PILOT TEST OF FATIGUE MANAGEMENT TECHNOLOGIES

David F. Dinges, PhD **Paper # 05-1234**
Professor of Psychology in Psychiatry
Chief, Division of Sleep and Chronobiology
Director, Unit for Experimental Psychiatry
University of Pennsylvania
School of Medicine
1013 Blockley Hall
423 Guardian Drive
Philadelphia, PA 19104
Phone: (215) 898-9949 Fax: (215) 573-6410
dinges@mail.med.upenn.edu

Greg Maislin, M.S., M.A.
Principal Biostatistician
Biomedical Statistical Consulting
1357 Garden Road
Wynnewood, PA 19096
Phone: (610) 645-5708 Fax: (610) 645-5734
gmaislin@biomedstat.com

Rebecca M. Brewster (Corresponding Author)
President and COO
American Transportation Research Institute
1800 Lake Park Drive, Suite 123
Smyrna, GA 30080
Phone: (770) 432-0628 Fax: (770) 432-0638
rbrewster@trucking.org

Gerald P. Krueger, PhD, CPE
Principal Scientist/Ergonomist
Krueger Ergonomics Consultants
4105 Komes Court
Alexandria, VA 22306-1252
Phone: (703) 768-3421 Fax: (703) 768-3421
JerryKrueg@aol.com

Robert J. Carroll, MS, CPE
U.S. Department of Transportation
Federal Motor Carrier Safety Administration
Office of Research and Technology
400 Virginia Avenue, SW, Suite 600
Washington, DC 20024
Phone: (202) 385-2388 Fax: (202) 385-2422
robert.carroll@fmcsa.dot.gov
WORD COUNT: 7,158

ABSTRACT

This study involved over-the-road testing of a set of four fatigue management technologies (FMT) in trucking operations in Canada and the United States. Technologies bundled into a single intervention came from four domains of fatigue management: 1) one that provided objective information on driver sleep need; 2) one that provided objective information on driver drowsiness; 3) one that provided objective information on lane tracking performance; and 4) one that reduced the work involved in controlling vehicle stability while driving. The objective was to determine how drivers reacted to such technologies, and whether FMT FEEDBACK would improve their alertness, especially during night driving, and/or increase their sleep time on workdays and/or non-workdays. A within-subjects cross-over design was used to compare the effects of FMT FEEDBACK to NO-FEEDBACK. Each driver underwent the two conditions in the same order: 2-weeks of NO FEEDBACK (control) followed by 2-weeks of FMT FEEDBACK (intervention). Data from the FMT devices and other driving performance variables were recorder every second the trucks were operating for the 28 days each driver was in the study, resulting in 8.7 million data records among the 38 drivers in the combined study phases. Support was found for FMT effects. During night driving, FMT FEEDBACK significantly reduced drivers' drowsiness ($p = 0.004$) and lane tracking variability ($p = 0.007$). However, there was evidence from probed PVT testing that these improvements may have had "cost" due to the effort (in attention and compensatory behaviors) required to respond to the information from the devices. In general, participants agreed that commercial drivers would benefit from FMT, but they were more positive about technologies that involved vehicle monitoring more so than driver monitoring.

INTRODUCTION

There are currently 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 at the U.S. Department of Transportation, 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 and Transport Canada, in cooperation with the American Transportation Research Institute, the study was tasked to develop an experimental design and instrumentation plan, and conduct a pilot test of commercial truck drivers' reactions to a combination of FMT, under current Federally-mandated hours-of-service in both Canada and the U.S. 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 in a single field trial that had two phases – one completed in Canada and a second phase completed in the U.S. The project involved an extensive over-the-road test of the combined set of 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 them with information feedback about changes in sleep need, in drowsiness, and in driving performance during their routine driving schedules. Specifically, the research sought to determine whether feedback from combined FMT would enhance drivers' alertness and performance at work, and increase their sleep times on workdays and/or non-work days. A secondary specific aim was to obtain drivers' reactions to the 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.

