Pergamon

Accid. Anal. and Prev., Vol. 28, No. 6, pp. 709–719, 1996

0001-4575/96 $15.00 + 0.00

PII: S0001-4575(96)00044-9

THE IMPACT OF WORK PRACTICES ON FATIGUE IN LONG DISTANCE TRUCK DRIVERS*

ANN M. WILLIAMSON, ANNE-MARIE FEYER and RENA FRISWELL†

National Institute of Occupational Health and Safety, GPO Box 58, Sydney, 2001, Australia

(Received 23 May 1996)

Abstract—Twenty seven professional truck drivers completed a 12 hour, 900 km trip under each of three driving regimes—a relay (staged) trip, a working hours regulated one-way (single) trip, and a one-way (flexible) trip with no working hours constraints. The results indicated that none of the driving regimes prevented fatigue and that the pattern of fatigue experienced during the trips appeared to be related to pretrip fatigue levels. Copyright © 1996 Elsevier Science Ltd

Keywords—Driver fatigue, Work practices, Driving regime

INTRODUCTION

Driver fatigue and its measurement

Driver fatigue is usually defined as a state of reduced physical or mental alertness which impairs performance of a range of cognitive and psychomotor tasks including driving (e.g. Moore-Ede et al. 1988; Haworth and Heffernan 1989). It occurs commonly among professional long distance truck drivers (Williamson et al. 1992) and can have serious consequences, both for the fatigued driver and for other road users (e.g. see Haworth et al. 1988, 1989). As well as increasing the likelihood that drivers will fall asleep at the wheel, fatigue impairs the ability to maintain lateral road position and to maintain constant speed (e.g. Brookhuis and de Waard 1993; Mackie and Miller 1978; Ranney and Gawron 1987).

Among those professional drivers who are highly skilled at the driving task, fatigue may be quite severe before routine driving performance is noticeably affected (Dinges and Graeber 1989; Lisper et al. 1973). At lesser levels of fatigue, however, decreases in physiological arousal, slowed sensorimotor functions and impaired information processing can diminish a driver's ability to respond effectively to unusual, unexpected, or emergency situations (Babkoff et al. 1988; Dinges and Graeber 1989; Monk and Folkard 1985; Mascord and Heath 1992; Rosa and Colligen 1992).

Therefore, to measure the impact of fatigue on drivers, researchers need to utilize not only direct indices of driving (such as steering control and speed maintenance), but also tests which tap the component perceptual, motor, and cognitive skills underlying driving performance. Despite being ambiguous measures of fatigue when used in isolation (Mascord and Heath 1992; Riemersma et al. 1977), measures of physiological arousal, such as heart rate, can provide converging evidence for the development of fatigue. Typically, heart rate has been found to slow with fatigue and prolonged driving, and the time between heart beats becomes increasingly variable (e.g. Aasman et al. 1987; Brookhuis and de Waard 1993; Fagerstrom and Lisper 1977). Because it is the subjective experience of fatigue that will, ultimately, determine whether fatigue reduction strategies are used, any assessment of occupational fatigue should also incorporate subjective measures.

Work practices and fatigue management

Fatigue increases with time spent driving (e.g. Harris and Mackie 1972) and is heightened during the midnight to dawn hours compared with other times of day. The dawn hours themselves (4–6 a.m.) are a period of particular vulnerability (Feyer et al. 1993; Hamelin 1987; Moore-Ede et al. 1988; Van Ouwerkerk 1987; Williamson et al. 1992). In addition, driving is a fairly monotonous task which, nonetheless, requires sustained vigilant monitoring. Such tasks are known to be fatiguing (Davies et al. 1983; Monk and Folkard 1985). Chronic partial sleep deprivation (CPSD) and acute sleep deprivation also contribute to driver fatigue. CPSD is a cumulative sleep decrement resulting from ongoing, suboptimal

*This paper was presented at the 7th World Conference on Transport Research, Sydney, Australia, 16-21 July 1995.
†Author for correspondence.

amounts of sleep (Monk and Folkard 1985; Tepas and Mahan 1989). It is engendered by the sort of prolonged and irregular hours that are typical of work practices in the long distance road transport industry (e.g. Hensher et al. 1991; Williamson et al. 1992).

Different work practices vary in the extent to which they incorporate factors that contribute to fatigue, and also in the opportunities they provide for drivers to remedy fatigue. As a result, some work practices may act to promote or combat fatigue more than others. Williamson et al. (1992) examined the differential impact of staged, two-up and single driving regimes on fatigue in truck drivers. Staged driving occurs when drivers from different points of origin rendezvous en-route and exchange trucks or loads. Both drivers then return with their new cargoes to their respective starting points. In the two-up regime, two drivers share the driving of a single truck, and in the single driving regime, one driver completes an entire trip alone. Williamson et al. (1992) found that staged drivers were less likely to have experienced fatigue on their last trip, were less likely to have experienced fatigue on at least half of their trips, and were less likely to perceive fatigue as a personal problem than either two-up or single drivers. However, when staged drivers did experience fatigue its onset tended to be somewhat earlier than that reported by the other drivers.

Williamson et al. (1992) suggested that three factors, trip flexibility, trip start time, and trip length, might account for the apparent contradiction between the lower incidence and faster onset of fatigue in staged drivers. The staged driving regime affords drivers less flexibility to schedule rest breaks to counteract early onset fatigue than do other regimes. Staged drivers were also more likely than other drivers to start their trips at night, so that the time of peak vulnerability to fatigue (4–6 a.m.) would have fallen earlier in their trips. On the issue of trip length, the staged drivers surveyed by Williamson et al. (1992) typically completed shorter trips than either single or two-up drivers, which might explain the lower overall incidence of fatigue reported by staged drivers.

