Pilot Test of Fatigue Management Technologies

David F. Dinges, Greg Maislin, Rebecca M. Brewster, Gerald P. Krueger, and Robert J. Carroll

a a a a
c ( ) c a a a a a
a c b a ca
a a a a b c a
b c a
b c a a ac a c
a c c c ab
b c a ac c c
a bac a c a
c a a a
b c c a c a c
bac bac ac c a
bac (c) b
bac ( ) a a c a
a c a ab c c c a
a ac a a a c a
a c
bac ca c (*p* =)
a a ac a ab (*p* =) a c
b c a c a a
a a a c b ca (a a c
a b a) a
c a a c a a a c c a
b a ab
c a

There are a growing number of technologies that purport to help drivers manage fatigue and drowsy driving (*1–3*). In addition to establishing their validity to detect fatigue, there is a critical need to determine whether feedback from such technologies during driving could affect the behavior or alertness of commercial motor vehicle operators. Building on previous work by the U.S. Department of Transportation (USDOT), a study was carried out on the effects of feedback from a group of fatigue management technologies (FMT) bundled as a single intervention. Sponsored by the Federal Motor Carrier Safety Administration (FMCSA) and Transport Canada, in cooperation with the American Transportation Research Institute (ATRI), the study was tasked to develop an experimental design and instrumentation plan and to conduct a pilot test of commercial truck drivers' reactions to a combination of FMT, under federally mandated hours of service in both Canada and the United States. Since it was neither cost-effective nor practical to conduct a separate study of each individual technology, the selected technologies were combined and tested as a set within a single field trial that had two phases: one in Canada and one in the United States. The project involved an extensive over-the-road test of the combined FMT. The objective was to determine how drivers, engaged in over-the-road trucking operations, reacted to FMT and whether the technologies would improve the alertness and fatigue awareness of commercial truck drivers by providing information feedback about changes in sleep need, in drowsiness, and in driving performance during routine driving schedules. Specifically, the research sought to determine whether feedback from combined FMT would enhance drivers' alertness and performance at work and increase sleep times on workdays or nonwork days. A secondary specific aim was to obtain driver reaction to FMT. It was hypothesized that deployment of FMT would result in improved driver alertness and performance while driving (Hypothesis I) and in increased sleep time (Hypothesis II) and under both current U.S. hours of service and Canadian hours of service.

D. F. Dinges, Division of Sleep and Chronobiology, Unit for Experimental Psychiatry, School of Medicine, University of Pennsylvania, 1013 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104. G. Maislin, Biomedical Statistical Consulting, 1357 Garden Road, Wynnewood, PA 19096. R. M. Brewster, American Transportation Research Institute, 1850 Lake Park Drive, Suite 123, Smyrna, GA 30080. G. P. Krueger, Krueger Ergonomics Consultants, 4105 Komes Court, Alexandria, VA 22306-1252. R. J. Carroll, Office of Research and Technology, Federal Motor Carrier Safety Administration, U.S. Department of Transportation, 400 Virginia Avenue SW, Suite 600, Washington, DC 20024.

Transportation Research Record: Journal of the Transportation Research Board, No. 1922, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 175–182.

METHODS

Criteria for FMT Selection

Technologies selected were bundled into a single intervention from four fatigue management domains: one providing objective information on driver sleep need, one providing objective information on driver drowsiness, one providing objective information on lane tracking performance, and a technology that reduced the work involved in controlling vehicle stability while driving. Although each technology is described separately, the effects of feedback from them was investigated as a single intervention encompassing all four. This was deliberate—the project was not designed or resourced to compare the impact of individual FMT to each other or to compare the effects of FMT in Canadian versus U.S. drivers. The selection of specific technologies was not an endorsement of their validity or reliability. Technologies were selected for use in the pilot study because (*a*) each was representative of one of the four fatigue management domains, (*b*) each was available for study through the cooperation of their respective developers, and (*c*) each could be implemented by using participating company trucks.

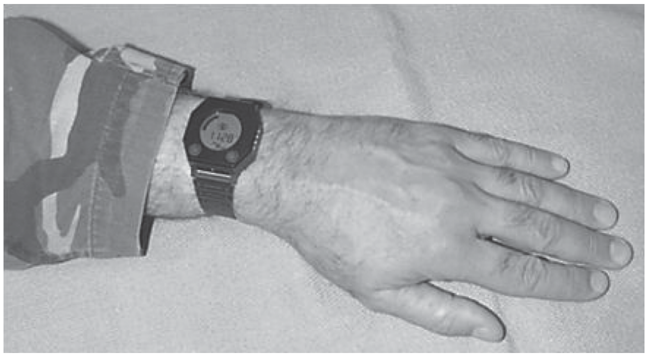

FIGURE 1 WRAIR SleepWatch.

SleepWatch

The technology selected for providing feedback to drivers on sleep need was the actigraphically based, wrist-worn SleepWatch (Precision Control Design, Inc., Ft. Walton Beach, Florida) shown in Figure 1, combined with an internal algorithm called the sleep management model from Walter Reed Army Institute of Research (WRAIR). Investigators at WRAIR developed the wrist-worn actigraph device used and the algorithm to detect sleep in actigraphy data (*4*, p. 149; 5–*8*). Wrist-worn actigraphic monitoring of drivers' rest–activity patterns, with feedback regarding estimated sleep need, was judged to be a potentially useful objective way to inform drivers of the development of cumulative sleep debt (*9*–*11*) and the need to obtain more sleep or take additional alertness-promoting countermeasures. SleepWatch displayed a clock and an analogue "performance fuel gauge" based on sleep need. When a button was pressed, an estimated numeric value of performance readiness was displayed as a percentage of from 0% to 100% performance (see Figure 1). The feedback aspects of the SleepWatch (i.e., the performance fuel gauge and the numeric value of performance readiness) were suppressed in the control (no-feedback) condition although objective data on sleep time were still collected by using the sleep management model.

