DROWSY DRIVER DETECTION AND WARNING SYSTEM FOR COMMERCIAL VEHICLE DRIVERS: FIELD OPERATIONAL TEST DESIGN, DATA ANALYSES, AND PROGRESS.

Paul Stephen Rau  
National Highway Traffic Safety Administration  
United States  
Paper Number 05-0192

ABSTRACT

Drowsiness among commercial vehicle drivers has been identified as the number one safety concern of commercial fleets at trucking summit meetings. Over the past 10 years, the U.S. Department of Transportation's National Highway Traffic Safety Administration (NHTSA) and its research partners have sought to quantify the loss of alertness among commercial vehicle drivers. This work led to the development of the world’s first unobtrusive and valid sensor of loss of alertness, and has been the benchmark for continuing international study. Replicated experiments have shown that the most valid measure of loss of alertness among drivers is the percentage of eyelid closure over the pupil over time (Perclos). Formerly pioneered by Dr. Walter Wierwille at the Virginia Polytechnical Institute and State University, using a manual observation technique[6], Perclos is now monitored in real time using machine vision technology in the vehicle. In order to estimate the highway safety benefit based on the effectiveness of the system, a Field Operational Test (FOT) is underway with long haul and express (i.e., overnight) fleet operations. This paper discusses the field test methodology, as well as the questions each analysis seeks to answer. A summary of the status of the project, the results to date, and a vision of future work for the deployment of this technology will be provided.

INTRODUCTION

Since 1996, research has been underway at the U.S. National Highway Traffic Safety Administration to develop, test, and evaluate a drowsy-driver detection and warning system for commercial vehicle drivers. Drowsiness is consistently identified in trucking summits as the number one safety concern of commercial vehicle drivers. As a result, numerous field studies and laboratory experiments were conducted, which produced the world’s first real-time and non-obtrusive means for detecting loss of alertness. Among all driver performance and bio-behavioral measures tested, the Percentage of Eyelid Closure Over Time (Perclos) reliably predicted the most widely recognized psycho-physiological index of loss of alertness. That index is a measure of the latency between a visual stimulus and a motor response. The latency is collected using the Psychomotor Vigilance Task (PVT), whereby a subject reacts to the onset of a display that counts milliseconds and then stops the counter by pressing a button. PVT is the most valid predictor of loss of alertness, previously validated for use in medical research[1].

The strong relationship between Perclos and PVT was consistent in all subsequent validation studies, which showed that the measure was invariant to individual lapsing style (subjects who might lapse earlier in the task v. later) and the passage of time (subjects who return months later to repeat the experiment.)[1]. We also learned that individuals possess a characteristic lapsing pattern; drowsy drivers progressively underestimate the passage of time and the extent of their drowsiness; external sensory stimulation triggered by an automatic detection system is not effective and performance measures like lane deviation alone do not reliably predict loss of alertness. We did find that providing a driver with valid and real-time feedback about their alertness is the most effective means to motivate a driver to initiate self-alerting strategies, which then improves vehicle handling[4]. The payoff of this program is that international efforts are underway to develop advanced drowsiness detection technologies that use the Perclos measure as a foundation. Monitoring driver condition is no longer elusive to measurement. Whereas, there exist challenges regarding implementation, a Field Operational Test is underway to begin understanding the highway safety benefits afforded by a system that provides real time feedback to heavy vehicle drivers about their alertness. This paper provides an overview of the experimental design, data analysis, and progress toward understanding those benefits.

Problem Size

Our current understanding of the drowsy driver problem in the United States is based on NHTSA's revised estimates for the 5-year period between 1989 and 1993[2,3]. An average annual total of 6.3 million police reported crashes occurred during this period. Of these, approximately 100,000 crashes per year (1.6% of 6.3 million) were identified on Police Crash Reports (PCR) where drowsiness was indicated, and from a review of “Drift-Out-Of-Lane” crashes not specifically indicated but which had drowsiness characteristics. Approximately 71,000 of all drowsy-related crashes involved non-fatal injuries, whereas 1,357 drowsy-related fatal crashes resulted in 1,544 fatalities (3.6% of all fatal crashes), as reported by the Fatality Analysis Reporting System (FARS). Nevertheless, many run-off-roadway...
crashes are not reported or cannot be verified by police, suggesting that the problem is much larger than previously estimated.

