U.S. Department
of Transportation
National Highway
Traffic Safety
Administration

www.nhtsa.gov

DOT HS 811 117 April 2009

Assessment of a Drowsy Driver Warning System for Heavy-Vehicle Drivers

Final Report

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank) DOT 811 117	2. REPORT DATE April 2009	3. REPORT TYPE AND DATES COVERED Final Report, August 29, 2007 – October 28, 2008
4. TITLE AND SUBTITLE Assessment of a Drowsy Driver Warning System for Heavy-Vehicle Drivers: Final Report		5. FUNDING NUMBERS DTNH22-05-D-01019, Task Order #18
6. AUTHOR(S) Part I Authors: Myra Blanco, Joseph L. Bocanegra, Justin F. Morgan, Gregory M. Fitch, Alejandra Medina, Rebecca L. Olson, Richard J. Hanowski, Brian Daily, and Richard P. Zimmermann Part II Authors: Heidi D. Howarth, Tara E. Di Domenico, Lawrence C. Barr, Stephen M. Popkin, and Kevin Green.		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Center for Truck and Bus Safety Virginia Tech Transportation Institute 3500 Transportation Research Plaza (0536) Blacksburg, VA 24061		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Department of Transportation National Highway Traffic Administration 1200 New Jersey Avenue SE. Washington, DC 20590		10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES This project was managed by Dr. Paul Rau of the National Highway Traffic Safety Administration.		
12a. DISTRIBUTION/AVAILABILITY STATEMENT This document is available to the public through the National Technical Information Service, Springfield, VA 22161.		12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Drowsiness has a globally negative impact on performance, slowing reaction time, decreasing situational awareness, and impairing judgment. A field operational test of an early prototype Drowsy Driver Warning System was conducted as a result of 12 years of field and laboratory studies by the National Highway Traffic Administration and the Federal Motor Carrier Safety Administration. This project included Control and Test groups. The final data set for the analysis consisted of 102 drivers from 3 for-hire trucking fleets using 46 instrumented trucks. Fifty-seven drivers were line-haul and 45 were long-haul operators. The data set contained nearly 12.4 terabytes of truck instrumentation data, kinematic data, and video recordings for 2.4 million miles of driving and 48,000 driving-data hours recorded, resulting in the largest data set ever collected by the U.S. Department of Transportation. In this study, 53 research questions were addressed related to safety benefits, acceptance, and deployment. Novel data reduction procedures and data analyses were used. Results showed that drivers in the Test Group were less drowsy. Drivers with favoring opinions of the system tended to have an increase in safety benefits. Results of the assessment revealed that the early prototype device had an overall positive impact on driver safety.

14. SUBJECT TERMS Acceptance, Behavior, Deployment, Driving Performance, Drowsiness, Fatigue, Safety Benefits, Sleep, Human Machine Interface, Warning System			15. NUMBER OF PAGES 465 16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

NSN 7540-01-280-5500 Standard Form 298 (rev. 2-89)
Prescribed by ANSI Std. 239-18, 298-102

TABLE OF CONTENTS

EXECUTIVE SUMMARY

BACKGROUND

Drowsiness has a globally negative impact on performance, slowing reaction time, decreasing situational awareness, and impairing judgment (Balkin et al., 2000; Van Dongen, Maislin, Mullington, & Dinges, 2003). Research with local/short-haul truck drivers found that approximately 20 percent of safety critical events (crashes and near-crashes) included driver drowsiness as a contributing factor (Hanowski, Wierwille, & Dingus, 2003). Because of this large impact of drowsiness on transportation safety, the U.S. Department of Transportation considers drowsy driving an important research area. This interest is embodied in the National Highway Traffic Safety Administration's and Federal Motor Carrier Safety Administration's (FMCSA) partnership with the research community to develop technological aids for drowsy drivers (Grace et al., 1999).

Based on the results of previous research into methods of alerting drivers (Comsis Corporation, 1996) and monitoring their levels of drowsiness and fatigue (Wierwille et al., 1994; Grace & Benjamin, 1999), a promising candidate for a drowsy driver detection system was developed (Ayoob, Grace, & Steinfeld, 2005). This system uses a near-infrared camera coupled with processing equipment to estimate the driver's percentage of eye-closure (PERCLOS), which has proven to be a reliable measure of driver drowsiness (Dinges & Mallis, 1998; Wierwille, 1999).

Following initial prototyping and laboratory testing, NHTSA and FMCSA sought to examine the possible safety benefits of placing a drowsy driver warning system (DDWS) into service. In order to answer this question, a field operational test (FOT) was performed. This FOT was performed using a prototype driver fatigue monitor (DFM) which was, at the time the study began, the system at the highest level of development and most viable for testing purposes. Data for this project was obtained in an on-road (naturalistic) manner beginning in September 2002. Hanowski et al. (2008) presents the details of the data collection effort, which serves as the source for the current assessment project. This project included both Control and Test Groups utilizing a quasi-experimental design. The final data set for the analysis consisted of 102[1] drivers (101 male, 1 female[2]) from three for-hire trucking fleets using 46 instrumented trucks. Of the drivers, 57 were line-haul (out-and-back) and 45 were long-haul (drivers on the road for approximately one week) operators. This resulting data set contained approximately 12.4 terabytes of video, truck instrumentation, and kinematic data for the 2.4 million miles of driving and 48,000 driving-data hours recorded. Load histories and sleep hygiene (actigraphy) data was collected as well.

ASSESSMENT ANALYSIS OVERVIEW

A data set composed of naturalistic parametric data and subjective data was originally collected by VTTI for NHTSA and FMCSA in a previous research effort, which allowed for a comprehensive examination of drowsiness in commercial motor vehicle (CMV) drivers. At its

[1] One driver with glasses was tested to assess system operation with glasses in a naturalistic setting. As expected, the DDWS was not able to reliably detect eyes in this real-world evaluation. As such, this driver will be removed from further analyses, so the final count is 102.
[2] This ratio of male to female drivers closely represents the actual distribution of commercial motor vehicle (CMV) drivers on the road.

time, this was the largest data set ever collected by the U.S. Department of Transportation. VTTI was tasked with analyzing the parametric data to determine if driver performance differed based on the assistance of a DDWS in conflict situations and evaluate the prototype DFM system's performance and capabilities (Part I of this report; Research Questions 1-6 below). VNTSC led the acceptance and deployment aspects of the assessment (Part II of this report; Research Questions 7-18 below). The research questions covered under this research effort are mentioned next and the main findings presented thereafter.

SAFETY BENEFITS

The most important benefit of an in-vehicle system is to promote safety. In the case of the DDWS the safety benefits could be drawn from multiple aspects, including reduction of on-the-job drowsiness, improved sleep hygiene, and ultimately a reduced number of safety-critical events (SCEs) related to drowsy drivers.

Research Question 1 - On-the-Job Drowsiness

To evaluate the question of a DDWS's impact on the incidence of on-the-job drowsiness, percentage of eye-closure (PERCLOS) data were collected during driving episodes matching a set of pre-defined conditions. Using this data, the following research questions were addressed:

RQ 1.1: Does the DDWS result in fewer episodes of drowsy driving?
RQ 1.2: Does the frequency of DDWS alerts decrease over time?
RQ 1.3: Do DDWS alerts have an impact on post-alert behavior?

Research Question 2 - Sleep Hygiene

The implementation of a DDWS may improve the salience of sleep's importance in safe driving and could prompt drivers to make changes to their sleeping habits. In order to address this issue, the following research questions were examined:

RQ 2.1: Does a DDWS influence drivers to get more sleep?
RQ 2.2: Do drivers using a DDWS achieve better quality sleep?

Research Question 3 - Involvement in Safety Critical Events

Safety critical events are situations within the driving task where a crash, near-crash, or crash relevant conflict occurs. Implementation of the DDWS may have an effect on driver involvement in such events. The investigation of this possible relationship is outlined as follows:

RQ 3.1: Does a DDWS affect involvement in safety critical events?
RQ 3.2: Does a DDWS affect involvement in at-fault safety critical events?

Research Question 4 – Human-Machine Interaction

The human-machine (driver-system) interface of the DDWS is of great importance to both acceptance on behalf of the driver and a reduction in overall system performance variance due to specifics of the DDWS interface. Thus, examining the characteristics of the DDWS operation is necessary.

> *RQ 4.1*: How do drivers operate the DDWS in a real-world environment?

Research Question 5 - Favoring Drivers - Follow-Up

The safety benefits that may be obtained via the DDWS are partially dependent upon optimal system performance. Understanding situations where optimal performance was not obtained will allow the research teams to both improve subsequent versions of the DDWS and better understand the system's current and potential operational envelope.

> *RQ 5.1*: How did the DDWS operate for drivers who rated the system favorably in the post-study survey?

Research Question 6 - At-Risk Driver - Follow-Up

Due to the large variability known to exist within the population of commercial heavy-vehicle drivers (Knipling, 2005), there are potential impacts on driver risk (Lancaster & Ward, 2002). To better explore these factors and gain a better understanding of what led to these drivers receiving more alerts, the following research question was examined.

> *RQ 6.1*: How did the DDWS operate for drivers who had significantly more alerts during the baseline?

DRIVER ACCEPTANCE

Acceptance depends upon the degree to which a driver perceives the benefits derived from a system as greater than the costs. In our assessment of the DDWS, we sought to determine the extent to which drivers were able to use the device easily and understand its functioning, find that it operated as intended and endorse it, and heed the feedback it provided. Ultimately, if drivers do not accept a technology such as this, they may be inclined to ignore its warnings or even generate additional risk by using the device in an inappropriate manner. Conceptualized in this way, it is critical to understand user acceptance on a number of levels, and evaluation of these elements includes a systematic assessment based on human factors principles as applied from the perspective of the technology user.

The research questions for this chapter included:

Research Question 7 – Ease of Use

The research questions below address DDWS usability, use patterns, and the degree of understanding and tolerance reported by those in the Test Group who experienced device feedback.

RQ 7.1: Does use of the DDWS create extra demands on the driver, such as added stress or increased fatigue?
RQ 7.2: For various degrees of fatigue, how often and what duration do drivers require to observe the device in order to understand its output? Additionally, how do assessments of device accuracy change under varying degrees of fatigue? *(not addressed due to data constraints)*
RQ 7.3: To what degree are drivers willing to tolerate false alarms? Also, what is their degree of reliance on the system and their perception of correct alarms?
RQ 7.4: To what degree were drivers able to recognize DDWS alerts?
RQ 7.5: Do drivers understand the DDWS' operational limitations?
RQ 7.6: To what extent are the DDWS controls easy and intuitive for drivers to use?
RQ 7.7: What actions do drivers take to improve their alertness based on the warnings they received? Under what circumstances do they take such actions or not take action?

Research Question 8 – Ease of Learning

The research questions below address the degree to which participants believed that the training they received enhanced their understanding of fatigue management, as well as of the DDWS and its application.

RQ 8.1: Were drivers able to retain information about device operation and the meanings and uses of its output?
RQ 8.2: How much time does it take for drivers to feel proficient with the DDWS and its output? How much time does it take for drivers to learn both the capabilities and limitations of the device?
RQ 8.3: Was the training drivers received on the DDWS complete, and was it understandable? Was the fatigue management training drivers received complete, and was it understandable?

Research Question 9 – Perceived Value

The research questions below address the utility of the DDWS in terms of its perceived ability to measure alertness state and details participants' perception of safety, health-related, and data confidentiality concerns pertaining to the device.

RQ 9.1: How frequently did drivers indicate they received appropriate warnings based on an accurate alertness assessment?
RQ 9.2: What is the degree to which drivers felt that the DDWS enhanced the effectiveness of their fatigue management program and practices?
RQ 9.3: Do drivers view the DDWS as a liability or invasion of privacy?
RQ 9.4: Did drivers feel that the DDWS effectively decreased instances of fatigued driving, thus keeping driving skills at an appropriate level for safety?
RQ 9.5: Do drivers feel that use of the DDWS will have an adverse effect on their health?
RQ 9.6: How do drivers evaluate their driving safety based on use of the DDWS?

Research Question 10 – Advocacy

The research questions below address assesses the Test Group's degree of reported satisfaction with the DDWS in the context of its usefulness, participants' willingness to endorse it, and potential future device use.

> *RQ 10.1*: How satisfied were drivers with the DDWS? How useful did drivers find the DDWS to be?
> *RQ 10.2*: Are drivers interested in having the DDWS purchased for their entire fleet?
> *RQ 10.3*: Are drivers interested in purchasing the DDWS for their truck, or sharing the cost of device with their employer?
> *RQ 10.4*: Are drivers willing to endorse the DDWS to drivers within and outside their own company?

Research Question 11 – Driver Changes

The research questions below address pre- and post-study levels of perceived fatigue, DDWS usage outcomes, and the degree to which Test Group participants used device output to alter their behavior during driving and non-driving periods.

> *RQ 11.1*: What are drivers' perceived levels of fatigue/alertness while driving with and without the aid of the DDWS?
> *RQ 11.2*: Did drivers initiate behaviors as a result of exposure to the DDWS, including unexpected uses for the device? If so, what were they?
> *RQ 11.3*: How much time do drivers spend monitoring the DDWS under various degrees of fatigue? Where do drivers reallocate time spent monitoring the DDWS? *(not addressed due to data constraints)*
> *RQ 11.4*: Was use of the DDWS feedback, potentially to adjust driving style, associated with health improvements (e.g., altered work/rest cycles)?
> *RQ 11.5*: What behavioral changes may have been brought about as a result of extended exposure to the DDWS?

FLEET MANAGEMENT ACCEPTANCE

Considering management-level perspectives is also important to the process of successful deployment of a device such as the DDWS. Not only must company management accept the technology, but they are also responsible for understanding its potential impact on their employees and company operations, as well as anticipated safety and economic benefits.

The research questions for this chapter included:

Research Question 12 – Perceived Driver Acceptance

The research questions below address the degree to which trucking company interviewees expected their drivers to support the use of such an alertness-monitoring device, in addition to their views regarding its various potential advantages and disadvantages.

> RQ 12.1: What are managers' personal opinions regarding driver acceptance of the device?
> RQ 12.2: What are fleet managers' opinions regarding driver acceptance of the device based on feedback from drivers?
> RQ 12.3: Based on information provided at the briefing, what are fleet managers' perceptions of the DDWS' capabilities, advantages, and disadvantages? Does fleet management believe that drivers will approve or disapprove of the DDWS?

Research Question 13 – Safety and Economic Benefit

The research questions below address trucking company interviewees' perceptions of the potential economic benefits of alertness-monitoring technologies, as attributed to increased safety, and how much they would be willing to spend for such a device. Respondents also offered insight pertaining to what types of insurance and federal incentives would make them more likely to recommend the purchase of this technology for their fleets.

> RQ 13.1: What are fleet managers' perceptions of the potential economic benefits, as attributed to increased safety, of the DDWS?
> RQ 13.2: What federal incentives would fleet management like to have associated with the adoption of the DDWS?
> RQ 13.3: What insurance incentives would fleet management like to have associated with the adoption of the DDWS?
> RQ 13.4: How much would fleet management be willing to pay for a DDWS?

Research Question 14 – Impact on Operation

The research questions below address interviewees' perceptions regarding the potential for trucking companies to monitor and use the data collected by such a device, the degree to which company access to device data might influence driver utilization and acceptance, and potential policies concerning required driver behavior following device warnings.

> RQ 14.1: How much training do fleet managers believe is required for their drivers to make the best use of the DDWS? And are they willing to provide it? *(not addressed due to data constraints; fleet management interviewees did not directly experience device.)*
> RQ 14.2: Does fleet management plan to monitor DDWS alerts? If so, how do they plan to use this information? What might the influence of this be on acceptance by drivers?
> RQ 14.3: What sort of policy, if any, does fleet management plan to implement regarding required driver behavior following DDWS alerts?
> RQ 14.4: To what extent is fleet management willing to modify their fatigue management programs (FMPs) based on FOT findings? *(not addressed due to data constraints; fleet management interviewees did not directly experience device.)*

Research Question 15 – Improvements

The research questions below address interviewees' perceptions regarding additional features or other improvements they would recommend for the concept device as it was described to them. Additionally, they were asked to comment on any concerns they could anticipate regarding device performance.

> RQ 15.1: What are fleet managers' concerns regarding performance of the DDWS?
> RQ 15.2: What features does fleet management desire in the DDWS? How much will fleet management pay for these features?
> RQ 15.3: Does fleet management seek other improvements in the DDWS? *(not addressed due to data constraints; fleet management interviewees did not directly experience device.)*

DEPLOYMENT

This chapter addresses short- and long-term deployment issues for alertness-monitoring technologies from the perspective of companies currently involved with the development and deployment of such devices. Additionally, it provides the views of a trucking-related organization representative regarding potential advantages and disadvantages of such technologies, driver acceptance and device deployment, as well as anticipated concerns of management and organized labor.

The research questions for this chapter included:

Research Question 16 – Introduction and Price Range

The research questions below address interviewees' perceptions regarding the market for alertness-monitoring devices, including the general availability of their product, pricing strategies, and views on long-term deployment.

> RQ 16.1: Is there a distinct introductory period anticipated for the DDWS? If so, how long will it last? What will be the introductory price for a DDWS? What will be the maximum penetration level of a DDWS in trucks (heavy vehicles) during the introductory period? When will suppliers and truck manufacturers consider deployment of the system?
> RQ 16.2: What will be the price range of DDWS deployments over the next 15 years? What is the expected rate of DDWS deployment over the next 15 years? Will the DDWS be available as an option for all trucks, or only more expensive models?

Research Question 17 – Perspectives of Trucking-Related Organizations

The research question below addresses the interviewee's perceptions regarding advantages and disadvantages of this type of technology, organized labor's expected views, and liability concerns. Additionally, attitudes pertaining to expected utilization, acceptance, and deployment of this technology were assessed.

> *RQ 17.1*: What are organized labor's concerns regarding the DDWS?

Research Question 18 – Additional Activities

The research question below addresses perceptions regarding activities that interviewees believed would help promote the deployment of an alertness-monitoring technology, as well as research undertakings they would recommend in the area of drowsy driving as related to crash avoidance. Additionally, the extent to which trucking fleets have already purchased or considered the use of alertness-monitoring devices was explored.

> *RQ 18.1*: Would there be a need for non-research activities beyond the FOT in order to expedite DDWS deployment, assuming it is desirable to expedite deployment? What else, if anything, should be done in drowsy-driver-related crash avoidance research?

SAFETY BENEFITS

Driver Drowsiness

There were inconclusive results as to the impact of the prototype DFM system on driver performance in conflict situations. Although precautions during data reduction were taken in an attempt to eliminate periods when the DFM would likely operate in an unreliable fashion, a high number of false alerts (not related to drowsiness) still occurred. A sampling of alerts revealed a discrepancy between the prototype DFM system's estimate of PERCLOS and the DDWS concept requirement. For example, certain driver actions (e.g., scanning mirrors) were included in the DFM's PERCLOS calculation. These make it difficult to draw any directly generalizable conclusions regarding the findings for the DDWS concept, but some general tendencies were observed.

With those precautions in mind, the analysis revealed that drivers in the Test Group had lower PERCLOS values overall as compared to other experimental conditions. Most of the valid DFM alerts were registered on dry two-lane highways with no adverse weather conditions present. These results suggest that providing the driver with feedback as to his or her level of arousal would lead to an overall reduction of instances of drowsy driving. Lastly, the prototype DFM system did not appear to impact driver's normal sleep pattern or sleep hygiene (quantity or quality), as measured by actigraphy.

A number of safety critical events were recorded over the course of the study. As expected, the majority of these were crash relevant conflicts and near-crashes. Regardless of the type of safety critical event, no statistically significant performance differences were present between those drivers with DFM feedback and those without.

Driver Subjective Ratings and At-Risk Follow-Ups

Two follow-up analyses were conducted. The first compared drivers who had a favoring opinion of the prototype system to those with a disfavoring opinion of the system. Those drivers who had a disfavoring opinion of the system tended to have a lower rate of valid alerts. Drivers with favoring opinions of the system tended to have an increase in safety benefits. Therefore the system did not benefit all drivers equally. Although the validity of DFM feedback may have been the influencing factor driving this difference, the very small sample size involved with this analysis precludes any sweeping conclusions.

The second follow-up analysis examined differences between drivers identified to be at risk of drowsiness-related events and those identified as having a lower risk of that type of event. Mixed results were found. Although the at-risk drivers had a reduction in drowsiness events, sleep quality data seemed to indicate that drivers with higher sleep quality (considering only sleep periods over 20 minutes) had a higher risk. As in the other follow-up analyses, this may be an artifact of the extremely low sample size identified for these analyses. However, the tendency presented by the data is one of interest. If drivers who sleep more attempt driving longer hours, the benefit of improved sleep hygiene may not reduce drowsiness-related events. Therefore it is important that a synergistic plan to coach drivers on the importance of sleep and safe driving conditions is implemented. If not, the benefit of longer sleep periods with higher sleep quality may not positively impact drivers who would like to take advantage of a DDWS. Moreover, if areas to safely park and rest are not available to drivers during their delivery runs, aspects such as improved sleep hygiene, accurate DDWS alerts, and comprehensive drowsiness management trainings might not yield the expected results.

Driver Interaction with the DFM

There were significant differences in the ways drivers interacted with the DFM. As would be expected with many in-vehicle devices, once initial driver preferences were determined the DFM tended to remain at those settings. That is, initial preferred warning sound and sensitivity level were frequently maintained. Some interesting patterns of use were apparent. One example is that drivers tended to use display brightness as a binary adjustment, using either the minimum or maximum of the full range of adjustments available to them.

Two user types became apparent. These two types were manifested as an interaction between the frequency of adjustments in DFM settings and the presence or absence of safety critical events within the driver’s shift. Those drivers with safety critical events in their shift tended to have an overall lower level of interaction with the sound, brightness, and sensitivity settings of the DFM as compared to drivers without a safety critical event in their shift. Although the frequency of interactions with the system declined for drivers with and without a safety critical event in their shift across the duration of the study, a statistical difference between the two groups remains. This could be viewed as a difference in driver engagement with the DFM, and suggests that a factor such as driver engagement with the DDWS may also influence driver performance.

PERFORMANCE AND CAPABILITIES

In support of the second major objective of this research effort, VTTI characterized the performance and capabilities of the DDWS. This was done in terms of DDWS operation under naturalistic conditions, system integrity in real-world operations, and the specific human-machine interaction.

The prototype DFM system used in the present study was the result of a long series of experiments by the NHTSA Drowsy Driver Technology Program since 1996. Subsequent versions of this system are currently available to commercial vehicle drivers and operators. The system is based on PERCLOS (Wierwille, 1999), a relatively commonly used metric for determining driver drowsiness. Information about driver slow eye-closures, used to generate the prototype DFM system's estimate of PERCLOS, was captured via near-infrared cameras mounted on the vehicle dashboard.

Although this system allowed for a relatively clear view of the driver's eyes, it was not an optimal placement for the device. This may have contributed to the relatively large number of alerts not related to drowsiness generated by the prototype DFM system. It is recommended that refinements be made to the device algorithms used to determine what information is included and excluded from driver slow eye-closures, along with a more optimal placement of the device (which would require vehicle manufacturer support). These should yield great benefits in the operation of a future DDWS.

DRIVER ACCEPTANCE

Driver acceptance of the DDWS was assessed with regard to participants' ability to use and understand the system and its feedback easily and the degree to which participants found that the system operated as intended, heeded its alerts, and were willing to endorse its use to others. As drawn from the full sample of 102, driver acceptance analyses were conducted using a subset of survey data from 48 participants (47 males, 1 female; 33 Test Group, 15 Control Group). Data collection entailed four separate paper-and-pencil surveys administered at various points during the FOT. A pre-participation survey was used in conjunction with additional screening procedures in order to select drivers for FOT participation. A pre-study survey was provided at the start of FOT participation, and post-study and debriefing surveys were completed following participation. Additionally, 14 Test Group participants contributed to post-FOT focus group sessions.

Driver acceptance of the DDWS was largely conditional. Generally, participants found the device easy and intuitive to operate and understand and were satisfied with the training they received to this end. Reports indicated that participants were most likely to utilize device feedback when they felt very tired, although the majority did not report feeling more alert during night driving when using the device. In response to drowsiness alerts, participants often talked on the phone or rolled down a window as opposed to stopping driving, which was the desired outcome.

Despite explanations provided to participants of the technology's operational shortcomings during daylight, dawn, and dusk, many found it difficult to disregard instances of perceived false alerts when evaluating the device. Attitudes prior to experiencing the DDWS were more optimistic regarding its ability to increase driving safety, whereas after device exposure, participant opinions were generally neutral regarding the extent to which their driving safety had improved. While reluctant to integrate DDWS output into a daily fatigue management routine, participants did indicate only infrequently ignoring the device, and most did not stop relying on it despite false alarms.

Survey data and focus group session outcomes revealed that overall satisfaction levels were less positive regarding device performance and more positive about device potential if its kinks were worked out. As such, survey findings showed that participants were hesitant to endorse this device for use by others or for their own continued use. Nevertheless, there existed a small subset of drivers who found the device to be particularly useful and who were more satisfied with its performance than the typical FOT participant. Uniformly, this extreme-favoring sample subset comprised long-haul drivers, while an opposing group of extreme-disfavoring participants comprised one long-haul and two line-haul drivers. Although beyond the scope of this effort, further investigation into the possibility that certain trucking operations or driver subgroups in the CMV-driving population are well-suited to successful use of a device such as the DDWS is warranted.

FLEET MANAGEMENT ACCEPTANCE

In an additional, separate data collection effort, we conducted telephone interviews with trucking company management in order to understand aspects of fleet management acceptance of a conceptual DDWS. Overall, interviewees expressed support for the potential of an alertness-monitoring device as a tool to improve safety, and consequently save money by reducing accidents and injuries. However, there was shared concern that employees would consider it a means for management to monitor their actions and that it would be utilized as basis for punishment or dismissal. Nevertheless it was anticipated that driver acceptance could be bolstered through clear discussion of management's intended use for the data, including informing drivers of potential safety hazards, as long as there were not punitive outcomes.

Despite several respondents indicating that they did not believe that use of such a device was necessary at their company, reduced insurance rates and government tax credits were cited as incentives that might make them more likely to purchase the technology, especially in light of rapidly rising fuel costs. The most frequently cited economic benefit of alertness-monitoring technologies was also cost-related, as interviewees anticipated that a safety benefit for the device would translate into reduced accident, injury, legal, and property damage expenditures. Concerns with regard to reliability, accuracy (in particular, false alerts), and maintainability of the device were noted as potentially critical system performance issues.

DEPLOYMENT

In an additional, separate data collection effort, we conducted telephone interviews, and in one case solicited information via email, from technology developers and a trucking advocacy organization in order to understand aspects of deployment of alertness-monitoring technologies. Both technology company interviewees reported that, as part of development, their respective systems had undergone extensive research, design, laboratory, and on-road testing. At interview time, one resulting technology had already been deployed, and the other was engaged in further development in conjunction with original equipment manufacturers. Although not having been directly exposed to alertness-monitoring technologies in the field, the trucking advocacy organization representative who was interviewed expressed support for the premise of such devices, despite believing that motor carriers would be more likely to focus available resources on other engineered systems geared toward reducing high crash-likelihood scenarios. As such, trucking company management staff provided additional ideas for promotional and deployment strategies for alertness-monitoring technologies, including documenting driver acceptance and device safety benefits through trial field testing at companies, educational seminars, and various economic incentives. In the case of organizations or among owner-operators for whom driver fatigue is not perceived as an issue, however, it was offered that no amount of promotion or examples of successful industry deployment would be likely to lead to immediate purchases of such technology.

CONCLUDING REMARKS

The assessment revealed interesting findings and challenges in data reduction and analysis. For example, the results demonstrated that the prototype device was attempting to provide alerts outside of the optimal operating envelope, providing for a large number of alarms not related to drowsiness, and for scenarios in which the device was not meant to operate, including driver distraction. Despite the prototype pitfalls, the DDWS concept seemed to flourish. The results of the assessment suggest the prototype device produced some indications of positive changes in driver safety. Therefore, further refinements in the algorithms, operational envelope, and driver-system interaction may promote stronger acceptance and impact in aspects such as on-the-job drowsiness. Drivers who favored the system seemed to receive a benefit from it; therefore, increasing the aspects that were favored and improving the aspects that were not will potentially help the drivers who need it the most (i.e., drivers at-risk of a higher number of drowsiness episodes). The DDWS concept appears to have merit and value to truck drivers, especially long-haul drivers, and trucking companies. Results of the driver acceptance portion of this study strongly point to the need for significant user involvement in system interface development, including “false alarm” suppression and proper integration into a viable fatigue management program. Providing alerts without providing drivers an appropriate level of prescriptive information regarding countermeasures, or situations where drivers find themselves unable to employ a referred countermeasure (e.g., no safe area to pull off the road), will likely lead to frustration and perhaps non-use, in particular if it is believed that the intent of the device is to allow blame-shifting to the drivers in the case of drowsiness-related safety critical events. Additionally, if the driver subpopulation of owner-operators is to benefit, the system will need to be packaged with a basic fatigue management protocol for this driver subpopulation. The fact that there are several manufacturers developing this technology is good news; if done well, there appears to be a waiting market to embrace this safety net device.

Part I: Safety Benefits and Performance/Capabilities

Myra Blanco
Joseph L. Bocanegra
Justin F. Morgan
Gregory M. Fitch
Alejandra Medina
Rebecca L. Olson
Richard J. Hanowski
Brian Daily
Richard P. Zimmermann

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AMI	Ambulatory Monitoring, Inc.
ANOVA	analysis of variance
ATI	Attention Technologies, Inc.
ATRI	American Transportation Research Institute
BL	break length
CB	citizen's band
CDL	commercial driver's license
CMV	commercial motor vehicle
CR	correct rejection
CRC	crash relevant conflict
DART	Data Analysis and Reduction Tool
DAS	Data Acquisition System
DDWS	Drowsy Driver Warning System
DFM	Driver Fatigue Monitor
DOT	U.S. Department of Transportation
DVI	driver-vehicle interface
EEG	electroencephalography
EMG	electromyography
EOG	electrooculography
FA	false alarm
FHWA JPO	Federal Highway Administration Joint Program Office
FMCSA	Federal Motor Carrier Safety Administration
FOT	field operational test
GPS	global positioning system
GSR	galvanic skin response
HRV	heart rate variability
LCL	lower confidence level
LTL	less than a truckload
MLM	multilevel modeling
NC	near-crash
NHTSA	National Highway Traffic Safety Administration
NSTSCE	National Surface Transportation Safety Center for Excellence
O-O Interval	sleep onset-offset interval
OSAS	obstructive sleep apnea syndrome
PERCLOS	percent of eye-closure
QA	quality assurance
REM	rapid eye movement
RQ	research question
SCE	safety critical event
SSP	scored sleep period
VNTSC	Volpe National Transportation Systems Center
UCL	upper confidence level
VTTI	Virginia Tech Transportation Institute

ACKNOWLEDGEMENTS

This research was funded by the National Highway Traffic Safety Administration under DTNH22-05-D-01019 Task Order #18. The authors wish to express their gratitude to Dr. Paul Rau, task order manager, for his technical direction, insight, advice, and support throughout the entire project.

The data used for this assessment study was funded under Contract # DTNH22-00-C-07007 Task Order #14. Support for this endeavor was also provided by the Federal Motor Carrier Safety Administration (FMCSA). Tim Johnson (now with NHTSA) oversaw the project for FMCSA.

The Federal Highway Administration (FHWA) Intelligent Transportation System (ITS) Joint Program Office (JPO) provided funding for this study. Virginia Tech also provided funding to support this effort. Dr. Tom Dingus, director of the Virginia Tech Transportation Institute (VTTI), facilitated this process.

The data collection would not have been possible without the support of three safety-conscious trucking companies. J.B. Hunt (Greer Woodruff), Pitt Ohio Express (Jeff Mercadante), and Howell's Motor Freight (Tim Shephard) allowed the research team access to vehicles and drivers. The authors wish to thank these companies, their employees, and support staff at the fleet locations.

Finally, this study could not have been conducted with the participation of the drivers. These three companies and their drivers are dedicated to the trucking profession and have safety as a core value.

CHAPTER 1. INTRODUCTION

PROJECT OVERVIEW

Drowsiness has a globally negative impact on performance, slowing reaction time, decreasing situation awareness, and impairing judgment (Balkin et al., 2000; Van Dongen, Maislin, Mullington, & Dinges, 2003). Research with local/short-haul truck drivers found that approximately 20 percent of safety critical events (SCE, including crashes and near-crashes) included driver drowsiness as a contributing factor (Hanowski, Wierwille, & Dingus, 2003). Because of this large impact of drowsiness on transportation safety, the U.S. Department of Transportation considers drowsy driving an important research area. This interest is embodied in the National Highway Traffic Safety Administration's partnership with the research community to develop technological aids for drowsy drivers (Grace et al., 1999). An important output of the NHTSA research program was a Drowsy Driver Warning System (Grace & Stewart, 2001; Wierwille et al., 2003).

Following the development of a prototype, NHTSA sought to ascertain the safety benefits of integrating a DDWS into in-service truck fleets. To answer this question, a field operational test was initiated in September 2002. This FOT, co-sponsored by NHTSA, the Federal Motor Carrier Safety Administration, and the Federal Highway Administration Joint Program Office, served as an assessment of the driver drowsiness warning system concept. This FOT was performed using a prototype device (the driver fatigue monitor) which was, at the time the study began, the system at the highest level of development and most viable for testing purposes. Hanowski et al. (2008) presents the details of the data collection effort, which serves as the source for the current assessment project.

FOT Experimental Design

The experimental design implemented in the study included Test and Control Groups. The Test Group followed an AB design, where A refers to the condition in which the DFM did not provide feedback to the driver (i.e., a baseline condition) and B refers to the condition where the DFM did provide feedback (i.e., an active mode where the system was fully functional). For the Control Group the DFM collected data over the duration of the participant's involvement but did not provide feedback to the driver.

It must be noted that, during the course of the study, the number of weeks for the baseline and active conditions were adjusted to compensate for data missed due to truck or other system malfunctions (these are anticipated occurrences with naturalistic data collection). Thus, the planned experimental model was modified after data collection began to accommodate the actual naturalistic data collection constraints. As such, the experience for drivers in the active condition was not the same for all drivers, and some, for example, had more baseline weeks than other drivers. The design provided guidance on the baseline/Active split, but issues associated with naturalistic data collection (including inconsistencies in the length of drivers' trips) made it impossible to have equal exposure for all drivers within the original allotted timeframe. At the completion of the study all drivers in the Test Group, with the exception of drivers who dropped out of the study before completion, had a baseline period of approximately two weeks (at a minimum) followed by the active condition of approximately nine weeks or more.

The experimental plan called for collection of data from 102 commercial vehicle drivers. This design was sensitive to the typical turnover rate of drivers in the trucking industry, which is estimated at 136 percent for large truckload carriers (Paz-Frankel, 2006). The experimental plan also called for the drivers to (1) wear no eyeglasses, (2) drive at night, and (3) include long-haul (i.e., drivers on the road for approximately one week) and line-haul (i.e., out-and-back) operations. The criteria in items 1 and 2 were based on results of a study that characterized and defined the operational constraints of the device (Wierwille et al., 2003). The experimental plan also called for the instrumentation (with the DFM and sophisticated data collection equipment) of 34 commercial tractor vehicles.

FOT Data Set

The final study resulted in a data set with the following characteristics:

- 102 drivers (101 male, 1 female; mean age = 40 yrs)
- Three for-hire trucking fleets and 46 instrumented trucks:
    - J.B. Hunt (19 trucks, 45 drivers)
    - Howell's Motor Freight (19 trucks, 41 drivers)
    - Pitt Ohio Express (8 trucks, 16 drivers)
- Test Group = 78 drivers; Control Group = 24 drivers
- Line-haul operations = 57 drivers; long-haul operations = 45 drivers
- Final data set includes:
    - ~12.4 terabytes (including video at ~30 Hz and truck instrumentation, kinematic, data at 10 Hz)
    - ~2.4 million miles of driving
    - ~48,000 driving-data hours
    - Load histories (i.e., destinations and type of goods distributed, as available)
    - ~190,000 hours of actigraphy data (measures of sleep quantity and sleep quality)
    - Driver and fleet manager surveys (pre-participation, pre-study, post-study, debriefing, fleet management)
    - Driver focus group results (post-study)

For this assessment the data for 96 drivers was used (73 Test and 23 Control). The participants eliminated from the assessment were not exposed to their experimental condition in a similar manner as the other drivers in their condition due to technical problems. Data from these drivers can be used for other types of future analyses, but the DFM portion of time in baseline does not comply with the data collection design.

OBJECTIVES

The objective of this work was to accomplish each of the high-level objectives outlined in the Volpe National Transportation Systems Center 2002 analysis plan (see Wilson, Popkin, Barr, & Hitz, 2002). Specifically, in this work the Virginia Tech Transportation Institute accomplished the objectives related to the safety benefits of the system and its associated performance and capabilities. Issues of driver acceptance, fleet management acceptance, and deployment are addressed by VNTSC in Part II of this document. Thus, the following objectives were determined for this research project:

- Assess safety benefits by evaluating:
    - Driver performance with and without DDWS assistance in non-conflict situations;
    - Driver performance with and without DDWS assistance in all conflict situations;
    - Near-collision experience with and without DDWS assistance; and
    - Driver performance versus drowsiness level as measured by the DDWS.
- Characterize the performance and capabilities of the DDWS in terms of:
    - Sensor performance;
    - Alert logic;
    - Driver-vehicle interface (DVI); and
    - Calibration, maintenance, adjustment, and reliability.

The information gained from this assessment not only provided insight into the safety benefits and driver performance of the tested DFM device, but it also provided valuable information to help guide the development of any such future system.

Safety Benefits Model

This section provides further definition of the safety benefits objective. In developing the research questions and analyses needed to assess the safety benefits associated with the DDWS, it was useful to develop a model showing the relationship between the DDWS and the anticipated benefits. In developing this model, the following basic assumptions were held:

1. The purpose of a PERCLOS-based DDWS is to provide the driver with timely feedback regarding an unsafe drowsy state (Wierwille, 1999).

2. Without drowsiness alerting information, it would be expected that drivers would have more frequent episodes of drowsy driving. However, if used appropriately, the use of the system should lead to fewer episodes of on-the-job driver drowsiness.

3. Sleep is the only true remedy for drowsiness.

4. Alert information providing feedback to the driver about his or her drowsiness state, coupled with a fatigue management plan that informs the driver about the importance of sufficient sleep, indicates to the driver that driving safety is being compromised.

5. Drivers will be positively influenced by their experience with the DDWS.

6. Research indicates that drowsiness is a contributing factor (not necessarily a causal factor) in 20 percent of SCEs (Hanowski, Wierwille, & Dingus, 2003). For some unspecified portion of SCEs, it is hypothesized that high alertness may have prevented the incident from occurring. Therefore, alert drivers would be expected to be involved in fewer critical incidents as compared to drowsy drivers.

Based on these assumptions, the DDWS-Safety Benefits model shown in Figure 1 was developed. The model incorporates the six assumptions made above and indicates that a valid and reliable DDWS would be expected to (1) reduce on-the-job drowsiness, (2) increase the amount of sleep drivers get, (3) reduce involvement in drowsiness-related SCEs, and (4) increase the experience drivers have in interacting with the device (i.e., the DVI). Therefore, these would be the four anticipated primary safety benefits associated with a valid and reliable DDWS.

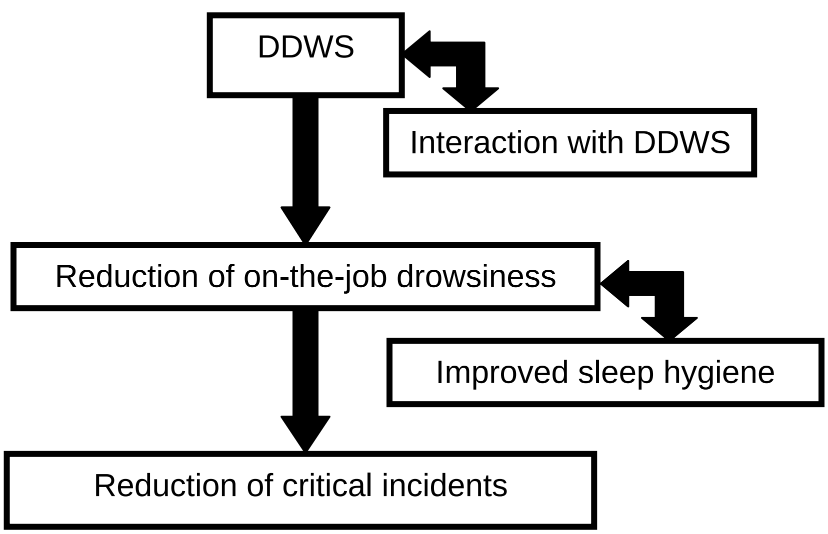


Figure 1. Modeling Safety Benefits Associated with a DDWS

Though not included in the model, it must be noted that all drivers in the study participated in a pre-study fatigue management course. For drivers in the Test Group, this discussion also included how to use the DFM, including the purpose, components, and use of the system. The fatigue management training lasted approximately two hours and was presented in small groups (i.e., two to six drivers) or in individual one-on-one sessions. The fatigue training was based primarily on a PowerPoint presentation entitled "Understanding Fatigue and Alert Driving" which, though modified, was previously developed by the American Transportation Research Institute (ATRI) with funding from FMCSA (see Hanowski et al., 2008). For drivers in the Test Group, DFM instruction immediately followed the fatigue management session.

DOCUMENT OVERVIEW

This document is divided into two main parts. Part I: Safety Benefits & Performance/Capabilities covers the safety benefits as well as performance and capabilities of a DDWS. It comprises five chapters:

- Chapter 1 introduces the concept of drowsiness as well as the DDWS FOT and states the project objectives.
- Chapter 2 provides an overview of the DDWS FOT data collection and concepts related to drowsiness and drowsiness detection.
- Chapter 3 presents the research questions answered in Part I of this report and operational definitions of the measurements used for the analyses performed to answer these questions.
- Chapter 4 shows the methods, results, and discussion for each research question, as well as an overall discussion of the main safety benefit issues of interest.
- Chapter 5 presents a discussion of the safety benefits analyses and aspects related to the performance and capabilities of the system.
- Overall conclusions are presented after Part II (i.e., Driver Assessment, Fleet Management Assessment, and Deployment).

CHAPTER 2. AN OVERVIEW OF NATURALISTIC DATA COLLECTION AND THE DROWSY DRIVER WARNING SYSTEM FIELD OPERATIONAL TEST

BACKGROUND

Although the network of commercial vehicle drivers in large trucks operates at a relatively high overall level of safety (Wang, Knipling, & Blincoe, 1999), the crash involvement of large trucks presents issues that passenger cars and light trucks do not. Large vehicles are represented in approximately 385,000 crashes a year (National Highway Traffic Safety Administration, 2008). These crashes are not only a burden to society in terms of economic losses suffered, but are also devastating in terms of the lives lost (large truck crashes represent over 8.2 percent of all traffic fatalities) and injuries suffered (approximately 80,000 in 2006, NHTSA, 2008).

Drowsiness is a known contributor to commercial vehicle crashes (Hanowski, Wierwille, & Dingus, 2003) and results in an overall reduction in levels of cognitive and physiological arousal (Balkin et al., 2000). It is estimated that driver drowsiness is a primary contributing factor in approximately 750 deaths and over 20,000 injuries involving commercial motor vehicles (CMV; Advocates for Highway and Auto Safety, 2001). Therefore, any increase in our understanding of such incidents is likely to greatly benefit public safety.

One problem posed by attempts to study the crash involvement of large vehicles is the availability of high-quality data. Despite the increased use of systems such as electronic logbooks and automated traffic monitoring, the majority of both near-crashes and crashes occurring on our national highway system are not recorded. Instead, what little information we are able to gather usually comes retrospectively, from police reports and state/federal crash databases. This information is not only collected after the event of interest (sometimes days afterward), but is also subject to the same reporting and recording biases that all human beings experience. Therefore, the drowsy driving problem may be largely underreported.

Specificity in the data recorded may cause additional problems, as the codes from police accident reports vary by state. For information about large-scale tendencies in crash involvement (such as the number of crashes and involvement by vehicle type), this rarely presents difficulty. However, little and limited information about the details of each crash are captured by this method of recording. Additionally, this method of analysis provides no information on the near-crash. Near-crashes, or situations in which the driver, roadway, or environment are able to ameliorate an imminent crash, may be even more informative to driver safety than crashes.

However, all in-depth attempts to study crash causation and the antecedents to such events require a source of reliable data. A source of reliable, quantitative data is absolutely necessary in order to better understand these factors contributing to crash involvement. This source of data must include epochs of driving occurring before any incident, as well as data about the event as it occurs. This method of data collection was largely infeasible due to technological barriers. With improvements in the power requirements, storage capabilities, speed, sensor resolution, and size of the equipment needed for such methods of data collection, the process of collecting continuous data from drivers on the road is now possible.

NATURALISTIC DATA COLLECTION

One powerful approach used by researchers is naturalistic data collection. As opposed to traditional epidemiological and experimental/empirical approaches, this *in situ* process uses drivers operating vehicles that have been equipped with specialized sensor, processing, and recording equipment, effectively rendering the vehicle a data collection device. The drivers operate and interact with these vehicles as they would during their normal driving routines; the data collection equipment is continuously recording numerous items of interest while the truck is driven.

This approach to data collection gives the researcher a powerful tool for examining a variety of questions. Because of the variety of data collected during each drive, researchers have the ability to examine the events occurring during, immediately prior to, and preceding any event of interest. This process also allows for the re-examination of collected data to incorporate more recent findings and theories into an analysis.

Naturalistic data collection methods require a sophisticated network of sensor, processing, and recording systems. This system provides a diverse collection of both on-road driving and driver (participant, non-driving) data, including measures such as driver input and performance (e.g., lane position, headway, etc.), four camera video views, and driver activity data. This information may be supplemented by subjective data, such as questionnaire data.

The centerpiece of this method is the Data Acquisition System. The DAS used in this study was designed and developed over the last decade by VTTI personnel. The system consisted of a Pentium-based computer that received and stored data from a network of sensors distributed to collect data of interest from the vehicle. Data were stored on the system's external hard drive, which could store several weeks of driving data before it needed replacement. This system has been successfully used in several on-road studies.

The DAS's sensors include a box containing the computer equipment necessary for obtaining data from the vehicle network, an accelerometer for determining longitudinal and lateral acceleration, a system that provides information on distance to lead vehicles (VORAD), an incident box that allows participants to flag incidents for the research team, a video-based lane-tracking system that measures lane-keeping behavior, and video recordings to validate any sensor-based findings. The video subsystem provides a continuous digital video record of events and situations occurring in and around the truck and trailer. This system allows for the synchronization, simultaneous display, and efficient archiving and retrieval of data. Four camera views are recorded, monitoring the driver's face, forward road view, and left and right side of the tractor trailer to observe the traffic actions of other vehicles around the truck. Additional system capabilities include system initialization equipment to automatically control system status and a GPS to collect information on vehicle position. Each of the sensor subsystems within the instrumented vehicle is independent with respect to the others, resulting in containment of sensor failures to the single sensor itself.

Due to the tremendous amount of information collected during these procedures, naturalistic data collection methods require significant post-collection processing. This is provided at VTTI's facilities, where secure machines on an isolated network allow for both processing and analysis of collected data.

AN INTRODUCTION TO DROWSINESS DETECTION

Driving is a continuous control task involving a great deal of visual information. It is commonly accepted that vision is the most important sense in the driving task (Sivak, 1996). Because of the great importance of vision in driving, any situation where visual attention is either removed or reduced during the driving task is likely to be associated with safety decrements. One of the more common factors in decreased driver visual information and control is drowsiness (Smiley, 2002). This is especially true in the domain of commercial truck drivers, where both line-haul and long-haul drivers are often subject to situations where drowsy driving is either likely or frequently occurs. Within local/short-haul truck driving alone, nearly 20 percent of crashes involved driver drowsiness as either a direct or contributing factor (Hanowski, Wierwille, & Dingus, 2003). Because of the magnitude of the problem, and the cost to society in terms of lives and capital lost, the National Highway Traffic Safety Administration and the Federal Motor Carrier Safety Administration have strived for the creation of a technological solution to drowsy driving through alerting drivers to their own levels of drowsiness (Grace et al., 1999).

Fatigue Versus Drowsiness

The terms *drowsiness* and *fatigue* are often used interchangeably (Dinges, 1995). For the purposes of the current evaluation, the researchers considered fatigue and drowsiness as two separate, but related, concepts. The exact definitions and connotations of each bring different contexts to the question of measurement and possible countermeasures to prevent impacts on driver performance and safety. Fatigue is defined as a global reduction in physical or mental arousal that results in a performance deficit (Williamson, Feyer, & Friswell, 1996) and results in a diminished capacity to perform a task (Brown, 1994). This is in contrast to drowsiness, which is defined as the physiological drive to sleep (Stutts, Wilkins, & Vaughn, 1999). Drowsiness is, in the common vernacular, often referred to as sleepiness.

Fatigue may be further divided into factors of physical aspects, perceptual aspects, or boredom and apathy (Desmond, Matthews, & Hancock, 1997). Physical fatigue consists of symptoms resulting from the strain and stress endured by the body and can include muscle stiffness, headache, and gastrointestinal discomfort. Perceptual fatigue focuses on symptoms involving the sense organs such as eye strain and ringing in the ears. Boredom and apathy are symptoms resulting from a loss of interest in the task or prolonged vigilance. Any one of these facets or a combination may be present when we consider an individual as fatigued.

Drowsiness, which may also be referred to as sleepiness, is a naturally occurring biophysiological process. Although it is possible for a person to be fatigued without being drowsy, drowsiness may be a product of a variety of factors, including fatigue. Therefore, the difference between the concepts of fatigue and drowsiness is important. One of the more dangerous outcomes of driver drowsiness is rapid onset microsleeps (Kloss, Szuba, & Dinges, 2003). These are periods lasting up to minutes in which the person loses consciousness and directly enters a sleeping period. People often have no memory of these events occurring and, of course, have limited control over the vehicle while in a microsleep. This makes the detection of drowsy drivers of paramount importance.

An Overview of Drowsiness Detection Methods

Two divergent approaches to the measurement of drowsy driving exist. These two methods differ largely on where the monitoring for drowsiness occurs. The first and most proximal to the driver is the measurement of physiological markers associated with drowsiness. The second is the measurement of aspects of driver (primary task) performance known to correlate with drowsy driving. Although both methods of monitoring provide valuable information, both include challenges to implementation (whether technological or otherwise), which provide challenges to the design and implementation of any DDWS.

Physiological Measures

The human body operates on a sleep-wake cycle of approximately 24 hours. This circadian rhythm is regulated largely by zeitgebers (exogenous cues for the body's time regulation) such as time of day and internal mechanisms such as body temperature, which lead to regular changes in level of arousal throughout the day (Moore-Ede, Sulzman, & Fuller, 1982; Webb, 1982). In a driving simulation experiment where participants drove in six sessions across a 24-hour period, Lenné, Triggs, and Redman (1997) found significant reductions in driver performance at the (clock) times of 6 a.m., 12 p.m., and 2 a.m. This suggests the correspondence of circadian rhythms with driver performance throughout the day. Associated with these cyclical variations in arousal are a variety of physiological markers that research has identified as possible predictors of drowsiness.

Ocular Measures

Blinks (Fast-Close)

Many of the characteristics of blinks and associated fast-close eye movements are of interest for the detection of drowsy drivers. Factors such as the rate and duration of blinks have a demonstrable relationship to drowsiness (Hyoki, Shigeta, Tsuno, Kawamuro, & Kinoshita, 1998). This is partially because of the basal relationship between the eyes and the central and peripheral nervous systems (Sirevaag & Stern, 2000).

Blinks are attractive as a potential metric for drowsiness for other reasons as well. Stern, Boyer, and Schroeder (1994) noted a positive relationship between blink rate and time on task. Initially, this suggests that blinks appear to be one of the more efficient and direct ways for measuring drowsiness. However, there is empirical evidence that large variations in an individual's blink rates are present. Ingre, Åkerstedt, Peters, Anund, & Kecklund (2006) examined drowsy driving and found a strong relationship between lateral control of a vehicle and blink patterns. Although this appears promising, much of the variance in their findings could also be explained by individual differences in driving.

Findings from the study of blinks have been used to develop various drowsy driver detection systems (RSSB, 2002; Johns, Tucker, & Chapman, 2005). Like most other indirect measurement systems, these devices depend on the use of a remote (or otherwise distal from the driver) sensor, typically an infrared camera, which monitors the presence and position of the pupil. Although they are typically non-intrusive to the driver (not requiring the driver to remained tethered to a control box or wear glasses with integrated cameras), they tend to suffer from a general inability to compensate for sudden, rapid fluctuations in illumination, which is typical in the driving environment. In addition, some systems are unable to adjust for situations where the driver's attention is focused somewhere besides the forward visual field for any length of time. These are issues which must be addressed to increase the practical value of systems based on eye blink behavior.

Slow Close Measures

Although the relationship between drowsy driving and roadway crashes was well understood prior to this work, some of the earliest research in finding a physiological-based method of assessing driver alertness came from research performed by C.W. Erwin (see Wierwille, 1999 for a review). Two of Erwin's initial approaches for predicting driver drowsiness were based on some of the oldest physiological measures available: heart rate variability (Volow & Erwin, 1973) and electrical measures such as electroencephalography and galvanic skin response (EEG and GSR, respectively; Erwin, 1976). Additionally, Erwin (1976, also see Erwin, Weiner, Hartwell, Truscott, and Linnoila, 1975) noted that perhaps the best overall predictor of drowsiness was what he termed *slow-ramp closures*, or prolonged closures (lasting over 1.0 *s*), observed during the onset of drowsiness.

However, Erwin's methods failed to operationally define slow-ramp closures in a manner suitable for use in testing field systems. Additionally, technological challenges such as the lack of automation in the processing of eye-closure information made implementation of this important research difficult at best. Due to advances in technology and with the understanding that continued research into the problem of drowsy driving may lead to significant safety benefits for the public, researchers began to revisit the questions of detecting drowsy driving through measuring a combination of variables including driver physiology and performance. Specifically, the physiological measurement included several different methods of quantifying driver slow eye-closure (Erwin's slow-ramp closure): PERCLOSE, EYEMEAS, and EYEMEAN (Skipper & Wierwille, 1986a). These measures all use slow eye-closures over a defined time interval but differ in their calculation and derivative values. Importantly, all of these measures specifically exclude rapid eye-closures such as blinks, leaving only the slow eye-closures for inclusion in the analysis.

PERCLOS is defined as the percentage (alternatively defined as the proportion) of time that the eyes are between 80 and 100 percent closed during a defined time interval. The onset value of 80 percent eye-closure was initially chosen under the assumption that this value would include significant coverage of the pupil and thus hamper the gathering of visual information (Wierwille, 1999). EYEMEAS is the sample mean square of instantaneous percentage of eye-closure over a defined time interval. EYEMEAN (also referred to as AVECLOSE and EYELID) is the sample mean of percentage closure over a defined time interval. These measures were demonstrated to reliably co-vary with performance-based indicators of drowsy driving (Skipper & Wierwille, 1986a, 1986b). Later work (Dingus, Hardee, & Wierwille, 1987) was able to demonstrate that PERCLOS has a significantly higher correlation to measures associated with drowsy driving, including lateral vehicle control measures of lane holding, yaw variance, steering velocity, and steering reversals. This led to the further exploration of slow-close eye measurements as a metric for detecting drowsiness in the driver.

PERCLOS has been experimentally examined in a variety of settings. Dinges and Grace (1998) compared a number of existing methods and algorithms for the detection of drowsy drivers against a psychomotor vigilance task. Their findings indicated that even though all technologies examined held promise for drowsy driver detection, PERCLOS had a higher correlation to the psychomotor vigilance task and higher overall coherence (both within test periods and minute-to-minute). Later work by Dinges, Maislin, Brewster, Krueger, and Carroll (2005) compared various proprietary fatigue management technologies (including actigraphy, PERCLOS, lane tracking monitors, and a steering centering device). They found that the PERCLOS monitor was successful in detecting bouts of drowsiness and provided feedback to drivers, which allowed them to adjust their sleep schedules accordingly.

Gaze

Examining the fixations and saccades associated with driving may provide information as to the driver's state of arousal. Two types of movements are of interest: smooth pursuit movements and saccades. These are usually measured via eye tracker (either coupled to the user's head or through cameras mounted directly in front of the person) or, if only the movements of the eye are of interest, by measuring the changes in the resting potential of the retina through small electrodes placed above and below the eye in electrooculography (EOG) techniques.

