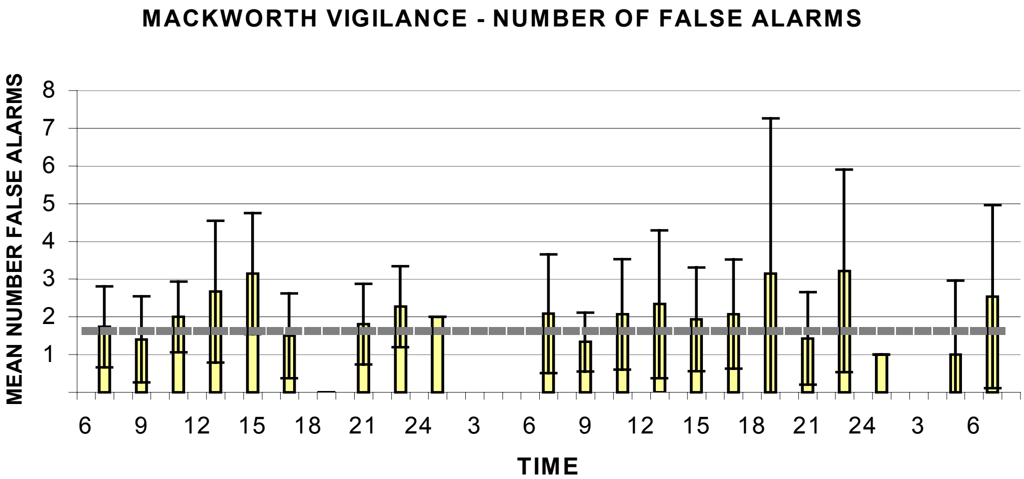
Finally, analysis of the number of false alarms in this test across the trip showed no significant trend overall ($F_{(4,11)}$=0.63,n.s.) and no significant within trip differences (beginning and end of day 1 ($F_{(1,14)}$=0.48,n.s.) beginning and end of day 2 ($F_{(1,14)}$=1.09,n.s.), between days 1 and 2 ($F_{(1,14)}$=0.19,n.s.) and between days 2 and 3 ($F_{(1,14)}$=0.45,n.s.).

Compared to the 0.05% BAC standard limit, reaction time and the number of misses (Figure 3) in the Mackworth task were better than the standard limit for most of the trip, although by the middle of the afternoon of the second day, reaction speed was lower than the 0.05% limit and the number of misses had increased to the standard level or worse by around mid-morning of the second day. Reaction time levels and misses returned to better than standard levels at the beginning of day 3, but misses increased to poorer than the standard again very early in day three. The number of false alarms in the Mackworth test were poorer than the standard from mid-afternoon on the first day and remained higher throughout most of the second day. They were better than the standard at the beginning of day 3, but increased to become worse than the standard early in day 3, by only the second test session of the day.

FIGURE 3: Mackworth Vigilance task measures with 95% confidence intervals. Dotted lines indicate 0.05%BAC equivalence level.





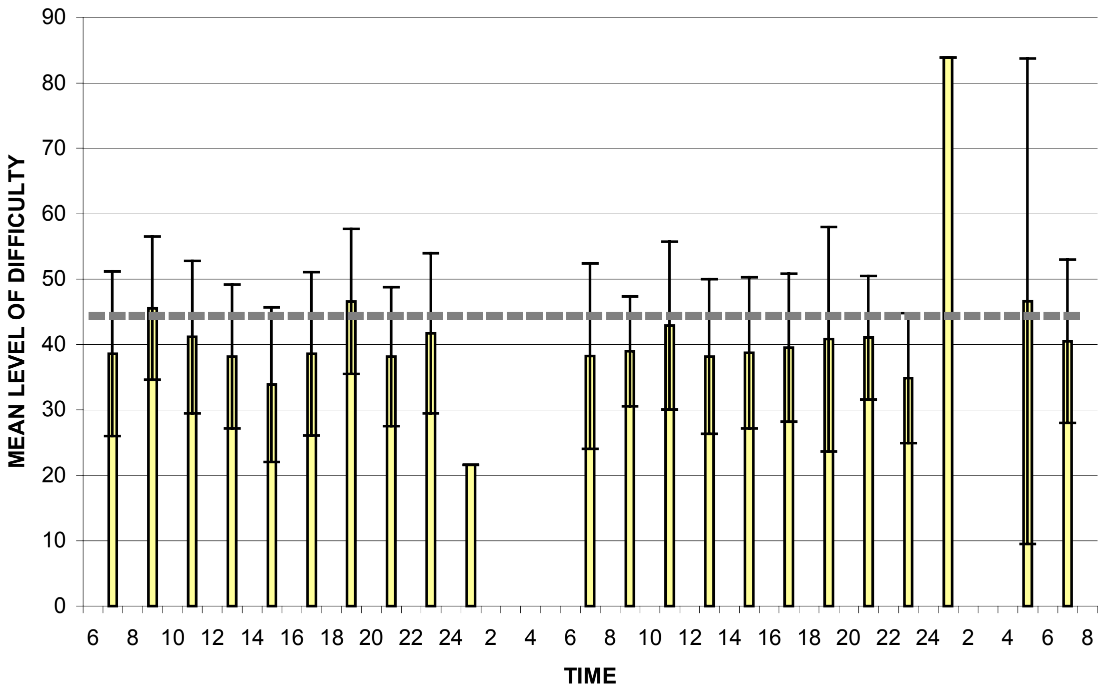


Tracking task

There was very little change in Tracking performance as shown by the level of task difficult shown across the trip (see Table 23). The overall main effect for time of test was not significant ($F_{(4,11)}$=0.08,n.s.) and within trip there was no evidence of any significant changes between test occasions (beginning and end of day 1 ($F_{(1,14)}$=0.03,n.s.) beginning and end of day 2 ($F_{(1,14)}$=0.12,n.s.), between days 1 and 2 ($F_{(1,14)}$=0.22,n.s.) and between days 2 and 3 ($F_{(1,14)}$=0.05,n.s.).

Using the 0.05% BAC as a comparison standard, performance across most of the simulated trip was poorer than the standard level (see Figure 4).

FIGURE 4: Unstable Tracking performance with 95% confidence intervals where higher performance indicates better performance. Dotted line indicates performance at 0.05%BAC equivalence level.



