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 <u>not</u> 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 monitor, 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:

Alert Disabled	Alert Enabled	
A A A	B B B B B B B B B	Experimental Group 26 drivers
A A A	A A A A A A A A A	Control Group 8 drivers
1 2 3	4 5 6 7 8 9 10 11 12	
	Week	

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

<u>DDWS Effect on Drowsiness</u>. 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 (P_3). The distributions of Perclos P_3 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 A_3B_9, 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 P_3, 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 P_3.

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-relations 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 (P_3), separately for observations made under DDWS-Off and DDWS-On conditions. Specifically, P_3 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 P_3.

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 P_3 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 P_3 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 P_3 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)?; and

3) Do differences in MOP's between nighttime DDWS-On and DDWS-Off driving vary according to driver state

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. Previous analyses do not distinguish between safe and risky driving. In this analysis, we begin with the crash statistics. Crash statistics provide the dominant drowsy driver pre-crash scenarios. The DAS supports the identification of these scenarios in the FOT data. Pre-crash scenarios are classified as combinations of vehicle movement and critical event pairs. For example, Going straight – Departed road; Going straight – Other vehicle slower; Going straight – Lost control; Negotiating a curve – Departed roadway; etc. Using crash statistics based on 1997-2001 GES data, drowsy crashes are classified and distributed among the scenarios.

Once itemized, our familiar framework of research questions is then applied, and then similarly analyzed as in the drowsiness and driver performance sections. For example, without restating the entire sequence, the first question is, "What is the nature of the *conflict attributes* and do these differ with and without the DDWS."

Conflict attributes used to characterize a conflict include:
1) Frequency – Expressed in conflicts per VMT.
2) Initial conditions –Typically the closing speed, binned over several ranges.
3) Response timing – Typically the time-to-collision for closing conflicts or time-to-departure for lateral conflicts.
4) Response intensity – Typically the longitudinal and or lateral acceleration.

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

[1]Dinges, D.F., Mallis, M.M., Maislin, G., and Powell, J.W., *Evaluation of Techniques for Ocular Measurement as an Index of Fatigue and the Basis for Alertness Management*. DOT HS 808 762, April 1998.

[2]Knipling, R.R., and Wang, J.S., *Crashes and Fatalities Related to Driver Drowsiness/ Fatigue.* National Highway Traffic Safety Administration, Research Note, November, 1994.

[3]Knipling, R.R., Wang, J.S., and Kanianthra, J.N., Current NHTSA Drowsy Driver R&D. *Proceedings of the 15th International Technical Conference on Enhanced Safety of Vehicles (ESV)*, Melbourne, Australia, May 13-17, 1995.

[4]Mallis, M.M., et al., *BioBehavioral Res[ponses to Drowsy Driving Alarms and Alerting Stimuli*, DOT HS 809 202, February 2000.

[5]Rau, P.S., *A Heavy Vehicle Drowsy Driver Detection and Warning System: Scientific Issues and Technical Challenges*, Stress, Workload and Fatigue, Hancock and Desmond (Eds), Earlbaum, 2001.

[6]Wierwille, W.W., Lewin, M.G., Fairbanks, RJ, and Neal, V.L., *Research on Vehicle-Based Driver Status/Performance Monitoring; 3 Volume Final Report*. Vehicle Analysis and Simulation Laboratory, Virginia Polytechnic Institute and State University, VPSIU Report No. ISE 96-06 through 96-08, Contract Number DTNH 22-91-Y-07266, NHTSA Report No. TBD, September, 1996.

AN OVERVIEW OF THE 100-CAR NATURALISTIC STUDY AND FINDINGS

Vicki L. Neale
Thomas A. Dingus
Sheila G. Klauer
Jeremy Sudweeks
Virginia Tech Transportation Institute
Michael Goodman
National Highway Traffic Safety Administration
United States
Paper Number 05-0400

ABSTRACT

A key to the development of effective crash countermeasures is an understanding of pre-crash causal and contributing factors. This research effort was initiated to provide an unprecedented level of detail concerning driver performance, behavior, environment, driving context and other factors that were associated with critical incidents, near crashes and crashes for 100 drivers across a period of one year. A primary goal was to provide vital exposure and pre-crash data necessary for understanding causes of crashes, supporting the development and refinement of crash avoidance countermeasures, and estimating the potential of these countermeasures to reduce crashes and their consequences.

The 100-Car Naturalistic Driving Study database contains many extreme cases of driving behavior and performance, including severe fatigue, impairment, judgment error, risk taking, willingness to engage in secondary tasks, aggressive driving, and traffic violations. The data set includes approximately 2,000,000 vehicle miles, almost 43,000 hours of data, 241 primary and secondary drivers, 12 to 13 months of data collection for each vehicle, and data from a highly capable instrumentation system including five channels of video and vehicle kinematics. From the data, an "event" database was created, similar in classification structure to an epidemiological crash database, but with video and electronic driver and vehicle performance data. The events are crashes, near crashes and other "incidents." Data was classified by pre-event maneuver, precipitating factor, event type, contributing factors, and the avoidance maneuver exhibited. Parameters such as vehicle speed, vehicle headway, time-to-collision, and driver reaction time are also recorded.