Smooth pursuit movements, or those movements where the eye's point of gaze is moving in a continuous fashion, are unlikely to be affected by drowsiness (Porcu, Ferrara, Urbani, Ballatreccia, & Casagrande, 1998). This is likely because of the strong presence of smooth pursuit movements during drowsiness (de'Sperati, 2005). Saccadic movements, which are those movements where the eye's point of gaze is rapidly shifting to the next fixation location (it is important to note that no visual information is being transmitted to the visual cortex during this time), has been found to relate to drowsiness (Porcu, Ferrara, Urbani, Ballatreccia, & Casagrande, 1998). Likewise, animal research has demonstrated a relationship between drowsiness and saccade amplitude, duration, and velocity. One interesting note is that a slow fluctuation in pupil diameter may be observed approximately 10 s before the onset of a microsleep. This fluctuation is not related to any gaze-specific coordinate (Heitmann & Guttkuhn, 2001). This method is impractical for active measurement due to current technological barriers. Additionally, the development of any system using gaze information to detect driver drowsiness has multiple problems to overcome, such as the eye disappearing from the camera's field of view when drivers look away from the forward visual field.

Electrophysiological Measures

Electroencephalography

The use of brainwave measurement techniques, such as EEG, has a long history in behavioral research. EEG has been demonstrated to reliably mark both level of alertness (Erwin, 1973; Artaud et al., 1994) and the transition from wake to sleep stages (Andreassi, 2000). Normal levels of alertness and arousal are indicated by beta band activity. Changes marking the onset of drowsiness are first noted in the alpha bands, with an increase in theta band activity indicating entry into sleep. EEG is frequently used in the measurement of driver levels of arousal (Brookhuis, Louwerens, & O'Hanlon, 1986; Brookhuis & de Waard, 1993; Brookhuis, 1995; Wylie, Schultz, Miller, Mitler, & Mackie, 1996). However, pragmatic issues make it difficult to implement such systems for use, as will be discussed below.

Although EEG has been reliable in marking drowsiness and fatigue within the driving task (Petit et al., 1990; Lal & Craig, 2000), several barriers to the creation of an effective DDWS exist. Primary is the possibility of resistance from drivers to the equipment required to measure brainwave activity. Measurement of brainwaves requires an electrode cap with several leads (usually between 16 and 32) to be positioned fairly accurately. The cap is typically tethered to a recording/processing system, which restricts driver movement to some extent. Furthermore, although EEG is able to demonstrate the transition between wakeful and sleep states, artifacts in the EEG signals may be produced by driver movements and vibrations of the vehicle. These have proven to be a barrier to the use of EEG in DDWS applications (Lal, Craig, Boord, Kirkup, & Nguyen, 2003).

Heart Rate Variability

An overall decline in heart rate is typically observed across periods of driving (Smiley, 2002). However, heart rate variability may prove to be a better measure of arousal. HRV is useful as a measure of both mental workload and stress. It has also been used as a physiological measure for driver drowsiness (Mulder, 1992). In early experiments, Volow and Erwin (1973) found that HRV (defined as the mean-square of heart rate) was a significant, yet unreliable, predictor of drowsiness onset. Similar findings were obtained by van der Berg, Neely, Wiklund, and Landström (2005), who found that HRV was sensitive to the transition between stages of sleep, but was a poor overall indicator of drowsiness onset. Boyle, Hill, Tippin, Faber, and Rizzo (2007) may have partially explained these prior findings by their examination of HRV in drivers with and without obstructive sleep apnea syndrome. Their findings indicated an increasing trend in HRV over time, but only for drivers with OSAS.

One, more unambiguous, finding regarding the use of HRV in detecting drowsy drivers was obtained by Byeon, Han, Min, Wo, Park, & Huh (2006). In the comparison of drivers at a normal level of arousal and those in a known drowsy state, drowsiness is indicated by a decreasing trend in the ratio of low frequency to high frequency signals in the analysis of HRV. However, using HRV as a sole detector of driver drowsiness would be problematic because of its sensitivity to issues such as mental workload, stress, and general arousal.

Galvanic Skin Response

Like most electrophysiological methods, galvanic skin response depends on the placement of electrodes on the person's exposed skin. This allows levels of electrical activity across the skin surface to be accurately recorded. The use of this information to determine levels of arousal has a long history. It is also well known that skin conductance, such as measured by GSR, is known to co-vary with the level of alertness (Nishimura & Nagumo, 1985). At a surface level, this makes GSR techniques highly attractive as a tool in detecting drowsiness.

However, several problems interfere with the implementation of GSR-based drowsy driver detection systems. Primary among these is the fact that drivers would be required to remain tethered to a system for measurements to be obtained. Even though the electrodes are comparatively small, on average 1 inch in diameter, they must be applied and removed from the skin. Additionally, GSR is sensitive to a variety of other factors not related to drowsiness. These include general cognitive arousal, mental workload, stress, and anxiety (Duley, 2006). Because of the uncertainty associated with the precision and specificity of GSR as a driver drowsiness detection metric, further development is needed before a practical system is ready.

Vehicle-Based Measures

Lateral Control

Steering Movements and Lane Position

Driver performance is an attractive measure for researchers since it may demonstrate a relationship between the driver's current state of arousal (e.g., drowsiness) and the primary task (driving). One of the driver performance measures with the strongest relationship to drowsy driving is the driver's lateral control of the vehicle. Lateral control is usually measured in terms of vehicle positioning (either lane position or lane crossing/violations), or the frequency and magnitude of steering inputs required to maintain the vehicle's position.

Successful operation of any vehicle requires continuous control of the vehicle's lateral position on the roadway surface. This is accomplished through steering wheel inputs, making both steering wheel movements and lane position attractive candidates for the monitoring and detection of drowsy driving. Boer's (2000) steering entropy provides a useful metric for the assessment of driver control. Steering entropy uses a time-series history of steering angle, comparing a predicted value to the observed value and capturing any steering correction of a significant magnitude (Boer, Rakauskas, Ward, & Goodrich, 2005). In a study assessing steering entropy during microsleeps (extremely brief-duration sleep), steering entropy proved to be a reliable measure of driver performance (Amit, Boyle, Boer, Tippin, & Rizzo, 2005). Unfortunately this research has not extended to the prediction of drowsy driving (a predecessor of microsleeps), where much more specific detection algorithms are required.

Lane positioning has proven to be a reliable and valid measure of driver performance (Boer, 2000), and it has demonstrated an association with drowsy driving (Rimini-Doering, Manstetten, Altmueller, Ladstaetter, & Mahler, 2001; Rimini-Doering, Altmueller, Ladstaetter, & Rossmeier, 2005). Some researchers (Oron-Gilad & Ronen, 2007) have speculated that the increase in lane position variance during drowsy driving is due to the driver attempting to reduce the demand imposed by the task of driving by manipulating their own level of performance, in effect attempting to cope for the dynamic nature of the task. Specifically, drowsy drivers may be differentiated via large amplitude, low frequency steering angle adjustments. This was mirrored in the findings of some researchers (Petit et al., 1990; Thiffault & Bergeron, 2003). However, questions regarding how to implement steering movement measures into any drowsy driver detection system have been raised (Fairclough, 1997). Some commercial devices using these methods of detection are now either in advanced development or currently available.

Steering wheel movements and lane positioning have a complex relationship. The correlation between lane position and (gross, continuous) steering wheel movements was not empirically supported in some early research (Sugarman & Cozad, 1972; Mackie & Miller, 1978). Later research specifically examining micro-movements have been able to demonstrate a correlation with lane positioning (Filiatrault, Cooper, King, Siegmund, & Wong, 1996). Also, existing automated lane positioning systems depend on clearly visible and consistently marked road lines being present. This is not always the case on the nation's highways, especially in areas undergoing maintenance or construction operations.

Velocity Control
In addition to findings indicating a strong relationship between drowsiness and the lateral control of a vehicle, there appears to be some relationship between velocity control and drowsiness. Moller, Kayumov, Bulmash, Nhan, and Shapiro (2006) described a gradual increase in driver's velocity (and the standard deviation of velocity) as a function of time of day. This effect has also been observed by researchers using sleep deprived drivers (Pizza, Contardi, Mostacci, Mondini, & Cirignotta, 2004), where daytime sleepiness is partially indicated by the frequency of exceeding the speed limit. Unfortunately, this method faces several problems before it may be implemented in any field system. Also, as noted by Evans (2004), drivers have very little incentive to drive exactly at the speed limit. Additionally, roadway demands force some degree of variability in the velocity of drivers' vehicles.

SENSOR PERFORMANCE

The Data Acquisition System

The data acquisition system used in this experiment is the result of over 15 years of development by VTTI. The DAS is a highly flexible centralized data collection device that has been successfully used in a number of naturalistic driving studies (e.g., Neale, Dingus, Klauer, Sudweeks, & Goodman, 2005). This system consists of a microcomputer which receives, processes, and stores data from modular sensors positioned throughout the vehicle. Besides being configured as a modular system, the DAS sensor array operates in a protected fashion; any single sensor failure does not stop data collection from any other sensor in the network. On-board storage is via removable hard drive, which allows for several weeks of data storage before downloading or hard drive replacement is necessary.

The main DAS unit is mounted in an inconspicuous location within the tractor, either under the passenger seat (for units without air ride seats), or in a side compartment. The network of sensors connected to the DAS includes a number of sensors not part of the current analysis. This allows for data to be collected and analyzed (or re-analyzed) as new or updated information is uncovered.

DFM/PERCLOS Monitor

Based on the results of previous research into methods of alerting drivers (Comsis Corporation, 1996) and monitoring their levels of drowsiness and fatigue (Wierwille et al., 1994; Grace & Benjamin, 1999), a promising candidate for a drowsy driver detection system was developed (Ayoob, Grace, Steinfeld, 2005; see Figure 2). This system uses a near-infrared camera coupled with processing equipment to estimate the driver's percentage of eye-closure (PERCLOS), which has proven to be a reliable measure of driver drowsiness (Wierwille, 1999).

Humans typically have very poor performance on assessing their own level of drowsiness (Dinges & Mallis, 1998). This device attempts to measure that initial period of the onset of drowsiness (Knipling & Wierwille, 1994) and alert the driver as to their state of drowsiness. The intent is to notify the driver before the onset of a microsleep, which allows the driver time to safely undertake countermeasures to drowsiness such as caffeine, napping, or other activities. An earlier simulator-based evaluation of the DDWS determined that drowsiness feedback resulted in improved driver performance, whereas various olfactory and vibro-tactile alerts did not have any discernable effect (Mallis et al., 2000).

Figure 2. The DFM Monitor

The system is mounted atop the dash of the tractor cabin within direct view of the driver. When the driver is detected at or above a set level of PERCLOS, visual and auditory alerts are provided. The driver acknowledges these through a response button on the top of the monitor device. Information about the rate, duration, and total number of drowsiness events is also given to the driver so that he or she may better understand their current state of arousal. A stable prototype of the driver fatigue monitor was constructed, and the manufacturer, Attention Technologies, Inc. (ATI), in conjunction with VTTI and the National Highway Traffic Safety Administration, sought to examine possible benefits of the DFM through a field operational test.

The Drowsy Driver Warning System estimates the driver's current level of drowsiness and triggers warnings when the system estimates that the driver's drowsiness is at or above certain pre-determined levels. The prototype system used in the present experiment is the DFM, which operates based on monitoring PERCLOS values (Wierwille et al., 2003), which has proven to be a reliable and valid method of detecting driver drowsiness. The DFM consists of a processing unit and an infrared camera, which is ideally located within a 20° angle from the driver's centerline. The DFM was developed by ATI. The most recent version of the unit (Generation 2) was used for the current experiment.

Road Scout Lane Tracker

VTTI has developed a lane tracking system, Road Scout, (Neale, Klauer, Dingus, Holbrook, & Peterson, 2001) which integrates with the DAS. Road Scout uses a single camera (black and white) coupled with a PC, frame grabber card, and a vehicle network interface for obtaining the vehicle's current speed. The grabbed video frames are processed algorithmically in real time (i.e., not stored on the Road Scout computer or DAS) and allow Road Scout to determine the vehicle's position relative to the lane markings present on the road.

Road Scout is able to determine the distance from the center of the truck to the left and right lane markings with a high degree of accuracy (estimated error: 2.0 inches average, 6.0 inches maximum), determine the offset between the truck's centerline and the road centerline (estimated maximum error < 1°), and provide status information for the road line markings.

The Road Scout lane tracker was not connected to the DFM for the purposes of this experiment. Instead, the information gathered by Road Scout was used for secondary analysis to determine road conditions, possible driver violations, and other information to help inform the analysis.

ALERT LOGIC AND THE DRIVER-VEHICLE INTERFACE

Alert Characteristics

For the current experiment, the DFM could operate in two different modes: Active and Dark. While in active mode the driver received feedback from, and could interact with, the DFM. The operational characteristics of the DFM while in active mode depended on certain conditions being met. These are described below. The other DFM operational mode was dark mode. This mode was activated by positioning the key-locked switch to the dark mode position (only the experimenters had access to this switch). In dark mode, the user interface was disabled (no information about alerts was provided to the driver); however, all data collection and processing of the DFM (including recording situations where warnings would have been given to the driver) continued. The DFM defaulted to low sensitivity (PERCLOS-5) while in dark mode.

Although drowsy driving is not restricted to any particular conditions such as time of day or operating conditions, the accurate functioning of the DFM system required specific conditions for the algorithm to function properly. For the system to activate, two conditions had to be met. First, the driver had to be traveling at 35 mph or higher. Second, illuminance reading must have been 50 lux or less. The first requirement allows for more efficient calculations of PERCLOS, as the driver should be making fewer gross head movements (keeping his or her eyes in view of the infrared camera). The illuminance requirement was due to difficulty in the DFM system's performance in daytime conditions. When these two conditions were met, the system became Active and (if set to do so) provided alerts to the driver. These alerts were based on the DFM's measure of PERCLOS (eyelid droop, as described above), using infrared cameras to observe the driver's eye. PERCLOS is a measure of how obscured the driver's eye (and thus, the pupil) is. As values of PERCLOS increase, the driver is generally considered drowsier. The system did not operate properly for drivers with glasses or sunglasses (i.e., eyes could not be covered).

For the purposes of data analysis, a third condition was added: the driver must be driving at the required speed in the required level of illuminance for at least 30 minutes. This condition did not interfere with the alerts given to the driver; it only affected the subsequent data analysis performed by VTTI.

The sensitivity of the system (a driver adjustable control, see below) was based on both the temporal window of the PERCLOS calculation and the PERCLOS sensitivity level at which the initial alert was triggered (Table 1). Full alert occurred at 12 percent PERCLOS in all sensitivity levels. The lowest sensitivity level, PERCLOS-5, was based upon a calculation period of 5 minutes. In this setting, initial alert occurred at 10 percent PERCLOS. Medium sensitivity, PERCLOS-3, was calculated over a 3-minute time period, with the initial alert occurring at 9 percent PERCLOS. The highest sensitivity level, PERCLOS-1, is calculated over a 1-minute time interval, with the initial alert occurring at 8 percent PERCLOS.

Table 1. Alert Sensitivity Levels

Sensitivity	Period	Initial Alert	Full Alert
PERCLOS-5 (Low)	5 min	10 percent PERCLOS	12 percent PERCLOS
PERCLOS-3 (Medium)	3 min	9 percent PERCLOS	12 percent PERCLOS
PERCLOS-1 (High)	1 min	8 percent PERCLOS	12 percent PERCLOS

If the DFM's estimate of PERCLOS was above the selected initial warning threshold (either PERCLOS-1, 3, or 5), an initial advisory alert was sounded. The PERCLOS calculation continued after this, and if it continued to rise to the full alert threshold the driver's selected auditory warning was repeated at the rate of once per second (a 1.0-Hz cycle) and accompanied by a visual alert on the DFM visual display. Alerts continued until the driver silenced them using the warning response button on the device.

Interface Characteristics

The DFM interface consists of a small box with a square camera mounting affixed to the top (Figure 3). The DFM is mounted on the dashboard of the truck. Whenever possible, the DFM was mounted to the right of the driver, as close to the centerline (Figure 4) as possible.

Figure 3. The DFM Mounted on the Dashboard of a Test Vehicle

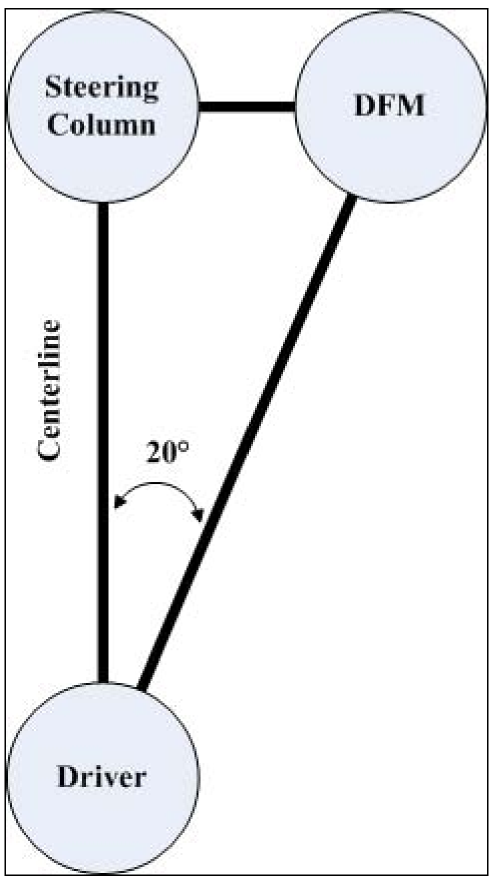


Figure 4. Placement of the DFM Relative to the Driver

However, in some cases adjustments to the DFM positioning were made to ensure the placement did not interfere with the driver's vision, steering wheel range of movement, or CB radio. A bracket was fabricated to ensure a solid and secure surface was available for the DFM in cases where the dashboard was too narrow for proper mounting of the DFM. Whenever possible, an angle of 20° to the right from the centerline was used as the maximum placement of the DFM.

Driver Adjustable Controls

The DFM had multiple driver-selectable adjustments. These included the sensitivity of the system, the brightness setting of the visual display, the sound used for auditory alerts, and the volume of the auditory alerts. Although these adjustments did not directly affect the operation of the system per se, driver selection of settings for these controls influenced the initial threshold for the alarm triggering (with the adjustment of the sensitivity) or the driver's perception of the alarm (with the adjustment of the visual and auditory characteristics).

Sensitivity

A switch on the front display panel of the DFM allowed for the user to select between three operational sensitivity levels. These were labeled as L (low), M (medium), and H (high) on the area surrounding the switch. The adjustment of this setting did not change the method of calculating the alerts provided to the driver. Instead, toggling the switch between the three settings changed the calculation period used to determine when the driver was drowsy.

As mentioned earlier, the DFM's operational state is partially dependent on the environmental condition present. When the driver's speed was below 30 mph and ambient illumination was greater than 100 lux, the DFM was in active-standby mode and displayed the word "Standby" on the visual display of the system. The driver adjustable controls were not illuminated during this time, and no feedback from the DFM was provided to the driver. Active mode with feedback to the driver occurred when speeds above 30 mph and ambient illumination less than 50 lux was detected. In this mode, a green system status light on the DFM visual display was active, the driver adjustable controls were illuminated and active, and the system provided alerts to the driver.

Display Brightness

The DFM display luminance was adjustable by the user. This setting was controlled by a knurled knob located on the front panel display. The DFM read this setting as an arbitrary value between zero (the minimum) and 255 (the maximum). In an initial test performed by VTTI, the DFM luminance settings were tested at five settings (equidistant within the range of adjustment), and the associated display luminance were measured. This testing was performed with the tractor's dash lights off and on. The results of these measurements are summarized in Table 2 and Table 3, respectively.

Table 2. Display Luminance (in Foot-Candles) With Dash Lights Off

Brightness Setting	1	2	3	4	5
Info Button	0	0.2	1.6	3.7	4.6
Sensitivity Button	0	1.2	3.4	6.6	9.2

Table 3. Display Luminance (in Foot-Candles) With Dash Lights On

Brightness Setting	1	2	3	4	5
Info Button	0	0.4	2.0	3.8	4.6
Sensitivity Button	0	1.3	3.5	6.8	9.6

Alert Sounds

The DFM allowed for the driver to select between six different alert sounds. Selection between the various auditory alerts was made via push-button switch on the left front display panel of the DFM. Drivers were able to toggle between the choices of alert sound by pressing the button, which played the selected alert sound for a period of 1.0 s.

Volume

Drivers were also able to control the volume of the auditory alerts. This selection was made via a knurled knob on the right-hand side of the front display panel. In order to obtain the alert characteristics across the various volume settings available to the driver, measurements on the A-weighted decibel scale were obtained. These measurements were taken with the engine stopped (ambient noise = 35.8 dBA) and running (ambient noise = 66.5 dBA), with the sound pressure meter at the approximate location of a driver's ear. These measures are summarized in Table 4 and Table 5.

Table 4. Sound Level of Alerts (in dBA) Without the Truck Engine Running

Volume Setting	1	2	3	4	5
Sound 1	59.9	62.0	65.3	73.1	76.2
Sound 2	45.8	57.5	61.1	64.2	65.6
Sound 3	66.4	77.2	79.6	80.1	81.9
Sound 4	66.7	68.8	79.3	80.0	81.2
Sound 5	64.5	79.2	79.9	81.9	82.5
Sound 6	63.3	76.2	76.4	78.1	79.1

Table 5. Sound Level of Alerts (in dBA) ith the Truck Engine Running

Volume Setting	1	2	3	4	5
Sound 1	66.8	70.6	71.9	75.4	75.9
Sound 2	66.7	68.1	70.8	71.6	71.8
Sound 3	71.1	77.2	79.1	81.1	81.4
Sound 4	70.5	79.0	79.5	80.3	80.5
Sound 5	70.6	77.3	80.2	81.8	82.8
Sound 6	69.3	76.6	79.7	81.7	82.8

Warning Responses

Warnings were acknowledged by the driver pressing the warning response button located on the DFM housing. This button stopped the warning and refreshed the visual display with information pertaining to the warning and the trip's past history. This information was given to the driver in the form of a bar display showing the duration of the longest eye-closure identified during the PERCLOS calculation period, and a numeric display of the total number of warnings and the overall warning rate (given in units of warnings per hour). The information on this visual display was provided for a period of 15 s or could be cleared from the display immediately by pressing the warning response button a second time.

THE FIELD OPERATIONAL TEST

Following initial prototyping and laboratory testing, NHTSA sought to examine the possible safety benefits of placing a DDWS into service. In order to answer this question, a field operational test was performed. Data for this project was obtained in an on-road (naturalistic) manner beginning in September 2002. This project included both Control and Test Groups using a quasi-experimental design.

The Control Group followed an A^9 design, where A (superscript refers to the prescribed number of weeks for that condition) refers to the condition in which the DFM collected data over the duration of the participant's involvement but never provided feedback to the driver. The Test Group experienced an A^2B^9 design, where A refers to the condition in which the DFM did not provide feedback to the driver (i.e., baseline and control conditions) and B refers to the condition where the DFM did provide feedback (i.e., the system was fully functional, or test conditions). In order to look for any adjustments in driving or sleeping behavior during the field operational test that were not attributable to the DFM, a baseline control condition was defined within the design.

Adjustments in the number of weeks for the baseline and active conditions were made with respect to data missing due to malfunctions in trucks or the DFM, and drivers not being able to meet with experimenters to switch the DFM to Active or finish participation at the exact time these milestones needed to happen. All these are an anticipated occurrence in any naturalistic data collection effort. Due to this accommodation for data collection, not all drivers experienced the exact same circumstances in the active condition, and some drivers had more baseline or active weeks than others. Other considerations in the adjustment of duration in the experimental conditions were made for inconsistencies in the length of drivers' trips. However, at the completion of the study all drivers (with the exception of those drivers leaving the study before completion) in the Test Group had a minimum baseline duration of two weeks followed by the active condition for approximately nine weeks (Figure 5).

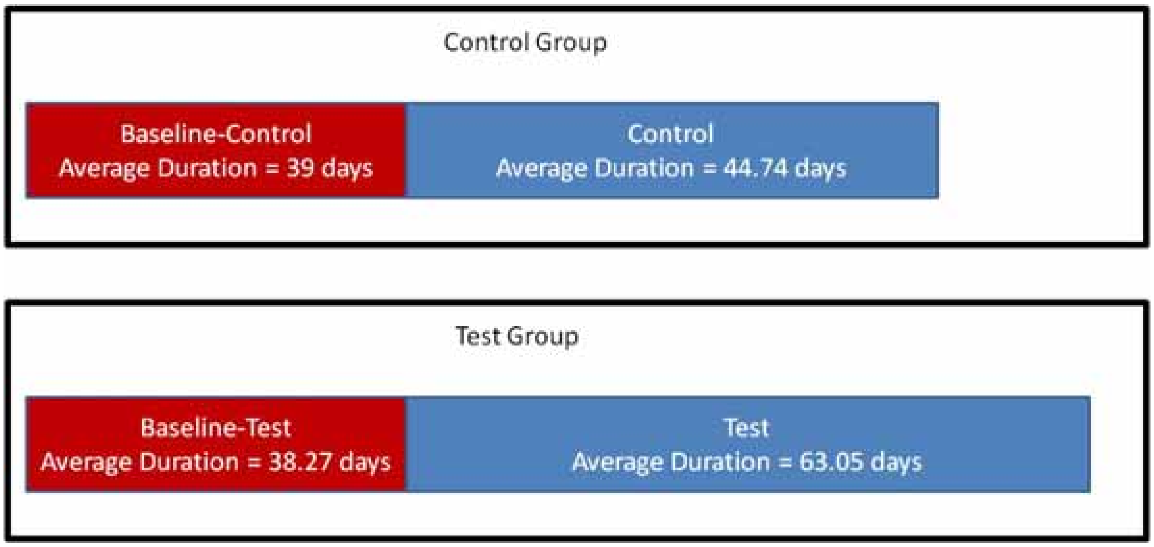


Figure 5. Final Timeline for Control and Test Groups

The experimental plan called for collection of data from 102 commercial vehicle drivers. This design was sensitive to the typical turnover rate (approximately 136 percent; Paz-Frankel, 2006) of drivers in the trucking industry. Some of the primary inclusion criteria for the study called for the drivers not to wear eyeglasses (a requirement of the prototype DFM system), to drive at night, and to include long-haul (i.e., drivers on the road for approximately one week) and line-haul (i.e., out-and-back) operations. The experimental plan also called for the instrumentation (with the DFM and sophisticated data collection equipment) of 34 commercial tractor vehicles.

The final study resulted in a data set with the following characteristics:

- 102[1] drivers (101 male, 1 female[2])
- Three for-hire trucking fleets and 46 instrumented trucks:
    - J.B. Hunt (19 trucks, 45 drivers)
    - Howell's Motor Freight (19 trucks, 41 drivers)
    - Pitt Ohio Express (8 trucks, 16 drivers)
- Experiment Groups:
    - Test Group = 78 drivers
    - Control Group = 24 drivers
- Operation Types:
    - Line-haul operations = 57 drivers
    - Long-haul operations = 45 drivers
- Final data set includes:
    - Approximately 12.4 terabytes (including video at ~30 Hz and truck instrumentation, kinematic data at 10 Hz) consisting of 2.4 million miles of driving and 48,000 driving-data hours
    - Load histories

[1] One driver with glasses was tested to assess system operation with glasses in a naturalistic setting. As expected, the DDWS was not able to reliably detect eyes in this real-world evaluation. As such, this driver will be removed from further analyses such that the final count is 102.
[2] This ratio of male to female drivers closely represents the actual distribution of CMV drivers on the road.

- Actigraphy data (measures of sleep quantity and sleep quality)

Participants

Companies

Three trucking companies participated in this study: Howell’s Motor Freight, J.B. Hunt, and Pitt Ohio Express. These companies provided a total of 46 trucks for instrumentation (Figure 6), and allowed for driver-participant recruitment at their terminals in Virginia and North Carolina. Of the three companies, two were primarily transporting dry goods using standard trailers and one hauled perishable goods in refrigerated trailers. Each company’s operations are profiled in the following sections.

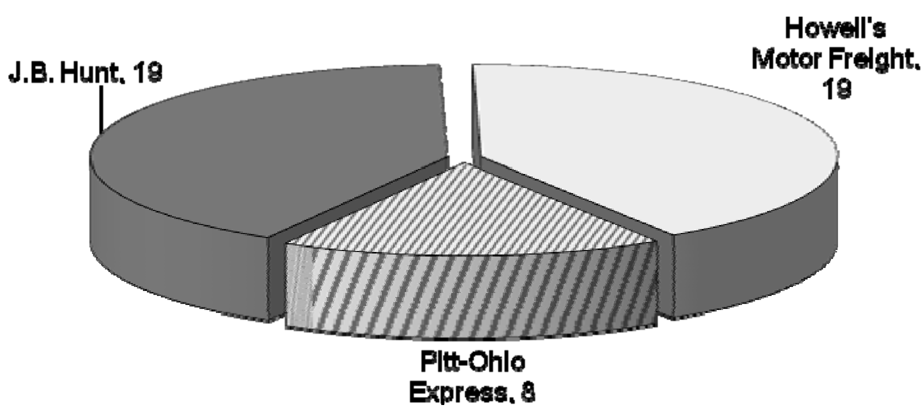


Figure 6. Proportion of Instrumented Truck for Each Participating Company

Howell’s Motor Freight. Howell’s Motor Freight, Inc., is a privately owned, Virginia-based, freight carrier providing temperature-controlled and dry truckload, less-than-a-truckload, and distribution services primarily to the wholesale/retail food and grocery industry and its suppliers. Howell’s currently operates 200 company-owned trucks and 500 refrigerated trailers. The trucks owned by the company include both day cabs and sleepers. The trailer fleet includes both 53-ft and 48-ft trailers. Drivers are usually assigned to particular trucks. Exceptions to this include assignment of new trucks and truck breakdowns. Howell’s operates seven terminal locations across five states. Howell’s temperature-controlled terminals in Roanoke, VA, Portsmouth, VA, Raleigh, NC, Charlotte, NC, and Atlanta, GA, provide full-service pool and LTL distribution as well as support for the truckload/line-haul division. The truckload/line-haul division operates in a territory that encompasses Delaware, Georgia, Illinois, Indiana, Kentucky, Maryland, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, Virginia, West Virginia, Missouri, and Alabama. The company also operates several stand-alone dedicated operations along with a logistics division for customer services. The two terminals that participated in the study were the Roanoke, VA, and Charlotte, NC, terminals.

Fifty-two company-owned trucks are based at the Roanoke terminal. Seventeen of these trucks were instrumented by VTTI for use in the study. Two of the 40 trucks based at Charlotte terminal were instrumented for use in the study. At the time of the study, operations at Howell's depended on 170 drivers. Of these, 42 were truckload drivers. Thirty-nine truckload drivers based in Roanoke, VA, and two based in Charlotte, NC, served as driver-participants in this study. Truckload drivers typically leave the terminal on Sunday or Monday and come back Friday or Saturday. Due to the nature of trucking operations, this schedule varies depending on the loads, due dates, etc. Some drivers may drive during weekends as well. Drivers report an approximate of 460 mi driven per day.

J.B. Hunt. J.B. Hunt Transport Services, Inc., is a major transport company headquartered in Lowell, AR, that hauls throughout the continental United States and portions of Canada and Mexico. Transport components include dry van, inter-modal, and dedicated contract services. J.B. Hunt operates approximately 10,600 trucks, along with approximately 26,087 trailers and 19,672 containers. The company employs approximately 12,800 drivers at 217 terminals. Six Virginia terminals participated in this study.

Fifteen of the 130 J.B. Hunt trucks based in the South Boston, VA, terminal were instrumented for use in this study. The South Boston terminal employs approximately 130 drivers, of whom 15 served as participants. J.B. Hunt was the dedicated service provider for a general goods store located in South Boston, Virginia. The majority of all loads hauled were performed during nighttime, with an approximate average of 500 mi per shift. During the course of the study, this dedicated contract was modified, and J.B. Hunt offered further participation in the study by transferring all instrumented trucks to a new location in Stuarts Draft, VA. Drivers also transferring to this new location continued their participation in the study.

J.B. Hunt was the dedicated service provider for a major chain store in Stuarts Draft, VA. The 15 instrumented trucks from the South Boston terminal were transferred to this terminal, bringing the total number of company trucks to 65. This terminal employed approximately 60 drivers, 16 of whom participated in the study. One extra truck was also instrumented to allow participation for an additional participant. The majority of all loads hauled from this terminal were performed at night, with an approximate average of 500 mi per shift running from Sunday mornings to Saturday evenings. After exhausting all the available and interested drivers for participation in the study, a transfer of 8 of the 15 instrumented trucks was made to a new location in Winchester, VA, for further participation recruitment.

J.B. Hunt was the dedicated service provider for another general goods store in Winchester. A total of 72 trucks were maintained by this terminal including 8 of the trucks transferred from Stuarts Draft. This terminal employed approximately 70 drivers, 12 of whom participated in the study. The majority of all loads hauled from this terminal were performed during nighttime, with an approximate average of 600 mi per shift running 7 days a week. After exhausting the population of available and interested drivers for participation in the study, J.B. Hunt offered further participation in the study by allowing VTTI to use three other small locations.

A dedicated service provider location was made available for driver recruitment in Winchester, VA. Serving a large home-improvement store, this terminal maintained 22 trucks (including the instrumented truck transferred in) and employed approximately 20 drivers, 1 of whom participated in the study. Approximately 80 percent of the deliveries were performed during the day, leaving 20 percent as overnight deliveries. The approximate average miles per shift were 400 mi. A second new location made available for driver recruitment was a J.B. Hunt dedicated service provider in Virginia Beach, VA. This terminal employed approximately 14 drivers, 2 of whom participated in the study. Approximately 40 percent of the deliveries were performed during the day, leaving 60 percent as overnight deliveries. A total of 17 trucks were maintained by this terminal including 2 of the instrumented trucks transferred specifically for use in the study. The approximate average miles per shift were 1,000 mi. The terminal generally performed deliveries from Sunday evening to Saturday morning.

The last location made available for driver recruitment was a J.B. Hunt dedicated service provider in Winchester, VA. This terminal employed five drivers, one of whom participated in the study. The majority of loads were performed on flatbed trailers. This small terminal delivered approximately 75 to 80 deliveries per week. A total of five trucks were maintained by this terminal including the instrumented truck use in the study. Drivers generally averaged about 400 mi per shift and work typically Sunday evening until Friday evening.

Pitt Ohio Express. Pitt Ohio Express headquarters are located in Pittsburgh, PA. The company is involved primarily in LTL operations, in which drivers go to deliver and return to the base location in a short period of time (usually less than 24 hours). The company has 20 terminals located mainly within Ohio and West Virginia. All the drivers have day trips and return to the terminal within 24 hours, driving approximately 500 mi per shift. Additionally, Pitt Ohio Express provides 14-ft van trucks for short-haul transportation at selected terminals. These van trucks were not employed for the DFM FOT.

Pitt Ohio employs a “slip-seat” operation for assigning trucks to drivers, in which one day and one night driver are paired up to drive the same truck. Each driver begins and ends their shift at the same terminal. Usually there is less than one hour between the daytime and the nighttime driver’s shift. The daytime drivers leave the terminal in the morning (approximately between 7 a.m. and 9 a.m.) and return in the evening (approximately between 7 p.m. and 8 p.m.); the nighttime drivers are informed about their departure time that evening, depending on delivery time, destination, and when their truck is expected to return to the terminal. The nighttime drivers return to the terminal the next morning (approximately between 5 a.m. and 8 a.m.). At this point the daytime driver gets into the truck again. The operation runs from Monday morning with the daytime drivers to Saturday morning with the nighttime drivers. All drivers participating in the study were nighttime drivers, and the instrumented trucks were day-cab type trucks (i.e., no sleeper in the cab). Drivers from two Pitt Ohio Express terminals (Roanoke and Richmond, VA) participated in the study.

As of June 2005, the Roanoke terminal employed 7 dispatchers and 54 drivers. Thirty-seven of the 54 drivers are daytime drivers, and the remaining 17 are nighttime drivers. Nine of the 17 nighttime drivers participated in the study. The Roanoke terminal has 24 tractors with single trailers and 12 straight (single unit) trucks. Four of the 24 tractors were instrumented by VTTI to be used in the study.

The Pitt Ohio Express Richmond terminal is base for more than 50 drivers. Approximately 48 of them are tractor-trailer drivers. Approximately 30 of those are daytime drivers, and 18 are nighttime drivers. The terminal also runs single unit trucks. Four trucks were instrumented by VTTI to participate in the study. Seven of the 18 nighttime drivers qualified and agreed to participate in the study. The primary routes for the participating drivers included driving from Richmond to terminals in Baltimore, MD, Cumberland, MD, and Roanoke, VA. The drivers typically logged between 5 and 9 hours of driving time per night, 5 nights per week. None of the drivers drove regularly on Saturday or Sunday nights.

Drivers

Data for a total of 103 drivers holding a Class-A commercial driver's license was collected for the study. One of these participant drivers wore glasses and was not counted in the final number of drivers in some instances (driver selection was based on screening for optimal detection by the DFM, based on a previous characterization study at VTTI). Of the 103, there were 101 male drivers and 1 female driver. The majority of drivers (46) came from J.B. Hunt, 41 drivers worked for Howell's Motor Freight, and the remaining 16 were from Pitt Ohio Express (Figure 7). Although some drivers were excluded from the analysis due to missing data, failures of the DAS or DFM, or other problems expected in the course of a naturalistic driving study, the information presented here provides demographic information on all drivers involved in the study.

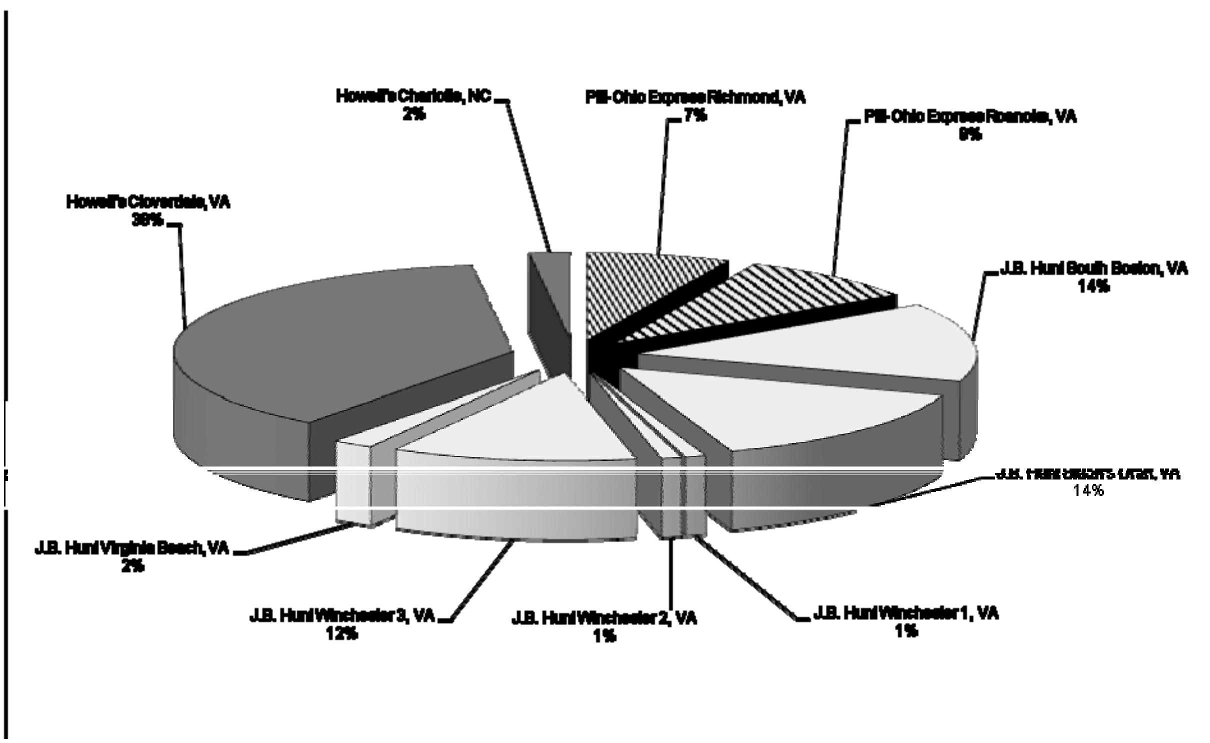


Figure 7. Proportion of Drivers per Company and Terminal Location

Drivers possessed a range of age (*M* = 40.0 years, *SD* = 8.24, range 24 and 60, see Figure 8) and driving experiences (*M* = 10.6 years, *SD* = 8.37), allowing for a more comprehensive examination of the DFM. Both line-haul (out-and-back) and long-haul (out for approximately one week) operations were represented in the sample. The majority of the participants were Caucasian-Americans (65%), with the remaining participants identifying as African-Americans (29%), Native-American (4%), Asian American (1%), or Hispanic-American (1%); see Figure 9. Ninety-four participants reported their highest achieved education level. Of these 94 participants, 49 percent reported having some type of education after high school, 36 percent had high school as their highest education level, and 15 percent did not finish high school education.

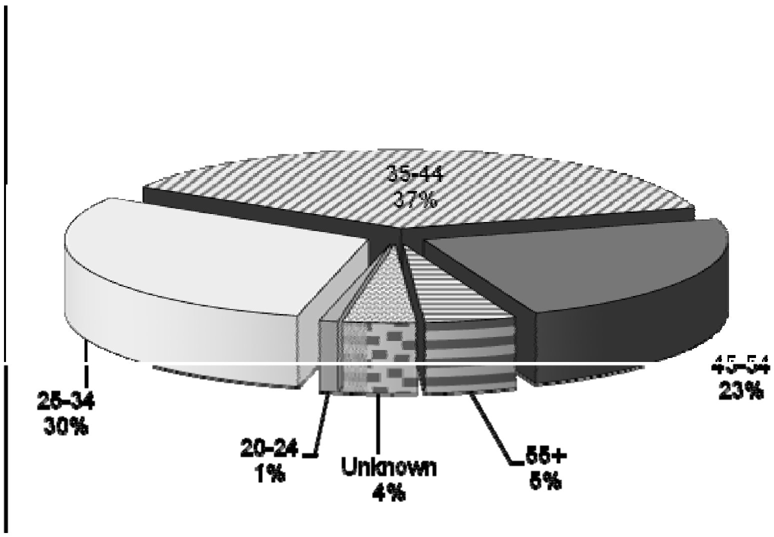


Figure 8. Driver Age Distribution

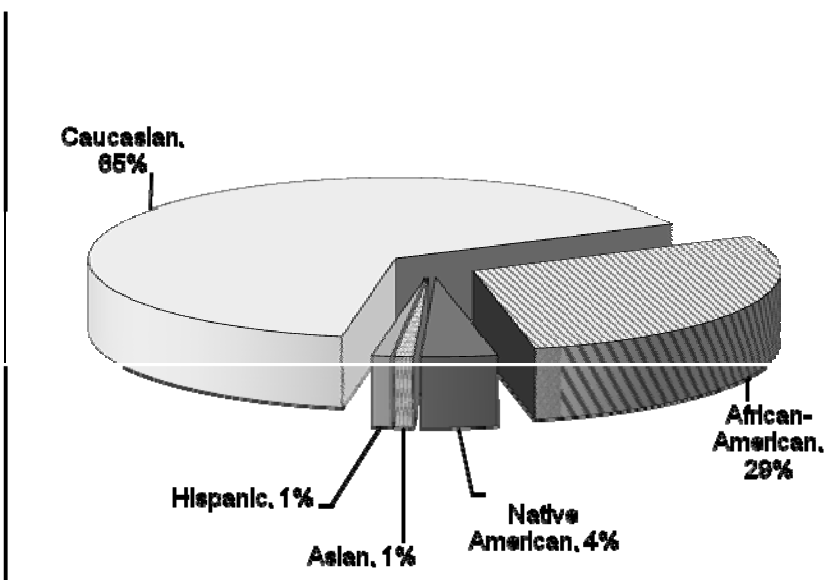


Figure 9. Driver Ethnicity Distribution

Inclusion criteria for drivers were based on the operational parameters of the DFM being tested: (1) engaged in night driving, (2) did not wear glasses while driving, (3) had a low risk of dropping out or leaving the company, (4) passed vision and hearing tests, and (5) retinal reflection of IR from the DFM was deemed optimal. All participation was completely voluntary. Drivers were compensated for their time with a total payment which varied depending on the length of their participation (10 weeks, $1,050; 14 weeks, $1,350; 16 weeks, $1,500). A Certificate of Confidentiality was obtained, granting confidentiality to research participants. This confidentiality is provided by the Public Health Services Act (§ 301(d), 42 U.S.C. 8241(d)), and ensures protection against compulsory legal process for personally identifiable research information.

Technical Problems

As with any prototype technology, the experimental units used in the field operational test had some degree of system failures and problems. Among the implemented systems, there were 86 incidences of technical problems. Reported problems were recorded and classified according to whether they were problems primarily of the program and software associated with the unit, or a problem with the hardware components of the unit. In the FOT, 28 program and software problems were observed (33% of the total number of technical problems observed; see Table 6). Units with these problems were returned to the manufacturer for repair.

Table 6. Frequency and Type of Program and Software Problems Observed

Program/Software-Related Problem	Frequency
Old software was installed	5
DFM cannot be activated	4
DFM does not produce any output/signal to the DAS	4
PERCLOS values were not changing correctly	4
Program in the microchip was damaged by sunlight	3
"Eyes found" was not working correctly	3
"Standby" was lit even when the truck was driving over 35 mph at night	2
DFM in dark mode, yet worked as if it was in active mode ("Standby" was lit even if DFM was in dark mode; DFM in dark mode presents warnings)	2
Sensitivity was set to high instead of low in dark mode; Sensitivity outputs included 4^{th} level even if only 3 levels (low, medium, high) were available	1
Total	**28**

Fifty-eight hardware problems (67% of the total number of technical problems) were observed. Of these, VTTI was able to repair 34 issues, while the remaining 24 problems required the devices to be returned to ATI for repair. Table 7 shows the frequencies of observed hardware problems during the FOT. The majority of program/software- and hardware-related problems were most likely normal wear and tear; however, some of the hardware-related problems may have been caused by drivers.

Table 7. Frequency and Type of Hardware Problems Observed

Hardware-Related Problem	Frequency
Speaker fell off inside of the DFM	24
DFM had rattle or loose screws	8
IR sensor in the DFM camera did not flash	7
Ambient light sensor attached to the DFM was damaged or fell off	4
DFM camera did not send the signals for view	3
Cable was broken and no signal could be sent from DFM (e.g., no auditory warning)	3
The size of hole for the buttons was inadequate and the buttons were stuck	3
The warning response button (to stop the auditory warning) became loose	3
DFM drew too much power and shut down the DAS	2
"Standby" was not lit in active mode	1
Total	**58**

Among the 41 DFM units used in the DFM FOT, 12 units (29%) did not have any technical problems while the remaining 29 units (71%) had at least one hardware- and/or program/software-related problem. Among the 29 units with at least one technical problem, three units were responsible for 21 (24.4%) of the program/software- or hardware-related malfunctions.

CHAPTER 3. RESEARCH QUESTIONS AND OPERATIONAL DEFINITIONS

RESEARCH QUESTIONS

Based on the objectives of this research effort, VTTI outlined a total of six key topics of interest necessary to examine the DDWS safety benefits. Each of these topics of interest was then operationalized as a research question (RQ), with individual components serving as sub-questions. They are delineated as follows.

Research Question 1 - On-the-Job Drowsiness

To evaluate the question of a DDWS's impact on the incidence of on-the-job drowsiness, PERCLOS data were collected during driving episodes matching a set of pre-defined conditions. Using these data, the following research questions were addressed:

RQ 1.1: Does the DDWS result in fewer episodes of drowsy driving?
RQ 1.2: Does the frequency of DDWS alerts decrease over time?
RQ 1.3: Do DDWS alerts have an impact on post-alert behavior?

Research Question 2 - Sleep Hygiene

The implementation of a DDWS may improve the salience of sleep's importance in safe driving and could prompt drivers to make changes to their sleeping habits. In order to address this issue, the following research questions were examined:

RQ 2.1: Does a DDWS influence drivers to get more sleep?
RQ 2.2: Do drivers using a DDWS achieve better quality sleep?

Research Question 3 - Involvement in Safety Critical Events

Critical incidents are situations within the driving task where a crash, near-crash, or crash relevant conflict occurs. Implementation of the DDWS may have an effect on driver involvement in such events. The investigation of this possible relationship is outlined as follows:

RQ 3.1: Does a DDWS affect involvement in safety critical events?
RQ 3.2: Does a DDWS affect involvement in at-fault safety critical events?

Research Question 4 – Human-Machine Interaction

The human-machine (driver-system) interface of the DDWS is of great importance to both acceptance on behalf of the driver and a reduction in overall system performance variance due to specifics of the DDWS interface. Thus, examining the characteristics of the DDWS operation is necessary.

RQ 4.1: How do drivers operate the DDWS in a real-world environment?

Research Question 5 - Favoring Drivers - Follow-Up

The safety benefits that may be obtained via the DDWS are partially dependent upon optimal system performance. Understanding situations where optimal performance was not obtained will allow the research teams to both improve subsequent versions of the DDWS and better understand the system's current and potential operational envelope.

> *RQ 5.1*: How did the DDWS operate for drivers who rated the system favorably in the post-study survey?

Research Question 6 - At-Risk Driver - Follow-Up

Due to the large variability known to exist within the population of commercial heavy-vehicle drivers (Knipling, 2005), there are potential impacts on driver risk (Lancaster & Ward, 2002). To better explore these factors and gain a better understanding of what led to these drivers receiving more alerts, the following research question was examined.

> *RQ 6.1*: How did the DDWS operate for drivers who had significantly more alerts during the baseline?

OPERATIONAL DEFINITIONS

To facilitate understanding as well as to maintain consistency in the analyses, several key concepts were defined for the current research effort. Any subsequent reference to the terms defined below may be considered to use these definitions.

PERCLOS

1. PERCLOS
Operational definition: A measure of slow eye-closure (eye droop). PERCLOS is defined as the percentage (alternatively defined as the proportion) of time that the eyes are between 80 and 100 percent closed during a defined time interval (Wierwille, 1999). The driver is considered drowsier as the amount of PERCLOS increases.

2. DFM PERCLOS Value
Operational definition: The PERCLOS value calculated by the DFM.

3. Manual PERCLOS Value
Operational definition: The manually calculated PERCLOS value for the 3.0-min time period preceding an event.

4. P1 or PERCLOS-1 (High Sensitivity)
Operational definition: PERCLOS value calculated over a period of 1 min; initial advisory tone occurs at 8 percent PERCLOS (4.8 s), with full warning occurring at 12 percent PERCLOS (7.2 s).

5. P3 or PERCLOS-3 (Medium Sensitivity)
Operational definition: PERCLOS value calculated over a period of 3 min; initial advisory tone occurs at 9 percent PERCLOS (16.2 s), with full warning occurring at 12 percent PERCLOS (21.6 s).

6. P5 or PERCLOS-5 (Low Sensitivity)
Operational definition: PERCLOS value calculated over a period of 5 min; initial advisory tone occurs at 10 percent PERCLOS (30.0 s), with full warning occurring at 12 percent PERCLOS (36.0 s).

Safety Critical Events

7. Safety Critical Events (SCE)
Operational definition: A safety critical event (SCE) is an event that may be classified as either a crash; crash: tire strike; near-crash; or crash-relevant conflict.

8. Crash
Operational definition: Any contact with an object, either moving or fixed, at any speed. Includes other vehicles, roadside barriers, objects on or off of the roadway, pedestrians, cyclists, or animals.

9. Crash Relevant Conflict
Operational definition: Any circumstance that requires a crash avoidance response on the part of the participant vehicle, any other vehicle, pedestrian, cyclist, or animal that is less severe than a rapid evasive maneuver (as defined in near-crash) to avoid a crash, or any circumstance that results in close proximity of the participant vehicle to any other vehicle, pedestrian, cyclist, animal, or fixed object where, due to apparent unawareness on the part of the driver(s), pedestrians, cyclists or animals, there is no avoidance maneuver or response. A crash avoidance response can include braking, steering, accelerating, or any combination of control inputs. Examples of potential crash-relevant conflicts include hard braking by a driver because of a specific crash threat, or proximity to other vehicles.

10. Crash: Tire Strike
Operational definition: Any contact with an object, either moving or fixed, at any speed where the contact occurs on the truck's tire only. No damage occurs during these events (e.g., a truck is making a right turn at an intersection and runs over the sidewalk/curb with a tire).

11. Event ID
Operational definition: A unique identification number assigned to an SCE of interest.

12. Near-Crash
Operational definition: Any circumstance requiring a rapid evasive maneuver by the participant vehicle or any other vehicle, pedestrian, cyclist, or animal to avoid a crash, or any circumstance that results in extraordinarily close proximity of the participant vehicle to any other vehicle, pedestrian, cyclist, animal, or fixed object where, due to apparent unawareness on the part of the drivers, pedestrians, cyclists or animals, there is no avoidance maneuver or response. A rapid

evasive maneuver is defined as steering, braking, accelerating, or any combination of control inputs that approaches the limits of the vehicle's capabilities.

Driver Fatigue Monitor

13. Driver Fatigue Monitor Output

Operational definition: The DFM data collection, identifying Sensitivity Level, Eyes Found, Operating Mode, Display Status, Display Mode, Warning Flag, Warning Sound, and Ambient Brightness.

14. Alert

Operational definition: Visual and auditory warning given to the driver when the DFM determined that the PERCLOS reached a predefined threshold. The DFM produced auditory and visual warnings. The driver received an initial advisory tone and a full warning, each of which is based upon the sensitivity level selected (see P1, P3, P5 PERCLOS).

15. Full Alert (Full Warning)

Operational definition: The second warning the driver receives. This warning follows several seconds after the initial advisory tone if the DFM continues to detect closed eyes. The following components are associated with full warnings:

- Auditory warning: The warning sound chosen by the driver. The driver halts this warning by pressing the OK button on the DFM.
- Visual warning: The three visual components of the DFM display: bar graph, rate, and total.

16. Initial Alert (Initial Advisory Tone)

Operational definition: The first warning received. This auditory warning will sound after a predetermined length of time (depending on the sensitivity setting; see definitions 4, 5, and 6: P1, P3, P5 PERCLOS, respectively). A single presentation of the tone is given and does not require driver interaction.

17. Invalid Alert

Operational definition: A DFM alert provided to the driver that was not due to slow eye-closure.

18. Valid Alert

Operational definition: A DFM alert provided the driver that was due to slow eye-closure.

Exposure and Experimental Design

19. Day-in-Study

Operational definition: The number of days since the driver began the study until the time of the alert.

20. Driving Hour

Operational definition: Number of hours driving from beginning of the shift until the alert, episode, or SCE occurred.

21. Episode

Operational definition: Portion of driving data that complies with all the conditions of interest: speed ≥ 35 mph, lux ≤ 10, and continuous driving ≥ 30 min.

22. Experimental Condition

Operational definition: The two main experimental groups are the Test Group and Control Group. The DFM was always available in the vehicles for both Test Groups. However, for the baseline it was set to dark mode (no alerts displayed to the driver). In certain experimental conditions the DFM provided an audible/visible alert (active mode) to the driver if condition of interest was met (see definition 21, Episode). These are the four experimental conditions:

- Test Group
  - i. Baseline Test: No alert is provided to the driver during these weeks (dark mode), but the data stream records when the alert would have occurred.
  - ii. Test: Alerts are provided to the driver based on the driver-selected settings during these weeks; alerts are stored as part of the data stream.
- Control Group
  - i. Baseline Control: A defined period of time for analysis purposes beginning with the start of the driver's participation. This group was created to provide an equivalent comparison group to the baseline test condition. Similar to baseline test condition, no alerts are presented to the driver, but they are recorded as part of the data stream.
  - ii. Control: No alert is provided to the driver during these weeks, but the time the alert would have occurred is recorded as part of the data stream.

23. Shift

Operational definition: A measure of the driver's time-on-task, defined by the following characteristics:

- Break Length (BL): Number of minutes the break speed needs to persist to be counted as a break. For this calculation BL > 10 min. If BL > 6 hours, the beginning of the next file is considered a new shift.
- Break Speed (mph): Any speed ≤ 5 mph for the time specified as break length (BL) constitutes a break. (Separates non-driving time from time at a stoplight.)
- Regulation time: Number of hours from beginning of shift before a new trip can start. Default was 14 hours. If the break happens at the 14-hour mark or after the 14 hours, the next file is considered a new trip.

24. Shift-in-Week

Operational definition: The number of shifts the driver has driven since the beginning of the work week until the time of the alert, episode, or SCE.

25. Week-in-Study
Operational definition: Number of weeks since the driver began the study until the beginning of the alert, episode, or SCE.

26. Week-in-Treatment
Operational definition: Number of weeks since the driver began using the DFM in active mode.

27. Weeks Since DFM Mode Change
Operational definition: Number of weeks since the DFM was changed from Dark to active mode until the beginning of the alert, episode, or SCE. A week is defined as 7 calendar days from the day the DFM was switched to Active.

