ANALYSIS OF FATIGUE LEVELS IN NEW ZEALAND TAXI AND LOCAL-ROUTE TRUCK DRIVERS

Prepared by:

Samuel G. Charlton, Ph.D.
Peter H. Baas
Brett D. Alley
and
Rebecca E. Luther

Prepared for:

The Road Safety Trust

February 2003

Correspondence or requests for information about this report should be directed to:
Dr. Samuel G. Charlton, TERNZ Ltd.,
353 Pencarrow Road, RD3, Hamilton,
New Zealand. E-mail samiam@ternz.co.nz

Contents

ANALYSIS OF FATIGUE LEVELS IN NEW ZEALAND TAXI AND LOCAL ROUTE TRUCK DRIVERS

Executive Summary

The House of Representatives inquiry into truck crashes found that, despite its importance, driver fatigue is a largely unrecognised problem in New Zealand. The goal of the present research programme was to find out how common driver fatigue is in New Zealand, and the degree to which New Zealand drivers suffer from fatigue and fatigue-related effects on their driving performance. To that end, this phase of the project was directed at collecting demographic, lifestyle, and driver fatigue data from an initial sample of New Zealand taxi and local-route truck drivers.

Previous phases of the project focused on the assessment of driver fatigue in New Zealand by testing a representative sample of 606 truck drivers at a variety of North Island sites along long-haul truck routes (Charlton & Baas, 2001). The results showed that a considerable number of truck drivers were operating in excess of the hours of service regulations. One-third of the drivers reported driving more than the maximum of 11 hours out of 24. Fully 50% of the logging, stock, and line-haul drivers exceeded the maximum hours of service. The three fatigue measures indicated significant levels of driver fatigue in the New Zealand transport industry. One in four of the drivers' self-ratings of fatigue were in the "tired" range, even though many of them were surveyed at the beginning of their shifts. Twenty-four percent of the sample failed one or more of the performance criteria on the psychomotor fatigue test. The results of the daytime sleepiness inventory showed that the drivers in our sample had somewhat higher levels of daytime sleepiness than do heavy goods vehicle operators in the UK. There was significant correspondence between the self-rating, psychomotor performance, and daytime sleepiness fatigue measures and Police Commercial Vehicle Investigation Unit (CVIU)-reported crash rates. As with results from Western Australia, the drivers in our sample typically felt that fatigue was more of a problem for other drivers than for themselves. Finally, the data also revealed that drivers on short-haul routes (under 250 km per day) and drivers over the age of 37 (the average age for truck drivers) were statistically more likely to fail the psychomotor performance test (27% and 34% failure rates for short-haul and older drivers respectively).

This report documents the data and findings gathered from an initial sample of local-route truck drivers and taxi drivers in the Auckland, Waikato, and Bay of Plenty regions. Data from a total of 95 local-route truck drivers and 102 taxi drivers were analysed in terms

of their daily activities, attitudes towards fatigue, and levels of driver fatigue. The results from this initial sample of local-route truck drivers and taxi drivers indicate that there are appreciable levels of fatigue in these sectors of the transport industry and are interesting in several respects. Firstly, as with our previous sample of truck drivers, appreciable numbers of local-route drivers (24%) reported driving longer that the 11-hour maximum allowed. A large percentage of the local-route drivers' self ratings of fatigue were in the "*tired*" range (33.8%). The daytime-sleepiness Epworth Sleepiness Scale (ESS) scores for the local-route drivers were slightly lower than those in the previous sample of short-haul drivers (5.27 vs 5.66) but not significantly so. The psychomotor performance test resulted in significantly different pass rates depending on the local-route drivers' freight types. The drivers of refrigerated freight had the highest pass rate at 90%. The general goods/local freight drivers pass rates were approximately equivalent to those obtained for short-haul drivers in the previous sample of truck drivers (i.e. 71.4% pass rate as compared to 73%). The waste removal drivers, however, were by far the worst in terms of the psychomotor performance test, achieving only a 56.2% pass rate.

As compared to the local-route truck drivers, the taxi drivers in our sample were older, less experienced, and worked longer hours. Forty-two percent of the taxi drivers reported driving more than the 11-hour maximum in the previous 24 hours. The taxi drivers' self-ratings of momentary fatigue were also higher, with 39.2% being in the "*tired*" range. The taxi drivers were much less proficient than the truck drivers at the speed management portion of the tracking task in the TOPS test. Although difficulties in speed management are indeed a key indicator of driver fatigue, the fact that the original TOPS performance criteria were validated with a sample of truck drivers calls into question the applicability of the full set of TOPS criteria as a psychomotor index of the degree of fatigue among our sample of taxi drivers. Although the original TOPS criteria may not be entirely applicable for use outside the truck-driving population, the performance data can be assessed in conjunction with the broader pattern to indicate a range for the incidence of taxi driver fatigue of 25.5% to 35.6%.

Background

The adverse effect of fatigue on human performance is a well-known experience to most of us. We encounter it to some degree in the course of our everyday lives. Brown (1994) has offered the following definition of driver fatigue: *"...the subjective experience of fatigue involves conflict between the desire to rest and the inclination (or perceived commercial pressure) to continue driving to their planned destination...The main effect of fatigue is a progressive withdrawal of attention from road and traffic demands...the withdrawal of attention will be involuntary and difficult, if not impossible to resist...Individuals so affected have been described as 'driving without attention' (DWA) because they are apparently oblivious to impending collisions..*(pp. 311-312)." The present study adopts the above use of the term driver fatigue, treating the phenomenon as a generalised subjective state resulting from a combination of task demands, environmental factors, arrangement of duty and rest cycles, and factors such as drivers' consumption of alcohol and medications. Of particular importance to the present study are the performance decrements in driving that arise from the psychological state of fatigue.

While it is difficult to quantify the contribution of driver fatigue to crash rates, a number of overseas studies have produced estimates. Vic Roads, the state roading authority in Victoria, Australia, has estimated that it is a factor in approximately 25% of all truck-related crashes. Further, it is believed that truck and car drivers are equally responsible for fatigue-related crashes (Vic Roads, 1995). In the United States, it is estimated that each year sleep-related crashes in transportation claim over 15,000 lives and cost more than 12 billion dollars a year in lost productivity and property damage (Caldwell, 2000; Rau, 1996). Other estimates place the incidence of fatigue in commercial driver crashes at somewhere between 1% and 56%, depending on whether the estimates are from safety researchers, transport regulatory agencies, or coroners' findings (Mitler, Miller, Lipsitz, Walsh & Wylie, 1997). Estimates of the incidence of fatigue-related motor crashes vary widely, primarily because fatigue leaves no direct physical evidence at the scene of a crash and thus must be inferred from the circumstances of the crash and from potentially unreliable reports from the individuals involved (Summala & Mikkola, 1994). Nonetheless, it is generally acknowledged that fatigue is significantly under-reported in official crash statistics, and is a high-priority safety issue for the transport industry (Moore & Brooks, 2000).

In New Zealand, the 1996 House of Representatives Report of the Transport Committee on the Inquiry into Truck Crashes found that: *"fatigue is likely to be a significant contributing factor in all types of crashes, not just truck crashes. Despite its importance, however, it is largely unrecognised as a problem in New Zealand."* Similarly, the Land Transport Safety Authority of New Zealand (LTSA) states that driver fatigue is a factor in 8% of fatal crashes and 5% of injury crashes annually (LTSA, 2000). At present the only measure in place in New Zealand with which to assess the incidence and extent of fatigue in drivers is the examination of the driving hours recorded in truck drivers' logbooks. Inasmuch as logbooks are used as the means of compliance checking for hours of service restrictions, and the Truck Crash Inquiry report recognised that the system was widely abused, there is a need to find alternative methods of determining the extent of the driver fatigue problem in New Zealand. Faced with the knowledge that fatigue is a serious problem for the New Zealand transportation system, and the lack of any reliable data on its incidence or impact, the present research programme attempted to identify a reliable means of measuring driver fatigue and then apply it in a large-scale sample of the transport industry.

As alluded to above, direct measures of fatigue are simply not possible. Thus, researchers have searched for measurable correlates and performance indicators of fatigue, with varying degrees of success. Psychophysiological methods such as EEG recordings, eyelid position and activity (Brookhuis, 1995, 2000; Stern, Boyer, & Schroeder, 1994) have shown promise but are intrusive, difficult to measure outside the laboratory, and suffer from relatively large individual differences. Another approach proposed by some researchers has been to measure accident precursors logically associated with fatigue (Brown, 1995; De Waard & Brookhuis, 1991). Measurement of "eyes-off-the road" time and lane-keeping ability has good logical correspondence to increased crash risk but the lack of an agreed-upon benchmark definition of impairment, and practical difficulties in data collection, have kept these measures at the level of discussion and demonstration.

Fatigue has well-documented adverse effects on multiple aspects of cognitive and psychomotor performance. As a result, part-task performance tests of cognitive and behavioural impairment associated with fatigue have been among the most successful measures of fatigue to date. These psychomotor tests have included a wide variety of measures including digit-span, memory, vigilance, divided attention, and eye-hand tracking tasks. Of these tests, the vigilance, divided attention, and tracking tasks have enjoyed the greatest acceptance by researchers and industry professionals. At least part of the reason for this acceptance is their clear relationship to the elements of driving. Tracking task performance closely parallels vehicle-steering and lane-keeping abilities, while divided attention and vigilance tests correspond to the attentional demands of traffic and road conditions. In a series of studies of driver performance in driving simulators (Stein, Paraseghian, Allen, & Miller, 1992; Stein, 1995), fatigue effects were found to be manifested in reliably measurable changes in drivers' ability to maintain their vehicle in the proper lane, to maintain appropriate speed, and in their ability to divide their attention. Although the use of part-task performance tests has generally necessitated laboratory measurement, the increasing power and portability of small computers has seen increasing field use of these tests (Charlton & Ashton, 1997).

Finally, various subjective measures of fatigue and sleepiness have been developed. These have ranged from formalised expert observation (by trained driving instructors or traffic safety officers) to self-rating scales and activity inventories completed by drivers. The success of these measures has been mixed. While expert observations of driving behaviour, or of fatigue correlates such as facial symptoms, have issues of inter-observer reliability, they do appear to possess good sensitivity if the criteria for impaired driving can be appropriately defined (Brookhuis, 2000). The implementation of expert observations as a measure is, however, fairly intrusive, and typically the knowledge that they are being observed has the effect of arousing drivers and masking their fatigue. Self-report inventories of sleep and fatigue have also met with mixed success. Some researchers have argued that drivers are not good assessors of their own momentary levels of fatigue (Bartlett, 1943, cited in Holding, 1983; Brown, 1994), with individuals tending to overestimate their levels of alertness (Rosekind et al., 1994). Recently, however, a number of researchers have found good correspondence between subjective sleepiness and driving impairment (Baulk, Axelsson, Reynor, & Horne, 1998; Maycock, 1995, 1997; Neville, Bisson, French, Boll, & Storm, 1994). In a recent study of subjective sleepiness it was observed that "major incidents" on a driving simulator (all four wheels out of the land) were preceded by self-awareness of increasing sleepiness as early as 40 minutes prior to the incidents (Horne & Reyner, 2000).