Regarding differences between cars and trucks, approximately 96% of annual drowsy driver crashes (96,000 total including 1,429 fatalities) involved drivers of passenger vehicles, whereas only 3.3% (3,300 total including 84 fatalities) involved drivers of combination-unit trucks. However, drowsiness was cited in more truck crash involvements (.82%) than passenger vehicle crashes (.52%). In addition, the risk of a drowsiness-related crash in a combination-unit truck’s operational life is 4.5 times greater than that of passenger vehicles, because of greater exposure (60K versus 11K miles/year), longer operational life (15 versus 13 years), and more night [2,3]. There is also a greater likelihood of injury in heavy vehicle crashes. Approximately 37% of the truck-related drowsy driver fatalities and 20% of the non-fatal injuries occurred to individuals outside the truck, compared to 12% of the fatalities and 13% of the non-fatal injuries from drowsy passenger vehicle drivers.

FIELD OPERATIONAL TEST

This field test is underway to collect and analyze driver performance and alertness data between August 2004 and August 2005. There are 102 commercial drivers and 34 single-unit heavy trucks. There will be 16 weeks of data collected from each driver. Fifty-one drivers from Howell’s Trucking Company will represent long haul (cross country) operations. The remaining drivers will represent overnight express operations, 6 from Pitt-Ohio (Pennsylvania Turnpike Operations) and 45 from J.B. Hunt (Virginia Interstate Highway Operations). This arrangement was decided based on the experimental design and analysis requirements to answer the key research questions of the FOT. Whereas, the process of data collection, reduction and transfer has begun, there are no results to report in this writing.

Participation

There are three main research partners involved with this field test. First, Dr. Rich Hanowski of the Virginia Poly-Technical and State University Transportation Institute (VTTI) provides leadership and expertise in the activity of conducting the field test. Activity includes vehicle instrumentation, subject scheduling, data acquisition, data reduction, special analyses, and transmission of data to the independent evaluator. Second, working in close coordination with the “conductor”, the “independent evaluator” role includes Dr. Bruce Wilson and Dr. Steve Popkin, from the Department of Transportation’s Volpe Center in Cambridge, Massachusetts. Volpe provides expertise in the experimental design and data analyses required to answer the objectives of this research. Lastly, Dr. Richard Grace, the developer of the Perclos sensor and president of Attention Technologies in Pittsburgh, PA., supplies the conductor with the required copies of the advanced Perclos sensor.

Test Objectives

Through this research we expect to learn about 1) the nature of the distribution of drowsiness in the population of heavy vehicle drivers, and how these groups differ in their performance with and without the warning system; 2) the effects of independent factors such as driver age, health, sleep patterns, road conditions, and type of trucking operation, etc.; 3) the effect of the warning system and independent factors on conflict driving, near collisions, and severe near collisions; and 4) fleet acceptance and deployment prospects.

This paper is organized to show how the experimental design and data analyses are structured to answer the safety benefits question of the Field Operational Test (FOT.) FOT questions 3 – 5, below, are the subject of a separate paper.

1). What are the safety benefits associated with device usage?
2). What performance and capabilities does the Drowsy Driver Warning System (DDWS) have?
3). Will drivers accept the device?
4). Will fleet management purchase the device?
5). What are the deployment prospects of the DDWS?

Preliminary Tests

In October 2003, three preparatory activities were completed in advance of the FOT to verify the operational condition of the prototype equipment. The three activities included a laboratory revalidation of the Perclos metric produced by a 2nd generation Perclos sensor, the development of a Perclos system user interface suitable for commercial vehicle operations, and a study of the response characteristic of the Perclos monitor in a heavy vehicle environment. Activities addressed concerns about using the device in an operational setting. Its usability depended on the capability of the camera to detect infrared light reflected back to the source at the camera from the drivers’ retina.