28. Time-of-Day
Operational definition: The time of day when alert, episode, or SCE of interest occurred. This figure uses the beginning of the alert, episode, or SCE as the time of interest.

29. Time-within-Shift
Operational definition: Number of hours from beginning of the shift until the alert, episode, or SCE occurred.

30. Day-of-Week
Operational definition: The day of the week in which the alert, episode, or SCE occurred.

31. Post-Alert Behavior
Operational definition: Actions the driver made during the 5-min epoch immediately following an alarm.

32. Time from Alarm to Stop
Operational definition: The amount of time from when the alarm is presented until the driver stops. Stop lights, stop signs, yielding, construction, or traffic are not considered a stop of interest.

33. Total Driving Time per 24-Hour Period
Operational definition: Total on-duty driving time per 24-hour period (midnight to midnight) starting with the driver's first full day of participation, in minutes.

34. Hours Driven
Operational definition: The number of hours driven by the participant during the study under the experimental condition of interest.

35. Number of Alerts for the Prior Hour
Operational definition: The number of DFM alerts one hour before the SCE of interest. This is a sum of initial and full triggers and is reported as an integer.

36. Number of Initial Trigger Alerts for the Prior Hour

Operational definition: The number of initial-trigger alerts one hour before the SCE of interest, reported as an integer.

37. Number of Full Trigger Alerts for the Prior Hour

Operational definition: The number of full-trigger alerts one hour before the SCE of interest, reported as an integer.

38. Number of Lanes

Operational definition: The number of travel lanes of the road the driver is on during the alert, episode, or SCE of interest. This is collected via visual inspection during data reduction.

39. Number of Valid Alerts for the Prior Hour

Operational definition: The number of valid DFM alerts one hour before the SCE of interest. This is a sum of valid initial triggers and valid full triggers, reported as an integer.

40. Number of Valid Initial Trigger Alerts for the Prior Hour

Operational definition: The number of DFM valid initial trigger alerts one hour before the SCE of interest. This is a sum of valid initial triggers and valid full triggers, reported as an integer.

41. Number of Valid Full Trigger Alerts for the Prior Hour

Operational definition: The number of DFM valid full alerts one hour before the SCE of interest, reported as an integer.

42. Frequency of Driver Input

Operational definition: A count of each instance of the driver changing the sensitivity level, warning sound, or display brightness setting.

43. Alert Duration

Operational definition: The temporal duration between the driver receiving a full alert and the moment the warning response button is pressed.

Sleep

44. Actigraph Data

Operational definition: Data obtained from an actigraph motion logging device. The device is roughly the size and shape of a wristwatch and constantly records motion (motor unit activity) of the wrist and arm. Periods with little to no motion logged are indicative of sleep.

45. Scored Sleep Period

Operational definition: The portion of time scored a sleep by the Cole-Kripke sleep algorithm using wrist motion inactivity.

46. Sleep Onset-Offset Interval (O-O Interval)

Operational definition: The period of time the driver is attempting to sleep. The O-O interval was derived from driver actigraphy data based on SSPs of 20 min or longer. SSPs occurring either within an hour of the last identified SSP or that end within 30 min of the beginning of the first 20 min SSP are grouped together. The O-O interval spans from the beginning of the first SSP to the end of the last SSP.

47. Sleep Quality

Operational definition: An estimate of a driver's overall quality of sleep based on the following measures:

- Minutes of sleep
- Sleep efficiency during sleep-onset/sleep-offset interval
- Number of awakenings
- Long sleep episodes (episodes of sleep lasting greater than 20 min)
- Longest sleep episode

48. Sleep Quantity

Operational definition: The amount of sleep the driver had in the 24-hr block from the beginning to the end of the measurement period.

CHAPTER 4. SAFETY BENEFITS

INTRODUCTION AND OVERVIEW

The purpose of a warning system for drowsy driving is initially to detect drowsiness, and optimally inform drivers as to their current state of arousal. For the system to be deemed effective, it must not only provide timely and noticeable warnings, but also truly assist drivers in making safety-conscious decisions regarding drowsy driving.

As stated before, there are several basic assumptions for examining the impact of a DDWS on driving safety. Primary among these is that the feedback provided to drivers must be timely and accurate (Wierwille, 1999). It is expected that drivers would adapt to such warnings and take one or more actions that may ameliorate this situation (Figure 10).

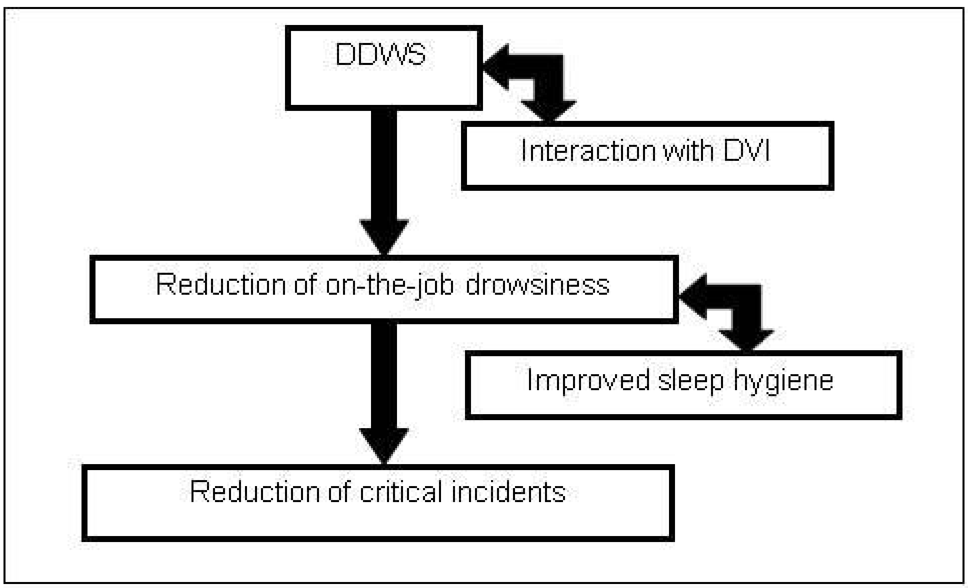


Figure 10. Modeling Safety Benefits Associated with a DDWS

These actions may include sleeping, taking a break from driving, ingesting caffeine, or some other stimulating action. The benefits model outlined in the introduction (see Chapter 1) predicts that drivers will experience a positive change in behavior after gaining experience with, and feedback from, the DDWS. That is, they will be positively influenced in terms of their sleep schedules, be more aware of drowsiness onset while driving, and be less likely to be involved in SCEs due to drowsiness.

The following are the findings of an examination of the safety benefits of a DDWS, with respect to the previously stated research questions. The methods used, results obtained, and a discussion of each major research question is presented below. In addition, a global view of the possible safety benefits of a DDWS is discussed.

EXPERIMENTAL CONDITIONS

Based on work by the VNTSC, the independent reviewer for this experiment and the original designer of the FOT experiment, the experimental design for the DDWS went through several iterations and modifications. Each of these iterations and modifications is documented in the data collection report for the DDWS FOT (Hanowski et. al, 2008). As the data collector for the study, VTTI performed the recruitment of fleets and driver participants and assigned the participants to the suggested conditions based on the VNTSC experimental design and modifications. The first implemented modification was composed of the Control Group (A^{12}, baseline for 12 weeks) and the Test Group (A^3B^9). The Test Group had a baseline (A) sequence for the first three weeks and a treatment (B) sequence for the last nine weeks. The Control Group had a baseline (A) for the duration of their participation. Modification 1 allowed up to four additional weeks per driver after the original 12 weeks of data collection were finished for additional data collection, if needed, to compensate for data missed due to truck or other system breakdowns.

The second experimental design modification prescribed by VNTSC was a similar design to the one described in Modification 1. Modification 2 retained the Control Group (A^9), but for only nine weeks with the potential of an additional week of data collection if needed. The Test Group for Modification 2 (A^2B^9) had a two-week baseline sequence (A) and kept the treatment sequence (B) of nine weeks. The Test Group for Modification 2 included three additional weeks. A similar proportion of line-haul and long-haul participants were recruited for each group, resulting in the final durations given below.

As is expected in a naturalistic study examining a prototype system, several vehicle and system breakdowns occurred. Therefore, the participation for some drivers was extended. These time extensions used the allocated additional weeks suggested by the experimental design. After reviewing the data collected for all 102 drivers participating in the study and all experimental conditions, some drivers' data were not included in the data analyses discussed next. Hence, a total of 96 drivers (73 in the test condition and 23 in the control condition) were included in the analyses presented in this chapter.

The average duration of the baseline condition for the Test Group (baseline test) was 38.92 days (Figure 11). This average duration represents the number of calendar days between the driver beginning the study and when the vehicle was available to switch the DFM from Dark to active mode. This average does not mean that data was available for all days, due to various system failures or drivers being off duty or not using the experimental vehicle. The baseline test mean duration represents the 73 drivers included in the Test Group and reflects an expanded baseline due to technical problems or not having access to the driver (due to their delivery schedule and other company logistics) to switch the system from Dark to Active. These 73 drivers represent 16 line-haul and 57 long-haul drivers. For the line-haul drivers, the mean baseline test duration was 37.13 days. Long-haul drivers' average baseline test was 39.4 days.

In order to measure the equivalence of the baseline test for the two different haul types represented in the Test Group, a t-test was used. The *p*-value obtained for this test was 0.7745 (using a Satterthwaite adjustment for unequal variance). There is no statistically significant difference in the baseline test duration between the two types of drivers represented in the Test Group. For the purpose of the safety analysis, a period of time matching the baseline test condition was selected for the Control Group (i.e., baseline control). For making a meaningful comparison, a 39-day period for baseline control condition was selected for all 23 control drivers to reflect the average length of the baseline Test condition. Therefore, this baseline control condition is defined as the 39-day period after the driver started participating in the FOT. Figure 11 illustrates the Control Group conditions in comparison to the Test Group conditions. The portion of the participation following the baseline for each group is referred to as Control and Test conditions, respectively. This split created four experimental conditions. The control condition lasted 44.74 days on average, while the test condition lasted an average of 63.05 days. Thus, the experimental design used for this FOT made the participation of the Control Group somewhat shorter than that of the Test Group. In summary, the effect of the imbalance was corrected statistically in the analysis. Therefore, the groups are effectively equivalent for subsequent comparisons.

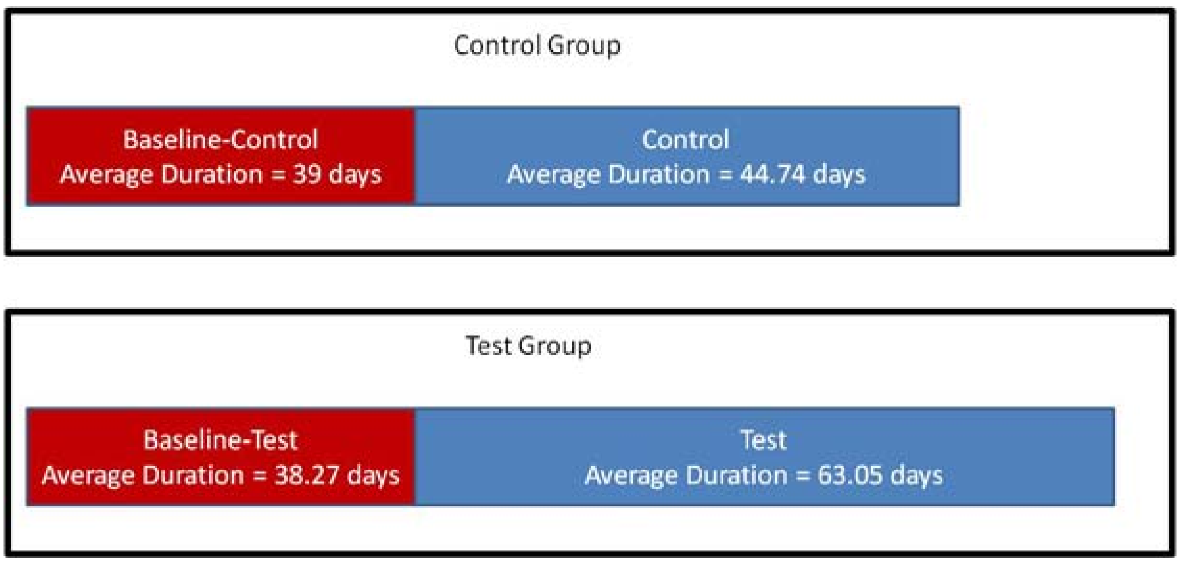


Figure 11. Average Length of Experimental Conditions

RESEARCH QUESTION 1: ON-THE-JOB DROWSINESS

For this Research Question (RQ) three sub-questions were answered:

RQ 1.1: Does the DDWS result in fewer episodes of drowsy driving?
RQ 1.2: Does the frequency of DDWS alerts decrease over time?
RQ 1.3: Do DDWS alerts have an impact on post-alert behavior?

Methods, analysis, measures of interest, results, and discussion related to these three questions are presented next.

Methods

Episodes of Interest

The purpose of this safety benefit analysis was to evaluate if a warning system that alerts drivers when drowsiness is present would result in less drowsiness, fewer alerts, and more frequent stops to take breaks across time. The focus of the analyses performed for Research Question 1 were the data within an "episode of interest" during the drivers' delivery runs. An episode of interest was operationally defined as an instance in which the participant was driving the test vehicle for at least 30 min at a speed of 35 mph or faster with an in-cab illuminance reading of 10 lux or lower. The criteria help reduce the number of instances where the system might operate unreliably due to extraneous lighting or driver movement conditions. The system used for this FOT should be considered a prototype as it was not an off-the-shelf production DDWS. At the time the FOT was performed, it was the best system available for this type of concept evaluation. Limitations existed in its operation (e.g., only worked in low illumination, driver could not be wearing glasses). The evaluation of the system was restricted to minimize instances where the known limitations would affect the operation of the drowsiness monitor concept. Another system limitation was discovered during the FOT data collection process. The DFM categorized any glance away from the forward roadway as eyes-closed, an action which artificially inflated the PERCLOS values used to calculate alert presentation. Therefore, an episode determination procedure was deemed necessary to eliminate, as much as possible, conditions such as stop-and-go traffic, driving in a parking lot, and other instances that might occur outside of the operating envelope of the DFM. By eliminating the type of conditions where the driver would be constantly monitoring the mirrors or surroundings, the assessment would focus on alerts presented due to drowsiness while the eyes were predominantly on the forward roadway.

Alert Validation

In addition to the episodes of interest defined above (which are relevant for all three questions under Research Question 1), valid alerts were identified through a visual inspection of data during the reduction process. Appendix A presents the protocol that was followed. The visual evaluation of alerts was performed for a sample of the alerts presented within the episodes of interest for each participant in the study. Even when the system was operating in the Dark mode, alerts were recorded, even though they were not presented to the driver. Only alerts during the Active mode were presented to the driver. The data reduction process included each week of participation for each driver. Data reductionists visually evaluated over 15,000 total alerts. This process categorized the alerts as valid or invalid. If the alert was deemed invalid the main reason was determined (see Appendix D for taxonomy of invalid alerts). A total of 721 valid alerts were obtained from this process.

The analyses performed to answer Research Questions 1.2 and 1.3 used the 721 valid alerts. Taking only valid alerts into consideration allowed for the analysis to focus specifically on the DDWS concept. Therefore, if the alert was considered a false alarm the alert was not used for Research Questions 1.2 or 1.3. A false alert was defined as anytime an initial or full alert was presented to the driver when either the DFM was not working properly (i.e., not tracking the driver's eyes) or the increase in the PERCLOS value was not due to a slow eye-closure (e.g., driver checking the mirrors). Therefore, an alert was considered to be valid if the alert was caused by slow eye-closures. If the alert was caused by anything other than a slow eye-closure (e.g., head turning to look at mirror, looking down for logbook), the alert was consider and invalid alert. To determine which activities caused invalid DFM alerts, the data reductionists examined the increase in PERCLOS values leading to an alert by watching the video and parametric data associated with that alert. If it was determined that the increased PERCLOS values were generally caused by eye-closures, the alert was marked as valid. In determining whether the increase in PERCLOS was caused by eye-closures or by the DFM's limitations to find the driver's eyes, the Number of Eyes Found variable was used. The DFM assumed that if no eyes were found, the driver was closing his or her eyes. If no eyes were found while the driver was looking forward and his or her eyes were open, it was assumed that the DFM was not locating the driver's eyes correctly. For more detail on the alert validation procedure please see Appendix A.

DFM Alert Validity Quality Control

A quality control examination of the data reductionists' assessment of DFM alert validity was performed. Two quality assurance (QA) inspectors reviewed validated alerts to verify the reductionists' judgment of DFM alert validity. The protocol describing how this was accomplished is presented in Appendix B. The QA inspectors reviewed the first 100 alerts assessed by each reductionist. If a discrepancy arose between the reductionists' and QA inspector's assessment of alert validity, then a final reviewer inspected the alert in question. The final reviewer was experienced with manual PERCLOS calculations. Thus, the final reviewer's judgment served as a mediator in the validity assessment process. The QA inspector then discussed the reductionists' assessment with the reductionist in order to provide feedback and improve reductionist performance. Once this review was performed, the reductionist continued reviewing the generated DFM alerts to assess their validity.

An inspection of reductionist performance relative to the final reviewer's assessment (considered to be the correct assessment) revealed that reductionist judgment of alert validity were highly conservative. In other words, when a reductionist judged an alert to be invalid, they were in agreement with the final reviewer 98 percent of the time. However, when a reductionist determined an alert was valid, they were in agreement with the final reviewer only 52 percent of the time. This decision threshold is illustrated in Figure 12, where a "Hit" occurs when the reductionist and final reviewer agree an alert is valid, a "Miss" occurs when a reductionist judges an alert to be invalid and the final reviewer disagrees, a "Correct Rejection (CR)" occurs when both the reductionist and final reviewer judge an alert as invalid, and a "False Alarm (FA)" occurs when the reductionist judges an alert to be valid when the reviewer disagrees. These results compelled the decision to have the QA inspectors review every alert that the reductionists judged to be valid. Therefore, the valid alerts used in the analyses, as presented in this report, were reviewed by a minimum of two reviewers.

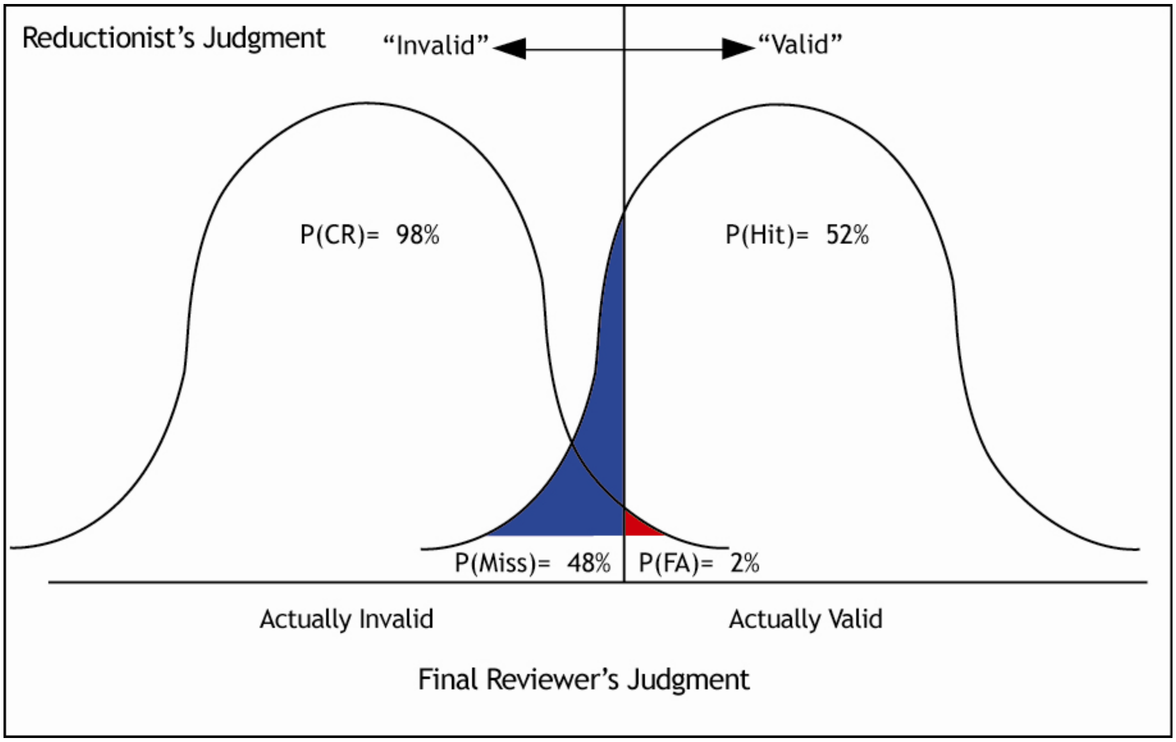


Figure 12. Probability of Reductionists Correctly Judging DFM Alerts as Valid

The validation procedure revealed that 90 out of the 96 drivers received alerts during the episode of interest. Six participants did not receive any alerts during the episode of interest due to DFM malfunction. A total of 74,638 alerts were present in the episodes of interest for the 90 drivers who received alerts. From all alerts present, a total of 15,386 individual alerts were evaluated (i.e., a sample of 20.61 percent of all alerts) and 721 of those were deemed valid alerts (i.e., 4.69 percent of sampled alerts were deemed valid alerts). Of the 90 drivers who received alerts, 31 drivers (34.44%) did not receive a single valid alert (i.e., the validity rate for alerts presented to these drivers was zero). The other 59 drivers received at least one valid alert. Figure 13 shows that for most of the drivers, the validity rate was at or below 5 percent, with a range of 0 to 41.7 percent, a median of 1.3 percent, and an overall validity rate of 4.69 percent. The drowsiness level of episodes without alerts was not examined. Therefore, no conclusions about the sensitivity or specificity of the system can be made in this study. It should be noted that although manual calculation of PERCLOS was not performed in this study, during the data reduction process the researchers evaluated if the automatic calculation seemed to match the eye-closures of the driver.

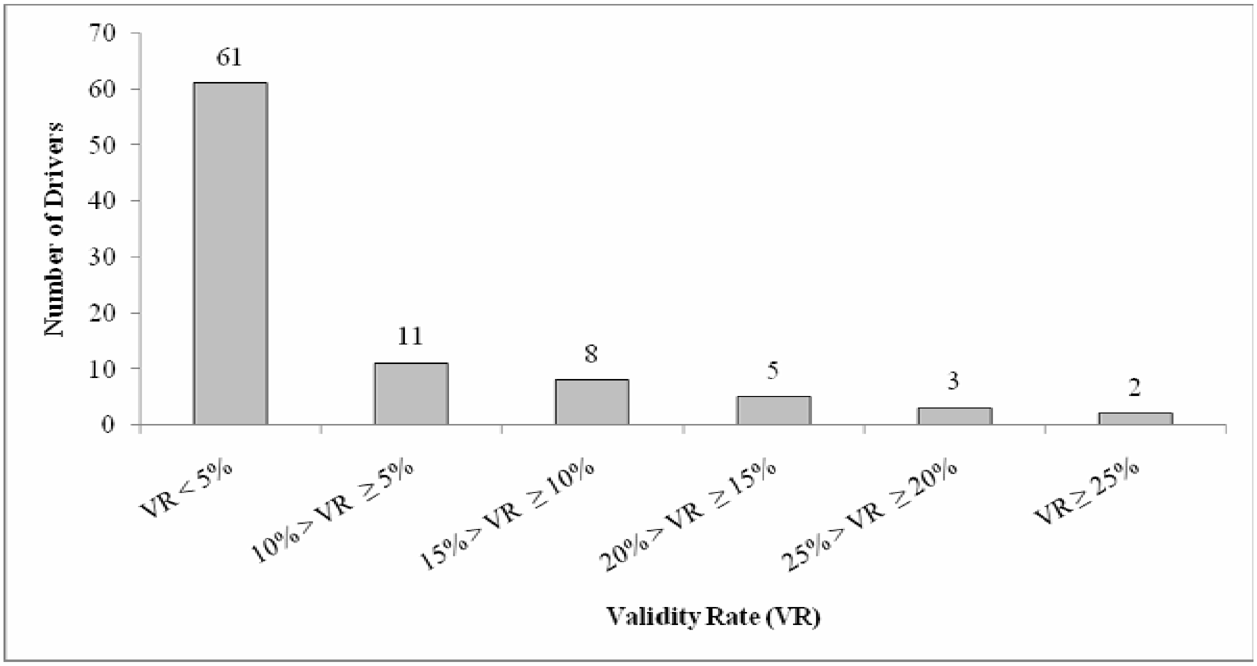


Figure 13. Validity Rate (VR) Frequency

Post-Alert Behavior Data Reduction

The data to investigate the effect of DFM alerts on driver behavior was collected using a process of manual data reduction. A team of 10 data reductionists were instructed to review 5 min of video footage after each of the 721 valid alerts. This video footage was reviewed using VTTI's Data Analysis and Reduction Tool. DART allows video footage to be advanced on a frame-by-frame basis, yielding a precise measurement of the occurrence of a behavior. The video footage for this study was recorded at approximately 30 Hz. Reductionists were instructed to look for behaviors that were anticipated to occur following the onset of drowsiness. Table 8 lists these behaviors as well as their operational definitions. Reductionists were also asked to record unanticipated behaviors if they were relevant to the driver's level of drowsiness.

Table 8. Post-Alert Behaviors

Post-Alert Behavior	Frame Where Record Sync Begins
Use Cell Phone	Cell phone first appears in the video image
Use CB Radio	Driver picks up the CB radio
Drink	Driver puts drink to mouth.
Eat	Driver puts food to mouth
Smoke/Chew Tobacco	Driver touches the cigarette or chewing tobacco to his/her mouth
Take Pill/Medication	Driver puts the medication in his/her mouth
Raise/Lower Window	Window begins to move
Reach to Front	Driver begins to move his/her arm
Reach to Side/Reach Up	Driver begins to move his/her arm
Reach Back	Driver begins to move his/her arm
Veer Off Road	Truck unintentionally leaves the lane, either to the left or right side (paint markings disappear under the truck as would occur in a lane change)
Talk/Sing/Laugh	Driver begins to move his\her mouth
Rub/Scratch Eyes	Driver touches his\her eyes
Excessive Blinking	Driver begins to blink at the start of the excessive blinking behavior
Look to Side/Up/Down	Driver's eyes/head initiate their fixation (when the eyes land on where they are looking) to the side
Raise/Lower Eyebrows	Driver begins to raise or lower his\her eyebrows
Scratch/Straighten Hair	Driver begins to touch his/her hair
Rub/Scratch/Hold Face	Driver begins to touch his/her face
Rub/Scratch/ Hold Neck	Driver begins to touch his/her neck
Yawn	Driver begins to open his/her mouth
Adjust Head Position	Driver begins to move his/her head in the adjustment
Shake Head	Driver begins to move his/her head when shaking it
Adjust Body Position	Driver begins to move his/her body in the adjustment
Stretch	Driver begins to move his/her body when stretching
Stop and Rest	Truck comes to a complete stop (based on the forward roadway video image) and driver remains in seat to rest
Stop and Leave	Truck comes to a complete stop (based on the forward roadway video image) and driver leaves the driver's seat
Restart	Truck begins to move after the driver has taken a break (stops for reasons other than rest, such as at a red light, do not count)
Right Lane Change	Truck crosses the lane markings when making a right lane change. Only the first right lane change should be marked
Left Lane Change	Truck crosses the lane markings when making a left lane change. Only the first left lane change should be marked
Press Brake Pedal	Brake pedal is pressed (obtained via vehicle network). Only the first brake pedal press should be marked
Exit Highway	If on a highway, the frame where the truck enters the exit ramp (based on the forward roadway image)
Enter Highway	If entering a highway, the frame where the truck enters the entrance ramp (based on the forward roadway image)
Turn (Left or Right)	Truck initiates the right or left turn (based on the forward roadway camera view). Only the first turn should be marked
Other	Any other behavior of note observed (also recorded in log along with the sync number)

Reductionists also recorded the environmental conditions present during the valid alerts. This included weather (e.g., no adverse weather, rain, sleet, snow, fog), roadway surface conditions (e.g., dry, wet, icy), traffic-way flow (e.g., road not divided, divided highway, one-way traffic), number of travel lanes, and traffic density (e.g., free flow, flow with restrictions, unstable flow, forced traffic flow).

A quality control examination of the reductionists' post-alert behavior reduction was performed. The first 10 alerts identified by each reductionist were inspected by a QA inspector who provided feedback to the reductionist. Since each alert was reviewed by a single reductionist, the QA inspector inspected 10 percent of the alerts generated for each driver by taking a random sample of the alerts for each week in the study. These alerts were then reviewed in accordance with the alert validation and quality assurance procedures (Appendix D).

Manual Versus DFM PERCLOS Calculations

Due to the low alert validity rate obtained during the alert validation process, an analysis using previously calculated manual PERCLOS values was performed as a secondary assessment. A total of 1,433 baseline events were evaluated in terms of eye-closure under a study funded by the National Surface Transportation Safety Center for Excellence. Manual PERCLOS was calculated for each one these baseline events (Wiegand, Hanowski, Olson, & Melvin, 2008). For the current comparison, the PERCLOS-3 values from the DFM were compared for baseline events to the PERCLOS-3 values manually calculated under the NSTSCE study. The manual PERCLOS-3 values were calculated by examining the preceding 3 min and 10 s of video data for each NSTSCE baseline epoch. Reductionist would then code the estimated manual PERCLOS sync by sync (each video frame) as eyes opened, eyes closed, or eyes not visible. Any rapid eye blinks were eliminated from the manual PERCLOS calculation; a rapid eye blink was any eye-closure that was $1/10^{th}$ of a second in length. If an event had more than 20 percent of the video coded as "eyes not found," the manual PERCLOS calculations were not preformed. For the current comparison, the PERCLOS-3 values from the DFM for each NSTSCE baseline event were compared to the PERCLOS-3 values manually calculated under the NSTSCE study.

A total of 294 baselines occurred within the episode of interest. The absolute difference between DFM and manual PERCLOS measures was calculated for each of the baseline events that occurred within the episode of interest. The average absolute difference was 7.10. This value was evaluated using a t-test (signed rank) and found to be statistically different from zero, $t(293) = 21682.5$, $p < 0.0001$. In addition, the Pearson correlation coefficient between the manual PERCLOS and the DFM PERCLOS was $r = 0.09$. This indicates that there is no linear correlation between the PERCLOS value manually calculated and the estimate calculated by the DFM. Therefore, the results presented in Research Question 1.1, which examined PERCLOS values across different factors of interest, showed both DFM PERCLOS and the manual PERCLOS values collected from the NSTSCE study. However, the reader should be cautious about comparing the results between these two values as each set of results should be considered to stand alone.

Analysis

In addition to descriptive statistics, several other analyses were conducted for this research question. Multi-level modeling was used to evaluate behavior changes over time. Specific applications of MLM will be discussed as they occur in the following sections. PERCLOS values were evaluated in order to evaluate changes in the level of drowsiness. For changes in post-alert behavior, the time it took the driver to stop driving after an alert onset was evaluated. In addition, a Poisson regression was used to analyze the number of valid alerts over time. A brief description of MLM and Poisson regression is presented next.

Multi-Level Modeling. Longitudinal data is the designation for data collected for the same participant over time. If changes over time are of interest, MLM is suggested for analysis. Several of the research questions in the safety benefit analysis are directed at evaluating behavior change over time, and the data were collected longitudinally. Therefore, MLM was used for the analysis of Research Questions 1.1, 2.1, and 2.2 and is discussed with its specific applications in the appropriate sections.

MLM is also known as mixed models, hierarchical linear models, and individual growth models. As a technique commonly used to investigate change, MLM was deemed appropriate for the present analysis. MLM is composed of several levels. The first-level model describes within-individual change over time, and the second-level model relates predictors to any between-individual difference in change (Singer & Willett, 2003). Figure 14 presents the first level of the MLM as presented by Singer (1998). The response variable (in this case PERCLOS) is a function of TIME at the i^{th} time period for the j^{th} participant. Similar to a linear regression, r_{ij} is the residual term of the i^{th} time period for the j^{th} participant and is assumed to be normally distributed with a mean of zero and a variance of σ^2.

$$Y_{ij} = \pi_{0j} + \pi_{1j}\left(TIME\right)_{ij} + r_{ij} \quad \text{where} \quad r_{ij} \sim N\left(0, \sigma^2\right)$$

Figure 14. First-Level Model

For the second level, the MLM departs from linear regression. Figure 15 presents the intercept term and the time effect from the first-level model (Singer, 1998). They are modeled as the response to a predictor variable. This assumes observations from the same participant are correlated. In linear regression techniques, this would violate the assumption of independence of observations; however, MLM assumes a correlation structure in the second level. Therefore, MLM allows for an accurate estimation of treatment effects and the differences that may exist between them.

$$\begin{aligned}\pi_{0j} &= \beta_{00} + \beta_{01}\left(Test - condition\right)_j + u_{0j} \\ \pi_{1j} &= \beta_{10} + \beta_{11}\left(Test - condition\right)_j + u_{1j}\end{aligned} \quad \text{where} \quad \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}\right]$$

Figure 15. Second-Level Model

Poisson Regression. The response variable for Research Question 1.2 was the count of valid alerts per week for each driver. Count data tends to follow a Poisson distribution (a discrete probability distribution representing the probability of a number of events occurring during a time period when events occur independently of one another with a known average rate; see Cameron & Trivedi, 1986; Lee, 1986). Therefore a Poisson regression model was developed to address this question. Parameters for the Poisson regression were estimated using generalized estimating equations, which assumed that observations from the same driver were correlated, and observations between drivers were independent.

Measures

Due to the magnitude of operations on even a single factor within the original DDWS FOT database, it was not feasible to manipulate the full database each time an analysis was needed. Therefore data sets were developed with the specific measurements of interest for each research question.

Data set for Research Question 1.1. The main focus of Research Question 1.1 was the evaluation of PERCLOS values across the different experimental conditions. Therefore, this data set includes the average DFM PERCLOS (using PERCLOS-3) during the episodes of interest for each day in the study, per each participant. Each driver has a DFM PERCLOS value for each day he or she participated in the study that the DFM was operational (i.e., non-standby mode). A total of 4,151 days are represented in this data set for the 95 drivers who are part of this analysis (one of the 96 drivers did not have any episodes of interest). On average, each driver had 44 days during which the DFM was operational and within the episode of interest.

In addition, this research question was supplemented using the PERCLOS-3 values calculated from NSTSCE baseline events. The primary difference between DFM PERCLOS values and the NSTSCE calculated values is that the DFM PERCLOS value is automatically calculated by the DFM, while the NSTSCE value was based on a frame-by-frame manual video data reduction performed as part of that research effort (Wiegand, Hanowski, Olson, & Melvin, 2008). Also the NSTSCE baseline events represent a single epoch, while the DFM PERCLOS values were averaged for each day when the DFM was operational and within the episode of interest.

The NSTSCE baseline events were randomly selected epochs. The number of baselines for a given driver was dependent on the weeks of participation. For each week of participation, there were 1.5 baseline events selected. These events were then reviewed by a data reductionist to confirm that the driver was in the vehicle and that the truck was in motion. A total of 1,433 baseline events were used for this supplemental analysis.

Data set for Research Question 1.2. Only the 721 valid alerts were part of this data set. In addition to the number of valid alerts per week for each driver, the length of the participation (total weeks of participation) was included as well. All the data were divided by experimental condition.

Data set for Research Question 1.3. Similar to the data set for RQ 1.2, this data set only considered valid alerts. A post-alert period of 5 min was analyzed to identify behaviors after a valid alert (Appendix C). The first occurrence of each behavior in that 5-min interval was recorded as part of this data set. A second data set was created for this research question to examine the post-alert stopping behavior of the drivers; the elapsed time from the point a driver received a valid alert until the driver took a break was calculated. The data were divided by experimental condition and included the following specific items:

- Valid Alert Level: the sensitivity setting selected by the driver when a valid DFM alert was generated.
- Environmental Conditions: the environmental conditions present when valid DFM alerts were generated. This includes the time of day, weather, roadway surface condition, traffic-way flow, number of travel lanes, and traffic density.
- Driver Characteristics: the characteristics of the drivers who received valid DFM alerts. This includes the driver's age, commercial driving experience, and truck operation type (i.e., line-haul or long-haul).
- Post-Alert Behavior: the specific behavior that was exhibited by drivers within 5 min of receiving a valid DFM alert.
- Elapsed Time from Alert to Observed Behavior: the amount of time in minutes that passed from a generated DFM alert to the occurrence of the behavior.

RQ 1.1: Does the DDWS result in fewer episodes of drowsy driving?

Results

In order to assess if the DDWS resulted in fewer episodes of drowsy driving, several measurements were evaluated. DFM PERCLOS was the main measure of interest for this question. If the DDWS had a positive effect and helped drivers identify when they were drowsy, the Test Group should have lower DFM PERCLOS values than the Control Group. A representative PERCLOS value was obtained for each day the driver participated in the study in order to perform this assessment. In addition, a supplemental analysis was performed using the 1,433 NSTSCE baseline events. These baseline events were used to evaluate drowsiness at random instances from each week of the study. The last portion of this results section looks at an MLM that compares DFM PERCLOS at the different experimental conditions as well as the Manual PERCLOS.

DFM PERCLOS. The following are descriptive statistics looking at the DFM PERCLOS value obtained for each day of participation. The data only include the daily average DFM PERCLOS values within the episodes of interest for the study.

Figure 16 presents an overview of the DFM PERCLOS values that characterize each of the experimental conditions. The minimum DFM PERCLOS value for all four conditions was zero. The maximum ranged from 67 to 96; baseline conditions for both were in the lower 70s. Recall that these values may be affected by scanning behaviors (i.e., eyes off of forward roadway). Therefore, having a high DFM PERCLOS value during a given day does not mean that the driver had his or her eyes closed for almost 3 min (PERCLOS-3) but that during a moving window of 3 min, the driver may have been drowsy, and the system was detecting that, or the measure was invalid, due to mirror checking, for example. However, this is the best measure of drowsiness that we have to examine the impact of the DFM on driver drowsiness over time. The central tendency measures give a more realistic perspective in this case; the mean and median DFM PERCLOS values range from 4 to 8, and the test condition is the lower one for both. For this general trend overview, only the mean values will be used from this point on. Due to the high number of alerts (approximately 75,000) presented during the episodes of interest, it was not feasible to visually validate all alerts. However, as part of this research effort a sampling was done to identify valid alerts for Research Questions 1.2 and 1.3, using the validation method described previously.

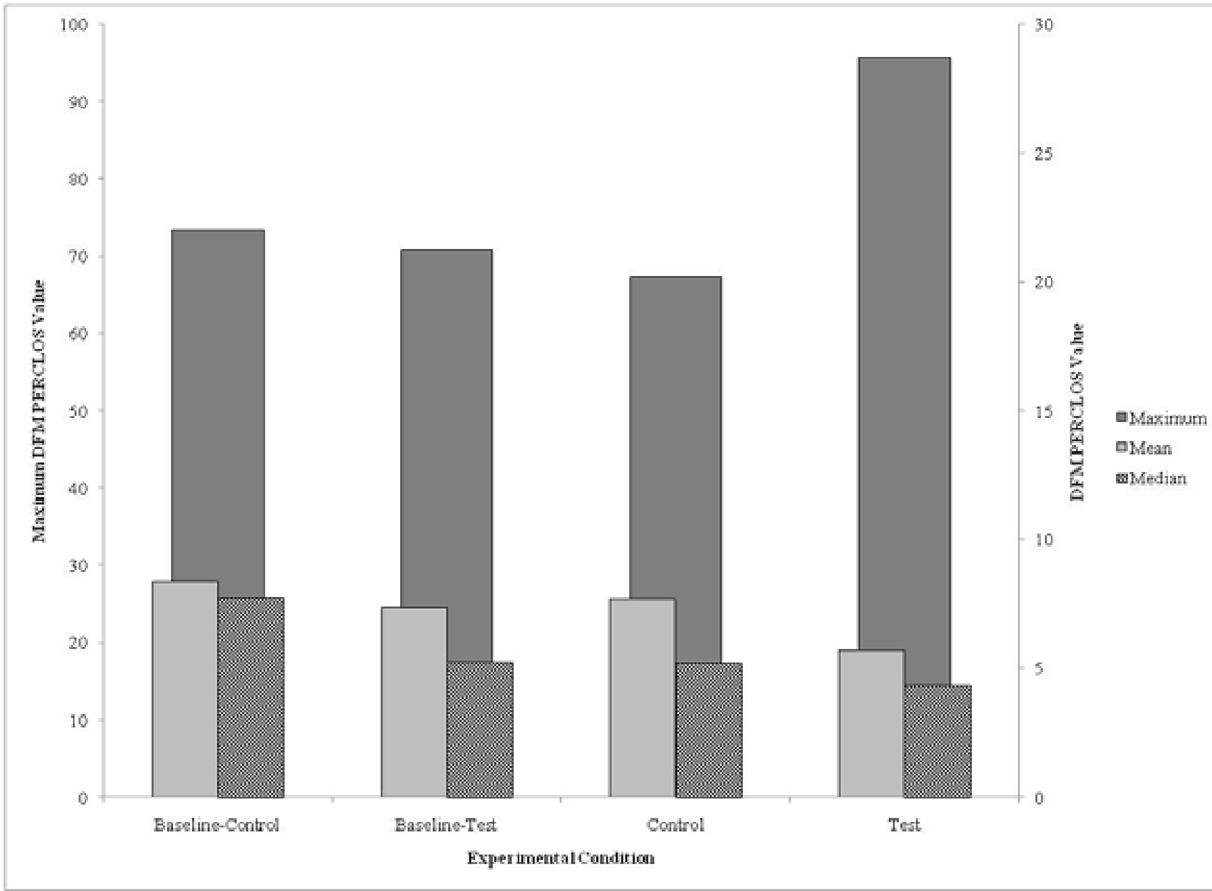


Figure 16. DFM PERCLOS Values by Experimental Condition

Evaluating mean DFM PERCLOS by taking into consideration the days drivers were in the study showed no trend other than the test condition having the lowest DFM PERCLOS values throughout the days in the study (Figure 17) no matter which day of the week the value was obtained (Figure 18). The 2 data points in Figure 17 that represent the baseline test condition for 70-79 days and 80-89 days are both composed of 7 observations each. These observations are all from the same participant (Driver 136). This driver never obtained a valid alert in the sampling process performed during alert validation. Although the DFM PERCLOS values for this driver appear to be unusually high most of the time, the DAS only collected the DFM output (not potential internal failures messages for the DFM software and hardware). Therefore, it is possible that these two data points represent potential outliers. However, there is no internal DFM data to confirm failure and eliminate them, and other parameters do not show this driver as a potential outlier. For this same figure, the control condition at the 120-129 days data point, which is close to zero, represents just one observation and was confirmed to be accurate. All other data points represent multiple observations.

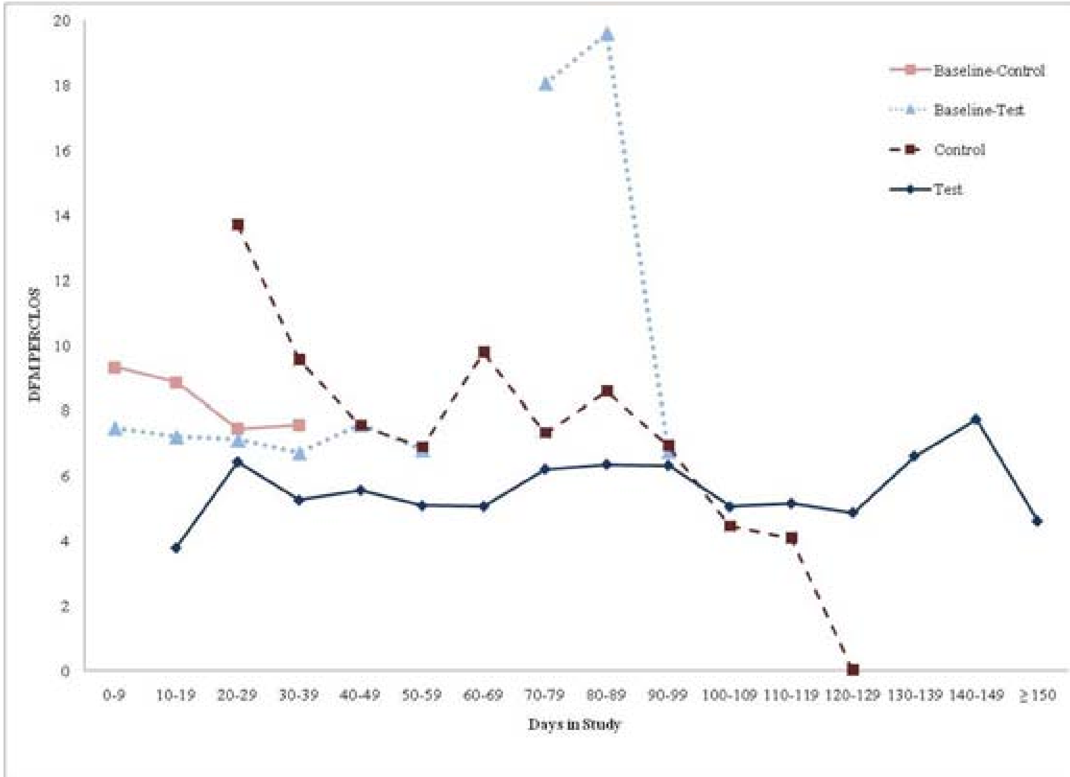


Figure 17. DFM PERCLOS Values by Day in Study

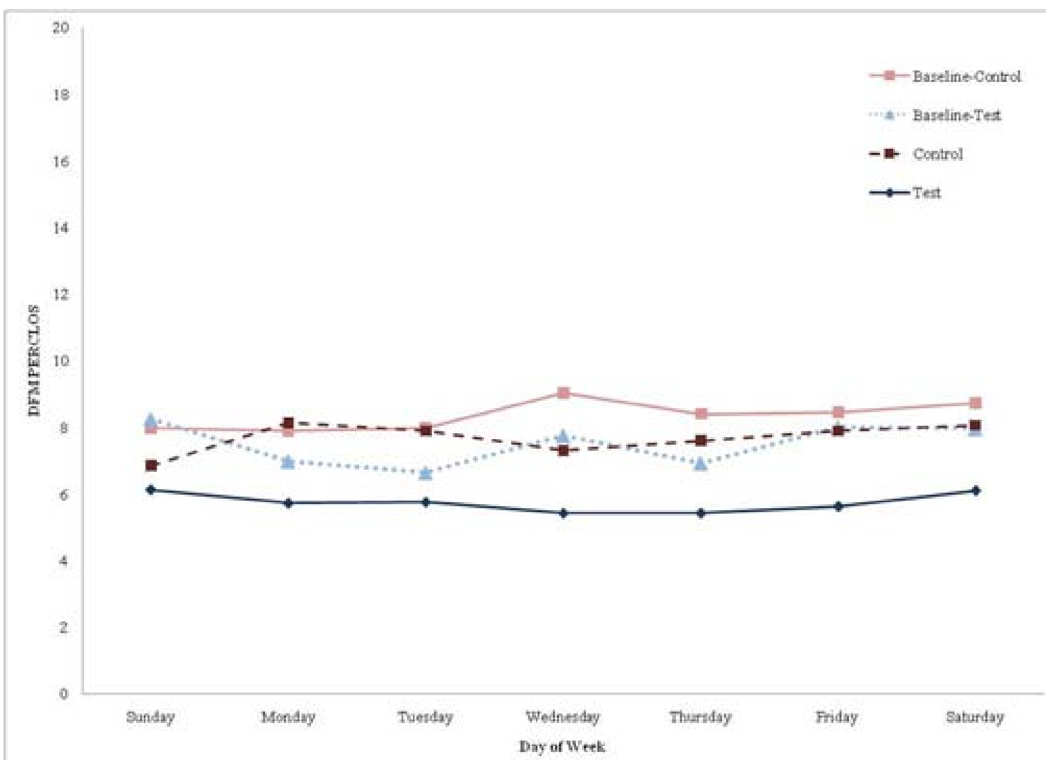


Figure 18. DFM PERCLOS Values by Day of Week

A time progression perspective is obtained by evaluating the DFM PERCLOS values by weeks of particiaption in the study (Figure 19) and by driving hours during the participation in the study (Figure 20). Figure 19 shows that in general the test condition presented a lower DFM PERCLOS value for most of the weeks in the study, with the baseline test condition as well as baseline control and control conditions showing a mean value over 4 and as high as 16. Most data points are composed of several hundred values; however, the last two datapoints for the baseline test condition in Figure 19 appear to represent potential outliers. These two values for weeks 8-11 and 12-15 reprsent 8 and 13 observations, respectively. Similar to the DFM PERCLOS results by day in study, a single participant (Driver 136) is driving these two points. This driver's data represents the majority of observations forming those data points. Similar to the previous discussion, there was not enough evidence to eliminate these observations.

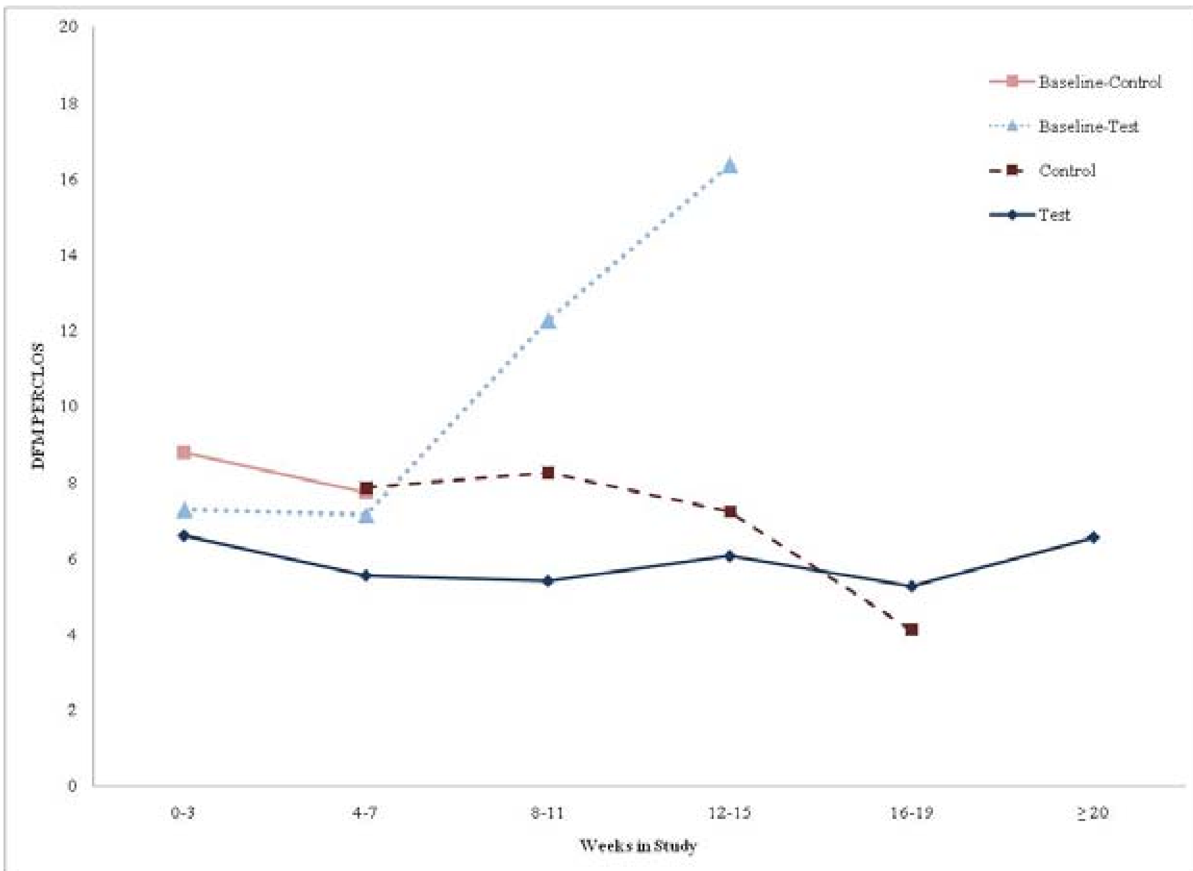


Figure 19. DFM PERCLOS Values by Week in Study

Figure 20 presents the DFM PERCLOS values as the driving hours progress throughout the study. This also presents a very similar trend to the DFM PERCLOS values by week, where for the majority of the time the test condition had a lower mean DFM PERCLOS value. At a few points in time, the control condition displayed a higher DFM PERCLOS; however for most of the instances this was not the norm. The one data point for baseline test that goes to zero represents only two observations for a given driver.

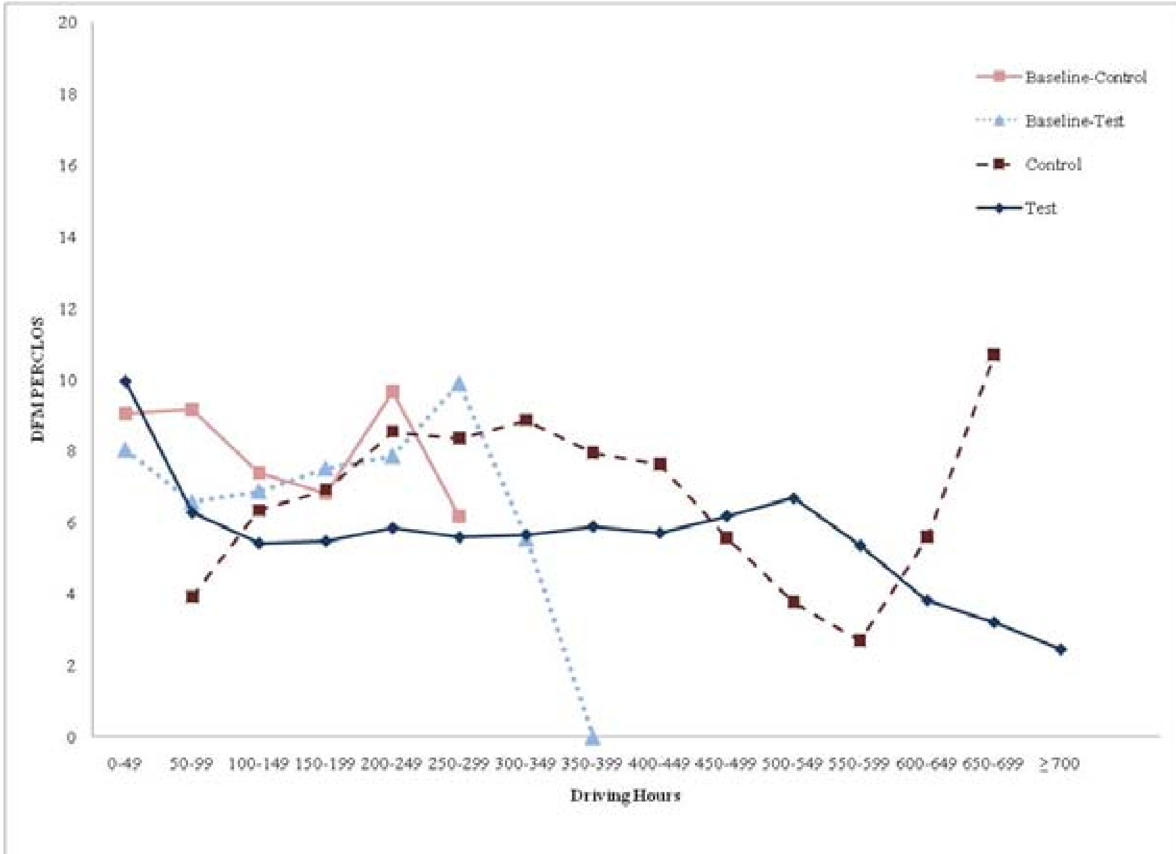


Figure 20. DFM PERCLOS Values by Driving Hours in the Study

Manual PERCLOS – Baseline Events. The DFM PERCLOS values presented above only examined episodes of interest, which primarily occurred during low illuminance conditions (i.e., night, dawn, dusk). The baseline events represent a sample of random points in time during the driver's participation in the study. Baseline data presents a global look at the driver participation, including high illuminance conditions, lower driving speeds, and other aspects that were used as part of the criteria to eliminate portions of the data for the DFM PERCLOS analysis due to DFM operation envelope limitations.

Figure 21 presents an overview of the manual PERCLOS values that characterize each experimental condition. The minimum manual PERCLOS value for the Test Group was zero, and the minimum value was less than 0.5 for the Control Group. The maximum manual PERCLOS value ranged from 22 to 31; the baseline control was the lower of the conditions, and all the other conditions were fairly similar. Note that these values are calculated by visually evaluating eye-closure based on the video available for the baseline event. These results suggest a more similar group of drivers and drowsiness level during regular operations. The mean manual PERCLOS value was approximately 6 for all experimental conditions, and the median ranged from 4 to 6.