CoPilot

The technology selected for providing drowsiness feedback to drivers was the CoPilot system (Attention Technologies, Pittsburgh, Pennsylvania) for monitoring percent eyelid closure (PERCLOS). USDOT-funded research in the laboratories of Wierwille et al. (*12*–*14*), Dinges et al. (*1*), Mallis et al. (*15*), and Dinges et al. (*16*) led to the discovery that slow eyelid closures were a highly reliable measure of lapses of attention caused by sleepiness or drowsiness, which led to the development of CoPilot, an infrared-based retinal reflectance monitor for eye closure detection by R. Grace of Carnegie Mellon University. CoPilot used a structured illumination approach and identified a driver's eyes by using two identical images with different sources of infrared illumination. The image of the face was passed through a beam splitter that reflected the image onto the lenses of a camera with an 850-nm filter and a camera with a 950-nm filter. The 850-nm filter yielded a bright-eye camera image (i.e., distinct glowing of the driver's pupils), as seen in Figure 2*a*. The 950-nm filter yielded a dark-eye image, as seen in Figure 2*b*. A third image enhanced the bright eyes by calculating the difference of the two images (Figure 2*c*). A driver's eyes were identified in this third image by applying a threshold determined adaptively by examining the average brightness in each video frame. The CoPilot infrared retinal reflectance device requires it to be operated at low ambient light levels. It was mounted on the truck dashboard, typically just to the right of the steering wheel (Figure 3). Feedback from the system was provided on a separate digital display box and consisted of a CoPilot proprietary algorithm score from 0 to 99, in which 0 indicated maximum eyelid closure and 99 indicated least eyelid closure. Eyelid closure feedback information was active during the 2 weeks drivers operated in the feedback condition. The numeric feedback from the PERCLOS system was disabled during the no-feedback condition, but PERCLOS information was still being recorded for analyses.

SafeTRAC

The technology selected for providing lane tracking feedback to drivers was SafeTRAC (Applied Perception and AssistWare Technology, Wexford, Pennsylvania). Lane tracking, which refers to monitoring the position of the vehicle in the driving lane and detection of lane drifting, weaving, or variability in tracking the lane, is a well-established measure of driving performance with a long history of use. In addition to having excellent face validity in driving safety, many studies of fatigue-related driving deficits have found variability in lane tracking to be one of the more sensitive measures of drowsiness and fatigue. SafeTRAC consisted of a video camera mounted on the windshield (Figure 4) and coupled to a small computer that continuously analyzed the image of the road, lane markings, and other roadway features. Lane departures, erratic movements, and other possible errors were detected. Intentional lane shifts indicated by the turn signal were designed to be ignored by the system. The SafeTRAC feedback monitor was mounted on the dashboard just to the left of the

(a) **(b)** **(c)**

FIGURE 2 Eye images taken by CoPilot: (*a*) bright-eye, (*b*) dark-eye, and (*c*) difference images.

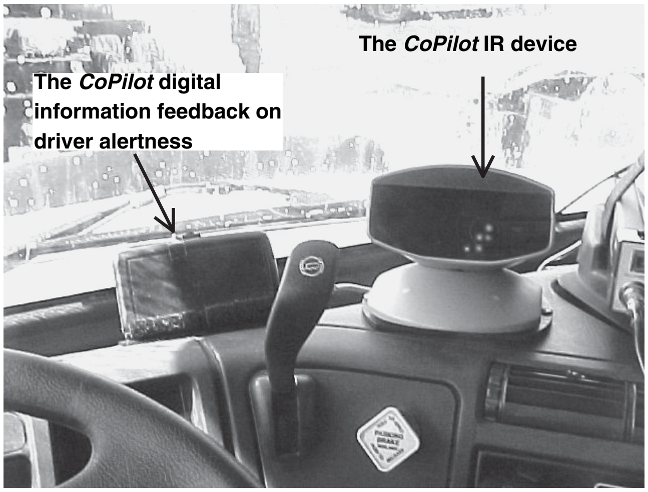

FIGURE 3 CoPilot infrared retinal reflectance monitor.

steering wheel. Feedback from the system consisted of a 0-to-99 scale, in which 0 indicated most erratic lane tracking and 99 indicated least erratic lane tracking, according to a proprietary algorithm. If a driver made an abrupt deviation from the lane without signaling, SafeTRAC provided an auditory warning signal. As with other FMT technologies, feedback information from the SafeTRAC device was active during the 2 weeks drivers operated their trucks in the feedback condition. The numeric feedback from the system was disabled during the 2-week no-feedback period, although it still collected objective data on lane tracking.