Dual Task

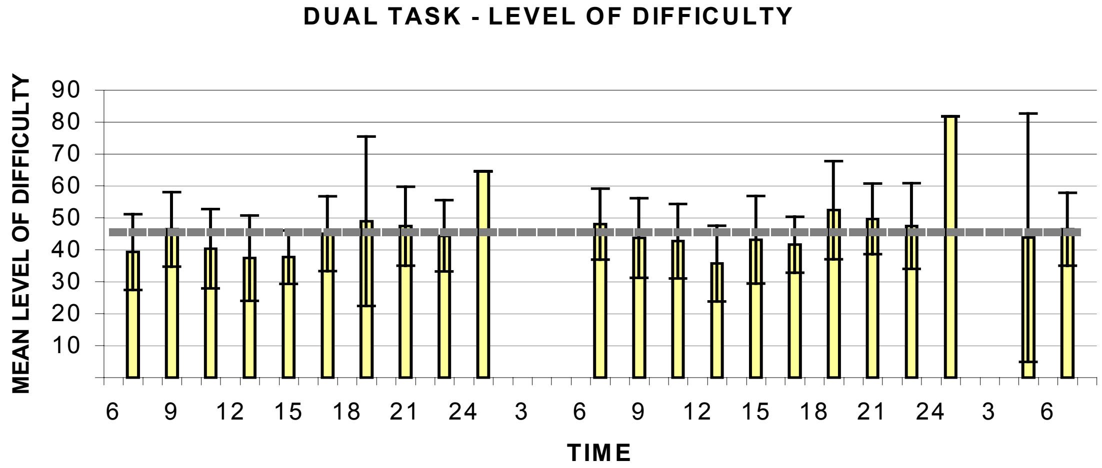
There was little change in any of the measures in the Dual Task across the trip (Table 23). Reaction time measures showed no overall main effect of time of test ($F_{(4,9)}$=1.41,n.s.) and there were no significant changes within trip (beginning and end of day 1 ($F_{(1,12)}$=0.02,n.s.) beginning and end of day 2 ($F_{(1,12)}$=0.05,n.s.), between days 1 and 2 ($F_{(1,12)}$=0.14,n.s.) and between days 2 and 3 ($F_{(1,12)}$=3.8,n.s.). Variability of reaction time showed similar results with no significant overall changes over time ($F_{(4,9)}$=0.20,n.s.). Post hoc contrasts of within trip changes also showed no differences between measurement occasions (beginning and end of day 1 ($F_{(1,12)}$=0.06,n.s.) beginning and end of day 2 ($F_{(1,12)}$=0.02,n.s.), between days 1 and 2 ($F_{(1,12)}$=0.007,n.s.) and between days 2 and 3 ($F_{(1,12)}$=0.81,n.s.). The number of missed signals also showed no significant overall changes over time ($F_{(4,11)}$=1.45,n.s.) with again no significant within trip changes (beginning and end of day 1 ($F_{(1,14)}$=2.30,n.s.) beginning and end of day 2 ($F_{(1,14)}$=1.00,n.s.), between days 1 and 2 ($F_{(1,14)}$=2.15,n.s.) and between days 2 and 3 ($F_{(1,14)}$=0.19,n.s.). Finally, there was no significant overall trend in the level of difficulty achieved in the tracking component of this task over test occasions across the trip ($F_{(4,11)}$=1.71,n.s.) and no significant changes within the trip (beginning and end of day 1 ($F_{(1,14)}$=1.28,n.s.) beginning and end of day 2 ($F_{(1,14)}$=0.18,n.s.), between days 1 and 2 ($F_{(1,14)}$=0.13,n.s.) and between days 2 and 3 ($F_{(1,14)}$=1.18,n.s.).

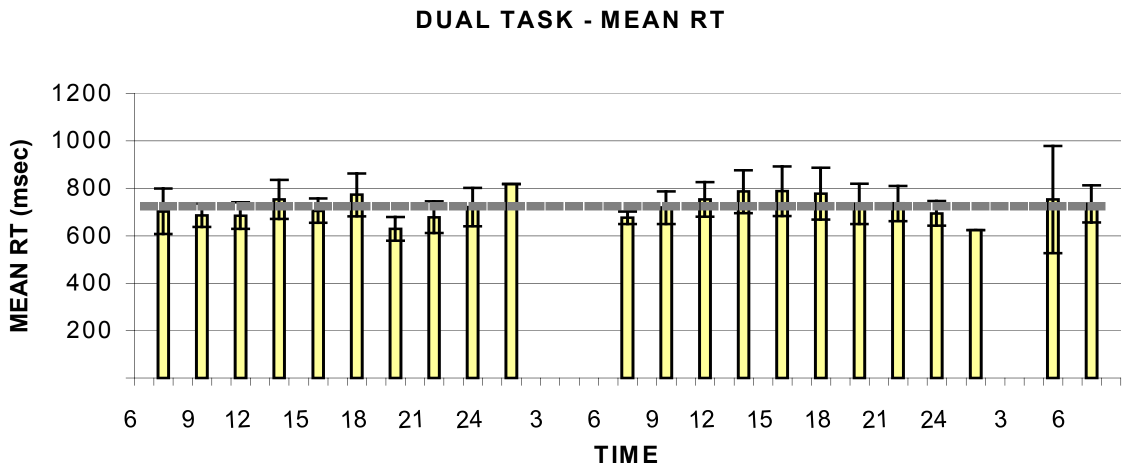
For this test, reaction speed was around the level of the 0.05% BAC standard throughout most of the trip and tended to become even slower in the afternoons of both day one and day 2 (see Figure 5). Similarly, for the tracking component of the test, performance was very close to the 0.05% BAC standard for most of the trip. Poorer than standard performance tended to occur in the early afternoon of both test days for this measure as well

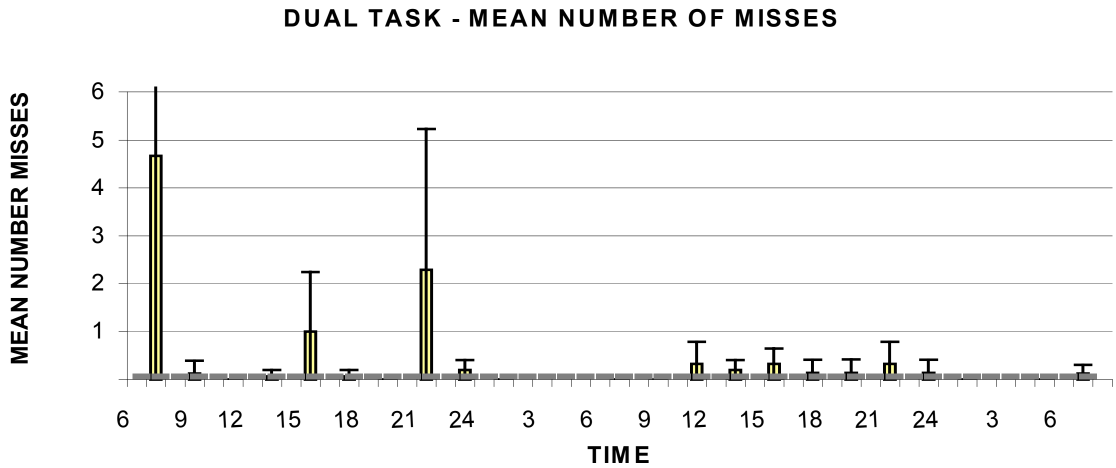
Symbol Digit Task

In this task there are two measures that reflect speed of responding, reaction time and, because this is a time limited test, the number of symbols. Both tests showed similar results (Table 23). There was a significant overall main effect of time of test

FIGURE 5: Dual Task measures with 95% confidence intervals. Dotted lines indicates 0.05%BAC equivalence level.