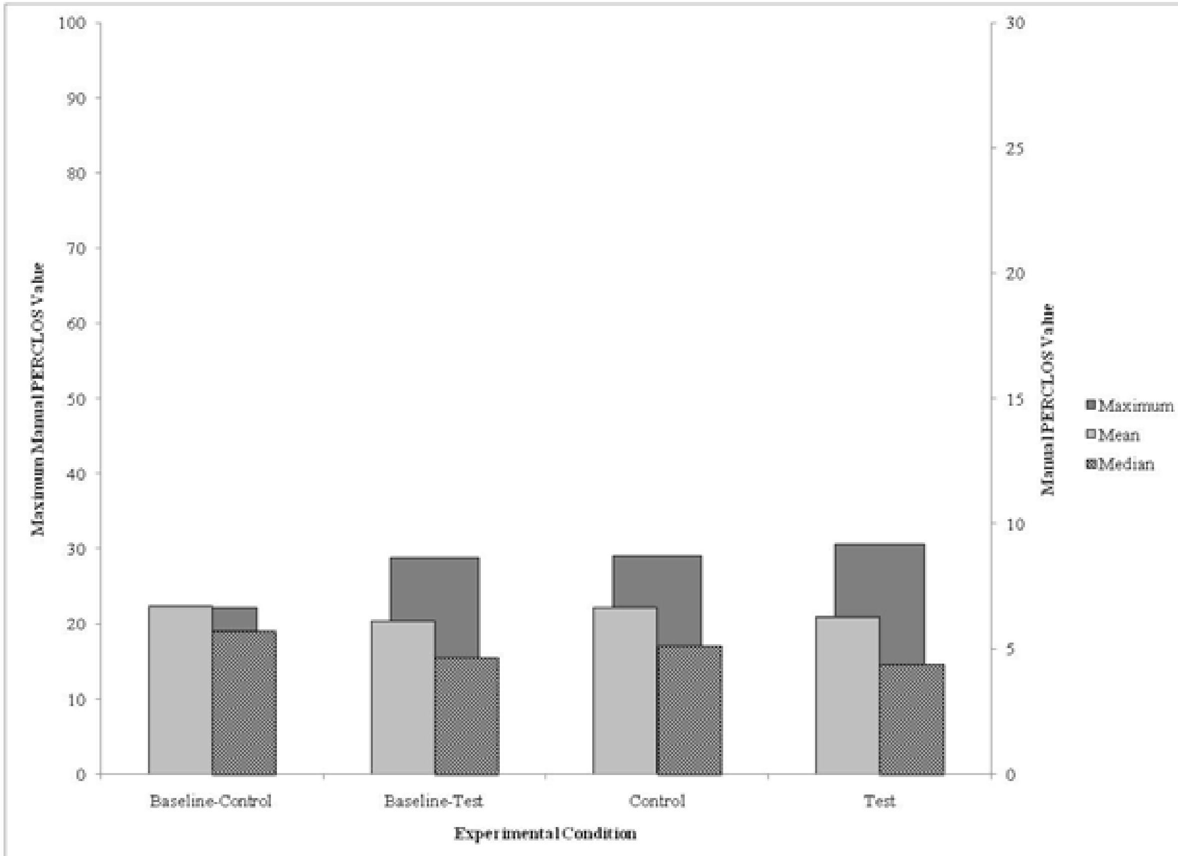


Figure 21. Manual PERCLOS Values by Experimental Condition

The mean manual PERCLOS evaluation by days the drivers were participating in the study (Figure 22) or the day of the week the value was obtained (Figure 23) shows no trend. As performed within the DFM PERCLOS data, a time progression assessment was performed by evaluating the manual PERCLOS values by weeks of particiaption in the study (Figure 24) and the driving hours during the participation in the study (Figure 25). Manual PERCLOS for the test condition was lower than the ones for the control condition from week eight onward. This trend was not found when the accumulated hours driven were examined (Figure 25). For this analysis the control condition and the test condition manual PERCLOS values varied depending on the week the driver was currently driving without any pattern of one of them being the lowest for an extended period of time.

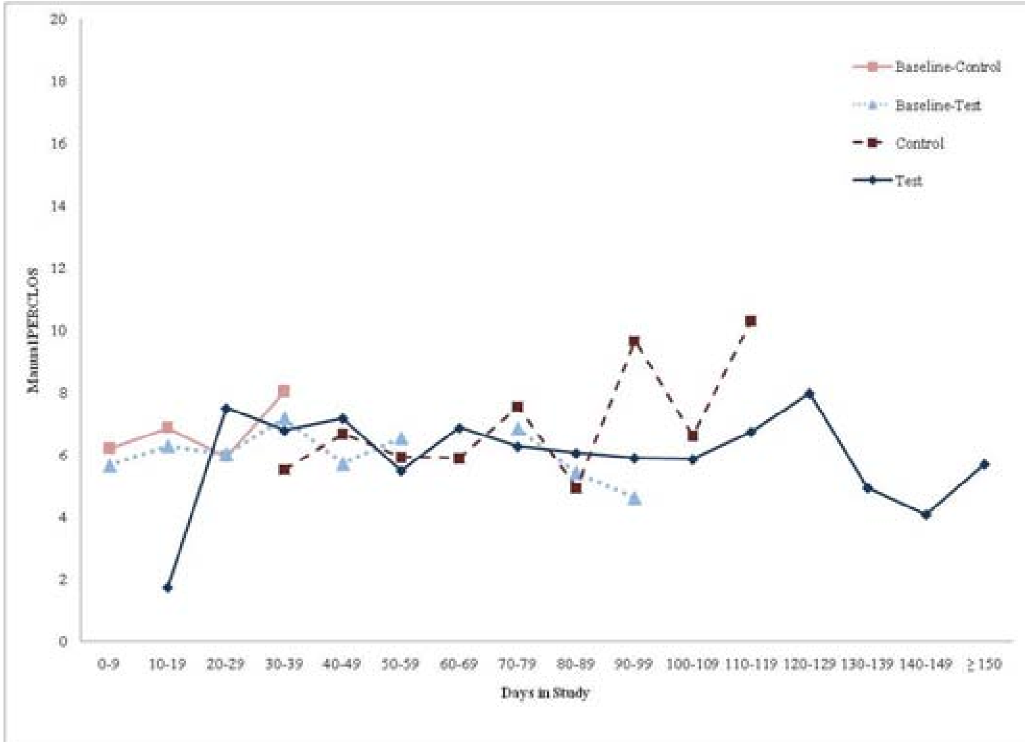


Figure 22. Manual PERCLOS Values by Day in Study

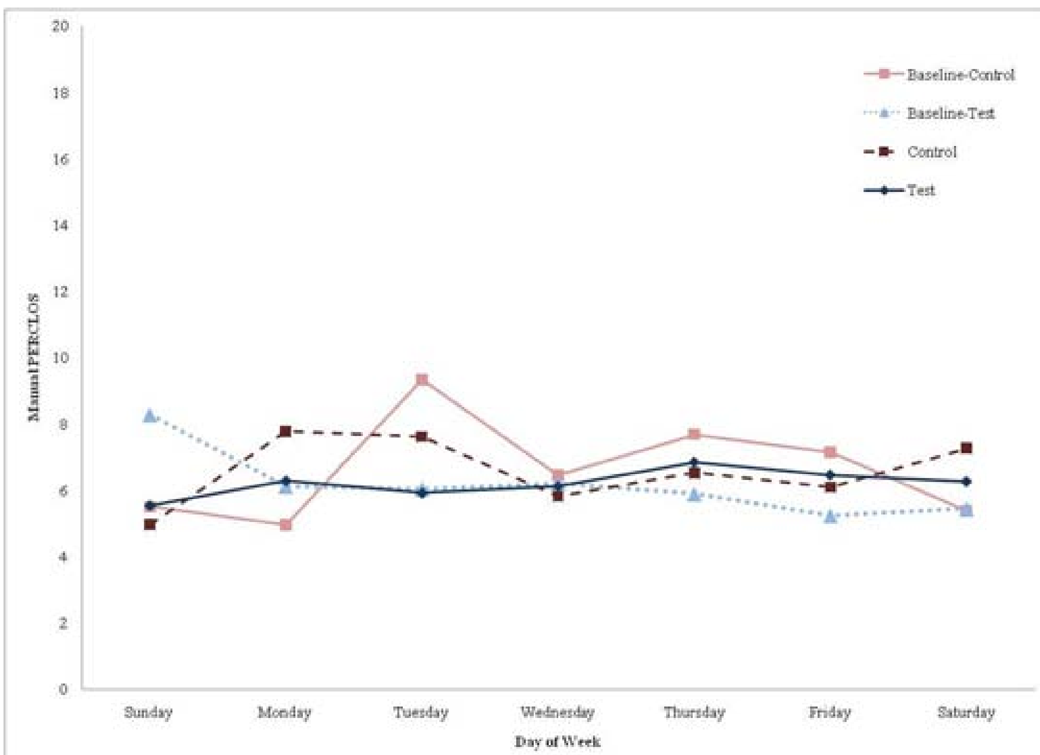


Figure 23. Manual PERCLOS Values by Day of Week

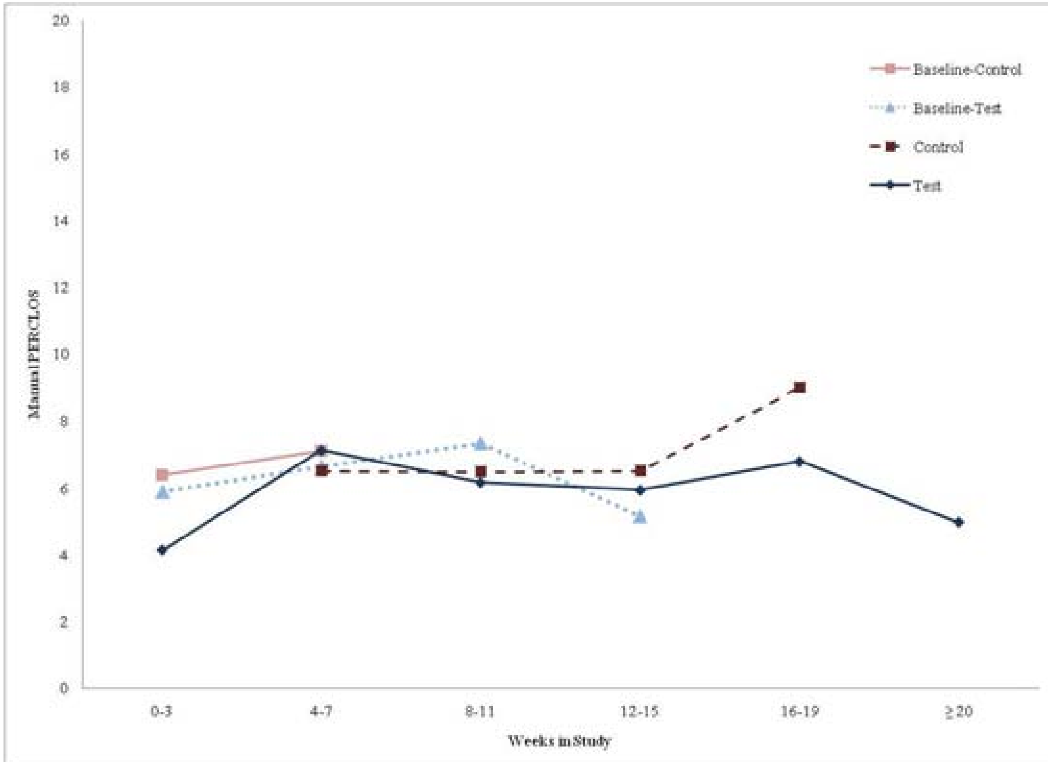


Figure 24. Manual PERCLOS Values by Week in Study

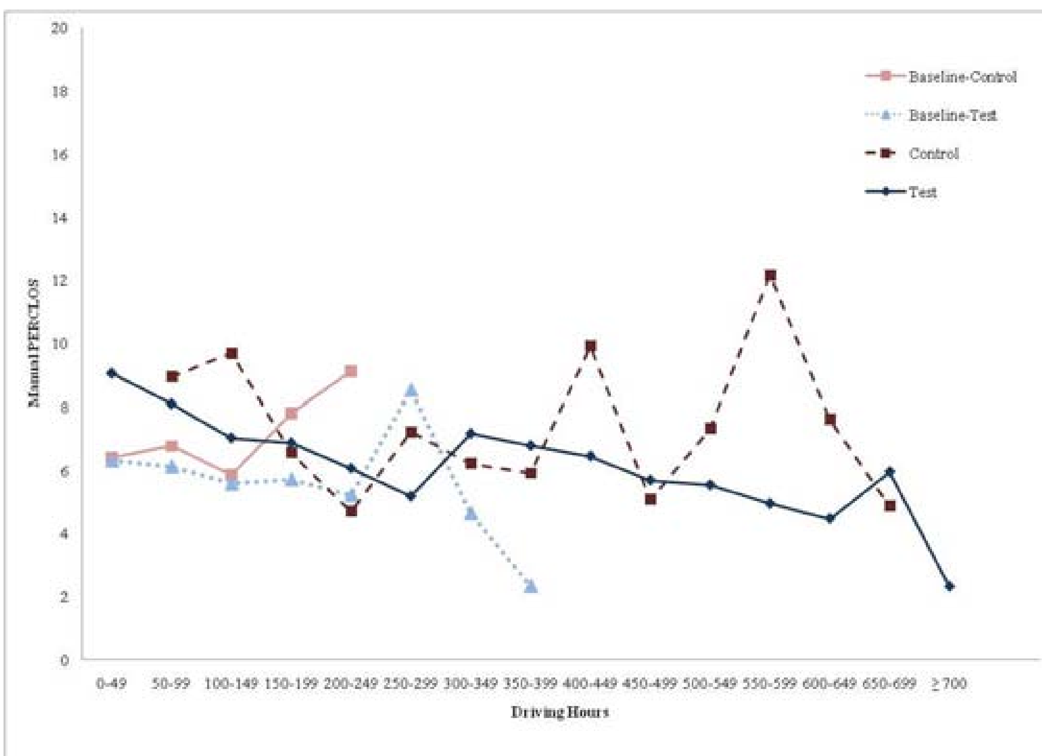


Figure 25. Manual PERCLOS Values by Driving Hours in the Study

Multi-Level Modeling. Two MLM analyses were performed, one taking into consideration DFM PERCLOS for each experimental condition, and the second using manual PERCLOS. The first set of results suggests that, over time, DFM PERCLOS is lower for the test condition (i.e., when the DDWS is providing feedback to the driver) than during the other conditions when the system is not providing feedback to the driver (Figure 26).

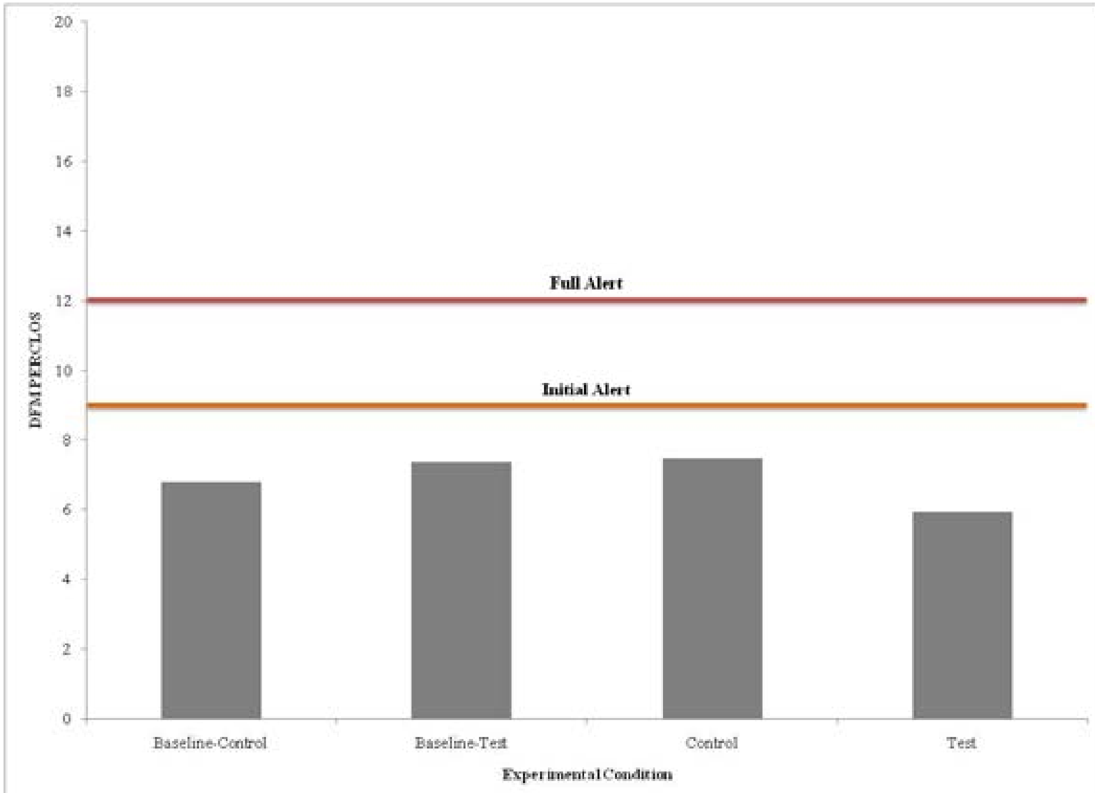


Figure 26. Estimated DFM PERCLOS Value by Experimental Condition Based on MLM

Table 9 presents a summary table with the comparisons of interest. There are no statistical differences between baselines for the Test and Control Groups. This suggests both groups would have similar PERCLOS levels if they did not have feedback about their drowsiness level. Additionally, the DFM PERCLOS values were not statistically different between the two artificial conditions (baseline control and control conditions) used for the Control Group. This split between baseline control and Control was performed for analysis purposes only, as drivers in the Control Group did not experience any alerts during their participation. The difference in PERCLOS values between Test and baseline test conditions was statistically significant ($p = 0.0077$), with the DFM PERCLOS in the test condition being 1.44 lower than during the baseline Test period. However, none of the estimated values reached the level where an initial alert would be needed. It should also be noted that number of days in the study was not a significant factor, and the interaction between the number of days in the study and the experimental condition was not significant for any level of the experimental condition.

Table 9. MLM Results Using DFM PERCLOS Values

Condition	Condition	Estimate	SE	df	*t* Value	Pr > |*t*|
Baseline Control	Baseline Test	-0.59	1.46	82	-0.4	0.6881
Baseline Control	Control	-0.70	0.93	82	-0.75	0.4584
Baseline Test	Test	1.44	0.53	82	2.73	0.0077
Control	Test	1.54	1.11	82	1.39	0.1688

The second set of results is based on manual PERCLOS. Recall that these values were calculated for random points in time throughout each week of participation in the study for each participant. The DFM's operation had illumination and speed restrictions and, therefore, the results based on manual PERCLOS present another perspective to the drowsiness evaluation. Figure 27 shows the estimated values for manual PERCLOS, and Table 10 presents a summary of the comparisons of interest. The focus of this analysis is the Test Group. The results suggest that there are no significant differences between the different experimental conditions in this model. Based on PERCLOS estimates for this model, drivers in the Test Group did not exhibit a change in the level of drowsiness throughout the study. The Control Group seems to present a higher PERCLOS level at the beginning of the study, but the change was not big enough to make this difference statistically significant when using an alpha level of 0.05. As before, the number of days in the study was not a significant factor, and the interaction between the number of days in the study and the experimental condition was not significant for any level of the experimental condition. Both control and test condition estimates were very similar (Table 10).

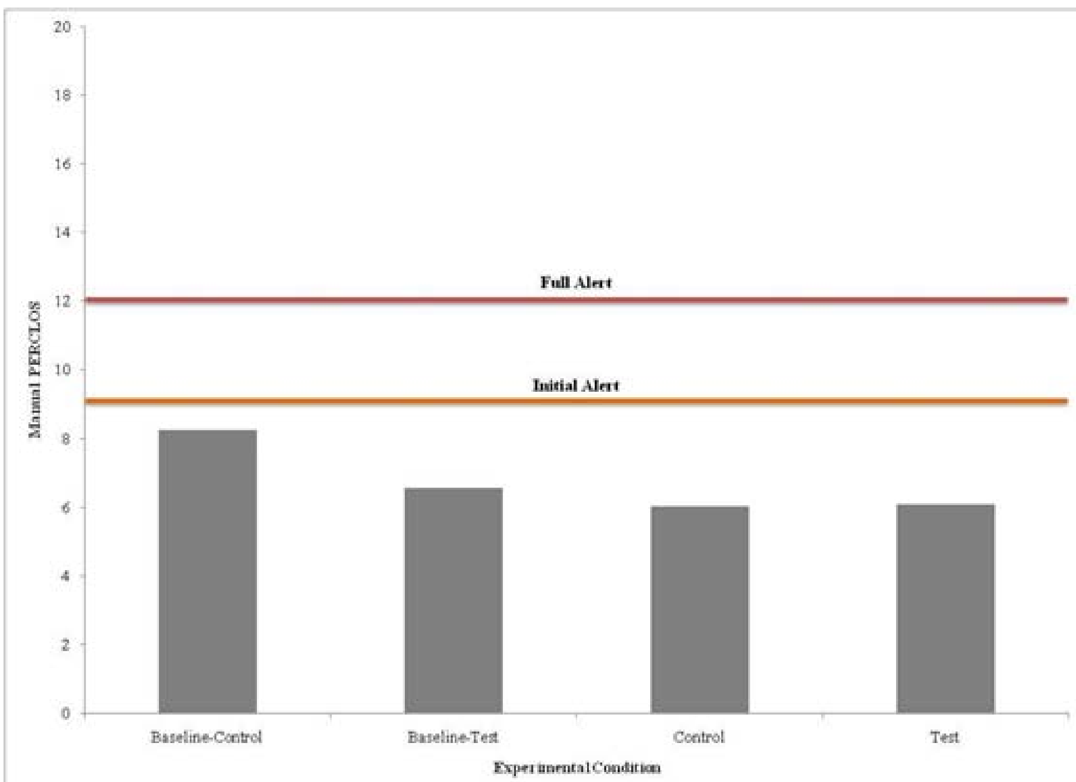


Figure 27. Estimated Manual PERCLOS Values by Experimental Condition Based on MLM

Table 10. MLM Results Using Manual PERCLOS Values

Condition	Condition	Estimate	SE	df	*t* Value	Pr > |*t*|
Baseline Control	Baseline Test	1.71	1.54	79	1.11	0.2705
Baseline Control	Control	2.22	1.18	79	1.89	0.0627
Baseline Test	Test	0.47	0.54	79	0.87	0.3873
Control	Test	-0.04	1.07	79	-0.04	0.9692

RQ 1.2: Does the frequency of DDWS alerts decrease over time?

Results

The results for the previous research question provided an overview of the on-the-job drowsiness level based on DFM PERCLOS and manual PERCLOS values. If the PERCLOS values did not reach a predetermined threshold, drivers did not receive an alert. Also, some alerts might have been due to increased driver situational awareness (e.g., monitoring the mirrors) or other system limitations (e.g., not capturing the driver's eyes properly). Therefore, the next analysis of the on-the-job drowsiness only considers valid alerts (i.e., visually verified valid alerts). These alerts represent a sample of all the alerts presented to the drivers during the episodes of interest.

Valid Alerts. The analysis performed in order to answer this research question was based solely on visually verified valid alerts, whether they were displayed to the driver or not. A total of 721 valid alerts were used for this analysis. Table 11 presents the split of valid alerts by experimental condition. The test condition had the higher proportion of valid alerts. The higher proportion of valid alerts happened during weekdays, primarily Monday through Wednesday (Figure 28). The valid alerts, as would be expected, were not equally distributed across the time of day (Figure 29); rather they were distributed across the episodes of interest. Therefore, the hours in the middle of the day are not covered (i.e., high illuminance readings).

Table 11. Proportion of Valid Alerts by Experimental Condition

Condition	Frequency	Percent
Baseline Control	46	6
Baseline Test	190	26
Control	47	7
Test	438	61
Total	721	100

Poisson Regression. As discussed in the methods section of this research question, a Poisson regression was used to evaluate the number of valid alerts per week in the study. The 721 valid alerts were grouped by the week the alert was given to the driver. Figure 30 and Figure 31 presents a compilation of all the valid alerts per days and weeks in the study, respectively.

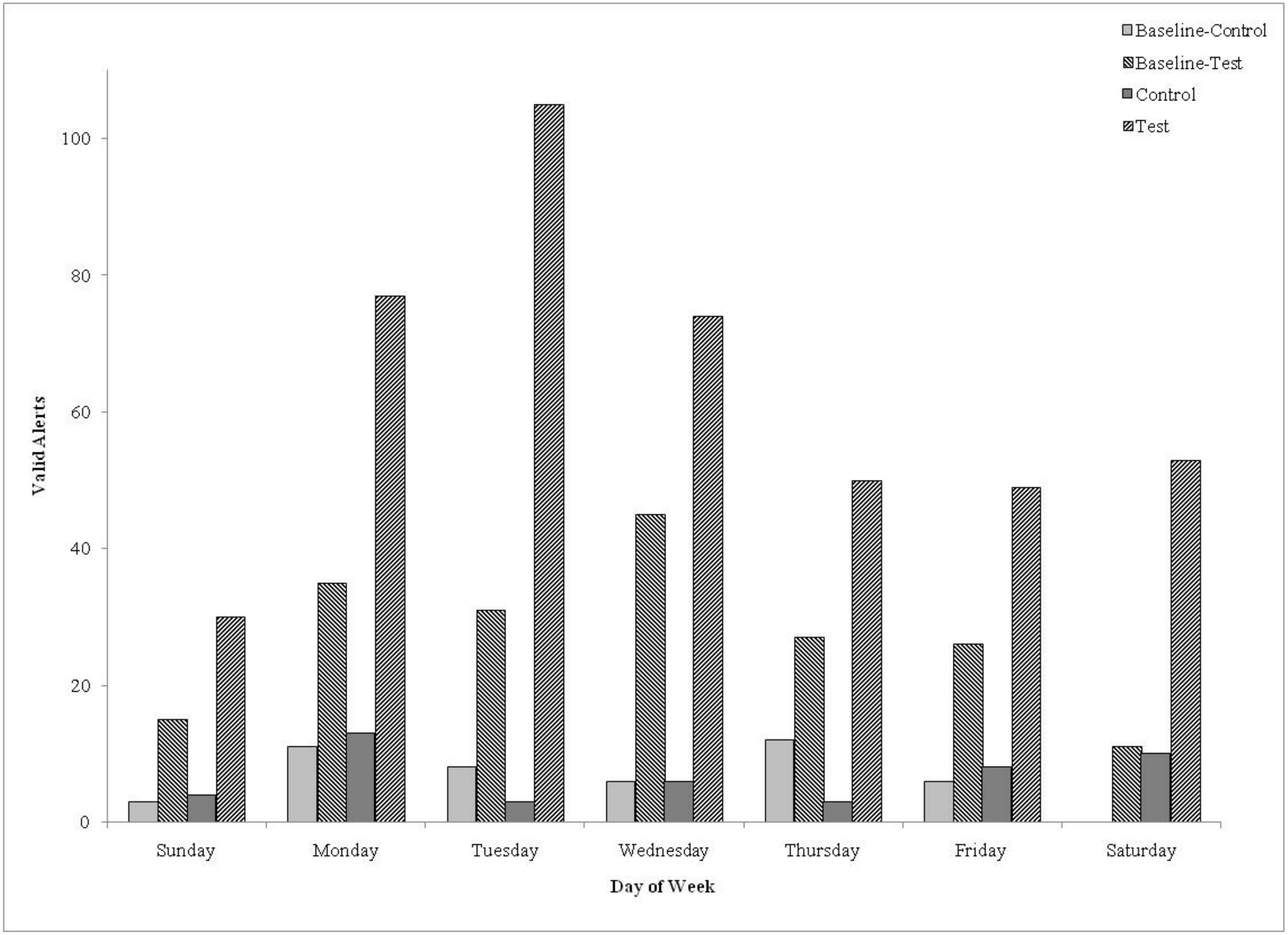


Figure 28. Number of Valid Alerts per Day of Week by Experimental Condition

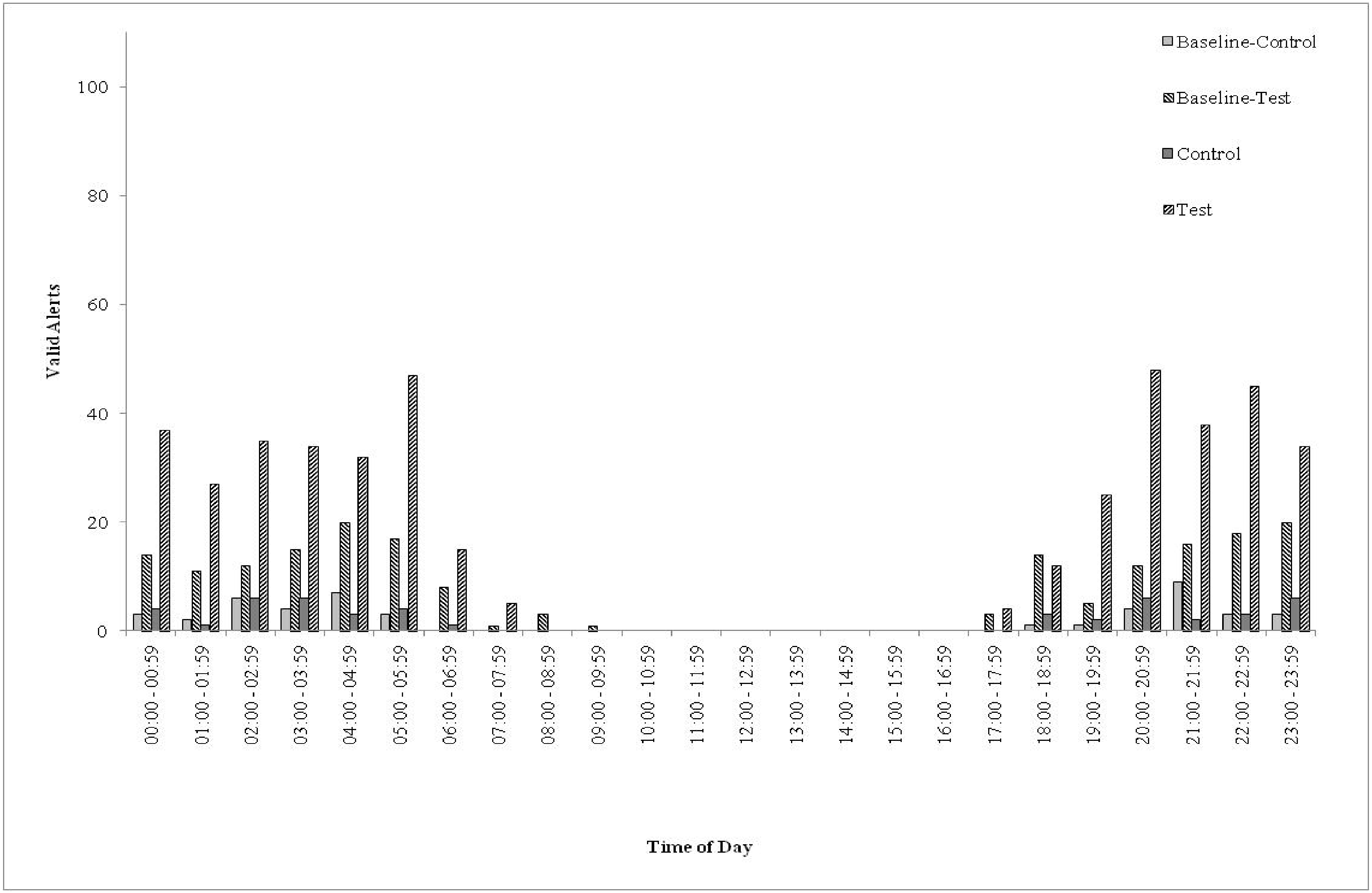


Figure 29. Number of Valid Alerts per Time of Day by Experimental Condition

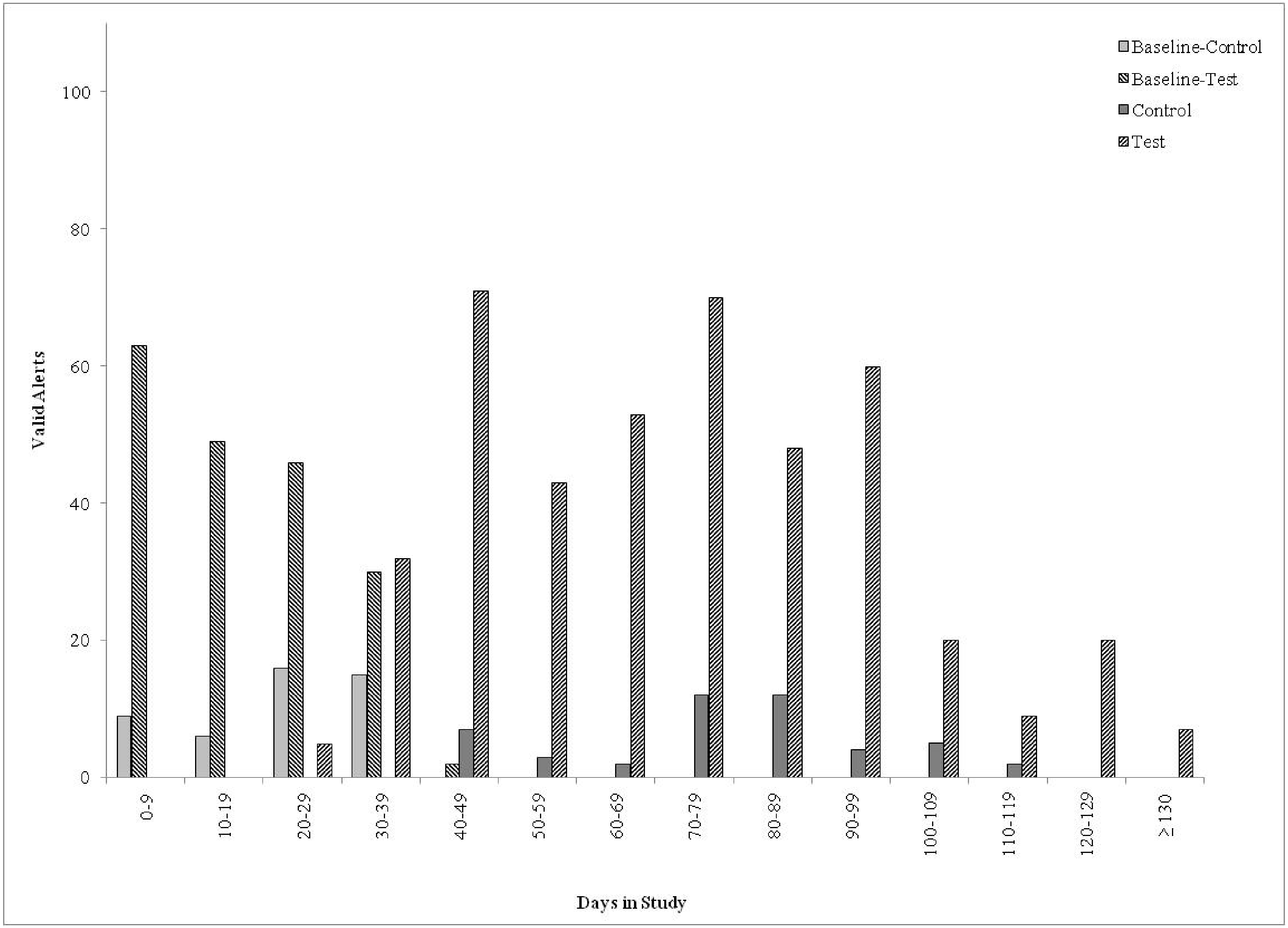


Figure 30. Valid Alerts per Day in the Study

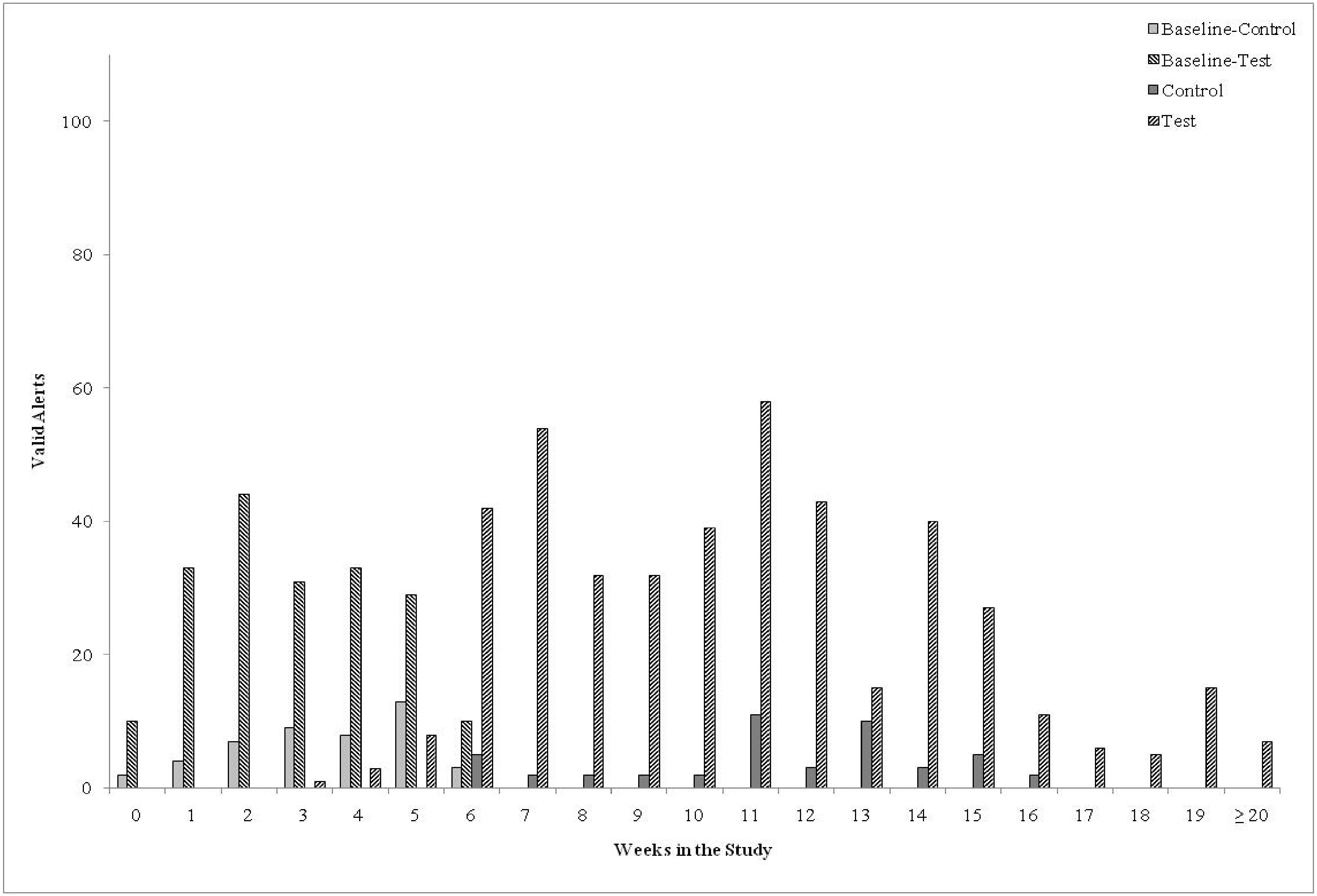


Figure 31. Valid Alerts per Week in Study

The Poisson regression suggests that a statistically significant difference is present between the experimental conditions ($X^2 = 10.98$, $p = 0.0118$). The regression estimates that drivers will obtain more than one valid alert per week if they are part of the Test Group but less than this if they are part of the Control Group (Figure 32). Based on the comparisons of interest shown in Table 12, the only statistically significant difference ($p = 0.0013$) was between the Test and control conditions. Based on this analysis the test condition would have more alerts per week if a DDWS was available to the drivers. As the actual difference between the two conditions is a fraction of an alert, for practical purposes these two conditions are very similar.

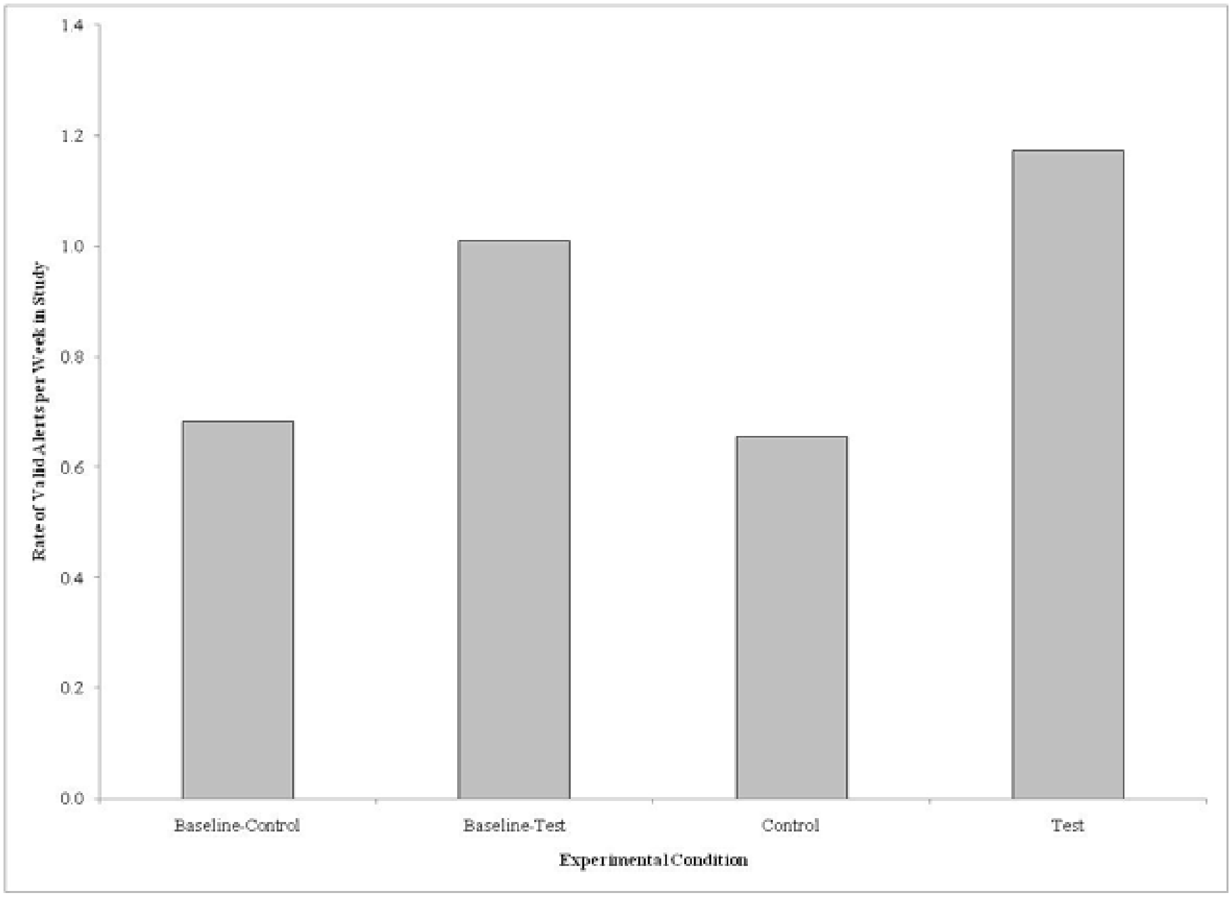


Figure 32. Estimated Alerts per Week by Experimental Condition

Table 12. Poisson Results: Experimental Condition Differences Based on Alerts per Week

Condition	Condition	Estimate	SE	DF	X^2	Pr > X^2
Baseline Control	Baseline Test	-0.3284	0.248	1	1.75	0.1855
Baseline Control	Control	0.0265	0.211	1	0.02	0.9000
Baseline Test	Test	-0.1628	0.1462	1	1.24	0.2653
Control	Test	-0.5178	0.1612	1	10.32	0.0013

RQ 1.3: Do DDWS alerts have an impact on post-alert behavior?

Results

Level of Alerts. The DFM sensitivity level was fixed at PERCLOS-5 (low) for the Control Group and baseline test condition. However, drivers in the test condition were allowed to change the system's sensitivity setting. Figure 33 shows the distribution of valid alerts generated in the test condition across the three DFM sensitivity settings (PERCLOS-5/low, PERCLOS-3/medium, and PERCLOS-1/high). The majority of the drivers (76%) in the test condition elected to have the DFM operate on the medium setting. The second most common setting was the low setting (21%). Only 3 percent of the valid alerts in the test condition were generated with the DFM set to the high setting.

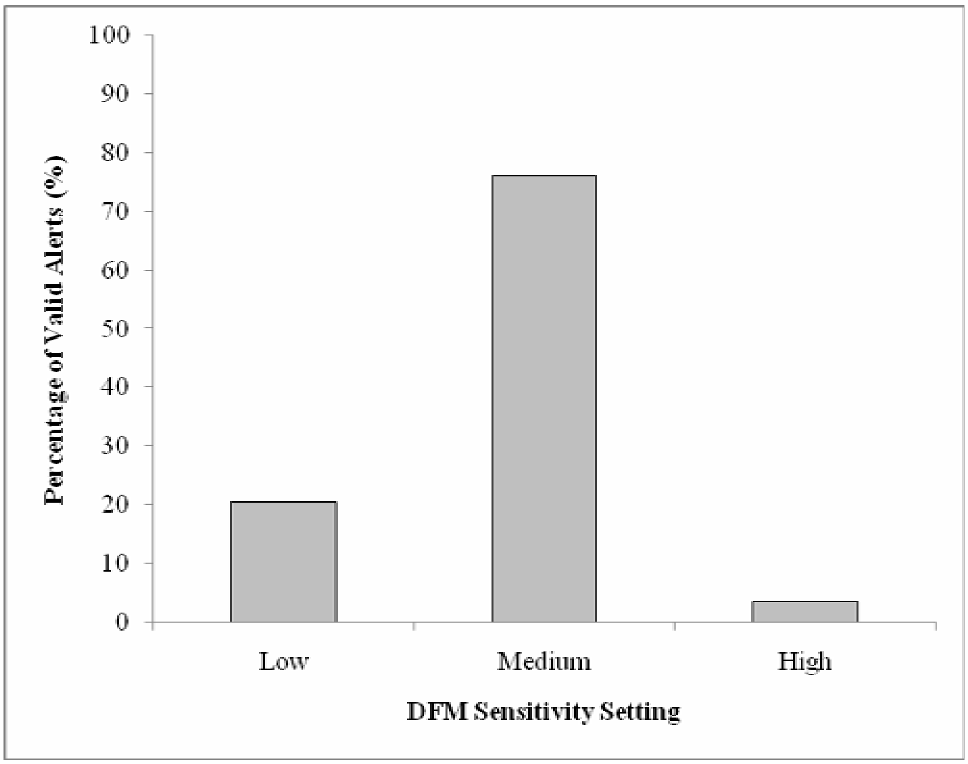


Figure 33. DFM Sensitivity Setting for the Generated Valid Alerts

Environmental Conditions Present during Generation of Valid DFM Alerts. Figure 34 presents the distribution of alerts across time. Although valid alerts were generated between the hours of 7 a.m. and 9 a.m., 99 percent of the valid alerts were generated between the hours of 6 p.m. and 7 a.m. The alerts occurred most frequently at 10 p.m. Their next highest occurrence was at 5 a.m., followed by 8 p.m.

The most frequently occurring valid DFM alerts were generated when the road was dry and no adverse weather conditions were present. Ninety-five percent of the valid DFM alerts were generated during clear weather conditions. Four percent were generated while it was raining (note that weather conditions for 1 percent of valid alerts could not be determined). The road was observed to be dry for 93 percent of the valid alerts and wet for 7 percent of the alerts.

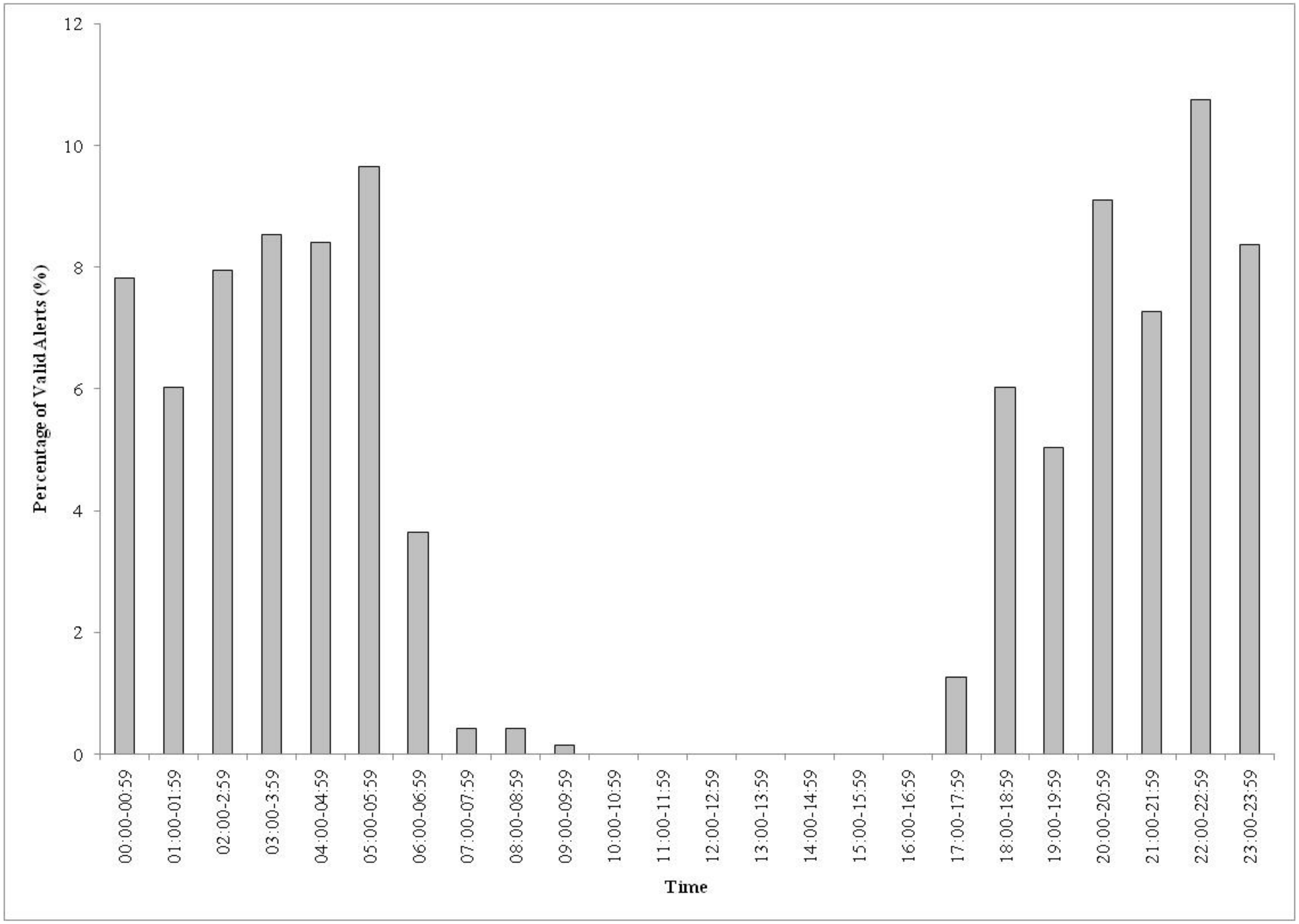


Figure 34. Distribution of Valid DFM Alerts Across Time

The majority of valid alerts (98%) occurred while drivers were traveling on a divided highway. Of the valid alerts generated while traveling on a divided highway, most (83%) occurred on two-lane highways (Table 13).

Table 13. Number of Travel Lanes Available During Valid DFM Alerts Presented on Divided Highways

Number of Travel Lanes	Percent
1	0.43
2	82.70
3	12.48
4	4.11
5	0.28

Figure 35 shows the age distribution of drivers receiving valid DFM alerts. Drivers 30 to 34 years old received the majority of alerts. This is contrasted with the drivers' commercial driving experience, with a mode of 4 years (Figure 36). Fifty-nine drivers received valid alerts; from those, 46 (78% of the sample) were long-haul drivers and 13 (22%) were line-haul drivers.

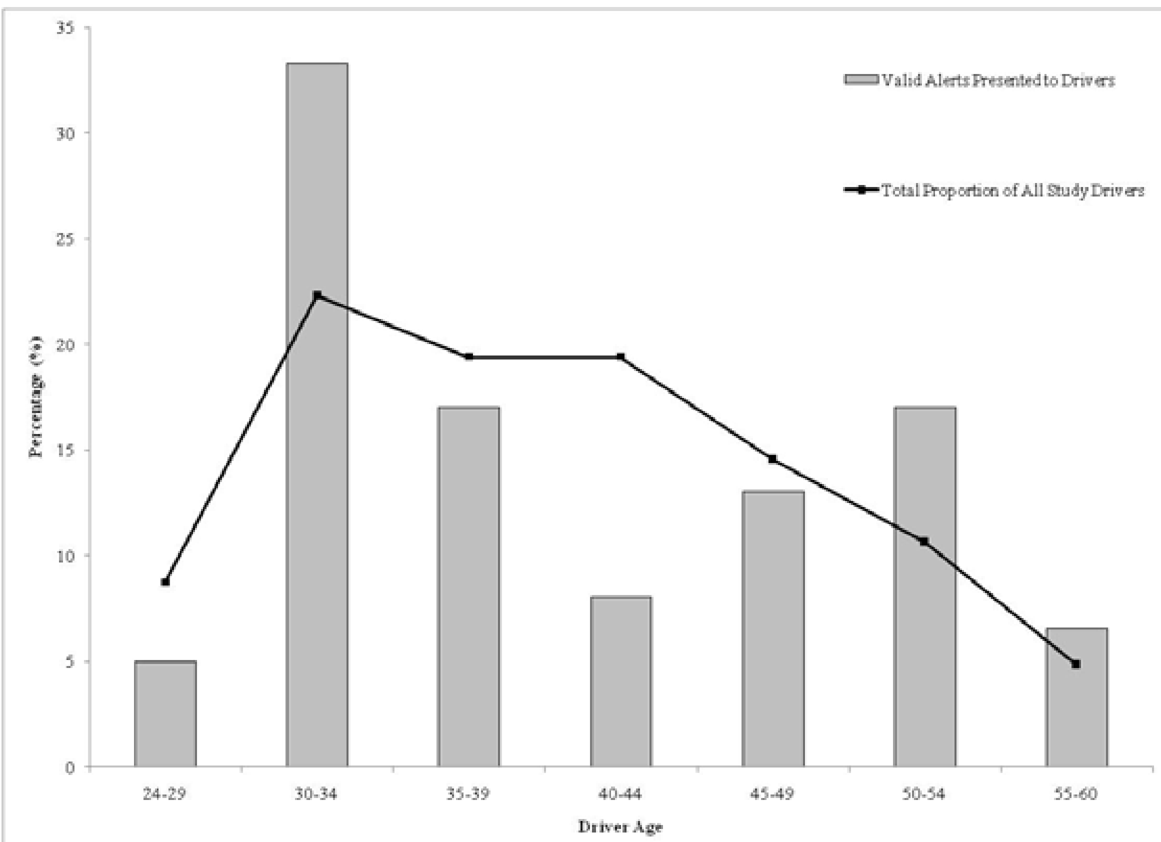


Figure 35. Age Distribution of Drivers Receiving Valid DFM Alerts

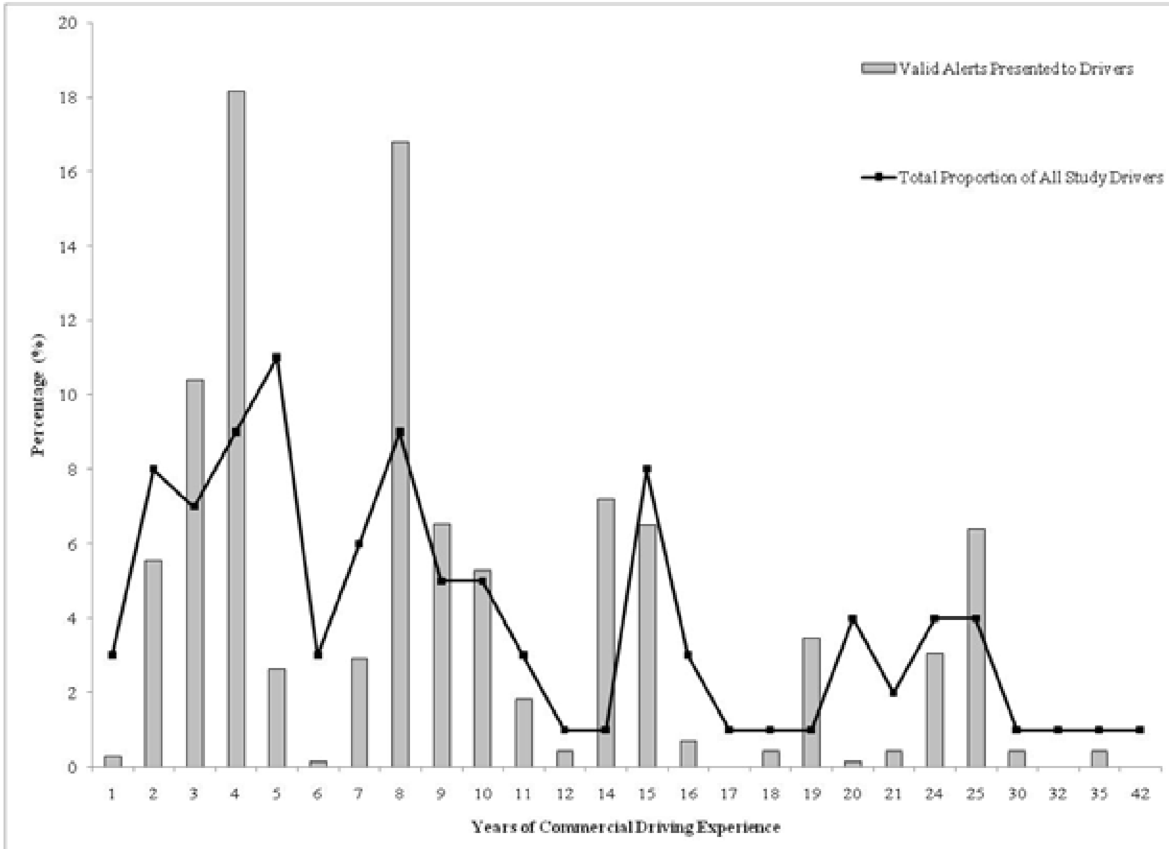


Figure 36. Commercial Driving Experience of Drivers Receiving Valid DFM Alerts

Alert Response – Stop to Rest Behavior. Figure 37 shows the elapsed time between when the driver received a valid DFM alert and when he or she decided to stop the vehicle for a period of 10 min or more. After a valid DFM alert was recorded by the system, drivers in the baseline control condition drove an average of 1 hour and 4 min before stopping the vehicle. Drivers in the control condition drove slightly longer (1 hour and 8 min) before stopping. The difference in elapsed time between the two conditions is not statistically significant ($t(25) = 0.34$, $p = 0.74$). A similar finding occurred for drivers in the Test Group. After receiving a valid DFM alert, drivers in the baseline test condition drove 59 min before stopping, while drivers in the test condition drove 1 hour and 6 min before stopping. Again, the difference between these values is not statistically significant ($t(25) = -2.05$, $p = 0.0515$). The data suggests that even in the case of valid DFM alerts the DFM did not have an effect on drivers' post-alert stopping behavior.