Perclos revalidation was successful and involved a replication of the prospective laboratory protocol, used in two previous validation efforts [1,4]. In a second effort, Attention Technologies convened focus groups separately composed of commercial drivers and design experts to determine the essential functionality of the interface. The redesign included visual displays showing the number of total lapses, the longest lapse during the previous...
measurement interval, and the length of roadway traversed during that lapse. Drivers would then acknowledge the lapsing by pressing a button on top of the device to silence the concurrent audible warning. Lastly, Dr. Weirwille, et al. of VTTI performed a systematic characterization study of the device detecting Perclos in trucks. The study measured the sensitivity of the device to retinal pigmentation (the ability of the eye to reflect infrared light) and to the refraction of light through eyeglasses. Sensitivity was sufficient for nighttime operation with a test for retinal reflectance as a requisite for subject participation.

EXPERIMENTAL DESIGN

Team experts articulated the experimental design, which was reduced to a written specification by Dr. Bruce Wilson and Dr. Steve Popkin of the Volpe Center, and Mr. Greg Maislin of Biomedical Statistical Consulting of Wynnewood, PA. The following is based on the written specification.

Alternative designs were evaluated with consideration for 1) the maximum statistical power required for the safety benefits estimation; 2) accommodating data loss; and 3) maximizing the statistical power for the driver acceptance analysis.

The selected design is represented as follows:

<table>
<thead>
<tr>
<th>Alert Disabled</th>
<th>Alert Enabled</th>
</tr>
</thead>
<tbody>
<tr>
<td>A A A</td>
<td>B B B B B B B B</td>
</tr>
<tr>
<td>Experimental Group</td>
<td>26 drivers</td>
</tr>
<tr>
<td>A A A A A A A A A A A A</td>
<td>Control Group</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10 11 12</td>
<td>Week</td>
</tr>
</tbody>
</table>

This design includes 26 experimental participants from each trucking operation. These participants begin the study in a 3-week baseline condition, and follow with a 9-week treatment condition. The initial period is to measure baseline behavior for estimating the main effects of sleepiness, performance and crash risk factors. This design also includes 8 control participants from each trucking operation who will be monitored for the duration of the experiment, but who will not receive DDWS alerts. These 34 drivers from each trucking operation comprise the ‘core’ drivers of the experiment. Whereas, the minimum useful baseline period for the experimental group is three weeks, the corresponding maximum duration of the treatment period is 9 weeks. Therefore, this design maximizes the exposure of subjects to the device, while retaining proper experimental control of the variance for statistical analysis.

ANALYSES

SAFETY BENEFITS

DDWS Effect on Drowsiness. Analyzing the effect of the DDWS on drowsiness is a key safety benefits estimation objective. This objective addresses three research questions:

1) What is the distribution of drowsy level Perclos, and do these differ with and without the DDWS?

2) Does the distribution of drowsiness vary by driver “trait” characteristics (e.g., age, health); driver “state” characteristics (e.g., quality of previous night’s sleep, elapsed time on duty, “circadian phase”); road conditions (e.g., road type, urban/rural); and type of operation (overnight express v. long haul)?; and

3) Do differences in drowsiness between nighttime DDWS-On and DDWS-Off driving vary according to driver state characteristics, road conditions, and type of operation?

This framework of questions applies to the analysis of drowsiness, as well as further analyses where the DDWS safety benefit can be observed, i.e. driving performance, conflict driving, and near collision driving

Drowsiness level with and without the DDWS

Drowsiness is measured using a 3-minute running average of slow eyelid closures, as assessed by the DDWS during nighttime driving (P3). The distributions of Perclos P3 are compared using condition-specific summaries, duty period-specific summaries, and within-duty-period-stratified summaries as follows:

The experimental condition-specific summary is the primary summary measure for characterizing driver drowsiness within a specific condition. For example, in the comparison of conditions A3B9, median values will be determined for the 3-week baseline DDWS-Off period and the 9-week primary DDWS-On period.