The aim of the current study, then, was to investigate more closely the relationship between staged driving and fatigue. More specifically, the study sought to examine the impact of the structural features of a twelve hour, 900 km trip on fatigue. Three different regimes, which were equated for length and start time, were compared—a staged trip and a single trip, both of which involved regulated working hours, and a maximally flexible regime, where drivers were exempted from the regulations governing hours of driving.

METHOD

Design

The major highway linking Australia's two most populated cities (Sydney and Melbourne) was selected as the standard route for the study. This road spans approx. 900 km, crosses 2 states, and is predominantly double-laned dual-carriageway through rural countryside.

A repeated measures design was used to assess the effects of type of driving regime on fatigue. Each driver was asked to complete 3 trips between Sydney and Melbourne as part of his normal work roster—one trip under each driving regime. The staged trip entailed driving from Sydney or Melbourne to the trip midpoint (Tarcutta), exchanging trucks or loads with a driver coming in the opposite direction, and then returning to the point of origin. The performance of both drivers in each staged pair was measured. The single one-way trips involved driving directly from Sydney to Melbourne, and the flexible one-way trips involved driving from Melbourne to Sydney. The refusal of authorities in one state to provide exemptions from the hours of driving regulations meant that the direction of flexible, and therefore single trips, could not be counterbalanced. Under the regulations, drivers on single and staged trips were obliged to break for 30 minutes after each five hour period. Under the flexible regime drivers could choose to take breaks as often or as rarely as they needed with no constraint on the time taken to complete the trip.

Scheduling of trips was determined by the company rosters. As a result, the staged trip was generally the first of the three. However, the order of single and flexible trips was fully counterbalanced.

Subjects

Subjects were 27 professional truck drivers drawn from two companies which routinely ran staged operations between Sydney and Melbourne in Australia. One company provided 19 of the drivers and the other provided 8. Fourteen drivers originated in Melbourne and the others lived in Sydney. All were men and drove staged trips between Sydney and Melbourne on a regular basis, but only 7 drove staged trips exclusively. On average, the drivers were 38.4 years old and had driven trucks commercially for an average of 15.9 years. The drivers appeared to have relatively healthy lifestyles. Smokers were in the minority (26.1%), and although most drivers drank alcohol (82.6%) they did so infrequently (73.7% drank once a week or less) and in moderation (52.6% drank 3 drinks or less per occasion). The majority (63.7%) exercised at least twice a week. None of the drivers reported symptoms indicative of sleep disorders (e.g. sleep apnoea).

Equipment, measurement and procedure

Full technical details may be found in Williamson et al. (1994).

Fatigue was indexed in a variety of ways. Physiological arousal was assessed continuously throughout the trip using a measure of heart rate—the interval between heart beats (interbeat interval). Two measures of driving performance (speed and steering variability) were also recorded continuously across the trip. Continuous on-board recording was achieved using a custom-built data logger. The sensors measuring heart rate, steering variability and speed fed directly into the logger which stored the data and provided a measure of time for each data sample. At the end of a trip, the data and time information were downloaded to a PC. The on-board recording apparatus was designed to be portable so that each driver could be measured in his regular truck.

Four tests of information processing capacity were chosen to tap component skills of the driving task. These measured perceptual sensitivity (critical flicker fusion task), simple manual reaction time, the capacity for sustained attention on a monotonous task (vigilance task), and the ability to respond to an unpredictable stimulus (unstable tracking task). The tests were conducted off-road, at the start, the midpoint, and the end of the trips using a portable, computer-assisted test battery developed by the senior authors. Midtrip tests were not conducted on flexible trips because to do so would have compromised the drivers' absolute freedom to schedule breaks.

Although the drivers were not available for extensive practice sessions before the study, each test administration was preceded by practice trials. The tests themselves were conducted in the following order:

(a) The critical flicker fusion task tested drivers' ability to detect the onset (descending version) or cessation (ascending version) of flickering in a small light. The higher the flicker frequency at detection, the more alert the driver. Four practice trials were given, and the order of ascending and descending trials was counterbalanced across subjects.
(b) The vigilance task required the driver to extinguish lights by pressing appropriate buttons. 600 lights were presented sequentially over a 10 minute period. Slowed reaction times and increased error rates reflect decreased attentiveness. Thirty practice trials were given. The length of the task made it impractical at midtrip testing.
(c) The simple manual reaction time task required drivers to respond as rapidly as possible to a visual signal. Two measures were available for this task. The time taken to initiate the response, referred to here as 'decision time', incorporates the perceptual and cognitive processes involved in the decision to respond. The time taken, subsequently, to execute the response reflects motor speed, and is denoted 'movement time'. Slowed times indicate reduced alertness. Ten practice trials preceded this task.
(d) The unstable tracking task tested the driver's ability to control the increasingly variable movements of a visual stimulus using a response dial. In this task, the amount of variability of stimulus movement was directly related to the changes in performance capacity of the driver. As the driver improved, the stimulus became more variable. The length of time that the driver was able to control the stimulus, and the level of stimulus instability (the level of difficulty) attained were measured. Alertness is associated with a longer time in control, and higher difficulty levels. Five practice trials were given.

In addition, drivers were asked to rate their subjective fatigue at the beginning and end of each trip and at the beginning of each break from driving, using the seven point Stanford Sleepiness Scale (Hoddes et al. 1973) and three visual analogue scales (fresh–tired, clear headed–muzzy headed, and very alert–very drowsy). Drivers carried a trip diary with them in the truck for this purpose, and were also asked to keep an ongoing record of the time of each break and the activities undertaken.

Additional questionnaires were used to collect information from the drivers about their work and rest activities in the week prior to the study, and about their general state of health. This information allowed the impact of pretrip factors on fatigue to be assessed.

