

The vision of the Office of Motor Carriers is to help move people, goods, and commercial motor vehicles on our Nation's highways in the most efficient, economical, and crash-free manner possible. The OMC research and technology program focuses on improving safety in interstate commercial motor vehicle operations and serves a trucking and motor coach industry that carries more than 40 percent of all intercity freight.

Studies are conducted in the following areas: commercial driver human factors, health, and performance needs; new and emerging driver and vehicle technologies; safety-related data collection and analysis needs; and performance-based changes to the Federal Motor Carrier Safety Regulations.

The OMC's technology research promotes safety by identifying, collecting, and communicating information about technological advances.



Office of Motor Carrier Research
and Standards
400 Seventh Street, SW
HCS-30; Room 3107
Washington, DC 20590

PERCLOS: A Valid Psychophysiological Measure of Alertness As Assessed by Psychomotor Vigilance

Introduction

Driver fatigue is a significant risk factor in commercial transportation. Fatigue from combinations of sleep loss, night driving, and prolonged work time contributes substantially to the number of motor vehicle crashes. Over the past decade, driver drowsiness/fatigue has been the subject of intensified Department of Transportation interest and activity; one major initiative has been to explore technology that will enhance commercial motor vehicle (CMV) driver fatigue management.

This Tech Brief summarizes an Intelligent Transportation System (ITS) study titled *Evaluation of Techniques for Ocular Measurement as an Index of Fatigue and as the Basis for Alertness Management*. The study was funded in part by FHWA's Office of Motor Carriers and managed by the National Highway Traffic Safety Administration (NHTSA). The project's goal was to evaluate the validity and reliability of several drowsiness-detection measures and technologies in a controlled laboratory setting, and to analyze the effects of alerting stimuli on drivers' alertness levels.

Of the drowsiness-detection measures and technologies evaluated in this study, the measure referred to as "PERCLOS" was found to be the most reliable and valid determination of a driver's alertness level. PERCLOS is the percentage of eyelid closure over the pupil over time and reflects slow eyelid closures ("droops") rather than blinks. A PERCLOS drowsiness metric was established in a 1994 driving simulator study as the proportion of time in a minute that the eyes are at least 80 percent closed. (Wierwille et al., 1994) Based on research by Wierwille and colleagues (1994), FHWA and NHTSA consider PERCLOS to be among the most promising known real-time measures of alertness for in-vehicle drowsiness-detection systems.

The results of this research support the development of a "first-ever" real-time drowsiness detection sensor that would measure the percentage of eyelid closure over the pupil, over time (i.e., PERCLOS).

Purpose

Because this study was concerned with finding a reliable and valid drowsiness-detection system, FHWA and NHTSA evaluated several promising drowsiness-detection technologies. These operator-centered, in-vehicle, fatigue-monitoring technologies record various biobehavioral dimensions of an operator, such as a feature of the eyes, face, head, heart, or brain electrical activity, while the operator is driving.

Practical systems for in-vehicle alertness monitoring may consist of a continuous indicator display of a driver's alertness level, provided by in-vehicle fatigue-monitoring technology, as well as warning messages to notify a driver when drowsiness is detected and alerting stimuli to restore driver alertness and performance. Thus, this research project was divided into two experiments:

- To evaluate the validity, sensitivity, and reliability of specified operator-centered, fatigue-detection devices for predicting vigilance lapses in alert and drowsy subjects; and
- To analyze effects of a combination of auditory and vibrotactile alerting stimuli on alertness lapses.

Research Methodology

Experiment 1

To obtain estimates of the validity of the selected drowsiness-detection technologies, researchers conducted a controlled laboratory experiment. Fourteen adult males remained awake in a laboratory for 42 hours, while working on a computerized test battery every 2 hours. The tests included a 20-minute psychomotor vigilance task (PVT), that required subjects to sustain attention and respond to a randomly appearing light on a computer screen by pressing a button. PVT performance lapses refer to the times when a subject failed to respond to the task in a timely manner (i.e., < 500 msec.); lapses were recorded each minute throughout the trial and then totaled for the entire 20 minutes.

PVT lapses were selected as the validation criterion variable because driving is a vigilance task requiring psychomotor reactions, and psychomotor vigilance has been previously validated to be very sensitive to fatigue from night work and sleep loss. Thus, PVT lapses were a valid index for evaluating candidate technologies.