Duty-period-specific summaries are computed over all nighttime epochs or measurement intervals (20 minutes) within each duty period. These statistics are used in mixed model analyses of variance that will always account for within-driver correlations across duty periods within driver and condition factors Duty period-specific mixed models admit “driver state” covariates such as prior sleep/wake history.

In further extensions of the analyses above, epochs within a duty cycle may be further stratified for groups of epochs
defined by characteristics that can vary within a duty period. Characteristics include those that reflect the homeostatic drive for sleep (reflected in elapsed time since the start of the duty period), those that reflect driving conditions such as rural/urban, road type, congestion, and other factors. For example, we will estimate the portion of the driving distance that drivers spend in each of four drowsy states: none, lo, medium, and high.

Other than analyses using Perclos P3, the DDWS provides a continuous record of the number of epochs the eyes were open (and closed). When the “number of epochs the eyes were closed” sample is divided by the total number of samples, a measure of the proportion of time that both eyes are closed can be obtained.

Since the variance of statistical estimates of proportions varies proportionally with its expected value, the arcsine variance stabilizing transformation will be employed in parametric analyses that assume variance homogeneity. These analyses would then be analogous to those performed on P3.

Drowsiness varying by independent factors

Drowsiness levels recorded by the DDWS will be summarized by relevant independent variables (i.e. age, job tenure, type of freight operation, type of driving) over clock time, consecutive workday, etc. Statistical exploration techniques will provide an analysis of the inter-relation between the selected independent variables and their effect on drowsiness levels, as measured by the DDWS. These analyses (e.g. factor analysis, correlation matrices, variance/co-variance tables) will determine the selection of variables used in subsequent hypothesis testing.

The second research question regarding the DDWS effect on drowsiness is: Does the distribution of drowsiness vary by driver “trait” characteristics (e.g., age, health); driver “state” characteristics (e.g., quality of previous night’s sleep, elapsed time on duty, “circadian phase”); and road conditions (e.g., road type, urban/rural)? As above, this question will be answered by an analysis of the interrelations between each candidate control variable and drowsiness (P3), separately for observations made under DDWS-Off and DDWS-On conditions. Specifically, P3 values for each subject under each condition will be summarized for all nighttime epochs to assess, for example, the contribution of driver traits in the variability of P3.

Differences in drowsiness by independent factors

Do differences in drowsiness between nighttime DDWS-On and DDWS-Off driving vary according to driver state characteristics, road conditions, and trucking operation? The first question in this section examined drowsiness level distributions and differences between them with and without the DDWS. This is a “Big Picture” type of question and hypothesis. The second question examined variations in drowsiness with a number of independent factors, e.g. variation in drowsiness from the quality of the previous night’s sleep. The second question identifies factors that affect drowsiness and, therefore which could confound the observed differences between drowsiness with the DDWS on and off. In order to control these variables, this 3rd question uses information from question 2 and provides a finer (stratified) analysis of drowsiness distributions. The finer stratified analysis allows us to test the hypothesis: Is the difference in P3 scores between DDWS-On and DDWS-Off the same across different characteristics?

Driving Performance. The questions and hypotheses of the next two sections are nearly identical to those of the drowsiness section, except that the topics change. In the next 2 sections, drowsiness is ignored while examining only the effect of the DDWS (Off v. On) on driver performance, conflicts, and near collisions.

Whereas, the relationship between P3 and drowsiness has been validated, it is possible that there will be beneficial (or adverse) effects of the DDWS on driving performance and crash risk that are not necessarily detected by changes in P3 for every driver. Thus, driving performance and crash risk need to be evaluated, in their own right, as potential outcomes that are affected by the DDWS.

In examining driver performance, the analyses focus on headway closing and lane keeping measures, since these crash categories largely characterize the drowsy driver crash statistics. The Measures of Performance (MOP’s) include 1) braking (number of events, peak deceleration distribution, duration of event); 2) closing (minimum range per event, minimum time to collision per event); 3) following (time gap v. vehicle speed); 4) lane changes (frequency, peak acceleration distribution, duration); 5) lane keeping (violation frequency, violation distance, boundary type, direction); and 6) speed maintenance (vehicle speed v. posted speed).