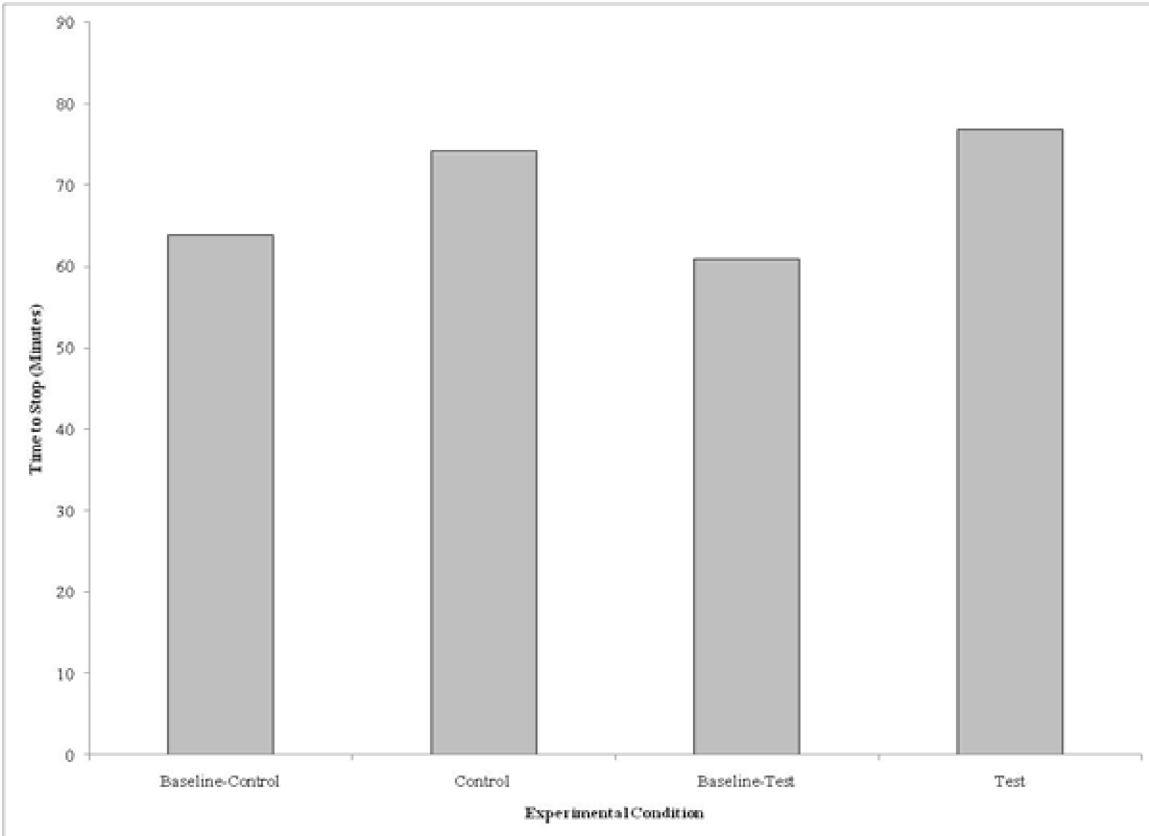


Figure 37. Elapsed Time from Valid Alert to Driver Stopping the Vehicle

Post-Alert Taxonomy. In order to better understand the nature of post-DFM-alert behaviors, ataxonomy of observed post-alert behaviors was created. Table 14 presents this information, with the right column listing those behaviors recorded after the occurrence of valid DFM alerts and the left column presenting the groupings under which they fall. Whether each type of behavior was a temporary measure or if it was considered a countermeasure to alleviate drowsiness is listed alongside each behavior group.

Table 14. Post-Alert Behavior Taxonomy

Post-Alert Behavior Group	Post-Alert Behavior
Adjust Body/Reach/Stretch (Temporary)	• Adjust Body Position • Adjust Head Position • Driver Moving/Stretching Mouth • Raise/Lower Eyebrows • Reach to Front • Reach to Side/Reach Up • Reach Back • Shake Head • Dancing
Touch Body (Temporary)	• Biting nails • Indirectly Touch Head • Rub/Scratch Eyes • Rub/Scratch/Hold Face • Rub/Scratch/Hold Neck • Scratch/Straighten Hair • Touch Body Below Head
Drink/Eat (Temporary)	• Drink • Eat
Driver Input (Temporary)	• Enter Highway • Exit Highway • Press Brake Pedal • Stop at Intersection • Right Lane Change • Stop at Toll Booth • Turn (Left or Right)
Use Communication Device (Temporary)	• Use CB Radio • Use Cell Phone
Stop Vehicle (Rest)	• Stop and Leave • Stop and Rest
Reading/Writing (Temporary)	• Reading • Writing
Raise/Lower Window (Temporary)	• Raise/Lower Window
Restart (Temporary)	• Continued driving
Smoke/Chew Tobacco (Temporary)	• Smoke/Chew Tobacco
Take Pill/Medication (Temporary)	• Take Pill/Medication
Talk/Sing/Laugh (Temporary)	• Talk/Sing/Laugh
Turn Light On/Off (Temporary)	• Turn Light On/Off
Veer Off Road (Temporary)	• Veer Off Road
Look to Side/Up/Down (Temporary)	• Look to Side/Up/Down

Post-Alert Behaviors by Test Condition. The range of post-alert behaviors observed was grouped by test conditions. For drivers in the Control Group, these frequencies of post-valid alert behaviors are presented in Figure 38. The behavior data is divided between alerts generated in the baseline control and control conditions. A Cochran-Mantel-Hanzel test ($CMH(1) = 0.5184$, $p = 0.4715$) did not reveal any statistically significant differences in the distribution of behavior frequency between the baseline control and control conditions. This test was used due to the repeated measures component of the analysis needed and to avoid assuming a normal distribution (Hollander & Wolfe, 1999). This is expected, as drivers in the Control Group could not perceive the DFM alerts (the alerts were recorded by the DAS but were not presented to drivers in either the baseline control or control conditions).

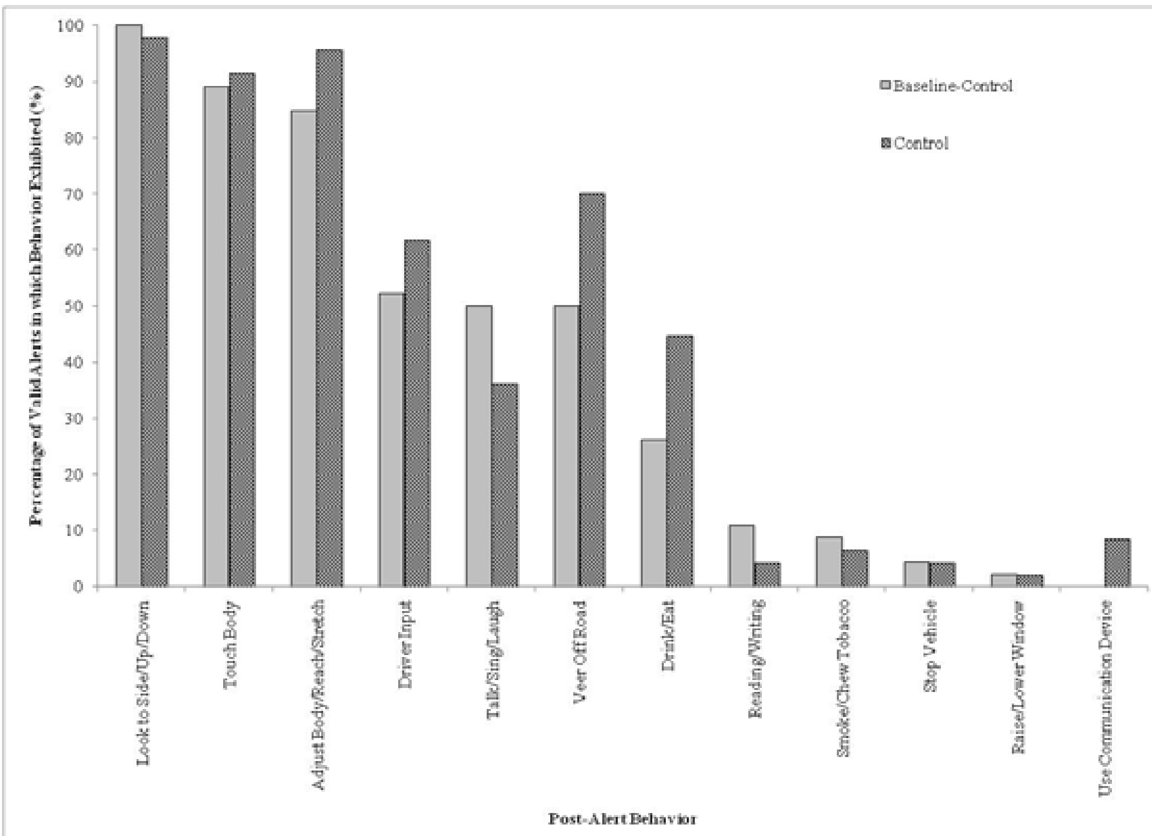


Figure 38. Percentage of Valid Alerts by Type of Behavior: Control Group

Figure 39 presents the frequency of behaviors that were observed after drivers in the Test Group received valid DFM alerts. Again, the behavior data is divided between alerts generated in the baseline test and test conditions. A Cochran-Mantel-Hanzel test ($CMH(1) = 0$, $p = 0.9992$) did not indicate the presence of any statistically significant differences for the experimental conditions within the Test Group in terms of the distribution of behavior frequencies. This finding suggests that the DFM did not have an effect on the frequency of the observed post-alert behaviors.

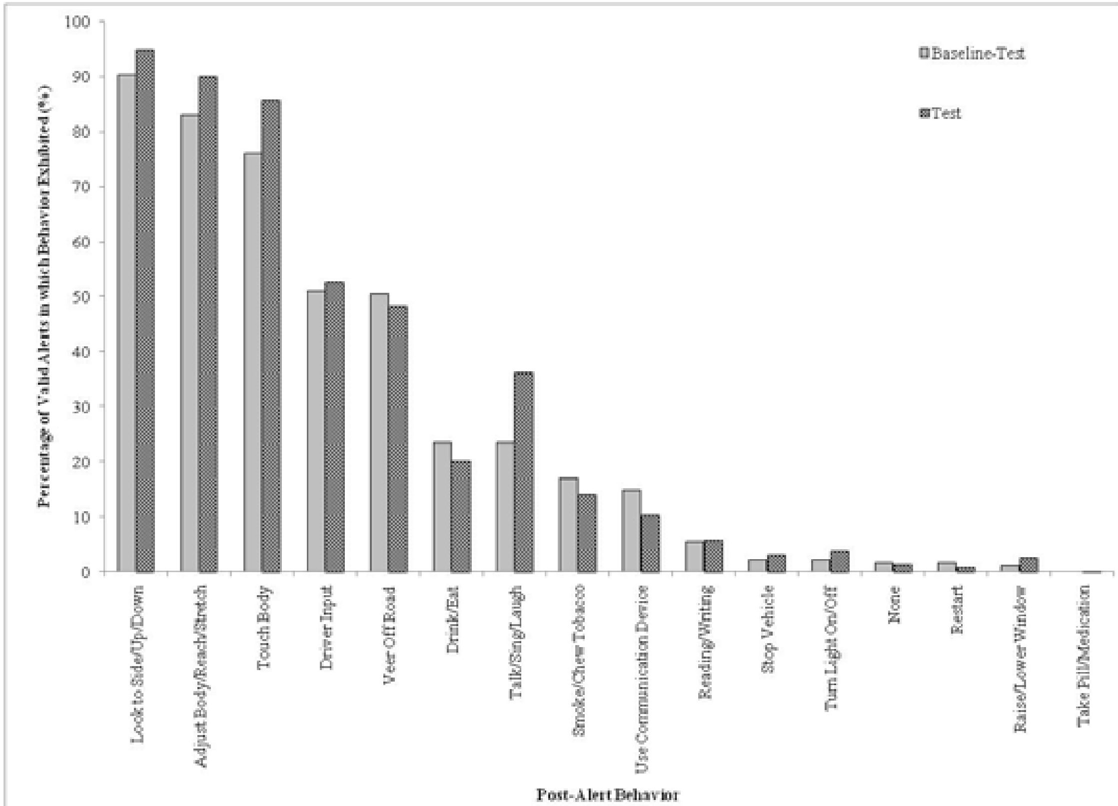


Figure 39. Percentage of Valid Alerts by Type of Behavior: Test Group

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

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

Elapsed Time from Alert to Behavior. Figure 40 presents the average elapsed time from a valid DFM alert being recorded to the observed behavior for alerts presented in the baseline controlcontrol condition and the control condition. MLM was not possible within this data set due to an insufficient amount of data in some specific levels of the analysis.

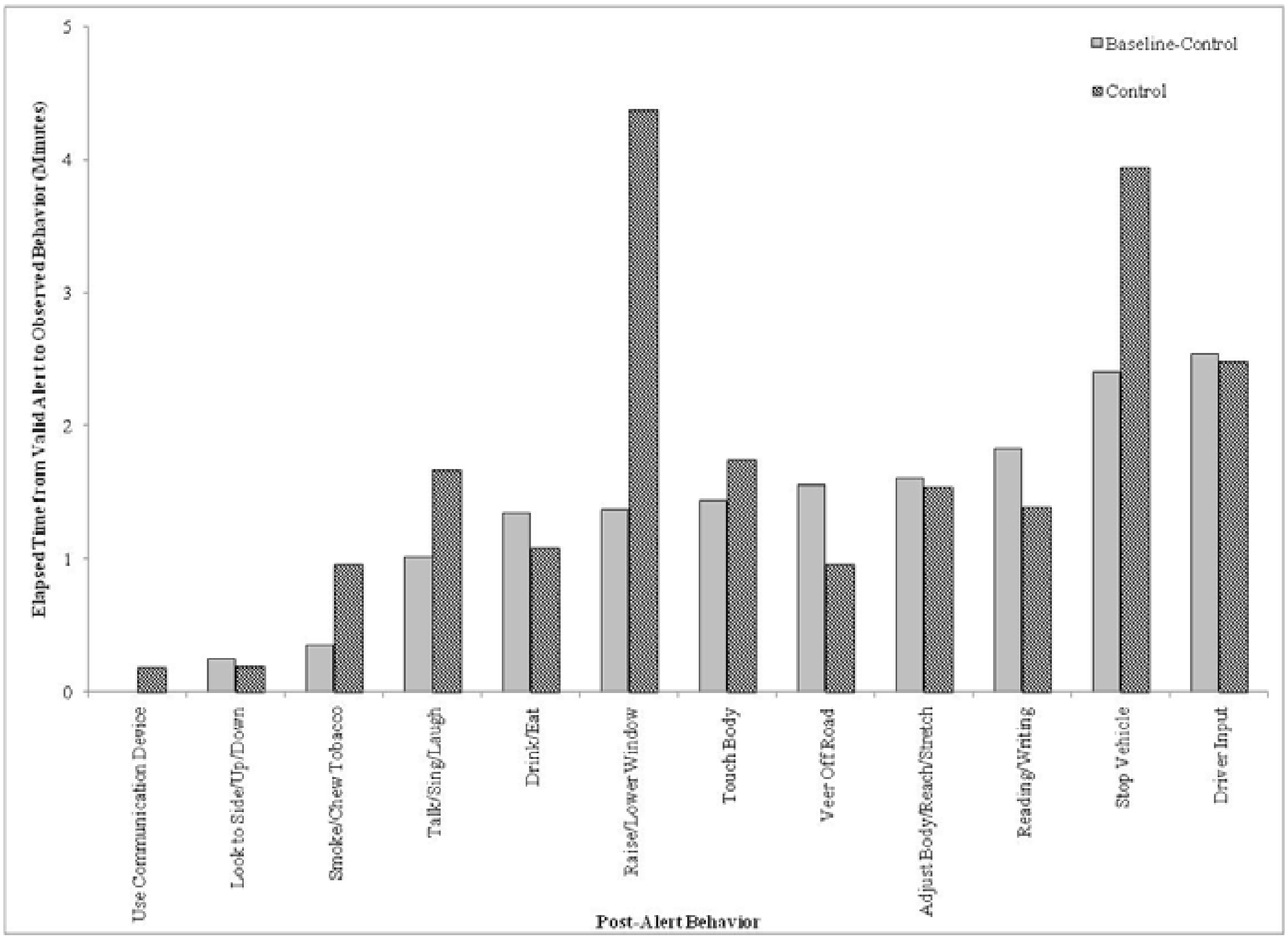


Figure 40. Elapsed Time from Valid DFM Alert to Observed Behavior: Control Group

Figure 41 presents the average elapsed time from a valid DFM alert to the observed behavior for alerts presented in the baseline test condition and the test condition. An MLM analysis was not possible owing to an insufficient amount of data.

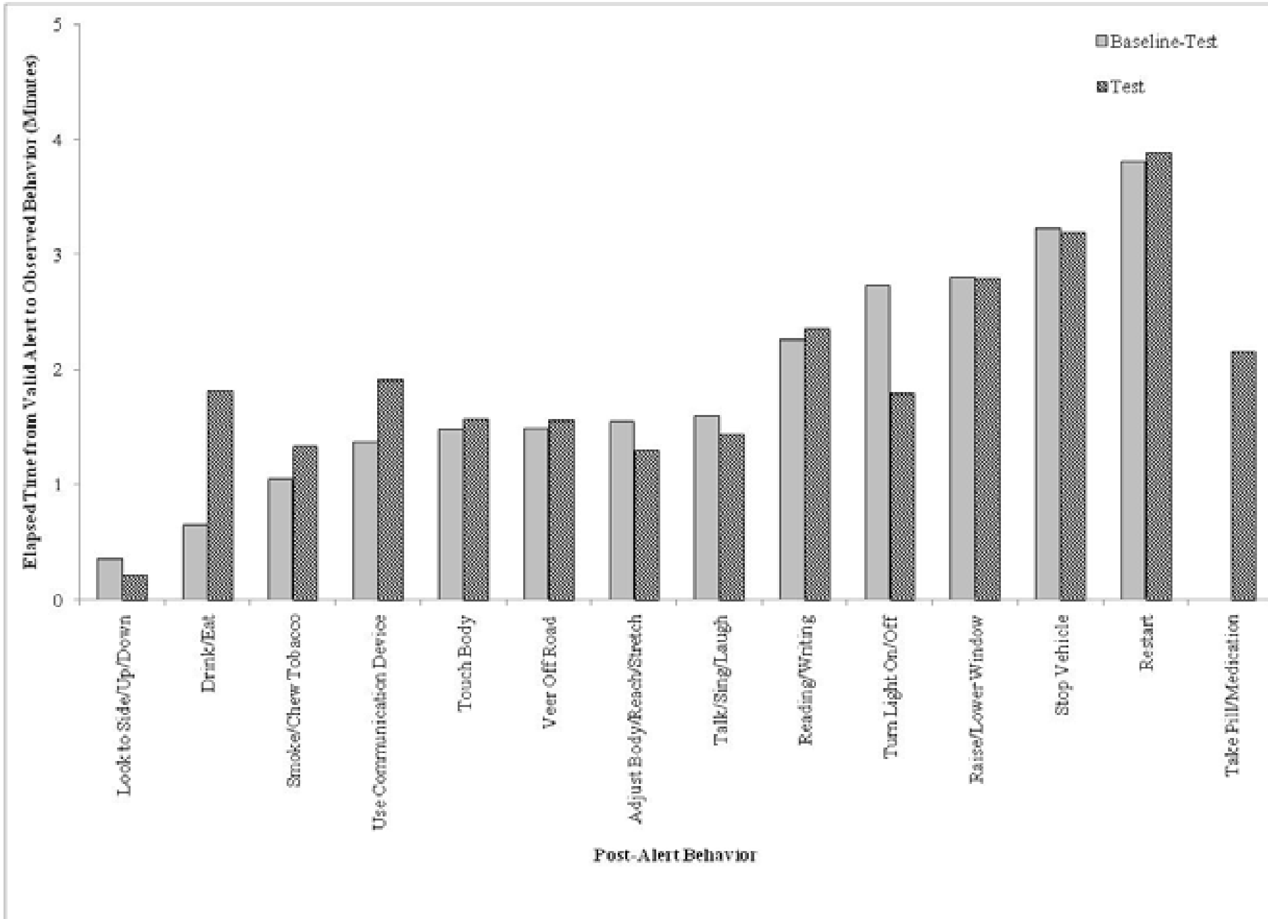


Figure 41. Elapsed Time from Valid DFM Alert to Observed Behavior: Test Group

Discussion

Drowsiness Assessment. This research question assessed how the level of drowsiness of the participants who were exposed to the DFM alerts (i.e., Test Group, specifically during the test condition) compares to the other experimental conditions where DFM feedback was not provided. In order to accomplish this portion of the assessment, the level of drowsiness the drivers exhibited throughout the study was evaluated. If the DFM was successful in providing the adequate feedback to the drivers, this would be reflected in lower PERCLOS levels during the test condition than during the other experimental conditions.

Based on the limitations and operational envelope of the prototype DFM, several considerations were made to appropriately evaluate the DDWS concept. The DDWS concept is based on the determination of a PERCLOS level. PERCLOS was defined by Wierwille et al. (1994) as a slow eye-closure where the eyelid covers more than 80 percent of the pupil. First, episodes of interest were identified in the continuous data files for each driver to try to eliminate the instances where the DFM might be operating unreliably due to extraneous conditions. Second, sampling of alerts was performed for each week of drivers' participation to assess any major problems that may have affected the results. The validation procedure revealed a discrepancy between the DFM estimate of eyes-closed and the DDWS concept requirement for the determination of the level of driver's drowsiness (i.e., PERCLOS level). Specifically, the DFM tended to increase the PERCLOS level if the DFM's camera was unable to capture the driver's eyes properly (e.g., when the driver was scanning the mirrors) or other situations such as if the camera went out of alignment (i.e., bumped by the driver while moving objects within the vehicle cab). Therefore, measures of central tendency for each driver were used for the analyses. This should avoid picking up extreme PERCLOS values artificially inflated by instances such as lane changes in an interstate when traffic is surrounding the research vehicle. During these instances the driver needs to scan the mirrors multiple times before deciding if it is appropriate to begin a lane change, which in turn increased the PERCLOS level calculated by the DFM.

Two independent evaluations were performed in order to obtain an accurate evaluation of the drowsiness level of the drivers participating in this study. The first evaluation took into consideration DFM PERCLOS (automatically calculated by the DFM during the episodes of interest), and the second used manual PERCLOS (calculated based on eye-closure observed in a video-based data reduction process for random points in time throughout the study). In addition, two types of analyses were performed for these two different measures of PERCLOS. The first portion of the analysis was performed by evaluating descriptive statistics, and the second used MLM. Evaluating the mean DFM PERCLOS by taking into consideration the days drivers were in the study does not show a dramatic reduction of drowsiness over time. The test condition represented the lowest PERCLOS values for most of the days in the study regardless of the day of the week in which the value was obtained. As driving hours progress throughout the study a similar trend for the test condition was obtained, in which the drivers exposed to the DDWS had a lower PERCLOS value. At a few points in time (depending on the unit of time examined) this trend switched with the control condition; but for most of the instances the Test Condtion exibited the lowerst PERCLOS level.

The second evaluation looked at the manual PERCLOS values. The mean manual PERCLOS evaluation, which considered the days the drivers were participating in the study, did not indicate the presence of a clear trend for the test condition. Similar to the DFM PERCLOS results, a time progression assessment was done by seperately evaluating the manual PERCLOS values by weeks and hours of driver particiaption in the study. Manual PERCLOS for the test condition was lower than the ones for the control condition from the eighth week on. This trend was not found when the accumulated driver hours were examined.

The MLM results based on DFM PERCLOS suggest that, over time, PERCLOS is lower for the test condition (i.e., when the driver is receiving DDWS feedback) than during the other conditions in which the system is not providing drowsiness feedback. However, the test condition results obtained from this model suggest a small reduction in PERCLOS level (i.e., 1.44 lower than during the baseline test condition). Based on the MLM model none of the experimental conditions reached the level where even an initial alert would be needed.

The MLM model based on manual PERCLOS represents the drowsiness level at any random point during the drive. In contrast with the DFM response, recall that the data used for this model was not restricted to nighttime driving or any specific criteria used for episodes of interest in the examination of DFM PERCLOS. The results based on the MLM for manual PERCLOS suggest there are no statistically significant differences between the experimental conditions or within

Test Group. Based on PERCLOS estimates for this model, drivers in the Test Group did not exhibit a statistically significant change in the level of drowsiness through the study.

In summary, the prototype used to evaluate the DDWS concept suggests that there is a statistically significant reduction in the level of drowsiness over time when a driver receives feedback on alertness level during restricted illumination conditions (i.e., dawn, dusk, night). However, when the evaluation is performed outside the operating envelope of the prototype system, a similar reduction in the level of drowsiness is not observed. The prototype system used to assess the DDWS concept only operated within a limited set of conditions (e.g., low illuminance, driver cannot wear glasses). In order to generalize these results to the larger population of interest, other conditions must be explored (e.g., glasses, daytime, not considering mirror scanning or distraction to increase PERCLOS level).

Alert Frequency. Drivers only received valid alerts when the PERCLOS value reached a predetermined threshold and the system was in active mode (i.e., test condition). Alerts during the dark mode (i.e., baseline test, baseline control, and control conditions) were recorded by the DAS but not presented to the drivers. As mentioned in the methods section for this research question, an alert validation was performed, and only valid alerts were used for this portion of the analysis. Most of the alerts (61 percent of all alerts) obtained were in the test condition. The highest proportions of valid alerts occurred during weekdays, mainly Monday through Wednesday, and were equally distributed at the times during the day when the prototype system was operational. The Poisson regression analysis suggests a statistically significant difference between the experimental conditions. The regression estimates that drivers will obtain more than one valid alert per week if they are part of the Test Group, but less than that if they are part of the Control Group. Based on this analysis, drivers would be more prone to need an alert if a DDWS was available than if it was not available to them. This is potentially an unintended consequence of having a system that will identify for the drivers when they need to rest and not having them self-regulate their sleep (i.e., they might be pushing their limit knowing that the system will let them know when to rest). However, the actual difference between the two conditions is just a fraction of an alert. Therefore, for practical purposes these two conditions are very similar.

Potentially, this result could be an artifact of the prototype limitations and not a response to actual drowsiness by the DDWS. Having more valid alerts might mean that the drivers were accommodating their posture or scanning behavior to reduce the number of false alarms during the test condition given that they were actually exposed to the alarm and could react to it. If this hypothesis is true the consequences could be twofold. Effects are positive if it reduced distractions that caused the eyes to be off of the forward roadway and/or negative if it reduced scanning of mirrors and situational awareness to reduce false alarms. However, eye glance behavior is outside of the scope of this analysis, and the hypothesis cannot be tested as part of this research effort.

Post-Alert Behavior. Most of the valid alerts for the test condition were in the PERCLOS-3 (medium) sensitivity level. Although the DFM setting was not investigated, for both valid and invalid DFM alerts, drivers elected to use the medium sensitivity level more often than the other two settings.

Valid DFM alerts occurred evenly across the hours during which DFM was operational. There was a slight increase in the number of valid alerts at 10 p.m. One explanation for this finding is that drivers' circadian rhythms are more strongly driving the urge to sleep at approximately this time of the evening (Schmidt, Collette, Cajochen, & Peigneux, 2007). During the data reduction process, it was determined that the majority of valid alerts also occurred on dry two-lane divided highways when no adverse weather conditions were present. This is an expected finding, as the nature of long-distance trucking makes driving in such conditions entirely common. As other researchers have noted (Thiffault & Bergeron, 2003), perhaps the monotonous conditions present in the largely unchanging driving environment are highly conducive to vigilance effects and therefore contribute to the onset of drowsiness.

A measurable impact of valid DFM alerts on drivers' post-alert behavior, both in terms of the frequency and timing of drivers' post-alert behaviors, was not found in this study. Nevertheless, the data indicate that drivers perform a myriad of behaviors when they are drowsy. It is possible that these behaviors were performed to counteract the state of vigilance imposed by the driving task. Behaviors that stimulate either cognitive or physiological faculties, such as seeking and perceiving new information or sensations, may serve to reduce the magnitude of the vigilance decrement and resulting drowsiness. Additionally, this serves as an indication that driver reactions to the DFM prototype were similar in magnitude to the responses observed in the control condition.

RESEARCH QUESTION 2: SLEEP HYGIENE

As part of this research question two sub-questions were answered:

> RQ 2.1: Sleep Quantity: Does a DDWS influence drivers to get more sleep?
> RQ 2.2: Sleep Quality: Do drivers using a DDWS achieve better quality sleep?

Methods, analysis, measures of interest, results, and discussion related to these two questions are presented next.

Methods

Actigraphy

The procedures necessary to accurately measure heavy-vehicle drivers' sleep without invading their privacy, disrupting their sleep, or overtly manipulating their sleep environment are still undergoing development. Sleep has traditionally been recorded and observed using laboratory-based measurement procedures requiring participants to sleep at a predetermined location (Mitler et al., 1997). These procedures include polysomnographic scoring, which is performed by attaching electrodes to participants to obtain central- and occipital-derivation electroencephalography (EEG) measures, recording movement of both eyes through electrooculography (EOG), and recording muscle activity of the chin through electromyography (EMG). Measuring the participant's respiratory air flow as well as pulse oximetry are often additional measures taken to assist in the assessment of sleep (Mitler et al., 1997). Although the measurements obtained are precise, the instruments required are intrusive and must be connected to the person as far as 90 min before the onset of sleep. As the equipment may become disconnected as the person moved during sleep, researchers must constantly monitor the equipment status and be prepared to reconnect the instruments if they become detached.

A less invasive and well-accepted alternative to laboratory sleep measurement is actigraphy (Cole et al., 1992; Jean-Louis et al., 2001). Actigraphy is the measurement of sleep by detecting motor activity while a participant attempts to sleep (Ancoli-Israel et al., 2003). This technique operates under the premise that motor activity is indicative of wakefulness, while motor inactivity is indicative of sleep. Using this less-invasive technique of actigraphy, the effects of the DFM system on truck driver sleep hygiene was investigated. An additional benefit of actigraphy is its ability to record continuously over extended periods. This makes actigraphy more reliable than sleep logs requiring participants to not only recall episodes and durations of sleep during the night, but also to remember to record their sleep activity on a very regular basis. Actigraphy is also considered more reliable than observations, which are limited in length (Ancoli-Israel et al., 2003).

Participants in this study wore an actigraphy device on the wrist of their non-dominant hand. The device used, a MicroMini Motionlogger Actigraph (Ambulatory Monitoring, Inc., AMI), is the approximate size and shape of a wrist watch (Figure 42). A piezoelectric accelerometer in the device detects wrist movement (categorized as movements where accelerations > 0.003 g) and generates a voltage in response to each movement. If the generated voltage surpasses a set reference voltage, each zero-crossing increments an activity count, which is stored in the device's memory. This data is then downloaded from the device onto a computer where it can be analyzed.

Figure 42. Actigraph Device Worn by Participants

The frequency of zero-crossings over time is used to quantify sleep. This calculation depends on the assumption that few zero crossings, signifying a low amount of wrist activity, is indicative of sleep (Jean-Louis et al., 2001). The Cole-Kripke sleep algorithm (Cole et al., 1992) was used to score the actigraphy data into individual sleep periods. The algorithm does not necessarily require complete wrist inactivity to score sleep. Brief wrist activity is accounted for and removed from the analysis. The Cole-Kripke sleep estimation algorithm is able to accurately detect sleep from actigraphy data approximately 88 percent of the time (as validated by a standard polysomnogram montage, two-channel EOG, and EMG; Cole et al., 1992).

Two caveats are necessary when using actigraphy data. The first is that because the algorithm detects change in movement as a function of time, it is possible that inactivity during waking hours will be scored as sleep, and motor activity occurring during restless sleep will be scored as waking (Tani et al., 2005). The second is that inactivity is only truly indicative of sleep when it is known with some certainty that the participant is attempting to sleep. This information is typically confirmed using sleep logs or direct observation (Jean-Louis et al., 2001). As direct observation and sleep logs were both outside the scope of this study, an algorithm to group scored sleep periods (SSPs) together was developed.

The algorithm used is based on the observation that clusters of SSPs are likely to be indicative of attempted sleep. The algorithm operates by first identifying an SSP that is longer than 20 min. Therefore, any SSP that was less than 20 min (i.e., naps) was not included in the total amount of sleep for a driver. The reason for excluding the ≤ 20-min sleep segments was to achieve a reliable sleep quality measure, which, it was thought, short sleep/nap periods might not provide. SSPs of any length that occur within an hour of the > 20-min anchor sleep period are grouped with the anchor sleep period. SSPs within an hour of the last SSP in the group are iteratively included. Additionally, SSPs whose conclusion occurs within a half hour of the beginning of the anchor sleep period are included in the grouping. SSPs whose conclusion is within a half hour of the beginning of the earliest sleep period are also iteratively added in. The elapsed time from the beginning of the first SSP in the group to the end of the last sleep period in the group forms the sleep onset-sleep offset, or O-O, interval (Figure 43). Note that sleep periods of less than 20 min were not included in the sleep quantity measure. One such period can be seen in the first row, to the right of the O-O interval shown in Figure 43. This procedure for assessing the O-O interval without sleep logs was investigated using collected sleep data from VTTI personnel. These personnel kept sleep logs for three days. The time they spent in bed was compared to the O-O intervals computed from their actigraphy data and was found to closely match. However, it should be mentioned that the sleep patterns of researchers at VTTI are likely different from those exhibited by commercial motor vehicle drivers.

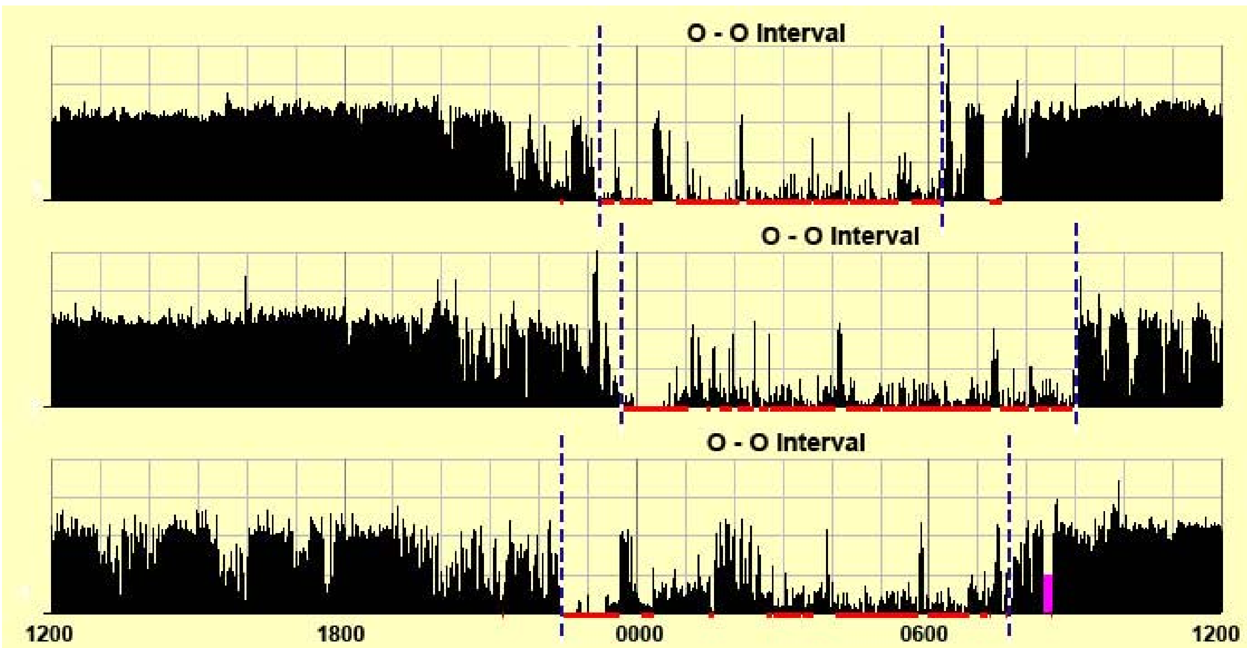


Figure 43. Example of Different O-O Intervals

After the O-O intervals were identified, the SSPs contained within the O-O intervals were summed and averaged across 24-hour periods. The average quantity of sleep reported per day is thus the sum of the SSPs contained in the O-O interval in a standard calendar day. This approach is slightly different from that used by Hanowski et al. (2007), who estimated quantity of sleep by summing all the SSPs in a 24-hour period (including those outside the O-O interval), and Greco et al. (2004), who estimated sleep quantity by summing SSPs over a 12-hour period.

Estimating sleep quality, however, demands consideration for the entire O-O interval. To avoid attributing those O-O intervals spanning two calendar days to both days, the sleep quality statistics pertaining to an O-O interval were assigned to the calendar day in which the O-O interval began.

It was not uncommon for drivers to remove their Actigraph watches over the course of the study. When this occurred, the watch remained still and may have misrepresented sleep data. To accommodate this possibility, actigraphy data was considered to be invalid when the activity level decreased from a high value (e.g., 200 or greater) to zero in 1 min, stayed at zero for over 10 min (flat line), and then jumped from zero to a higher level. Days in which more than

20 percent (4.8 hours) of the Actigraph data was marked as bad were removed from the sleep analyses. These data were removed because it could not be accurately determined whether the participants were sleeping or awake during this time period.

Measures

Before listing the measures of sleep quantity and quality, it is worth explaining the sleep processes. Sleep is organized into cycles of rapid eye movement (REM) and non-REM (or nREM) stages (Rechtschaffen & Siegel, 2000). These stages consist of REM and four stages of nREM sleep. People generally cycle through the stages of sleep in the following order: stage 1, a low activity transition; stage 2, periodic neuronal activity against continual low-level EEG activity; stage 3, increasing high-amplitude EEG activity; and stage 4, high-amplitude EEG activity in the prefrontal cortex. These four nREM stages are followed by REM sleep, which resembles nREM stage 1 in terms of its EEG profile. Moving from stages 1 to 3 typically takes 70-80 min, while moving from stages 1 to 4 typically takes 90-110 min. Periods of REM sleep typically last between 5 and 10 min.

This cycle repeats four to six times per night. It is noteworthy that brain activity during REM sleep resembles that observed during wakefulness. Although the functions of sleep are still largely unknown, it is believed that REM sleep facilitates neural maturation, improves mental health, and may facilitate learning and memory. It is also the sleep stage where most dreams occur.

A lack of deep sleep and REM sleep can reduce sleep quality (Tasali, Leproult, Ehrmann, & Van Cauter, 2008). Sleep interruptions forcing the sleep cycle to restart reduce sleep quality. Such interruptions have also been shown to lead to a deficit of leptin production in the human body, the hormone responsible for suppressing appetite (Spiegel, Tasali, Penev, & Van Cauter, 2007). This is why it is believed that sleep-quality-deprived individuals experience a higher prevalence of obesity. The continuity (a form of quality) of sleep is therefore equally as important as the quantity of sleep. The following measures are used to assess the quantity and quality of sleep obtained by heavy-vehicle drivers:

Measuring Sleep Quantity. The quantity of sleep obtained by drivers was investigated by analyzing the average amount of total SSPs in an O-O interval for a calendar day.

Measuring Sleep Quality. The quality of sleep obtained by drivers was investigated by averaging the following four measures across each valid day the O-O interval begins:

1. Sleep Efficiency: the percentage of time in the O-O interval that is spent asleep.
2. Number of Awakenings: the number of discontinuities between SSPs in an O-O interval.
3. Number of Sleep Periods Longer than 20 min: a count of the SSPs whose duration extended past 20 min. As noted, sleep periods that were less than 20 min were excluded.
4. Longest Sleep Period in O-O Interval: the longest stretch of uninterrupted sleep that occurred in the O-O interval.

These measures were adapted from Greco et al. (2004) to assess sleep quality. They are also recommended by Ambulatory Monitoring (the Actigraph watch manufacturer) as a method of assessing sleep quality via the Actigraph watch. The suggested methods were adapted to include only sleep periods 20 min or longer in order to accurately evaluate sleep quality during the periods of interest.

RQ 2.1: Sleep Quantity: Does a DDWS influence drivers to get more sleep?

Results

The average quantity of sleep truck drivers attained in a 24-hour period was calculated as a function of experimental condition (Table 15). On average, drivers received 5.47 hours of sleep while in the baseline controlcontrol condition, and an average of 5.74 hours of sleep per period following entry into the control condition. The same was true for drivers in the Test Group. On average, drivers received 5.52 hours of sleep in the baseline test condition and 5.66 hours of sleep per period after moving into the test condition. An MLM analysis of sleep quantity did not reveal statistically significant differences between the baseline control and Control conditions ($t(82) = 1.66$, $p = 0.10$). A similar analysis was used to investigate possible differences between the baseline test and test condition. Again, statistically significant differences were not found ($t(82) = 0.99$, $p = 0.33$). It cannot be said that either the DFM studied or factors of experimental demand encouraged drivers to obtain more sleep.

Table 15. Mean Sleep Quantity by Experimental Condition

Experimental Condition	Number of Days	Mean Sleep Quantity in 24 hrs (h)	Standard Error (h)
Baseline Control	595	5.47	0.12
Control	801	5.74	0.10
Baseline Test	1,826	5.49	0.06
Test	3,483	5.67	0.05
Overall	6,705	5.61	0.03

RQ 2.2: Sleep Quality: Do drivers using a DDWS achieved better quality sleep?

Results

Drivers’ quality of sleep was assessed by analyzing the continuity of their sleep during the O-O interval. Table 16 presents drivers’ sleep efficiency, which is the percentage of time in the O-O interval that was scored as sleep, by experimental condition. On average, drivers spent 82 percent of their time sleeping when they intended to sleep. An MLM analysis of sleep efficiency did not reveal statistically significant differences between the baseline control and control conditions ($t(82) = -0.11$, $p = 0.91$). A similar analysis was used to investigate possible differences between the baseline test and test condition. Again, statistically significant differences were not found ($t(82) = -0.63$, $p = 0.53$).

Table 16. Mean Sleep Efficiency by Experimental Condition

Experimental Condition	Number of Days	Mean Sleep Efficiency	Standard Error
Baseline Control	505	81	0.59
Control	680	82	0.53
Baseline Test	1,604	82	0.34
Test	3,041	82	0.26
Overall	5,830	82	0.18

Interruptions in the continuity of sleep reduce the total amounts of deep and REM sleep obtained. The average amount of times drivers awoke during an O-O interval by experimental condition is presented in Table 17. Awakenings are defined here as actigraphy activity surpassing a pre-defined level of actigraphy activity that can occur when one sleeps. The data indicate that drivers woke, on average, 11 times per O-O interval. This suggests that these drivers did not attain continuous sleep. An MLM analysis of the number of awakenings did not reveal statistically significant differences between the baseline control and control conditions ($t(82) = 0.90$, $p = 0.37$). A similar analysis was used to investigate possible differences between the baseline test and test condition. Again, statistically significant differences were not found ($t(82) = -1.8$, $p = 0.08$).

Table 17. Mean Number of Awakenings by Experimental Condition

Experimental Condition	Number of Days	Mean Awakenings	Standard Error
Baseline Control	505	12.93	0.44
Control	680	12.41	0.33
Baseline Test	1,604	11.79	0.24
Test	3,041	11.37	0.17
Overall	5,830	11.74	0.12

Closer inspection reveals that drivers had an average of five SSPs per O-O interval lasting longer than 20 min (Table 18). An MLM analysis of the number of SSPs lasting longer than 20 min did not reveal statistically significant differences between the baseline control and control conditions ($t(82) = 0.27$, $p = 0.79$). A similar analysis was used to investigate possible differences between the baseline test and test conditions. Again, statistically significant differences were not found ($t(82) = 0.76$, $p = 0.45$).

Table 18. Mean Number of SSPs Longer than 20 min in O-O Interval by Experimental Condition

Experimental Condition	Number of Days	Mean #SSPs Longer than 20 min	Standard Error
Baseline Control	505	5.25	0.14
Control	680	5.32	0.12
Baseline Test	1,604	4.84	0.07
Test	3,041	4.67	0.05
Overall	5,830	4.85	0.04

The average longest SSP per O-O interval was found to be 136 min (Table 19). An MLM analysis of sleep quantity did not reveal statistically significant differences between the baseline control and control conditions ($t(82) = -0.66$, $p = 0.51$). A similar analysis was used to investigate possible differences between the baseline test and test conditions. This time, however, a statistically significant difference was found ($t(82) = -2.05$, $p = 0.04$). On average, the longest SSP of drivers in the test condition was 9.8 min longer than drivers in the baseline test condition.

Table 19. Mean Longest SSP in an O-O Interval by Experimental Condition

Experimental Condition	Number of Days	Mean Longest SSP (min)	Standard Error (min)
Baseline Control	505	124.95	4.50
Control	680	130.73	4.16
Baseline Test	1,604	131.89	2.23
Test	3,041	141.69	1.86
Overall	5,830	136.19	1.31

Discussion

The purpose of the sleep hygiene analysis was to compare drivers' sleep quality based on a defined sleep period across the different experimental conditions. The analysis performed was not able to find statistical or practical differences between the sleep quantity and quality of the experimental conditions in this study.

This study identified a subset of DFM alerts that were verified to correctly identify drowsy driving. However, the sleep analyses performed in this section go beyond the scope of these alerts and consider all the sleep the participants obtained over the course of their involvement in the DDWS FOT. Since 4.69 percent of the DFM alerts that were evoked correctly identified drowsiness, it is foreseeable that the myriad of false alarms had a greater effect on drivers' willingness to consider the DFM alerts. Hence, statistical differences were not found across time for the different experimental conditions. It is possible that drivers habituated to the DFM's invalid alerts over the time they were exposed to them. Therefore, a DDWS capable of accurately and reliably identifying drowsiness, and potentially even obtaining data from devices such as Actigraph watches, may stand to have a significant effect in driver alertness when this concept is further developed.

The analyses performed served to provide an overall sleep hygiene comparison for the experimental conditions. The reader is referred to Hanowski, Hickman, Fumero, Olson, and Dingus (2007) for a detailed sleep quantity analysis of this same data set. This previous research found that the mean sleep quantity per 24-hour period (midnight centered using the Cole-Kripke algorithm) for drivers was 6.15 hours ($SD = 1.36$). However, due to differences in analysis methods (specifically in data inclusion/exclusion criteria used for sleep quality analysis in the current study), each analysis should stand alone. Primarily, the sleep quantity scores reported include only sleep obtained within the O-O interval. This excludes short naps/micro-sleeps of under 20 min if they are not followed or preceded by another sleep period (i.e., other small naps or longer sleep period) as suggested in the methods for this research question. Although naps of 10 min or longer may have been beneficial to performance in terms of ameliorating drowsiness

(Buxton & Hartley, 2001), their exclusion in the present analysis served to reduce any uncontrolled variation in sleep quality measures specifically calculated for this study. This exclusion criteria was chosen because drivers would be unable to complete a sleep cycle in the given amount of time ($t < 20$ min; see Rechtschaffen & Siegel, 2000), thus allowing the analysis to obtain a higher degree of data accuracy specifically sought by Research Question 2.

There are other differences between the sleep data sets analyzed by Hanowski, Hickman, Fumero, Olson, and Dingus (2007) and the current study. First, Hanowski et al. (2007) report sleep data collected from 73 drivers, while the sleep of 16 extra drivers are analyzed in the current study. Secondly, sleep anomalies (such as when the driver did not sleep at all during a 24-hour period), known vacation periods (prolonged periods off-duty), and inaccurate data (when the Actigraph watch was removed and sleep or wake was indeterminable) were treated differently based on the objective of the research questions that these two different research efforts were answering. It is noteworthy that a current study using naturalistic driving data from 100 long-haul and line-haul drivers (Blanco et al., 2008) and the algorithm delineated in this research question found that overall sleep of CMV drivers to be 6.6 hrs.

RESEARCH QUESTION 3: INVOLVEMENT IN SAFETY CRITICAL EVENTS

As part of this research question (RQ) two sub-questions were addressed:

> RQ 3.1: Does a DDWS affect involvement in safety critical events (SCEs)?
> RQ 3.2: Does a DDWS affect involvement in at-fault SCEs?

Methods, analysis, measures of interest, results, and discussion related to these two questions are presented next.

Methods

The objective of Research Question 3 was to determine if the deployment of a DDWS will result in safety benefits, in the form of a reduction of driver involvement in SCEs. Estimation of the DDWS safety benefits involved comparing SCEs during the different experimental conditions. In addressing this question three data sets were of interest: all SCEs, SCEs within the DFM operating envelope, and SCEs when the truck driver was at fault. The safety benefits model presented in Chapter 1 predicted that an effective DDWS will produce a positive change in driver behavior after that driver obtains feedback. This will, in turn, result in a reduction of SCEs. The general method of estimating these potential benefits includes calculating the SCE frequency and rates by type of SCE, experimental condition, and whether the DFM was operational or not.

The analysis for the current research question used a data set composed of three types of SCEs:

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

Determining an SCE and its characteristics was performed as part of previous research efforts. For details on how the SCEs were determined, and the data reduction process performed to gather all the information needed to characterize these events, refer to Hickman et al. (2005) and Wiegand, Hanowski, Olson, and Melvin (2008). However, it should be noted that the much more demanding and illustrative SCE of crashes and near-crashes are the focus of this analysis.

Operating Envelope

The DFM had two different operating modes: Active and Dark. While the system was in active mode, the driver received feedback from the DFM and could interact with the system and its various control features. The system went into standby mode, in which the DFM gave no feedback, depending on ambient illumination and vehicle speed. Thus, for this research question, the operating envelope of interest was defined as those times when the DFM was not in standby. The SCEs within the DFM's operating envelope were of greatest interest.

Exposure Measure

In evaluating SCEs, finding a reasonable measure of exposure is very important especially if some characteristics of the SCEs are related not only to environmental and vehicle characteristics but also to driver-related factors. Multiple exposure measures were considered in this study, including the number of calendar days of participation, number of days or weeks the driver participated, and the number of hours driven for each experimental condition. The number of hours driven for each experimental condition is perhaps the most appropriate measure of exposure with which to evaluate SCEs. Hours driven has been used in the past for other SCE evaluations (Hanowski, Olson, Bocanegra, & Hickman, 2007), and it is a common exposure measure when evaluating drowsiness, fatigue, and other similar topics of interest.

The hours driven for each driver, over a 24-hour period, were computed for this research question. This calculation was based on the driving data files for the 96 participants included in this analysis and took into consideration each experimental condition start and end date. Table 20 presents a summary of hours driven under each experimental condition. If the system was in standby (i.e., outside of the operating envelope), it did not provide alerts. Table 21 presents the hours driven for each experimental condition when the DFM was within the operating envelope (not in standby). Note that only the alerts in the test condition were actually presented to the driver. The alerts in the baselines and the control condition were used as a comparison, but were not presented to the drivers.

Table 20. Total Hours Driven by Experimental Condition

Experimental Condition	Hours Driven and Percentage of Total
Baseline Test	10,292 (26%)
Test	20,291 (52%)
Baseline Control	4,035 (10%)
Control	4,605 (12%)
Total	39,223

Table 21. Hours Driven by Experimental Condition Within the Operating Envelope

Experimental Condition	Hours Driven and Percentage of Total
Baseline Test	3,501 (25%)
Test	7,289 (52%)
Baseline Control	1,532 (11%)
Control	1,614 (12%)
Total	13,936

In order to compare the SCE under the different experimental conditions, an SCE rate per hours driven was calculated. The SCE rate (units: rate per 100 hours driven) was calculated as:

$$\text{SCE Rate} = 100 \times (\text{SCE}/\text{H})$$

where SCE is the number of safety critical events and H represents the number of hours either within or outside the operating envelope for the experimental condition of interest.

Drowsiness Measures

Two surrogate measures of driver drowsiness were used for this research question: manual PERCLOS and a subjective driver behavior evaluation. Manual PERCLOS values, as explained in Research Question 1, were calculated based on video data reduction for different studies and included manual PERCLOS for SCEs and baseline events (Wiegand, Hanowski, Olson, & Melvin, 2008). Manual PERCLOS was used for this analysis instead of DMF PERCLOS given that the SCE could happen at any point in time (e.g., high illumination) and the DFM PERCLOS was not operational if the speed and illuminance thresholds were not met. Therefore, in order to be consistent across SCEs, manual PERCLOS was used for this analysis. However, it is important to note that SCE did not only occur within the operational envelope (specifically the speed and illumination requirements) of the prototype DFM system. Using DFM PERCLOS as the measure of identifying drowsiness in SCEs would exclude a significant portion of all SCEs, thus skewing any subsequent analysis. Therefore, all SCE were examined using the manually calculated PERCLOS measures.

The second measure of interest was a subjective assessment of driver behavior. This measure was obtained during the data reduction process performed to characterize each of the SCEs (Hickman et al., 2005). The researchers performing the data reduction coded up to four factors related to the experimental vehicle/driver believed to have relevance to the occurrence of the SCE. If more than four relevant factors were observed, the data reductionists were instructed to select what they believed to be the four most important in relation to the SCE.

Vehicle at Fault

As part of the data reduction process for SCEs, multiple characteristics were of interest. VTTI developed a data directory that takes into consideration aspects from multiple databases of interest (e.g., NHTSA’s General Estimate System and Fatal Analysis Reporting System, FMCSA’s Large Truck Crash Causation Study). This data directory included aspects of the SCE related to the vehicle, driver, and driving environment and roadway. Similar to the driver behavior assessment discussed above, the vehicle at fault is one of the items included in this directory. The SCE was attributed to the driver of the vehicle determined to be at fault. Although fault typically carries legal connotations, here the term is used only to indicate that the vehicle/driver was assigned the SCE. However, in regard to the issue of which vehicle/driver is predominantly at fault, the current study was limited by the fact that the vehicle instrumentation included tractor-mounted sensors (e.g., forward radar), but no trailer instrumentation (e.g., rearward radar). In addition, the dynamic sensor triggers used to capture events were based primarily on evasive maneuvers by the truck and would not flag events in which the only evasive maneuver was performed by the other vehicle. For example, events in which the truck made evasive maneuvers following (longitudinal or lateral) encroachment toward another vehicle were likely to be flagged and captured, but the opposite scenarios involving encroaching non-participant vehicles were unlikely to be captured.