On the day of departure, a member of the research team set up the on-board measuring equipment on the relevant truck, and met the driver an hour before his scheduled departure time. On his first trip, the driver provided background information on his recent work history and health, and the use of the heart rate, trip diaries and other on-board equipment was explained. The driver completed the pretrip off-road testing session and departed. On arrival at Tarcutta (trip midpoint), the driver was met by another member of the research team, who administered the midtrip off-road testing session. The driver was met again on arrival at his destination and the posttrip off-road testing was completed. The researcher collected the driver's trip diary and stripped the equipment from the truck.

RESULTS AND DISCUSSION

A more complete analysis of the results of this study was reported in Williamson et al. (1994).

Missing data

Missing data arose as a result of occasional technical problems, and because one pair of drivers declined to be measured on their staged trip. Consequently, repeated measures statistical analyses were based upon reduced sample sizes and the number of missing data points varied for different measures. Because there was no appreciable difference between the data of the total available sample and those of the reduced samples, descriptive statistics have been presented for the total number of available subjects on each measure.

Trip starting time and length

All of the observed trips took place overnight. Table 1 confirms that most trips began in the early evening and night, between 16:00 and 23:59, and took approx. 12 hours to complete. Importantly, the three driving regimes did not differ significantly in starting time, and although, on average, staged trips took longer to complete than flexible trips [$F(2,17)=3.94$, $p=0.04$], the trip lengths differed by only 40 minutes.

Scheduling of breaks

The number of breaks taken (Table 1) increased across flexible, single and staged trips, suggesting an increasing need for rest as a function of driving regime. However, breaks were taken after similar periods of driving for the three trip types (Table 1). The longest drive period (4.5 hours) routinely occurred before the first break and the shortest drive period (2.5–3 hours) preceded the second break. The minor variations in the length of the second and third drive periods suggest that, in the middle of the trip drivers on staged regimes were tiring most rapidly, whereas toward the end of the trip drivers on flexible regimes were tiring most rapidly. Taken together, these data offer some evidence that across trip types the drivers were tiring most quickly in the middle of their trips, but this effect was strongest under the staged regime. Average break lengths were also similar across trip types, varying at most by approx. 30 minutes. Examination of the minor variations in break length revealed little change in the lengths of the breaks taken across the staged trip, but a tendency for flexible trips to include a shorter second break, and for single trips to include a relatively short third break. Thus, drivers appeared to require less recovery time in the middle of flexible trips and toward the end of the single trips than at other times.

In summary, the pattern of break-taking suggests increased fatigue around the middle of the trips and greater overall fatigue on staged trips. However, these conclusions need to be treated cautiously because break taking may have occurred for reasons other than providing a rest from driving, and at times of

Table 1. Descriptive statistics on trip timing and the scheduling of breaks

Measures		Trip type: Staged	Single	Flexible
Start time (%)				
	16:00–19:59	44.0	44.0	37.0
	20:00–23:59	56.0	44.0	48.0
Mean (SD) start time (h:min)		19:55 (1:43)	20:22 (2:23)	20:34 (2:35)
Mean (SD) trip length (h:min)		12:36 (0:55)	12:10 (1:12)	11:59 (0:40)
Number of breaks taken (%)	0	0	0	7.4
	1	29.2	32.0	33.3
	2	20.8	44.0	48.1
	3	45.8	24.0	11.1
	4	4.2	0	0
Mean (SD) number of breaks		2.3 (0.9)	1.9 (0.8)	1.6 (0.8)
Mean break length (h:min)				
	Break 1	00:38	00:44	00:33
	Break 2	00:40	00:50	00:22
	Break 3	00:37	00:17	00:40
Mean length of drive periods (h:min)				
	Drive 1	04:23	04:32	04:30
	Drive 2	02:29	02:48	02:52
	Drive 3	03:20	03:45	03:02

convenience, necessity or habit. In addition, drivers on single and staged trips were asked to stop at midtrip for a cognitive testing session. In contrast to the staged regime, the single regime did not otherwise demand this break. Consequently, the present data may overestimate break-taking on single trips and, to the extent that the break data reflect fatigue, underestimate the alertness of drivers on single trips compared to drivers on staged trips. The subjective ratings provide more direct measures of fatigue.

Subjective fatigue

Both the Stanford Sleepiness Scale ratings and the visual analogue scale ratings (Table 2) revealed a marked increase in fatigue between the beginnings and ends of the trips for all trip types. Drivers tended to feel most fatigued on staged trips and least fatigued on single trips, however, this pattern was in evidence before driving commenced and at posttrip had not been modified by the intervening driving regime.

Drivers reported comparable increases in fatigue between the start of the trip and the first break under all three regimes. At the first break, neither staged nor flexible fatigue levels differed from single trip fatigue. On flexible trips drivers reported higher fatigue at break 2 than break 1. Fatigue at break 2, was also higher on flexible trips than on single and staged trips. At break 3, the numbers of drivers with single and flexible data were too small for meaningful analysis.

Off-road testing

Vigilance (Table 3) was only performed during the pre and posttrip testing sessions. To reduce any impact of 'warm-up' effects on performance, only the last 300 trials were used for the main analyses. Reaction times for all trips decreased between pre and posttrip suggesting increased alertness, however, the frequency of errors did not change. Reaction times were slower and errors more numerous for staged trips compared to single trips at both measurement sessions. In addition, drivers made more errors before and after single trips compared to flexible trips but there was no analogous difference in reaction time. The improved performance at posttrip compared with pretrip, and on single trips compared with staged trips may simply have reflected the amount of practice acquired on the task as a result of the order of testing, and the unbalanced order of trip types. However, no evidence of practice effects was found when the first 300 and the last 300 trials within the 10 minute tests were compared, arguing against an interpretation of the session and trip type results in terms of practice. The fact that performance accuracy differed between flexible trips and single trips, despite their fully counterbalanced order, also suggests that practice *per se* is not sufficient to account for the vigilance results.