METHODS

CRITERIA FOR FMT SELECTION — Building on previous work at the U.S. Department of Transportation, the project bundled into a single intervention technologies from four domains of fatigue management: 1) a technology that provided objective information on driver sleep need; 2) a technology that provided objective information on driver drowsiness; 3) a technology that provided objective information on lane tracking performance; and 4) a technology that reduced the work involved in controlling vehicle stability while driving. Although each technology is described separately below, 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 fatigue management technologies to each other, or to compare the effects of FMT in Canadian versus U.S. drivers (except by way of establishing consistent findings in both countries). The selection of specific technologies was also not an endorsement of their validity or reliability. We also did not require that technologies were either under development or commercially available. Technologies were selected for use in the pilot study because 1) each was representative of one of the four domains of fatigue management; 2) each was available for study through the cooperation of their respective developers; and 3) each could be implemented with the trucks of the participating companies.

SLEEPWATCH® — The technology selected for providing feedback to drivers on their need for sleep was the actigraphically-based, wrist-worn *SleepWatch®* (Precision Control Design, Inc., FL) shown in Figure 1, combined with an internal algorithm entitled the "Sleep Management Model" from the Walter Reed Army Institute of Research (WRAIR). Investigators at WRAIR developed the wrist-worn actigraph device used and algorithm to detect sleep in actigraphy data.[4-8] 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 and/or take additional alertness-promoting countermeasures. *SleepWatch®* displayed a clock and an analog "performance fuel gauge" based on sleep need. When a button was pressed an estimated numeric value of "Performance-Readiness" was displayed as a percentage from 0-100% 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 (see STUDY DESIGN below) while still collecting objective data on sleep time using the "Sleep Management Model."

COPILOT® — The technology selected for providing feedback to drivers on their drowsiness was the *CoPilot®* (Attention Technologies, Pittsburgh, PA) system for monitoring percent eyelid closure (PERCLOS). U.S. DOT-funded research in the laboratories of Wierwille[12-14] and Dinges[1,15-16] led to the discovery that slow eyelid closures were a highly reliable measure of lapses of attention due to sleepiness/drowsiness which led to the development of *CoPilot®*, an infrared-based retinal reflectance monitor for eye closure detection by R. Grace at Carnegie Mellon University. *CoPilot®* used a structured illumination approach and identified a driver's eyes 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 one 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 2A. The 950 nm filter yielded a "dark-eye" image as seen in Figure 2B. A third image enhanced the bright eyes by calculating the difference of the two images (Figure 2C). 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 dashboards of trucks, 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, where 0 indicated maximum eyelid closure and 99 indicated least eyelid closure. Eyelid closure feedback information was active during the 2-weeks drivers operated their trucks 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 feedback to drivers on their lane tracking was the *SafeTRAC®* (Applied Perception and AssistWare Technology, Inc., Wexford, PA). 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 lane tracking 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 steering wheel. Feedback from the system consisted of a 0 to 99 scale, where 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®* also 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 while still collecting 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, Inc., San Antonio, TX). Unlike the other FMT technologies that were designed to provide feedback to drivers on their 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 due to 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 was tested in the pilot study was the *HPCS* system. The *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. 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, the *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.

OTHER NON-FMT DATA RECORDING TECHNOLOGIES **—** Trucks of volunteer drivers were instrumented with the *Accident Prevention Plus* (*AP+*) on-board 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 also completed a daily diary on their work-rest activities, and performed the 10-minute Psychomotor Vigilance Task (PVT)[17], twice daily—midway in each trip and at the end of each trip—as an independent validation of their level of behavioral alertness.