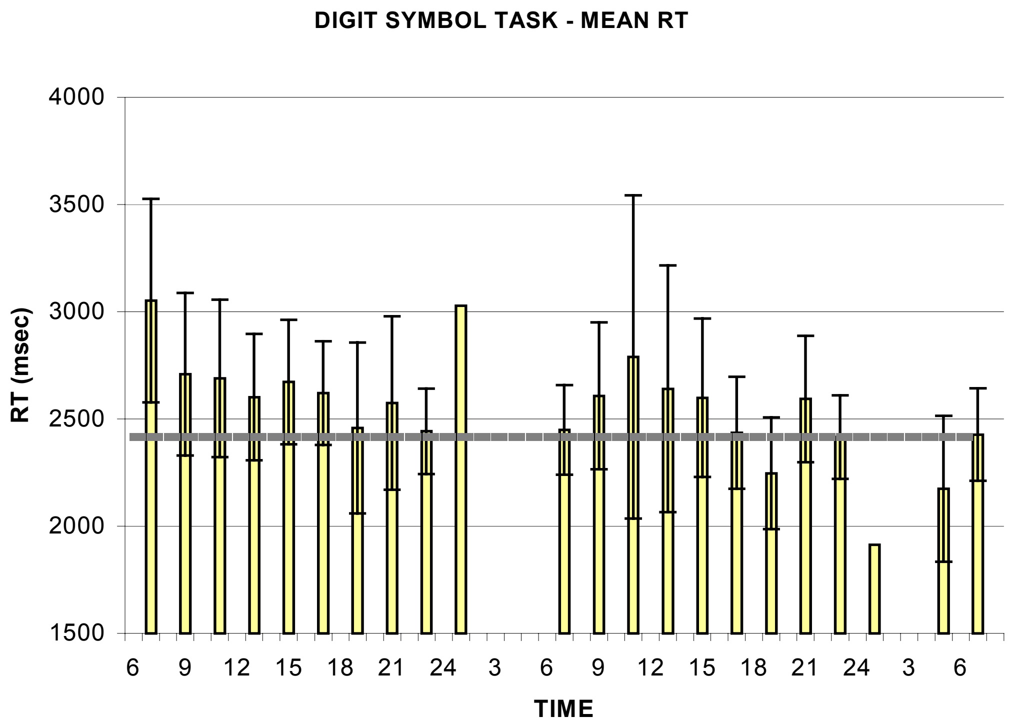
for the reaction time measure, (univariate $F_{(2,30)}$=6.76,p<0.005, although the multivariate test was not significant ($F_{(4,11)}$= 2.45,p<0.11), and for the number of symbols presented ($F_{(4,11)}$= 6.73,p<0.005). For both measures, there was also a significant post hoc contrast showing that at the beginning of day 1 drivers were slower to respond and, as a result, processed fewer symbols compared to their performance at the end of day 1 (for reaction time $F_{(1,14)}$=8.38,p<0,01 and for number of symbols $F_{(1,14)}$=13.42, p<0.003). Performance remained at the same level for the remainder of the study for both measures (between end day 1 and beginning of day 2, reaction time, $F_{(1,14)}$=0.73,n.s., number of symbols, $F_{(1,14)}$=0.30,n.s.; the beginning of day 2 and the end of day 2 reaction time, $F_{(1,14)}$=1.85,n.s.; number of symbols, $F_{(1,14)}$=1,49,n.s., and the end of day 2 and the beginning of day 3, reaction time, $F_{(1,14)}$=0.31,n.s.; number of symbols, $F_{(1,14)}$= 0.10,n.s.).

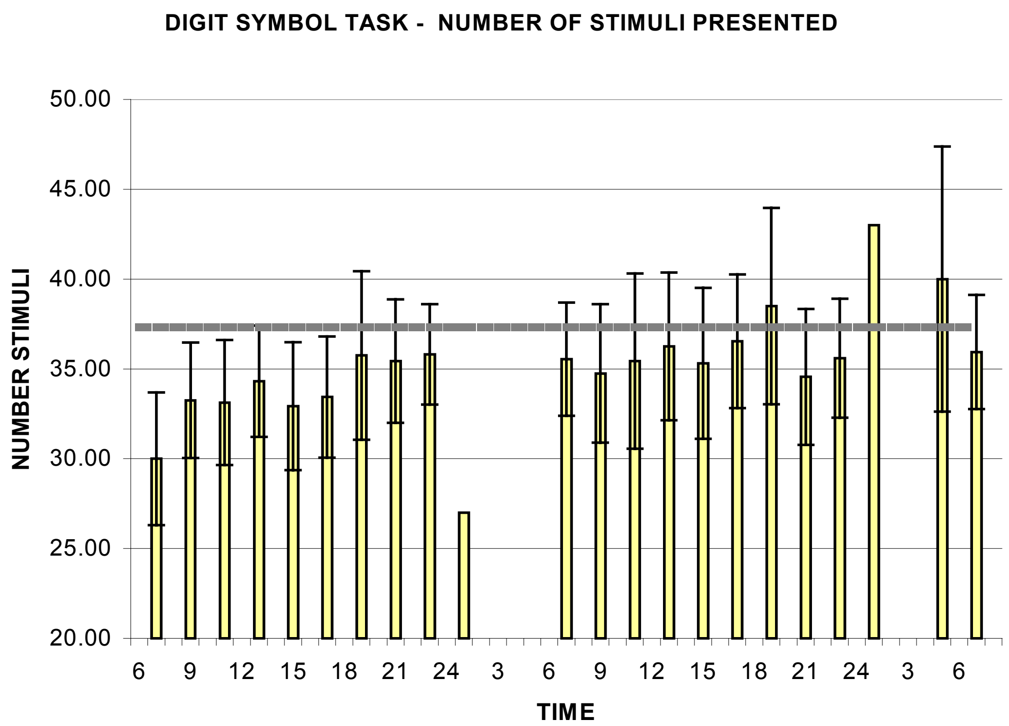
The amount of variability in reaction time showed a similar pattern to these reaction time measures although it was not statistically significant. The overall multivariate test was not significant, ($F_{(4,11)}$= 1.49,n.s.), but the univariate test showed that there was a trend to differences between test occasions ($F_{(2,23)}$= 3.23,0.05<p<0.10). This trend was reflected in the post hoc analysis with a trend for a difference between the
start and end of day 1 ($F_{(1,14)}$= 4.12,p<0.06) and no further change in variability for the next four measurement occasions (end day 1 and start day 2, ($F_{(4,11)}$= 0.13,n.s); across day 2, ($F_{(4,11)}$= 0.31,n.s.) and between end of day 2 and start of day 3, ($F_{(4,11)}$= 1.46,n.s.)

The accuracy measure in this test, percentage of correct responses, showed that performance levels were high across the entire trip. Drivers started the trip making few errors in this test and this did not change over the trip (overall multivariate $F_{(4,11)}$= 0.21,n.s.; and post hoc contrasts, across day 1, $F_{(1,14)}$= 0.15,n.s., between end of day 1 and beginning of day 2, $F_{(,14)}$= 0.04,n.s.; across day 2, $F_{(1,14)}$= 0.43,n.s.; and between end of day 2 and beginning of day 3, $F_{(1,14)}$= 0.52,n.s.).

Comparison of reaction speed in this task with the alcohol standard showed that from the beginning and across the simulated trip drivers performed more slowly than the alcohol standard at almost all measurement occasions (see Figure 6). This could be seen for both measures of reaction speed; reaction time and the number of symbols presented.

FIGURE 6: Symbol Digit measures with 95% confidence intervals. Dotted line indicates 0.05%BAC equivalence level.