This paper presents the 100-Car Naturalistic Driving Study method, including instrumentation and vehicle characteristics, and a sample of study results. Presented analyses address the driver characteristics, the role of inattention and distraction in rear-end and lane change events. In addition, the methodological attributes of naturalistic data collection and the implications for a larger-scale naturalistic data collection effort are provided.

INTRODUCTION

Although the crash rate is declining, the number of driving related deaths is approximately 43,000 per year. While the development of mechanistic safety features, such as seat belts, air bags, and collapsible steering wheels, have been extremely important in lowering the vehicle-related death rate, it is plausible that the next significant decrease in roadway fatalities will require systems to assist drivers in preventing crashes. However, driver assistance systems require a more precise understanding of the driver behaviors prior to an adverse driving event to be more effective.

Data collected to study driver behavior have historically relied on epidemiological, simulator, and test track studies. While these are valuable techniques that certainly have their place in the study of driver behavior, they are not well suited to explain the combination of factors leading to an adverse driving event. For example, a police crash report form might list the cause of a rear-end collision as "following too close." However, contributing factors might be fatigue, distraction, traffic backed up from the intersection, and/or a blind corner leading up to the same intersection. For this hypothetical case, there are both driver and infrastructure related causes of the event. Likewise, simulator and test track studies cannot mimic the combination of complex driving environments and the simultaneous array of driver behaviors that lead to many events.

As demonstrated in only a small handful of studies, naturalistic data collection fills the gap in current data collection methods. "Naturalistic" data includes data from a suite of vehicle sensors and unobtrusively placed video cameras. The drivers are given no special instructions, no experimenter is present, and the data collection instrumentation is unobtrusive. This naturalistic data collection method was applied to study fatigue and resulting driver performance in truck drivers making local/short haul deliveries [1] In this study, 42 drivers drove 4 instrumented vehicles while they made deliveries. The study resulted in approximately 1000 hours of data that included five video views and a host of vehicle sensor data.

In a long-haul truck driving study, naturalistic data was collected from 56 single and team drivers who drove one of two instrumented vehicles [2]. Data was collected to assess sleep quality, driver alertness, and driver performance on normal revenue-producing trips averaging up to eight days in length. This data collection effort resulted in 250 hours of data that was triggered based upon vehicle sensor data. The results showed that single drivers suffered the worst bouts of fatigue and had the most severe critical incidents (by about 4 to 1).

A key to the development of effective crash countermeasures is an understanding of pre-crash causal and contributing factors. This research effort was initiated to provide an unprecedented level of detail concerning driver performance, behavior, environment, driving context and other factors that were associated with critical incidents, near crashes and crashes for 100 drivers across a period of one year. A primary goal was to provide vital exposure and pre-crash data necessary for understanding causes of crashes, supporting the development and refinement of crash avoidance countermeasures, and estimating the potential of these countermeasures to reduce crashes and their consequences.

The 100-Car Naturalistic Driving Study (100-Car Study) was the first instrumented vehicle study undertaken with the primary purpose of collecting large-scale naturalistic driving data. Unique to the 100-Car Study was that the majority of the drivers drove their own vehicles (78 out of 100 vehicles). There is every indication that the drivers rapidly disregarded the presence of the instrumentation, as is indicated by the resulting database containing many extreme cases of driving behavior and performance including: severe fatigue, impairment, judgment error, risk taking, willingness to engage, aggressive driving, and traffic violations (just to name a few). These types of driving events have been heretofore greatly attenuated by other empirical techniques.

Due to the scale of the 100-Car Study and the fact that private vehicles were instrumented, new techniques had to be created and existing methods modified to make the study successful. The data collection effort resulted in the following data set contents:

- Approximately 2,000,000 vehicle miles
- Almost 43,000 hours of data
- 241 primary and secondary drivers participated
- 12 to 13 month data collection period for each vehicle
- Five channels of video and many vehicle state and kinematic variables

This paper presents a sample of the analysis results from the 100-Car Study data collected. The full study report is available through the National Highway Traffic Safety Administration [3].

METHOD

Instrumentation

The 100-Car instrumentation package was engineered by VTTI to be rugged, durable, expandable, and unobtrusive. It constituted the seventh generation of hardware and software, developed over a 15 year period that has been deployed for a variety of purposes. The system consisted of a Pentium-based computer that received and stored data from a network of sensors distributed around the vehicle. Data storage was achieved via the system's hard drive, which was large enough to store data for several weeks of driving before requiring data downloading.