By replacing the drowsiness questions and hypotheses of the previous section with MOP’s, the questions and hypotheses about driving performance, conflict driving, and near-collision driving are stated as follows:

1) What is the nature of the MOP’s distributions, and do these differ with and without the DDWS?

2) Does the distribution of MOP’s vary by driver “trait” characteristics (e.g., age, health); driver “state” characteristics (e.g., quality of previous night’s sleep, elapsed time on duty, “circadian phase”); road conditions (e.g., road type, urban/rural); and type of operation (overnight express v. long haul)?

3) Do differences in MOP’s between nighttime DDWS-On and DDWS-Off driving vary according to driver state characteristics, road conditions, and trucking operation?
characteristics, road conditions, and type of operation?

Performance measure distributions with and without the DDWS

Each MOP is similarly analyzed. MOP’s selected for analysis are identified through exploratory techniques including factor analysis, correlation, and covariance analyses. As an example, for the number of lane boundary violations, this measure is the median number of lane violations normalized by vehicle miles traveled (VMT) for each driver. The related hypothesis is: “Driver’s median lane-boundary-violation frequencies (violations/VMT) are lower when the device is active compared to when the device is inactive.” This analysis provides a separate understanding of the distribution of MOP’s, and for identifying the most significant measures that might explain the performance benefit of the DDWS. This analysis is performed in advance of a subsequent univariate step used to understand how a single performance measure is mediated by factors such as driver traits.

Performance measures and independent factors

In the next step, each selected measure of performance will be summarized with respect to all independent variables (i.e. age, job, tenure, type of freight operation, type of driving), over clock time, consecutive workday, etc. This level of analysis separately addresses each MOP. Univariate statistics, including Analysis of Variance, and Multiple Correlation and Regression procedures will provide the analysis framework. These results will be used to understand the main and interaction effects of independent factors upon each MOP, in the presence v. absence of the DDWS.

Differences in measures with independent factors

Multivariate Regression and Correlation (MRC) analysis will examine the importance and interaction of each trait in predicting single, combination, or interacting MOP’s. The framework will provide a comprehensive means to develop a well-specified model of driver performance, as well as identifying the most significant relationships that explain the performance benefits of the DDWS. Results of the previous 2 steps will assist to identify independent variables that may serve as covariates in the MRC analysis in order to reduce the number of the most important factors in the model that explain the performance benefit of the DDWS.

Conflict Driving Analyses

Analyses of conflict driving follow the steps, as previously described for drowsiness and driver performance. The analysis framework includes the familiar progression, i.e. Conflicts with and without the DDWS, Conflict attributes and independent factors, and Differences in attributes with independent factors.”

An example analysis for “Conflicts with and without the DDWS” begins by measuring the conflict frequency for a going-straight-and-closing scenario. By computing the median number of conflicts normalized by VMT for each driver, the following hypothesis may be tested: “Drivers’ median going-straight-and-closing conflict frequency (conflicts/VMT) are lower when the DDWS is active compared to when the DDWS is inactive.” Other conflicts are similarly reduced, including near collision driving.

The next step in the analysis, “Conflict attributes and independent factors”, involves exploring each measured conflict attribute with respect to relevant independent variables, e.g. traits, driver state characteristics, road conditions. As before, this level of analysis involves univariate statistical procedures to understand how independent variables explain variability observed in each conflict attribute. The analysis is separately performed under DDWS-On and DDWS-Off conditions for each attribute. The outcome of this step, including the initial exploratory procedures (factor analysis, correlation, covariance, etc.), provides an understanding of each measure separately and is the basis for selecting factors used in the subsequent
multivariate analyses that address differences among attributes from independent factors.

