# A HEAVY VEHICLE DROWSY DRIVER DETECTION AND WARNING SYSTEM: SCIENTIFIC ISSUES AND TECHNICAL CHALLENGES

**Paul Stephen Rau**
National Highway Traffic Safety Administration
United States
Paper Number 98-S2-P-30

## ABSTRACT

Even though loss of alertness has been detected in laboratory driving simulators with impressive accuracy, there are numerous scientific issues and technical challenges associated with developing a field-operational drowsiness detection and warning system. The key ***scientific issues*** are related to the development of fieldable detection models and warning systems. Issues include model validation, individualized versus generalized monitoring, and detection and warning versus activity-based maintenance. The key ***technical challenges*** are related to system operability and acceptance. Challenges include system upkeep and calibration, driver and vehicle compatibility, risk compensation and migration, alertness restoration, and operational reliability. This paper provides an overview of the drowsy driver problem in the United States, a description of NHTSA's drowsy driver technology program, and an introduction to some of the scientific issues and technical challenges that confront system deployment.

## INTRODUCTION

Research is underway at the U.S. National Highway Traffic Safety Administration to develop, test, and evaluate a prototype drowsy driver detection and warning system for commercial vehicle drivers (1996-1998). Even though the loss of alertness in drivers has been detected in laboratory driving simulators with impressive accuracy (Wierwille et al, 1996), there remain numerous scientific issues and technical challenges associated with field deployment. This paper provides an overview of the drowsy driver problem in the United States, a description of NHTSA's drowsy driver technology program, and an introduction to some of the scientific issues and technical challenges that confront system deployment.

The ***scientific issues*** discussed are related to the development of detection models and warning systems. The discussion includes the issues of model validation, individualized versus generalized monitoring, and detection and warning versus activity-based maintenance. ***Technical challenges*** relate to system operability and acceptance, including system upkeep and calibration, driver and vehicle compatibility, risk compensation and migration, alertness restoration, and operational reliability.

While the list of issues and challenges is not exhaustive, it provides an initial framework suggestive of deployment alternatives as the detection and warning system is developed. The prototype development team is presently charged to fully understand these concerns, and to complete the initial prototype by the end of fiscal year 1998. Ultimately, the final system seeks to reduce the annual numbers of injuries and deaths associated with drowsiness.

### Problem Size

Currently, our 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 (Knipling et al, 1995). 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 can not 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. Nevertheless, 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 driving (Knipling & Wang, 1994). 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.

## Drowsy Driver Technology Program

The objective of NHTSA's Drowsy Driver Technology Program is to develop, test, and evaluate a prototype drowsy driver detection and warning system for commercial motor vehicle drivers. The program began in fiscal year 1996 and is scheduled to continue through fiscal year 1998. One of the key tasks of the program is to develop drowsiness detection models and algorithms based on field data. However, laboratory-based experiments will also be conducted to suggest sensors and algorithms for further validation in the context of over-the-road driving. There are a variety of university, industry, and government partners associated with the laboratory and field study elements of the program.

First, in partnership with the University of Pennsylvania (funded by the Federal Highway Administration's Office of Motor Carriers), candidate sensors are being validated by monitoring sleep deprived subjects in a controlled laboratory setting. Subjects undergo vigilance and cognitive tests while deprived of sleep. Specifically, polysomnographic and performance measures are collected continuously; subjects are either "alerted" or "not-alerted" about their drowsiness as they become drowsy over a 20 hour period. Alerted and unalerted conditions are experimentally comparable because the presence or absence of an alerting stimuli could alter the response characteristic of certain devices. As another part of the validation process, "blind" data from the experiments are provided to the vendors of each device to determine when the drowsiness episodes occurred (prospective phase). Successful device vendors from the prospective phase receive algorithms from each of the other device vendors, as an opportunity to improve the detectability of their respective methods (retrospective phase).

Second, NHTSA's principal industry partner for building the prototype system is Carnegie Mellon Research Institute (CMRI), in Pittsburgh, Pennsylvania. CMRI is the technical lead on the project and has outfitted several commercial trucks (courtesy of Pitt-Ohio Express, Inc.) with numerous sensors and an automated data collection system. Field studies are designed to unobtrusively monitor commercial truck drivers over 10 hour overnight express runs. In the procedure, numerous performance and behavioral measures are collected as the foundation for developing detection models. This field work is guided by drowsiness detection procedures, which were developed under NHTSA sponsorship over a five year period, based on driving studies in simulators (Wierwille et al, 1996). However, a number of new detection model and algorithm approaches are also being developed and tested from the new field data, including the measures from sensors validated in the laboratory phase.