The critical flicker fusion task (Table 3) showed mixed results. On the descending version of the task, drivers' performance was not significantly affected by

Table 2. Analysis of fatigue ratings

Measures			Trip type: Staged	Single	Flexible	Significance
Mean (SD) Stanford Sleepiness Scale						
	Pretrip		0.73 (1.03)	0.46 (0.67)	0.73 (0.77)	Trip type $F(2,42)=5.33$ $p<0.009$
	Posttrip		2.09 (1.15)	1.23 (1.07)	1.77 (0.92)	Time in trip $F(1,21)=62.14$ $p<0.001$
Stanford Sleepiness Scale (%)						
	Pretrip	Ratings 1 and 2	79.2	91.6	84.6	
		Ratings 3 and 4	16.7	8.3	15.4	
		Ratings 5 and 6	4.2	0	0	
	Break 1	Ratings 1 and 2	60.9	75.0	66.7	
		Ratings 3 and 4	30.4	16.6	14.3	
		Ratings 5 and 6	8.6	8.4	19.1	
	Break 2	Ratings 1 and 2	71.5	50.0	61.6	
		Ratings 3 and 4	21.4	50.0	15.4	
		Ratings 5 and 6	7.1	0	23.1	
	Posttrip	Ratings 1 and 2	29.2	52.0	40.7	
		Ratings 3 and 4	62.5	48.0	51.8	
		Ratings 5 and 6	8.4	0	7.4	
Mean (SD) combined visual analogue scale						
	Pretrip		21.7 (23.4)	11.7 (9.1)	15.5 (12.8)	Trip type $F(2,20)=6.22$ $p<0.008$
	Break 1		34.2 (24.0)	26.5 (21.3)	29.1 (23.9)	Pre vs post trip $F(1,21)=94.34$ $p<0.001$
	Break 2		26.4 (17.2)	3 2.9 (19.7)	40.1 (29.6)	
	Posttrip		48.2 (18.7)	33.9 (16.6)	41.2 (16.8)	

Table 3. Off-road test results, showing statistically significant effects

			Trip type			
			Staged	Single	Flexible	Significance
Vigilance						
	Mean (SD) reaction time (msec)					
		Pretrip	568.5 (55.5)	518.3 (47.9)	527.6 (84.9)	Trip type $F(2,10)=5.08$ $p<0.05$
		Posttrip	543.8 (47.5)	512.3 (69.0)	515.9 (54.5)	Time in trip $F(1,11)=16.97$ $p<0.002$
	Mean (SD) number of errors					
		Pretrip	7.5 (4.4)	5.1 (3.7)	3.3 (2.8)	Trip type $F(2,22)=3.88$ $p<0.04$
		Posttrip	6.7 (7.7)	4.0 (6.1)	4.2 (3.7)	
Ascending critical flicker fusion						
	Mean (SD) thresholds (hz)					
		Pretrip	38.5 (5.7)	41.3 (5.0)	44.0 (3.3)	Trip type $F(1,10)=4.16$ $p<0.07$
		Midtrip	39.6 (5.2)	39.7 (3.9)	—	Trip type by time in trip $F(2,9)=4.12$ $p<0.05$
		Posttrip	40.1 (5.7)	40.7 (4.9)	43.7 (3.4)	
Unstable tracking						
	Mean (SD) time on task (sec)					
		Pretrip	10.0 (2.6)	10.5 (2.0)	10.6 (2.1)	Time in trip $F(2,13)=4.38$ $p<0.04$
		Midtrip	8.6 (2.1)	10.0 (1.4)	—	Trip type by time in trip $F(1,14)=3.98$ $p<0.07$
		Posttrip	9.9 (2.8)	11.5 (3.3)	10.0 (2.4)	
	Mean (SD) level of difficulty					
		Pretrip	5.1 (1.0)	5.5 (0.8)	5.4 (0.9)	Trip type $F(1,14)=5.61$ $p<0.03$
		Midtrip	4.2 (0.9)	4.8 (0.6)	—	Time in trip $F(2,14)=12.38$ $p<0.001$
		Posttrip	5.2 (1.0)	5.7 (1.3)	5.1 (1.1)	

either the type of driving regime or the point in the trip at which the test was administered. Conversely, on the ascending version of the test, drivers' sensitivity to visual change was lower before the staged trip than before the single trip. However, sensitivity at mid and posttrip did not differ between the two trip types. On single trips, drivers demonstrated a drop in sensitivity at midtrip followed by recovery of sensitivity at posttrip, whereas on staged trips, drivers showed a linear increase in sensitivity across the trip. Flexible trips yielded similar performance to single trips. These data suggest that drivers were more fatigued at the start of staged trips than at the start of the other trips, but their level of alertness improved over the trip. On single trips alertness appeared to dip at midtrip, a finding not consistent with the operation of practice effects. The fact that only ascending task performance varied between trip types and measurement occasions also argues against an explanation of the staged results in terms of practice. When ascending and descending trials were combined so that trial order could be examined directly, a weak 'warm-up' effect [$F(1,10)=4.89$, $p=0.05$] was observed within each individual staged and single test session, but there was no systematic improvement across test sessions. Nor did trial order produce performance differences between the single and staged trip types, suggesting that practice alone cannot account for the pattern of ascending task results.

On the unstable tracking task (Table 3), drivers again showed a dip in performance at midtrip. This was true for both staged and single trips and for both the time on task and difficulty measures. Single drivers also improved their time on task at posttrip relative to pretrip. On staged trips, drivers consistently performed at lower levels of task difficulty than on single trips. Single and flexible trips did not differ. Once again, an explanation of the results in terms of practice effects was contraindicated by the midtrip performance slump. In addition, although staged trips consistently produced the poorest performance, within-session 'warm-up' effects, suggestive of practice effects, were only observed in the pretrip testing session, whereas on single trips 'warm-up' effects were evident at both pre and posttrip [time on task $F(1,14)=11.72$, $p=0.004$; task difficulty $F(1,14)=4.23$, $p=0.06$]. This result does not support a central role for practice effects in the staged trip data. Rather, it suggests that drivers were less able to capitalize on task experience during the staged trip than during the later single trips, implicating fatigue as a factor on staged trips.