Howard Power Center Steering System

The technology selected for reducing the physical work of controlling vehicle stability while driving was the Howard Power Center Steering (HPCS) system (River City Products, San Antonio, Texas). Unlike the other FMT technologies that were designed to provide feedback to drivers on behavioral alertness relative to fatigue based in sleep and circadian biology, the HPCS system was designed to lessen physical fatigue associated with drivers fighting the steering wheel in cross winds. Heavy-vehicle stability and control problems contribute to the work of driving a truck, inducing fatigue because of the often continuous amount of driver steering corrections needed to counteract the unstable behavior of the castered truck wheels. The physical workload associated with fighting the steering wheel in cross winds is particularly fatiguing to neck and shoulder muscles. There was a need to determine whether a technology that lessened this physical workload on drivers would result in less fatigue. The technology that best fulfilled this requirement and that was tested in the pilot study was the HPCS system. HPCS involved a hydraulic device attached to a truck's tie rod and steering system to reduce the physical demands of driving. The system consisted of two principal components: the hydraulic power centering cylinder and the air-activated hydraulic pressure accumulator. The normal operation of the system was automatic and required little attention from the driver. The driver controlled the desirable hydraulic pressure on a panel by adjusting air pressure, which increased or decreased effectiveness of the system. The system was turned on and off by the driver via a switch the driver pressed to release air pressure in the accumulator. Unlike the SleepWatch, the CoPilot drowsiness monitor, and the SafeTRAC lane tracker, HPCS did not provide numeric feedback. Rather, this system was turned on in the feedback condition, and it was off in the no-feedback condition. When the system was turned on, drivers could feel the steering wheel stability relative to when the system was turned off. As with the measurements made by other FMT technologies, steering wheel variability was recorded electronically in both the feedback (HPCS turned on) and no-feedback (HPSC turned off) conditions. Figure 5 displays HPCS as used in the project trucks.

FIGURE 4 SafeTRAC lane-tracking monitor.

Other Non-FMT Data Recording Technologies

Volunteer drivers' trucks were instrumented with the Accident Prevention Plus (AP+) onboard recording device (black box) to continuously record a range of truck motion variables (speed, lateral acceleration, etc.) as well as information from three of the FMT devices (CoPilot, SafeTRAC, HPCS). Volunteer drivers completed a daily diary on work–rest activities and performed the 10-min psychomotor vigilance task (PVT) (*17*) twice daily—midway in each trip and at the end of each trip—as an independent validation of behavioral alertness levels.

Education on Alertness and Fatigue Management

In addition to training in the use of all these technologies, drivers received education on alertness and fatigue management before driving the instrumented trucks at the beginning of the 2-week FMT no-feedback portion of the study and at the beginning of the 2-week FMT feedback portion of the study. Drivers were provided a 3-h course entitled "Mastering Alertness and Managing Driver Fatigue" (sponsored by FMCSA and ATRI), which was prepared for this study (*18*). The course was taught to four drivers at a time, 2 to 3 days before they were issued an instrumented truck. The education module encouraged drivers to be responsible for alertness levels at all times throughout the study. Since all drivers in the study received it as part of risk mitigation, it was not varied between feedback and no-feedback conditions. It likely increased drivers' acceptance of the FMT.

Human Factors Structured Interview Questionnaire

Following completion of the study, drivers were debriefed and completed the human factors structured interview questionnaire, in which

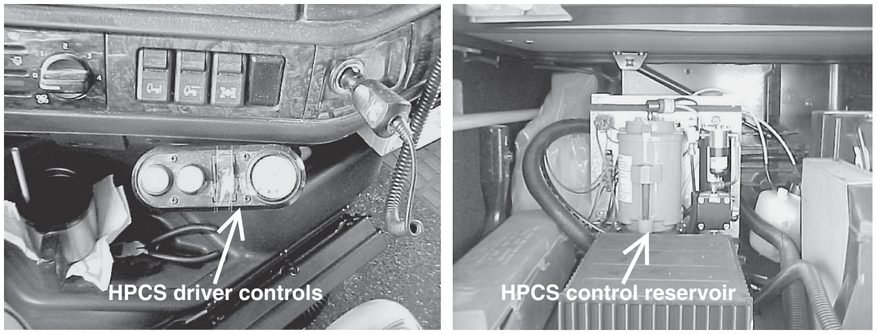

FIGURE 5 HPCS.

they reported reactions to all interventions, measures, and technologies used in the study.

Experimental Design

A within-subjects crossover design was used in both phases (countries) of the study to compare the effects of feedback from combined FMT with no feedback from FMT. The design did not require manipulating or controlling what the participating companies and drivers did, what schedules the drivers adhered to, or what operating practices they followed. Rather, the FMT intervention and data collection were applied to existing routine trucking operations. Thus, for comparisons of the effects of FMT feedback versus no feedback, volunteer drivers served as their own controls—undergoing both conditions under nearly identical circumstances (i.e., a given truck driver drove comparable trucks and schedules during both feedback and no-feedback conditions). A crossover design is efficient and has a number of advantages over an independent-groups design. It ensures roughly the same intersubject variability across both conditions, it provides an opportunity for subjects to explicitly compare and contrast conditions, and it requires fewer subjects than an independent-groups design, which makes it more feasible from both cost and time line perspectives. On the downside, a crossover design necessarily burdens a smaller group of subjects with more recording time than would be the case in an independent-groups design. If too burdensome, subjects may fail to complete all conditions. This occurred to some extent in both phases of the present study but was not a major problem.