Analysis

An analysis of variance was used to test the equality of means of the SCE rates of the different experimental conditions. A mixed factor design was used for this analysis. The repeated measures, or within-subject, portion of that type of ANOVA model is a robust set of statistical techniques used when all members of a sample are measured under a number of different conditions. As the sample is exposed to each condition, the measurement of the dependent variable is repeated. In cases where there is a great deal of variation between sample members, error in variance estimates from standard ANOVAs is large. Repeated measures of each sample member provide a way of accounting for this variance, thus reducing error variance.

For this analysis, the dependent variable used was the SCE rate. Drivers who did not experience any SCE were included in the analysis as well with a SCE rate of zero. The between-subjects factor portion was experimental group (i.e., Test and Control Groups). The baseline Reference denotes if the participant was within the baseline or not during a given SCE; this is the within-subject factor for the model. In the current mixed factor design used for the ANOVA, the effects of interest were the following:

- Between-subjects factor: experimental group (Levels: Test Group, Control Groups)
- Within-subject factor: baseline reference (Levels: during baseline, after baseline)
- Interactions: experimental conditions (Levels: baseline test, test, baseline control, control)

In addition to the ANOVAs, Chi-Square or Fisher tests were performed, as appropriate, to evaluate frequency of SCE at different levels.

RQ 3.1: Does a DDWS affect involvement in safety critical events?

Results – All Safety Critical Events

In order to evaluate the impact of the DDWS in involvement in SCEs, several analyses were performed taking into consideration all SCEs (i.e., within and outside the DFM's operating envelope). A total of 1,124 SCEs correspond to the 96 drivers who were considered in this study. A first look at the data included all the SCEs both within and outside of the DFM's operating envelope. A more detailed look focused on the SCE within the operating envelope. Figure 44 shows a summary of the distribution of SCEs by:

- Number of SCEs;
- Hours driven for that set of SCEs; and
- Event type: crash (C), near-crash (NC), and crash relevant conflict (CRC).

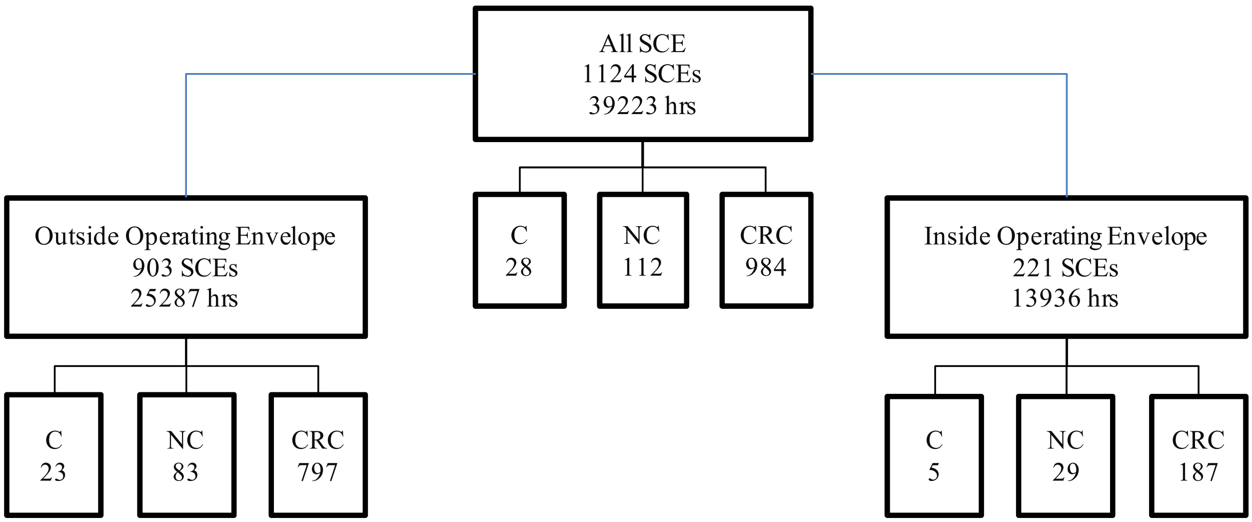


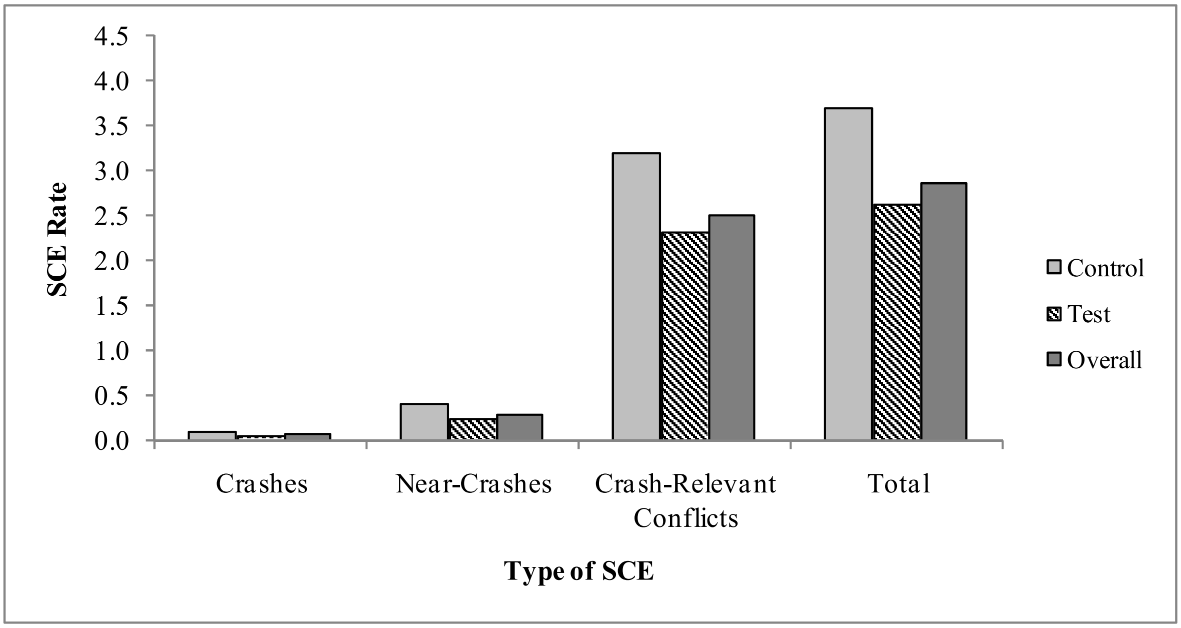
Figure 44. Distribution of SCEs by Operating Envelope Status and Type of Event – All SCEs

Overview of All Safety Critical Events by Experimental Group. The first set of results presented herein was focused on all the SCEs as they occurred in the two experimental groups. Table 22 shows the SCEs divided by the experimental group the driver was part of and the frequency of the three event types. The Control Group represents 28 percent (318/1,124) of all SCEs and 22 percent of the hours driven (as per Table 20, collapsing over the two control conditions). Crashes, near-crashes, and crash relevant conflicts represent approximately 3, 10, and 87 percent, respectively, of the SCEs for each experimental group. A Chi-square test indicated there was no statistical difference between the frequency of SCEs for the two experimental groups (χ^2 (2, N = 1,124) = 0.27, p = 0.88). These results also indicated that there was no statistically significant relationship between the experimental group and the Event Type.

Table 22. Frequency of SCE by Event Type and Experimental Group – All SCEs

Experimental Group	Crashes	Near-Crashes	Crash-Relevant Conflicts	Total
Control	8	34	276	318 (28%)
Test	20	78	708	806 (72%)
Total	28 (2%)	112 (10%)	984 (88%)	1,124

The SCE rate was computed for each experimental group as the number of SCEs per 100 hours driven (Figure 45). Even though the differences in the SCE rate per experimental group were not statistically different (discussed in SCE rate ANOVA results section below), the Test Group experienced the lowest SCE rates for all event types.



Overview of All Safety Critical Events by Experimental Condition. The SCEs were also evaluated by experimental condition (Figure 46). Of the 318 SCEs that occurred in the Control Group, 145 of them (46%) occurred during the control condition, which accounted for 53 percent of the hours driven for the drivers in that group. In the Test Group, 556 of the SCEs (69%) occurred in the test condition, and the hours driven during this period of time accounted for 66 percent of the total hours driven for that group. The homogeneity Chi-square test indicated no difference in the frequency of SCEs between the experimental conditions. These results also indicate no statistically significant relationship between the experimental condition and the event type (χ^2 (6, N = 1,124) = 6.6342, p = 0.3562).

The proportion of crashes to total SCEs for the Control Group ranged from 4 percent in the baseline control condition to 1 percent in the control condition. For the Test Group the proportion of crashes was much closer for the baseline test and test conditions, 3 and 2 percent, respectively.

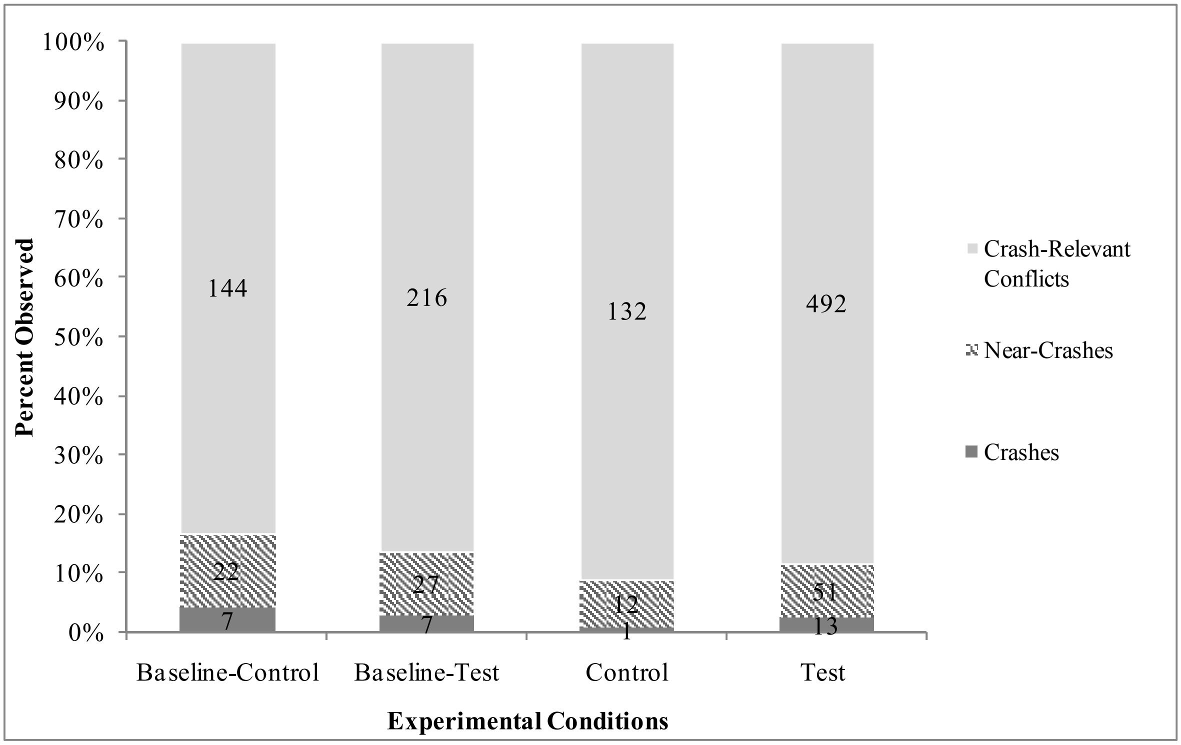


Figure 46. Frequencies and Percentage of SCE by Experimental Condition – All SCEs

Figure 47 compares the SCE rates as the number of SCEs per 100 hours driven by experimental condition. The rate of crashes for the Control Group varies from 0.17 for the baseline control (composed of seven crashes) to 0.02 for the control condition (composed of 1 crash). For the Test Group, the rates for the crashes and near-crashes are 5 percent lower for the test condition than for the baseline test condition.

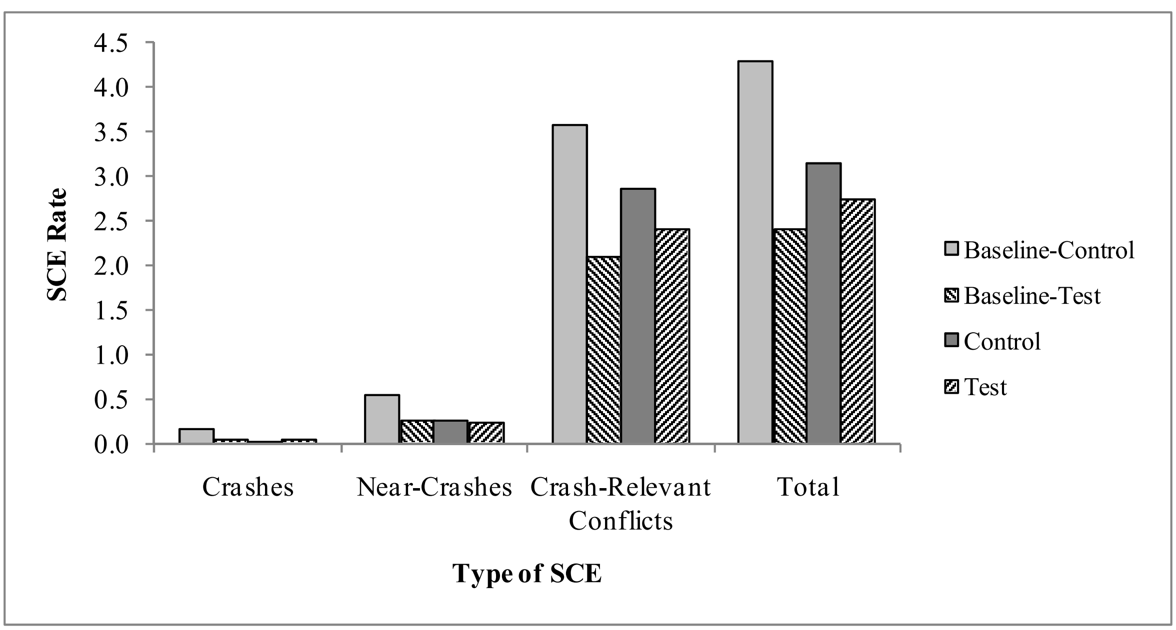


Figure 47. SCE Rate by Experimental Condition – All SCEs

Two surrogate measures of driver drowsiness were employed to characterize level of drowsiness during SCEs. The first measure evaluated was manual PERCLOS. The second was a subjective observation of drowsiness-related driving behavior prior to the SCE.

Manual PERCLOS for All Safety Critical Events. Of the 1,124 events, 355 SCEs were lacking sufficient video data to calculate manual PERCLOS (Wiegand, Hanowski, Olson, & Melvin, 2008). Therefore, 769 SCE were used in the descriptive statistics calculated to characterize the experimental conditions based on manual PERCLOS. The mean manual PERCLOS value for the SCEs was 5 with a maximum observed manual PERCLOS value of 26. Table 23 shows the descriptive statistics for the manual PERCLOS values at the moment of SCE occurrence by experimental condition. The mean values are higher for the test condition than for the control condition. However, this was the case for their respective baselines as well. The minimum for the Control and Test Group was zero. The maximum values ranged from 10 to 26, where the baseline control had the lowest value and the baseline test the highest. Mean and maximum manual PERCLOS values for the SCE in each experimental condition were the same or lower than the manual PERCLOS values presented for the baseline events analyzed as part of Research Question 1.

Table 23. Manual PERCLOS Value for the SCE by Experimental Condition – All SCEs

Experimental Condition	Total SCE	SCE With Manual PERCLOS Available	Minimum	Maximum	Mean
Baseline Control	173	100	0	10	3
Baseline Test	250	185	0	25	6
Control	145	95	0	18	3
Test	556	389	0	26	5
Total	1124	769	-	-	-

Table 24 presents the frequency of SCEs with manual PERCLOS values above 12 (drowsiness threshold used for this study) by experimental condition. Sixty-one percent (691/1,124) of SCEs had manual PERCLOS values below 12. Given that approximately 32 percent of the SCEs did not have manual PERCLOS available and the majority of the SCEs had a PERCLOS value below 12, no further statistical analyses were performed.

Table 24. SCEs With Manual PERCLOS Above 12 by Experimental Condition – All SCEs

Experimental Condition	PERCLOS <12	PERCLOS ≥12	PERCLOS N\A	Total
Baseline Control	100	0	73	173
Baseline Test	151	34	65	250
Control	92	3	50	145
Test	348	41	167	556
Total	691 (61%)	78 (7%)	355 (32%)	1124

Drowsiness-Related Driving Behavior Evaluation for All Safety Critical Events. As mentioned earlier, all SCEs were evaluated and driver behavior was categorized if the driver appeared drowsy, sleepy, asleep, fatigued, or showed signs of reduced alertness that occurred during the period of time leading to the SCE. Table 25 shows the frequency of drowsiness related SCEs by experimental condition. Of the 1,124 SCEs analysts identified, 143 (13%) involved a drowsiness-related behavior. However, there was no statistical difference among the experimental conditions based on the proportion of SCEs that were drowsiness versus non-drowsiness related (χ^2 (3, N = 1,124) = 0.85, p = 0.84).

Table 25. Frequency of Drowsiness-Related SCE by Experimental Condition – All SCEs

Experimental Condition	Not Drowsy	Drowsy	Total
Baseline Control	154	19	173 (15%)
Baseline Test	216	34	250 (22%)
Control	128	17	145 (13%)
Test	483	73	556 (50%)
Total	981 (87%)	143 (13%)	1,124

Table 26 shows the frequency of drowsiness-related SCEs by experimental condition, and Table 27 presents these values as an SCE rate per 100 hours driven. The percentage of near-crashes for the baseline test condition is 26 percent (7/27) compared to 20 percent (10/51) for the test condition; however, this reduction is not statistically significant (χ^2 (1, N = 708) = 0.41, p = 0.51). The total SCE rates that are drowsiness related for the baseline test and test conditions remains relatively unchanged, with values of 0.33 and 0.36, respectively, per 100 hours driven. Additionally, the test condition presented the lowest rate for the near-crashes. The percentage of crash relevant conflicts remained virtually the same.

Table 26. Drowsiness Related SCE by Experimental Condition and Event Type – All SCEs

Experimental Condition	Drowsiness Related	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Baseline Control	No	6	21	127	154
Baseline Control	Yes	1	1	17	19
Baseline Control	Total	7	22	144	173
Control	No	1	10	117	128
Control	Yes	0	2	15	17
Control	Total	1	12	132	145
Baseline Test	No	7	20	189	216
Baseline Test	Yes	0	7	27	34
Baseline Test	Total	7	27	216	250
Test	No	11	41	431	483
Test	Yes	2	10	61	73
Test	Total	13	51	492	556

Table 27. SCE Rate by 100 Hours Driven for Drowsiness-Related SCE by Experimental Condition and Event Type – All SCEs

Experimental Condition	Drowsiness Related	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Baseline Control	No	0.15	0.52	3.15	3.82
Baseline Control	Yes	0.02	0.02	0.42	0.47
Baseline Control	Total	0.17	0.55	3.57	4.29
Control	No	0.02	0.22	2.54	2.78
Control	Yes	0.00	0.04	0.33	0.37
Control	Total	0.02	0.26	2.87	3.15
Baseline Test	No	0.07	0.19	1.84	2.10
Baseline Test	Yes	0.00	0.07	0.26	0.33
Baseline Test	Total	0.07	0.26	2.10	2.43
Test	No	0.05	0.20	2.12	2.38
Test	Yes	0.01	0.05	0.30	0.36
Test	Total	0.06	0.25	2.42	2.74

Driver at Fault. The proportion of drivers at fault is displayed in Figure 48 by experimental condition. The driver of the experimental vehicle (V1) was judged to be at fault between 71 and 76 percent of the time. Results showed that there were no statistically significant differences in the distribution of SCEs for the different experimental conditions (χ^2 (6, N = 1,124) = 3.7573, p = 0.7095). The 836 SCEs in which V1 was at fault were used to address Research Question 3.2.

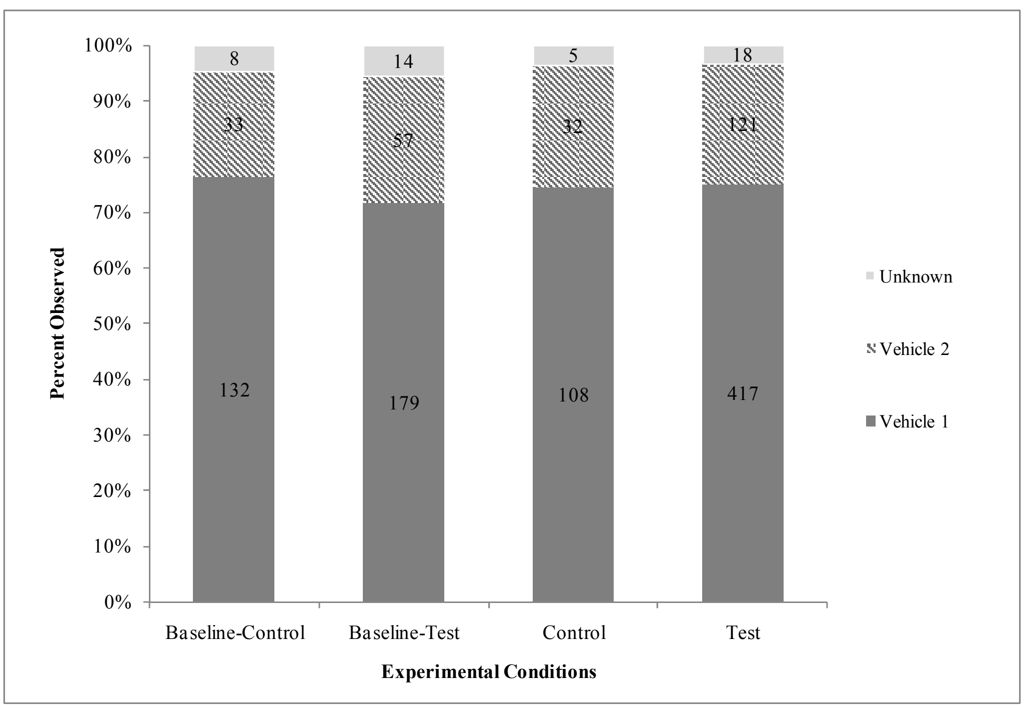


Figure 48. Vehicle at Fault by Test Condition

Safety Critical Event Rate ANOVA Results. A mixed factor design was used for the ANOVA to test the equality of SCE rate means for the different experimental conditions (Table 28). Based on the results presented in the ANOVA summary table, there were no statistically significant differences in terms of SCE rate for the experimental group or baseline Reference. However, the interaction showed a statistical difference. Table 29 presents the SCE rates per experimental condition (i.e., interaction). The SCE rate for the baseline control condition was higher than the baseline test, but that was not the case when compared to the other experimental conditions.

Table 28. ANOVA Summary Table – All SCEs

Source	DF	SS	MS	F -Value	Pr > F
Between Subjects					
Experimental Group (EG)	1	35.50	35.50	1.75	0.19
Participant/EG	83	1678.91	20.23		
Within Subject					
Baseline Reference (BR)	1	0.41	0.41	0.10	0.75
BR x EG	1	16.11	16.11	3.98	0.05
BR x Participant/EG	83	336.13	4.05		

Table 29. Descriptive Statistics for SCE Rate per Experimental Condition – All SCEs

Experimental Condition	Drivers (N)	Minimum	Maximum	Mean	SE
Baseline Control	21	0	20.6	4.2	1.1
Baseline Test	64	0	14.1	2.4	0.3
Control	21	0	13.5	3.4	0.8
Test	64	0	17.5	3.1	0.5

Results – Safety Critical Events Within the Operating Envelope

In order to better evaluate the impact of the DFM in the involvement of SCEs, several analyses were performed taking into consideration SCEs within the DFM's operating envelope. Of the 1,124 SCEs, 221 events (approximately 20%) occurred within the operating envelope.
Figure 49 shows a summary of the distribution of SCEs by:

- Number of SCEs;
- Hours driven for that set of SCEs;
- Event type: crash (C), near-crash (NC), and crash relevant conflict (CRC);
- Experimental Group: control and test; and
- Experimental Condition: baseline control, control, baseline test, and test.

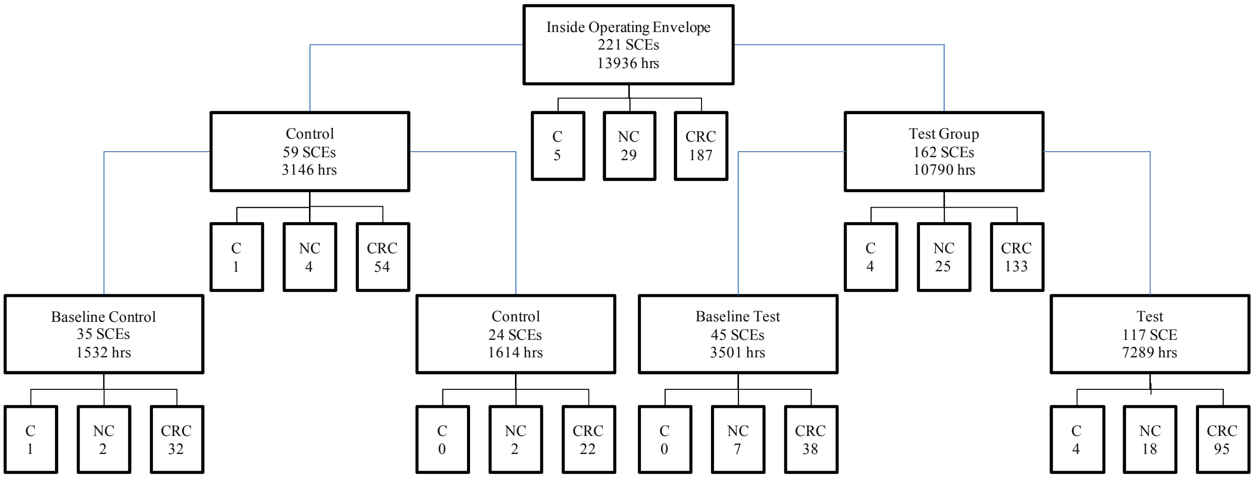


Figure 49. Distribution of SCEs by Experimental Condition and Event Type – Within Operating Envelope

Overview of Safety Critical Events Within the Operating Envelope by Experimental Group. The first set of results presented herein was focused on SCEs that occurred in the two experimental groups. Table 30 shows the distributions of events for the Test and Control Groups. The Control Group represents 27 percent of all SCEs and 23 percent of the hours driven (as shown in Table 21). Chi-square test results indicated that both groups have the same distribution of SCEs ($\chi 2$ (2, N = 221) =3, p = 0.21). These results also indicated that there was no statistically significant relationship between the experimental group and the event type.

Table 30. Frequency of SCE by Event Type and Experimental Group – Within Operating Envelope

Experimental Group	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Control	1	4	54	59 (27%)
Test	4	25	133	162 (73%)
Total	5	29	187	221

The SCE rate was computed for each experimental group as the number of SCEs per 100 hours driven (Figure 50). There were no statistically significant differences in SCE rate per experimental group (discussed in SCE Rate ANOVA Results section below).

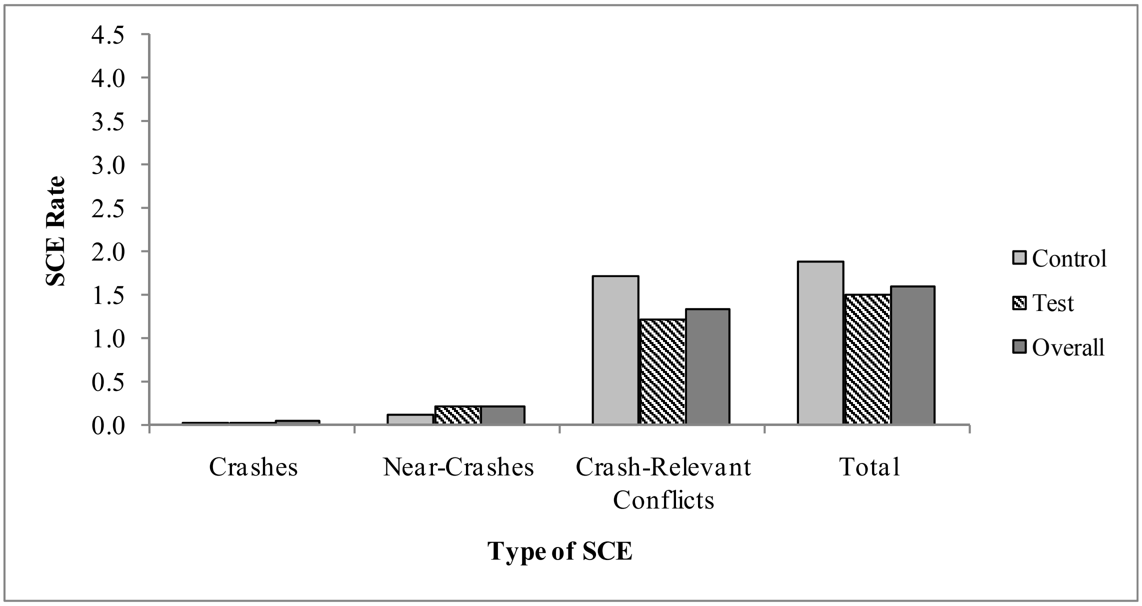


Figure 50. SCE Rates per 100 Hours Driven by Event Type – Within Operating Envelope

Overview of Safety Critical Events Within the Operating Envelope by Experimental Condition. The SCEs were also evaluated by experimental condition (Figure 51). Of the 59 SCEs that occurred in the Control Group, 24 (40%) occurred in the control condition, which accounted for 51 percent of the hours driven for the drivers in that group. In the Test Group, 117 of the SCEs (72%) occurred in the test condition, which accounted for 67 percent of the hours driven for the drivers in that group. Results showed that there was no statistical difference in the distribution of events among the various experimental conditions (Fischer's exact test $p = 0.59$)

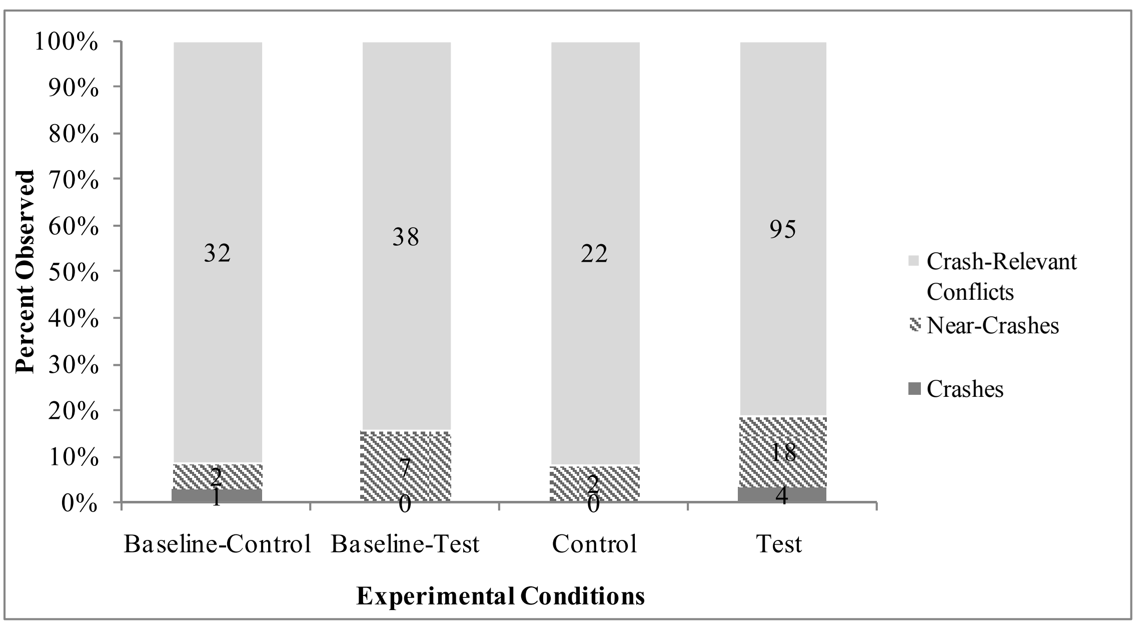


Figure 51. Frequencies and Percentage of SCE by Experimental Condition – Within Operating Envelope

Figure 52 compares the SCE rates as the number of SCEs per 100 hours driven by experimental condition where the DFM was not in standby mode. There was no statistically significant difference between experimental conditions (discussed in sections below).

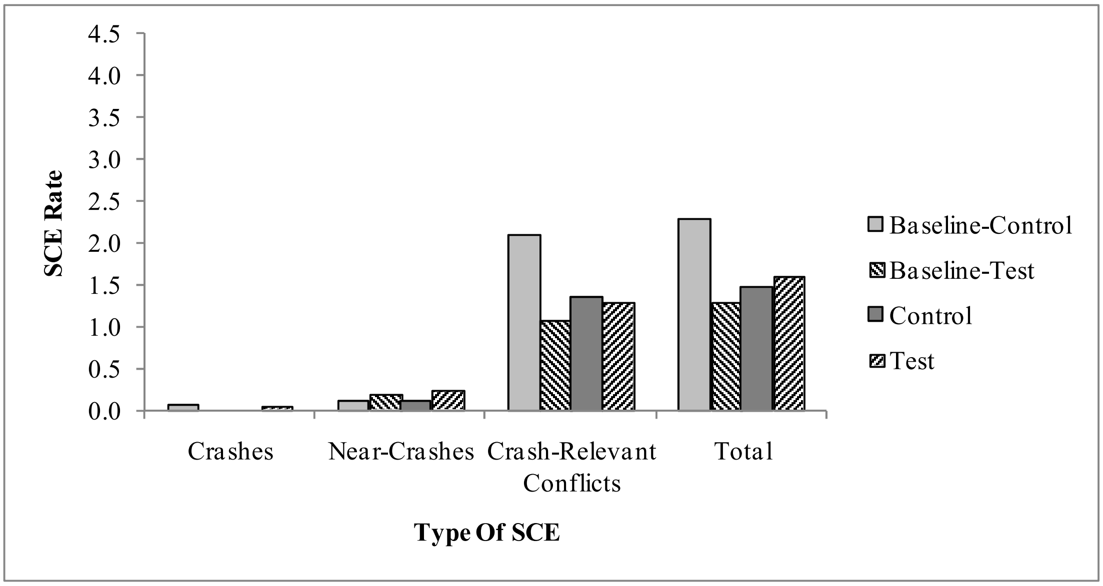


Figure 52. SCE Rate by Experimental Condition – Within Operating Envelope

Manual PERCLOS of Safety Critical Events Within the Operating Envelope. Two surrogate measures of driver drowsiness were employed to characterize level of drowsiness during SCEs. The first measure evaluated was the manual PERCLOS. Of the 221 events, 42 SCEs (19%) were lacking sufficient video data to calculate manual PERCLOS (Wiegand, Hanowski, Olson, & Melvin, 2008). Therefore, 179 SCEs were used in the descriptive statistics calculated to characterize the experimental conditions based on manual PERCLOS (Table 31). The manual PERCLOS mean for SCEs within the operating envelope was 7 (34 percent higher than the collapsed score for all SCEs), and its maximum value was 26. The maximum values ranged from 10 to 26, where the baseline control had the lowest value and the baseline test had the highest. Mean and maximum manual PERCLOS values for the SCE in each experimental condition were the same or lower than the manual PERCLOS values presented for the baseline events analyzed as part of Research Question 1. The means for the Control Group were lower than for the Test Group. However, the results suggested that there was no correlation between the experimental condition and manual PERCLOS value (Fischer's exact test, $p = 0.5136$)

Table 31. Manual PERCLOS Value for the SCEs in Each Experimental Condition – Within Operating Envelope

Experimental Condition	Total SCE	SCE With Manual PERCLOS Available	Minimum	Maximum	Mean
Baseline Control	35	25	1	10	3
Baseline Test	45	37	1	24	8
Control	24	19	1	17	5
Test	117	98	0	26	7
Total	221	179	-	-	-

Table 32 presents the frequency of SCEs with manual PERCLOS values above 12 (drowsiness threshold used for this study) by experimental condition. Results indicated that 68 percent (150/221) of SCEs had PERCLOS values below 12. Given that 19 percent of the SCEs did not have manual PERCLOS available, and the majority of the SCEs had a value under 12, no further statistical analysis was performed.

Table 32. SCEs With Manual PERCLOS Above 12 by Experimental Condition – Within Operating Envelope

Experimental Condition	PERCLOS <12	PERCLOS ≥12	PERCLOS N/A	Total
Baseline Control	25	0	10	35
Baseline Test	26	11	8	45
Control	18	1	5	24
Test	81	17	19	117
Total	150 (68%)	29 (13%)	42 (19%)	221

Drowsiness-Related Driving Behavior Evaluation of Safety Critical Events Within the Operating Envelope. As mentioned earlier, all SCEs were evaluated and driver behavior was categorized if the driver appeared drowsy, sleepy, asleep, fatigued, or showed signs of reduced alertness that occurred during the period of time leading to the SCE. Table 33 shows the frequency of drowsiness-related SCEs by experimental condition. Of the 221 events, analysts identified 77 (35%) in which the driver exhibited a drowsiness-related behavior. This is 22 percent more than when all SCEs were considered. However, Chi-square test results showed that the distribution of SCEs by experimental condition of drowsiness-related and not drowsiness-related SCEs was not statistically significant (χ^2 (3, N = 221) = 1.03, p = 0.79). For the test condition, only 41 events were considered drowsiness-related. Of these 41 events, 2 were crashes, 7 were near-crashes, and 32 were classified as crash relevant conflicts.

Table 33. Frequency of Drowsiness-Related SCEs by Experimental Condition – Within Operating Envelope

Experimental Condition	Not Drowsy	Drowsy	Total
Baseline Control	24	11	35
Baseline Test	27	18	45
Control	17	7	24
Test	76	41	117
Total	144 (65%)	77 (35%)	221

Table 34 shows the frequency of drowsiness-related SCEs by experimental condition, and Table 35 presents these values as an SCE rate per 100 hours driven. The drowsiness-related SCE rates for the baseline test and test conditions remained relatively unchanged, with values of 0.51 and 0.56 per 100 hours driven, respectively.

Table 34. Drowsiness-Related SCEs by Experimental Condition and Event Type – Within Operating Envelope

Experimental Condition	Drowsiness-Related	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Baseline Control	No	0	2	22	24
Baseline Control	Yes	1	0	10	11
Baseline Control	Total	1	2	32	35
Control	No	0	1	16	17
Control	Yes	0	1	6	7
Control	Total	0	2	22	24
Baseline Test	No	0	3	24	27
Baseline Test	Yes	0	4	14	18
Baseline Test	Total	0	7	38	45
Test	No	2	11	63	76
Test	Yes	2	7	32	41
Test	Total	4	18	95	117

Table 35. Drowsiness Related SCEs by Experimental Condition and Event Type – Within Operating Envelope

Experimental Condition	Drowsiness Related	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Baseline Control	No	0.00	0.13	1.44	1.57
Baseline Control	Yes	0.07	0.00	0.65	0.72
Baseline Control	Total	0.07	0.13	2.09	2.29
Control	No	0.00	0.06	0.99	1.05
Control	Yes	0.00	0.06	0.37	0.43
Control	Total	0.00	0.12	1.36	1.49
Baseline Test	No	0.00	0.09	0.69	0.77
Baseline Test	Yes	0.00	0.11	0.40	0.51
Baseline Test	Total	0.00	0.20	1.09	1.29
Test	No	0.03	0.15	0.86	1.04
Test	Yes	0.03	0.10	0.44	0.56
Test	Total	0.05	0.25	1.30	1.61

Number of Vehicles Involved in Safety Critical Events Within the Operating Envelope. Table 36 presents the number of vehicles involved in SCEs by experimental condition. Most of the SCEs within the operating envelope involved two vehicles (54%), compared to 67 to -80 percent if all SCEs were considered. The percentage of SCEs within the operating envelope that involved only one vehicle (i.e., experimental vehicle) was 34 percent, with higher percentages for the baseline test and the Test conditions. The number of events that involved the experimental vehicle plus a pedestrian, cyclist, or animal accounted for 11 percent, compared with 3 percent when all the SCEs were considered. Kendall's Tau-b value (τ_b = -0.0438) and Spearman correlation (ρ = -0.0048) suggest that the number of vehicles involved was not associated with the experimental condition.

Table 36. Number of Vehicles Involved in SCEs by Experimental Condition – Within Operating Envelope

Experimental Condition	Single Vehicle	Two Vehicles	More than Two Vehicles	Pedestrian, Cyclist, or Animal	Total
Baseline Control	9	21	0	5	35
Baseline Test	19	21	0	5	45
Control	4	18	0	2	24
Test	43	60	2	12	117
Total	75 (34%)	120 (54%)	2 (1%)	24 (11%)	221

Driver at Fault. The proportion of drivers at fault is displayed in Figure 53 by experimental condition. The driver of the experimental vehicle (V1) was judged to be at fault in 65 to 70 percent of SCEs. Results show that there is not a statistically significant difference in the distribution of SCEs for the different experimental conditions ($\chi 2$ (6, N = 221) = 3.57, p = 0.0733). The 149 SCEs in which V1 was at fault was used to address Research Question 3.2.

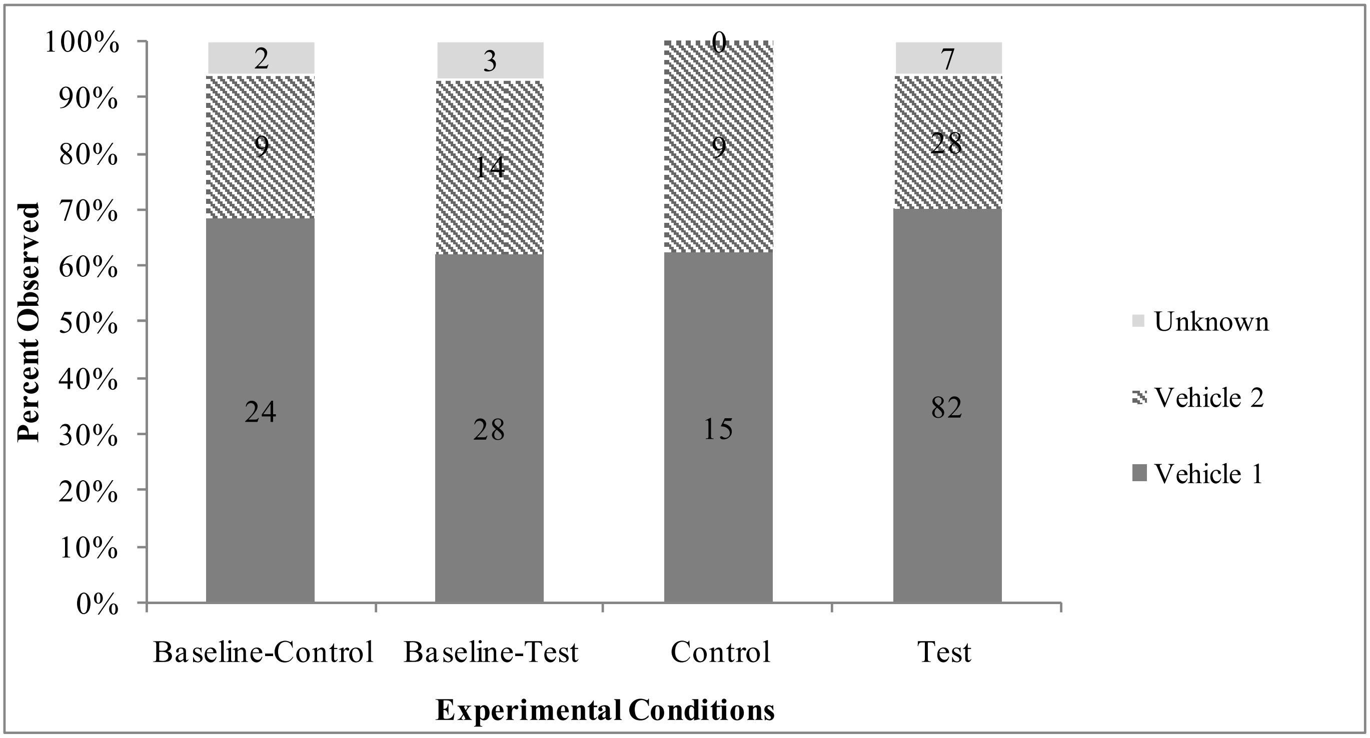


Figure 53. Number of Vehicles at Fault by Experimental Condition – Within Operating Envelope

Safety Critical Events Within the Operating Envelope that Received Alerts. The DFM recorded 300 alerts in the hour prior to the 221 SCEs. However, only 38 of those 221 SCEs involved the driver receiving alerts (33 received full alerts). Table 37 and Table 38 show the descriptive statistics for the alerts (number of SCE and number of alerts received). The percentage of SCEs with an alert, both initial and full alerts, was lower for the Test than the control condition.

Table 37. Descriptive Statistics for Number of Initial Alerts 1 Hour Before SCE by Experimental Condition –Within Operating Envelope

Experimental Condition	Total SCE	SCE That Received Alerts	Percentage of SCEs with Alerts	Minimum (No. Alerts)	Maximum (No. Alerts)	Mean (No. Alerts)
Baseline Control	35	4	11	1	18	9
Baseline Test	45	4	9	3	11	7
Control	24	8	33	2	56	16
Test	117	22	19	1	23	5
Total	221	38	-	-	-	-

Table 38. Descriptive Statistics of Full Alerts 1 Hour Before SCE by Experimental Condition –Within Operating Envelope

Experimental Condition	Total SCE	SCE That Received Alerts	Percentage of SCEs with Alerts	Minimum (No. Alerts)	Maximum (No. Alerts)	Mean (No. Alerts)
Baseline Control	35	4	11	1	17	9
Baseline Test	45	3	7	3	10	7
Control	24	7	29	2	56	18
Test	117	19	16	1	23	5
Total	221	33	-	-	-	-

Sleep Quantity Before Safety Critical Event for Safety Critical Events Within the Operating Envelope. The amount of sleep in the 24 hours prior to the SCE was calculated for all drivers with available actigraphy data. As all drivers did not have actigraphy data available due to device failures and drivers removing the device, only a subset of events had sleep data available during the period of interest. As discussed previously, the method of calculating sleep quantity was different than that found in other sleep studies and involved removing all nap periods that were less than 20 min (even though sleep periods of 20 min or less may be beneficial to immediate performance, they were removed from the present analysis to maximize the related sleep quality data). As such, the total sleep values provided a lower number than has been found in previous studies (Blanco et al., 2008) and other studies using this same data set (Hanowski et al., 2007). Nonetheless, based on those data and the method used, drivers in the baseline control condition slept an average of 5.3 hours prior to an SCE. This is very similar to the 5.2 hours of sleep obtained by drivers in the baseline test Condition prior to an SCE. Drivers in the control condition had the lowest average (4.9 hours). Drivers in the test condition had the highest average with 6.2 hours of sleep prior to an SCE (Table 39). Both the Spearman Correlation ($\rho = 0.15$) and the Kendall's Tau-b ($\tau_b = 0.12$) suggest that the sleep quantity before an SCE was not associated with the experimental condition. Similar results were found when evaluating the data by event type, suggesting that the amount slept in the 24 hours before an SCE was not associated with the type of SCE ($\rho = 0.04$), Kendall's Tau-b ($\tau_b = 0.04$).

Table 39. Descriptive Statistics for Sleep 24 Hours Before Safety Critical Events by Experimental Condition –Within Operating Envelope

Experimental Condition	Total SCE	SCE With 24-h Actigraphy Data	Minimum (hrs)	Maximum (hrs)	Mean (hrs)
Baseline Control	35	23	1.0	10.4	5.3
Baseline Test	45	36	0.8	13.6	5.2
Control	24	19	0.8	8.5	5.0
Test	117	92	0.6	16.0	6.2
Total	221	170	-	-	-

SCE Rate ANOVA Results. A mixed factor design was used for the ANOVA to test the equality of SCE rate means for the different experimental conditions (Table 40). Based on the results presented in the ANOVA summary table, there were no statistically significant differences found in terms of SCE rate for the main effects or interaction.

Table 40. ANOVA Summary Table – Within the Operating Envelope

Source	DF	SS	MS	F -Value	Pr > F
Between Subjects					
Experimental Group (EG)	1	14.88	14.88	1.02	0.31
Participant/EG	82	1191.89	14.54		
Within Subject					
Baseline Reference (BR)	1	5.28	5.28	0.68	0.41
BR x EG	1	15.05	15.05	1.95	0.17
BR x Participant/EG	82	632.03	7.71		

RQ 3.2: Does a DDWS affect involvement in at-fault safety critical events?

Results – At-Fault Safety Critical Events

In order to evaluate the impact of the DDWS on involvement in SCEs, several analyses were performed taking into consideration the SCE, both within and outside the DFM's operating envelope, where the participant driver was at fault. A total of 836 SCEs were represented by the 96 drivers who were considered for this study. A first look at the data included all the SCEs that occurred both within and outside the DFM's operating envelope. A more detailed look focused on the SCEs within the operating envelope. Figure 54 shows a summary of the distribution of SCEs by:

- Number of SCEs;
- Hours driven during for that set of SCEs; and
- Event type: crash (C), near-crash (NC), and crash relevant conflict (CRC).

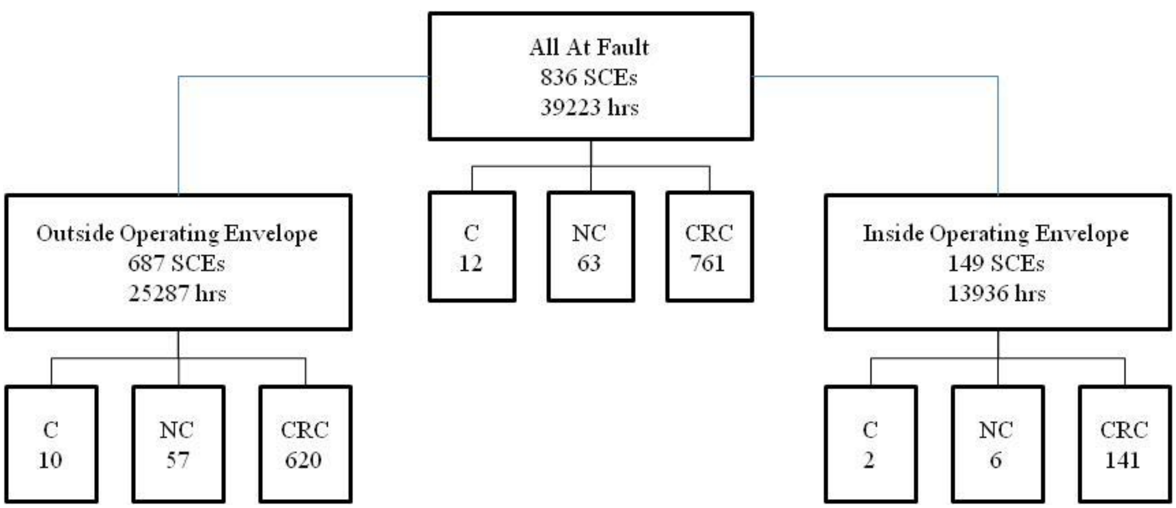


Figure 54. Distribution of SCEs by Operating Envelope Status and Type of Event – At-Fault SCEs

Overview of At-Fault Safety Critical Events by Experimental Group. The first set of results focused on the SCEs in which the participant driver was at fault. Of the 1,124 SCEs, 836 were associated with the driver of the experimental vehicle causing the precipitating factor for the SCE. These 836 SCEs consisted of 12 crashes, 63 near-crashes, and 761 crash relevant conflicts. Table 41 shows the SCEs as a function of the experimental group. Crashes, near-crashes, and crash relevant conflicts represent approximately 1, 8, and 91 percent, respectively, of the SCEs for the experimental groups. Results indicate that both experimental groups had the same distribution of SCEs (χ^2 (2, N = 836) = 0.38, p = 0.83). These results also indicated there was no statistically significant relationship between the experimental group and event type.

Table 41. Frequency of SCE by Event Type and Experimental Group – At-Fault SCEs

Experimental Condition	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Control	3	20	217	240
Test	9	43	544	596
Total	12 (1%)	63 (8%)	761 (91%)	836

The SCE rate was computed for each experimental group as the number of SCEs per 100 hours driven (Figure 55). Even though the differences in SCE rate per experimental group were not statistically different (discussed in the SCE Rate ANOVA Results section below), the Test Group experienced the lowest SCE rates for all event types.

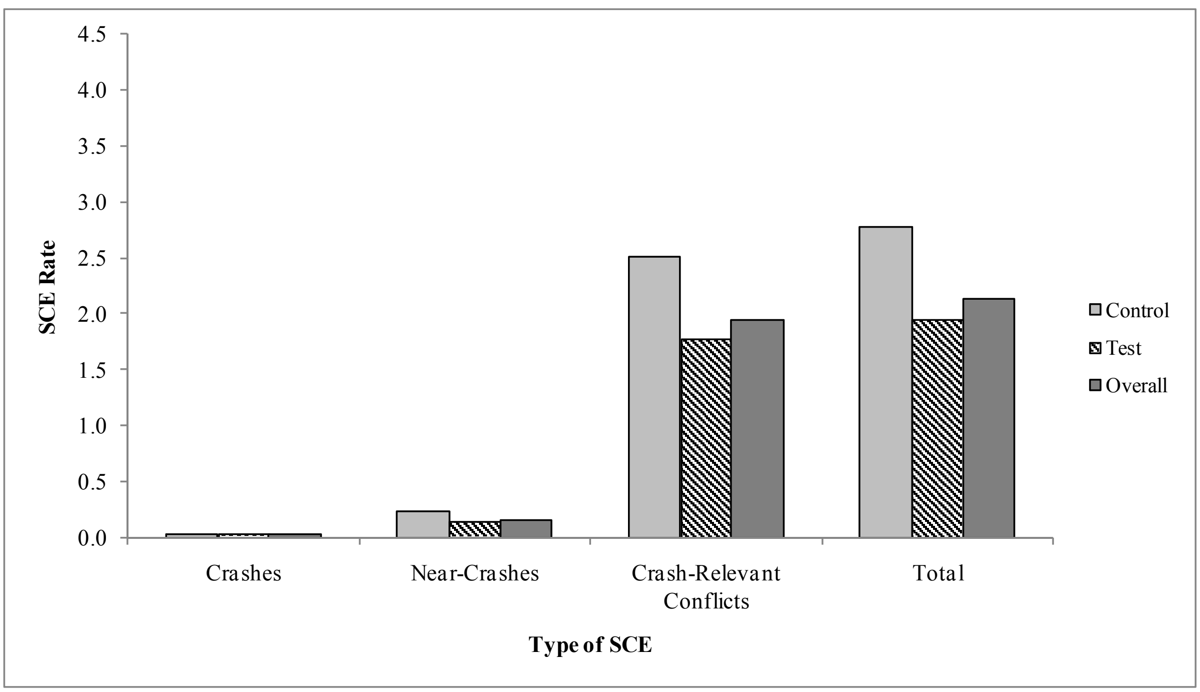


Figure 55. SCE Rates per 100 Hours Driven by Event Type – At-Fault SCEs

Overview of At-Fault Safety Critical Events by Experimental Condition. The SCEs were also evaluated by experimental condition (Figure 56). Of the 240 SCEs that occurred in the Control

Group, 108 SCEs (45%) occurred during the control condition, which accounted for 53 percent of the hours driven. In the Test Group, 417 of the SCEs (70%) occurred in the test condition, and the hours driven accounted for 67 percent of the total hours driven for that group. The homogeneity Chi-square test indicated no difference in the frequency of SCEs between the experimental conditions. These results also indicated no statistically significant relationship between the experimental condition and the event type (χ^2 (6, N = 836) = 4.91, p = 0.55).