Each of the sensing subsystems in the car was independent, so that any failures that occurred were constrained to a single sensor type. Sensors included a vehicle network box that interacted with the vehicle network, an accelerometer box that obtained longitudinal and lateral kinematic information, a headway detection system to provide information on leading or following vehicles, side obstacle detection to detect lateral conflicts, an incident box to allow drivers to flag incidents for the research team, a video-based lane tracking system to measure lane keeping behavior, and video to validate any sensor-based findings. The video subsystem was particularly important as it provided a continuous window into the happenings in and around the vehicle. This subsystem included five camera views monitoring the driver's face and driver side of the vehicle, the forward view, the rear view, the passenger side of the vehicle, and an over-the-shoulder view for the driver's hands and surrounding areas. An important feature of the video system is that it was digital, with software-controllable video compression capability. This allowed synchronization, simultaneous display, and efficient archiving and retrieval of 100-Car data. A frame of compressed 100-Car video data is shown in Figure 1.

The modular aspect of the data collection system allowed for integration of instrumentation that was not essential for data collection, but which provided the research team with additional and important information. These subsystems included automatic collision notification that informed the research team of the possibility of a collision; cellular communications that were used by the research team to communicate with vehicles on the road to determine system status and position; system initialization equipment that automatically controlled system status; and a GPS positioning subsystem that collected information on vehicle position. The GPS positioning subsystem and the cellular communications were often used in concert to allow for vehicle localization and tracking.

Figure 1. A compressed video image from the 100-Car data. The driver's face (upper left quadrant) is distorted to protect the driver's identity. The lower right quadrant is split with the left-side (top) and the rear (bottom) views.

The system included several major components and subsystems that were installed on each vehicle. These included the main Data Acquisition System (DAS) unit that was mounted under the package shelf for the sedans (Figure 2) and behind the rear seat in the SUVs.

Doppler radar antennas were mounted behind special plastic license plates on the front and rear of the vehicle (Figure 3). The location behind the plates allowed the vehicle instrumentation to remain inconspicuous to other drivers.

Figure 2. The main Data Acquisition System (DAS) unit mounted under the "package shelf" of the trunk.

Figure 3. Doppler radar antenna mounted on the front of a vehicle, covered by one of the plastic license plates used for this study.

The final major components in the 100-Car hardware installation were mounted above and in front of the center rear-view mirror. These components included an "incident" pushbutton box which housed a momentary pushbutton that the subject could press whenever an unusual event happened in the driving environment. Also contained in the housing was an unobtrusive miniature camera that provided the driver face view. The camera was invisible to the driver since it was mounted behind a "smoked" Plexiglas cover.

Mounted behind the center mirror were the forward-view camera and the glare sensor (Figure 4). This location was selected to be as unobtrusive as possible and did not occlude any of the driver's normal field of view.

Figure 4. The incident push button box mounted above the rearview mirror. The portion on the right contains the driver face/left vehicle side camera hidden by a smoked plexiglass cover.

Subjects

One-hundred drivers who commuted into or out of the Northern Virginia/Washington, DC metropolitan area were initially recruited as primary drivers to have their vehicles instrumented or receive a leased vehicle for this study. Drivers were recruited by placing flyers on vehicles as well as by placing newspaper announcements in the classified section. Drivers who had their private vehicles instrumented (78) received $125.00 per month and a bonus at the end of the study for completing necessary paperwork. Drivers who received a leased vehicle (22) received free use of the vehicle, including standard maintenance, and the same bonus at the end of the study for completing necessary paperwork. Drivers of leased vehicles were insured under the Commonwealth of Virginia policy.

As some drivers had to be replaced for various reasons (for example, a move from the study area or repeated crashes in leased vehicles), 109 primary drivers were included in the study. Since other family members and friends would occasionally drive the instrumented vehicles, data were collected on 132 additional drivers.

A goal of this study was to maximize the potential to record crash and near-crash events through the selection of subjects with higher than average crash- or near-crash risk exposure. Exposure was manipulated through the selection of a larger sample of drivers below the age of 25, and by the selection of a sample that drove more than the average number of miles. The age by gender distribution of the primary drivers is shown in Table 1. The distribution of miles driven by the subjects during the study appears as Table 2. As presented, the data are somewhat biased compared to the national averages in each case, based on TransStats, 2001 [4]. Nevertheless, the distribution was generally representative of national averages when viewed across the distribution of mileages within the TransStats data.

One demographic issue with the 100-Car data sample that needs to be understood is that the data were collected in only one area (i.e., Northern Virginia/Metro Washington, DC). This area represents primarily urban- and suburban driving conditions, often in moderate to heavy traffic. Thus, rural driving, as well as differing demographics within the U.S., are not well represented.

Table 1. Driver age and gender distributions.