The goal of the present research programme was to find out how common driver fatigue was in New Zealand and the degree to which drivers suffered from fatigue-related effects on their driving performance. Phases I and II of the research programme involved the development of a driver fatigue survey and a performance-based fatigue test that were used to collect data from a representative sample of truck drivers throughout the North Island of New

Zealand. The survey asked drivers how many hours they had driven, the amount of sleep they had had in the previous 48 hours, how sleepiness affected them, and about the level of fatigue they felt at that moment. The psychomotor test involved a short drive on a driving simulator to measure their vehicle control and reaction times. The survey form took an average of 10 minutes to fill out and the simulator test took 10-12 minutes to complete, for a total of approximately 20 minutes of driver time.

During Phase I, the roadside survey was trialled on 100 truck drivers in the Waikato District at truck depots, rest stops, and cargo terminals throughout the day and night. During Phase II of the study, data were collected from 506 truck drivers at a variety of North Island sites along long-haul truck routes, including truck stops and depots in Northland, Auckland, Bay of Plenty, Gisborne, Hawke's Bay, Taranaki, Wanganui, and Wellington. The results indicated that there were significant levels of fatigue in the New Zealand transport industry. One in four of the drivers' self-ratings of fatigue were in the "tired" range, even though many of them were surveyed at the beginning of their shifts. The psychomotor test also indicated a very high level of fatigue in the sample. Overall, 24% of the sample failed one or more of the psychomotor performance criteria. Psychomotor performance was found to be significantly related to the amount of rest and sleep, shift length, and the number of driving days per week. Finally, the data also revealed that drivers on short-haul routes (under 250 km per day) and drivers over the age of 37 (the average age for truck drivers) were statistically more likely to fail the psychomotor performance test (27% and 34% failure rates for short-haul and older drivers respectively). This report documents the data and findings gathered from Phase III of the programme: an initial sample of local-route truck drivers and taxi drivers in the Auckland and Waikato regions.

Survey Methodology

The goals of the taxi and local-route truck driver fatigue survey were to: 1) identify key demographic and work/rest patterns; 2) collect information on drivers' attitudes towards fatigue and propensity for daytime sleepiness, for comparison with other studies of driver fatigue; 3) to obtain self-assessments on drivers' momentary levels of fatigue; and 4) to collect performance data on fatigue-related driving impairment. Earlier phases of the study were directed at developing the written questionnaires, adapting the performance test, and testing the data collection methodology. The details of the development work were documented in the Phase I report "*Fatigue and Fitness for Duty of New Zealand Truck*

Drivers Phase I Report: Initial Driver Sample and Concept Demonstration" (Charlton, Baas & Ashton, 1998). The characteristics of the survey instrument (written questionnaire and performance test) and their adaptation for use in the present study are summarised below.

Written questionnaire

In order to minimise the disruptive effects of the testing protocol on the drivers' schedules, it was desirable to make the questionnaire short enough to complete in 10-15 minutes. The questions were selected by reviewing a variety of prior surveys related to fatigue and truck driving. The resulting questionnaire was reviewed by independent researchers in the field and ultimately used with long-haul truck drivers. The demographic portion of the questionnaire was then modified for use with taxi drivers and local-route drivers. The questionnaire used to collect data for this study is shown at Appendix A.

The demographic portion of the questionnaire contained approximately one dozen questions about the drivers' ages, their years of professional driving experience, their types of employment, vehicle types, average workday lengths and typical driving distances. The demographic questions were followed by three questions on the degree to which driver fatigue was perceived as a hazard to road safety, for the purposes of comparison with prior studies (Hartley et al., 1996) and in order to determine any potential relationship between these attitudes and driving schedules and driver performance.

The second page of the survey contained a rating scale intended to capture driver estimates of their own levels of momentary fatigue. The rating scale was adapted from the USAF School of Aerospace Medicine Crew Status Survey (CSS) which has been employed in studies of operator workload and fatigue in a variety of aviation and command and control systems (Charlton, 1996). The fatigue scale was followed by an activity inventory which tallied the drivers' time spent actually driving, as well as their sleep periods, mealtimes, physical exercise and freight-loading duties, time spent engaged in any desk work, rest periods, and any partying or drinking over the preceding 48 hours. This activity survey was also adapted from a USAF School of Aerospace Medicine instrument that has been developed to study the activity and rest cycles of aircrews, medical teams, field air traffic controllers, and personnel in other extended-duration duties. (Neville, Bisson, French, Boll and Storm, 1994)

The last page of the survey contained eight questions on the degree to which the drivers were likely to feel sleepy in various situations. These questions, known collectively as the Epworth Sleepiness Scale or ESS (Maycock, 1995, 1997), were included to provide another point of comparison with the momentary fatigue ratings and the activity inventory. The sleepiness scale, while not an indicator of a driver's momentary sleepiness or fatigue, is a good indicator of overall sleep debt and has been used in several studies linking the likelihood of daytime sleepiness with accidents by car drivers and heavy goods vehicle drivers.

Psychomotor performance test

The psychomotor performance test was based on driving simulator hardware and software purchased from Systems Technologies Inc. of Hawthorne California. The hardware consisted of a Pentium™ computer equipped with a 34020 TIGA graphics board and 20-inch monitor for displaying the driving scenario; a Metrabyte M5312-4 optical encoder interface card, throttle/brake pedal controller and active steering controller; a sound board and amplified stereo speakers for presenting audio feedback and instructions to the participants; a VGA display card and 14-inch monitor for displaying control information to the experimenter; and a printer. The equipment was configured and installed in a caravan for easy transport and set-up at the data collection sites (See Figure 1).

Figure 1. Fatigue survey caravan.

The software consisted of the commercially-available Truck Operator Proficiency System (TOPS) testing software. TOPS is based on a dual-axis sub-critical tracking task (maintaining speed and steering in a controlled but unstable environment, a virtual roadway affected by the appearance of random wind gusts requiring steering correction), and a tertiary or side-task requiring driver monitoring and periodic responses. In the course of its development, TOPS passed through three verification and validation stages (Stein et al., 1992): baseline testing of the device on long-haul truck drivers (to establish driver acceptance, reliability and ease of use), development of pass/fail criteria for driver performance (based on a discriminant analysis of 40 performance measures taken from three separately-sampled sets of long-haul truck drivers), and field testing to correlate TOPS performance with actual driving performance and physiological measures of decreased alertness (i.e. EEG, EOG and EMG).

The TOPS performance index algorithm was defined, such that the resulting criteria would have a fatigued driver failure rate of at least 50% (correct detection of fatigued drivers) with a non-fatigued failure rate of only 5% (failure by non-fatigued drivers). These criteria were selected to maintain an acceptably low rate of falsely identifying a driver as fatigued, while still detecting the 50% of drivers most adversely impaired by fatigue. Further, since the test was designed for use in selective enforcement stops (testing drivers suspected of being impaired) and not in random testing applications, the operational false positive rate is purportedly much lower than 5% (Stein, 1995). As an aside, it should be noted that different algorithms were obtained for impairment due to fatigue versus impairment due to alcohol. The alcohol data showed impairment on similar variables but the magnitude of the effects was different.

Because the testing scenario so closely resembles the operational reality of driving, TOPS has enjoyed very good operator acceptance where it has been employed. As with all fitness for duty tests, when a driver testing paradigm has clear relevance to “real world” driving situations, and the safety implications associated with passing or failing the test are readily apparent, driver acceptance is readily obtained (Miller, 1976). The original TOPS driving scenarios underwent various modifications (i.e. road markings, left-side driving, display of metric rather than Imperial speedometer units) to make them more relevant to New Zealand drivers. The resulting performance test scenario consisted of an eight-minute testing session comprising a straight-road scene and 27 to 30 (depending on the driver’s speed) divided-attention events. As with the original TOPS studies, the divided- attention events consisted of symbols presented in the side mirrors to which the driver responded by indicating for a left turn, right turn or pressing the horn button (as appropriate for the type of symbol displayed).

The test scenario was divided into four two-minute data collection blocks for analysis purposes. As was the case in the original TOPS studies, data from the first two-minute block were excluded from the analysis. A variety of driver performance data were collected throughout the test scenario. Of particular interest were the performance variables used to calculate a pass/fail score by the TOPS performance index algorithm. Table 1 lists the 20 performance measures used; the 40 variables used in the algorithm consisted of the mean value and the standard deviation for each measure across the data collection blocks.

Calculation of pass/fail scores was based on five performance index coefficients (linear combinations of the 40 performance variables), such that a driver's performance was transformed according to the five performance indices and compared to established performance criteria for each of the indices. One of the driver's data files is shown at Appendix B. The five indices, although composed of different weightings of the 40 performance variables, can be characterised as focussing on the following five general categories: curvature error variability, divided attention response time variability, throttle activity variability, steering activity variability, and longitudinal speed variability. A driver was required to obtain a passing score on each of the five performance indices in order to receive a passing score for the trial as a whole. The criteria used in the present study were the same as used in the original TOPS studies, with the exception of the removal of a sixth performance index and criterion which was used to detect driving impairment associated with blood alcohol levels in the later TOPS validation trials.

Data collection procedure

For the taxi driver sample, two methodological approaches were tried. In the first, taxi drivers were approached on the taxi ranks and given flyers asking them to come to the testing caravan (which was parked in a central location) during a break or at the end of their shifts. Drivers approached in this way were reluctant to participate – only 5 drivers were tested over a twelve-hour period. A more effective method of recruitment was to approach taxi companies and to have the depot managers and dispatchers approach potential participants as drivers returned to the depot. It was emphasised to the drivers that their participation was still voluntary, however they were more willing to participate when they realised that their companies supported the research. All of the local-route truck drivers were recruited in this way. In total, 25 companies were approached and offered the opportunity to participate in the research. Of these, six declined to participate, stating that they were either too busy or were not interested, leaving a sample that involved 19 companies (seven taxi and twelve local route).

Drivers expressing a willingness to participate were shown the informed consent form that guaranteed confidentiality of their simulator performance and responses to the survey questions. Drivers were asked to sign a copy of the form and were then asked the Driver Fatigue Survey questions by one of the experimenters. Completion of the survey took an average of 10-15 minutes and was followed by the experimenter showing the driver to the caravan for the driving simulator performance portion of the test.

Table 1. Performance test measures.
Number of correct divided attention responses
Number of incorrect divided attention responses
Number of divided attention responses with no response
Number of road departures (collisions)
Average time for a divided attention response (seconds)
RMS for a divided attention response
Average lane deviation (feet)
RMS lane deviation
Average steering wheel rate (degrees/sec.)
RMS steering wheel rate
Average vehicle heading error (degrees)
RMS vehicle heading error
Average curvature error (1/foot)
RMS curvature error
Average throttle activity (g's/sec)
RMS throttle activity
Average longitudinal acceleration (g's)
RMS longitudinal acceleration
Average longitudinal speed (miles/hour)
RMS longitudinal speed

The driving performance test was begun by seating the drivers in front of the monitor and simulator controls and showing them how to adjust the seat so that they were comfortable and could easily reach the hand and foot controls. This was followed by the presentation of a two-minute orientation scenario which automatically provided visual and auditory instructions on what to expect, and practice in "driving" the simulated vehicle and in responding to the divided attention symbols. After completion of the orientation scenario the drivers were given a final opportunity to ask questions and the eight-minute performance test was conducted. At the end of the performance test drivers were thanked for their participation, provided with a LTSA factsheet on driver fatigue, and given complimentary chocolate bars. The performance-testing portion of the survey took an average of 12-15 minutes to complete per driver.

