ASSESSMENT OF A DROWSY DRIVER WARNING SYSTEM (DDWS) FOR HEAVY VEHICLE DRIVERS

Paul S. Rau, Ph.D., CPE
National Highway Traffic Safety Administration
Gregory M. Fitch, Ph.D.
Joseph L. Bocanegra, M.S.
Myra Blanco, Ph.D.
Richard J. Hanowski, Ph.D.
Virginia Tech Transportation Institute
United States
Paper Number: 09-0569

ABSTRACT

Drowsiness has a globally negative impact on human performance by slowing response time, decreasing situational awareness, and impairing judgment. This paper reports the findings of a Field Operational Test (FOT) of an early prototype Drowsy Driver Warning System (DDWS). Fifty-three research questions were addressed related to performance, capabilities, acceptance, and deployment. The FOT included control and test groups utilizing an experimental design suitable for a field test. The dataset for the analysis consisted of 102 drivers from 3 for-hire trucking fleets using 46 instrumented trucks. Fifty-seven drivers were line-haul and 45 were long-haul operators. The data set contained nearly 12.4 terabytes of video, truck instrumentation, and kinematics data for 2.4 million miles of driving and 48,000 driving-data hours recorded, resulting in the largest data set ever collected by the U.S. Department of Transportation. When considering the operational window of the Driver Fatigue Monitor, results showed that the drivers in the Test Group had lower drowsy measurement values, and that drivers who received feedback from the system had an overall reduction of drowsy driver instances. Whereas, the experimental design was specified to support the statistical reliability of potential findings, the dataset was largely diminished from eyes-off-road time from driver distraction and normal mirror checking tasks, which were incorrectly sensed by this early prototype as drowsy episodes. As a result, no statistically reliable safety benefit was observed. However, novel data reduction procedures were able to extract data during the time periods in which the system was accurately detecting drowsiness, and analysis of these data indicated a slight reduction in critical unsafe driving events related to drowsiness. As a result, while there is some indication that a DDWS may be a promising concept, the particular prototype used in this field test to implement the concept needs significant improvement and further study.

INTRODUCTION

The National Highway Traffic Safety Administration (NHTSA), the Federal Motor Carrier Safety Administration (FMCSA), and the Intelligent Transportation System Joint Program Office (ITS JPO) sought to investigate the potential safety benefits offered by deploying a drowsy driver warning system (DDWS) into fleet service. The Virginia Tech Transportation Institute (VTTI) was contracted to perform a Field Operational Test (FOT) of a working DDWS prototype developed by Attention Technologies, Inc. (ATI). The prototype, termed the Driver Fatigue Monitor (DFM), was designed to alert drivers using auditory and visual alerts when drowsiness was detected. The DFM assessed drowsiness using the percentage of eye-closure (PERCLOS) measure developed by (Wierwille, 1999). PERCLOS refers to the percentage of time that the driver's eyes are between 80 and 100 percent closed during a defined time interval. The premise here is that the driver's pupils become covered, and therefore do not perceive visual stimuli, when the eyes are greater than 80 percent closed. Since visual information can no longer be gathered at this point, this is considered to be critical in terms of safe driving. DFM devices were installed in a fleet of heavy vehicles. Driver performance with these devices was investigated. Fifty-three research questions related to performance, capabilities, acceptance, and deployment were addressed. This paper highlights the findings.

METHODS

Apparatus

The DFM (Figure 1) was mounted on the dash of the truck cab. By illuminating the driver's face with

infra-red (IR) light, the DFM used a camera and a processor to determine the percentage of time that the driver's eyelids were closed more than a pre-set threshold. The pre-set thresholds were 1 minute, 3 minutes, or 5 minutes depending on what DFM sensitivity setting was selected.

Figure 1. The DFM monitor mounted on the truck's dash.

Driver behavior was continuously recorded using four video cameras: one pointed at the driver's face, one pointed at the forward roadway, one pointed down the left side of the tractor-trailer, and one pointed down the right side of the tractor-trailer (Figure 2). The four video images were multiplexed into a single image prior to being recorded (Figure 3).