	Gender		
Age N % of total	Female	Male	Grand Total
18-20	9 8.3%	7 6.4%	16 14.7%
21-24	11 10.1%	10 9.2%	21 19.3%
25-34	7 6.4%	12 11.0%	19 17.4%
35-44	4 3.7%	16 14.7%	20 18.3%
45-54	7 6.4%	13 11.9%	20 18.3%
55+	5 4.6%	8 7.3%	13 11.9%
Total N Total Percent	43 39.4%	66 60.6%	109 100.0%

Table 2. Actual miles driven during the study.

Actual miles driven	**Number of Drivers**	**Percent of Drivers**
0-9,000	29	26.6%
9,001-12,000	22	20.2%
12,001-15,000	26	23.9%
15,001-18,000	11	10.1%
18,001-21,000	8	7.3%
More than 21,000	13	11.9%

A goal of the recruitment process was to attempt to avoid extreme drivers in either direction (i.e., very safe or very unsafe). Self reported historical data indicate that a reasonably diverse distribution of drivers was obtained.

Vehicles

Since 100 vehicles had to be instrumented with a number of sensors and data collection hardware, and since the complexity of the hardware required a number of custom mounting brackets to be manufactured, the number of vehicle types had to be limited for this study. Six different vehicle models were selected based upon their prevalence in the Northern Virginia area. These included five sedan models (Chevrolet Malibu and Cavalier, Toyota Camry and Corolla, and Ford Taurus) and one SUV model (Ford Explorer). The model years were limited to those with common body types and accessible vehicle networks (generally 1995 to 2003). The distribution of these vehicle types was:

- Toyota Camry – 17%
- Toyota Corolla – 18%
- Chevy Cavalier – 17%
- Chevy Malibu – 21%
- Ford Taurus – 12%
- Ford Explorer – 15%

Classification of events

Table 3 provides definitions of traffic "events" that served as a basis for the classifications that follow. The distinction between *near crashes* and *incidents* was based on the subjective assessment of reviewers in concert with kinematic and proximity data associated with adjacent vehicles or objects.

RESULTS

Table 4 shows the relative frequency of crashes, near-crashes, and incidents for each conflicts type. Of the 82 crashes, 13 either occurred while the system was initializing after the vehicle ignition was started (approximately 90 seconds), or has incomplete data for other reasons (e.g., camera failure), leaving a total of 69 crashes for which data could be completely reduced. These data also included 761 near-crashes and 8,295 incidents. The first eight conflict types shown in Table 4 accounted for all of the crashes, 87 percent of the near-crashes and 93 percent of the incidents.

Table 3. Classification of Events.

Event Category	Definition
Crashes	Any contact between the subject vehicle and another vehicle, fixed object, pedestrian pedacyclist, animal
Near Crashes	Defined as a conflict situation requiring a rapid, severe evasive maneuver to avoid a crash.
Incidents	Conflict requiring an evasive maneuver, but of lesser magnitude than a near crash

It is important to note that all of the crashes, including low speed collisions that were not police reported, are shown in Table 5. A "crash" was operationally defined as "any measurable dissipation or transfer of energy due to the contact of the subject vehicle with another vehicle or object." A benefit of the naturalistic approach is that it was possible to record all of these events; however the severity of the crashes must be delineated to better understand the data. Thus, the 69 crashes are parsed into the following four crash categories. Note that 75 percent of the single vehicle crashes were low-g force physical contact or tire strikes; in other words, most of the crashes involved very minor physical contact.

- Level I: Police-reported air bag deployment and/or injury
- Level II: Police-reported property damage only
- Level III: Non-police-reported property damage only
- Level IV: Non-police-reported low-g physical contact or tire strike (greater than 10 mph)

Since it was possible to detect all crashes regardless of severity, it is interesting to note the large number of drivers who experienced one or more collisions during the 12 to 13 month data collection period. Of all drivers, 7.5% of drivers never experienced an event of any severity. In contrast, 7.4% of the drivers experienced many incidents and 3 or 4 crashes. Thus, a handful of subjects were either very risky drivers or very safe, with the majority of drivers demonstrating a relatively normal distribution of events across the data collection period.

Table 4. Number of crashes, near-crashes, and incidents for each conflict type.

Conflict Type	Crash	Near-crash	Incident
Single vehicle	24	48	191
Lead-vehicle	15	380	5783
Following vehicle	12	70	766
Object/obstacle	9	6	394
Parked vehicle	4	5	83
Animal	2	10	56
Vehicle turning across subject vehicle path in opposite direction	2	27	79
Adjacent vehicle	1	115	342
Other	0	2	13
Oncoming traffic	0	27	184
Vehicle turning across subject vehicle path in same direction	0	3	10
Vehicle turning into subject vehicle path in same direction	0	28	90
Vehicle turning into subject vehicle path in opposite direction	0	0	1
Vehicle moving across subject vehicle path through intersection	0	27	158
Merging vehicle	0	6	18
Pedestrian	0	6	108
Pedalcyclist	0	0	16
Unknown	0	1	3

Table 5. Crash type by crash severity level.