The goal was to test all drivers at the ends of their shifts, and this was possible with the taxi drivers as their schedules were somewhat more flexible than those of the local-route truck drivers. Due to the time pressures on local-route drivers, we found that it was impossible to test them exclusively at the ends of their shifts, and a few of these drivers were tested during breaks, before and after loading between runs. In the case of the local-route truck drivers, the amount of driving immediately prior to the survey ranged from zero to 13 hours, averaging 7.01 hours (std. dev. of 3.03) across all drivers. The taxi drivers' hours ranged from two to 14 hours of work just prior to testing (an average of 9.75 hours, std. dev. of 2.61). This was only slightly less than the drivers' reported average shift length (an average of 10.3 hours, std. dev. of 1.77). The number of fares the taxi drivers reported taking that day ranged from two to 20, with an average of 15.43 fares (std. dev. of 9.49). The participants' hours of driving immediately prior to taking part in the survey are shown in Figure 2. Of the 202 drivers tested, the data from five local-route truck drivers were discarded due to: incomplete data (3); eyesight difficulties (1); and a manager who was not a regular driver (1).

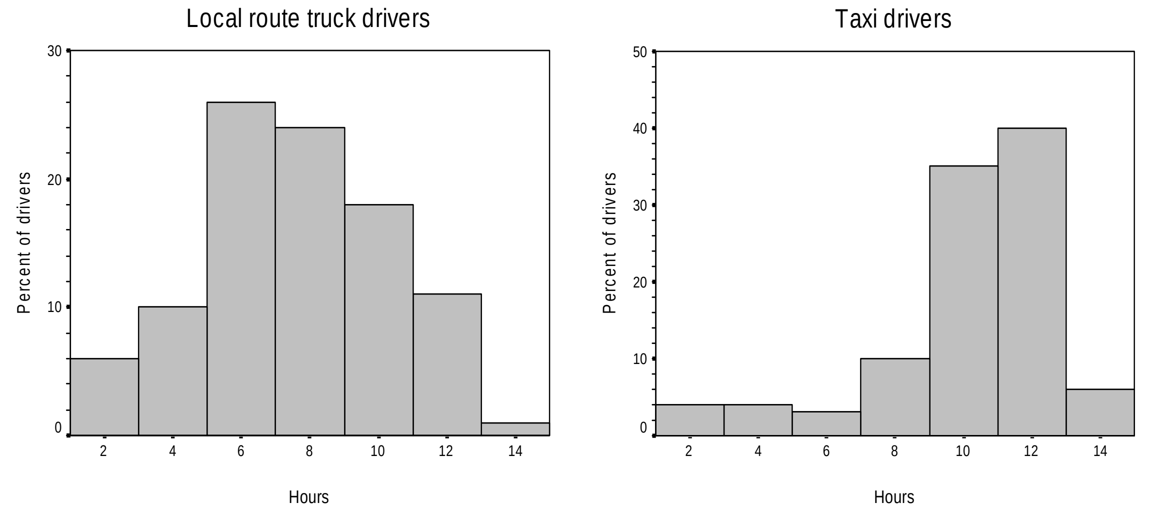

Figure 2. Participants' driving hours prior to survey.

Results: Local-route truck drivers

Driver demographics

The average age of the 95 drivers in the local-route sample was 38.89 years (ranging from 23 to 61 years, standard deviation of 8.98). Driving experience averaged 13.31 years (ranging from less than 1 year to 42 years, standard deviation of 10.14 years). All but one of the local-route drivers participating in the survey were male.

As shown in Figure 3, the majority of the local-route truck drivers sampled were company employees. The next largest category, subcontractors, made up 35% of the sample, while drivers working for owner/drivers and one driver describing himself as a "freelancer" comprised the rest of the sample. Figure 4 shows the type of vehicle driven and the freight contents reported by the drivers. Forty-five percent of the drivers reported driving a truck between 5 and 10 tonnes, 40.4% drove a larger truck, and 14.9% reported driving a vehicle smaller than 5 tonnes. As can be seen in the figure, the largest freight category was general goods (53.7%) followed by waste removal (33.7%) and refrigerated goods (10.5%). The drivers reported an average distance driven per shift ranging from 38 km to 500 km (an average of 201.80 km, std. dev. of 94.42).

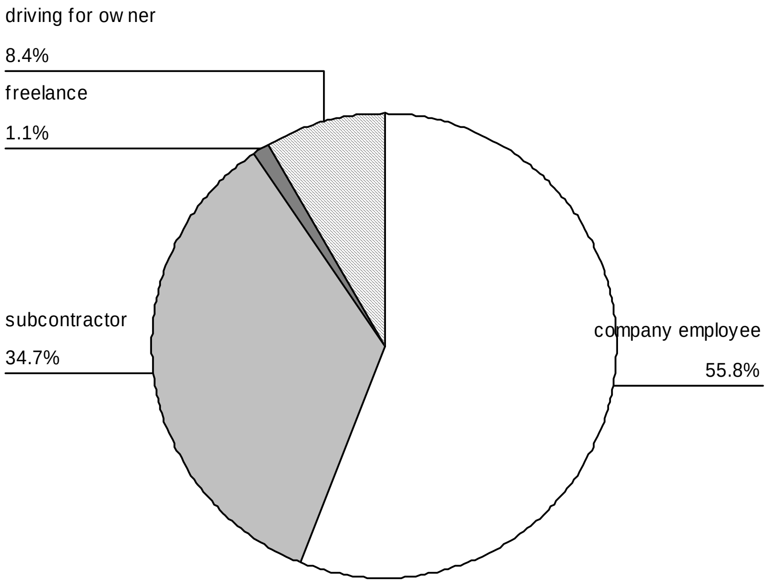

Figure 3. Local-route truck drivers' employment status.

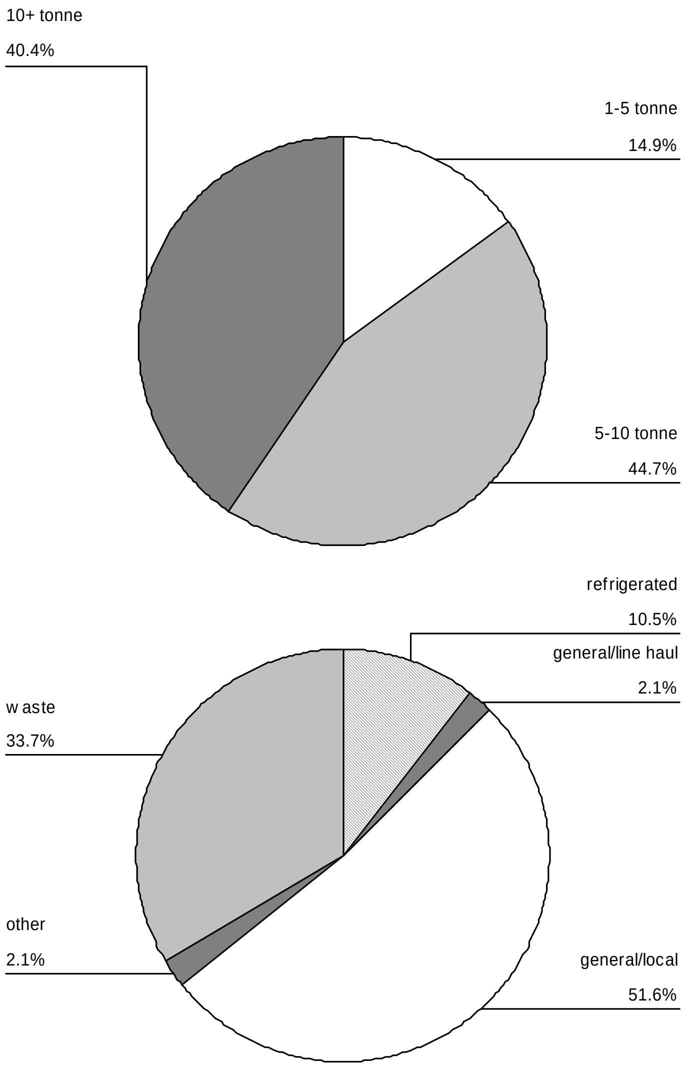

Figure 4. Local route truck drivers' vehicles & freight types.

Driver activities and hours of work

When asked their typical number of days worked per week, the local-route drivers reported an average of 5.35 (ranging from 3 to 6 days, std. deviation of 0.52), with an average shift length of 9.96 hours. Examining the activity data from the 48 hours prior to the survey shows that the number of hours spent driving in the previous 24 hours ranged from zero to 24 hours (an average of 9.45 hours, std. deviation of 2.93 hours). A total of 24% of the drivers reported more than 11 hours of driving during the previous 24 hours. The total number of hours of driving in the previous 48 hours ranged from three to 26 (an average of 14.78 hours, std. deviation of 6.10). There were some differences in the driving demands for the different freight types: waste removal drivers reported somewhat longer shifts and greater distances

(10.02 hrs and 228.94 km) than refrigerated goods drivers (9.90 hrs and 134.90 km) and general goods/local drivers (9.81 hrs and 190.71 km).

Looking at the off-duty activities of the drivers, the average amount of sleep reported for the previous 24 hours was 6.87 hours (std. deviation of 1.74), and 12.50% of the drivers reported their last rest/sleep period was less than the required 9 hours. The full activity data from the drivers are shown in Table 2. When asked if they drove to a fixed company schedule, 88.5% of the drivers answered "yes." Eighty-two percent of the drivers said, however, that they could stop and rest when they wanted to. Finally, 92.7% of the local-route drivers said that they had loaded the freight they were carrying that day.

Table 2. Local-route truck driver activity data

	Mean	Std. Deviation	Minimum	Maximum
Work days per week	5.353	0.520	3.00	6.00
Distance per shift (km)	199.000	90.819	38.00	500.00
Shift length (hr)	9.939	1.613	5.80	14.00
Hours driving in past 24hrs	9.448	2.934	0.00	20.00
Hours driving in past 48hrs	14.781	6.101	3.00	26.00
Length of last duty shift	9.333	2.888	3.00	14.00
Hours sleeping in past 24hrs	6.865	1.739	2.00	11.00
Hours sleeping in past 48hrs	14.990	2.565	7.00	21.00
Length of last sleep	6.708	1.829	2.00	11.00
Length of last rest & sleep	15.882	8.432	4.00	39.00
Meals in past 24hrs	2.208	0.753	.00	4.00
Meals in past 48hrs	4.167	1.171	1.50	7.50
Physical work/exercise past 24hrs	0.583	1.751	.00	14.00
Physical work/exercise past 48hrs	1.188	2.459	.00	18.00
Desk work in past 24hrs	0.271	1.021	.00	5.00
Desk work in past 48hrs	0.396	1.261	.00	8.00
Relaxing in past 24hrs	6.521	3.078	.00	13.00
Relaxing in past 48hrs	15.1986	6.247	.00	30.00
Partying in past 24hrs	0.208	0.994	.00	8.00
Partying in past 48hrs	0.823	2.484	.00	15.00

Fatigue measures

There were three principal fatigue measures included in the survey: the CSS self-report fatigue scale; the ESS daytime sleepiness inventory; and the TOPS performance test. The local-route truck drivers' CSS self-ratings of levels of momentary fatigue ranged from 1 "Fully alert" to 6 "Extremely tired." The median rating for the sample was 3, "Somewhat fresh." This was the same as the median rating reported for the previous sample of 596 truck drivers (Charlton & Baas, 2001). As shown in Figure 5, drivers working night shifts had a median fatigue rating of 4, "a *little tired*", but as only ten of the 95 local-route drivers were on a night shift (all in the general goods/local freight category), this sample size was not large enough to determine whether this was a statistically-reliable difference. In all, 33.8% of the local-route drivers rated themselves as being "*a little tired*" or worse. A stepwise multivariate regression analysis predicting the CSS self-ratings from the various activity measures found that the hours of driving prior to the survey, the number of hours relaxing in the past 24, the length of the last sleep and the number of meals in the past 48 hours were the strongest predictors of fatigue self-ratings: $\underline{F}_{(5,\ 89)} = 2.766$, $\underline{p} < .05$, $R^2 = .134$, stepwise criterion for inclusion was $\underline{p} \leq 0.05$.