CAUSAL ANALYSES

The objective of the causal analysis is to determine if the observed data support the existence of a mediating factor that explains the relationship between independent variables and an outcome variable. Mediating variables explain why an antecedent variable (independent variable) affects a consequent variable (dependent variable). For example, sleepiness is proposed to reduce driving performance, which in turn increases the rate of driving conflicts. The introduction of the intervening variable (driving performance) transforms one proposition into 2 linked propositions; from: Sleepiness leads to increases in driving conflicts, to: Sleepiness leads to driving performance reductions – driving performance reductions lead to increases in driving conflicts.

The analysis of causes will include an exhaustive exploration of candidate relations that might suggest causality. For example, the analysis will examine whether the relationship between sleepiness and conflict is mediated by driving performance; and whether the relationship between DDWS and improved driving performance is mediated through sleepiness. Results of previous statistical analyses will help identify the most likely candidates to examine for causal relationships.

This phase of the analysis is an extension of previously described exploratory, univariate, and multivariate statistical methods. However, the extension of these models to understand causality includes an analysis of covariance. In these procedures, factors are entered separately as an extension of the model containing specific continuous variables. There are optional arrangements that include adding factors as a collective group, or even as a combined representation (such as the first un-rotated principal component of a factor analysis). In these statistical model structures, using covariate extensions, we can explicitly test for any co-linear effects of intervening variables.

CRASH ESTIMATION

Crash estimation techniques depart from the analysis system previously discussed. Using crash estimation, DDWS safety benefits will be estimated using the measures of 1) the number of heavy-vehicle crashes prevented, and 2) the number of heavy-vehicle fatalities prevented. These numbers will then be expressed in economic terms (U.S. dollars saved due to crash and fatality reduction.)

The technique used will apply a “crash forecast” method by comparing forecasts of crashes with the DDWS deployed, by adjusting FOT data (conflicts and near collisions) with GES and CDS data (pre-crash scenarios and crashes without DDWS deployed).

Whereas, the latter estimation methods are based on crash conflicts, other possible approaches will include indirect methods. Indirect methods to estimate crash probabilities may include crash prevention boundary analysis, Extreme value theory, Monte Carlo simulation, and/or Severity index.

A comparative assessment of the various indirect techniques may be found elsewhere. However, the Monte Carlo method is perhaps the most suitable for the DDWS evaluation. In this method, there will be distributions formed from the FOT data before and after device activation. We will use these distributions to predict the conditional crash probabilities and crash frequencies.

PROJECT STATUS

FOT conductor and independent evaluator activities of the project are performing at an outstanding level. Each team has been challenged by changes in fleet operation, including the loss of drivers and relocation of fleet operations. However, both activities are operating on-cost and on-schedule, with completion expected in August 2005.

FUTURE WORK

Depending on a favorable outcome of this FOT, operational concepts for fleet deployment will need to be defined. In working with the Federal Motor Carrier Safety Administration (FMCSA) throughout the development of this FOT, there have been discussions that have considered the use of this device in some capacity in parallel with hours of service rules. The combination would provide a means for performance based monitoring. There are numerous other concepts, ranging from its use as a stand-alone feedback system to a system that provides alertness data to a dispatcher for altering a delivery schedule as required.

Whereas, the DDWS of this study operates between dusk and dawn, there exist international efforts to improve the detection capability of the Perclos sensor. These new systems will enable studies for understanding daytime drowsiness, and fatigue that is suddenly experienced when transitioning between levels of activity.

Future work may also include a continuation of the FOT in order to better estimate both the safety benefits and DDWS usability when the technology is deployed. A period of continued testing will ensure that our crash estimation models receive an optimal exposure to crash events in order for these models to produce the most statistically reliable benefits estimates.
CONCLUSIONS

Through the experimental design and data analyses of this FOT, further understanding is expected about highway safety benefits, fleet acceptance, operational utility, and fatigue management practices. We believe that drowsy impaired driving can be successfully mediated by advanced technology. We expect that when combined as one component of a fleet’s fatigue management strategy, the public safety benefit will be greatly multiplied. Finally, the learning accomplished by this research should assist the development of similar systems for passenger vehicle drivers, where we observe the largest prevalence of the fatigue crash problem.

REFERENCES