Conflict Type	Total	Level I	Level II	Level III	Level IV
Single vehicle	24	1	0	5	18
Lead-vehicle	15	1	3	5	6
Following vehicle	12	2	2	5	3
Object/obstacle	9	0	1	3	5
Parked vehicle	4	0	0	2	2
Animal	2	0	0	0	2
Oncoming vehicle turning across subject vehicle path	2	1	1	0	0
Adjacent vehicle	1	0	0	1	0

Characterization of Driver Inattention

Historically, driver distraction has been typically discussed as a secondary task engagement. Fatigue has also been described as relating to driver inattention. In this study, it became clear that the definition of driver distraction needed to be expanded to a more encompassing 'driver inattention' construct that includes *secondary task engagement* and *fatigue* as well as two new categories, '*Driving-related inattention to the forward roadway*' and '*non-specific eye glance*'. '*Driving-related inattention to the forward roadway*' involves the driver checking rear-view mirrors or their blind spots. This new category was added after viewing multiple crashes, near-crashes, and incidents for which the driver was clearly paying attention to the driving task, but was not paying attention to the *critical aspect* of the driving task (i.e., forward roadway) at an inopportune moment involving a precipitating factor.

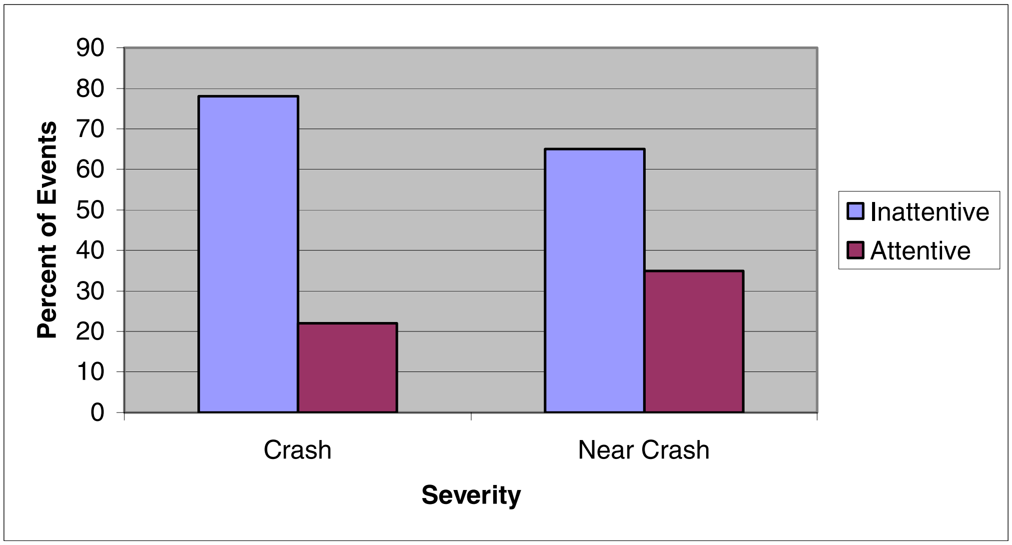

Figure 5. Percentage of events for attention by severity level.

A second analysis of the crashes and near-crashes in the 100-Car database was also conducted using the eye glance analysis performed manually by data reductionists. The '*non-specific eyeglance away from the forward roadway'* describes cases for which drivers glanced, usually momentarily, away from the roadway, but at no discernable object or person. For this project, eye glance reduction was accomplished for crash and near-crash events only, so this category can only be used for the more severe events. The four inattention categories identified above and considered together, suggested that driver's glances away from the forward roadway potentially contribute to a much greater percentage of events than has been previously thought. As shown in Figure 5, 78 percent of the crashes and 65 percent of the near crashes had one of these four inattention categories as a contributing factor.

An analysis of these types of inattention revealed that secondary task distraction was the largest of the four categories. The sources of inattention that generally contributed to the highest percentages of events (Figure 6) were wireless devices (primarily cell phones) internal distractions, and passenger-related secondary tasks (primarily conversations). It is important to note that "exposure," the frequency and duration of inattention associated with each source of inattention, is not considered in these data. Since it is exposure that determines the overall risk of a distraction source, an analysis of frequency of device use is currently being conducted for a future report that will allow calculations of event rates to determine estimates of the relative risk associated with these tasks.

Figure 7 shows a breakdown of the wireless device tasks and associated events. For these data, all of the crashes (about 8.7 percent of total study crashes) and a majority of the near crashes and incidents occurred during a cell phone conversation, although the dialing task was relatively high in term of total conflicts and was associated with the largest number of near crashes for this source of inattention. Although these data are important in that they represent the factors that contribute to events, they also highlight the need for the exposure data described above to establish the degree of risk.