The focus of the study was not on comparing Canadian and U.S. operations but rather on comparing drivers during the FMT feedback and no-feedback conditions. Each driver underwent the conditions in the same order: 2 weeks of no feedback (control condition) occurred first, followed by 2 weeks of feedback (intervention condition). Condition order was not counterbalancing because providing the no-feedback condition after the feedback condition would have involved a change in driver behavior carried over from the feedback condition. In contrast, by providing the no-feedback condition first, drivers engaged in normal driving practices for 2 weeks, although driving performance, drowsiness, and sleep need were still recorded by the relevant FMT technologies (i.e., FMT devices were recording but not providing feedback). The no-feedback condition therefore served as a baseline against which the FMT feedback intervention was compared.

Volunteer Drivers

A total of $n = 39$ drivers volunteered for the study ($n = 27$ from Canada, $n = 12$ from the United States). One driver dropped out after being empanelled, which reduced the Canadian sample to $n = 26$ (20 males, six females) and the total sample to $n = 38$. Demographic characteristics of the volunteers as they pertain to truck driving experience are shown in Table 1. More drivers were empanelled than the target sample size of $n = 24$ because of the need to compensate for the loss of data caused by equipment failure. Equipment failure during the 4-week data acquisition study reduced specific comparisons between feedback and no-feedback conditions on some variables to sample sizes ranging between $n = 15$ and $n = 25$ drivers in the Canadian study phase and between $n = 7$ and $n = 12$ drivers in the U.S. study phase. Therefore, when study phases are combined, the hypothesis-testing sample size ranged between $n = 22$ and $n = 38$, depending on the variable being analyzed. As shown in Table 1, most participating drivers were middle-aged males with many years of long-haul driving experience. Drivers were solicited for participation after the protocol, procedures, and informed consents were reviewed and approved by the Canadian Research Ethics Board and by the WRAIR institutional review board.

TABLE 1 Characteristics of Participating Truck Drivers

Country	*n*	Sex	Age Mean (yr)	Age Range (yr)	Years at Company (mean)	Years at Company (range)	Years Driving Large Trucks (mean)	Years Driving Long Haul (mean)	Miles Driven Past Year (mean)
Canada	20	M	45.4	22–58	4.6	< 0.5–17	16.6	11.3	> 109K*
Canada	6	F	35.3	22–50	4.0	< 0.5–15	2.1	1.6	> 76K
U.S.	12	M	46.9	32–57	11.5	6.5–18	23.7	18.0	> 99K
All drivers	38	84% male	44.2	22–58	6.7	< 0.5–18	16.6	11.9	> 100K

*Based on $n = 18$ (data missing from 2 male drivers)

Data Quality Control

Given the extraordinarily large volume of data gathered in the study, it was necessary to determine data management and variable extraction procedures that would ensure quality control of the data. Of particular concern was the need to use procedures that avoided including erroneous data values, especially data corrupted by equipment failure in the field. [Although all the equipment accompanied drivers during 4 weeks of work, no investigator or study technicians were present while drivers were on the road, and hence no one was present to prevent data loss or corruption from equipment damage due to environmental conditions (vibration, heat, cold, rain, snow, ice) in which it was deployed.] Data were carefully segregated into three broad categories: (*a*) all AP+ data with no records excluded, (*b*) AP+ data records in which speed was at least 30 mph, and (*c*) AP+ data for speed ≥30 mph, artifacts eliminated, and records within measurement range. Thus, final cleaned analysis samples from Canada and the United States were defined on the basis of the subset of drivers with sufficient data under both conditions (feedback and no feedback), restricting attention to records recorded at speeds of at least 30 mph, after excluding additional data found to be invalid, following careful examination of driver-specific distributions.

Study Phase 1 took place under Canadian HOS and involved a Canadian trucking company. Volunteer drivers operated single tractor-trailer units with sleeper berths, and approximately 26% of their driving was conducted during nighttime hours (74% in daylight hours). Study Phase 2 took place under U.S. HOS and involved a U.S. trucking company. Volunteer drivers operated tandem tractor-trailer units without sleeper berths, and approximately 93% of their driving was conducted during nighttime hours (7% in daylight hours). The differences between the Canadian and U.S. trucking companies were in part a function of which companies agreed to be part of the study as well as the study goal to expressly study companies for which night driving was both a minority (Study Phase 1) and a majority (Study Phase 2) of trucking operations. For these reasons, the Canada study phase and the U.S. study phase were analyzed separately for the effects of FMT feedback on driving and alertness outcomes before being combined.

Statistical Methods

For each outcome variable recorded by the AP+ system, four analyses were performed to assess if there was a significant change from the no-feedback condition to the feedback condition within Study Phase 1 in Canada and again within Study Phase 2 in the United States. The first of the statistical methods was unweighted analysis for means and standard deviations values across all records for a specific driver under a specific condition (no feedback and feedback). Mean values were compared for the following outcome variables: CoPilot measures of PERCLOS during night hours and SafeTRAC alertness score. Standard deviations were compared for lateral distance, steering wheel movements, and front wheel movements. Then within-driver change scores were computed between no-feedback and feedback conditions. Paired *t*-tests were performed to assess the statistical significance of the changes in means or standard deviations as appropriate.

The second statistical method introduced two weighting factors. First, when the within driver and condition mean, median, standard deviation, and interquartile range values were computed, records were replicated if they corresponded to more than 1 s in duration. In this way, records with durations that were 3 s contributed a weight three times greater than records with durations of 1 s. Even accounting for record duration, drivers varied greatly for total duration of data in the cleaned analysis sample. Drivers with greater total durations under both conditions contributed more information about intervention effects. In contrast, a driver with a short duration under one of the conditions contributes less information about within-driver changes. To account for this, and to optimize the ability to consider both within-subjects and between-subjects sources of variance, mixed model analyses of variance were used to compare mean (duration weighted) values between the no-feedback and feedback conditions, weighting by the total number of available records (separately by condition). All mixed model analyses were implemented by using the Proc. Mixed procedure available in SAS.