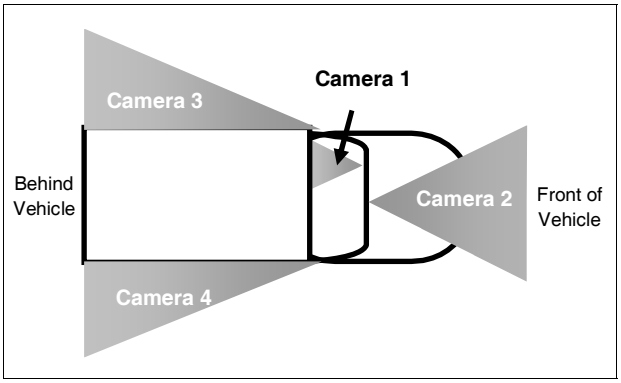

Figure 2. Camera directions and approximate fields of view.

Figure 3. Camera views multiplexed into a single image.

A vehicle-onboard-radar (VORAD) unit was mounted on the truck's front bumper and was used to detect objects in front of the truck (Figure 4). The VORAD unit allowed the distance to, and relative velocity of, lead vehicles to be continuously measured. A measure of the truck's time-to-collision (TTC) to a lead vehicle could be derived from this data. The VORAD unit was used for passive data collection only and did not display range information to the driver.

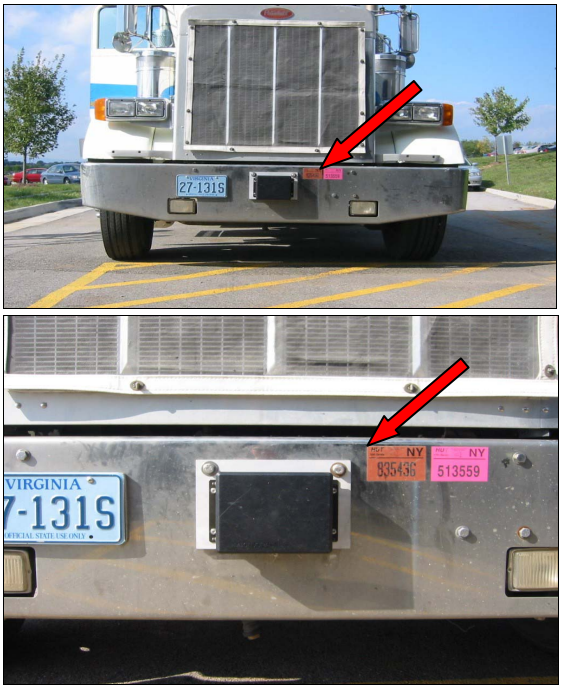

Figure 4. Vorad unit on the front of the truck.

The DFM data, video images, VORAD data, as well as data collected from various vehicle sensors were recorded by the data acquisition system (DAS) whenever the truck was on and in motion. The DAS consisted of a Pentium-based computer that received and stored data from a network of sensors distributed around the vehicle. Data were stored on the system's external hard drive, which could store several weeks of driving data before it needed to be replaced. DASs were either mounted under the passenger seat (Figure 5) or in the truck's rear storage compartment (Figure 6).

Figure 5. Encased computer and external hard drive installed under the passenger's seat.

Figure 6. Encased computer and external hard drive installed in the truck's rear storage compartment.

Subjects

The DFM device was installed in 46 trucks from three for-hire trucking fleets. A total of 102 drivers (101 males and one female) participated. Fifty-seven drivers were line-haul (i.e., out-and-back) operators and 45 drivers were long-haul (i.e., drivers on the road for approximately one week) operators.

Safety Benefit Model

In designing the DDWS FOT, the following assumptions were made regarding the safety benefits offered by the DFM prototype:

- The purpose of a PERCLOS-based DDWS is to provide the driver with timely feedback regarding an unsafe drowsy state (Wierwille, 1999).
- Without drowsiness alerting information, it would be expected that drivers would have more frequent episodes of drowsy driving. However, if used appropriately, the use of the system should lead to fewer episodes of on-the-job driver drowsiness.
- Sleep is the only true remedy for drowsiness.
- Alert information providing feedback to the driver about his or her drowsiness state, coupled with a fatigue management plan that informs the driver about the importance of sufficient sleep, indicates to the driver that driving safety is being compromised.
- Drivers will be positively influenced by their experience with the DDWS.
- Research indicates that drowsiness is a contributing factor (not necessarily a causal factor) in 20 percent of Safety Critical Events (SCEs) (Hanowski, Wierwille, & Dingus, 2003). For some unspecified portion of SCEs, it is hypothesized that high alertness may have prevented the incident from occurring. Therefore, alert drivers would be expected to be involved in fewer critical incidents as compared to drowsy drivers.