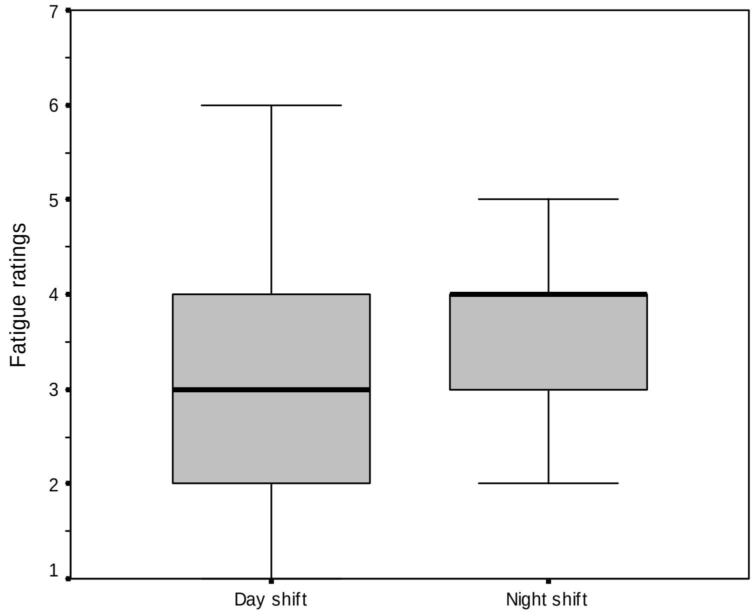

Figure 5. Median, interquartile range, and range of CSS ratings for day and night-shift local-route drivers.

In response to questions on how great a problem fatigue was for themselves and other drivers, most local-route drivers in the sample thought that fatigue was a greater problem for other drivers than for themselves (see Figure 6). This is essentially the same pattern of results obtained in surveys of truck drivers in Western Australia and New Zealand (Hartley et. al., 1996; Charlton & Baas, 2001). In Hartley et al.'s (1996) study of truck drivers in Western Australia, "other" drivers were seen as having a problem with fatigue *always* or *often* by 35.8% of drivers, and fatigue was seen as a problem for themselves *always* or *often* by only 10% of drivers. In the previous survey of New Zealand truck drivers, 21.4% rated fatigue as being a problem for other drivers *always* or *often*, whereas only 3.4% rated it as a problem for themselves *always* or *often* (Charlton & Baas, 2001). In our sample of local-route drivers, only 2.1% rated fatigue as *often* a problem for themselves, whereas 34.8% rated it *always* or *often* a problem for other drivers. Similarly, 21.9% of the drivers rated fatigue as *never* a problem for themselves, while only 10.1% rated it as *never* a problem for others. When asked if fatigue was dangerous for drivers, however, 88.5% of the drivers answered *always* with a further 8% of drivers answering *often*.

Drivers' ratings of fatigue as a problem for others were highly correlated with their ratings of fatigue as a problem for themselves; r = .360, p < .01. The drivers' ratings of fatigue as a problem for themselves also correlated well with their CSS ratings; r = .347, p < .01. Further, their ratings of fatigue danger were also correlated with CSS ratings, such that drivers rating fatigue as a danger were more likely to rate their own levels of fatigue as higher; r = .211, p < .05. Finally, there was a trend for older drivers to rate fatigue as less of a problem than did the younger drivers, suggesting a lower level of fatigue awareness among older drivers; r = -.186, p< .05.

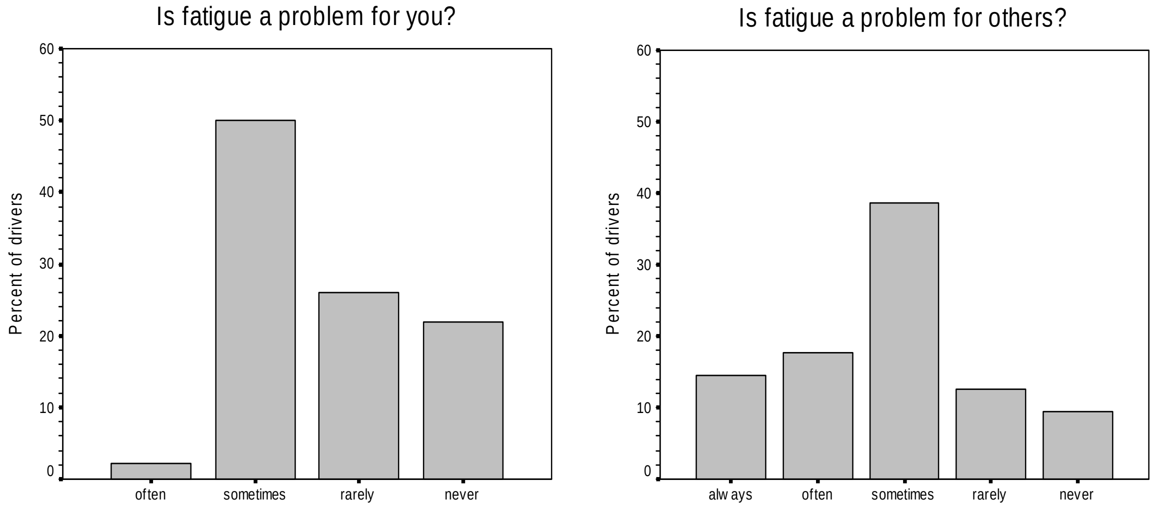

Figure 6. Local-route drivers' ratings of fatigue as a problem for drivers

The second fatigue measure, the Epworth Sleepiness Scale (ESS), was used an indicator of chronic sleepiness and sleep debt as opposed to the momentary fatigue measured by the CSS. The local-route drivers' ESS scores are shown in Figure 7; the average score was 5.27 (std. dev. = 3.67), lower than the average of 6.13 in the previous survey of New Zealand truck drivers and the 5.70 reported for heavy goods drivers in the UK (Charlton & Baas, 2001). This measure of chronic sleepiness is, however, not too much different from the average ESS score of 5.66 reported by the short-haul truck drivers (250 km or less) in our previous sample of New Zealand drivers. A multiple regression analysis predicting ESS ratings found that the best predictors were the drivers' ages and their years of driving experience; $\underline{F}_{(2,\ 88)} = 3.694$, $\underline{p}<.05$, $R^2 =.077$, stepwise criterion for inclusion was $p \leq 0.05$. Figure 8 shows the relationship between the ESS scores and drivers' ages and years of experience. ESS scores generally increased co-linearly with driver age and years of experience until the age of 50, or 30 years of driving experience (the inverse of the relationship between CSS ratings and age/experience). Interestingly, the drivers' ESS scores were correlated with their ratings of how often fatigue was a problem for themselves, r = .270, $\underline{p} < .01$, but not with the CSS ratings.

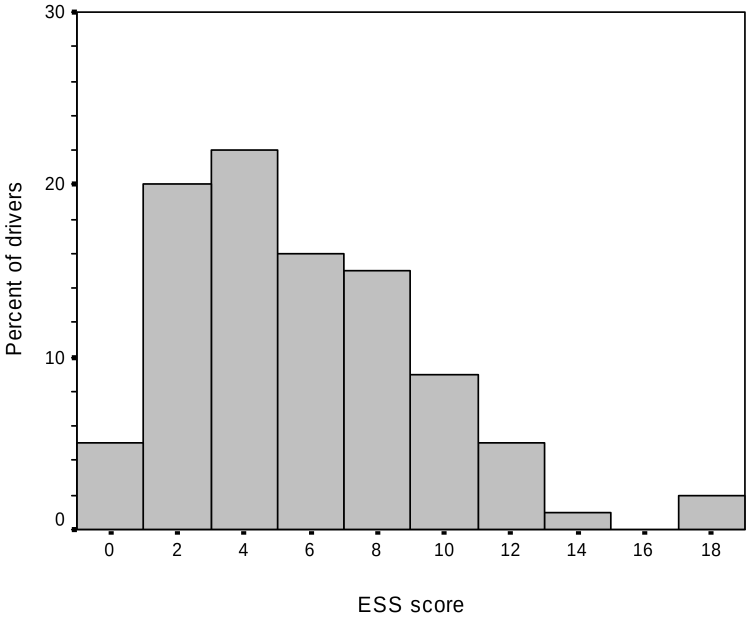

Figure 7. Distribution of local-route drivers' Epworth Sleepiness Scale (ESS) scores

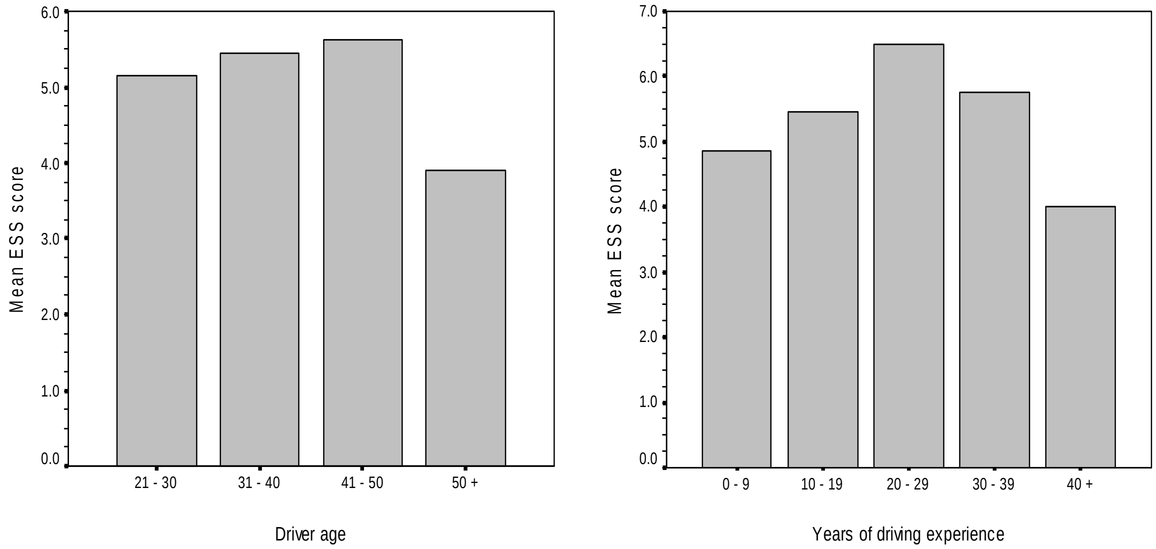

Figure 8. Epworth Sleepiness Scale (ESS) scores as a function of driver age and years of driving experience

The third fatigue measure was the TOPS psychomotor performance test. Across all freight category types, 33.7% of the local-route truck drivers failed one or more of the TOPS criteria. This failure rate is substantially higher than the 27% failure rate previously reported for truck drivers with routes under 250 km (Charlton & Baas, 2001), but corresponds reasonably well with the number of drivers rating themselves as “a little tired” or worse on the CSS (33.8%). A nonparametric correlation coefficient computed for the CSS ratings and the TOPS pass/fail criterion confirmed a significant relationship between these two fatigue measures; Spearman’s rho = .185, $\underline{p} < .05$. There was no corresponding relationship between TOPS and the chronic sleepiness measure, the ESS.