The analyses were repeated to summarize the no-feedback and feedback distributions of CoPilot PERCLOS during night hours and SafeTRAC alertness score by median values rather than mean values, to provide summaries of the center of these distributions that are less sensitive to outliers and skewness. Similarly, AP+ lateral distance, AP+ steering wheel movements, and AP+ front wheel movements were summarized by using interquartile ranges (IQR) instead of standard deviations. The IQR is defined as the difference between the 75th percentile value and the 25th percentile value and is less influenced by extreme values than the standard deviation. Both the paired *t*-test and the mixed model weighted analyses were performed on the median and the interquartile range for each variable (which are the nonparametric alternatives to the mean and standard deviation).

Mixed model analysis of variance was used to assess the significance of the intervention effect (no feedback versus feedback), controlling for time-of-day category (day, evening, night). The initial model included fixed effects for time of day (morning, evening, night), presence versus absence of feedback, and time of day by feedback interaction. It also included a random effect for driver to account for correlations within driver. The interaction model (i.e., feedback condition, time of day, time of day by feedback condition) was used to compute an adjusted intraclass correlation (ICC). The ICC is the proportion of total variance explained by systematic differences among drivers after accounting for time-of-day and feedback condition effects. The model used to determine the ICCs was used to examine whether differences between responses obtained during the no-feedback and feedback conditions varied by time of day. A *p*-value of 0.10 was used because of the low power inherent in tests for interaction. If $p \geq 0.10$, then the interaction terms were removed from the model and the feedback effects and time-of-day effects were tested as main effects in the ANOVA model. If $p < 0.10$, it was concluded that differences between the no-feedback and feedback conditions significantly varied by time of day. Therefore, separate mixed models were used to test for feedback effects at each time-of-day interval (day, evening, night). Daily mean values were analyzed for variables derived from Sleep-Watch. Mixed model analyses of variance were used to assess the significance of the fixed intervention effect. Random effects included between- and within-driver variance, which were used to compute ICCs. Descriptive statistics were used to analyze the drivers' daily diary and postexperimental responses to the human factors structured interview questionnaire.

RESULTS

Data from the FMT devices and other driving performance variables gathered on the AP+ black box recorder every second the trucks were operating for the 28 days each driver was in the study resulted in 8,737,705 total records among the 38 drivers in the combined study phases, which reduced to 6,683,855 data records among 29 drivers

(Canada, $n = 20$; United States, $n = 9$), when data analyses were confined to artifact-free records in which speed was at least 30 mph (i.e., highway driving). Equipment failure resulted in a loss of approximately 25% of the data. Even with this attrition, the data set and remaining sample sizes were adequate for hypothesis testing. Although rough road conditions in the operating trucks caused some data loss, the final data set was among the most extensive on truck driver alertness and truck performance ever recorded. In addition, data acquired from the drivers' daily diaries, their 933 PVT performance tests, their 1.2 million minutes of SleepWatch actigraphic data, and their extensive responses and comments to the human factors structured interview questionnaire resulted in millions of additional data records. Many of the latter variables could be analyzed by using all 38 drivers who completed the study. Key findings are summarized here relative to the primary hypotheses and to other key findings and recommendations relevant to fatigue management in long-haul trucking.

Hypothesis I: FMT Feedback Will Improve Driver Alertness or Reduce Driver Drowsiness or Both

Phase I: Canadian Drivers

There was marginal evidence to support the hypothesis that FMT feedback will improve driver alertness or reduce driver drowsiness. Drowsiness as measured by the CoPilot index of PERCLOS during night hours was modestly lower under the feedback condition compared to the no-feedback condition ($p = 0.094$). Drivers' subjective sleepiness ratings taken before and after PVT performance tests at night also indicated they were less sleepy ($p = 0.009$), although Canadian drivers spent only a minority of time in night driving. However, the SafeTRAC index of driver alertness and drivers' PVT performance lapses during daytime trials showed effects opposite those found for nighttime driving. There was a slight reduction in SafeTRAC alertness during the daytime in the feedback condition relative to the no-feedback condition among Canadian drivers ($p = 0.013$) and an elevation of PVT lapses ($p = 0.0004$). Hence there was no consistent finding in support of Hypothesis I in the Phase 1 data.

Phase 2: U.S. Drivers

There was evidence in support of Hypothesis I in the Phase 2 data. This phase focused more extensively on drivers who primarily drove at night (73% of the time), when sleepiness would be expected to be more of a problem. There was clear evidence of greater alertness in the feedback condition during night driving than in the no-feedback condition at night from both the SafeTRAC index of driver alertness ($t = 2.67$, $df = 8$, $p = 0.028$) and the CoPilot index of PERCLOS ($t = 2.70$, $df = 8$, $p = 0.027$). Although only a statistical trend, lane tracking variability also improved with feedback during night driving in the U.S. study phase ($p = 0.083$).