A safety benefits model that relates the DDWS to the anticipated benefits was developed (Figure 7). The model incorporates the six assumptions and indicates that a valid and reliable DDWS would be expected to: 1) reduce on-the-job drowsiness, 2) increase the amount of sleep drivers get, and 3) reduce involvement in drowsiness-related SCEs.

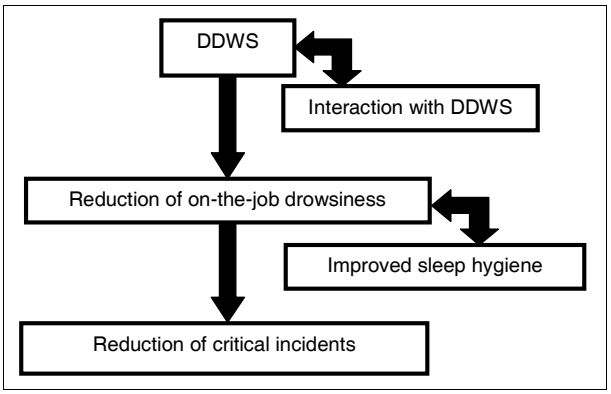

Figure 7. Modeling Safety Benefits Associated with a DDWS

Experimental Design

The DDWS FOT used a quasi-experimental design that included both Control and Test groups. The Control Group followed an A^9 design, where A (superscript refers to the prescribed number of weeks for that condition) refers to the condition in which the DFM collected data over the duration of the participant's involvement but never provided feedback to the driver. The Test Group experienced an A^2B^9 design, where A refers to the condition in which the DFM did not provide feedback to the driver (i.e., Baseline and Control Conditions) and B refers to the condition where the DFM did provide feedback (i.e., the system was fully functional, or Test Conditions). In order to look for any adjustments in driving behavior during the FOT which were not attributable to the DFM, a Baseline Control Condition was defined within the design.

Adjustments in the number of weeks for the Baseline and Active Conditions were made with respect to data missing due to malfunctions in trucks or the DFM, and drivers not being able to meet with experimenters to switch the DFM to Active or finish participation at the exact time these milestones needed to happen. All these are an anticipated occurrence in any naturalistic data collection effort. Due to this accommodation for data collection, not all drivers experienced the exact same circumstances in the Active Condition, and some drivers had more Baseline or Active weeks than others. Other considerations in the adjustment of duration in the Experimental Conditions were made for inconsistencies in the length of drivers' trips.

However, at the completion of the study all drivers (with the exception of those drivers leaving the study before completion) in the Test Group had a minimum Baseline Condition duration of two weeks followed by the Active Condition for approximately nine weeks (Figure 8).

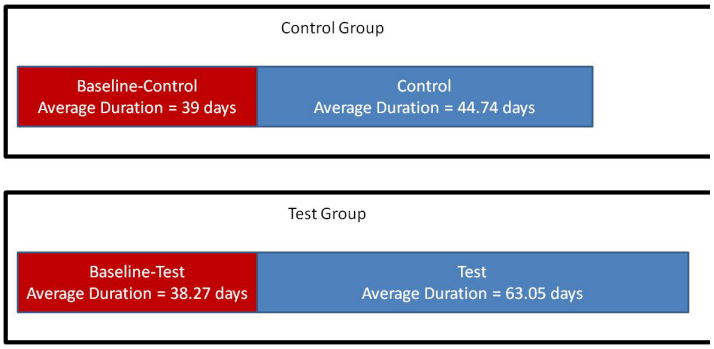

Figure 8. Final timeline for Control and Test Groups.

The dataset produced from this extended data collection effort contained approximately 12.4 terabytes of video and parametric data that encompassed 2.4 million miles of driving and 48,000 driving-data hours.

RESULTS

The purpose of this safety benefit analysis was to evaluate whether DDWS feedback leads to fewer occurrences of drowsiness, fewer alerts, and more frequent stops to take breaks across time and fewer safety critical events (SCEs). The results of this investigation are summarized below.

Drowsiness

The DFM prototype was found to provide a statistically significant reduction in the level of drowsiness over time when a driver received feedback on alertness level during restricted illumination conditions (i.e., dawn, dusk, night) (p = 0.0077). However, when the evaluation was performed outside the operating envelope of the prototype system, a similar reduction in the level of drowsiness was not observed. This is because the DFM prototype only operated within a limited set of conditions. First, the truck had to be travelling at least 35 mph for the DFM to activate. Second, the ambient lighting had to be lower than 50 Lux. Third,

drivers could not wear glasses since eyewear prevented the DFM from detecting eyes. And finally, the DFM often assessed glances away from the forward roadway (e.g., checking mirrors) as eye-closures since the drivers' eyes could not be tracked when this occurred (i.e., false alarms). This last condition was a significant issue that the DDWS FOT identified with the DFM prototype. In order to generalize these results to the larger population of interest, these other conditions must be explored (e.g., glasses, daytime, not considering mirror scanning or distraction to increase PERCLOS level).