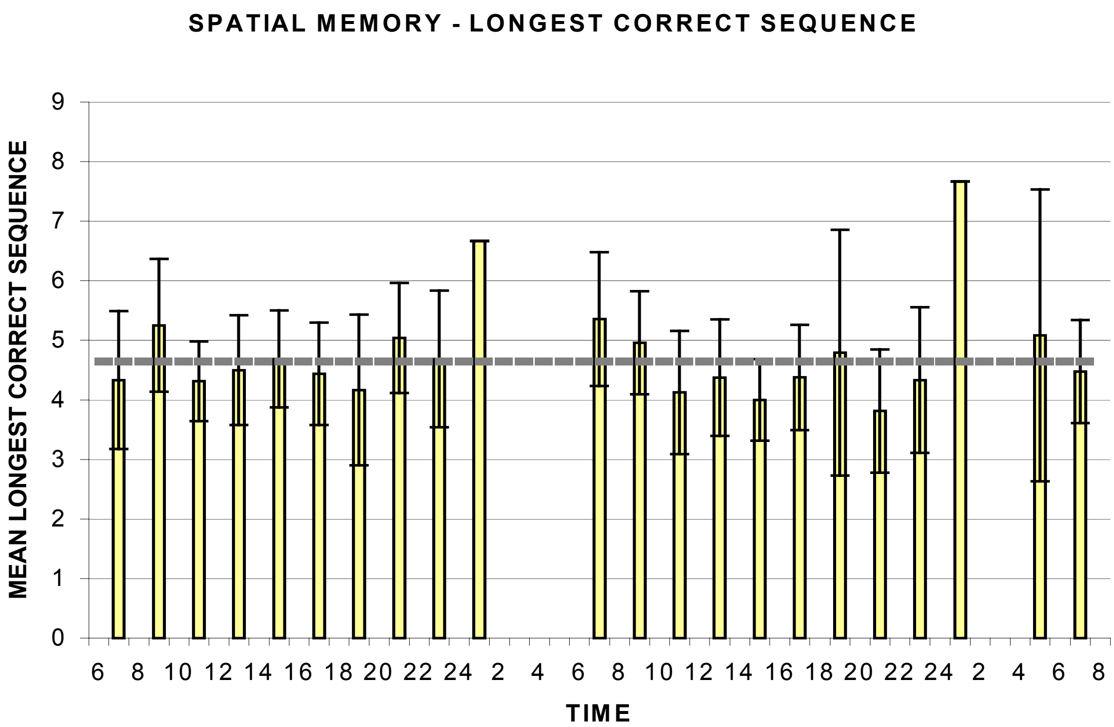


Sequential Spatial Memory

Performance on this task is shown in Table 23. There was no significant effect of time of test on the ability to recall items in the Sequential Spatial Memory test ($F_{(4,10)}$= 1.99,n.s). Post hoc contrasts confirmed this with no significant change across the first three measurement occasion (across day 1 $F_{(1,13)}$= 1.27,n.s. and between the end of day 1 and the beginning of day 2, $F_{(1,13)}$= 0.79,n.s.), however the number of items recalled fell significantly across the second day of the trip ($F_{(1,13)}$= 7.45,$p<0.02$) and remained at the same lower level until the end of the trip, at the beginning of day 3 ($F_{(1,13)}$= 0.76,n.s.).

When the spatial recall was compared to alcohol-standardised performance levels, recall performance was poorer than the 0.05% BAC standard from the beginning of the trip, and was equal to or poorer than this level for most of the trip (Figure 7).

FIGURE 7: Sequential Spatial Memory task with 95% confidence intervals. Dotted line indicates 0.05%BAC equivalence level.



DISCUSSION

Based on the subjective ratings of fatigue, together with the results on a number of the performance tests, it can be concluded that this simulated trip does not allow for adequate fatigue management. Fatigue ratings increased across each day of the trip, but the recovery from fatigue that should occur after a significant block of sleep did not occur after the second sleep period. Previous work with two-up drivers doing very long trips (Feyer, Williamson, & Friswell, 1995) demonstrated that a long sleep, strategically placed within the trip provides drivers with an efficient method for managing fatigue. At the beginning of day 3, drivers were just as tired as they had been at the end of driving on each of the previous two days. Performance test results show similar results. By the middle of the second day drivers were much more likely to miss signals in both the Simple Reaction Time and Mackworth Vigilance tests. They also showed poorer recall in the Sequential Spatial Memory test at this time. In the same way as for fatigue, each of these measures also showed no improvement after the long recovery sleep at the end of day 2. It seems from these results that this trip produced fatigue sufficient to influence performance in a way that would be very likely to affect driving since it particularly interfered with drivers ability to pick up signals. Just as importantly though, the trip did not allow for recovery from fatigue when it occurred. Drivers began day 3 tired and less able to perform. These results demonstrate that after a long trip, a break of only 6 hours is not sufficient to achieve recovery from fatigue.

It is unlikely that chronic fatigue due to the demands of work in the past week played a large role in these findings. The large range of work-rest experience in the past 7 days before the study made it possible to examine the influence of work-rest experiences immediately before the simulated trip. As would be expected, there were lawful or predictable relationships between the amount of work, and the amount and quality of sleep and fatigue at the beginning of the simulated trip. There was also a trend for performance to be better, at least for reaction time, in drivers who had less work in the past week. However, the relationship between fatigue and performance was not as expected, since performance on the Simple Reaction Time and Mackworth Vigilance tests was better at the beginning of the trip for drivers who reported greater fatigue at that time. It is possible that higher or lower fatigue ratings at the beginning of testing only reflect differences between drivers use of the scale, that is some drivers tend to always rate higher than others. This effect disappears for all subsequent ratings across the trip because drivers tend to use these ratings in a comparative way to reflect how they feel now compared to the way they felt for the last rating. In any case, over the trip fatigue ratings and performance followed a very similar pattern with performance deficits only occurring at times when fatigue was highest.

Using the laboratory-developed performance equivalent for 0.05% BAC as a standard was an attempt to aid the interpretation of the changes in performance over the trip. If the standard is useful, performance which is better than the 0.05% BAC performance standard should be judged as safe and performance that is poorer should be regarded as unsafe. Performance for the Mackworth Vigilance task was better than the standard for the first day and a half of the trip, but by mid afternoon on the second day, performance on this test had deteriorated to be poorer than the standard and tended to stay that way for the rest of the trip . This reinforces the earlier analysis and indicates that the changes in the ability to detect signals in this test become of substantive concern for safety in the latter part of the trip.

For a number of tests, performance on the simulated trip was equivalent to or poorer than the standard from the beginning of the trip. In the response speed measures of the Simple Reaction Time test, Dual Task, and Symbol Digit test, in the Tracking measures and in the Sequential Spatial Memory test, drivers started and continued the simulated trip with performance that was poorer than seen in the laboratory when their blood alcohol levels were 0.05% and above.

This is a surprising finding since it suggests that performance on these measures should be judged unsafe from the very beginning of the trip through to the end. This finding cannot be attributed to fatigue however, because it does not mirror the pattern of fatigue across the trip and because fatigue was lowest at the beginning of the trip. The most likely explanation for these findings is that drivers took a strategic approach to the simulated trip. Drivers paced themselves by adopting the strategy of trading off speed for accuracy in the beginning of the trip because they were aware that the trip was a very long one. By being slow and careful in the beginning of the trip they may be better able to cope with the demands of the long trip and to stave off fatigue for longer. Consistent with this explanation, the only task that did not show evidence of poorer performance from the outset, the Mackworth Vigilance test, is also the only task in which it is not possible to trade off speed for accuracy. The primary task in this test is to first detect an infrequent signal, only then is a response required. The pace of signal presentation is not under the control of the subject and speed in responding to the signal does not affect any aspect of the test. Therefore if drivers were taking a conservative approach to managing their performance, it would not affect the Mackworth Vigilance test, but could affect all of the other tests. This was indeed the case.

Additional support for this interpretation comes from the laboratory study (Williamson et al, in press, CR189) where, in the Symbol Digit test in particular, drivers tended to emphasise accuracy rather than speed compared to non-driving controls.

Performance in the tracking task both alone and as part of the Dual Task showed a tendency to be poorer than the standard in the mid to late part of each afternoon as did the reaction time measure of the Dual Task. Analysis of the time of test failed to show any significant differences between test occasions that might be related to fatigue. These small changes in performance are likely to be due to circadian influences that are known to adversely affect performance in the midnight to dawn period and in the mid afternoon (Craig, 1987) rather than responses due to fatigue.