Finally, another government agency partner is the Naval Health Research Center (NHRC) in San Diego, California. NHRC provides special expertise in monitoring drowsiness from a recently developed method of processing electroencephalograph (EEG) signals. NHRC's role on the team is to assist in the development of field-based drowsiness detection models, and to provide a psychophysiological index of drowsiness previously developed under contract with the Office of Naval Research. The validity of the NHRC drowsiness detection metric will also be examined under the prospective and retrospective phases of the laboratory study.

## SCIENTIFIC ISSUES - DETECTION MODELS AND WARNING SYSTEMS

### Model Validation

Model validation is the principal activity of the program. These models derive their ability to detect changes in alertness from relationships among factors, the correlations between which are built up from data collected during observed levels of alertness. As a result, models represent relationships among the conditions required for drowsiness to be detected. For example, conditions may include drifting out of lane, excessive lane deviations, drift and jerk steering, percentage of eye closure, etc. Thus, a model might specify that a certain magnitude of deviation within a lane can be expected from a certain percentage of eye closure. In addition to prediction, models also specify the relative importance of relationships among the measures such that we might also gain an improved understanding of the important behavioral and performance components of driving.

**<u>Performance and Physiology</u>** - As the program goal is to develop a prototype system, one of our most important considerations is implementation. For example, we do not expect that commercial drivers will accept a system that requires a driver to don a cap wired with electrodes. Nevertheless, a model could be based on a measure like EEG if shown to be valid. Such a "gold" standard or yardstick by which drowsiness can be measured is important for building models that relate specific changes in physiology to driving performance. Thus, one option is to detect drowsiness based on performance inputs alone, once a strong relationship between driving performance and physiology has been established.

Another modeling option is to base the detection on a valid psychophysical index alone, if it could be measured unobtrusively. Specifically, ocular movement will soon be measured unobtrusively from within the vehicle (a 1998 NHTSA Small Business Innovative Research program initiative). This capability might provide direct access to an ocular index of drowsiness. Thus, the validation of an ocular index of drowsiness might result in: 1) models that relate driving performance to ocular measures, 2) models that relate ocular measures to other previously validated psychophysical indices of drowsiness, and/or 3) models that relate driving performance to ocular and/or other valid indices.

**Normative Weighting and Event-Driven Models** - It is possible that quantitative models alone can *not* be produced from the measures obtained in the field study. Therefore, it is an option to explore improving the quantitative models with various qualitative data related to normative trends in drowsy driving. For example, according to data from NHTSA's General Estimates System (GES), police reported drowsy related crashes occur most frequently between 1:00 a.m. and 5:00 a.m., and again in late afternoon between 3:00 p.m. and 6:00 p.m. Information is also available regarding the number of drowsy related crashes, based on the number of hours driven. Therefore, in a normatively-weighted model, a qualitative rule could be used to mediate the alarm/warning threshold of a data-driven model according to population trends.

Similarly, model validity might also be improved using knowledge about events that occur during the particular time-line of travel. For example, information about the frequency of stops, the duration of stops, regularity of speed, number of passengers, changes in air flow and temperature, and noise levels, could be measured and used to modify the detection capability of the model. Such an algorithm would detect departures from previously determined normative levels. For example, the alarm/warning threshold of a detection system could be lowered when there is an absence of an environmental change; a monotonous environment might indicate a pre-condition for drowsiness. Therefore, the validity of a quantitative detection model might be improved using qualitative information about the population of drivers and/or about the experience of a particular driver. It is also possible that the most useful detection model might be based on the qualitative information alone.

In sum, there are numerous modeling opportunities, all of which offer promise in producing an operational system. As a result, the prototype system could be based on some combination of driving performance, ocular behavior, and/or the inclusion of normative and event based heuristics.

### Individualized vs Generalized Models

Individualized versus generalized models are distinguished as those which either detect loss of alertness in a single driver or among all drivers, respectively. The issue is that quantitative models utilize the response data from only a small sample of drivers. Therefore, predictions about a larger population of drivers must be derived statistically. Nevertheless, any large differences among individual drivers could overwhelm any otherwise significant effect related to the group. It is, therefore, likely that a detection model could be improved by using specific knowledge about an individual driver. Moreover, individualized models could include normative or event information, as previously described. Lastly, some technologies have been shown to detect an individual drowsy "signature". For example, certain classes of neural networks can learn baseline driver behavior, and then warn the driver regarding departures from normal "alert" patterns. Both group-based and individualized models are potential outcomes of the research.