When the last half of the simple reaction time trials were analysed, no effects of either trip type or the point in the trip at which the test was conducted were evident. Because practice on this task may have obscured an opposing effect of fatigue, evidence for practice effects was sought. Comparison of the first and last halves of the trials revealed no within-session practice effects on the movement time measure, and weak evidence for a within-session 'warm-up' effect in decision time on staged trips compared to single trips [$F(1,11)=4.44$, $p=0.06$] was counterbalanced by a trend for decision times on staged trips to become slower across test sessions [$F(1,11)=4.08$, $p=0.07$].

In summary, drivers' performance on the vigilance, critical flicker fusion and unstable tracking tasks suggested heightened fatigue at the start of the staged trip, and for vigilance and unstable tracking, this impaired performance under the staged regime was maintained across the course of the trip. These findings are consistent with the higher fatigue ratings given by drivers before and after the staged trip. The possible confounding role of practice is of concern whenever cognitive tests are conducted. In the present study, effort was made to minimize or isolate practice effects by providing practice trials before every testing session, by basing analyses on the last half of the test trials in any session, and by examining the extent and pattern of within-session 'warm-up' effects and between-session changes in performance. For each of the cognitive tests used, there appear to be good reasons for rejecting practice as the main explanation of the observed results.

Both critical flicker fusion and unstable tracking tasks revealed performance improvements at the end of the trip relative to midtrip. The apparent recovery of test performance at the end of the trips is not consistent with the drivers' subjective fatigue levels at post trip. It is likely, however, that arriving at the final trip destination is an arousing experience particularly when it involves negotiating city streets and traffic. Although such rearousal effects would be temporary they would most probably have been in effect during the off-road testing session which was completed immediately upon the drivers arrival. Other researchers have also observed performance recovery effects occurring in anticipation of the end of a long experiment (e.g. Babkoff et al. 1985; Riemersma et al. 1977). Alternatively, performance recovery at post trip might be a function of the time of day at which testing was conducted, such that the early morning midtrip test (mean time = 1:54 a.m.) showed the worst performance. However, such an account would also predict poor performance at the posttrip testing session which typically occurred during the dawn period (mean time = 6:05 a.m.), an effect which was not observed.

Physiological arousal

Figures 1 and 2 summarize heart measures during the second and third hours of the trip and the second and third last hours of the trip for each trip type. The first and final hours were not used because they spanned periods of city driving. The intercept values reflect the magnitude of the interbeat intervals (Fig. 1) and their variability (Fig. 2) at the start of each period, whereas the slope values indicate the extent and direction of change within each period.

It is believed that heart activity and alertness are related, such that increased alertness is associated with shorter and less variable interbeat intervals (e.g. Brookhuis and de Waard 1993; Mascord and Heath 1992). On this basis, the data suggest that drivers were less alert at the beginning of flexible trips than at the beginning of staged and control trips, but that early in the trip, alertness was decreasing most quickly on staged trips. Across the flexible trip, drivers' alertness improved to a level above that attained at the end of staged and single trips. Conversely, staged and single regimes produced a drop in alertness over the trip. Within the 2 hour period at the end of the trip, variability in heart activity decreased slowly on all trip types, suggestive of increasing alertness. However, flexible regimes witnessed a very rapid rise in the magnitude of interbeat intervals over this period, suggesting that despite a higher level of alertness the drivers were tiring more quickly at the end of the flexible trip than they did toward the end of single and staged trips.

Driving performance

The deviations between successive steering wheel angles were calculated to yield a measure of steering angle variability for analysis. Figure 3 shows the magnitude of steering wheel deviations (intercepts) at the start of the second hour and the third last hour of the trip. The slope values indicate the rate of change in steering deviations across the subsequent two hour periods.

Because fatigue is thought to be reflected in larger steering variations (e.g. Brookhuis and de Waard 1993), these data suggest that drivers were more alert toward the start of flexible trips compared to single and staged trips, but under all regimes drivers' alertness levels were increasing over the first few hours. Furthermore, staged and single trips produced an increase in alertness over the course of the trip, whereas alertness seemed to drop over flexible trips. Interestingly, although drivers on flexible trips appeared most fatigued at the third last hour of the trip, their levels of alertness were increasing very rapidly over the two hour period which followed. Overall, the steering data are unusual amongst the