Alert Frequency

As suggested, there were numerous DFM alerts that were, in fact, false alarms. That is, they were generated when the driver was not actually drowsy. Since false alarms were typically generated when the driver's eyes could not be identified, there was an interest in determining the number of DFM alerts that were generated when drivers were *actually* drowsy. A sample of DFM alerts were visually inspected and validated. Most of the valid alerts obtained (61 percent of all valid alerts) were in the Test Condition. The highest proportions of valid alerts occurred during weekdays, mainly Monday through Wednesday, and were equally distributed at the times during the day when the DFM prototype was operational. Using a Poisson regression analysis, a statistically significant difference between the experimental conditions was found ($p = 0.0013$). The regression estimates that drivers will obtain more than one valid alert per week if they experience DFM feedback (i.e., are part of the Test Group), but less than that if they are not exposed to DFM feedback (i.e., are part of the Control Group). The implications of this finding are that drivers are more prone to behave in a way that generates a valid DFM alert when they receive DFM feedback than they are if DFM feedback is not available. This is potentially an unintended consequence of a system that notifies drivers when to rest instead of drivers self-regulating their sleep. Perhaps drivers were pushing their ability to stay vigilant knowing that the DFM device would let them know when to rest. It should be noted, however, that the actual difference between the two conditions is just a fraction of an alert. Therefore, for practical purposes these two conditions are very similar.

Post-Alert Behavior

The length of the time that elapsed from drivers receiving a valid DFM alert to them pulling over to rest for 10 minutes or longer was investigated (Figure 10). Drivers in the Baseline Control Condition drove an average of 1 hour and 4 min before stopping the vehicle. Drivers in the Control Condition drove slightly longer (1 hour and 8 min) before stopping. The difference in elapsed time between the two conditions was not statistically significant ($t(25) = 0.34$, $p = 0.74$). A similar finding occurred for drivers in the Test Group. After receiving a valid DFM alert, drivers in the Baseline Test Condition drove 59 min before stopping, while drivers in the Test Condition drove 1 hour and 6 min before stopping. Again, the difference between these values is not statistically significant ($t(25) = -2.05$, $p = 0.0515$). The data suggest that, even in the case of valid DFM alerts, the DFM did not have an effect on drivers' post-alert stopping behavior.

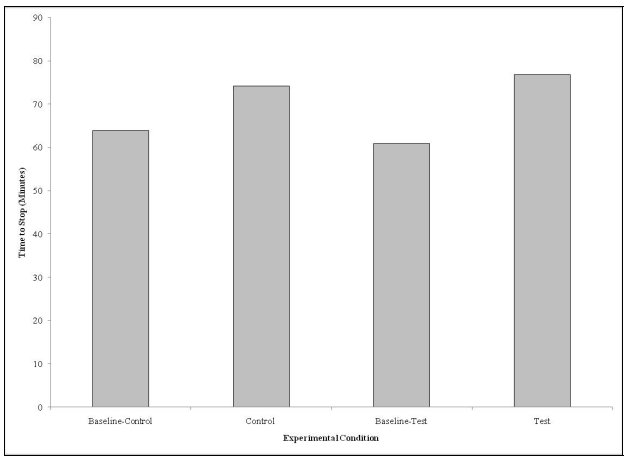

Figure 10. Elapsed Time from Valid Alert to Driver Stopping the Vehicle

The types of behaviors that occurred within the five minutes after each valid alert were examined to investigate wether DFM feedback affected the frequency of these post-alert behaviors. Figure 11 shows the frequency of post-alert behaviors for drivers in the Control Group, while Figure 12 shows the frequency of post-alert behaviors for drivers in the Test Group. Although a statistically significant

effect of the DFM on post-alert behavior was not found (CMH(1) = 0, p = 0.9992), a directionally opposite change trend between Baseline and its corresponding Experimental Condition (i.e., Baseline Control and Control, Baseline Test and Test) occurred for the talk/sing/laugh post-alert behavior. Fewer drivers talked/sang/laughed in the Control Condition compared to the Baseline Control condition, while a greater number of drivers talked/sang/laughed in the Test Condition compared to the Baseline Test condition. Furthermore, the frequency of the "veer off road" (defined as a loss of vehicle control due to various physiological or psychological causes) post-alert behavior was observed to increase from the Baseline Control to the Control Conditions, while it remained the same level between the Baseline Test and Test Conditions. The increase in talking behavior between Control and Test Conditions, as well as the decrease in veering off road behavior, may have arisen from the DFM alerts being generated while drivers were drowsy.