The drivers’ freight category was the demographic factor with the most profound relationship to the TOPS performance results. Figure 8 shows the TOPS failure rates for each of the freight categories: 10% of the refrigerated freight drivers (one in ten) failed to meet the performance criterion whereas the failure rate for the general goods/local category was 28.6% (14 of 49 drivers) and 43.8% for the waste removal drivers (14 of 32 drivers). The “other” “general/line haul” freight categories contained only two drivers apiece, and their failure rates were 50% and 100% respectively. A Pearson chi-square analysis indicated that the differential TOPS pass rate as a function of freight type was marginally significant; $\chi^2 = 8.712$, df = 4, $\underline{p} < .069$. For comparison, the right-hand panel of Figure 9 shows the drivers’ average CSS ratings by freight type and, as can be seen, the drivers’ self-ratings of fatigue generally correspond to the pattern of TOPS failure rates.

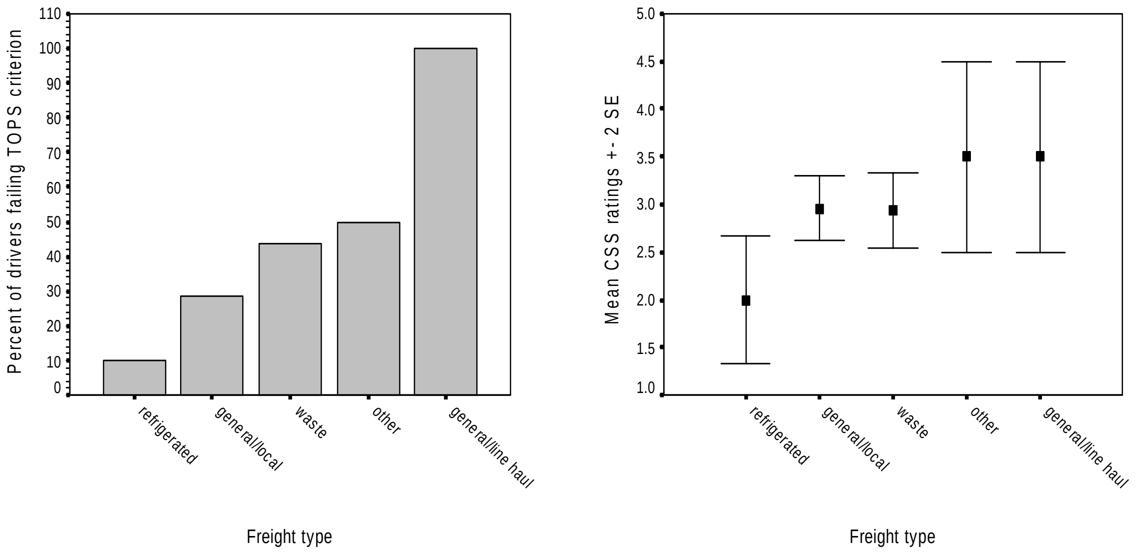

Figure 9. TOPS performance test failure rates and CSS ratings as a function of freight type

A discriminant analysis predicting TOPS pass/fail rates across all freight categories from the demographic and activity measures found that the best predictors were the drivers' years of driving experience and their average shift length; $\underline{F}_{(2, 88)} = 15.413$, $\underline{p}<.001$. Figure 10 shows the percentage of drivers meeting the TOPS performance criterion as a function of years of driving experience and average shift length. As indicated in the figure, drivers with fewer years of experience and shorter shifts were more likely to pass the TOPS performance test. The pass rates ranged from 82.6% for drivers with nine years of experience or less, to 25% for drivers with 30–39 years of experience; neither of the two drivers with more than 39 years of experience met the TOPS performance criterion[1]. A Pearson chi-square analysis indicated that the differential pass rate for various levels of driving experience was significant; $\chi^2 = 16.943$, df = 4, $\underline{p} < .01$. Similarly, only 47.6% of drivers with an average shift length of greater than 11 hours met the TOPS criterion, as compared to 71.6% of drivers with shifts of 11 hours or less. A Pearson chi-square analysis indicated that this difference too was significant; $\chi^2 = 4.210$, df = 1, $\underline{p} < .05$.

---

[1] Years of driving experience were, of course, highly correlated with driver age, r = .708, but the discriminant analysis indicated that experience was the better predictor of TOPS performance, hence its use here.

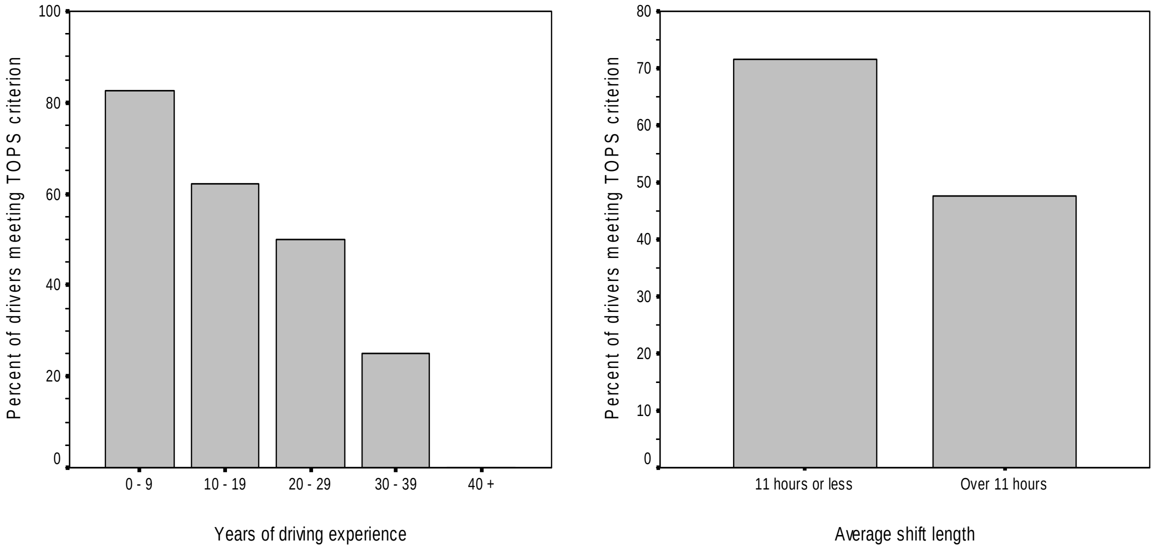

Figure 10. TOPS performance test pass rates as a function of years of driving experience and average shift length

Finally, a discriminant analysis predicting TOPS performance from the 40 psychomotor performance measures collected, identified nine significant predictors: RMS vehicle heading error; RMS longitudinal speed; RMS longitudinal speed variability; the number of divided attention events missed; RMS divided attention response time variability; curvature error variability; variability of incorrect divided attention responses; average longitudinal acceleration; and average vehicle heading error; $\underline{F}_{(9, 85)} = 15.038$, $\underline{p}<.001$.

Results: Taxi drivers

Driver demographics

The average age of the 102 taxi drivers sampled was 47.92 years (ranging from 25 to 71 years, standard deviation of 10.67). Driving experience averaged 6.55 years (ranging from less than one year to 38 years, standard deviation of 7.74 years). Seventy-nine (77.5%) of the drivers participating in the survey were male.

As is shown in Figure 11, only 5.9% of the drivers sampled were company employees. The largest categories, drivers working for owner/drivers and subcontractors, made up 48% and 46.1% of the sample respectively. The majority of the drivers (54.9%) reported driving a car for their job, 26.5% drove a van, and 18.6% reported driving both. The average distance driven per shift ranged from 33 km to 350 km (an average of 195.49 km, std. dev. of 49.99).

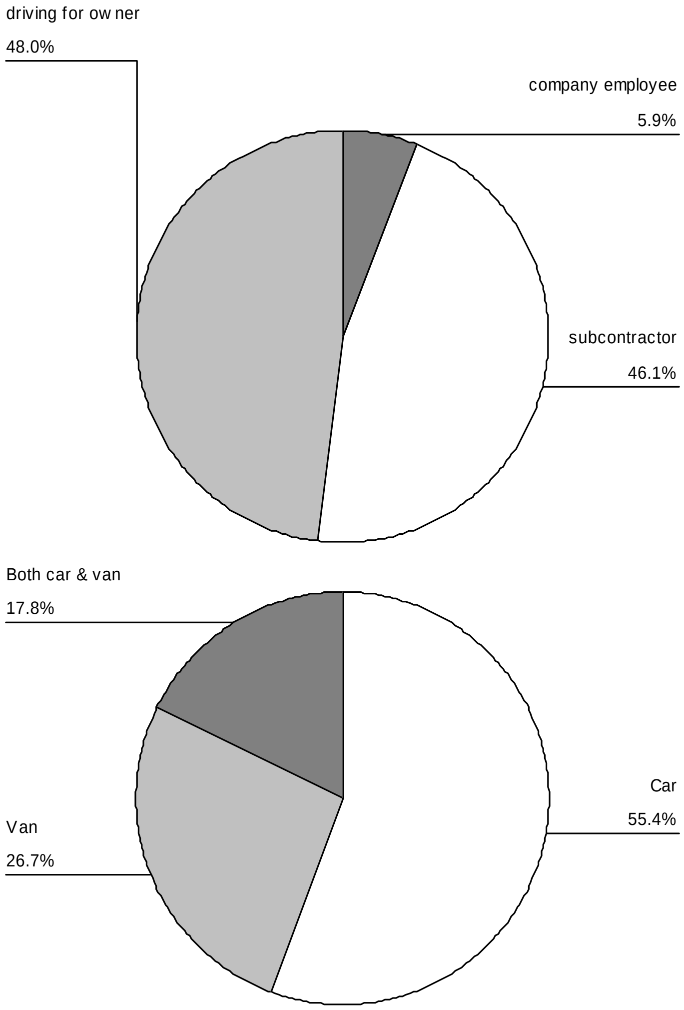

Figure 11. Taxi drivers' employment status and type of vehicle

Driver activities and hours of work

When asked their typical number of days worked per week, the participating drivers reported an average of 5.05 (ranging from one to six days, std. deviation of 1.13). As mentioned above, the drivers reported a typical shift length averaging 10.3 hours. Examining the activity data from the 48 hours prior to the survey, however, shows that the number of hours spent driving in the previous 24 hours ranged from zero to 16 hours (an average of 10.68 hours, std. deviation of 2.62 hours). A total of 42.2% of the drivers reported more than eleven hours on the job during the previous 24 hours. The total number of hours of driving in the previous 48 hours ranged from two to 31 (an average of 18.36 hours, std. deviation of 6.19). The full activity data from the drivers are shown in Table 3.