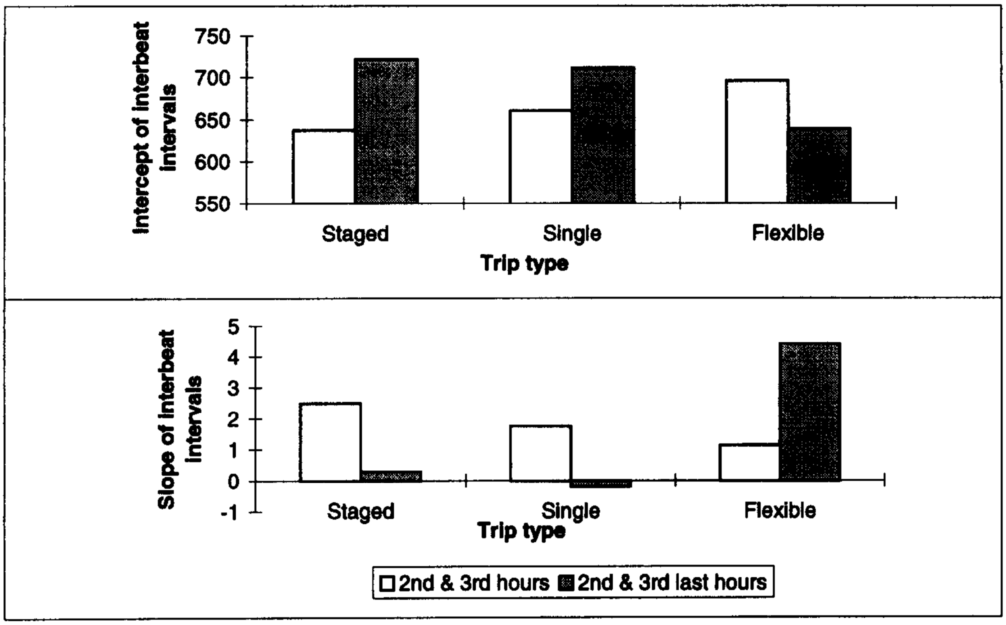

Fig. 1. Slopes and intercepts (msec) of the lines of best fit for interbeat interval over the 2nd and 3rd hours (start) and the 2nd and 3rd last hours (end) of each trip type.

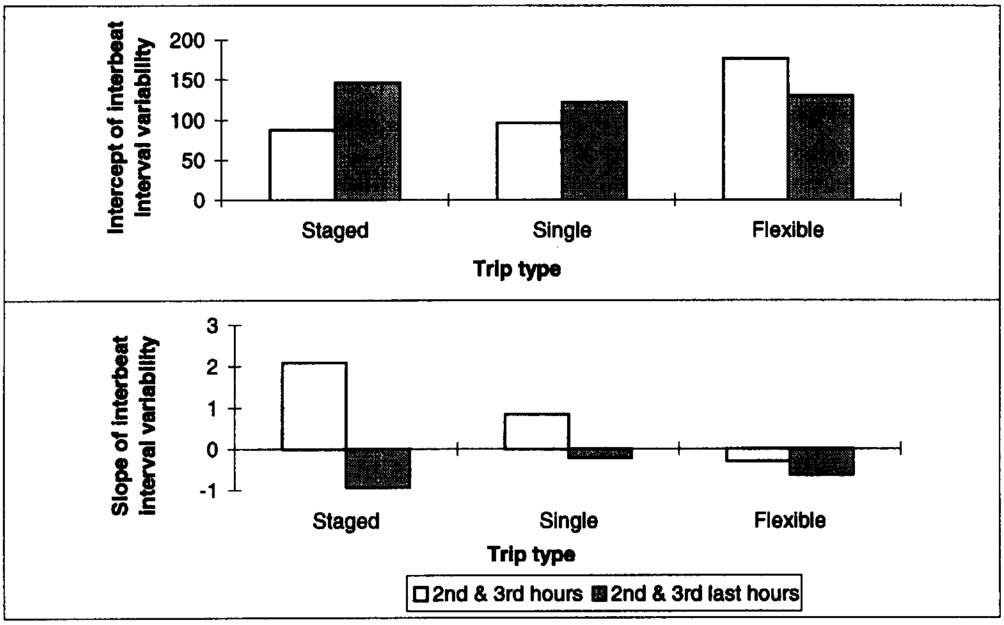

Fig. 2. Slopes and intercepts (msec) of the lines of best fit for interbeat interval variability over the 2nd and 3rd hours (start) and the 2nd and 3rd last hours (end) of each trip type.

measures taken in that they do not show poorer performance under the staged regime than under the single regime.

Speed data showed very little variation across driving regimes or times in the trip, except that driving was faster at the start of flexible trips than at other times. To the extent that speed reflects fatigue, this finding concurs with other measures, particularly steering, that drivers were more alert at the beginning of flexible trips than at other times. It should be noted that the trucks employed in the study were fitted with devices which warned drivers when they reached the legal speed limit. The stability of speeds over time and trips may be explained, then, if drivers simply maintained near-limit speeds whenever possible.

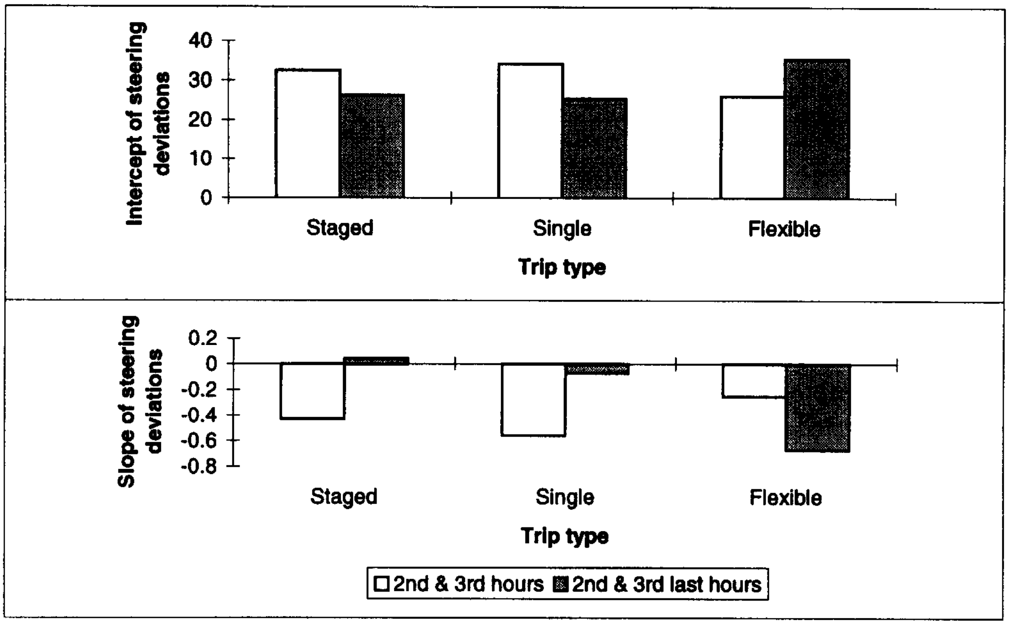

Fig. 3. Slopes and intercepts for the lines of best fit for steering wheel deviations over the 2nd and 3rd hours (start) and the 2nd and 3rd last hours (end) of each trip type. Intercepts are in shaft encoder units (SEUs), where SEU/19 = degrees.