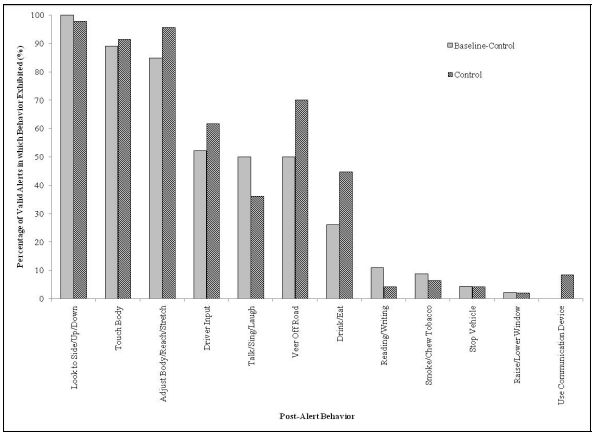

Figure 11. Percentage of valid alerts by type of behavior: Control Group.

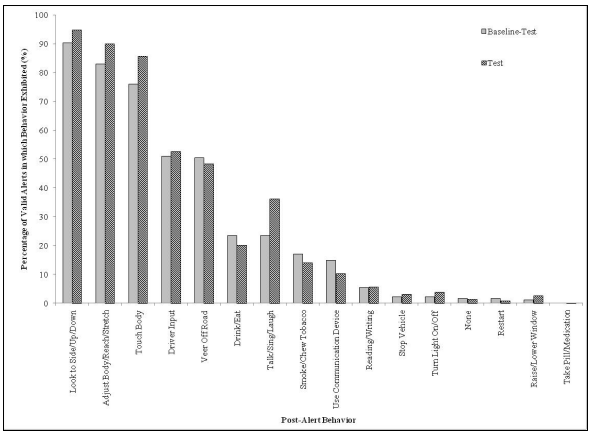

Figure 12. Percentage of valid alerts by type of behavior: Test Group.

In general, looking away from the forward roadway (situations in which the driver looks to the side, up, or down) was the most common behavior observed following presentation of a valid DFM alert. Although it is not possible to infer the frequency of such behavior from the data available, it does indicate that the behavior of gazing forward was broken. The process of seeking new visual information and the visual processing which follows may assist in raising the level of cognitive arousal. The second most common post-alert behavior was to adjust one's body position through shifting in the seat, reaching forward, or stretching out. This behavior may also raise physiological arousal levels. The third most common behavior involved the driver touching his or her face, either by rubbing, scratching, or otherwise holding the face. This behavior is interesting since the facial skin is highly sensitive to touch (Boff & Lincoln, 1988). Touching one's face may act as both a physiological and cognitive arousal mechanism due to the high degree of innervations in the area and the large amount of somatosensory cortex dedicated to this information (Gardner & Kandel, 2000). It is possible these frequently observed behaviors were manifestations, either conscious or not, of drivers' desire to raise their overall levels of arousal above a state of drowsiness.

Involvement in Safety Critical Events

The safety benefits offered by the DFM prototype were assessed by determining whether drivers

receiving DFM feedback were involved in fewer SCEs. Three types of SCEs were considered:

> Crash: Any contact with an object, either moving or fixed, at any speed. Contact could be with other vehicles, roadside barriers, objects on or off of the roadway, pedestrians, cyclists, or animals.
> Near-Crash: Any circumstance requiring a rapid evasive maneuver by the participant vehicle, any other vehicle, pedestrian, cyclist, or animal to avoid a crash. A rapid evasive maneuver was defined as a steering, braking, accelerating, or any combination of control inputs that approached the limits of the vehicle's capabilities.
> Crash-Relevant Conflict: Any circumstance that required a crash avoidance response on the part of the participant vehicle, any other vehicle, pedestrian, cyclist, or animal that was less severe than a rapid evasive maneuver (as defined above). A crash avoidance response could include braking, steering, accelerating, or any combination of control inputs.