EDUCATION ON ALERTNESS AND FATIGUE MANAGEMENT **—** In addition to training in the use of all technologies listed above, drivers also received *Education on Alertness and Fatigue Management* before they drove with 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 an approximately 3-hour course entitled “Mastering Alertness and Managing Driver Fatigue,” (sponsored by Federal Motor Carrier Safety Administration and the American Transportation Research Institute), which was prepared for this study and taught by G. Krueger.[18] The course was taught to four drivers at a time, 2-3 days before they were issued their instrumented trucks. The education module encouraged drivers to be responsible for their 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 fatigue management technologies.

HUMAN FACTORS STRUCTURED INTERVIEW QUESTIONNAIRE — Following completion of the study drivers were debriefed and completed the *Human Factors Structured Interview Questionnaire* in which they reported their reactions to all interventions, measures and technologies used in the study.

EXPERIMENTAL DESIGN — A within-subjects cross-over 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 actually followed. Rather, the FMT intervention and data collection were applied to existing routine trucking operations. Thus, for comparisons of the effects of FMT FEEDBACK vs. 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 cross-over design is efficient and has a number of advantages over an independent-groups design. It ensures roughly the same inter-subject 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 timeline perspectives. On the downside, a cross-over 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 Canada and U.S. operations, but rather to comparing drivers during the FMT FEEDBACK and NO-FEEDBACK conditions. Each driver underwent the two conditions in the same order: 2-weeks of the NO FEEDBACK (control condition) occurred first, followed by 2-weeks of the 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 their normal driving practices for 2 weeks, although their 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 U.S.). One driver dropped out after being empanelled, which reduced the Canadian sample to n = 26 (20 males, 6 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 due to the need to compensate for the loss of data due to 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 combining study phases, the hypothesis-testing sample size ranged between n = 22 and n = 38, depending on the variable being analyzed. As shown in Table 1, the majority of participating drivers were middle-aged males with many years experience driving long-haul. 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 Institutional Review Board of The Walter Reed Army Institute of Research.

Table 1: Characteristics of the Canadian and U.S. Truck Drivers Participating in the Pilot Study.

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 last 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
TOTAL	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 utilize procedures that avoided including erroneous data values (especially data corrupted by equipment failure in the field—it is important to keep in mind that while 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 the environmental conditions [e.g., vibration, heat, cold, rain, snow and ice] in which it was deployed). Data were carefully segregated into three broad categories: 1) All AP+® data with no records excluded; 2) AP+® data records in which speed was at least 30 mph; and 3) AP+® data for speed $\geq$ 30 mph, artifacts eliminated and records within measurement range.

Thus, final cleaned analysis samples from both Canada and the U.S. 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 in which 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 in which 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 difference between Canadian and U.S. trucking companies were in part a function of which companies agreed to be part of the study, as well as our goal to expressly study companies in 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 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 the study phase 1 in Canada, and again within study phase 2 in the U.S. The first of the statistical method 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 computing the within driver and condition mean, median, standard deviation, and interquartile range values, records were replicated if they corresponded to more than 1 second in duration. In this way, records with durations that were 3 seconds contributed a weight 3 times greater than records with durations of 1 second. Even accounting for record duration, drivers varied greatly with regard to the total duration of data in the cleaned analysis sample. Drivers with greater total durations under both conditions contribute more information with regard to 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 using the Proc. Mixed procedure available in SAS.

The analyses were repeated summarizing the NO FEEDBACK and FEEDBACK distributions of *CoPilot®* PERCLOS during night hours and *SafeTRAC®* "alertness" score by median values rather than mean values, in order 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 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 influence by extreme values than the standard deviation. Both the paired t-test and 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 analyses of variance was used to assess the significance of the intervention effect (NO FEEDBACK vs. FEEDBACK), controlling for time-of-day category (day, evening, night). The initial model included fixed effects for time-of-day (morning, evening, night), presence vs. 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 intraclass correlation is the proportion of total variance explained by systematic differences among drivers after accounting for time-of-day and feedback condition effects. The same model used to determine the ICC's 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 employed 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 $p < 0.10$, we 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 the *SleepWatch®.* Mixed model analyses of variance were used assess the significance of the fixed intervention effect. Random effects included between and within driver variance, which were used to compute intraclass correlations. Descriptive statistics were used for analyzing the drivers' daily diary and post-experimental 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 and U.S. n = 9), when confining data analyses 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. While rough road conditions in the operating trucks caused some data loss, the final dataset 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 using all 38 drivers who completed the study. Key findings are summarized briefly below relative to the primary hypotheses and to other key findings and recommendations relevant fatigue management in long-haul trucking.