Combined Canadian and U.S. Data

Composite results from pooling data from the two study phases yielded strong support for Hypothesis I. During night driving, feedback from fatigue management technologies significantly reduced slow eyelid closures (PERCLOS) as measured by CoPilot ($t = –3.24$, $n = 25$, $p = 0.004$), increased the SafeTRAC estimate of driver alertness ($t = 3.49$, $n = 24$, $p = 0.002$), and decreased lane tracking variability ($t = –2.96$, $n = 24$, $p = 0.007$).

Hypothesis II: FMT Feedback Will Increase Driver Sleep Time

Phase 1: Canadian Drivers

Within the Canada study phase, none of the SleepWatch actigraphy outcomes demonstrated systematic differences between the no-feedback and feedback conditions. There was also no evidence from drivers' daily diaries to support the hypothesis that FMT feedback resulted in increased sleep time relative to no feedback.

Phase 2: U.S. Drivers

Within the U.S. study phase, there was a significant increase in the number of SleepWatch actigraphically identified sleep episodes but not sleep duration in the feedback condition relative to the no feedback. There was also no evidence from drivers' daily diaries of increased sleep time.

Combined Canadian and U.S. Data

There was no support for Hypothesis II when SleepWatch data were combined between study phases.

Sleep on Workdays Versus Nonworkdays

Not surprisingly, drivers in both countries slept significantly more on nonworkdays than on workdays. During the no-feedback 2-week period of the Canadian study phase, drivers averaged 7 h 17 min of sleep per 24-h period on nonworkdays compared to 6 h 15 min on workdays, a mean difference of 1 h 2 min ($p = 0.023$). Similarly, during the feedback 2-week period of the Canadian phase, drivers averaged 7 h 31 min of sleep per 24 h on nonworkdays compared to 6 h 14 min on workdays, a mean difference of 1 h and 17 min ($p = 0.0005$). Comparable results were obtained in the U.S. study phase. During the no-feedback 2-week period, U.S. drivers averaged 6 h 32 min of sleep per 24 h on nonworkdays compared to 5 h 14 min on workdays, a mean difference of 1 h 18 min ($p = 0.018$). Similarly, during the feedback period, U.S. drivers averaged 7 h 32 min sleep compared to 5 h 1 min on workdays, a mean difference of 2 h 31 min ($p = 0.0004$). These are relatively large differences in 24-h sleep durations, suggesting that drivers developed sleep debts across the work week.

Effect of FMT Feedback on Nonworkdays Sleep

Although mean sleep duration was significantly less for U.S. drivers compared to Canadian drivers ($F_{1,28} = 7.50$, $p = 0.011$), when SleepWatch actigraphically identified sleep duration per 24 h was analyzed for both study phases, separating workdays and nonworkdays, there was clear evidence in support of Hypothesis II. In contrast to workdays, for which FMT feedback had no effect on sleep time, there was a significant increase in mean sleep duration during nonworkdays in the feedback condition relative to no feedback in both the Canadian drivers ($t = –2.55$, $df = 15$, $p = 0.023$) and the U.S. drivers ($t = –2.88$, $df = 10$, $p = 0.018$). Drivers in both study phases increased their nonworkday sleep durations by an average of 45 min per day over sleep duration on nonworkdays in the no-feedback condition ($F_{1,25} = 4.39$, $p = 0.046$).

Other Key Findings

Cost for Being More Alert with FMT Feedback?

As summarized, during FMT feedback, alertness improved significantly during driving in the U.S. study phase, which involved driving at night 93% of the time. However, there was also consistent evidence that PVT performance worsened and subjective sleepiness ratings increased during the feedback period of the U.S. study relative to the no-feedback period. U.S. drivers' nighttime PVT performance lapses per trial during the no-feedback and feedback conditions averaged 3.12 and 4.59, respectively ($t = 2.83$, df = 11, $p = 0.016$). Similar findings were obtained during daytime driving periods in the Canada study phase, when 74% of driving occurred. During daytime PVT test trials, the mean number of lapses per trial during the no-feedback and feedback conditions was 1.95 and 3.89, respectively ($t = 4.49$, df = 16, $p = 0.0004$). The feedback condition was also associated with slower median PVT reaction times during night driving in the U.S. phase ($t = 5.14$, df = 11, $p < 0.0001$) and during day driving in the Canada phase ($t = 3.54$, df = 16, $p = 0.003$). Drivers' ratings of their sleepiness on a post-PVT visual analogue scale also revealed greater sleepiness in the feedback condition than in the no-feedback condition during nighttime PVT tests of the U.S. study phase (3.29 versus 5.33; $t = 6.63$, df = 11, $p < 0.0001$). These findings suggest that FMT feedback in drivers who operate primarily at night may have alertness-promoting benefits during driving, but such feedback may also create a modest cost for the added effort (in attention and compensatory behaviors) required to respond to the information from the devices, and cost may manifest as slightly worse performance and greater subjective sleepiness when drivers perform a demanding vigilance-based reaction time task such as the PVT (while not driving).

Do Drivers Prefer Vehicle-Based Measures of Alertness?

In general, drivers agreed that commercial drivers would benefit from fatigue management aids (Canada, 88%; United States, 100%). Descriptive analyses of driver responses to the human factors structured interview questionnaire at the end of the 2-week no-feedback period, and again at the end of the 2-week feedback condition period, revealed clear preferences of both Canadian and U.S. drivers for fatigue management training and certain fatigue management technologies. Drivers were uniformly positive about the education on alertness and fatigue management course given at the beginning of each study phase. Among technologies designed to detect alertness or drowsiness, drivers gave higher ratings to SafeTRAC, medium ratings to the SleepWatch, and low ratings to the CoPilot. Among all FMT technologies deployed, however, drivers were significantly more enthusiastic about the benefits of the HPCS system and SafeTRAC than they were about SleepWatch and CoPilot. It is noteworthy that HPCS and SafeTRAC both interface with the vehicle, whereas SleepWatch and CoPilot interface with the driver. It may be that truck drivers prefer fatigue management to be through vehicle monitoring rather than through driver monitoring. More research is needed to understand what influences commercial drivers' attitudes toward feedback by technology (*19*).