Inattention for Rear End Lead-Vehicle Scenarios

Of particular interest in the analyses of rear-end conflict contributing factors was the prevalence of distraction. An important aspect in rear-end crash countermeasure development is the degree to which an un-alerted driver can be warned and make a proper response. Of course, the 100-Car data can provide great insight into the degree to which distraction is an issue in such conflicts. The important finding in this regard is that 93 percent of all lead vehicle crashes (13 out of 14) involved *inattention to the forward roadway* as a contributing factor (Figure 8). Note also that a majority (68 percent) of the near crashes have inattention identified as a contributing factor.

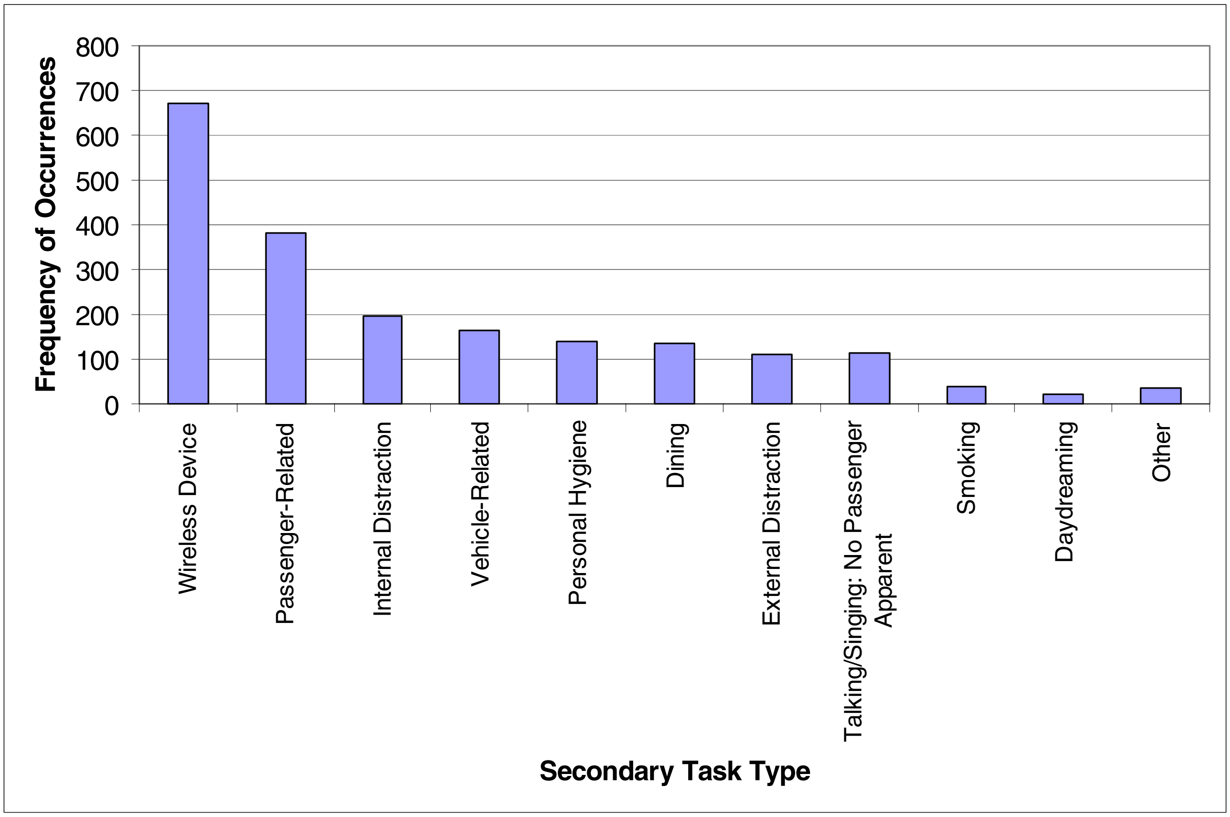

Figure 6. Frequency of occurrence of secondary tasks for crashes, near crashes and incidents.