Table 3. Taxi driver activity data

	Mean	Std. Deviation	Minimum	Maximum
Work days per week	5.054	1.128	1.00	6.00
Distance per shift (kms)	195.490	90.819	33.00	350.00
Shift length (hrs)	10.338	1.766	2.00	14.00
Number of fares today	15.429	9.487	2.00	40.00
Hours driving in past 24hrs	10.676	2.622	2.00	16.00
Hours driving in past 48hrs	18.363	6.185	2.00	31.00
Length of last duty shift	8.412	4.662	.00	15.00
Hours sleeping in past 24hrs	8.109	2.366	2.00	13.00
Hours sleeping in past 48hrs	15.333	4.349	2.00	24.00
Length of last sleep	8.584	2.338	2.00	15.00
Length of last rest & sleep	15.882	8.432	4.00	39.00
Meals in past 24hrs	1.770	0.681	.00	3.50
Meals in past 48hrs	3.265	1.254	.50	6.50
Physical work/exercise past 24hrs	0.235	1.902	.00	8.00
Physical work/exercise past 48hrs	0.804	2.670	.00	21.00
Desk work in past 24hrs	0.814	2.456	.00	15.00
Desk work in past 48hrs	1.520	4.507	.00	26.00
Relaxing in past 24hrs	4.069	2.611	.00	11.00
Relaxing in past 48hrs	10.216	6.044	.00	24.00
Partying in past 24hrs	0.039	0.396	.00	4.00
Partying in past 48hrs	0.157	0.931	.00	7.00

Moving to the off-duty activities of the drivers, the average amount of sleep reported for the past 24 hours was 8.11 hours (std. deviation of 2.37). Looking at the total length of their last sleep and rest period, 6.90% of the sample reported that their last rest/sleep period was less than the required nine hours. The average amounts of rest and sleep reported by the drivers did not appear unreasonable, although there were drivers reporting as little as two hours of sleep in the past 48 hours. When asked if they drove to a fixed company schedule, 67.6% of the drivers answered “yes.” Ninety-eight percent of the drivers said, however, that they could stop and rest when they wanted to.

Fatigue measures

The first of the three principal fatigue measures included in the survey, the CSS self-report fatigue ratings of momentary fatigue, ranged from 1 "Fully alert" to 5 "Moderately tired" for the taxi drivers. The median rating in the sample was a 3, “Somewhat fresh,” the same as the median rating obtained for the local-route truck drivers. As shown in Figure 12, the median CSS rating for drivers on night shift was 4, “a *little tired*”, as compared to the median of 3 for the day shift. Analysis of variance indicated that this difference in the CSS ratings of day and night-shift taxi drivers was statistically significant; $F_{(1,100)} = 13.065$, $\underline{p} < .001$. There were no significant differences in the drivers’ CSS ratings due to gender, employment category or vehicle type. A stepwise multivariate regression analysis identified two significant driver activity measures predicting the CSS self-ratings: hours of sleep in the past 24 hours and the number of hours relaxing in the past 24 hours; $\underline{F}_{(2, 95)} = 7.224$, $\underline{p} < .001$, $R^2 = .132$, stepwise criterion for inclusion was $\underline{p} \leq 0.05$.

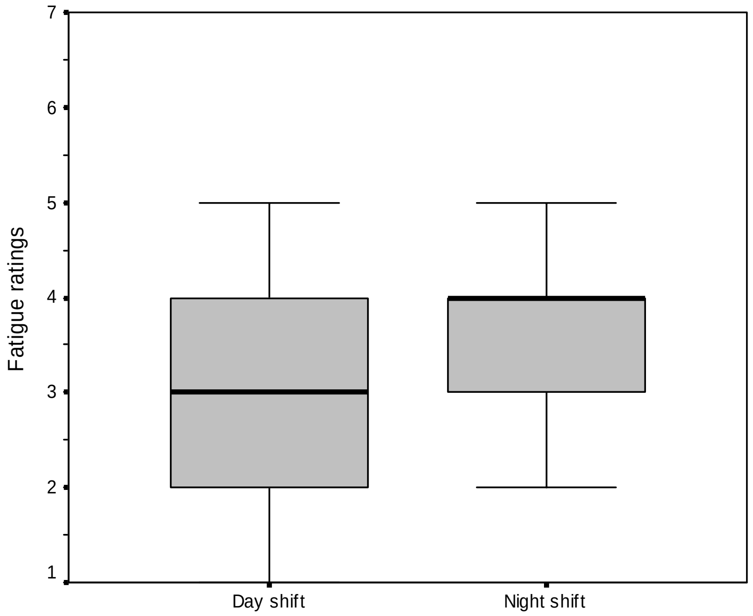

Figure 12. Median, interquartile range, and range of CSS ratings for day and night-shift taxi drivers.

In response to questions on how great a problem fatigue was for themselves and other drivers, most taxi drivers in the sample thought that fatigue was a greater problem for other drivers than for themselves (see Figure 13). When asked if fatigue was dangerous for drivers, however, 79% of the drivers answered *always,* with 6% and 9% of the drivers answering *often* or *sometimes* respectively. This is the same pattern of results obtained with the local-route truck drivers and in previous surveys of truck drivers in Western Australia and New Zealand (Hartley et al., 1996; Charlton & Baas, 2001). Only 2% of the taxi drivers rated fatigue as *often* a problem for themselves, whereas 25.6% rated it *always* or *often* a problem for other drivers. As was the case with the CSS results, there was a significant difference in the ratings of day and night-shift taxi drivers with regard to their ratings of whether fatigue was a problem for them ($F_{(1.100)} = 7.298$, $\underline{p} < .01$); the night-shift drivers had a median rating of 3 (*sometimes*) as compared to the day-shift drivers' median rating of 4 (*rarely*). Of some concern was the high proportion of drivers rating fatigue as *never* a problem for themselves (34.3%), and *never* a problem for others (23.8%), indicating either a fairly low level of fatigue awareness or some inhibition in their responses to the survey questions.

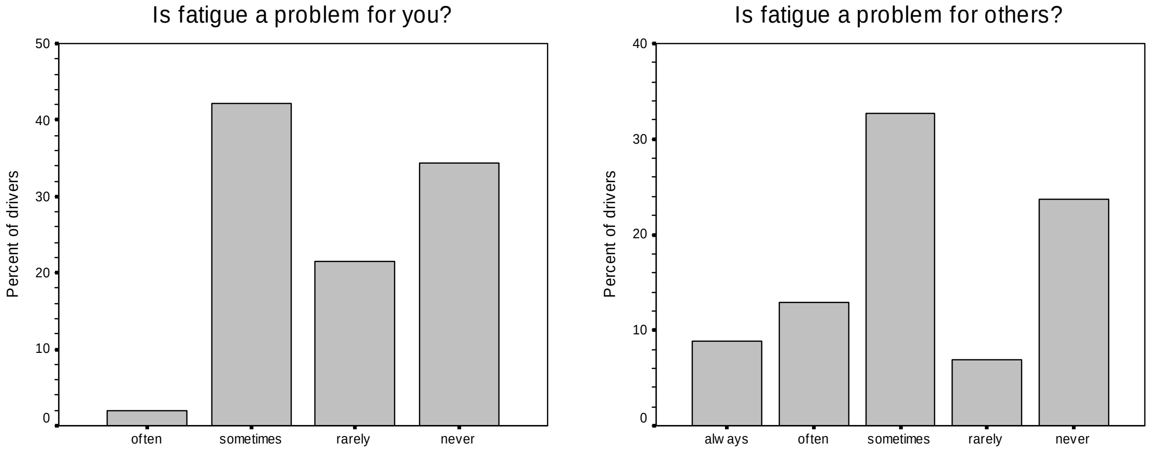

Figure 13. Taxi drivers' ratings of fatigue as a problem for drivers

The taxi drivers' results for the Epworth Sleepiness Scale (ESS) were much lower than the scores reported by truck drivers; the taxi drivers' average score was 3.9 (std. dev. = 3.29) as compared to the truck drivers' average of 6.13 in New Zealand and 5.70 in the UK (Charlton & Baas, 2001). As shown in Figure 14, the 78 male taxi drivers reported significantly higher ESS scores than did the 23 female drivers; $\underline{F}_{(1, 99)} = 4.427$, $\underline{p}<.05$. Unlike the fatigue ratings, there was no statistically-significant difference in the day and night shift taxi drivers' ESS ratings ($F_{(1,99)} < 1$). A stepwise multiple regression analysis predicting ESS ratings found two reliable predictors: the drivers' average shift length and their number of days per week of driving; $F_{(2, 97)} = 5.591$, $p<.01$, $R^2 = .103$, stepwise criterion for inclusion was $p \leq 0.05$. The right-hand panel of Figure 14 compares the ESS ratings for the 52 drivers reporting an average shift length of less than 11 hours with those for the 49 drivers reporting an average shift length of 11 hours or more. As can be seen in the figure, drivers working shifts of 11 hours or more reported higher levels of daytime sleepiness (average ESS score = 4.92, std. dev. = 3.63) than did the drivers working shorter shifts (average ESS score = 2.94, std. dev. = 2.63). An analysis of variance indicated that this difference was significant; $F(1,99) = 9.886$, $p < .01$.[2]

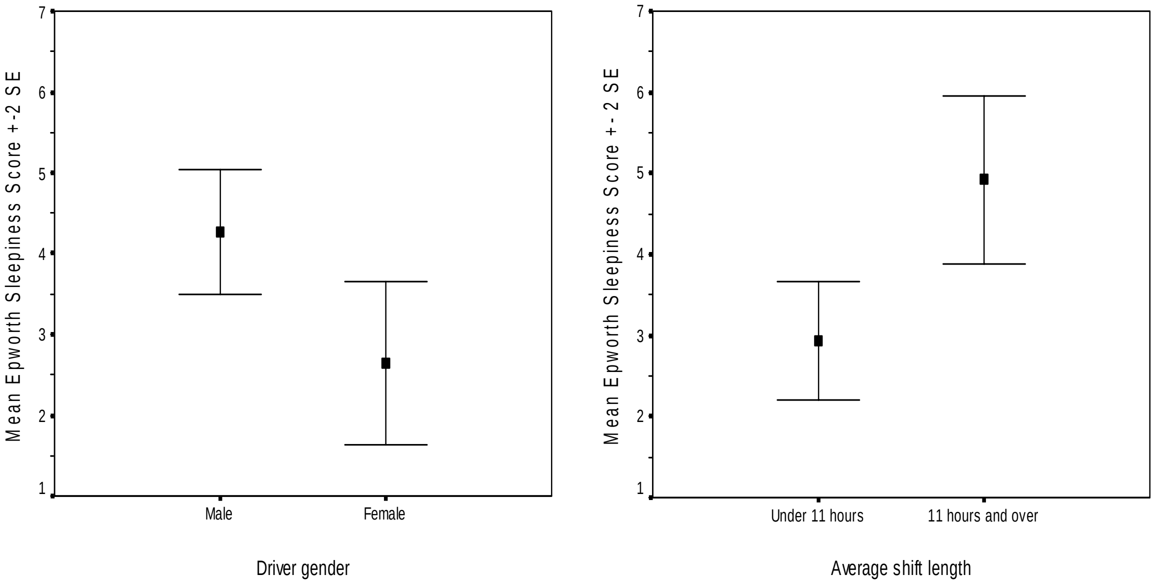

Figure 14. Taxi drivers' Epworth Sleepiness Scale scores

When compared with truck drivers, the taxi drivers displayed significant difficulties in managing the speed of their simulated vehicle during the TOPS psychomotor performance test. As can be seen in Figure 15, only 37.3% of the taxi drivers were able to pass Criterion 1 of the TOPS test, a performance standard based predominantly on longitudinal speed and longitudinal speed variability. Examining TOPS pass rates after removing Criterion 1, 74.5% of the taxi drivers were able to meet all four remaining criteria. This failure rate of 25.5% is

---

[2] Note that this analysis is somewhat different from the one used for the local-route truck drivers where shift lengths of 11 hours and under were compared to shifts of over 11 hours. When the same distinction is made for the taxi drivers the difference in ESS scores is not as profound, albeit still statistically significant; $F_{(1,99)} = 4.462$, $p < .05$.

comparable to the 24% of drivers showing appreciable levels of fatigue in the previous sample of long-haul truck drivers (Charlton & Baas, 2001), and perhaps represents a reasonable lower-bound fatigue estimate for taxi drivers. A multivariate regression predicting the drivers' pass rates on Criteria 2 to 5 showed that the length of the drivers' last rest and sleep was the best predictor of this measure of fatigue; $\underline{F}_{(1,\ 72)} = 3.633$, $\underline{p} < .06$ $R^2 = .048$, stepwise criterion for inclusion was $\underline{p} \leq 0.05$.