### Detection and Warning vs Activity-Based Maintenance

Detection and warning versus activity-based maintenance is an issue that contrasts the detection modeling approach of our program, with an "activity-based" approach that requires continuous driver interaction. For example, there are several devices that could alert the driver when a specific behavior fails. One device sounds an alarm when any change in steering wheel motion stops. Presumably, moments of motionless steering may indicate that the driver has fallen asleep. However, there are numerous differences in driving style and roadway conditions that result in motionless steering. Thus, when avoiding frequent alarms during normal periods of motionless steering, steering could become erratic and unsafe.

Another device measures a driver's reaction time to a small light, which is illuminated following a *random* elapsed period of time. A button must be pressed within 3 seconds after the small light is illuminated or else a buzzer sounds. This secondary task forces the driver to monitor a specific location inside the vehicle for the random occurrence of a single light. Monitoring a random event, not related to vehicle operation, could dangerously divide attention away from the roadway and mirrors. Still another device comprises two alarms; one, if a button is not pressed before an *adjustable* time interval, and two, if a second button is not pressed within another adjustable time period following the occurrence of the first alarm.

There are many compromises to driver safety in using activity-based alertness maintenance devices. Nevertheless, some form of activity-based device, which does not interfere with safe driving, might provide a useful countermeasure to drowsiness. Perhaps a future system might offer some combination of passive detection and alarm/warning methods, with an activity-based system.

## TECHNICAL CHALLENGES - OPERABILITY & ACCEPTANCE

### System Upkeep & Calibration

System upkeep and calibration is perhaps the most important technical challenge in designing a generally useful system. The system must be easy to learn, easy to use, and easy to maintain. However, certain sensors might be more difficult for passenger vehicle owners to maintain and calibrate on a regular basis. For commercial carriers, upkeep and calibration might be achieved during regular periods of maintenance.

The difficulty of this challenge depends on which sensors are required to support a valid detection model. For example, a camera-based lanekeeping system might require regular lens cleaning and alignment checking. Thus, for a non-technically oriented consumer, the system might also require a performance monitoring and fault localization (PMFL) device to automatically inform drivers if their system performance degrades. Our challenge, then, is that regardless of how valid and reliable in detecting drowsiness, the fielded system must be easily maintained and calibrated.

### Driver-Vehicle Compatibility

Driver-vehicle compatibility is presented not so much as a challenge of system design, but as an activity for building engineering models of driver-vehicle interaction. There exist various guidelines on vehicle *interfaces*, but there are no known models that specifically address driver-vehicle *interaction*. Such models would constitute computational methods for predicting human performance in vehicles. As a start, cognitive models of driver-vehicle interaction could arise from, as well as contribute to the existing wealth of knowledge from cognitive science and cognitive psychology. The challenge is to focus that knowledge as an organized framework of methods, whereby quantitative models of driver/vehicle interaction may be developed. Other specialty areas in human factors have previously begun this process. For example, in the area of user-computer interaction, there exist a number of models, which characterize the interaction (not necessarily the interface) between users and computer systems. Many of the basic components of previous models of human-machine dialog could also be applied to develop predictive models of driver-vehicle interaction. The present research contributes to this knowledge base, specifically with regard to the models developed that specify relationships between physiology and performance.

### Risk Compensation & Migration

Risk compensation and migration relate to diminished operational effectiveness due to the misuse of a countermeasure by drivers, as well as external sources of probability <u>not</u> associated with the detectability of a device. First, risk compensation refers to the undesired use of a countermeasure that reduces a driver's awareness of the actual risks associated with certain risky driving behaviors. For example, depending on how the system reports loss of alertness, drivers may use the information to continue driving. It is well known that drivers are often motivated to keep driving, even under impaired levels of drowsiness. Drivers will persist in driving drowsy for many reasons including proximity to their destination, safety concerns about sleeping at rest areas, lodging alternatives, and delays in schedule, etc. The technical challenge is to minimize risk compensation through the design of the user interface. For example, a continuous "fuel gauge" display of alertness might encourage drivers to continue driving, whereas a single threshold alarm would communicate that falling asleep at the wheel is imminent.

Second, risk migration refers to externally determined probabilities, which can affect the overall performance of the system. For example, there have been informal reports suggesting that with roadside rumble strips, there are fewer run-off-road crashes for those road segments that contain rumble strips. However, overall, the same number of crashes occur on that particular highway. It is as though the incidences "migrate" to subsequent road segments without the rumble strips. There are no models to predict this phenomenon, but it suggests that there are other probabilities involved that could influence the effectiveness of the system. Therefore, part of the challenge is confronting problems that are unexpected.