Overall these findings imply that this long trip would not be safe on the road. The simulated trip did not differ greatly from trips allowed under the regulated regime in terms of either length of driving allowed or the proportion of the trip in rest. The simulated trip did differ, though, in that all activity was either work or rest related to overcoming the effects of work. Drivers in this study tended to drive all day with a number of shorter breaks which were at most no longer than the time required to eat a meal, then have a night sleep. If this schedule was effective at managing fatigue, drivers should have shown sufficient recovery after the long night break to be able to drive without significant fatigue or performance effects for the next day. As seen from the results, recovery from the fatigue of the first day was only partial and did not occur at all on the third day. A likely problem with the simulated trip is that, by the time drivers had prepared themselves for sleep and then prepared themselves for work the next morning, the schedule allowed too little time for sleep per se. In addition, lack of recreational rest may also have been a problem. As the simulated trip only allowed for breaks designed to deal with the physiological needs of eating and sleeping, one of its problems may have been that it did not schedule any time for other forms of rest such as changing to another activity. The results imply that the pattern of work and rest should be amended either by allowing longer rest and sleep periods between days on the trip and/or longer break times during the day or by reducing the length of the trip for a single driver.

GENERAL CONCLUSIONS

These studies were important for three reasons. Most notably, these findings provide evidence about the effectiveness of two different work-rest regimes and suggest some general principles for the management of fatigue. The results showed that trips that are conducted under the current working hours regime of up to 14 hours allowable in each trip, do not produce significant levels of fatigue nor significant adverse effects on performance. The results of both studies show, however, that increasing hours of work produce both fatigue and poorer performance, even over the week. Further work is needed to be specific about how many hours work can be done and still remain safe. The simulation study suggests that up to 16 hours of work may be done without producing significant adverse effects on performance, but that such a long trip cannot be sustained for longer than one trip. This finding needs to be replicated on the road as overall performance levels in this simulation were considerably poorer than expected.

These studies also showed the importance of taking into account the effects of accumulation of fatigue over consecutive trips. The simulation revealed that recovery from fatigue is dependent on the accumulated level of fatigue and that recovery may not even occur when too much fatigue has accumulated. The on-road study showed that even at relatively low levels of fatigue, the amount of sleep and the number of breaks determine fatigue and performance. The results imply that management of chronic fatigue will be achieved best by controlling the amount and pattern of rest breaks rather than simply the number of working hours permitted.

The second reason that these findings are important is that the observed performance effects are of relevance to driving. Performance functions, such as reaction speed and the ability to consistently react quickly, and the ability to detect infrequent signals, were affected most by work-rest experiences and fatigue. Decrements in these types of functions are certainly likely to affect driving ability. It is also significant that the same performance effects, increases in the number of missed signals in the Mackworth Vigilance test and more variable reaction speed in the Simple Reaction Time test, were found across work periods in both studies. This suggests that these functions are most sensitive to fatigue.

Lastly, these studies are important because they demonstrate a method for systematically evaluating the effects of different work-rest patterns on fatigue and performance. One of the major problems for management of fatigue in the long distance road transport industry has been a lack of scientific evidence for recommending effective limits for work or patterning of rest. The results of these studies and of the previous laboratory study demonstrate a method that is sensitive enough to pick up effects of variations in the work-rest schedule and variations in fatigue levels. This method will therefore enable informed judgements to be made about how work and rest can be arranged to reduce the problem of driver fatigue. This method can be applied to assess other work-rest schedules, to evaluate their effectiveness for managing fatigue and to develop alternative schedules where a particular schedule is shown to be ineffective.

REFERENCES

Arnold, P.K., Hartley, L.R., Penna, F., Hochstadt, D., Corry, A., & Feyer, A.-M. (1996). *Fatigue in the Western Australian transport industry: Part two - the drivers' perspective.* Perth: Institute of Research, Safety, and Transport, Murdoch University.

Craig, A., Davies, D.R., & Matthews,G. (1987). Diurnal variation, task characteristics, and vigilance performance. *Human Factors, 29:6,* 675-684.

Dinges, D.F., Whitehouse, W.G., Orne, E.C, & Orne, M.T. (1988). The benefits of a nap during prolonged work and wakefulness. *Work and Stress*, *2*, 139-153.

Feyer, A.-M., Williamson, A.M., & Friswell, R. (1995). Strategies to combat fatigue in the long distance road transport industry. Stage 2: Evaluation of two-up operations. *CR 158.* Canberra: Federal Office of Road Safety.

Monk, T.H. (1994). Circadian rhythms in subjective activation, mood, and performance efficiency. In M.H. Kryger, T. Roth, & W.C. Dement (Eds.). *Principles and practice of sleep medicine (2nd ed.)*. Philadelphia: W.B. Saunders. (pp.321-330).

Pilcher, J.J., & Huffcutt, A.I. (1996). Effects of sleep deprivation on performance: A meta-analysis. *Sleep*, *19:4*, 318-326.

Williamson, A.M., Feyer, A.-M., Coumarelos, C., & Jenkins, T. (1992). Strategies to combat fatigue in the long distance road transport industry. Stage 1: The industry perspective. *CR 108.* Canberra: Federal Office of Road Safety.

Williamson, A.M., Feyer, A.-M., Mattick, R.P., Friswell, R., & Finlay-Brown, S. (2000, in press). Developing measures of fatigue: Using an alcohol comparison to validate the effects of fatigue on performance. *Accident Analysis and Prevention.*

Williamson, A.M., Feyer, A.-M., Friswell, R., & Finlay-Brown, S. (in press). Development of measures of fatigue: Using an alcohol comparison to validate the effects of fatigue on performance. *CR 189.* Canberra: Australian Transport Safety Bureau.

APPENDIX 1: CONSENT FORM FOR ON-ROAD EVALUATION STUDY

Participant Information and consent form

ON ROAD EVALUATION STUDY

1998

University of New South Wales, School of Psychology,
New Zealand Occupational and Environmental Health

ON ROAD EVALUATION STUDY

1998

Driver fatigue is a major safety issue in the long distance road transport industry in Australia mainly because of the long distances that have to be travelled. One of the options currently being explored to manage fatigue better is a move away from regulated working hours to more flexible Fatigue Management Programmes (FMPs). An FMP aims to ensure that companies manage all of the risk factors that contribute to heavy vehicle fatigue, including things like rostering systems and the length of shifts expected of drivers. Fatigue Management Programs need to be evaluated in operation to ensure that they really do help to manage driver fatigue. The aim of this study is to determine the effects on fatigue and performance of standard regulated shifts. This information will be used as a baseline against which your company's FMP plan can be compared.

What is involved?

You will be asked to complete 2 short performance tests at the beginning and end of each day's work and at the start of each rest break you take during your trips this week. The tests take about 7 minutes to complete. At the beginning and end of your first day and at the end of your last day this week, we would like you to complete a somewhat longer version of the tests. These take about 20 minutes. The tests have been shown to measure the effect of fatigue in drivers. We would also ask you to keep a brief record of how you feel and what you do during work shifts. Lastly, we will ask you to provide some background information about your lifestyle, health, and recent work and sleep patterns.

All the information you provide will be confidential. In fact, once all the information has been collected, we will not be keeping your name at all.

Your participation in this project is voluntary and you are free to withdraw at any time without penalty or prejudice. Please note that your decision to participate will have no bearing on your employment and your personal results will not be shown to your employer.

If you have any questions about the study please do not hesitate to contact Ann Williamson, Samantha Brown or Rena Friswell on (02) 9385 3806.

If you wish to complain about any aspect of the conduct of this research project please contact Mrs Margaret Wright, Executive Officer, Ethics Secretariat, University of New South Wales on (02) 9385 4234.

ON ROAD EVALUATION STUDY

1998

You are invited to participate in the evaluation of the effects on fatigue and performance of standard regulated shifts. If you wish to participate, please complete the consent form below.

Consent Form

I, ____, agree to participate in the evaluation of the effects on fatigue and performance of standard regulated shift.