Speed management was, however, an important indicator of fatigue for the taxi drivers; a stepwise multivariate regression predicting the CSS self-ratings indicated that average longitudinal speed was the best psychomotor predictor of the drivers' CSS ratings; it yielded an $\underline{F}_{(1,\ 99)} = 5.566$, $\underline{p} < .05$, Pearson's $r = .231$. Acknowledging this importance, the question then becomes how best to form an alternative psychomotor criterion that reflects the importance of speed management, while allowing for the poorer performance of the taxi drivers.

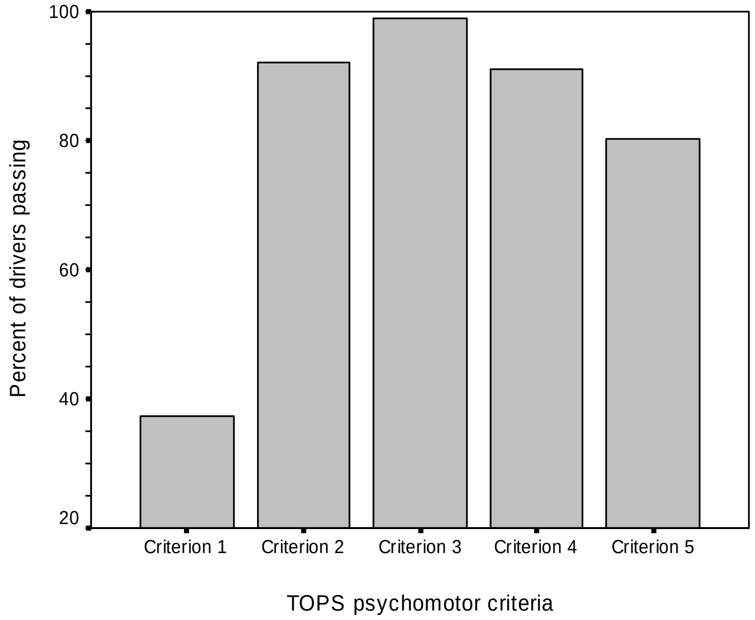

Figure 15. Pass rates for each TOPS criterion

A simple approach to establishing a new performance criterion was based on identifying drivers with average longitudinal speeds more than one standard deviation above the mean AND with longitudinal speed variability less than one standard deviation below the mean. Figure 16 shows the drivers' longitudinal speeds as a function of their CSS ratings, which displayed a significant correlation ($r = .231$, $\underline{p} < .05$); average longitudinal speeds were higher and longitudinal speed variability was lower for drivers with high self-ratings of fatigue on the CSS scale. Applying this new performance criterion to the sample of taxi drivers, 7.8% of drivers failed the longitudinal speed criterion, resulting in an overall psychomotor fatigue rate of 29.4% when combined with the existing TOPS criteria 2 to 5.

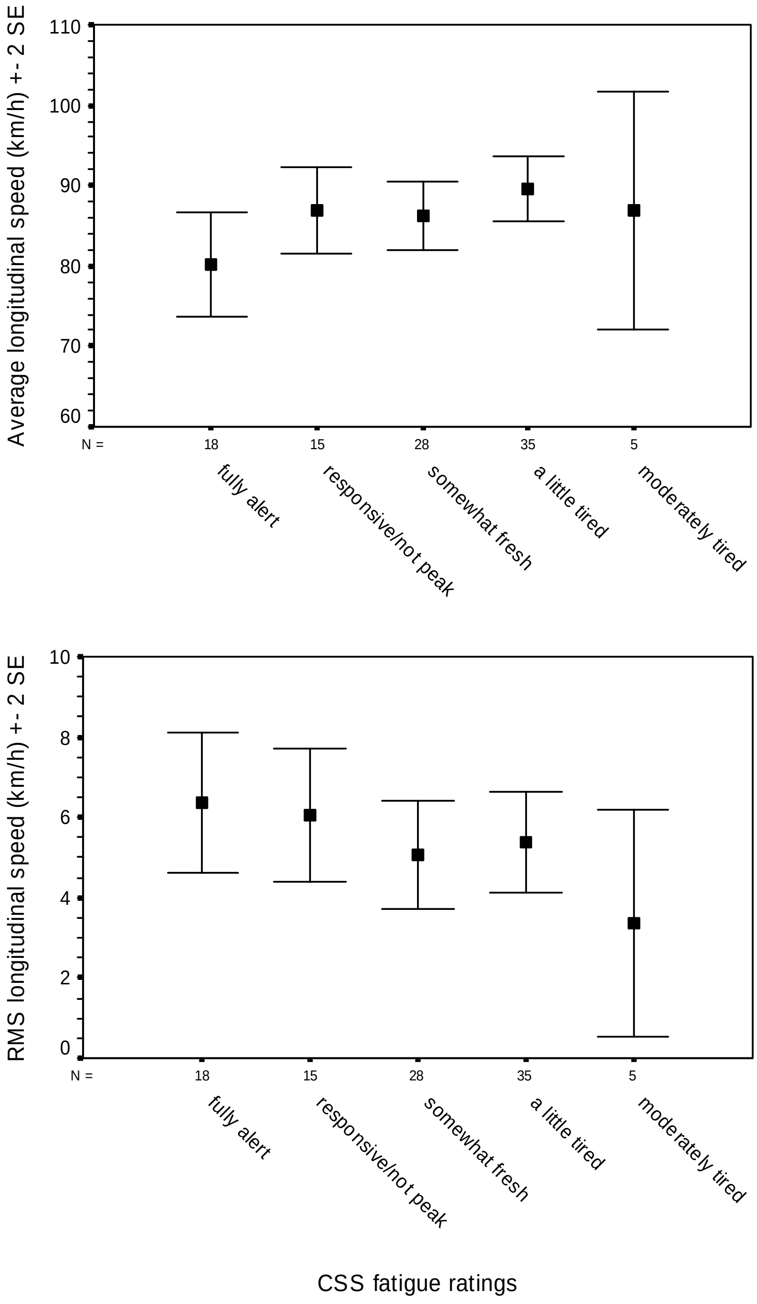

Figure 16. Drivers' average longitudinal speeds and longitudinal speed variability as a function of their self-ratings of fatigue

An alternative approach to identifying the levels of fatigue in the sample was to use a discriminant analysis to build a performance criterion based on the observed performance measures with the highest predictive relationships to the fatigue ratings. In the present sample those performance measures included: average longitudinal speed; steering wheel activity variability (deg./sec.); curvature error variability; vehicle heading error variability; the mean number of collisions (all four wheels off the simulated carriageway); and average throttle activity. Together, these variables accounted for 22% of the variance in the CSS ratings and produced a significant predictive model; $F_{(6,94)} = 4.321$, $\underline{p} < .001$. Based on this model of six performance measures, the discriminant analysis predicted that 32.67% of the drivers would be classified as "moderately tired" and the performance of the sampled drivers would yield an overall fatigue rate of 35.64% when the model was used to reclassify the results of the revised 5-criteria TOPS performance test.

This new six-variable performance model was significantly correlated to several activity measures (as well as the fatigue self-ratings). A stepwise multiple regression analysis predicting the new performance model identified seven significant activity measures: average distance driven per shift; hours on duty in the previous shift; the time spent partying in the past 48 hours; the average shift length; the time spent relaxing in the past 24 hours; the number of days worked per week; and the time spend engaged in physical work or exercise in the past 24 hours; $F_{(7,68)} = 4.784$, $\underline{p} < .001$, $R^2 = .348$, stepwise criterion for inclusion was $\underline{p} \leq 0.05$.

Implications

The results from this initial sample of local-route truck drivers and taxi drivers indicate that there are appreciable levels of fatigue in these sectors of the transport industry and are interesting in several respects. Firstly, as with our previous sample of truck drivers, appreciable numbers of local-route drivers (24%) reported driving longer than the 11-hour maximum allowed. Although the local-route truck drivers' median CSS rating of their own fatigue was the same as that in the previous sample of truck drivers, a larger percentage of the drivers were in the "*tired*" range (33.8% as compared to 24.7%). This difference could reflect the fact that most of the drivers in the current sample were tested at the end of their shifts whereas testing occurred at a variety of times throughout the shifts of the drivers in the previous sample. The daytime sleepiness (ESS) scores for the local-route drivers were slightly lower than those in the previous sample of short-haul drivers (5.27 vs 5.66) but not significantly so.

The third fatigue measure, the psychomotor performance test, resulted in significantly different pass rates depending on the drivers' freight type. The drivers of refrigerated freight had the highest pass rate at 90%. The general goods/local freight drivers' pass rate was approximately equivalent to that obtained for short-haul drivers in the previous sample of truck drivers (i.e. 71.4% as compared to 73%). The waste removal drivers, however, were by far the worst in terms of the psychomotor performance test, achieving only a 56.2% pass rate. It is impossible to ascertain whether this was indicative of much higher rates of fatigue in the waste removal drivers or some lower level of skill in operating the computer-based TOPS tracking task. Waste removal drivers did, however, report the longest average shift length (10.016 hrs, std. dev. = 1.388 as compared to 9.811, std. dev. = 1.853 for general/local) and the longest distance per shift (228.938 kms, std. dev. = 86.617 as compared to 190.714, std. dev. = 94.195 for general/local). The results also indicated that years of driving experience and shift length were negatively correlated with psychomotor performance in that local-route truck drivers with more years of experience and those driving longer than 11 hours on an average shift were less likely to pass the psychomotor performance test. The reason for the driving experience effect is not immediately apparent, but this does correspond in part to the finding that older drivers' psychomotor performance was poorer as reported for the previous sample of truck drivers.

As compared to the local-route truck drivers, the taxi drivers in our sample were older, less experienced, and worked longer hours. Forty-two percent of the taxi drivers reported driving more than the 11-hour maximum in the previous 24 hours. The taxi drivers' self-ratings of momentary fatigue were also higher; 39.2% of the drivers' ratings were in the "*tired*" range. These latter two findings are particularly noteworthy in the light of the data collectors' impressions that the taxi drivers were somewhat less forthcoming in their answers than the truck drivers, and typically understated their driving hours and fatigue levels. Their ESS daytime sleepiness scores were much lower than those reported for other samples of drivers, lending credence to the suggestion that the taxi drivers' survey responses tended to understate actual levels of fatigue and driving hours.

The taxi drivers were much less proficient than the truck drivers at the speed management portion of the tracking task in the TOPS test. Although difficulties in speed management are indeed a key indicator of driver fatigue, the fact that the original TOPS performance criteria were validated with a sample of truck drivers calls into question the applicability of the full set of TOPS criteria as a psychomotor index of the degree of fatigue among our sample of taxi drivers. Setting aside the speed management portion of the task, and using only a partial TOPS criterion as a lower bound for psychomotor fatigue, produced an estimate of 25.5% for fatigued taxi drivers in our sample. An alternative psychomotor criterion based on a discriminant function of six performance measures resulted in an estimate of 35.6% fatigued drivers; perhaps a more realistic estimate of fatigue incidence which correlated well with the subjective CSS ratings and with seven of the activity measures associated with fatigue in the previous sample of truck drivers. Thus, while the original TOPS criteria may not be entirely applicable for use outside the truck-driving population, the performance data can be assessed in conjunction with the broader pattern to indicate a range for the incidence of taxi driver fatigue of 25.5% to 35.6%.