HYPOTHESIS I: FMT FEEDBACK WILL IMPROVE DRIVER ALERTNESS AND/OR REDUCE DRIVER DROWSINESS

PHASE 1: CANADIAN DRIVERS — There was marginal evidence in support of the hypothesis. 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 the opposite effects than those found for night time 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 I 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 CANADA 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 CANADA AND U.S. DATA — There was no support for Hypothesis I when *SleepWatch®* data were combined between study phases.

SLEEP ON WORKDAYS VERSUS NON-WORKDAYS — Not surprisingly, drivers in both countries slept significantly more on non-workdays than on workdays. During the NO FEEDBACK 2-week period of the Canada study phase, drivers averaged 7 hours and 17 minutes sleep per 24h period on non-workdays compared to 6 hours and 15 minutes on workdays, a mean difference of 1 hour and 2 minutes ($p = 0.023$). Similarly, during the FEEDBACK 2-week period of the Canada phase, drivers averaged 7 hours and 31 minutes of sleep per 24 hours on non-workdays compared to 6 hours and 14 minutes on workdays, a mean difference of 1 hour and 17 minutes ($p = 0.0005$). Comparable results were obtained in the U.S. study phase. During the NO FEEDBACK 2-week period U.S. drivers averaged 6 hours and 32 minutes of sleep per 24 hours on non-workdays compared to 5 hours and 14 minutes on workdays, a mean difference of 1 hour and 18 minutes ($p = 0.018$). Similarly, during the FEEDBACK period, U.S. drivers averaged 7 hours and 32 minutes sleep compared to 5 hours and 1 minute on workdays, a mean difference of 2 hours and 31 minutes ($p = 0.0004$). These are relatively large differences in 24-hour sleep durations, suggesting that drivers developed sleep debts across the workweek.

EFFECT OF FMT FEEDBACK ON SLEEP ON NON-WORKDAYS — 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 hours was analyzed for both study phases, separating workdays and non-workdays, there was clear evidence in support of hypothesis I. In contrast to workdays, where FMT FEEDBACK had no effect on sleep time, there was a significant increase in mean sleep duration during non-workdays in the FEEDBACK condition relative to the NO FEEDBACK in both the Canadian drivers ($t = -2.55$, $df = 15$, $p = 0.023$) and U.S. drivers ($t = -2.88$, $df = 10$, $p = 0.018$). Drivers in both study phases increased their non-workday sleep durations by an average of 45 minutes per day over sleep duration on non-workdays in the NO FEEDBACK condition ($F_{1,25} = 4.39$, $p = 0.046$).

OTHER KEY FINDINGS

A "COST" TO BEING MORE ALERT WITH FMT FEEDBACK?—As summarized above, 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 analog 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 vs.

5.33; t = 6.63, df = 11, p < 0.0001). These findings suggest the possibility 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" to the added effort (in attention and compensatory behaviors) required to respond to the information from the devices, and that "cost" may manifest itself as slightly worse performance and greater subjective sleepiness when performing 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%; U.S. 100%). Descriptive analyses of drivers' 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 *Howard Power Center Steering®* system and *SafeTRAC®*, than they were about *SleepWatch®* and *CoPilot®*. It is noteworthy that *Howard Power Center Steering®* and *SafeTRAC®* both interface with the vehicle, while *SleepWatch®* and *CoPilot®* interface with the driver. It may be that truck drivers prefer fatigue management be carried out by way of vehicle monitoring more so than driver monitoring. More research is needed to understand what influences commercial drivers' attitudes toward feedback by technology.[19]

A FUTURE FOR FMT TECHNOLOGIES—Overall, participant drivers were positive toward the FMT approach in general and felt that if such technologies could be further improved, they would be of benefit in helping manage fatigue and alertness.