The physiological and driving performance measures are difficult to interpret because they do not present a consistent picture of the drivers' levels of alertness. For staged and control trips, heart activity, like the subjective ratings, indicated increasing fatigue over the trip, whereas steering deviations, suggested the opposite. Similarly, for flexible trips, heart activity indicated decreasing fatigue over the trip, unlike the subjective fatigue ratings, whereas steering deviations suggested an increase in fatigue. These results are further complicated by the speed and direction of changes within the early and later hours of the trip. Clearly, the measures of steering and speed reflect a combination of factors, including not only fatigue, but also the geometry of the road, the characteristics of the freight load, and the handling ease of the truck. In this way, these measures may be less reliable predictors of fatigue than more direct indices. This problem is compounded in the current study because the direction of the different trip types was not fully counterbalanced. Indeed, the precise stretch of road traversed during each of the epochs selected for analysis could not be standardized, either across trip types or across drivers within each trip type.

Interpretation of the cardiovascular data is further complicated because heart activity can show a marked initial elevation which is unrelated to fatigue, followed by habituation over the first few hours of measurement (e.g. Riemersma et al. 1977). This phenomenon, rather than fatigue, may explain the more rapid decrease in heart rate over the second and third hours of the staged trip compared to the later trip types, where drivers were more familiar with the study protocol. Moreover, heart activity has been shown to reflect mental effort as well as arousal (e.g. Aasman et al. 1987; Backs and Seljos 1994; Roscoe 1992). Thus, it is possible that the effort required to maintain driving performance as fatigue increases paradoxically produces patterns of heart activity more consistent with alertness than fatigue. Despite these ambiguities, both physiological arousal and driving performance suggested that the development of fatigue within a trip may follow a different course under the flexible regime compared with the staged and single regimes.

Over all measures, the pattern of results suggests that fatigue increased across all trip types. Drivers were more fatigued prior to their staged trip compared to other trip types, and may have been somewhat less fatigued early in flexible trips. On a number of measures, the lower level of alertness observed before staged trips was sustained throughout the trip. There is clearly a need to identify those factors affecting the drivers' pretrip fatigue and performance.

Pre trip experiences

Although drivers completed background questionnaires and received instructions in the hour prior to the first, usually staged trip, it seems unlikely that this would be sufficient to significantly increase subjective fatigue prior to the staged trip. The possibility that pretrip fatigue differences were due to acute sleep

Table 4. Descriptive statistics for pretrip sleep

Measures	Trip type: Staged	Single	Flexible
Mean (SD) length of pretrip sleep (h:min)	6:18 (2:41)	6:47 (2:22)	6:03 (1:22)
Mean (SD) interval between sleep and trip start (h:min)	3:52 (2:17)	3:15 (1:28)	2:50 (1:35)

deprivation can be discounted because all trips were preceded by equivalent amounts of sleep (Table 4). On flexible trips, sleep was more recent than on other trips but this finding does not explain the consistent differences in fatigue and performance between single and staged trips.

Due to operational constraints, the staged trip was almost always the first trip completed, and there was often a break of at least 24 hours between the staged and one-way trips. It is possible, then, that the lower initial levels of alertness of the drivers on staged trips were related to chronic partial sleep deprivation (CPSD), and that drivers were able to accrue recovery sleep before their first one-way trip. In line with this hypothesis, drivers averaged only 6.3 hours of sleep per day in the week preceding the staged trip compared to community norms of 7.5 hours (Hyppa et al. 1991). In addition, only half (51.8%) of their sleep occurred during the night hours. These figures suggest that drivers may well have been suffering the effects of CPSD before their staged trip. Unfortunately, no data were available that would confirm whether drivers were able to gain recovery sleep in the interval between the staged and one-way trips. Indeed, in the sleep periods directly preceding the one-way trips drivers' average sleep duration was similar to that habitually obtained prior to the staged trip. Clearly, however, CPSD and the sleep/work context in which a trip occurs may be critical determinants of fatigue, especially on relatively short trips like those studied here. The role of these factors deserves further investigation.

Conclusion

This study employed subjective, physiological, and performance measures to examine the relationship between the characteristics of staged driving and the development of fatigue. Although there was some evidence that fatigue developed differently within the three driving regimes (staged, single, and flexible), the levels of fatigue experienced by drivers increased markedly over all of the trips. None of the regimes demonstrated any overall advantage in combating fatigue compared to the other regimes. Whether these findings would generalize to longer trips, where there is greater potential for flexibility in break-taking to be effective, deserves further investigation. Clearly, however, even relatively short 12-hour trips are tiring, and effective strategies for fatigue reduction need to be identified. In the current study, the pretrip level of fatigue appeared to be an important determinant of later fatigue. This finding raises questions about the ongoing work schedules under which long distance drivers operate, and highlights the need to allow adequate rest and recuperation between trips and between blocks of trips to prevent chronic sleep loss and to reduce fatigue.

Acknowledgements—This project was funded by the Australian Federal Office of Road Safety. The authors wish to acknowledge David Leslie for his contribution to data collection and the compilation of the original report. The views expressed in this article are those of the authors and do not necessarily reflect those of the National Occupational Health and Safety Commission.

REFERENCES

Aasman J.; Mulder G.; Mulder L. Operator effort and the measurement of heart-rate variability. Human Factors 29:161–170; 1987.