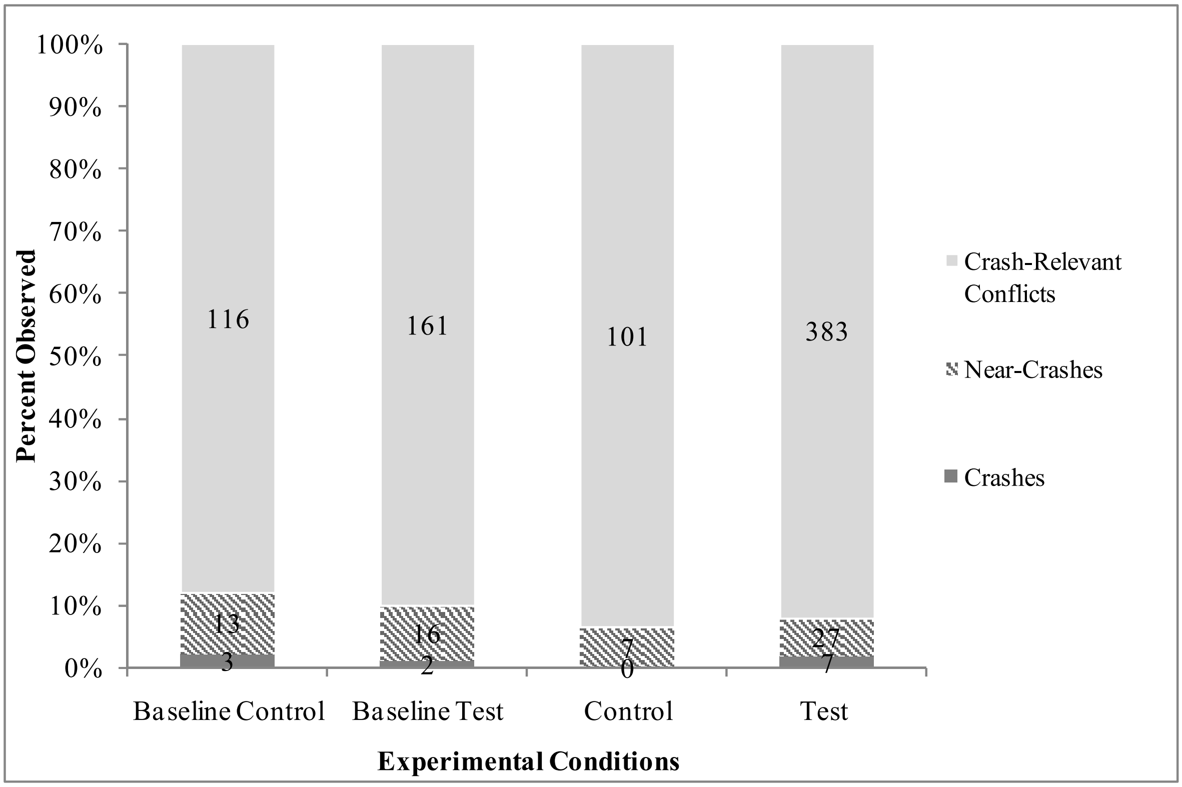


Figure 56. Frequency and Percentage of SCEs by Experimental Condition – At-Fault SCEs

Figure 57 compares the SCE rates looking at the number of SCEs per 100 hours driven by experimental condition. Even though there were no statistically significant differences in SCE rate per experimental group (discussed in the SCE Rate ANOVA Results section below), the test condition experienced the lowest SCE rates for near-crashes and crash relevant conflicts as compared to the control condition.

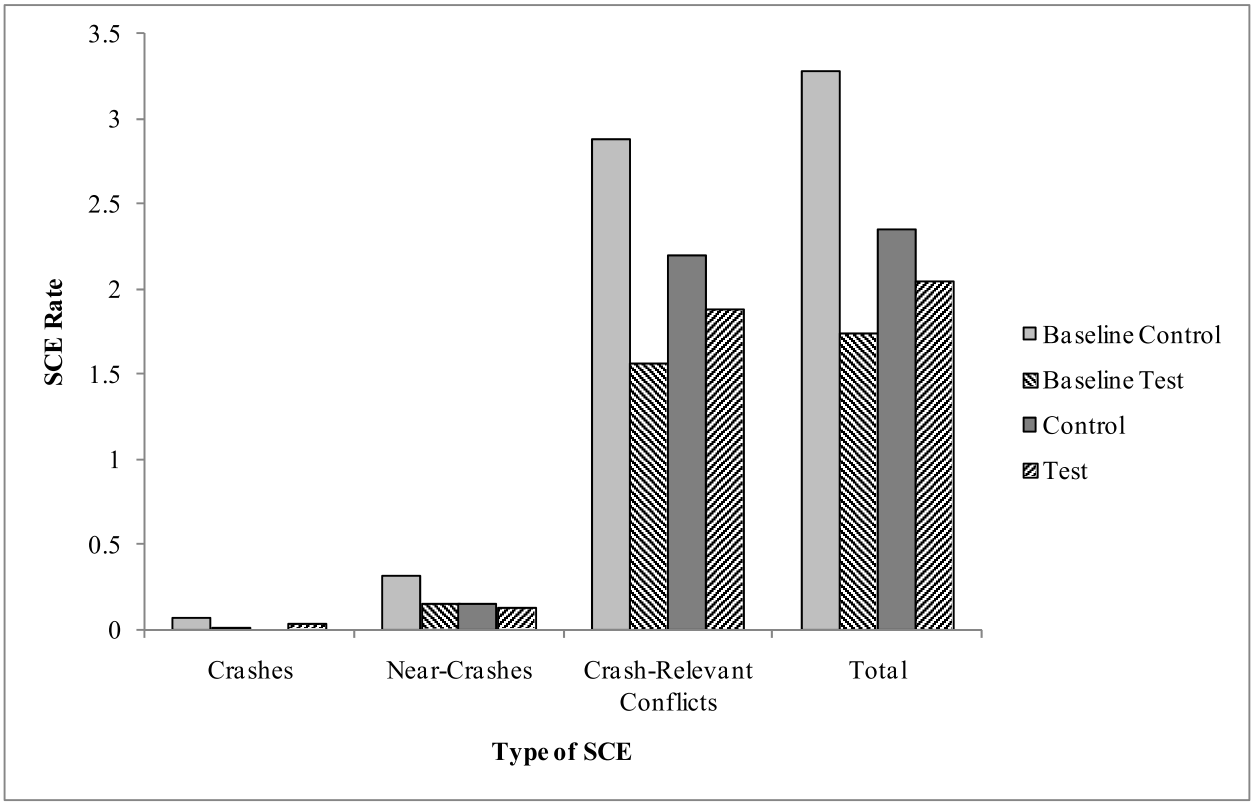


Figure 57. SCE Rate by Experimental Condition – At-Fault SCEs

Manual PERCLOS for At-Fault Safety Critical Events. Two surrogate measures of driver drowsiness were employed to characterize level of drowsiness during at-fault SCEs. The first measure evaluated was manual PERCLOS, and the second was the subjective rating of drowsiness related driving behavior. Of the 836 events, 268 SCEs were lacking sufficient video data to calculate manual PERCLOS. Therefore, 568 SCEs were used in the descriptive statistics calculated to characterize the experimental conditions based on manual PERCLOS.

Table 42 shows the descriptive statistics for the manual PERCLOS values at the moment of SCE occurrence as a function of experimental condition. The mean values were higher for the test condition than for the control condition. The minimum for the Control and Test Group was zero. The maximum values ranged from 10 to 26, where the baseline control had the lowest value and the Test the highest value. Mean and maximum manual PERCLOS values for the SCE in each experimental condition were the same or lower than the manual PERCLOS values presented for the baseline events analyzed as part of Research Question 1.

Table 42. Manual PERCLOS Values for the SCEs in Each Experimental Condition – At-Fault SCEs

Experimental Condition	Total SCE	SCEs With Manual PERCLOS Available	Minimum	Maximum	Mean
Baseline Control	132	83	0	10	3
Baseline Test	179	128	0	25	6
Control	108	74	0	18	3
Test	417	283	0	26	5
Total	836	568	-	-	-

Table 43 presents the frequency of SCEs with manual PERCLOS values above 12 (drowsiness threshold used for this study) by experimental condition. Sixty-two percent of SCEs have manual PERCLOS values below 12. Given that approximately 32 percent of the SCEs did not have manual PERCLOS available and the majority of the SCEs had a value below 12, no further statistical analyses were performed.

Table 43. SCEs With Manual PERCLOS Above 12 by Experimental Condition – At-Fault SCEs

Experimental Condition	PERCLOS <12	PERCLOS >12	PERCLOS N/A	Total
Baseline Control	83	0	49	132
Baseline Test	105	23	51	179
Control	73	1	34	108
Test	256	27	134	417
Total	517 (62%)	51 (6%)	268 (32%)	836

Drowsiness-Related Driving Behavior Evaluation for At-Fault Safety Critical Events. As mentioned earlier, all SCEs were evaluated and driver behavior was categorized if the driver appeared drowsy, sleepy, asleep, fatigued, or showed signs of reduced alertness that occurred during the period of time leading to the SCE. Table 44 shows the frequency of drowsiness-related SCEs by experimental condition. Of the 836 SCEs, analysts identified 98 events (12%) in which the driver exhibited a drowsiness-related behavior. However, there was no statistical difference among the experimental conditions based on the proportion of SCEs that were drowsiness-related versus not drowsiness-related (χ^2 (3, N = 836) = 1.8, p = 0.61).

Table 44. Frequency of Drowsiness-Related SCEs by Experimental Condition – At-Fault SCEs

Experimental Condition	Not Drowsy	Drowsy	Total
Baseline Control	118	14	132
Baseline Test	156	23	179
Control	99	9	108
Test	365	52	417
Total	738	98	836

Table 45 shows the frequency of drowsiness-related SCEs by experimental condition, and Table 46 presents these values as an SCE rate per 100 hours driven by experimental condition. The total SCEs for the Test Group is the higher; however, the SCE rate for the baseline test and test condition remains relatively unchanged with values of 0.20 and 0.26 per 100 hours driven, respectively, and it is lower than the one for the baseline control condition. The baseline control condition experienced the higher SCE rate.

Table 45. Drowsiness-Related SCEs by Experimental Condition and Event Type – At-Fault SCEs

Experimental Condition	Drowsiness Related	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Baseline Control	No	3	13	102	118
Baseline Control	Yes	0	0	14	14
Baseline Control	Total	3	13	116	132
Control	No	0	6	93	99
Control	Yes	0	1	8	9
Control	Total	0	7	101	108
Baseline Test	No	2	14	140	156
Baseline Test	Yes	0	2	21	23
Baseline Test	Total	2	16	161	179
Test	No	7	24	334	365
Test	Yes	0	3	49	52
Test	Total	7	27	383	417

Table 46. Drowsiness-Related SCEs by Experimental Condition and Event Type – At-Fault SCEs

Experimental Condition	Drowsiness Related	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Baseline Control	No	0.07	0.32	2.53	2.92
Baseline Control	Yes	0.00	0.00	0.35	0.35
Baseline Control	Total	0.07	0.32	2.87	3.27
Control	No	0.00	0.13	2.02	2.15
Control	Yes	0.00	0.02	0.17	0.20
Control	Total	0.00	0.15	2.19	2.35
Baseline Test	No	0.02	0.14	1.36	1.52
Baseline Test	Yes	0.00	0.02	0.20	0.22
Baseline Test	Total	0.02	0.16	1.56	1.74
Test	No	0.03	0.12	1.65	1.80
Test	Yes	0.00	0.01	0.24	0.26
Test	Total	0.03	0.13	1.89	2.06

SCE Rate ANOVA Results. A mixed factor design was used for the analysis of variance (ANOVA) to test the equality of SCE rate means for the different experimental conditions (Table 47). Of the 1,124 SCEs, the experimental vehicle's driver was determined to be at fault in 836. Based on the results presented in the ANOVA summary table, there were no statistically significant differences in terms of SCE rate for the experimental group or Baseline Reference. However, the interaction showed a borderline statistical difference. Table 48 presents the SCE rates per experimental condition (i.e., interaction). The SCE rate for the Baseline control condition was higher than the Baseline Test, but that was not the case when compared to the other Experimental Conditions.

Table 47. ANOVA Summary Table – At-Fault SCEs

Source	DF	SS	MS	F -Value	Pr > F
Between Subjects					
Experimental Group (EG)	1	22.42	22.42	1.61	0.21
Participant/EG	83	1157.18	13.94		
Within Subject					
Baseline Reference (BR)	1	0.01	0.01	0.01	0.94
BR x EG	1	9.65	9.65	3.93	0.05
BR x Participant/EG	83	203.94	2.46		

Table 48. Descriptive Statistics of SCE Rates by Experimental Condition – At-Fault SCEs

Experimental Condition	Drivers (N)	Minimum	Maximum	Mean	SE
Baseline Control	21	0.0	16.7	3.1	0.9
Baseline Test	64	0.0	11.0	1.7	0.3
Control	21	0.0	10.4	2.6	0.7
Test	64	0.0	15.5	2.3	0.4

Results – At-Fault Safety Critical Events Within the Operating Envelope

In order to better evaluate the impact of the DDWS in the involvement of SCEs, several analyses were performed taking into consideration the SCEs within the DFM's operating envelope. Of the 846 SCEs, 149 events (approximately17%) occurred within the operating envelope. Figure 58 shows a summary of the distribution of SCEs by:

- Number of SCEs;
- Hours driven for that set of SCEs;
- Event type: crash (C), near-crash (NC), and crash relevant conflict (CRC);
- Experimental Group: Control and Test; and
- Experimental Condition; Baseline Control, Control, Baseline Test, and Test.

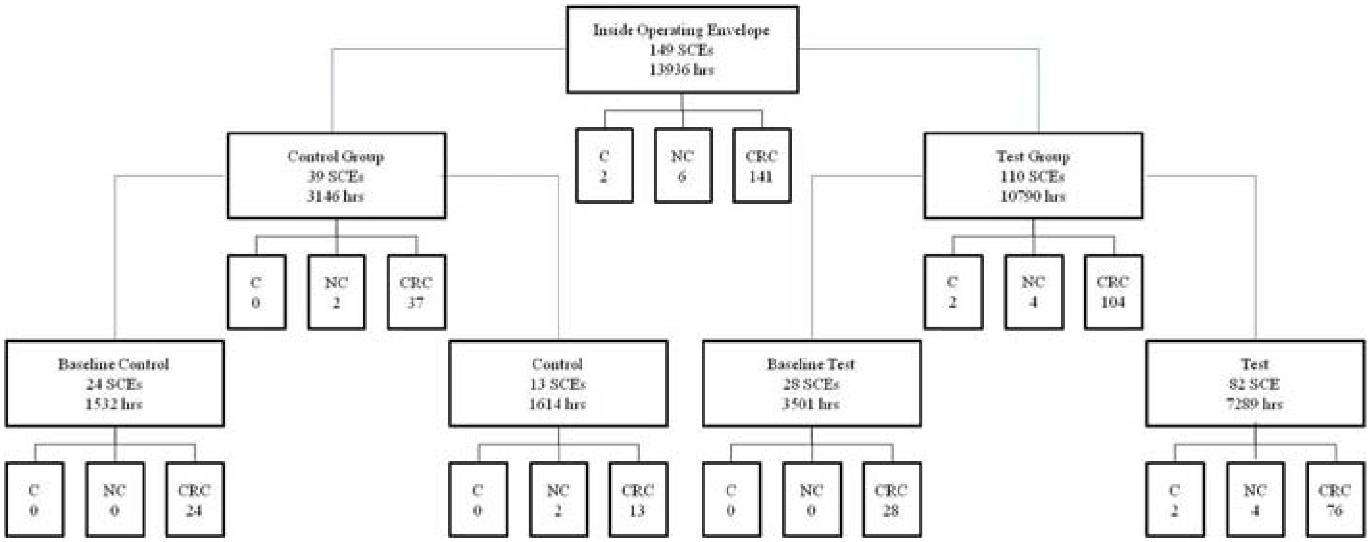


Figure 58. Distribution of SCEs by Experimental Condition and Type of Event – At-Fault Within the Operating Envelope

Overview of At-Fault Safety Critical Events Within the Operating Envelope by Experimental Group. The first set of results focused on SCEs as they occurred in the two experimental groups. Table 49 shows the distributions of events for the Test and Control Groups. The Control Group represents 26 percent of all SCEs and 22 percent of the hours driven. Chi-square test results indicated that both groups had the same distribution of SCEs, (Fischer test $p = 0.81$). These results also indicated that there was no statistically significant relationship between the experimental group and the event type.

Table 49. Frequency of SCEs by Event Type and Experimental Group – At-Fault Within the Operating Envelope

Experimental Condition	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Control	0	2	37	39 (26%)
Test	2	4	104	110 (74%)
Total	2	6	141	149

The SCE rate was computed for each experimental group as the number of SCEs per 100 hours driven (Figure 59). Even though the differences in SCE rate per experimental group are not statistically different (discussed in the SCE Rate ANOVA Results section below), the Test Group had the lowest SCE rates for all event types.

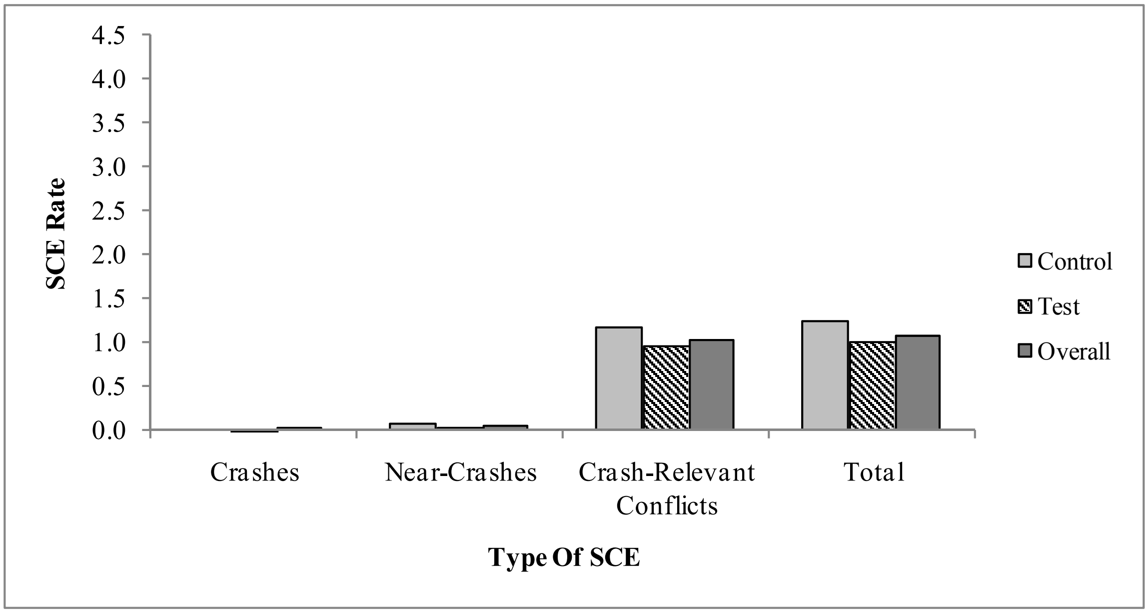


Figure 59. SCE Rates per 100 Hours Driven by Event Type – At-Fault Within the Operating Envelope

Overview of At-Fault Safety Critical Events Within the Operating Envelope by Experimental Condition. The SCEs were also evaluated by experimental condition (Figure 60). A total of 39 SCEs occurred in the Control Group, with 62 percent of those occurring in the baseline control condition. Of the 110 SCEs of the Test Group, only 28 (25%) occurred in the baseline test condition. It is important to note that there were no crashes or near-crashes in the baseline Control and baseline test conditions, and that the control condition only had two near-crashes. Test condition had two crashes and four near-crashes; however, the exposure for this condition was twice that of the baseline test and 4.5 times that of the control condition.

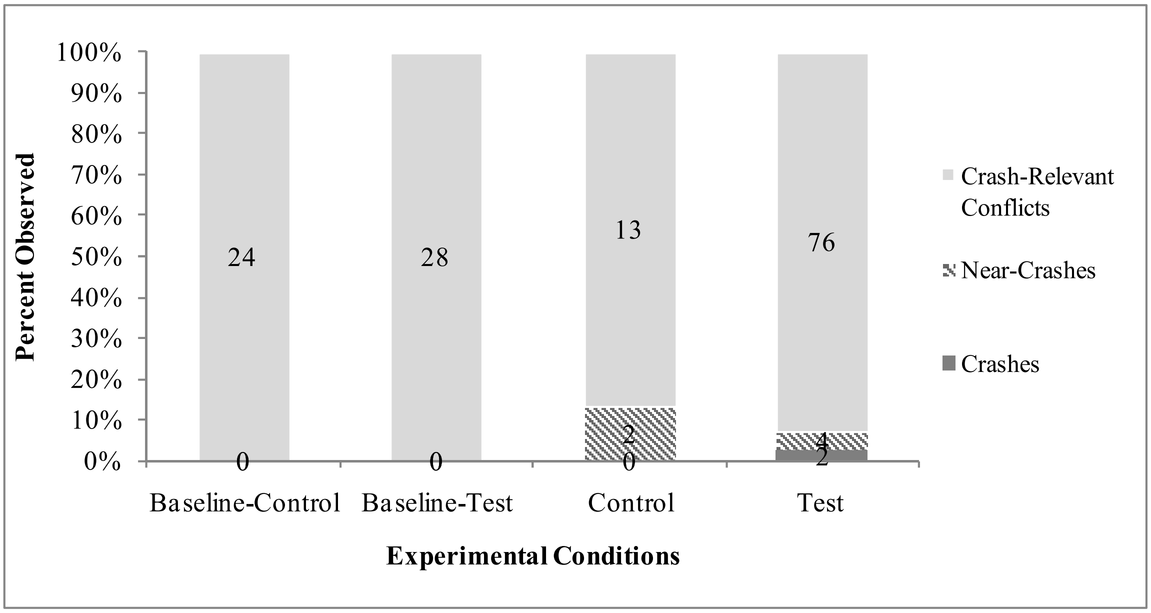


Figure 60. Frequency and Percentage of SCEs by Experimental Condition– At-Fault Within the Operating Envelope

Figure 61 compares the SCE rates using the number of SCEs per 100 hours driven by experimental condition. There were no statistically significant differences in SCE rate per experimental group (discussed in the SCE rate ANOVA results section below).

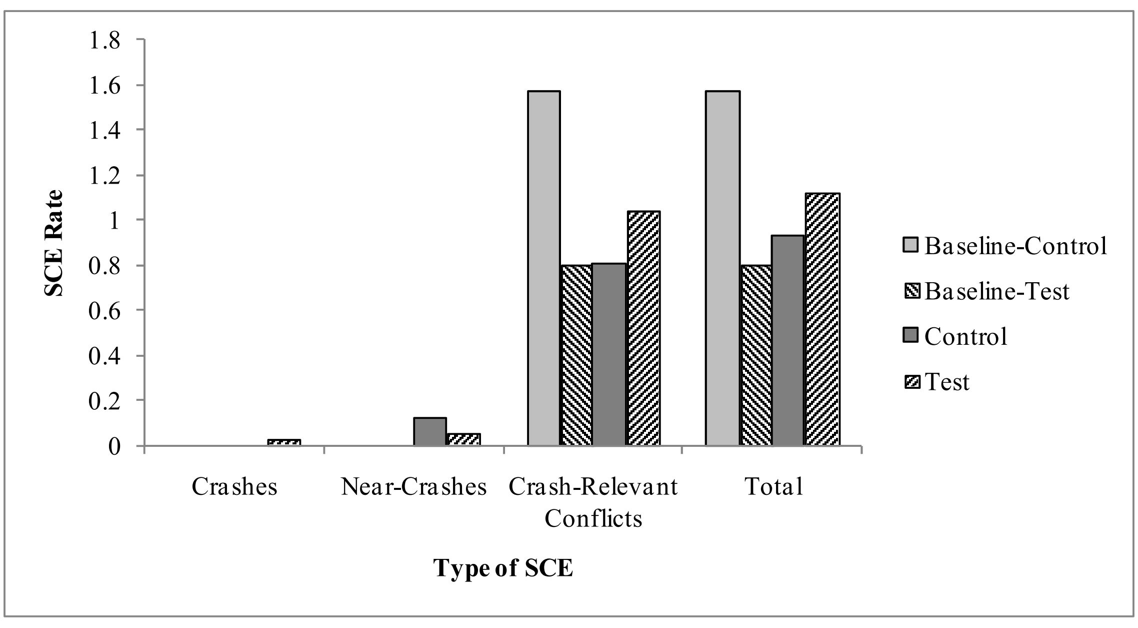


Figure 61. SCE Rate by Experimental Condition – At-Fault Within the Operating Envelope

Manual PERCLOS of At-Fault Safety Critical Events Within the Operating Envelope. Two surrogate measures of driver drowsiness were employed to characterize level of drowsiness during SCEs. The first measure evaluated was manual PERCLOS, and the second measure was subjective analysis of drowsiness-related driving behavior. Of the 149 events, 29 SCEs (19%) were lacking sufficient video data to calculate manual PERCLOS. Therefore, 120 SCEs were used in the descriptive statistics calculated to characterize the experimental conditions based on manual PERCLOS. The mean manual PERCLOS value for the SCEs was 7, with a maximum manual PERCLOS value of 23. Table 50 shows the descriptive statistics for the manual PERCLOS values at the moment of SCE occurrence, by experimental condition. The mean values are higher for the test conditions than for the control conditions. However, there are no differences when these values are compared to their respective baselines. Fischer test results confirmed that there was no correlation between manual PERCLOS value and the experimental condition (Fischer's exact test $p = 0.52$). Mean and maximum manual PERCLOS values for the SCE in each experimental condition were the same or lower than the manual PERCLOS values presented for the baseline events analyzed as part of Research Question 1.

Table 50. Manual PERCLOS Value for the SCEs in Each Experimental Condition – At-Fault Within the Operating Envelope

Experimental Condition	Total SCE	SCEs With Manual PERCLOS Available	Minimum	Maximum	Mean
baseline Control	24	18	1	10	3
baseline Test	28	23	1	23	9
Control	15	12	1	11	6
Test	82	67	0	23	7
Total	149	120	-	-	-

Table 51 presents the frequency of SCEs with manual PERCLOS values above 12 (drowsiness threshold used for this study) by experimental condition. Results indicated that 68 percent of SCEs have manual PERCLOS values less than 12. For the data available, the Control Group does not have any SCEs with manual PERCLOS values greater than 12. On the other hand, 12 percent of SCEs for the Test Group had manual PERCLOS values of 12 or greater. For the test condition, 10 SCEs have manual PERCLOS values greater than 12 (all crash-relevant conflicts). Given that 20 percent of the SCEs did not have manual PERCLOS and the majority of the SCEs had a value below 12, no further statistical analyses were performed.

Table 51. SCEs With Manual PERCLOS Above 12 by Experimental Condition – At-Fault Within the Operating Envelope

Experimental Condition	PERCLOS <12	PERCLOS ≥12	PERCLOS N/A	Total
Baseline Control	18	0	6	24
Baseline Test	15	8	5	28
Control	12	0	3	15
Test	57	10	15	82
Total	102 (68%)	18 (12%)	29 (20%)	149

Drowsiness-Related Driving Behavior Evaluation for At-Fault Safety Critical Events Within the Operating Envelope. As mentioned earlier, all SCEs were evaluated and driver behavior was categorized if the driver appeared drowsy, sleepy, asleep, fatigued, or showed signs of reduced alertness that occurred during the period of time leading to the SCE. Table 52 shows the frequency of drowsiness-related SCEs by experimental condition. Of the 149 SCEs, analysts identified 48 SCEs (32%) in which the driver exhibited a drowsiness-related behavior. However, there was no statistical difference among the experimental conditions based on the proportion of SCEs that were drowsiness-related versus not drowsiness-related (χ^2 (3, N = 149) = 3.3322, p = 0.0.3432). Twenty-eight SCEs were considered drowsiness-related for the test condition. Of these 28 SCEs, 2 were crashes and 26 were crash relevant conflicts.

Table 52. Frequency of Drowsiness-Related SCEs by Experimental Condition – At-Fault Within Operating Envelope

Experimental Condition	Not Drowsy	Drowsy	Total
Baseline Control	17	7	24
Baseline Test	17	11	28
Control	13	2	25
Test	54	28	82
Total	101 (68%)	48 (32%)	149

Number of Vehicles Involved in At-Fault Safety Critical Events Within the Operating Envelope. Table 53 presents the number of vehicles involved in SCEs as a function of experimental condition. Most of the SCEs involved two vehicles (54%). This represents the same proportion of SCEs if all the SCEs within the operating envelope were considered and not just those at-fault. The number of SCEs that involved only one vehicle (the experimental vehicle) was 44 percent. Kendall's Tau b value (τ_b = -0.0288) and Spearman correlation (ρ = -0.0309) suggested that the number of vehicles involved was not associated with the experimental condition.

Table 53. Number of Vehicles Involved in SCEs by Experimental Condition – At-Fault Within the Operating Envelope

Experimental Condition	Single Vehicle	Two Vehicles	More than Two Vehicles	Pedestrian, Cyclist, or Animal	Total
Baseline Control	8	16	0	0	24
Baseline Test	16	12	0	0	28
Control	4	11	0	0	15
Test	38	42	2	0	82
Total	66 (44%)	81 (54%)	2 (1%)	0	149

At-Fault Safety Critical Events Within the Operating Envelope that Received Alerts. The DFM recorded 224 alerts in the hour prior to the 149 SCEs. However, only 25 of those 149 SCEs received alerts (21 received full alerts). Table 54 and Table 55 show the descriptive statistics for the alerts (number of SCEs and number of alerts received). The percentage of SCEs that received initial and full alerts was lower for the test condition as compared to the control conditions.

Table 54. Descriptive Statistics for Number of Initial Alerts One Hour Before SCE by Experimental Condition – At-Fault Within the Operating Envelope

Experimental Condition	Total SCEs	SCEs that Received Alerts	Percentage of SCEs with Alerts	Minimum (No. Alerts)	Maximum (No. Alerts)	Mean (No. Alerts)
Baseline Control	24	2	8	8	8	8
Baseline Test	28	2	7	5	11	8
Control	15	6	40	2	56	17
Test	82	15	18	1	23	6
Total	149	25	-	-	-	-

Table 55. Descriptive Statistics for Number of Full Alerts One Hour Before SCE by Experimental Condition – At-Fault Within the Operating Envelope

Experimental Condition	Total SCEs	SCEs that Received Alerts	Percentage of SCEs with Alerts	Minimum (No. Alerts)	Maximum (No. Alerts)	Mean (No. Alerts)
Baseline Control	24	2	8	8	8	8
Baseline Test	28	2	7	3	10	7
Control	15	5	33	2	56	20
Test	82	12	15	1	23	7
Total	221	21	-	-	-	-

Sleep Quantity Before At-Fault Safety Critical Events Within the Operating Envelope. The amount of sleep in the 24 hours before the SCE was computed for all the drivers with available actigraphy data. As all drivers did not have actigraphy data available due to device failures and drivers removing the device, only a subset of events had sleep data available during the period of interest. Based on the actigraphy data, and the method described previously of removing all naps that were not at least 20 min in duration, drivers in the baseline control condition slept an average of 5.4 hours before an at-fault SCE, similar to drivers in the baseline test Condition. Drivers in the control condition had the lowest average, 4.7 hours. Drivers in the test condition had the highest average, with 6.5 hours of sleep prior to the SCE (Table 56). Both the Spearman Correlation ($\rho = 0.1857$) and the Kendall's Tau-b ($\tau_b = 0.1484$) suggest that the time slept before at-fault SCE was not associated with the Experimental Condition. However, it should be noted that small adjustments in the quantity of sleep may have beneficial effects, even if those effects are not reflected in the statistical analysis.

Table 56. Descriptive Statistics of Sleep 24 Hours Before SCEs by Experimental Conditions – At-Fault Within the Operating Envelope

Experimental Condition	Total SCEs	SCEs with 24-h Actigraphy Data	Minimum (hrs)	Maximum (hrs)	Mean (hrs)
Baseline Control	24	16	1.0	8.8	5.4
Baseline Test	28	22	0.8	10.0	5.0
Control	15	13	0.8	8.5	4.7
Test	82	69	0.6	16.0	6.5
Total	149	120	-	-	-

SCE Rate ANOVA Results. A mixed factor design was used for the ANOVA to test the equality of SCE rate means for the different experimental conditions (Table 57). Based on the results presented in the ANOVA summary table, there were no statistically significant differences in terms of SCE rate for the main effects or interaction.

Table 57. ANOVA Summary Table – At-Fault Within Operating Envelope

Source	DF	SS	MS	F -Value	Pr > F
Between Subjects					
Experimental Group (EG)	1	8.28	8.28	0.80	0.37
Participant/EG	82	849.40	10.36		
Within Subject					
Baseline Reference (BR)	1	2.62	2.62	0.53	0.47
BR x EG	1	5.14	5.14	1.04	0.31
BR x Participant/EG	82	407.35	4.97		

Discussion

The analyses presented in this chapter evaluated the safety benefits of the DFM as a function of a reduction of driver involvement in SCEs. The analyses took into consideration all SCEs, SCEs within the DFM operating envelope, and SCEs when the truck driver was at fault during the different experimental conditions. A total of 1,124 SCEs were analyzed, including 28 crashes, 112 near-crashes, and 984 crash relevant conflicts. Of these 1,124 SCEs, 221 occurred within the operating envelope of the DFM. Statistical tests considering all the SCEs showed no statistically significant difference between the Control and Test Groups. Moreover, no statistically significant differences in SCE distribution were observed for drivers during the baseline and the test conditions.

Two surrogate measures of driver drowsiness were used for this research question: manual PERCLOS and a subjective drowsiness-related driver behavior evaluation. Over 60 percent of the SCEs had a manual PERCLOS value below 12 (i.e., the majority of SCEs were under the drowsiness threshold). The second measure of interest was a subjective assessment of driver behavior. This measure was obtained during the data reduction process performed to characterize each of the SCEs (Hickman et al., 2005). Drivers exhibited a drowsiness-related behavior in 143

SCEs (13%). However, there were no statistical differences among the experimental conditions based on SCE in which drivers exhibited drowsiness-related behavior.

Of the 1,124 SCEs, 221 were within the DFM's operating envelope (i.e., the DFM could have potentially provided an alert if the conditions were met). However, very few alerts were registered preceding the SCE. Moreover, there were no statistically significant differences for the mean number of SCEs in the different experimental conditions even when drowsiness-related SCEs were considered.

In addition SCE exposure was also considered in the analyses. The hours driven for each driver were calculated, and an SCE rate per 100 hours driven was developed. The ANOVA indicated that there were no statistically significant differences even when exposure was considered (i.e., SCE rates) for the various experimental conditions. The results also showed no statistically significant differences between experimental groups.

A more in-depth analysis focused on the SCEs in which the truck driver was considered at fault. The at-fault SCEs represents 836 (74%) of the 1,124 SCEs considered. The 836 SCEs consisted of 12 crashes, 63 near-crashes, and 761 crash relevant conflicts. Similar analyses to the ones delineated above were performed for this subset of data (e.g., within operating envelope, sleep quantity, vehicles involved, PERCLOS). The results when only at-fault SCEs were considered were consistent with results already presented for all SCEs; that is, no statistically significant differences in terms of number of SCEs, SCE rate, sleep, number of vehicles involved, number of alerts, manual PERCLOS, and drowsiness-related behavior existed for the experimental conditions of interest or between the experimental groups.

RESEARCH QUESTION 4: HUMAN-MACHINE INTERACTION

As part of this section the following research question (RQ) was addressed:

> RQ 4.1: How do drivers operate the DDWS in a real-world environment?

Methods, analysis, measures of interest, results, and discussion related to this question are presented next.

Methods

In order to examine the interaction between drivers and the DFM in a real-world environment, each interaction that a driver had with the DFM was flagged within the data set. The DFM user interface is shown in Figure 62, with the user controls labeled. Note that all controls except for the DFM dark/active switch are user adjustable (Table 58).

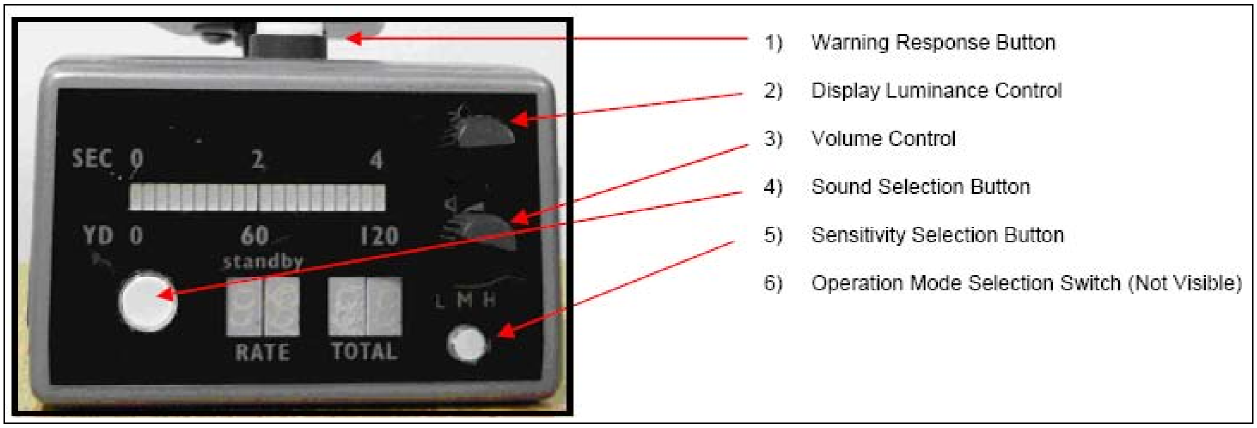


Figure 62. DFM User Interface and Controls

Table 58. DFM User Adjustable Control Description

Control	Control Type	User Setting Range
Warning response button	Momentary push button	Acknowledgement of alarm
Display luminance control	Knurled knob	Continuous adjustment
Volume control	Knurled knob	Continuous adjustment
Sound selection button	Momentary push button	Sound type: 1, 2, 3, 4, 5
Sensitivity selection switch	Momentary push button	Low (Dark default) Medium (Active default) High
Operation mode selection	Keyed switch	Dark, Active (not user adjustable)

During data collection, the DAS recorded any driver adjustment of a DFM setting. Therefore the data set was examined for driver adjustments of the sensitivity setting, warning sound, and display brightness, and also for when the warning response button was pressed.

Data Set

The data set for Research Question 4 included data from 96 drivers. In addition to the nine drivers removed from all analyses, an additional four drivers were specifically removed from the Research Question 4 analysis due to their incorrect use of the DFM (preventing it from properly switching between active and standby modes), preventing accurate data collection. For example, these drivers allowed the alert to continue to sound so that they did not have to repeatedly press the warning response button. Doing this prevents the DFM from correctly changing into standby mode, causing the DFM to stay in active mode when it should not. In addition, 10 drivers left the study prior to the DFM being switched into active mode; these drivers did not have data for inclusion in the present analysis. For the 82 remaining drivers, VTTI developed software (DART) that scanned all of the files recorded for each driver. Every instance of driver interaction with the DFM was flagged for those periods where the DFM was set to active mode (i.e., the driver was able to receive feedback from the DFM) and the device was not in standby mode (i.e., the driver was traveling over 35 mph, and ambient illumination was less than 50 lux). This resulted in 1,531 sensitivity level changes, 608 warning sound changes, 354 display brightness changes, and 41,298 warning response button presses.

Once the driver inputs were obtained, an analysis was performed examining the frequency of driver changes for each one of the three settings (sensitivity level, warning sound, or display brightness) and to obtain mean durations of how long the alert sounded before the driver pressed the warning response button. In addition, the amount of time that each setting was set in each level was calculated across all drivers to see if there were preferred settings.

The second analysis of the human-machine interface examined differences in user behaviors and interactions with the device in two periods. The first period included driving shifts where the driver had an SCE and those where the driver did not have an SCE. To perform this analysis, a unique identifier for the individual shift (shift ID) was assigned to each of the driver inputs. The shift ID was calculated by chronologically ordering all driver files and assigning shifts based on units of time driven by each individual driver. Shift time continued to accumulate until the driver took a break (i.e., truck speed = 0 mph) for six hours or more. Following this, the next file began the start of a new shift.

A shift ID was also assigned to each of 108 SCEs that occurred while the DFM was both in active mode and not in standby mode. Each shift ID of the driver inputs was cross-checked against the shift ID of the safety-critical events. This process yielded 114 sensitivity level changes, which were concurrent with an SCE, 35 warning sound changes, 21 display brightness changes, and 2,292 warning response button presses. The data were then assigned to categories of whether or not an SCE occurred in the shift, and then analyzed for frequency and duration of setting changes.

Measures

The following measures were used in the analysis of Research Question 4:

- Frequency of Driver Input: number of times the driver changed the sensitivity level, warning sound, or display brightness.
- Alert Duration: temporal duration, in seconds, between the driver receiving a full alert and when he or she pressed the warning response button.

Results

Mean Duration of Settings

The temporal duration of the levels of each of the three settings (sensitivity level, warning sound, and display brightness) were calculated across all drivers to determine if preferred settings existed. The mean duration for each of the three sensitivity levels was calculated across the 82 drivers. The DFM was most frequently set to the medium sensitivity level (Figure 63). This was an expected occurrence because when the DFM was in dark mode and no alerts were given to the driver (even though they were being recorded by the system), the default sensitivity level was low. As soon as it was set to active mode, the sensitivity default moved to medium. When the DFM was changed to active mode, the driver most often left the sensitivity level at medium and did not change it. When the driver did change the sensitivity level, it was most often changed to the low sensitivity. The low sensitivity provides alerts less frequently.

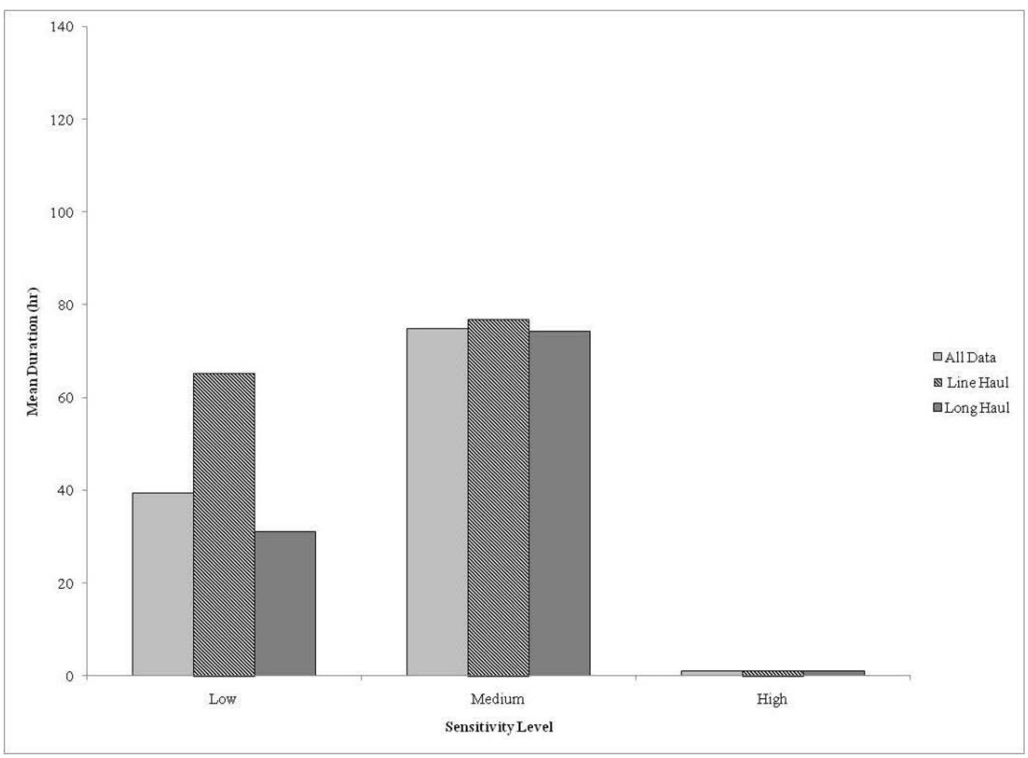


Figure 63. Mean Duration of Each Sensitivity Level Setting During the Test Condition

The DFM had six different built-in warning sounds selectable by the user. The driver was able to press the sound selection button (illustrated in Figure 62) to hear a preview of each sound. Similar to the adjustment of sensitivity level, when the DFM was in dark mode, the warning sound defaulted to the first sound. Note that no sound was actually presented to the driver during dark mode. Figure 64 shows that drivers most often left the sound on the default sound setting and did not change it. The next most popular sound was number 6, while the least popular sound was number 4.

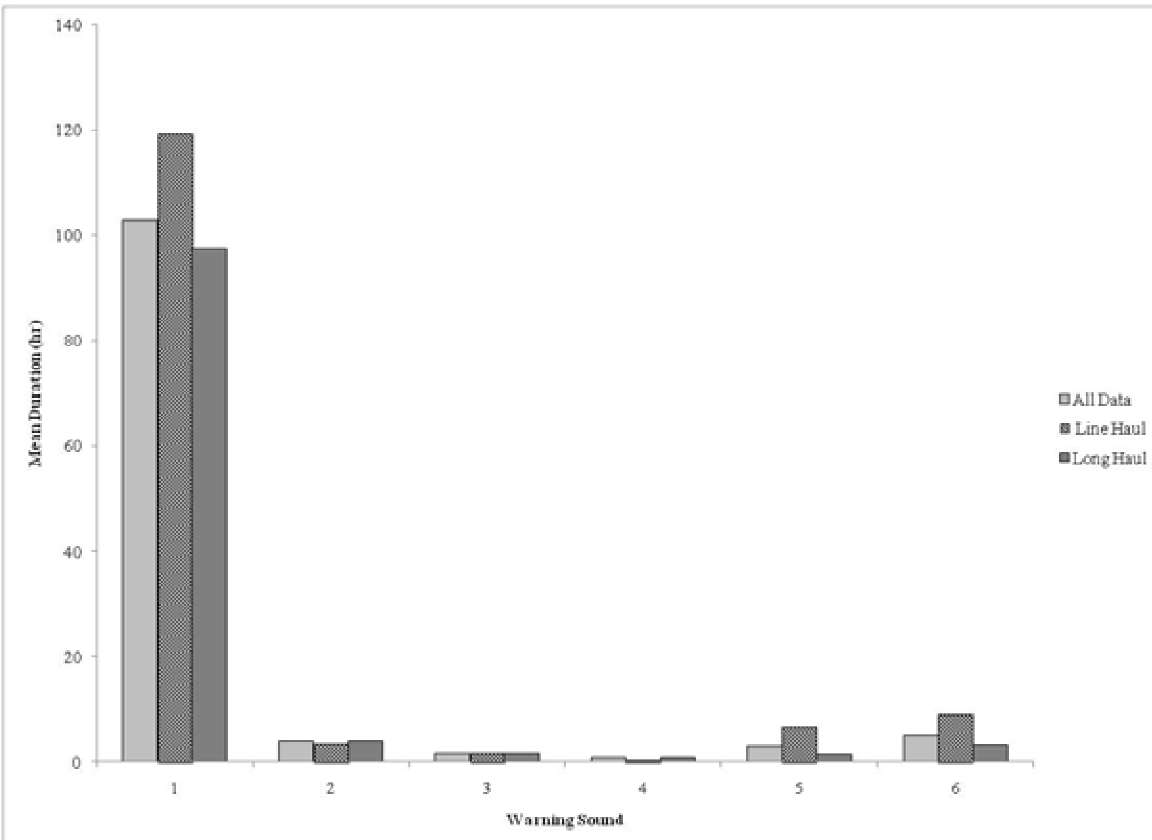


Figure 64. Mean Duration of Each Warning Sound Setting

The driver was able to use a knurled knob on the DFM to change the display brightness. The light level of the display brightness was an arbitrary number and ranged from 0 (dark) to 255 (bright). As it was possible to adjust the knurled knob while the DFM was in dark mode, there was no default setting for display brightness. Across all drivers, the display brightness was most often set between 201 and 255 (brightest) followed closely by settings between 0 and 50 (darkest). Figure 65 presents the results for this setting.

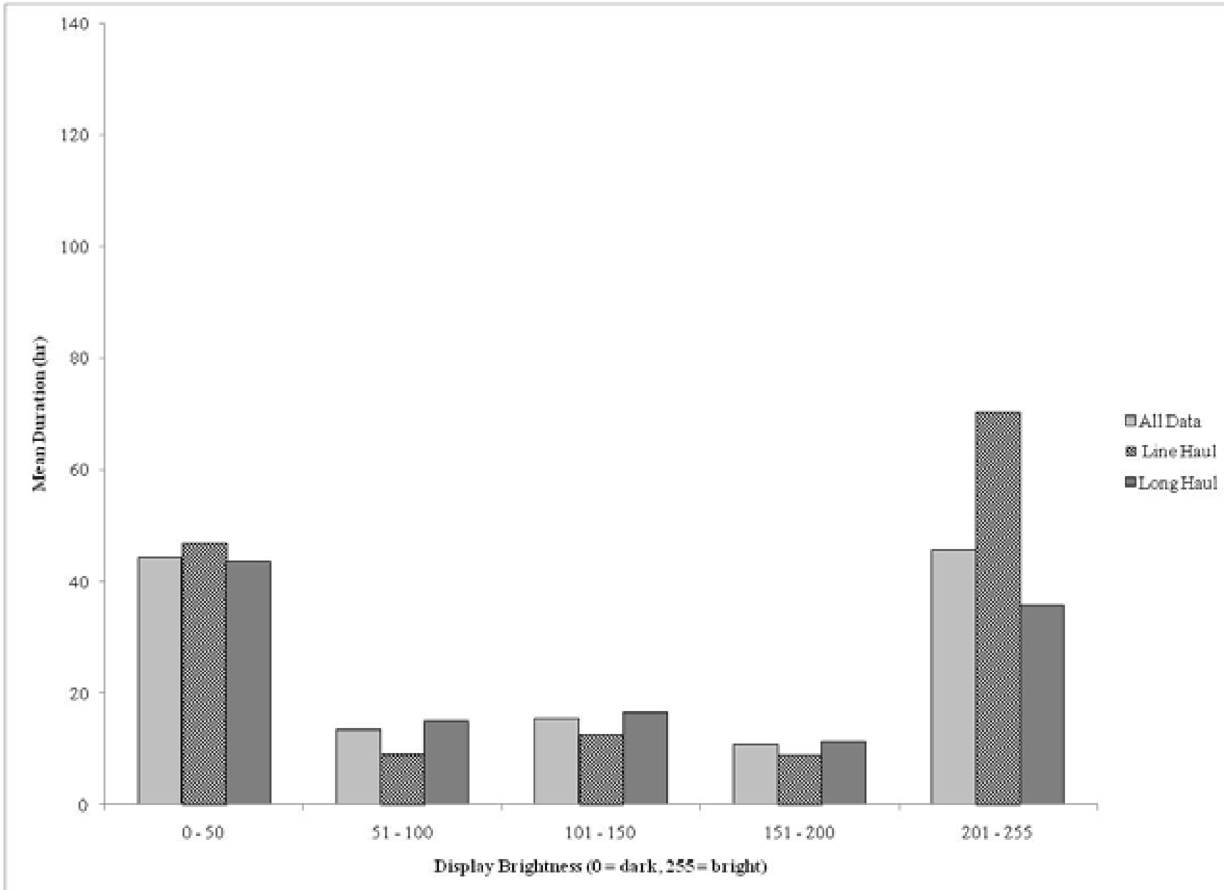


Figure 65. Mean Duration, in Hours, of Each Display Brightness Setting

Frequency of Driver Input

Each driver input to the DFM was analyzed. The data was analyzed for changes in sensitivity level, warning sound, and display brightness and then compared to exposure data in order to examine possible longitudinal effects.

Each driver input was categorized by the time of day that the change was made (as shown in Table 59). Exposure data was calculated using only time periods during the test condition, when the DFM was in active mode (i.e., during time periods when the driver was able to manipulate the DFM). This allowed a conclusion to be drawn as to whether an increase in the frequency of driver input in a particular hour was a unique event or merely due to an increased frequency of driving during that hour.

Table 59. Frequency of Sensitivity Level, Warning Sound, and Display Brightness Changes as a Function of Time of Day

Time	Sensitivity Level	Warning Sound	Display Brightness	Opportunities
0:00 - 0:59	87	38	20	2033
1:00 - 1:59	82	29	22	1,749
2:00 - 2:59	73	25	13	1,664
3:00 - 3:59	89	15	13	1,582
4:00 - 4:59	74	21	15	1,512
5:00 - 5:59	79	20	19	1,311
6:00 - 6:59	62	28	8	917
7:00 - 7:59	30	23	7	369
8:00 - 8:59	5	4	2	77
9:00 - 9:59	1	0	0	66
10:00 - 10:59	5	0	1	93
11:00 - 11:59	2	3	1	106
12:00 - 12:59	3	2	2	118
13:00 - 13:59	1	1	0	100
14:00 - 14:59	4	1	0	117
15:00 - 15:59	7	5	0	120
16:00 - 16:59	28	17	6	245
17:00 - 17:59	74	44	28	757
18:00 - 18:59	96	55	27	1,108
19:00 - 19:59	144	63	33	1,596
20:00 - 20:59	162	61	49	2,372
21:00 - 21:59	169	49	46	2,585
22:00 - 22:59	150	53	23	2,694
23:00 - 23:59	104	51	19	2,416
Totals	1,531	608	354	25,707

There was an increase in changes to the sensitivity level settings that did not mirror the exposure data (measured as the opportunities available to drive in that hour) for the block of time between 3:00 and 3:59 (Figure 66). One explanation for this is the circadian rhythm low that typically occurs between 2 a.m. and 4 a.m. during which the body naturally tends to be drowsy (Stutts, Wilkins, & Vaughn, 1999). The driver's level of drowsiness may be increasing during this time, producing an increased amount of alerts from the DFM, and possibly causing the driver to change the sensitivity level to a lower setting to decrease the frequency of alerts.

Because the DFM goes into standby when the ambient illumination surpasses 100 lux, the system should not be operational during daylight hours. A relatively small amount of driver inputs occurred between 8 a.m. and 5 p.m., when one would reasonably expect daylight conditions. This was due to transient readings obtained by the light meter, which recorded a false low ambient illumination level. The frequency of driver input for each change in the sensitivity level setting, warning sound, and the display brightness, plotted against time of day, is displayed in Figure 66.

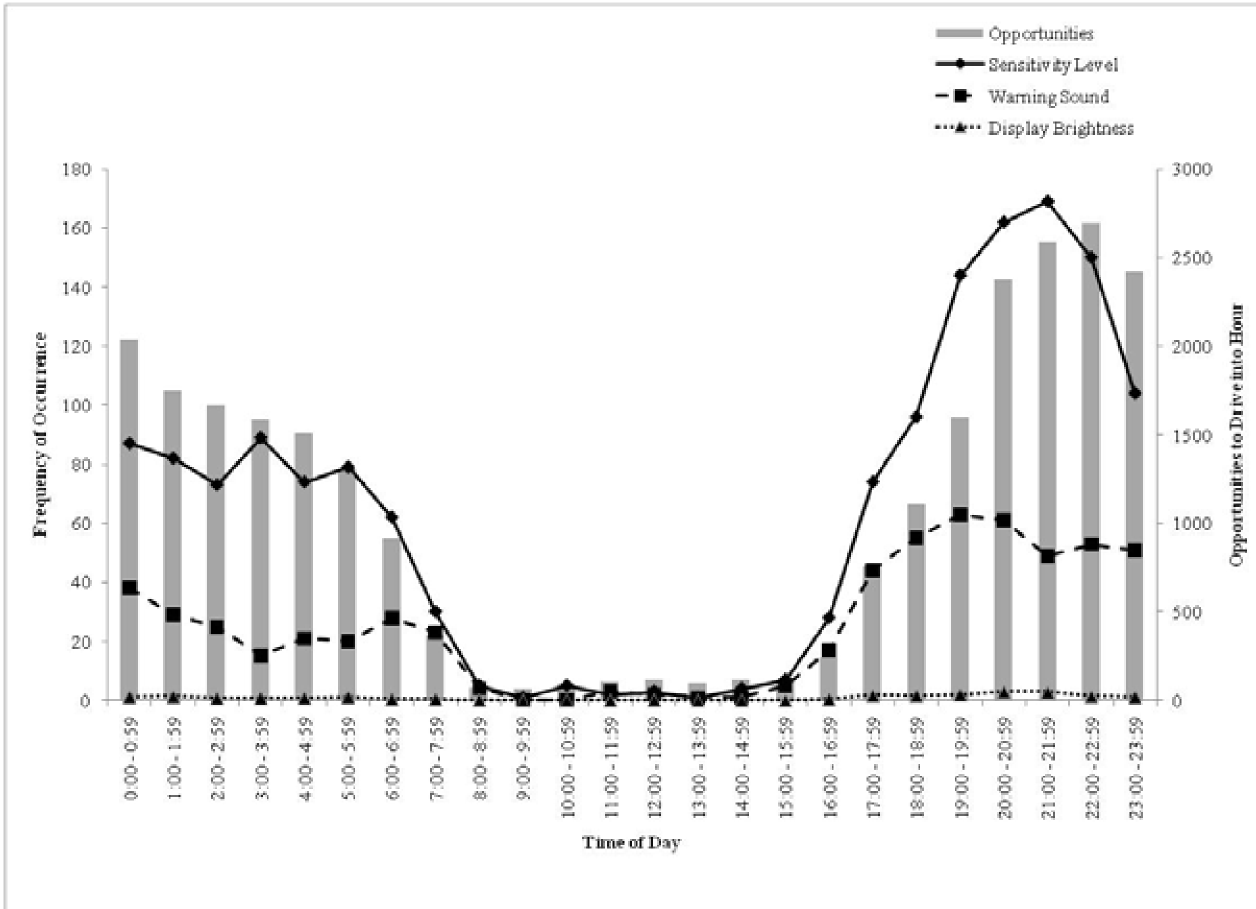


Figure 66. Frequency of Driver Input as a Function of Time of Day

The 2,493 driver inputs were also analyzed as a function of driving hour (Table 60). To determine driving hour, the DART software package sorted all files for an individual driver in chronological order and then grouped files into shifts (indicated by a unique shift ID) by the following method:

- The first shift began with the first file for a driver.
- DART calculated driving time as long as the truck was in motion (i.e., speed > 0 mph).
- Any time the truck stopped (speed = 0 mph) for greater than six hours it was assumed that the driver was taking a break, and the end of the driver's shift was marked.
- The next file with speed greater than 0 mph would be the start of a new shift.