RECOMMENDATIONS FOR FUTURE WORK OUTSIDE THE SCOPE OF PROJECT

CONTINUE DEVELOPMENT OF FMT TECHNOLOGIES—There is enough evidence to support the case for continued development of FMT technologies. But these should not solely be in the area of 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 a host of 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. Post-experimentally, drivers rated the course content and knowledge they gained as being from 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 divers indicated they perceived benefit from the course and that they would like to have more of this type of didactic to help teach them how to manage their 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 towards 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 the Psychomotor Vigilance Task 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 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, and positive and negative predictability of a shorter-duration PVT test (e.g., 3-5 minutes) relative to truck driver fatigue.

BARRIERS TO DRIVERS OBTAINING ADEQUATE SLEEP DURING WORKDAYS NEED TO BE IDENTIFIED—One of the more striking outcomes of the project was the finding that drivers in both countries were routinely averaging between 5 hours and 6¼ hours of sleep per day during workdays, despite very different work schedules. Recent scientific work, some of it by U.S. DOT on volunteer truck drivers, shows that severe sleep debt and deficits in behavioral alertness can develop within a few days at these sleep durations. The fact that project participants markedly increased their sleep durations on non-workdays also supports the view that they were suffering sleep debts. Much more needs to be understood about the factors that determine when and where drivers obtain sleep on workdays and non-workdays; the barriers to obtaining adequate sleep on workdays; and the factors that convince them to get more recovery sleep on non-workdays.

ACKNOWLEDGEMENTS

The Fatigue Management Technologies Pilot Test was sponsored by the Federal Motor Carrier Safety Administration (FMCSA) and Transport Canada (TC), in cooperation with the American Transportation Research Institute (ATRI). Project leadership was provided by Robert J. Carroll (FMCSA), Sesto Vespa (TC), and Rebecca Brewster (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. In addition to the authors, the following contributed to the research. Robert Hachadoorian (Biomedical Statistical Consulting); Adrian Ecker, Donald Terry, John Powell (University of Pennsylvania School of Medicine); Daniel Redmond, Thomas Balkin, Gregory Belenky, Greg Lounsberry (Walter Reed Army Institute of Research); Todd Jochem, Dean Pomerleau (Applied Perception and Assistware Technology, Inc.); Richard Grace, Robert Engel (Attention Technologies); B. Red Chester, Dee Howard, Ruben Clayton (River City Products); Allan Schwartz, Jean Paul Daveau, Patrick Painter, Pierre Pommarel (Accident Prevention Plus); Daniel Einwechter, Enno Jakobson, Robert Halfyard, Paul Clague, Steve Kratz, 26 Challenger Motor Freight drivers (Challenger Motor Freight); Robert Petrancosta, Edwin Gilmore, Keith Peters, Joseph Fyda, Mark Stets, Farris Scott, Anthony Martucci, Jarrod Stokes, 12 Con-Way Express drivers (Con-Way Central Express); William C. Rogers, Ronald Knipling.