The safety benefits model predicts that an effective DDWS will produce a positive change in driver behavior after that driver obtains feedback. This will, in turn, result in a reduction of SCEs. The analyses took into consideration all SCEs, SCEs within the DFM operating envelope, and SCEs when the truck driver was at fault during the different Experimental Conditions. A total of 1,124 SCEs were analyzed, including 28 crashes, 112 near-crashes, and 984 crash relevant conflicts. Of these 1,124 SCEs, 221 occurred within the operating envelope of the DFM. Statistical tests considering all the SCEs showed no statistically significant difference between the Control and Test Groups (χ^2 (2, N = 1,124) = 0.27, p = 0.88). Moreover, no statistically significant differences in SCE distribution were observed for drivers during the Baseline and the Test Conditions.

Additional analyses were conducted to further examine connections between SCEs and drowsiness. However, it is important to note that not all SCEs occurred within the operational envelope of the DFM prototype (specifically the speed and illumination requirements of the system). Using the DFM's PERCLOS as the measure to identify drowsiness in SCEs would exclude a significant portion of all SCEs, thus skewing any subsequent analysis. Therefore, all SCEs were examined by manually computing the driver's PERCLOS measures. In doing so, it was found that over 60 percent of the SCEs occurred when the driver was alert. A second behavioral analysis was conducted to characterize each of the SCEs (see Hickman et al., 2005). Drowsiness-related behaviors were observed in a total of 143 SCEs (13 percent). However, no statistical differences between the Experimental Conditions were present. Although drowsy driving plays a large role in crashes on the highway system, in some cases up to 20 percent of all SCE's (Hanowski, Wierwille, & Dingus, 2003), they do not represent the majority of these events. Additionally, crashes are relatively rare events as evidenced by the 28 observed crashes in this study. Therefore, the results of the present work are not entirely unexpected. These findings do not indicate any lessening of the magnitude of the drowsy driving problem from the DFM, they only illustrate the difficulty in studying a serious but rare event.

DISCUSSION

Drowsiness has a globally negative impact on performance, slowing response time, decreasing situation awareness, and impairing judgment (Balkin et al., 2000; Van Dongen, Maislin, Mullington, & Dinges, 2003). A DDWS that notifies drivers to rest when they become drowsy stands to improve highway safety. The DDWS FOT investigated the effects of implementing the DFM prototype in a multitude of heavy vehicles in a real revenue-producing environment. Overall, there was some evidence that the DFM prototype was successful in reducing levels of driver drowsiness. However, these findings were limited to the DFM prototype's operating envelope, such as low luminance, speeds greater than 35 mph, drivers not wearing eyeglasses, and drivers keeping their gaze on the forward roadway. The evaluations that were performed when the conditions fell outside the operating envelope did not show significant changes in driving behavior. Drivers were not reliably found to rest sooner, change their in-vehicle behavior, or reduce their involvement

in SCEs when receiving valid DFM. DDWSs must therefore address these conditions if changes in driving behavior are to occur, and an improvement in highway safety is to be observed.

REFERENCES

Balkin, T., Thome, D., Sing, H., Thomas, M., Redmond, D., Wesensten, N., Williams, J., Hall, S., & Belenky, G. (2000). *Effects of Sleep Schedules on Commercial Motor Vehicle Driver Performance*. Washington, DC: Federal Motor Carrier Safety Administration, USDOT.

Boff, K.R., & Lincoln, J.E. (Eds.). (1988). *Engineering Data Compendium: Human Perception and Performance* (Vol. II). Wright-Patterson Air Force Base, Ohio: Harry G. Armstrong Aerospace Medical Research Laboratory.

Gardner, E.P., & Kandel, E.R. (2000). Touch. In E.R. Kandel, J.H. Schwartz, & T.M. Jessell (Eds.), *Principles of Neural Science* (4th Ed.), New York: McGraw-Hill Medical, 451-471.

Hanowski, R.J., Wierwille, W.W., & Dingus, T.A. (2003). An on-road study to investigate fatigue in local\short haul trucking. *Accident Analysis and Prevention, 35*(2), 153-160.

Van Dongen, H.P., Maislin, G., Mullington, J.M., & Dinges, D.F. (2003). 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*(20), 117-126.

Wierwille, W.W. (1999). Historical perspective on slow eyelid closure: Whence PERCLOS? *Ocular Measures of Driver Alertness Technical Conference Proceedings*, Herndon, VA, 130-143.