In the light of the above findings, future research in this area might reasonably be directed towards three areas of inquiry. Firstly, the development and validation of psychomotor performance criteria for fatigued drivers should be undertaken outside the population of truck drivers previously investigated. A second development effort might focus on the validation of performance criteria for older drivers, a population with higher than average psychomotor failure rates in both the present study and the previous survey. These new criteria could make use of, not only subjective ratings and tracking task performance, but also converging measures such as expert ratings of driver performance (as was used in the original TOPS validation trials), as well as tracking a few individuals' psychomotor performance as they are deprived of sleep for up to 48 hours, and obtaining comparison performance data at various levels of blood alcohol, as has been suggested by previous researchers (Dawson & Reid, 1997; Fairclough & Graham, 1999). The result of this validation work would be psychomotor performance criteria for fatigue that are more widely applicable to a greater range of New Zealand drivers.

Following one or both of the criteria development efforts described above, future work might also be directed towards obtaining larger representative samples of taxi drivers, local-route drivers, and other driving populations of interest throughout New Zealand. These surveys could not only be used to assess the incidence of fatigue among New Zealand drivers, but could also serve to educate the road-using public regarding the effects and dangers of fatigue. With some modification to the testing protocols, a wealth of data regarding the driving performance of the driving public would become available to researchers and road safety professionals.

References

Brookhuis, K. (1995) Driver impairment monitoring by physiological measures. In: L. Hartley (ed.) Fatigue and driving: Driver impairment, driver fatigue, and driving simulation. London: Taylor & Francis.

Brookhuis, K. (2000) Defining criteria for driver impairment. In: Proceedings of 4th International Conference on Fatigue and Transportation. Western Australia: Institute for Research in Safety and Transport, Murdoch University.

Brown, I.D. (1994) Driver fatigue. Human Factors, 36, 298-314.

Brown, I.D. (1995) Methodological issues in driver fatigue research. In: L. Hartley (ed.) Fatigue and driving: Driver impairment, driver fatigue, and driving simulation. London: Taylor & Francis.

Caldwell, J.A. (2000) Efficacy of stimulants for fatigue management: The effects of Provigil and Dexedrine on sleep-deprived aviators. In: Proceedings of 4th International Conference on Fatigue and Transportation. Western Australia: Institute for Research in Safety and Transport, Murdoch University.

Charlton, S.G., Baas, P.H., and Ashton, M.E. (1998, August) Fatigue and Fitness for Duty of New Zealand Truck Drivers: Initial Driver Sample and Concept Demonstration (TERNZ Technical Report) Report to Land Transport Safety Authority /Road Safety Trust. Hamilton, NZ: Transport Engineering Research NZ Ltd.

Charlton, S.G. and Baas, P.H. (2001) Fatigue, work-rest cycles, and psychomotor performance of New Zealand truck drivers. New Zealand Journal of Psychology, 30(1), pp. 32-39.

Dawson, D. and Reid, K. (1997) Fatigue, alcohol and performance impairment. Nature, 388, 235.

De Waard, D. and Brookhuis, K. (1991) Assessing driver status: A demonstration experiment on the road. Accident Analysis & Prevention, 23, 297-307.

Fairclough, S.H. and Graham, R. (1999) Impairment of driving performance caused by sleep deprivation or alcohol: A comparative study. Human Factors, 41, 118-128.

Hartley, L.R., Arnold, P.K., Penna, F., Hochstadt, D., Corry, A. & Feyer, A-M. (1996) Fatigue in the Western Australian transport industry. Part one: The principle and comparative findings. Report No. 117. Perth: Institute for Research in Safety and Transport, Murdoch University.

Holding, D. (1983) Fatigue. In: R. Hockey (ed.) Stress and fatigue in human performance. Chichester, England: John Wiley & Sons.

Horne, J.A. and Reyner, L.A. (2000) Sleep related vehicle accidents: Some guides for road safety policies. In: Proceedings of 4th International Conference on Fatigue and Transportation. Western Australia: Institute for Research in Safety and Transport, Murdoch University.

Maycock, G. (1995) Driver Sleepiness as a Factor in Car and HGV Accidents. TRL Report 169. Crowthorne, Berkshire: Safety and Environment Resource Centre, Transport Research Laboratory.

Maycock, G. (1997) Sleepiness and driving. The experience of heavy goods vehicle drivers in the UK. Journal of Sleep Research, 6, 238-244.

Mitler, M.M., Miller, J.C., Lipsitz, J.J., Walsh, J.K. & Wylie, C.D. (1997) The sleep of long-haul truck drivers. The New England Journal of Medicine, 337:11, 755-761.

Moore, B. and Brooks, C. (2000) Heavy vehicle driver fatigue: A policy adviser's perspective. In: Proceedings of 4th International Conference on Fatigue and Transportation. Western Australia: Institute for Research in Safety and Transport, Murdoch University.

Neville, K.J., Bisson, R.U., French, J., Boll, P.A. & Storm, W.F. (1994) Subjective fatigue of C-141 aircrews during operation desert storm. *Human Factors,* 36(2), 339-349.

Rau, P.S. (1996) NHTSA's Drowsy driver research program. Washington DC: National Highway Traffic Safety Administration.

Rosekind, M.R., Gander, P.H., Miller, D.L., Gregory, K.B., Smith, R.M., Weldon, K.J., Co, E.L., McNally, K.L. & Storm, W.F. (1994) Fatigue in operational settings: Examples from the aviation environment. Human Factors, 36, 327-338.

Stein, A.C. (1995) Detecting fatigued drivers with vehicle simulators. In: L. Hartley (ed.) Fatigue and driving: Driver impairment, driver fatigue and driving simulation. London: Taylor & Francis.

Stein, A.C., Parseghian, Z., Allen, R.W. & Miller, J.C. (1992) High risk driver project: Theory, development and validation of the Truck Operator Proficiency System (TOPS). Volume 2: Report. Hawthorne, CA: Systems Technology, Inc. (Technical report 2417-1).

Stern, J.A., Boyer, D. & Schroeder, D. (1994) Blink rate: A possible measure of fatigue. Human Factors, 36, 285-297.

Summala, H. & Mikkola, T. (1994) Fatal accidents among car and truck drivers: Effects of fatigue, age, alcohol consumption. Human Factors, 36, 315-326.

LTSA (March, 2000) Fatigue and Driver Alertness: Factsheet 24. Wellington, New Zealand: Author.

New Zealand House of Representatives (August, 1996) Report of the Transport Committee on the inquiry into truck crashes. Author

United States Transportation Safety Board (1995) Factors that Affect Fatigue in Heavy Truck Accidents. Author

Vic Roads (September, 1995) Safety First, Victoria's Road Safety Strategy 1995-2000. Author

Appendix A – Taxi & Local Route Truck Driver Fatigue Survey Form

TERNZ Driver Fatigue Study
Information and Consent Form

We are conducting this study of driver fatigue in conjunction with the Road Safety Trust. What we are interested in learning is how common of a problem driver fatigue is in New Zealand and the degree to which NZ drivers suffer from fatigue-related effects. Our study of driver fatigue uses the term fatigue to refer to the general feeling of tiredness resulting from a combination of task demands, environmental factors, arrangement of duty and rest cycles, and factors such as consumption of alcohol and medications.

Our fatigue study uses two kinds of fatigue measures: 1) a short survey asking about the hours you have driven, the amount of sleep you have had in the past 48 hours, how sleepiness affects you, and the level of fatigue you feel at the moment, and 2) a short drive on a driving simulator to measure your vehicle control and reaction times. We are looking to measure some drivers at the beginning of their shift, some at the middle of their shift, and some at the end of their shift. **This informed consent form will, in writing, guarantee absolute confidentiality of your simulator performance and responses to the survey questions**. The survey will take approximately 5 to 10 minutes to fill out and the simulator test will take approximately 8 to 10 minutes to complete, for a total of approximately 15-20 minutes.

Researcher's Copy

I have received and read an information sheet about this research project. I have had the chance to ask any questions and discuss my participation. I understand that my individual answers and performance will be kept confidential. Any questions have been answered to my satisfaction.

I agree to participate in this research project and I understand that I may withdraw at any time.

Participant's Name: __________________ Signature: _________________ Date: ______

- -

Participant's Copy

I have received and read an information sheet about this research project. I have had the chance to ask any questions and discuss my participation. I understand that my individual answers and performance will be kept confidential. Any questions have been answered to my satisfaction.

I agree to participate in this research project and I understand that I may withdraw at any time.

Participant's Name: __________________ Signature: _________________ Date: ______

Date _________
Time _________
Location _________

Driver Questionnaire

Age ______ Height _________ Weight __________ M/F

How long have you been driving a truck/taxi for a living? _______ years

Are you:
_______ a company employee driver
_______ an owner driver subcontracting to a transport co
_______ a free lance owner driver
_______ a driver working for an owner driver
_______ other (specify) _____________________________________

(SHT) What type of freight are you carrying? ___________________________________

What vehicle types do you drive (taxi, van, small truck (<1 tonne), large truck?
____

Are you driving to a fixed company schedule? *yes/no*

Can you stop and rest when you want? *yes/no*

(SHT) Did you load the freight you are carrying or help load it? *yes/no*

(Taxis) How many fares have you had today?

What is your average workday length? _________ hours

What is your average number of days per week? _________

What is the average distance you drive during each day? ___________

How would you classify the traffic density that you have been working in today?

VERY LIGHT LIGHT MEDIUM HEAVY VERY HEAVY

Is tiredness or fatigue a problem for you when you drive?

(*circle the appropriate word)* always often sometimes rarely never

Do you think tiredness or fatigue is a problem for other drivers?

(*circle the appropriate word)* always often sometimes rarely never

Do you think tiredness or fatigue is dangerous on the road?

(*circle the appropriate word)* always often sometimes rarely never

Fatigue Survey

Place an X in the box that best describes how you feel right now.

	Fully alert, wide awake, extremely peppy
	Very lively, responsive, but not at peak
	Okay, somewhat fresh
	A little tired, less than fresh
	Moderately tired, let down
	Extremely tired, difficult to concentrate
	Completely exhausted, unable to function effectively, ready to drop

Activity Survey

Place a mark along the timeline below to show when each activity occurred and for how long

For example: Driving: X-------------X
Sleeping: X----------------X
Meal: L or H

(A line may be easier to do rather than a number of crosses, as long as it is clear.)

Time of day

	6	8	10	12	2	4	6	8	10	12	2	4	6	8	10	12	2	4	6	8	10	12	2	4	6
	am			noon						midnight						noon						midnight			am
Driving																									
Sleeping																									
Meal (H for big meal L for light)																									
Physical work/exercise																									
Desk work/sedentary																									
Relaxing/TV/reading																									
Partying/drinking																									

Sleepiness Survey

How likely are you to doze or fall asleep in the following situations, as opposed to just feeling tired?
(This refers 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. Tick the box for the most appropriate answer for each situation)

Situation	Would never doze	Slight chance of dozing	Moderate chance of dozing	High chance of dozing
Sitting and reading				
Watching TV				
Sitting inactive in a public place (theatre or meeting)				
As a passenger in a car for an hour without a break				
Lying down 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				