Future for FMT Technologies

Overall, participant drivers were positive toward the FMT approach in general and thought that if such technologies could be further improved, they would help manage fatigue and alertness.

RECOMMENDATIONS FOR FUTURE WORK OUTSIDE SCOPE OF PROJECT

Continue Development of FMT Technologies

There is enough evidence to support the case for continued development of FMT technologies. However, these should not be used only for driver monitors. Vehicle-based monitoring should also get increased attention, as truck drivers appear to have some preference for this mode of fatigue management.

Drivers Want Alertness and Fatigue Management Courses

Despite differences in country of operation, hours of service, type of trucks, and many other factors, U.S. and Canadian drivers had surprisingly similar views toward the FMT project. They were positive toward the alertness and fatigue management training course provided in the study. Postexperimentally, drivers rated the course content and knowledge gained as "good" to "very helpful" (highest rating); 83% to 96% indicated the course lessons were used by them during the FMT study and that they intended to continue to use them. Qualitative comments from drivers indicated they perceived benefit from the course and would like to have more of this type of didactic to help teach them how to manage fatigue. This is impressive given that these were largely seasoned long-haul drivers who appeared not to be inhibited about reporting that they can still learn about fatigue and ways to manage it. These positive views toward fatigue management training suggest that some segments of the trucking industry are likely to welcome fatigue management programs.

PVT Should Be Developed as a Fitness-for-Duty Test

Although PVT was not discussed with drivers as either an FMT technology or a fitness-for-duty test, a majority of drivers in both countries indicated when asked that the PVT could be used as a personal checking system on a driver fitness-for-duty system, if it could be reduced in duration. Drivers' generally positive view of the PVT as a potential fitness-for-duty device suggests that efforts should be made to attempt to validate the sensitivity of and positive and negative predictability of a shorter-duration PVT test (e.g., 3 to 5 min) relative to truck driver fatigue.

Barriers to Drivers Obtaining Adequate Sleep During Workdays Must Be Identified

One of the more striking outcomes was the finding that drivers in both countries were routinely averaging between 5 h and 6.25 h of sleep per day during workdays, despite very different work schedules. Recent scientific work, some of it by USDOT on volunteer truck drivers, shows that severe sleep debt and deficits in behavioral alertness can develop within a few days at these sleep durations. That project participants markedly increased sleep durations on nonworkdays also supports the view that they were suffering sleep debts. Much more must be understood about the factors that determine when and where drivers obtain sleep on workdays and nonworkdays, the barriers to obtaining adequate sleep on workdays, and the factors that convince drivers to get more recovery sleep on nonworkdays.

ACKNOWLEDGMENTS

The FMT pilot test was sponsored by FMCSA and Transport Canada (TC), in cooperation with ATRI. Project leadership was provided by Robert J. Carroll of FMCSA, Sesto Vespa of TC, and Rebecca Brewster of ATRI. The principal investigator was David F. Dinges of the University of Pennsylvania School of Medicine. The principal biostatistician was Greg Maislin of Biomedical Statistical Consulting. The following contributed to the research: Robert Hachadoorian of Biomedical Statistical Consulting; Adrian Ecker, Donald Terry, and John Powell of the University of Pennsylvania School of Medicine; Daniel Redmond, Thomas Balkin, Gregory Belenky, and Greg Lounsberry of WRAIR; Todd Jochem and Dean Pomerleau of Applied Perception and Assistware Technology; Richard Grace and Robert Engel of Attention Technologies; B. Red Chester, Dee Howard, and Ruben Clayton of River City Products; Allan Schwartz, Jean Paul Daveau, Patrick Painter, and Pierre Pommarel of Accident Prevention Plus; Daniel Einwechter, Enno Jakobson, Robert Halfyard, Paul Clague, and Steve Kratz of Challenger Motor Freight; Robert Petrancosta, Edwin Gilmore, Keith Peters, Joseph Fyda, Mark Stets, Farris Scott, Anthony Martucci, and Jarrod Stokes of Con-Way Central Express; and William C. Rogers and Ronald Knipling.