I acknowledge that I have read the statement above outlining the study, and that the statement has been explained to my satisfaction. I have been given the opportunity to ask any questions relating to any possible physical or mental harm I might suffer as a result of my participation, and have received satisfactory answers.

I understand the information that I provide will be strictly confidential, and that only the study's research team will have access to information that identifies me with my responses.

I also understand that I am free to withdraw my consent and stop my participation at any time without prejudice.

____________________________ ____________________________
(Signature of participant) (Signature of witness)

__________________ __________________
(Date) (Date)

APPENDIX 2: BACKGROUND QUESTIONNAIRE FOR ON-ROAD EVALUATION STUDY

Code Number:

PARTICIPANT

BACKGROUND INFORMATION

ON ROAD EVALUATION STUDY

1998

Fatigue Management Survey

As part of our research on the best ways to manage fatigue in the long distance road transport industry, we need to find out about the people participating in the study. In particular we need to collect some general information on your lifestyle, health and work history.

All information you give to us will be CONFIDENTIAL and ANONYMOUS. You will be assigned a code number so that your name will not appear on any of your results.

On the following pages there are some questions about these matters that we would appreciate you filling in as carefully as possible.

THANK YOU FOR YOUR HELP

Today's date: _____________

Current time: _____________ am / pm ?

1. What is your: Age: *(Please tick)*

< 20 years	( )
20 – 29 years	( )
30 - 39 years	( )
40 – 49 years	( )
50 –59 years	( )
60 or more years	( )

Sex: *(Please circle)* M F

2. Are you:

	Please tick
married or living in a defacto relationship?	( )
widowed, separated or divorced?	( )
single?	( )

3. How long have you been driving a truck for a living? _________ years

4. How far did you continue with formal education? *(Please tick)*

To Primary school level	( )
To High school Year 7, 8, 9, or 10 level	( )
To High school Year 11 or 12 level	( )
To Tafe level	( )
To College or University level	( )

5. How much experience have you had using personal computers?

	Please tick
None	( )
A little	( )
Frequent user	( )

6. Do you suffer any of the following health problems? *(Please circle)*

Diabetes	Yes	No
Asthma/Hayfever	Yes	No
Sleep disorders eg sleep apnea	Yes	No
Stomach or digestive problems	Yes	No
Liver or kidney problems	Yes	No
Heart or circulation problems eg angina, high blood pressure	Yes	No
Headaches or migraines	Yes	No

7. Do you smoke cigarettes?

No ( )

Given up ( )

How long ago did you give up? _______________ years/months

Yes ( )

How many do you smoke on average per day? ___________ cigarettes

8. Do you drink caffeinated drinks? Yes ( )

No ( )

If YES, what sorts of caffeinated drinks do you *usually* consume?

____

How many of these drinks do you have on average per day?

____

9. How often do you usually drink alcohol? *(Please tick)*

Every day	( )
2-3 times a week	( )
Once a week	( )
1-2 times a month	( )
Rarely	( )
Never	( )

If you do drink alcohol, how many standard drinks do you usually drink at one time? (*Please tick)*

One drink	( )
2-3 drinks	( )
4-5 drinks	( )
more than 5 drinks	( )

1 drink =	***1 middy beer or*** ***1 glass wine or*** ***1 nip spirits***
1 can beer =	***1.5 drinks***

10. When you are sleeping, how often do you:

		Please tick
Snore loudly ?	always	( )
	often	( )
	sometimes	( )
	rarely	( )
	never	( )
Stop breathing ?	always	( )
	often	( )
	sometimes	( )
	rarely	( )
	never	( )
Move around a lot ?	always	( )
	often	( )
	sometimes	( )
	rarely	( )
	never	( )

11. Do you have difficulty getting to sleep ? Yes ( )

No ( )

12. Do you have difficulty staying asleep once you are asleep ?

Yes ( )

No ( )

13. Do you have difficulty preventing yourself from falling asleep during the day ?

always ( )
often ( )
sometimes ( )
rarely ( )
never ( )

14. Have you had your adenoids removed ? Yes ( )

No ( )

Please continue
over page

15. How likely are you to DOZE OFF OR FALL ASLEEP, in contrast to just feeling tired, in the following situations?

These situations refer to your usual way of life in recent times. Even if you have not done some of these things recently try to work out how they would have affected you.

Use the following scale to choose the MOST APPROPRIATE NUMBER for indicating how likely it is you would have dozed off in each situation

0	Would never doze
1	Slight chance of dozing
2	Moderate chance of dozing
3	High chance of dozing

Situation	Chance of Dozing
Sitting and reading	_____
Watching TV	_____
Sitting inactive in a public place (eg. In a movie theatre or at a meeting)	_____
As a passenger in a car for an hour without a break	_____
Lying down to rest in the afternoon when circumstances permit	_____
Sitting and talking to someone	_____
Sitting quietly after a lunch without alcohol	_____
In a car, while stopped for a few minutes in traffic	_____

16. Do you usually work: *(Please tick)*

    Mostly days? (6:00 to 18:00) ( )

    Mostly nights (18:00 to 6:00)? ( )

    Days and nights about equally? ( )

17. Was the last week a typical working week for you? Yes ( )

    No ( )

    If No, what was unusual about it? *(e.g., on holidays, sick, on light duties etc)*

    ____

18. In the last 7 days (not counting today):

    How many hours did you work? __________ hours

    How many of these were at night (i.e. 18:00 to 6:00)? __________ hours

    How many trips did you drive? __________ trips

    How long was your *last trip* in terms of:

    Kilometres? __________ km

    Total duration? __________ hours

    Hours spent working __________ hours

19. When did your last work shift end?

    Time: ____________ *am/pm* *Day:* ___________ *Date:* ___________

20. In total, how much sleep have you had since then? __________ hours

21. How long was your last *substantial* sleep ? __________ hours

22. When did this sleep end?

Time: ____________ *am/pm* *Day:* ___________ *Date:* ___________

23. How would you rate the quality of this sleep?
(Please draw a cross at the point which most closely describes the quality of your sleep)

Very poor quality |————————————————| *Very good quality*

24. How did you feel when you awoke from this sleep?

(Please draw a cross at the point which most closely describes how refreshed you felt)

Not at all refreshed |————————————————| *Very refreshed*

25. Have you had any naps since your last substantial sleep ?

Yes ( )

No ( ) If No, go to question 28

If Yes, please record the length of the nap and the time you woke up in the table below. *(If you have had more than 3 naps, please record the others on the back of this page.)*

	Length of nap	*End of nap*	
	hours : minutes	date	time
Nap 1	:		am/pm
Nap 2	:		am/pm
Nap 3	:		am/pm

26. How would you rate the quality of your *last* nap ?

(Please draw a cross at the point which most closely describes the quality of your sleep)

Very poor quality |————————————————| *Very good quality*

27. How did you feel when you awoke from your *last* nap ?

(Please draw a cross at the point which most closely describes how refreshed you felt)

Not at all refreshed |————————————————| *Very refreshed*

28. When did you last eat a meal?

Time: ____________ *am/pm* *Day:* ___________ *Date:* ___________

Was this meal *(Please tick)*:

Light ( )

Moderate ( )

Large ( )

Have you snacked since then?

Yes ( )

No ( )

29. If applicable, when did you last have a caffeinated drink? *(eg. Coffee, tea, coke)*

Time: ____________ *am/pm* *Day:* ___________ *Date:* ___________

How many caffeinated drinks have you had today? ___________ drinks

30. If applicable, when did you last have an alcoholic drink?

Time: ____________ *am/pm* *Day:* ___________ *Date:* ___________

How many alcoholic drinks did you have on that occasion?