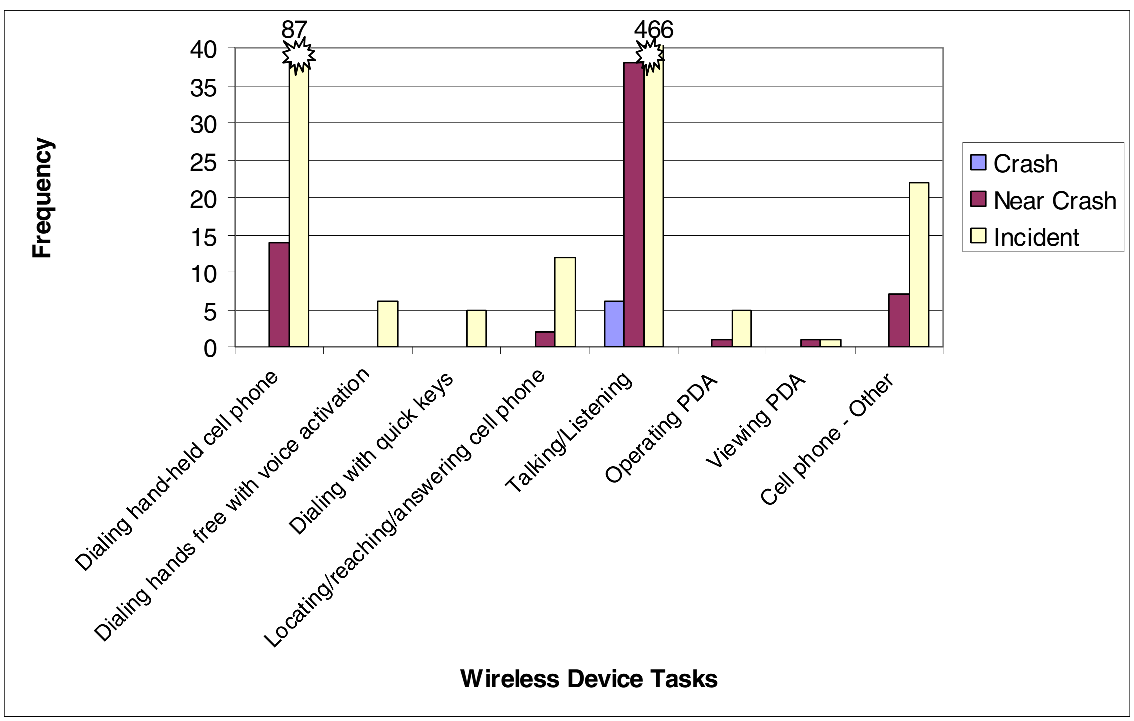

Figure 7. Frequency of occurrences in which the contributing factor was wireless device use by level of severity.

Figure 9 shows the frequency of each source of inattention for each of the secondary tasks. This allows comparison of the actual contribution of each of these sources of inattention to lead vehicle conflicts. Wireless devices (primarily cell phones, but also including PDAs) were the most frequent contributing factor for lead vehicle events, followed by passenger-related inattention. The trend was very similar for near-crashes. Interior distractions were the most frequent source of inattention for crashes.

While cell phone use contributed much more frequently to incidents and near-crashes than any other secondary task, cell phone use did not contribute to any lead vehicle conflict crashes. Nevertheless, cell phone use did contribute to other types of crashes, such as run off road, single vehicle conflict (driver ran into a barricade), and following vehicle conflict (subject vehicle was struck).

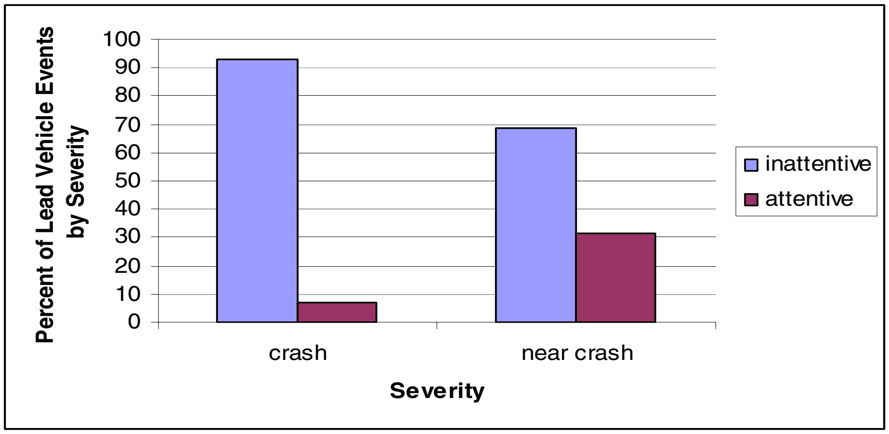

Figure 8. Percent of lead vehicle events for which inattention was listed as a contributing factor (includes the non-specific eye glance events for crashes and near crashes).