REFERENCES

1. Dinges, D. F., M. Mallis, G. Maislin, and J. W. Powell. *Evaluation of Techniques for Ocular Measurement as an Index of Fatigue and the Basis for Alertness Management.* NHTSA, U.S. Department of Transportation, 1998, pp. 1–113.
2. Krueger, G. *Technologies and Methods for Monitoring Driver Alertness and Detecting Driver Fatigue: A Review Applicable to Long-Haul Truck Driving.* Federal Motor Carrier Safety Administration, U.S. Department of Transportation, June 2004.
3. Dinges, D. F., and M. M. Mallis. Managing Fatigue by Drowsiness Detection: Can Technological Promises Be Realized? In *Managing Fatigue in Transportation,* Elsevier Science, Kidlington, Oxford, England, 1998, pp. 209–229.
4. Redmond, D. P., H. C. Sing, G. W. Harding, and F. W. Hegge. Human Rest Activity Rhythms and Wrist Movement Recording-Systems. *Chronobiologia,* Vol. 10, No. 2, 1983, p. 149.
5. Redmond, D. P., and F. W. Hegge. Observations on the Design and Specification of a Wrist-Worn Human Activity Monitoring System. *Behavior Research Methods, Instruments, and Computers,* Vol. 17, No. 6, 1985, pp. 659–669.
6. Redmond, D. P., H. C. Sing, G. W. Harding, and F. W. Hegge. Disturbed Human Rest Activity Cycles: Automated Data Reduction and Rhythmic Analysis of Wrist Activity Recordings. *Chronobiologia,* Vol. 12, No. 3, 1985, pp. 267–267.
7. Belenky, G., T. J. Balkin, D. P. Redmond, H. C. Sing, M. L. Thomas, D. R. Thorne, and N. J. Wesensten. Sustained Performance During Continuous Operations: The US Army's Sleep Management System. *Managing Fatigue in Transportation: Proceedings of the 3rd Fatigue in Transportation Conference, Fremantle, Western Australia* (L. Hartley, ed.), Elsevier Science, Kidlington, Oxford, England, 1998, 77–85.
8. Hursh, S. R., D. P. Redmond, M. L. Johnson, D. R. Thorne, G. Belenky, T. J. Balkin, W. F. Storm, J. C. Miller, and D. R. Eddy. Fatigue Models for Applied Research in Warfighting. *Aviation Space and Environmental Medicine,* Vol. 75, No. 3, Supp. 2, 2004, pp. A44–A53.
9. Dinges, D. F., F. Pack, K. Williams, K. A. Gillen, J. W. Powell, G. E. Ott, C. Aptowicz, and A. I. Pack. Cumulative Sleepiness, Mood Disturbance, and Psychomotor Vigilance Performance Decrements During a Week of Sleep Restricted to 4–5 Hours Per Night. *Sleep,* Vol. 20, 1997, pp. 267–277.
10. Van Dongen, H. P. A., G. Maislin, J. M. Mullington, and D. F. Dinges. The Cumulative Cost of Additional Wakefulness: Dose-Response Effects on Neurobehavioral Functions and Sleep Physiology from Chronic Sleep Restriction and Total Sleep Deprivation. *Sleep,* Vol. 26, 2003, pp. 117–126.
11. Belenky, G., N. J. Wesensten, D. R. Thorne, M. L. Thomas, H. C. Sing, D. P. Redmond, M. B. Russo, and T. J. Balkin. Patterns of Performance Degradation and Restoration During Sleep Restriction and Subsequent Recovery: A Sleep-Dose Response Study. *Journal of Sleep Research,* Vol. 12, 2003, pp. 1–12.
12. Wierwille, W. W., and L. A. Ellsworth. Evaluation of Driver Drowsiness by Trained Raters. *Accident Analysis and Prevention,* Vol. 26, No. 5, 1994, pp. 571–581.
13. Wierwille, W. W., L. A. Ellsworth, S. S. Wreggit, R. J. Fairbanks, and C. L. Kirn. *Research on Vehicle-Based Driver Status/Performance Monitoring: Development, Validation, and Refinement of Algorithms for Detection of Driver Drowsiness.* DOT-HS-808-247. NHTSA, U.S. Department of Transportation, 1994.
14. Wierwille, W. W. Historical Perspective on Slow Eyelid Closure: Whence PERCLOS? In *Technical Proceedings of Ocular Measures of Driver Alertness Conference* (R. J. Carroll, ed.). MC-99-136. FHWA, U.S. Department of Transportation, 1999, pp. 31–53.
15. Mallis, M., G. Maislin, N. Konowal, V. Byrne, D. Bierman, R. Davis, R. Grace, and D. F. Dinges. *Biobehavioral Responses to Drowsy Driving Alarms and Alerting Stimuli. Final Report to Develop, Test and Evaluate a Drowsy Driver Detection and Warning System for Commercial Motor Vehicle Drivers,* NHTSA, U.S. Department of Transportation, 2000.
16. Dinges, D. F., N. J. Price, G. Maislin, J. W. Powell, A. J. Ecker, M. M. Mallis, and M. P. Szuba. Prospective Laboratory Re-Validation of Ocular-Based Drowsiness Detection Technologies and Countermeasures. In *Drowsy Driver Detection and Interface Project.* DTNH 22-00-D-07007. NHTSA, U.S. Department of Transportation, 2002.
17. Dinges, D. F., and J. W. Powell. Microcomputer Analyses of Performance on a Portable, Simple Visual RT Task During Sustained Operations. *Behavior Research Methods, Instruments, and Computers,* Vol. 17, No. 6, 1985, pp. 652–655.
18. O'Neill, T. R., G. P. Krueger, and S. B. Van Hemel. *Fatigue and the Truck Driver: Instructor's Guide for a Fatigue Outreach Training Course for America's Trucking Industry: A 4-hr Course for Safety and Risk Managers.* American Trucking Associations, Alexandria, Va., 1996.
19. Roetting, M., Y. H. Huang, J. McDevitt, and D. Melton. When Technology Tells You How You Drive: Truck Drivers' Attitude Toward Feedback by Technology. *Transportation Research Part F: Traffic Psychology and Behaviour,* Vol. 6, No. 4, 2003, pp. 275–287.

The Truck and Bus Safety Committee sponsored publication of this paper.