___________ drinks

31. Are you currently taking any medication? Yes ( )

No ( )

If Yes, what? ____

APPENDIX 3: EXCERPTS FROM ON-ROAD DIARY FOR ON-ROAD EVALUATION STUDY

Code:

ON ROAD DIARY

ON ROAD EVALUATION STUDY

1998

Instructions

On the following pages, we would like you to record when you take each break from work over the next 6 days, and to rate how tired you feel at the beginning and end of these breaks. This includes rest and meal breaks, overnight breaks or breaks between trips. It is important that you try to make the ratings at the time of the break. There is one page for each break.

To make the ratings, simply put a mark somewhere on each scale line to show how you feel. For example, on the scale of hungriness below, if you were only a bit hungry you might put the mark as shown

Very hungry *Not at all hungry*

|————————————————————|

Please complete the ratings on this page, for the start of the week, now. The last page contains similar ratings to be completed at the end of the week.

Start of the week

Work start time: __________ am or pm Date: ____________

Start time of first trip: __________ am or pm Date: ____________

Please rate how you feel now on the following scales

Fresh *Tired*

|————————————————————|

Clear-headed *Muzzy-headed*

|————————————————————|

Very alert *Very drowsy*

|————————————————————|

BREAK 1

BREAK START time: __________ am or pm Date: _________

Have you done the handheld tests for the start of this break?
(Please tick) Yes ( )
No ( )

Since your last break, how long have you spent:

Driving? ___________ hours

Doing other work? ___________ hours

Please rate how you feel now on the following scales

Fresh *Tired*

Clear-headed *Muzzy-headed*

Very alert *Very drowsy*

BREAK END time: __________ am or pm Date: _________

Please rate how you feel now on the following scales?

Fresh *Tired*

Clear-headed *Muzzy-headed*

Very alert *Very drowsy*

Did you sleep during this break? Yes ( )
No ( )

If Yes,

How long was your sleep? ______________ hours

How would you rate the quality of this sleep?

Very poor quality *Very good quality*

How did you feel when you woke?

Not at all Refreshed *Very Refreshed*

BREAK 30

BREAK START time: __________ am or pm Date: _________

Have you done the handheld tests for the start of this break?
(Please tick) Yes ( )
No ( )

Since your last break, how long have you spent:

Driving? ___________ hours

Doing other work? ___________ hours

Please rate how you feel now on the following scales

Fresh *Tired*

Clear-headed *Muzzy-headed*

Very alert *Very drowsy*

BREAK END time: __________ am or pm Date: _________

Please rate how you feel now on the following scales?

Fresh *Tired*

Clear-headed *Muzzy-headed*

Very alert *Very drowsy*

Did you sleep during this break? Yes ( )
No ( )

If Yes,

How long was your sleep? ______________ hours

How would you rate the quality of this sleep?

Very poor quality *Very good quality*

How did you feel when you woke?

Not at all Refreshed *Very Refreshed*

End of week

Please rate how you feel now on the following scales

Fresh ——————————————————— *Tired*

Clear-headed ——————————————————— *Muzzy-headed*

Very alert ——————————————————— *Very drowsy*

How many trips did you drive this week? ________ trips

End time of last trip: _________ am or pm date: ________

Work end time: _________ am or pm date: ________

How would you describe your workload this week? *(please tick)*

Much less than usual ( )
Less than usual ( )
About the usual level ( )
Greater than usual ( )
Much greater than usual ( )

Overall, how would you describe your fatigue levels this week? *(please tick)*

Much less than usual ( )
Less than usual ( )
About usual ( )
Greater than usual ( )
Much greater than usual ( )

Do you have any other comments that you would like to make about your work or your fatigue this week?

____

____

____

____

____

____

____

Many thanks for your participation

APPENDIX 4: PALMTOP COMPUTER INSTRUCTIONS FOR ON-ROAD EVALUATION STUDY

Instructions for using handheld testers

(Please refer to diagrams of the handheld tester on the back of this page.)

1. Plug the keypad into the handheld tester, with the bump on the plug pointing up.

2. Press the ON button on the handheld tester.

   If the tester is ready for use, the last line of writing on the screen should be: E:\PIPS>

3. Use the keyboard on the handheld tester to type:

   TEST then leave a space by pressing the long 'space' key at the bottom of the keyboard, and then type your subject code as shown on the front of your On Road Diary.

   Make sure there is a space between the word TEST and your subject code.

   If you make a mistake while typing, use the 'backspace' key to erase the problem and then re-type.

   If you do not press any keys on the handheld tester for several minutes, the screen will go blank. If this happens, just press the ON button to bring the screen back to where you left it

4. Press the ENTER key on the keypad.

   If a "bad command" message appears on the screen followed by the E:\PIPS> message, simply redo step 3.

5. The tester will now ask you whether you have changed the batteries today.

   Press the 1 or 2 keys on the keypad to answer no or yes.

   If you have not changed the batteries, the program will ask you to turn the tester off and to replace the batteries. See instructions below and diagram over the page for changing batteries. Once you have changed the batteries, you should start from step 2 again.

   If you have already changed the batteries, the program will start the tests. Simply follow the instructions on the screen

6. When the tests are finished the E:\PIPS> message will return.

7. Press the ON button on the handheld tester again to switch the machine OFF.

8. Unplug the keypad from the tester.

Changing the batteries.

1. Make sure that the handheld tester is switched OFF and the keypad is unplugged.

2. Turn the machine upside down and push the battery cover off.

3. Replace the old AA batteries with new ones.

4. Slide the battery cover back on, taking care that it closes properly

5. Please keep the used batteries in the bag provided

If you have any problems using the tester, call Sam Brown, or Rena Friswell, or Ann Williamson on 02 9385 3806 or 0414 772 114.

APPENDIX 5: CONSENT FORM FOR SIMULATION STUDY

Participant Information and consent form

LONG SHIFT SIMULATION STUDY

WAGGA JUNE-JULY 1998

University of New South Wales, School of Psychology,
New Zealand Occupational and Environmental Health

LONG SHIFT SIMULATION STUDY

WAGGA JUNE-JULY 1998

Driver fatigue is a major safety issue in the long distance road transport industry in Australia mainly because of the long distances that have to be travelled. One of the options currently being explored to manage fatigue better is a move away from regulated working hours to more flexible Fatigue Management Programmes (FMPs). An FMP aims to ensure that companies manage all of the risk factors that contribute to heavy vehicle fatigue, including things like rostering systems and the length of shifts expected of drivers. Fatigue Management Programs need to be evaluated in operation to ensure that they really do help to manage driver fatigue. The aim of this study is to determine the level of fatigue and the effect s on performance of an extended (48 hour) shift when drivers are permitted to pattern their work and rest as they would under real operational conditions.

What is involved?

You will be asked to complete a set of performance tests every two hours, when you are not sleeping, during the 48 hours of long shift simulation. The tests have been shown to measure the effect of fatigue in drivers. We will also ask you to keep a brief record of how you feel and what you do during work shifts. Lastly, we will ask you to provide some background information about your recent work and sleep patterns.

All the information you provide will be confidential. In fact, once all the information has been collected, we will not be keeping your name at all.

We will arrange with your employer to make you available for these test days at normal rates of pay. Your participation in this project is voluntary and you are free to withdraw at any time without penalty or prejudice. Please note that your decision to participate will have no bearing on your employment and your personal results will not be shown to your employer.

If you have any questions about the study please do not hesitate to contact Ann Williamson, Samantha Brown or Rena Friswell on (02) 9385 3806.

If you wish to complain about any aspect of the conduct of this research project please contact Mrs Margaret Wright, Executive Officer, Ethics Secretariat, University of New South Wales on (02) 9385 4234.