REFERENCES

1. Dinges, D.F., Mallis, M., Maislin, G., Powell, J.W.: Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management. U.S. Department of Transportation, National Highway Traffic Safety Administration, Contract No. DTNH22-93-D-07007. pp. 1-113, 1998.
2. Krueger, G. Technologies And Methods For Monitoring Driver Alertness And Detecting Driver Fatigue: A Review Applicable To Long-Haul Truck Driving. Technical Report for American Transportation Research Institute and Federal Motor carrier Safety Administration, June 2004.
3. Dinges, D.F., Mallis, M.M.: Managing fatigue by drowsiness detection: Can technological promises be realized? In: Hartley, L. (Eds.) Managing Fatigue in Transportation, Elsevier Science Ltd., Kidlington, Oxford, UK, pp. 209-229, 1998.
4. Redmond, D.P., Sing, H.C., Harding, G.W., Hegge, F.W.: Human rest activity rhythms and wrist movement recording-systems. Chronobiologia 10(2):149-149, 1983.
5. Redmond, D.P. & Hegge, F.W.: Observations on the design and specification of a wrist-worn human activity monitoring system. Behavior Research Methods, Instruments, and Computers, 17(6):659-669, 1985.
6. Redmond, D.P., Sing, H.C., Harding, G.W., Hegge, F.W.: Disturbed human rest activity cycles - automated data reduction and rhythmic analysis of wrist activity recordings Chronobiologia 12 (3): 267-267, 1985
7. Belenky, G., Balkin, T.J., Redmond, D.P., Sing, H.C., Thomas, M.L., Thorne, D.R. & Wesensten, N.J.: Sustained performance during continuous operations: The US Army’s sleep management system. In Hartley, L.R. (Ed.) Proceedings of the Third International Conference on Fatigue and Transportation: Managing Fatigue in Transportation. Fremantle, Western Australia. Oxford UK: Elsevier Science Ltd., 77-85, 1998.

8. Hursh, S.R., Redmond, D.P., Johnson, M.L., Thorne, D.R., Belenky, G., Balkin, T.J., Storm, W.F., Miller, J.C., Eddy, D.R.: . Fatigue models for applied research in warfighting. Aviation Space and Environmental Medicine 75(3):A44-A53 Suppl. 2, 2004.
9. Dinges, D.F., Pack, F., Williams, K., et al. Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. Sleep 20:267-277, 1997.
10. Van Dongen H.P.A., Maislin, G., Mullington, J.M., et al. The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep 26:117-126, 2003.
11. Belenky, G., Wesensten, N.J., Thorne, D.R., et al. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep-dose response study. Journal of Sleep Research 12:1-12, 2003.
12. Wierwille, W.W. & Ellsworth, L.A. (1994). Evaluation of driver drowsiness by trained raters. *Accident Analysis and Prevention*, 26, 5, 571-581.
13. Wierwille, W.W., Ellsworth, L.A., Wreggit, S.S., Fairbanks, R.J., & Kirn, C.L. (1994). Research on vehicle-based driver status/performance monitoring: Development, validation, and refinement of algorithms for detection of driver drowsiness. (NHTSA Technical Report No. DOT-HS-808-247). Washington, DC: National Highway Traffic Safety Administration.
14. Wierwille, W.W.: Historical perspective on slow eyelid closure: Whence PERCLOS? In: Carroll, R.J. (Ed.) *Technical Proceedings of Ocular Measures of Driver Alertness Conference,* Herndon, VA. (FHWA Technical Report No. MC-99-136). Washington, DC: Federal Highway Administration, Office of Motor Carrier and Highway Safety, pp. 31-53, 1999.
15. Mallis, M., Maislin, G., Konowal, N., Byrne, V., Bierman, D., Davis, R., Grace, R., Dinges, D.F.: 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 sponsored by the National Highway Traffic Safety Administration, Federal Highway Administration, Office of Motor Carriers, 2000.
16. Dinges D.F., Price NJ, Maislin G, Powell JW, Ecker AJ, Mallis MM, Szuba MP: Prospective Laboratory Re-Validation of Ocular-Based Drowsiness Detection Technologies and Countermeasures. Subtask A in Report: Wierwille WW, Hanowski RJ, Olson RL, Dinges DF, Price NJ, Maislin G, Powell JW, Ecker AJ, Mallis, MM, Szuba MP, Ayoob E, Grace R, Steinfeld A.: NHTSA Drowsy Driver Detection And Interface Project, DTNH 22-00-D-07007; Task Order No. 7; September, 2002.
17. Dinges, D.F., Powell, J.W.: Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Behavior Research Methods, Instruments, and Computers 17(6):652-655, 1985.
18. O'Neill, T.R., Krueger, G.P. & Van Hemel, S.B. 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. Alexandria, VA: Star Mountain, Inc. & The American Trucking Associations in partnership with the U.S. Department of Transportation Federal Highway Administration. September, 1996.
19. Roetting, M., Huang, Y.H., McDevitt, J. and Melton, D.: When technology tells you how you drive - truck drivers' attitude toward feedback by technology. Transportation Research Part F: Traffic Psychology and Behaviour, 6(4): 275-287, 2003.