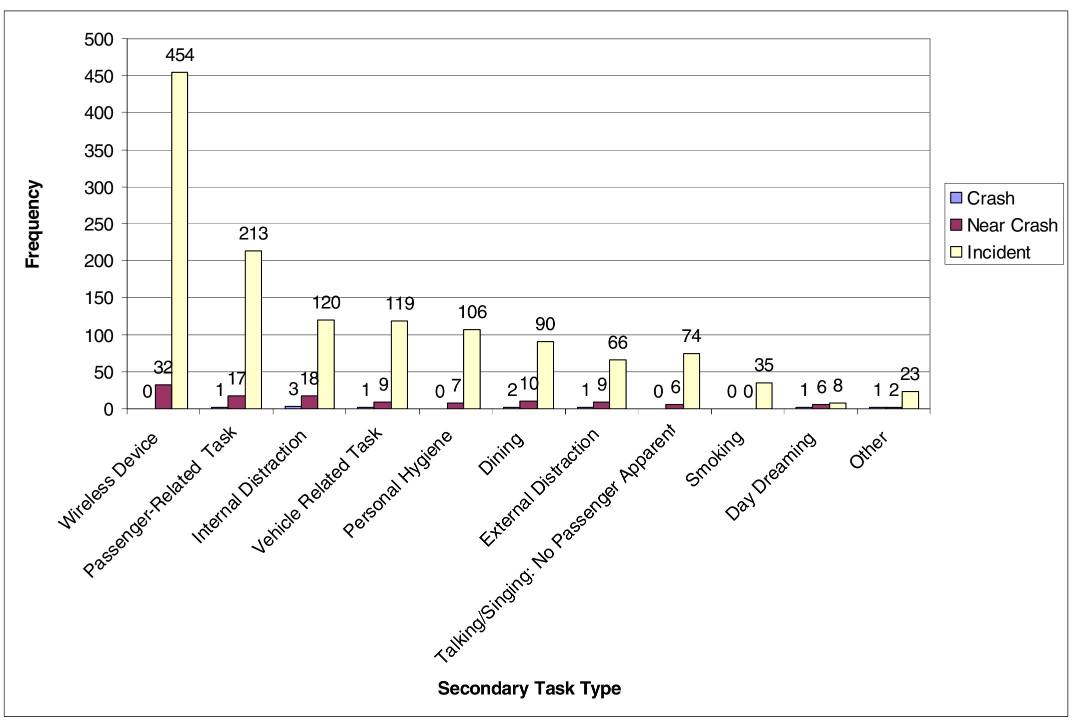

Figure 9. Total frequency of secondary task type by severity.

SUMMARY AND CONCLUSIONS

The event database that was created during the 100-Car Study can be useful for a variety of purposes; for example, evaluation of risky driving behavior and crash risk, calculation of relative risk of engaging in secondary tasks, and evaluation of driver response to lead vehicle brake lights. To facilitate this process, the initial event database will be made publicly accessible via the Internet. In addition, the initial event database can be expanded to address additional issues, since all of the video and electronic data for the entire study have been archived. The 100-Car Study contract specified ten objectives or goals that would be addressed through the initial analysis of the event database. However, as of the time of this writing, there are three additional data reduction and analysis efforts underway for the purpose of addressing another eight goals, and there is considerable interest in using the data for even more purposes. Progressing toward this potential for a multi-purpose, highly flexible and adaptable tool for driving safety may be the most important aspect of this study.

Despite the massive scope of the current effort, it was designed to serve as an exploratory study to a determine the feasibility, value, and methods for initiating a larger, more representative study. From an epidemiological viewpoint, the study was small with the presence of 15 police-reported and 82 total crashes, including minor collisions. Furthermore, drivers were represented from one area of the country (Northern Virginia/Washington, DC metro area). One purpose of a large-scale study would be to have a statistically representative sample of crashes (perhaps 2,000) and a more representative driver/environment sample.

The challenge of a large-scale study is not only the expense of such data collection but the management and analysis of such a large body of data. Nevertheless, it is believed that a large-scale database would be an enormous asset and would be used by transportation researchers for many years to gain insight and understanding into a wide array of driving behavior issues and potentially serve as a basis for decision making and program development within both the government and business sectors. This belief is based upon the robustness of the study results and the expectation that these data will continue to be analyzed and the results made available, from a variety of researchers and research organizations. Clearly, these data can provide unique insights into issues that have eluded the highway safety community for years.

REFERENCES

1. Hanowski, R. J., Wierwille, W. W., Garness, S. A. and Dingus, T. A. (2000). *Impact of local/short haul operations on driver fatigue: Final project report* (DOT-MC-00-203). Washington, DC: Federal Motor Carrier Safety Administration.

2. Dingus, T. A., Neale, V. L., Garness, S. A., Hanowski, R. J., Keisler, A. S., Lee, S. E., Perez, M. A., Robinson, G. S., Belz, S. M., Casali, J. G., Pace-Schott, E. F., Stickgold, R. A., Hobson, J. A. (2001). *Impact of sleeper berth usage on driver fatigue: Final project report* (PB2002-107930). Springfield, VA: NTIS.

3. Dingus, T. A., Klauer, S.G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., and Jermeland, J. (In press). *The 100-car naturalistic driving study, Phase II - results of the 100-car field experiment* (Contract No. DTNH22-00-C-07007). Washington, DC: National Highway Traffic Safety Administration.

4. Bureau of Transportation Statistics, TransStats 2001 (http://www.transtats.bts.gov/Databases.asp?Mode_ID=3&Mode_Desc=Highway&Subject_ID2=0)