LONG SHIFT SIMULATION STUDY

WAGGA JUNE-JULY 1998

You are invited to participate in the evaluation of the level of fatigue and the effect s on performance of an extended (48 hour) shift when drivers are permitted to pattern their work and rest as they would under real operational conditions. If you wish to participate, please complete the consent form below.

Consent Form

I, ____agree to participate in the evaluation of level of fatigue and the effect s on performance of an extended (48 hour) shift when drivers are permitted to pattern their work and rest as they would under real operational conditions.
I acknowledge that I have read the statement above outlining the study, and that the statement has been explained to my satisfaction. I have been given the opportunity to ask any questions relating to any possible physical or mental harm I might suffer as a result of my participation, and have received satisfactory answers.

I understand the information that I provide will be strictly confidential, and that only the study's research team will have access to information that identifies me with my responses.

I also understand that I am free to withdraw my consent and stop my participation at any time without prejudice.

____________________________	____________________________
(Signature of participant)	(Signature of witness)
__________________	__________________
(Date)	(Date)

APPENDIX 6: BACKGROUND QUESTIONNAIRE FOR SIMULATION STUDY

Code Number:

PARTICIPANT

BACKGROUND INFORMATION

Long shift simulation study

June 1998

Fatigue Management Survey

As part of our research on the best ways to manage fatigue in the long distance road transport industry, we need to find out your work and rest patterns in the last few days.

All information you give to us will be CONFIDENTIAL and ANONYMOUS. You will be assigned a code number so that your name will not appear on any of your results.

We would appreciate you filling in the following questions as carefully as possible.

THANK YOU FOR YOUR HELP

How many hours have you worked in the last 7 days ? __________ hours

How many of these were at night
(i.e., between 6 pm and 6 am) ? __________ hours

When did your last shift end ? time __________ am/pm

day ________________ date ______________

How long was your last shift ? __________ hours

When did you last have an alcoholic drink? ________________________

Are you currently taking any medication? Yes ( ) No ( )

If YES, what ? ____

What have you eaten today ?

____

How long ago did you last eat ? ____________________________

Have you had any drinks containing caffeine today ? eg. Coffee, tea, coke

Yes ( )

No ( )

If YES, What did you have ? ________________________

How much did you drink ? ____________________________

APPENDIX 7

COMPARISON OF LAPTOP AND PALMTOP TESTS

One of the problems for testing performance on the road is that it is difficult to find suitable measures. Safety considerations limit testing while the driver is operating the vehicle and there are also practical problems in having researchers administering tests at regular intervals in a long trip. One solution is to develop small portable test devices that can be used by the drivers to test themselves at regular intervals in the trip.

Palmtop computer versions were developed for a selection of the tests used in the earlier laboratory study. Tests were chosen based on the results of the laboratory study regarding their sensitivity for fatigue and on how readily they could be adapted for administration on a palmtop computer.

This analysis was an attempt to validate the results of performance tests conducted on-road using the palmtop computer by comparing them with the results of performance tests conducted off the road using a laptop computer. The palmtops were being assessed for their suitability as on-road testing devices for later field studies. It was necessary therefore to choose tests that balanced the clearest effects due to fatigue in the laboratory study with the shortest administration time. Two performance tests were chosen for presentation on the palmtops – the Simple Reaction Time task and a shortened version of the Mackworth Vigilance task.

The aim of this section of the project was to compare the results of the laptop and palmtop versions of the same tests to determine whether the palmtop versions could be used reliably in the place of the laptop versions

METHOD

The subjects for this comparison were the 16 drivers who participated in the simulation study. The simulation was chosen for this analysis since it was possible to have drivers do laptop and palmtop tests at approximately the same time under supervision. For this comparison the laptop and palmtop tests occurring within one hour of each other were paired. For most drivers this involved between 10 and 14 pairs of scores. Paired T-tests were then used to compare the scores from each type of test for each subject. The results of the T-tests were then compared across subjects.

RESULTS

Baseline subjective fatigue ratings

Table 24 compares baseline fatigue ratings made on the laptop and palmtop computers. Baseline ratings made on the laptops are an average of pre and post test ratings, whereas the palmtop ratings occurred before the tests only. Despite this, the two sets of ratings are very similar in terms of both average fatigue levels and intersubject variability.

TABLE 24: Baseline fatigue ratings (/100) on laptop and palmtop computers

	Laptop	Palmtop	Test result
Averaged fatigue			
• Mean	32.53	35.94	t(15)=1.34, ns
• *SD*	*19.98*	*19.19*	

Comparison of performance using the laptop and palmtop computers

Simple Reaction Time

Comparison between subjects of the results of the T-tests showed that for all subjects, except one, palmtop reaction time results were significantly slower than laptop result and variability of reaction time results was higher for the palmtop test for three-quarters of drivers (Table 25).

TABLE 25: Comparisons of performance on the laptop and palmtop tests during the simulation study.

	Mean Laptop – Palmtop difference		% subjects showing significant difference			
			Lap > Palm		Lap < Palm	
	Mean	SD	p<0.05	p<0.10	p<0.05	p<0.10
SRT						
• RT	-104.4	57.3	0.0	0.0	93.8	93.8
• SD	-27.3	20.6	6.3	6.3	68.8	75.0
• # missed	-1.8	1.6	0.0	0.0	44.0	56.0
MACKWORTH						
• RT	-39.9	190.8	12.5	12.5	6.3	6.3
• SD	45.8	124.6	12.5	18.8	0.0	0.0
• # missed	-2.8	7.6	0.0	0.0	6.3	6.3
• # false alarms	-3.8	8.2	0.0	0.0	18.8	31.3

Analysis of the differences between laptop and palmtop versions of the Simple Reaction Time tests showed a consistent trend towards slower and more variable reaction time on the palmtop computer. On average the reaction time was 101.4msecs (s.d.=57.3) slower for the palmtop version with an average increase in variability between responses of 27.3msecs. Similarly, accuracy was higher on the laptop test.

Mackworth Vigilance test

Comparison across subjects of individual T-tests for each of the measures on the Mackworth Vigilance test showed no consistent pattern of results. For the reaction time and variability of reaction time, only three drivers showed differences between palmtop and laptop test and the differences were not consistently in the same direction. For the number of missed signals, only one driver showed a difference between palmtop and laptop tests. For the false alarm measure, most drivers showed no difference between palmtop and laptop tests, but for a significant minority there was a considerably greater number of false alarms with the palmtop test compared to the laptop test.

DISCUSSION

For most measures there were no systematic differences between palmtop and laptop versions of the same tests. The major exception was the Simple Reaction Time test for which performance was considerably slower in the palmtop version. Overall the extent of slowing was around 20 percent. The variability of reaction speed also differed for the two versions. The visual display of the palmtop version of this test was more difficult to see than on the laptop version since the screen was so much smaller. It is not surprising then, that reaction speed was slower and more variable for the palmtop version.

Performance on both versions of the Mackworth Vigilance test was very similar. There was little difference between the reaction speed and variability measures and the number of missed responses between the two versions of the test. The number of false alarms was somewhat higher for the false alarms measure of this test. Possible reasons for this difference are that in the palmtop version the display was more difficult to see which resulted in drivers responding erroneously more often when there were no signals.

In conclusion, it seems that the main differences between palmtop and laptop versions lie in the Simple Reaction Time test. While the overall magnitude of the difference in reaction speed is not inconsiderable, it is not likely to have a major effect on the overall sensitivity of the test to fatigue. If anything, because both of the palmtop tests are slightly more difficult to perform because of the smaller display, it is likely that the palmtop version will be more sensitive to the onset of fatigue. Drivers are more likely to make errors and false alarms when the signal is harder to discern.