Babkoff H.; Thorne D.R.; Sing H.C.; Genser S.G.; Taube S.L.; Hegge F.W. Dynamic changes in work/rest duty cycles in a study of sleep deprivation. Behav. Res. Meth. Instrum. Comput. 17:604–613; 1985.

Babkoff H.; Mikulincer M.; Caspy T.; Kempinski D.; Sing H. The topology of performance curves during 72 hours of sleep loss: a memory and search task. Q. J. Expl Psychol. 40 (A):737–756; 1988.

Backs R.W.; Seljos K.A. Metabolic and cardiorespiratory measures of mental effort: The effects of level of difficulty in a working memory task. Int. J. Psychophysiol. 16:57–68; 1994.

Brookhuis K.; de Waard D. The use of psychophysiology to assess driver status. Ergonomics 36:1099–1110; 1993.

Davies, D. R.; Schackleton, V. J.; Parasuraman, R. Monotony and boredom. In: Hockey, R., Ed. Stress and Fatigue in Human Performance. Chichester: John Wiley; 1983.

Dinges, D.; Graeber, R. Crew fatigue monitoring. Flight Safety Foundation, Flight Safety Digest October; 1989.

Fagerstrom, K.O.; Lisper, H.O. Effects of listening to car radio, experience, and personality of the driver on subsidiary reaction time and heart rate in a long-term driving task. In: Mackie, R.R., Ed. Vigilance: Theory, Operational Performance, and Physiological Correlates. New York: Plenum Press; 1977.

Feyer, A-M.; Williamson, A.M.; Jenkin, R.A.; Higgins, T. Strategies to combat fatigue in the long distance road transport industry: The bus and coach perspective. Report No. CR 122. Canberra, Australia: Federal Office of Road Safety; 1993.

Hamelin, P. Lorry drivers' habits in work and their involvement in traffic accidents. Paper presented at the Second CEC Workshop on Irregular and Abnormal Hours of Work, Brighton, 16–18 January 1987.

Harris, W.; Mackie, R.R. A study of the relationships among fatigue, hours of service, and safety of operations of truck and bus drivers. Report 1727-2. Goleta, California: Human Factors Research Inc; 1972.

Haworth, N. L.; Heffernan, C. L. Information for development of an educational program to reduce fatigue-related truck accidents. Report No. 4. Melbourne, Australia: Monash University Accident Research Centre; 1989.

Haworth, N. L.; Triggs, T. J.; Grey, E. M. Driver fatigue: Concepts, measurement and accident countermeasures. *Report No. CR 72*. Canberra, Australia: Federal Office of Road Safety; 1988.

Haworth, N. L.; Heffernan, C. L.; Horne, E. J. Fatigue in truck crashes. Report No. 3. Melbourne, Australia: Monash University Crash Research Centre; 1989.

Hensher, D. A.; Battellino, H. C.; Gee, J. L.; Daniels, R. F. Long distance truck drivers' on-road performance and economic reward. Report No. CR 99. Canberra, Australia: Federal Office of Road Safety; 1991.

Hoddes E.; Zarcone V.; Smythe H.; Phillips R.; Dement W.C. Quantification of sleepiness: a new approach. Psychophysiology 10:431–436; 1973.

Hyppa M.T.; Kronholm E.; Mattlar C.E. Mental well being of sleepers in a random population sample. Brit. J. Med. Psychol. 64:25–34; 1991.

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

Mackie, R.R.; Miller, J.C. Effects of hours of service, regularity of schedules and cargo loading on truck and bus driver fatigue. Report No. DOT-HS-5-01142. Goleta, CA: Human Factors Research Inc; 1978.

Mascord D.J.; Heath R.A. Behavioural and physiological indices of fatigue in a visual tracking task. J. Safety Res. 23:19–25; 1992.

Monk, T. H.; Folkard, S. Shiftwork and performance. In: Folkard, S. and Monk, T.H., Eds. Hours of work. Chichester: John Wiley and Sons; 1985.

Moore-Ede, M.; Campbell, S.; Baker, T. Falling asleep behind the wheel: Research priorities to improve driver alertness and highway safety. Proceedings of Federal Highway Administration Symposium on Truck and Bus Driver Fatigue, Washington, D.C., 29–30 November 1988.

Ranney, T.; Gawron, V. J. Driving performance as a function of time on the road. Proceedings of the Human Factors Society 31st Annual Meeting; 1987.

Riemersma, J. B. J.; Sanders, A. F.; Wildervanck, C.; Gaillard, A. W. Performance decrement during prolonged night driving. In Mackie, R.R. Ed. Vigilance: Theory, Operational Performance, and Physiological Correlates. New York: Plenum Press; 1977.

Rosa R.; Colligen M. Application of a portable test battery in the assessment of fatigue in laboratory and worksite studies of 12-hour shifts. Scand. J. Work Environ. Hlth 18:113–115; 1992.

Roscoe A.H. Assessing pilot workload. Why measure heart rate. Biol. Psychol. 34:259–287; 1992.

Tepas D.I.; Mahan R.P. The many meanings of sleep. Work and Stress 3:93–102; 1989.

Van Ouwerkerk, F. Relationships between road transport working conditions, fatigue, health and traffic safety. Report No. VK 87-01. Haren, Netherlands: Traffic Research Centre, University of Groningen; 1987.

Williamson, A. M.; Feyer, A.-M.; Coumarelos, C.; Jenkins, A. Strategies to combat fatigue in the long distance road transport industry. Stage 1: The industry perspective. Report No. CR 108. Canberra, Australia: Federal Office of Road Safety; 1992.

Williamson, A. M.; Feyer, A.-M.; Friswell, R.; Leslie, D. Strategies to combat fatigue in the long distance road transport industry. Stage 2: Evaluation of alternative work practices. Report No. CR 144. Canberra, Australia: Federal Office of Road Safety; 1994.