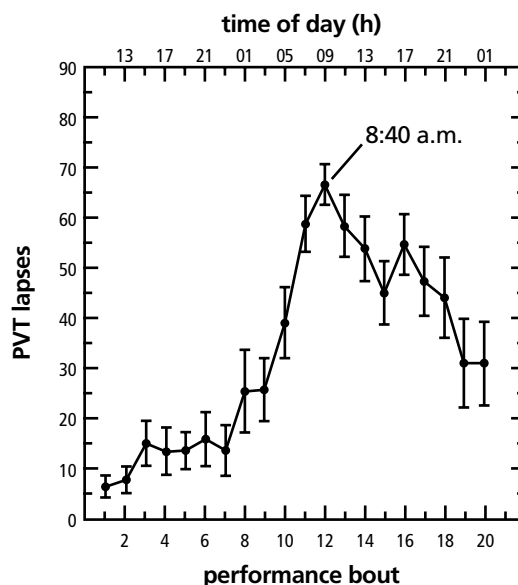
Six technologies were tested during the research program: two electroencephalographic (EEG) algorithms, which detect changes in brain wave activity; two eye blink monitors, which measure fatigue based on eye blink activity; a head-position monitoring device, which detects fatigue based on head motion; and PERCLOS. These six technologies yielded a total of nine drowsiness metrics, since both PERCLOS and the head-position monitoring device had more than one. PERCLOS had three drowsiness metrics:

- P70, the proportion of time the eyes were closed at least 70 percent;
- P80, the proportion of time the eyes were closed at least 80 percent; and
- EYEMEAS (EM), the mean square percentage of the eyelid closure rating.

Researchers simultaneously recorded PVT performance and the output of these drowsiness-detection technologies every 2 hours throughout the 42-hour period. The study design permitted an estimate of the degree of similarity, or coherence, between PVT

Figure 1.

Mean number of psychomotor vigilance task performance lapses per 20-minute test bout, across 20 test bouts.



performance lapses and the alertness/drowsiness output of each technology. Each technology was time-locked to PVT performance to test coherence between PVT lapses and the technology's drowsiness metric. In order for a technology to demonstrate a high coherence with PVT lapses, its drowsiness metric had to demonstrate regular covariation with performance lapses across the entire 42-hour period, within and between subjects.

To collect the PERCLOS measure, a low-light, closed-circuit television was used to monitor the subject's entire face and to record eyes and eyelids during PVT performance testing. Video recordings were collected on all 14 subjects, 10 of which were scorable (in 4 subjects, glare from other equipment prevented PERCLOS scoring). Trained scorers viewed the recordings and rated the degree to which the drivers' eyes were closed from moment to moment.

Experiment 2

The second part of the study evaluated the effects of alerting stimuli on PVT performance and PERCLOS scores. Experiment 2 was conducted 6 months after Experiment 1 was completed. Researchers wanted to find out if systematically delivered alerting stimuli could reduce drowsiness and its consequences for hypovigilance. Four of the original 14 subjects participated in this experiment, which also required the subjects to remain awake in a laboratory for 42 hours while performing a psychomotor vigilance task every 2 hours.

Table 1.
Bout-to bout coherence for individual subjects (Pearson correlation coefficients).

Subject	PERCLOS			EEG Algorithms		Head-Position Metrics		Eye Blink Monitors	
	P70	P80	EM	1	2	1	2	1	2
1	0.89	0.92	0.89	*	*	0.83	0.82	0.10	*
2	0.85	0.83	0.84	*	0.40	*	*	0.71	0.20
3	0.95	0.97	0.95	*	*	0.91	0.85	0.90	*
4	0.84	0.83	0.83	*	*	-0.54	0.20	0.54	0.77
5	0.94	0.94	0.95	0.54	*	*	*	-0.10	*
6	0.95	0.96	0.94	0.57	0.95	*	*	0.39	0.54
7	0.92	0.92	0.92	0.36	0.31	*	*	0.54	0.32
8	0.55	0.67	0.70	0.66	0.84	*	*	0.50	0.85
9	0.78	0.77	0.71	*	*	0.23	0.13	0.67	*
10	0.95	0.97	0.95	*	*	0.87	0.65	-0.48	*
11	*	*	*	*	*	*	*	0.31	0.79
12	*	*	*	*	*	*	*	0.34	*
13	*	*	*	*	*	*	*	0.14	*
14	*	*	*	*	*	*	*	0.17	*

* Data not available or remaining to be analyzed.

During this experiment, auditory and vibrotactile alerting stimuli were delivered during each 20-minute PVT performance bout, but were not contingent on actual PVT lapses. Subjects received more frequent stimuli during the hours when PVT lapses were expected to be the greatest, as recorded in Experiment 1 (i.e., after 22-26 hours of waking). Vibrotactile stimuli were delivered through a hand-held box with a duration of 5 seconds. Auditory stimuli consisted of three different recorded messages received through headsets.

Findings

Experiment 1

The study design effectively induced fatigued performance due to the combined influence of sleep loss and circadian nocturnal phase. As expected, the average number of PVT lapses increased significantly during the 42-hour test trial, as shown in **figure 1**.

Nearly all of the technologies showed potential for detecting drowsiness by predicting lapses in at least one subject or a subset of subjects, but only PERCLOS correlated highly with PVT lapses both within and between subjects. PERCLOS not only had the highest coherence of the technologies tested, but correlated more highly with PVT lapses than did the subjects' ratings of their own sleepiness.

Complete data for any given technology were often

available on only a subset of the 14 subjects due to unscorable results, a limited amount of certain equipment, and unreliable data storage/retrievability for some devices.