LIST OF FIGURES

FIGURE 1 Walter Reed Army Institute of Research (WRAIR) *SleepWatch®*

2A: Bright-Eye Image **2B: Dark-Eye Image** **2C: Difference Image**

FIGURE 2 Examples of Eye Images Taken by the *CoPilot®*

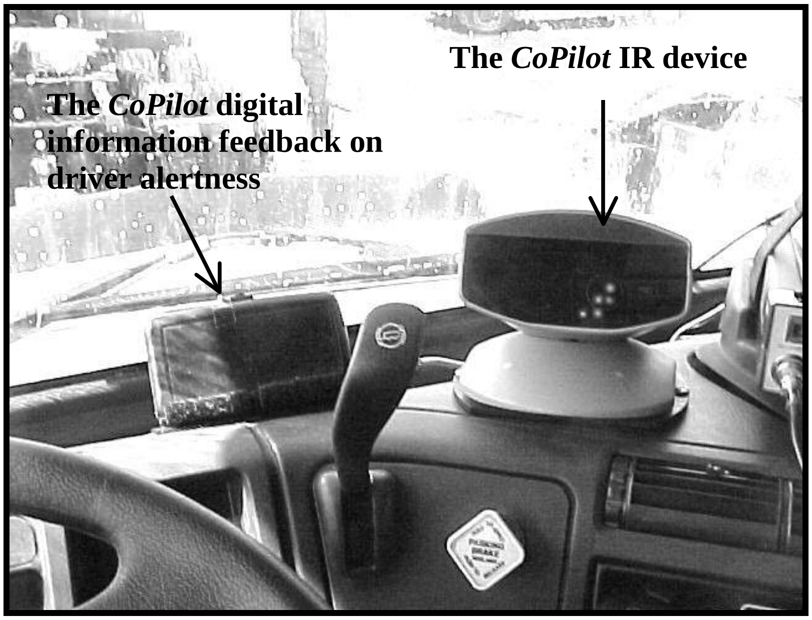

FIGURE 3 The *CoPilot*® Infrared Retinal Reflectance Monitor

FIGURE 4 *SafeTRAC®* Lane-Tracking Monitor

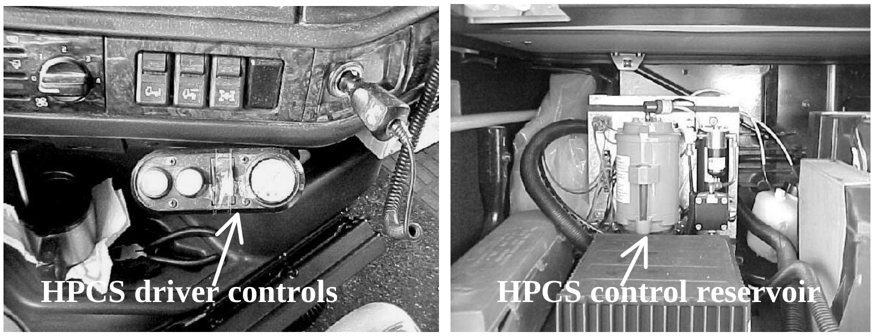

FIGURE 5 Howard Power Center Steering® (HPCS)