Bout to Bout Coherence

Bout-to-bout coherence refers to the correlation between the total number of visual performance lapses in a 20-minute PVT test bout and the drowsiness-detection results from a given technology. Each subject had 20 test bouts for comparison across the 42-hour test period. Bout-to-bout coherence was calculated within each subject using the technology available for that subject. **Table 1** provides bout-to-bout coherence measures for individual subjects and technologies. PERCLOS had the highest average bout-to-bout coherence of all of the devices.

Minute to Minute Coherence

Minute-to-minute coherence refers to the correlation between the total number of performance lapses in each minute of every 20-minute PVT bout and the results of a given drowsiness detection technology. Minute-to-minute coherence was consistently lower than bout-to-bout coherence. (Minute-to-minute statistics are not shown in this Tech Brief.) It appeared that all technologies had a better prediction of PVT lapses when sampling a longer (20-minute) rather than briefer (1-minute) period.

Lower Lapsers vs. Higher Lapsers

As expected, all subjects eventually became sleepy

Researcher

This study was performed by David F. Dinges, Ph.D., Department of Psychiatry, University of Pennsylvania School of Medicine; and Richard Grace, Ph.D., Carnegie Mellon Research Institute. Contract No. DTNH22-93-D-07007.

Distribution

This Tech Brief is being distributed according to a standard distribution. Direct distribution is being made to the Regions and Divisions.

Availability

The study final report is now available. Copies may be obtained from the National Technical Information Service, Telephone: (703) 605-6000.

Key Words

driver fatigue, drowsiness, fatigue, fatigue management, drowsiness-detection, PERCLOS, driver performance.

Notice

This Tech Brief is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The Tech Brief provides a synopsis of the study's final publication. The Tech Brief does not establish policies or regulations, nor does it imply FHWA endorsement of the conclusions or recommendations. The U.S. Government assumes no liability for its contents or their use.



U.S. Department of Transportation
Federal Highway Administration

October 1998

Publication No. FHWA-MCRT-98-006

and experienced increased PVT lapsing during the 42-hour period of waking. However, a subset of the subjects accounted for a disproportionate number of PVT lapses. Subjects who had an average of 38 or more PVT lapses per bout were referred to as "higher lappers" (HL). These subjects would begin lapsing almost immediately after beginning the task. Subjects with an average of 33 or less lapses per bout were "lower lappers" (LL). LL subjects maintained alertness for some time before beginning to lapse. HL subjects were 42 percent of the subjects, but accounted for 69 percent of all PVT lapses.

When bout-to-bout coherence was analyzed to compare data for HL subjects and LL subjects, researchers found that some technologies were inconsistent when assessing the drowsiness levels between the two groups. For example, one of the eye-blink technologies had a significantly lower bout-to-bout coherence for the LL subjects than for the HL subjects. Importantly, PERCLOS yielded comparably high bout-to-bout coherence between the two groups. Despite large inter-subject variability in PVT lapse rates, PERCLOS appeared to reliably predict lapsing across both groups.

Experiment 2

PVT lapse data and PERCLOS P80 data were compared between Experiments 1 and 2 for each of the four subjects. Comparison of PVT lapses in each minute prior, during, and following stimulation revealed no evidence that auditory or vibrotactile stimuli had much effect on PVT lapses. The alerting stimuli provided in this experiment did not markedly reduce lapses in drowsy subjects beyond the minute in which the alert occurred, suggesting only a narrow window of opportunity for a drowsy driver to safely leave the roadway.

Another notable finding was the intra-subject consistency in the time-course of the progression of drowsiness during the 42-hour test periods. Individual subjects exhibited drowsiness progressions in Experiment 2 that were highly similar to those shown in Experiment 1, even though the experiments were conducted 6 months apart. Although there were only four repeat subjects, this finding supports the notion of significant, durable differences in individual susceptibility to drowsiness during sleep deprivation. Indeed, each subject had a characteristic "signature" of drowsiness progression. Finally, PERCLOS reliably correlated with vigilance lapses in this second experiment.

Further Research

Attempts to transition an on-line, automated version of PERCLOS to a realistic over-the-road environment are currently underway. NHTSA has developed models of drowsy driver performance by relating PERCLOS and driving performance using over-the-road data from CMV drivers. FHWA and NHTSA are developing plans to understand the operational safety benefits of a drowsiness detection sensor in actual driving environments.

Through the ITS Intelligent Vehicle Initiative, NHTSA is currently managing a follow-up to this study to evaluate the effectiveness of various potential elements of the driver-vehicle interface (DVI) of in-vehicle driver alertness monitoring devices. Some of the components that will be assessed are real-time gauges, informational alarms and warnings, and alerting stimuli. The study will make recommendations regarding optimal DVI design elements for driver alertness monitors.

Reference

Wierwille, WW, Ellsworth, LA, Wreggit, SS, Fairbanks, RJ, Kirn, CL: *Research on vehicle-based driver status/performance monitoring: development, validation, and refinement of algorithms for detection of driver drowsiness*. National Highway Traffic Safety Administration Final Report: DOT HS 808 247, 1994.