This allowed for the determination of driving time when the input to the DFM was made. All driver input data were grouped by driving hour ranging from 1 to 11 hour(s). Note that 11 hours was the maximum number of hours a driver is allowed to drive within a given shift (FMCSA, 2005).

Table 60. Frequency of Sensitivity Level, Warning Sound, and Display Brightness Changes as a Function of Driving Hour

Driving Hour	Sensitivity Level	Warning Sound	Display Brightness	Opportunities
1	363	104	75	2,143
2	144	55	34	2,426
3	138	53	26	2,457
4	133	57	29	2,397
5	163	67	20	2,449
6	122	63	17	2,206
7	96	51	22	2,095
8	91	41	26	1,855
9	66	16	28	1,558
10	54	19	20	1,170
11	49	19	12	897
Totals	1,419	545	309	21,653

It can be seen in Figure 67 that the largest number of driver inputs occurred within the first hour of driving (363 sensitivity level changes, 104 warning sound changes, and 75 display brightness changes). A likely explanation is that the driver is adjusting the DFM settings after first starting his or her shift. The frequency of inputs for the remaining hours drops off as the frequency of driving in each hour decreases.

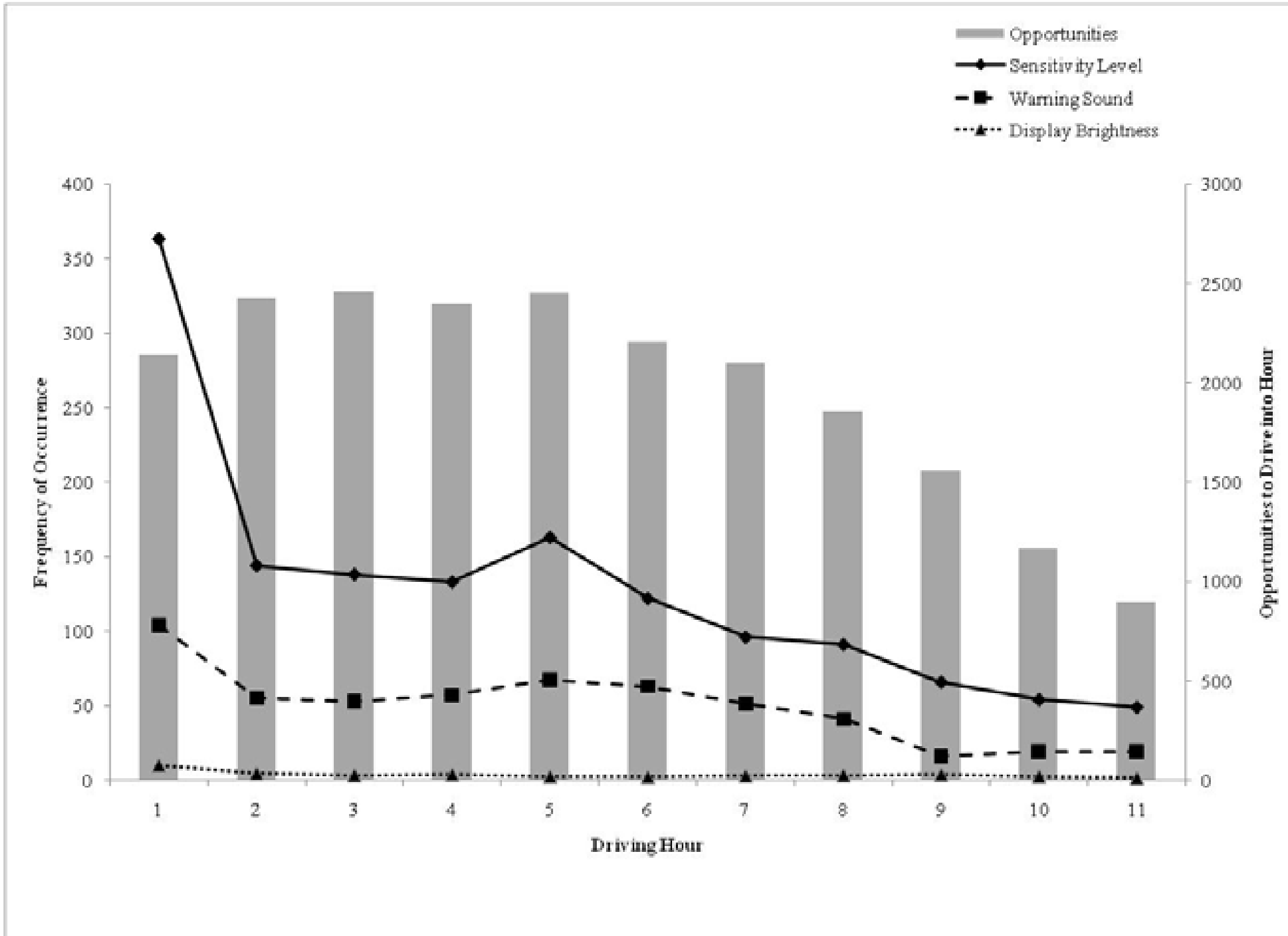


Figure 67. Frequency of Driver Input as a Function of Driving Hour

Next the DFM driver input data were analyzed as a function of weeks since the DFM mode was changed (Table 61). The DFM was kept in dark mode, in which alerts were recorded but not relayed to the driver, for approximately the first four weeks of participation. The date that the DFM was changed from dark mode to active mode is considered to be $t = 0$, and one week is calculated as seven days from that date.

When the data were examined as a function of changes in settings, against the number of weeks in the study, it was observed that the highest frequency of changes occurred during the first week (198 changes). This is likely due to the drivers testing out and familiarizing themselves with the different settings and levels. Figure 68 shows the data graphed as a function of the number of weeks since the DFM changed from dark mode to active mode.

Table 61. Frequency of Sensitivity Level, Warning Sound, and Display Brightness Changes as a Function of Weeks Since DFM Mode Change

Week Since DFM Mode Change	Sensitivity Level	Warning Sound	Display Brightness	Opportunities
1	198	144	71	80
2	143	56	32	76
3	134	64	23	68
4	148	29	36	76
5	154	61	23	73
6	183	56	34	69
7	137	26	30	65
8	127	53	27	65
9	83	13	11	53
10	71	33	12	53
11	60	20	8	38
12	51	29	23	25
13	29	10	12	17
14	7	9	0	8
15	6	3	9	6
16	0	2	2	4
17	0	0	1	2
Totals	1,531	608	354	778

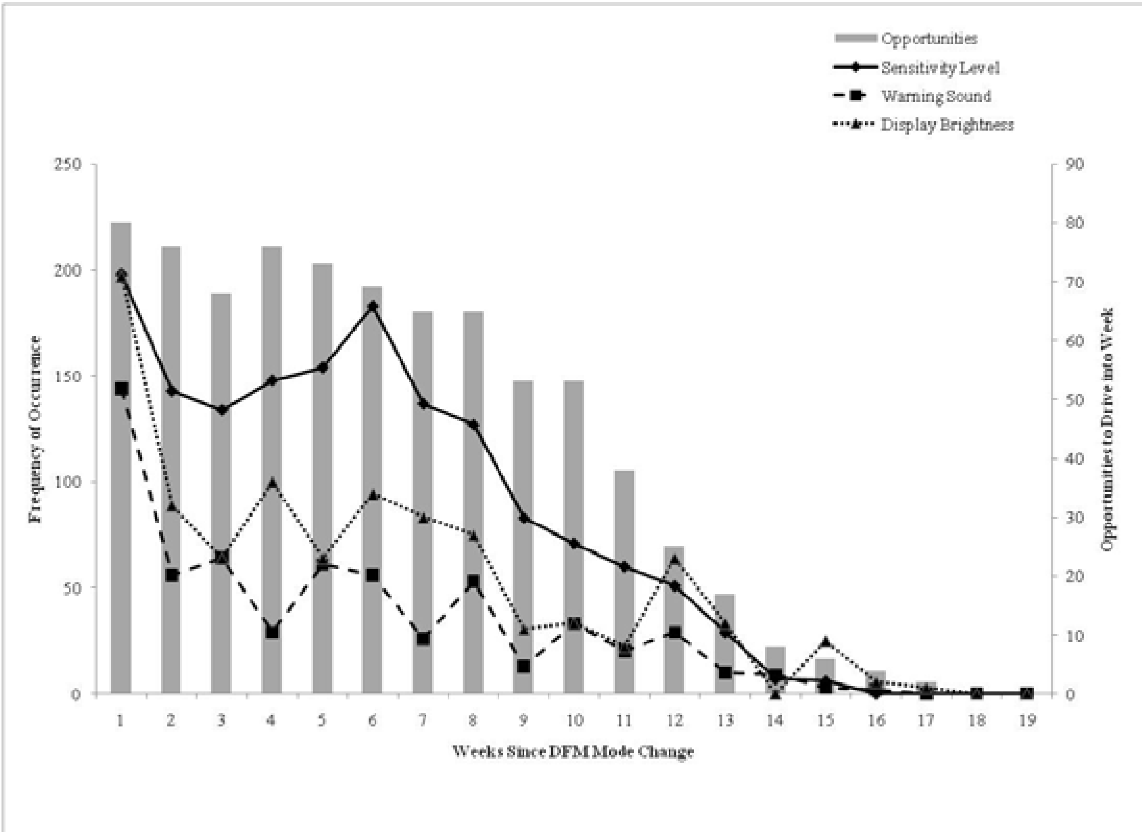


Figure 68. Frequency of Driver Input as a Function of Weeks Since DFM Mode Change

Alert Duration

In addition to identifying each occurrence of a driver-changed setting, the temporal duration (in seconds) of auditory alerts was calculated. Irrespective of sensitivity settings, once the DFM estimated PERCLOS at 12 percent, an auditory channel alert was produced. The driver had to press the warning response button to stop the alert. In order to calculate the mean alert duration grouped by time of day, the driver file was time stamped upon both the alert activation and the driver's press of the warning response button. This allowed for the alert duration to be calculated as the difference between these two times.

The mean duration of alerts was first analyzed as a function of time of day (Table 62). Figure 69 shows that alert duration increased in length at the 3:00-3:59 time block (51.89 s). One possible explanation for this is that the driver's reaction time is slowed due to drowsiness during this typical low-arousal period within the circadian rhythm. An alternative explanation arises from driver behaviors noted during the data reduction process (Hickman et al., 2005). Many analysts noted drivers reducing the volume of the DFM as far as possible, and occasionally raising their radio volume. This may have served as a conscious effort to mask the auditory alert. This would also allow the driver to let the alert continue to sound so that it would not act as a distraction or nuisance to the driver. This behavior may have resulted in drivers with an increased level of drowsiness. The increased drowsiness probably led to an increased frequency of alerts with this characteristic increased of alert duration for alerts that were ignored.

A second increase in alert duration is present for the noon–12:59 p.m. time block (32.90 s), when the DFM should not have been operational due to high ambient light levels. Further investigation revealed this was caused by the presence of a single high value that increased the mean value during the period. This driver was traveling through heavy rain, which likely caused the light level recorded by the light meter to drop below threshold, allowing the DFM to produce alerts. Review of the video data supported the conclusion that the driver was concentrating while driving through heavy rain and allowed the alerts to continue so as to not divert his attention from the driving task. This conclusion is further supported by the fact that the driver stopped the alert as soon as he passed through the area of heavy rain.

Table 62. Mean Alert Duration as a Function of Time of Day

Time of Day	Duration (s)	N
0:00 - 0:59	34.67	2290
1:00 - 1:59	41.52	1869
2:00 - 2:59	42.25	1658
3:00 - 3:59	51.89	1454
4:00 - 4:59	38.85	1945
5:00 - 5:59	29.25	2341
6:00 - 6:59	20.60	2427
7:00 - 7:59	16.75	1407
8:00 - 8:59	28.15	284
9:00 - 9:59	21.70	31
10:00 - 10:59	4.73	38
11:00 - 11:59	4.86	57
12:00 - 12:59	32.90	24
13:00 - 13:59	9.11	18
14:00 - 14:59	6.66	21
15:00 - 15:59	12.60	25
16:00 - 16:59	7.59	280
17:00 - 17:59	19.90	1928
18:00 - 18:59	24.81	2760
19:00 - 19:59	19.82	3596
20:00 - 20:59	16.89	5514
21:00 - 21:59	20.55	4363
22:00 - 22:59	26.96	3676
23:00 - 23:59	26.97	3292

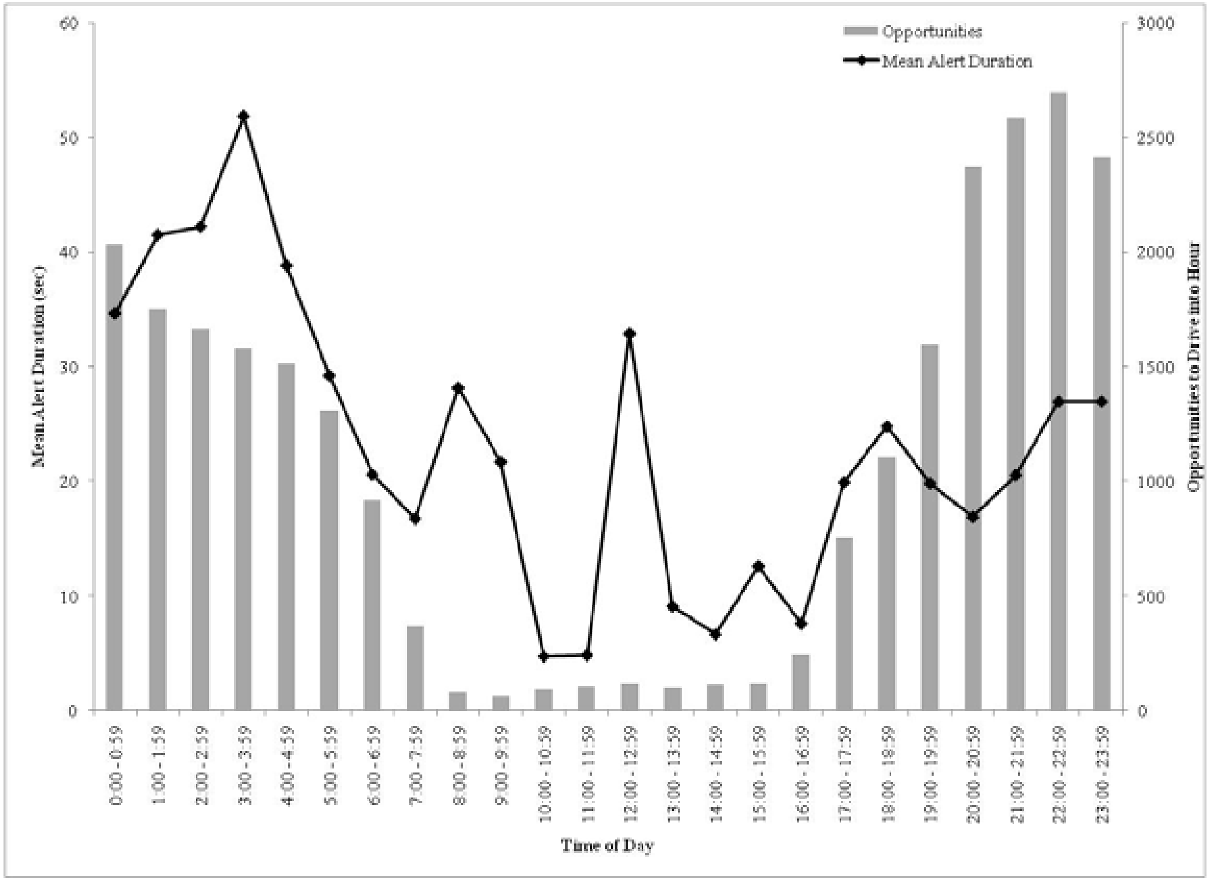


Figure 69. Mean Alert Duration as a Function of Time of Day

The mean alert duration was calculated as a function of driving hour, a measure of the driver's time on task (Table 63). A distinct pattern can be observed in Figure 70. Driver's mean alert response time was highest between the third and fifth hour of driving (34.96 s and 33.29 s, respectively). Mean alert durations were lower at the beginning of the drive, rose to the peak observed between the third and fifth hours, and then began to fall again as driving hour continued. It should be noted that there was an inverse relationship between driving hour and the number of drivers driving. This is partially due to the 11-hour maximum Hours-of-Service regulation, so fewer drivers are present as the time-on-task approaches the 11-hour legal limit. However, even when this is accounted for, a drop in mean alert duration was observed. This suggests that drivers were responding more rapidly to alerts as their driving time increased, which may be indicative of the need to have a stimulus to increase their arousal. Additional potential explanations include fatigue onset effects or simply experiencing the upward tracking of performance commonly seen toward the end of vigilance decrements with known end-points.

Table 63. Mean Alert Duration as a Function of Driving Hour

Driving Hour	Duration (s)	N
1	25.79	3,748
2	29.66	3,461
3	34.96	3,046
4	32.62	3,245
5	33.49	3,603
6	33.28	3,463
7	31.80	3,384
8	27.98	3,489
9	19.12	3,097
10	15.24	2,398
11	16.47	1,826

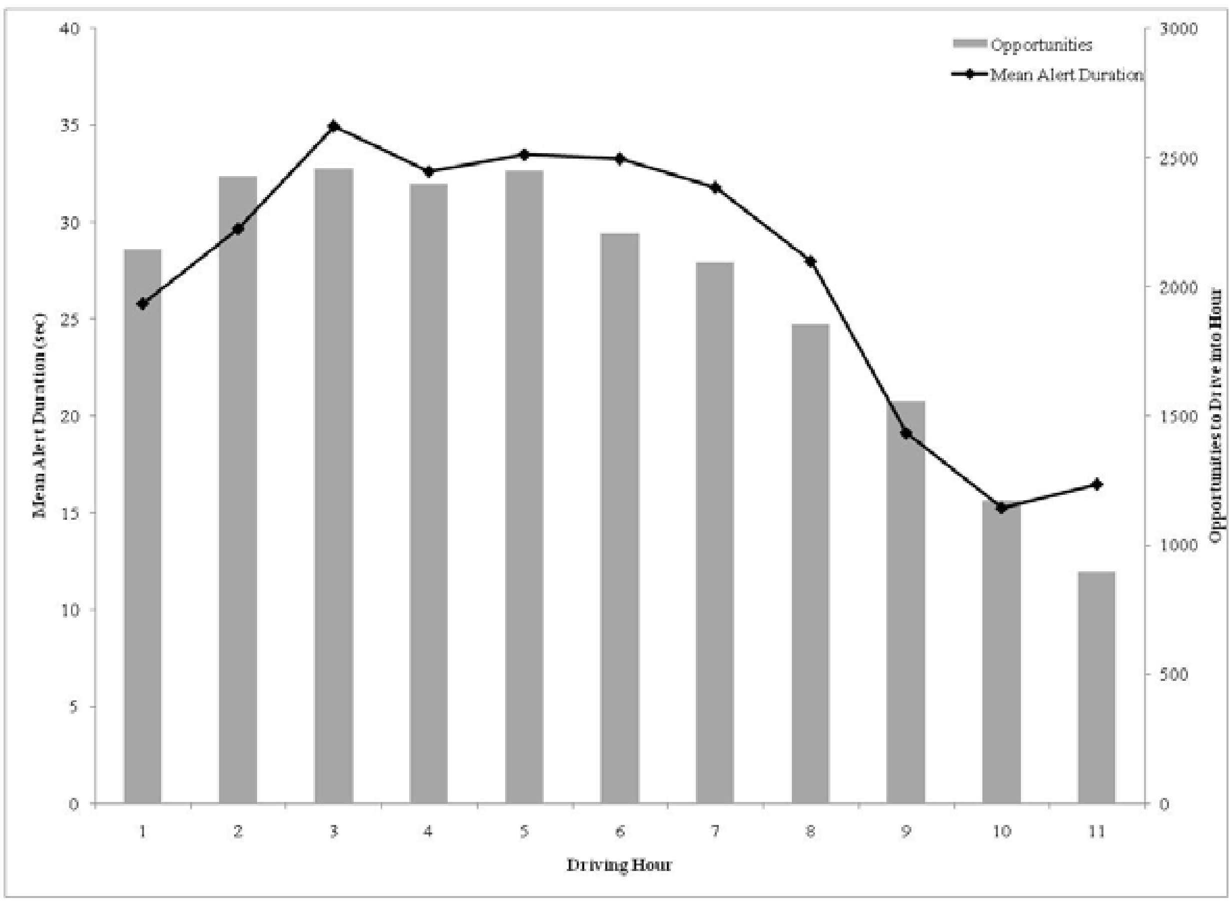


Figure 70. Mean Alert Duration as a Function of Driving Hour

Alert duration grouped as a function of number of weeks since the DFM mode was changed from dark to active was plotted (Figure 71). There is a large increase during the 14th week of data collection. Further examination revealed this was the result of a single driver who chose to completely reduce the volume of DFM auditory alerts and to allow alerts to continue to sound. The behavior from this driver also contributed to the increase in mean alert duration in weeks 12 and 13; the driver did not drive past week 14. However, this driver was not removed from the analysis due to the fact the driver represented approximately 75 percent of the observations during week 14. Removing this driver from the analysis would create a similar misrepresentation of the data in the opposite direction.

Table 64. Mean Alert Duration as a Function of Weeks Since DFM Mode Change

Weeks Since DFM Mode Change	Duration (s)	N
1	50.55	3,399
2	18.70	4,179
3	12.84	4,044
4	16.23	3,337
5	22.94	4,159
6	21.06	4,779
7	18.64	3,410
8	27.42	4,242
9	23.58	2,553
10	27.49	2,875
11	11.95	2,045
12	47.05	1,041
13	86.08	531
14	231.87	342
15	36.81	295
16	15.66	65
17	1.90	1
18	0	0
19	3.50	1

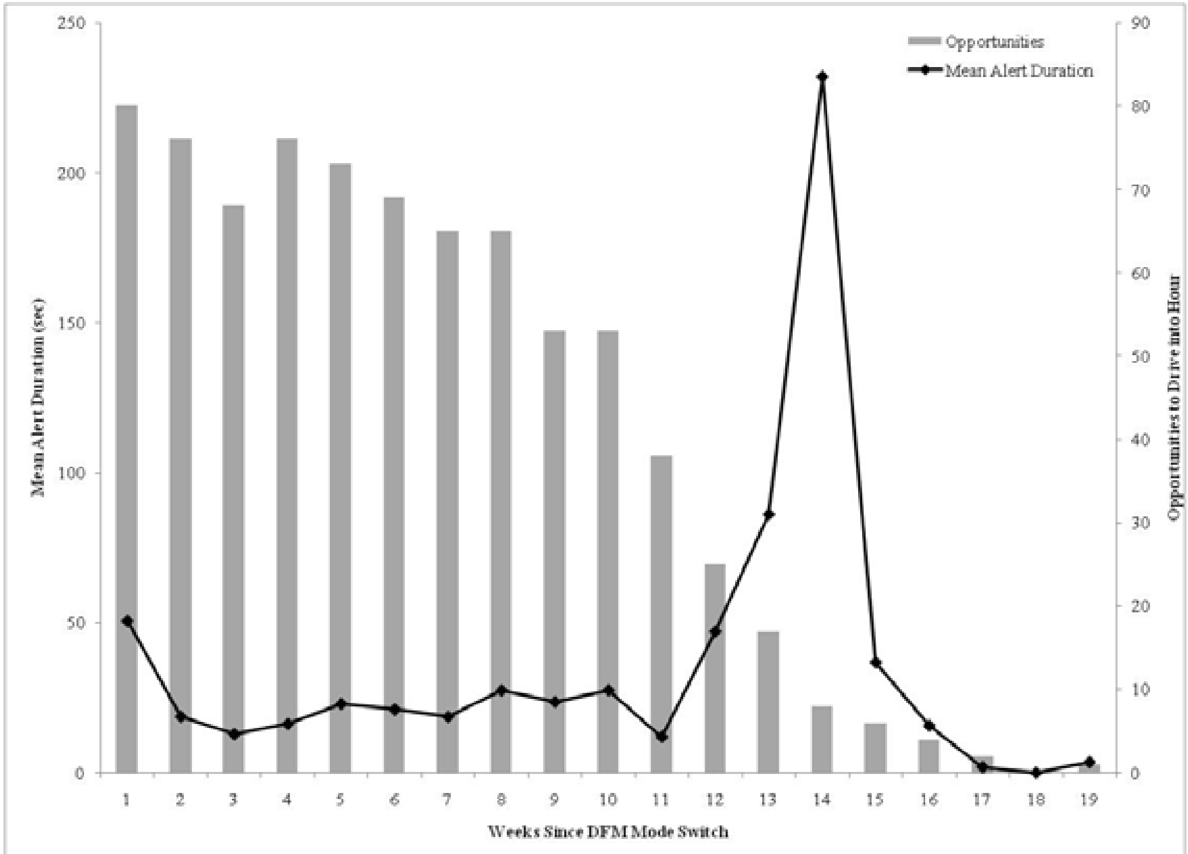


Figure 71. Mean Alert Duration as a Function of Weeks Since DFM Mode Change

Frequency of Driver Inputs With Regard to Safety Critical Events

A second analysis examined if there were differences in driver input when comparing shifts with and without an SCE. The frequency of driver interactions with the DFM in terms of the sensitivity level, warning sound, and display brightness were calculated as a function of time of day and whether or not drivers had at least one SCE recorded for their shift (Table 65).

The primary distinction present is a difference in system interaction between those drivers with an SCE in their shift and those without. Drivers with at least one SCE in their shift had a much lower level of interaction with the DFM in terms of all three adjustments examined (see Figure 72 to Figure 74), across all hours. Drivers with no SCE in their driving shift displayed a time of day pattern for their adjustments and interactions with the DFM. During the 15:00-21:00 time blocks, a very rapid increase in the frequency of driver interactions with DFM sensitivity level (Figure 72) was observed (peak value at 21:00-21:59). Similar trends occurred in the warning sound (Figure 73) selection (peak value at 19:00-19:59) and display brightness (peak value at 20:00-20:59). From these results it can be concluded that most interactions with DFM adjustments occurred until approximately the 8:00-8:59 time block, where almost all DFMs shifted into standby due to the illumination threshold. After that, the trend starts to increase again in the evening hours when the system became operational.

Table 65. Frequency of DFM Setting Changes by SCE in Shift as a Function of Time of Day

Hour	Sensitivity Level (SCE in Shift)	Sensitivity Level (No SCE in Shift)	Warning Sound (SCE in Shift)	Warning Sound (No SCE in Shift)	Display Brightness (SCE in Shift)	Display Brightness (No SCE in Shift)	Opportunities
0:00 - 0:59	11	76	2	36	4	16	2,033
1:00 - 1:59	7	75	4	25	2	20	1,749
2:00 - 2:59	8	65	3	22	1	12	1,664
3:00 - 3:59	4	85	0	15	0	13	1,582
4:00 - 4:59	5	69	1	20	0	15	1,512
5:00 - 5:59	4	75	0	20	1	18	1,311
6:00 - 6:59	1	61	0	28	0	8	917
7:00 - 7:59	2	28	0	23	0	7	369
8:00 - 8:59	0	5	0	4	0	2	77
9:00 - 9:59	0	1	0	0	0	0	66
10:00 - 10:59	0	5	0	0	0	1	93
11:00 - 11:59	0	2	0	3	0	1	106
12:00 - 12:59	0	3	0	2	0	2	118
13:00 - 13:59	0	1	0	1	0	0	100
14:00 - 14:59	1	3	0	1	0	0	117
15:00 - 15:59	0	7	0	5	0	0	120
16:00 - 16:59	7	21	1	16	1	5	245
17:00 - 17:59	14	60	6	38	3	25	757
18:00 - 18:59	6	90	3	52	1	26	1,108
19:00 - 19:59	10	134	0	63	1	32	1,596
20:00 - 20:59	9	153	5	56	3	46	2,372
21:00 - 21:59	10	159	1	48	2	44	2,585
22:00 - 22:59	9	141	5	48	1	22	2,694
23:00 - 23:59	8	96	4	47	1	18	2,416
Totals	116	1,415	35	573	21	333	25,707

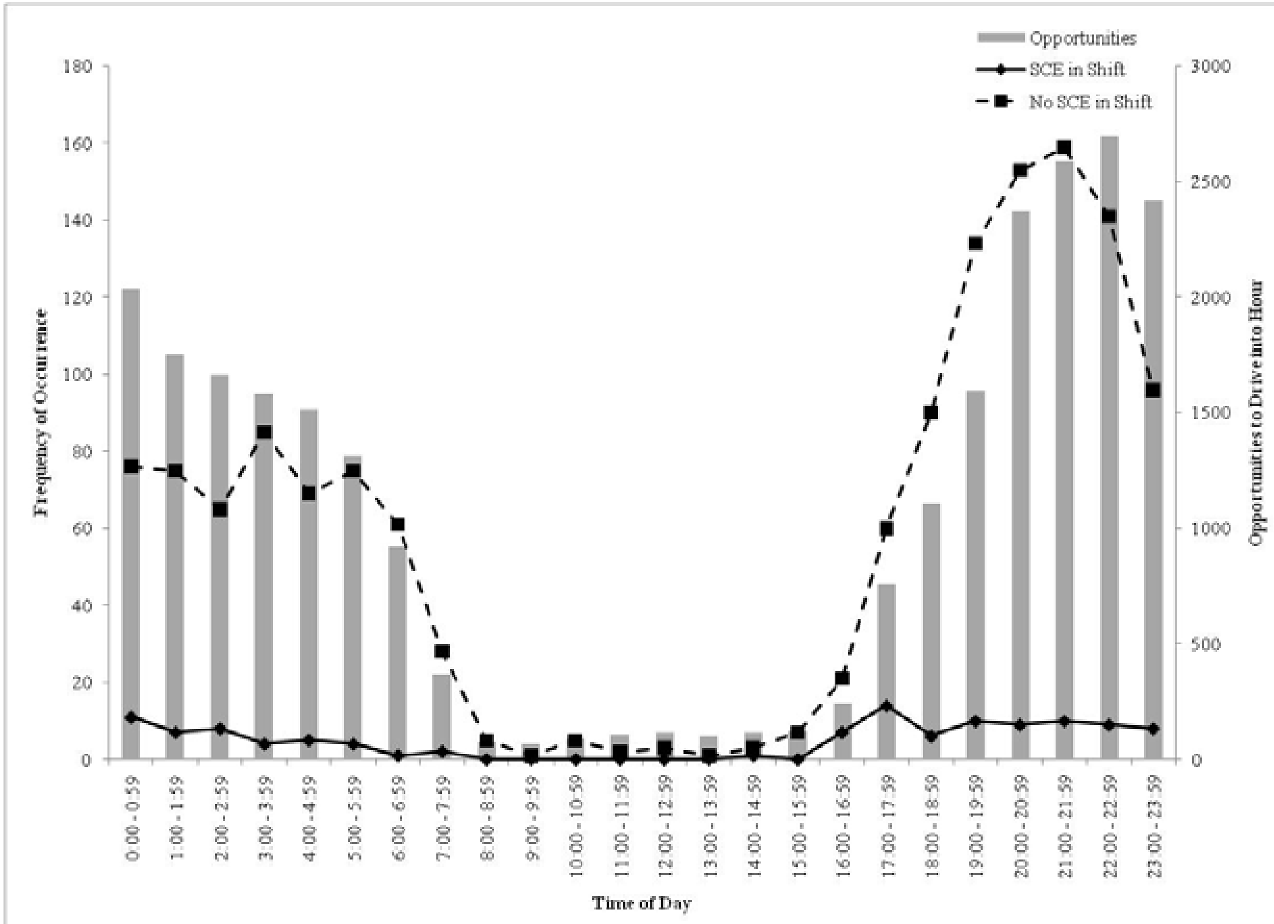


Figure 72. Frequency of Sensitivity Level Changes by SCE in Shift as a Function of Time of Day

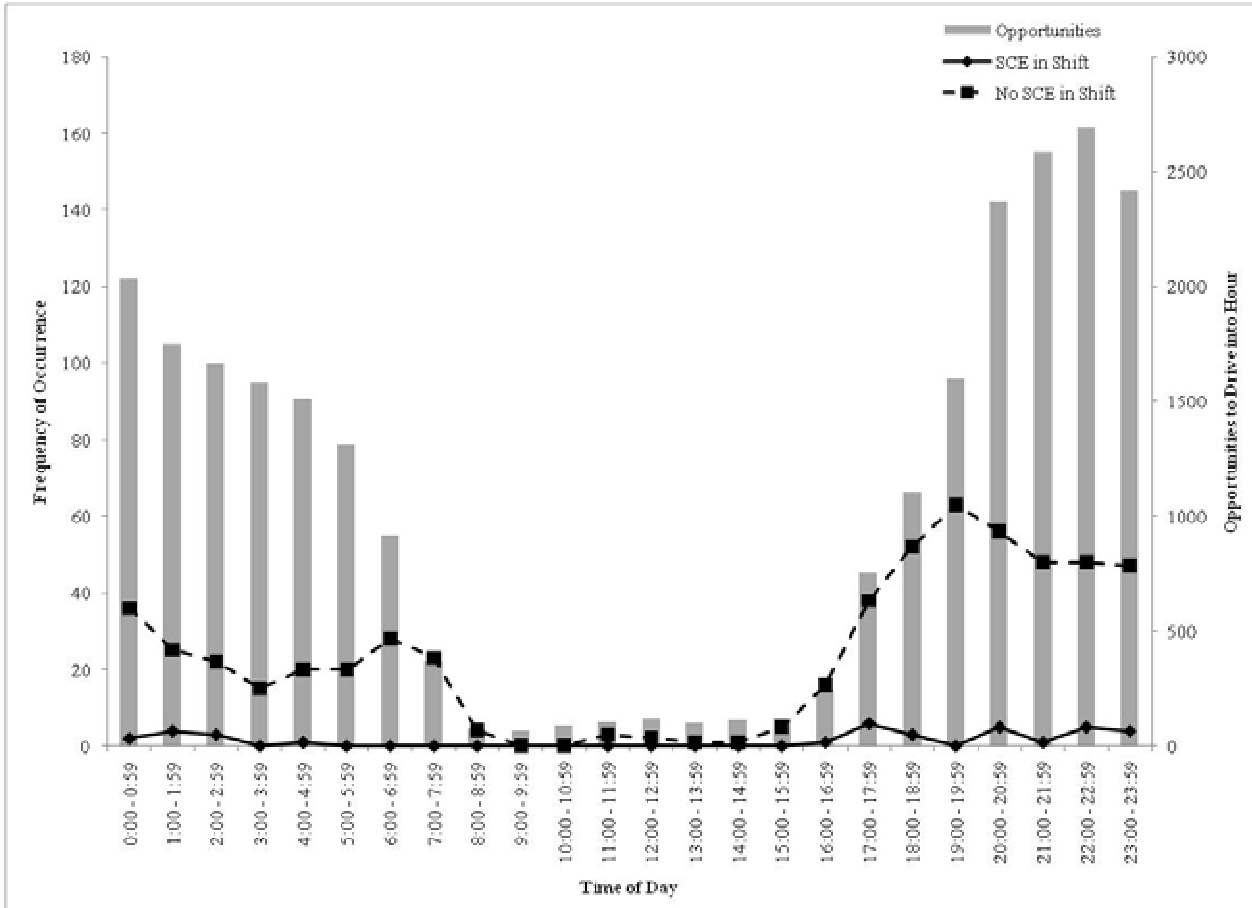


Figure 73. Frequency of Warning Sound Changes by SCE in Shift as a Function of Time of Day

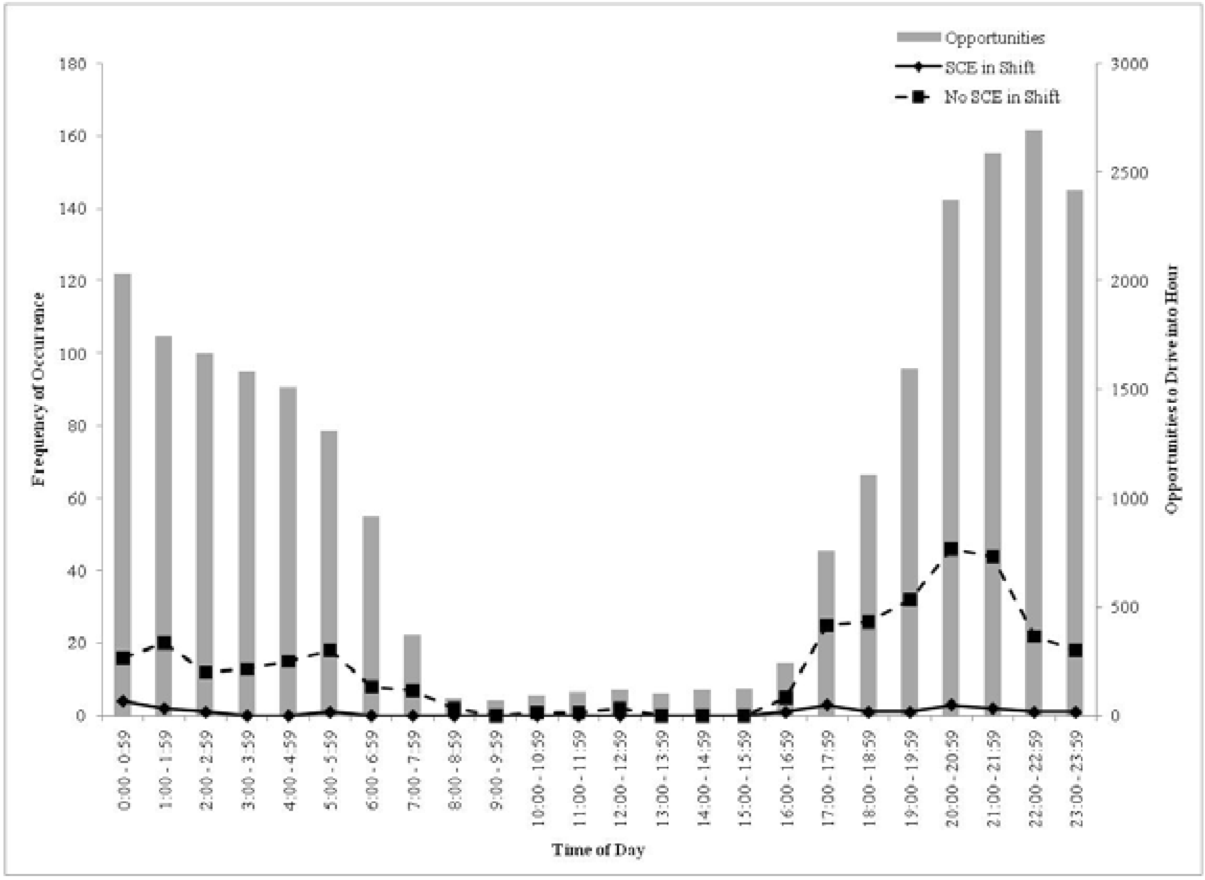


Figure 74. Frequency of Display Brightness Changes by SCE in Shift as a Function of Time of Day

When driver interactions with the DFM were calculated as a function of driving hour (and further decomposed into whether or not the driver had any SCE recorded for their shift; see Table 66), patterns similar to those observed across time of day were present. Drivers with an SCE recorded in their shift had very low levels of interaction with the DFM (regardless of adjustment type) across all driving hours. Drivers with no SCE recorded in their shift displayed a pattern of interacting with the system the most in the initial hours of driving. In general, over the course of the shift, driver interaction with the system gradually decreased.

Table 66. Frequency of Sensitivity Level, Warning Sound, and Display Brightness Changes by SCE in Shift as a Function of Driving Hour

Driving Hour	Sensitivity Level (SCE in Shift)	Sensitivity Level (No SCE in Shift)	Warning Sound (SCE in Shift)	Warning Sound (No SCE in Shift)	Display Brightness (SCE in Shift)	Display Brightness (No SCE in Shift)	Opportunities
1	23	340	6	98	3	72	2,143
2	12	132	2	53	1	33	2,426
3	15	123	5	48	4	22	2,457
4	13	120	5	52	4	25	2,397
5	14	149	6	61	3	17	2,449
6	8	114	7	56	0	17	2,206
7	7	89	2	49	2	20	2,095
8	7	84	1	40	2	24	1,855
9	8	58	0	16	1	27	1,558
10	1	53	0	19	1	19	1,170
11	1	48	0	19	0	12	897
Totals	109	1,310	34	511	21	288	21,653

Specifically, when driver interactions with the DFM sensitivity level was calculated as a function of driving hour (Figure 75), for those drivers with no SCE in their shift, a sharp drop between the initial hour (1) and all subsequent hours was observed. From hour 2 on, the gradual decline was present. This suggests that non-SCE drivers are initially searching for an optimal sensitivity level and generally remaining with the initially chosen sensitivity level unless adjustments are needed later in the driving shift.

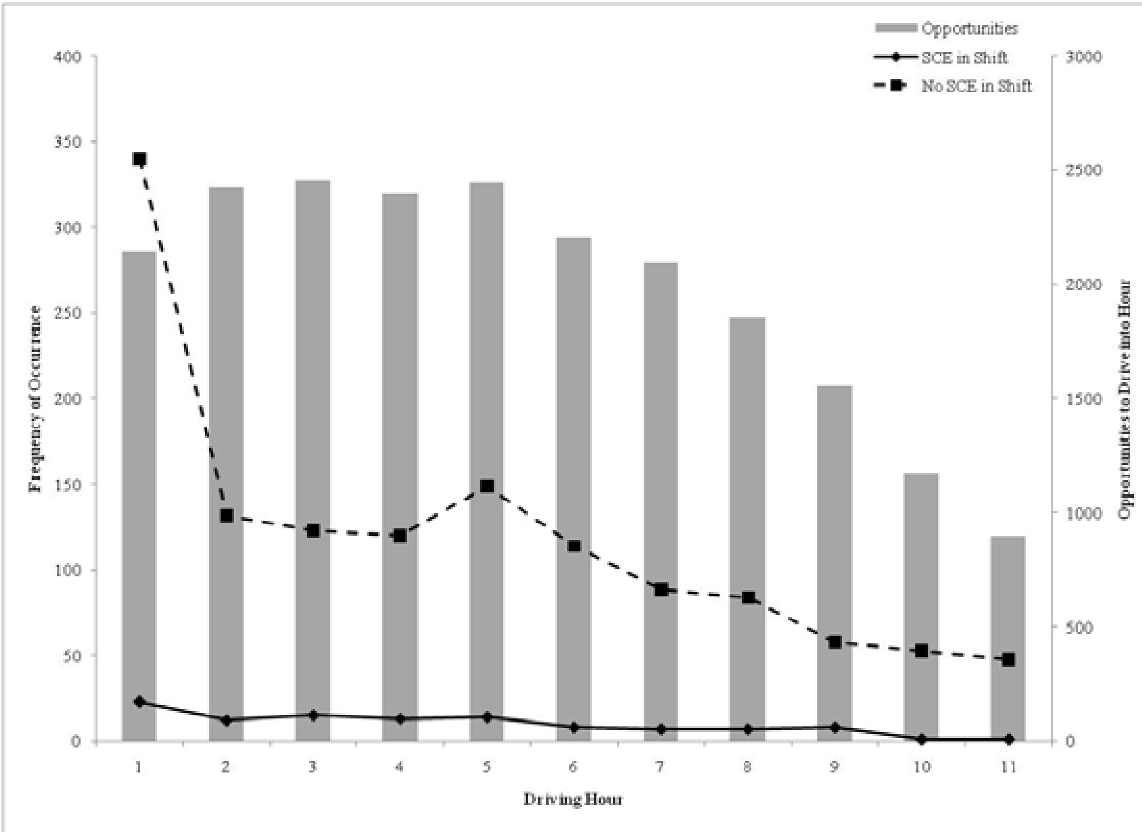


Figure 75. Frequency of Sensitivity Level Changes by SCE in Shift as a Function of Driving Hour

A similar pattern is present in the examination of driver interactions with the DFM warning sound selection as a function of driving hour (Figure 76); for those drivers with no SCEs in their shift, a sharp drop between the initial hour (1) and all subsequent hours was observed. However, drivers still had a relatively high degree of interaction with the warning sound selection throughout most of the shift (until approximately the eighth driving hour). After the eighth hour, interaction with warning sound selection dropped. As seen in other examinations, drivers with SCEs in their shift had a much lower level of interaction with the DFM's various warning sound selections across all driving hours.

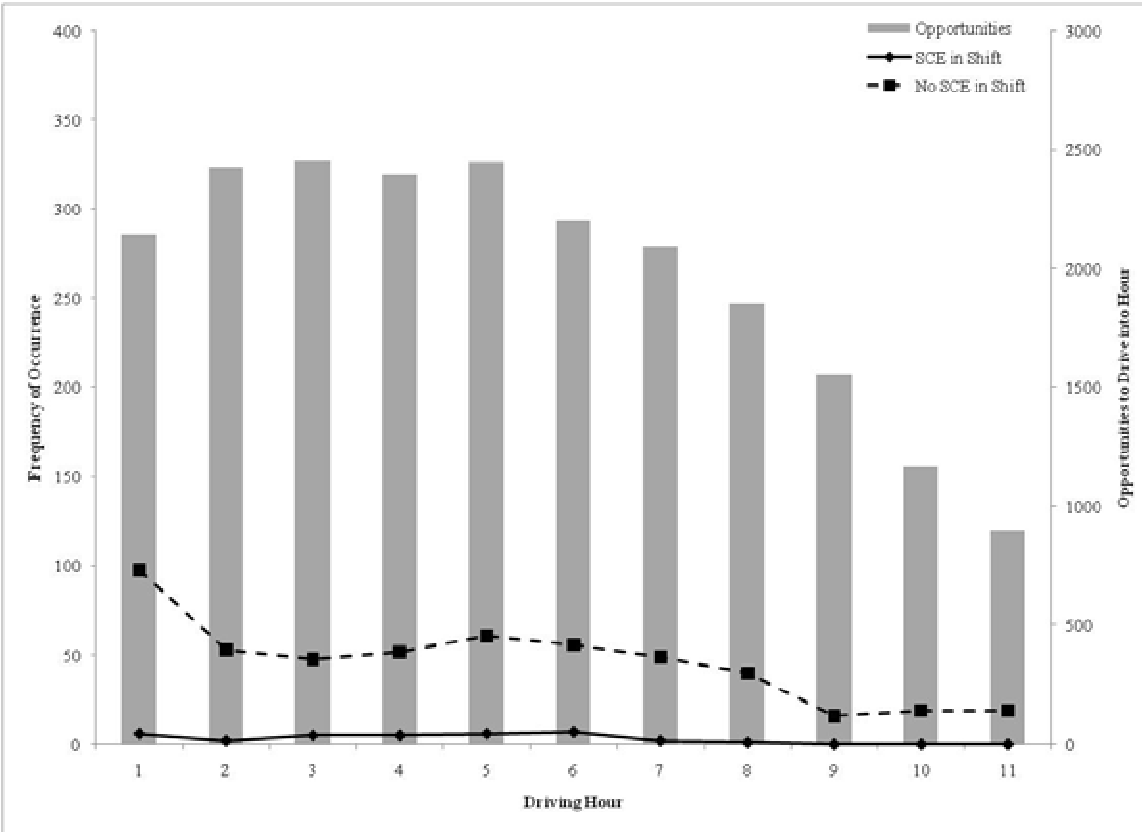


Figure 76. Frequency of Warning Sound Changes by SCE in Shift as a Function of Driving Hour

Driver adjustments to the DFM display brightness were most frequently observed in the initial hour of driving for drivers with no SCE in their shift (Figure 77). From this initial value, a sharp drop-off in the interactions with DFM display brightness was present. This suggests that drivers were making minor adjustments to display brightness as they drove. Additionally, drivers with SCEs in their shift did not display any distinct usage pattern of DFM brightness controls. Only minor adjustments were made by these drivers, primarily between the second and sixth hour of driving.

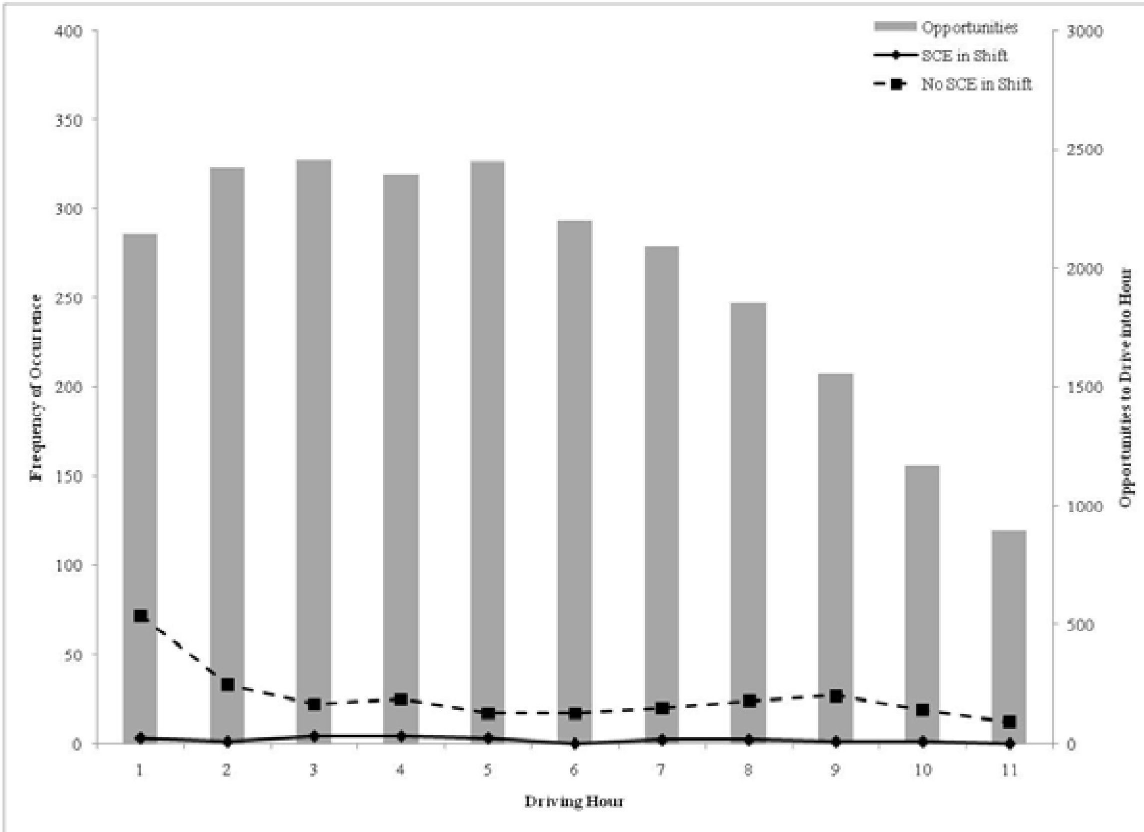


Figure 77. Frequency of Display Brightness Changes by SCE in Shift as a Function of Driving Hour

Driver interaction with the DFM adjustments for sensitivity (Figure 78), warning sound (Figure 79), and display brightness (Figure 80) were calculated as a function of weeks in the study. Drivers with SCEs in their shifts tended to have much less interaction with the DFM, regardless of the week within the study. Drivers with no SCEs in their driving shift tended to have a higher frequency of interaction with DFM settings in the initial weeks of the study, and then a reduction in their interactions with the settings as the study progressed. This trend was most noticeable in the driver adjustments of warning sound and brightness level. A more gradual reduction in non-SCE driver adjustments in sensitivity level was present. This suggests that these drivers were able to quickly distinguish their preferred settings of warning sound and display brightness. However, these drivers were making more frequent adjustments to the sensitivity level in order to find an optimal level.

Table 67. Frequency of Sensitivity Level, Warning Sound, and Display Brightness Changes by SCE in Shift as a Function of Weeks Since DFM Mode Change

Weeks Since DFM Mode Change	Sensitivity Level (SCE in Shift)	Sensitivity Level (No SCE in Shift)	Warning Sound (SCE in Shift)	Warning Sound (No SCE in Shift)	Display Brightness (SCE in Shift)	Display Brightness (No SCE in Shift)	Opportunities
1	17	181	9	135	4	67	80
2	10	133	1	55	1	31	76
3	6	128	5	59	3	20	68
4	10	138	2	27	4	32	76
5	18	136	2	59	2	21	73
6	12	171	2	54	0	34	69
7	15	122	2	24	2	28	65
8	10	117	10	43	3	24	65
9	4	79	0	13	1	10	53
10	1	70	0	33	0	12	53
11	1	59	1	19	0	8	38
12	8	43	0	29	0	23	25
13	4	25	1	9	1	11	17
14	0	7	0	9	0	0	8
15	0	6	0	3	0	9	6
16	0	0	0	2	0	2	4
17	0	0	0	0	0	1	2
Totals	116	1,415	35	573	21	333	778

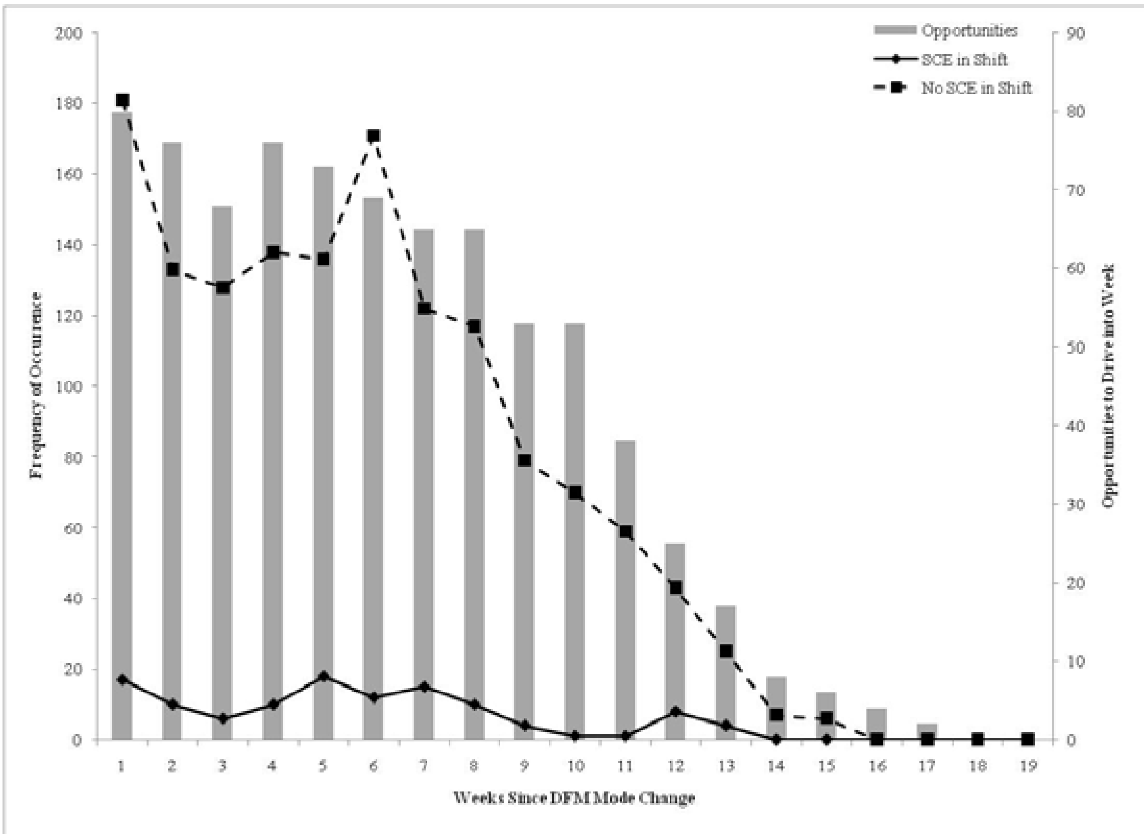


Figure 78. Frequency of Sensitivity Level Changes by SCE in Shift as a Function of Week

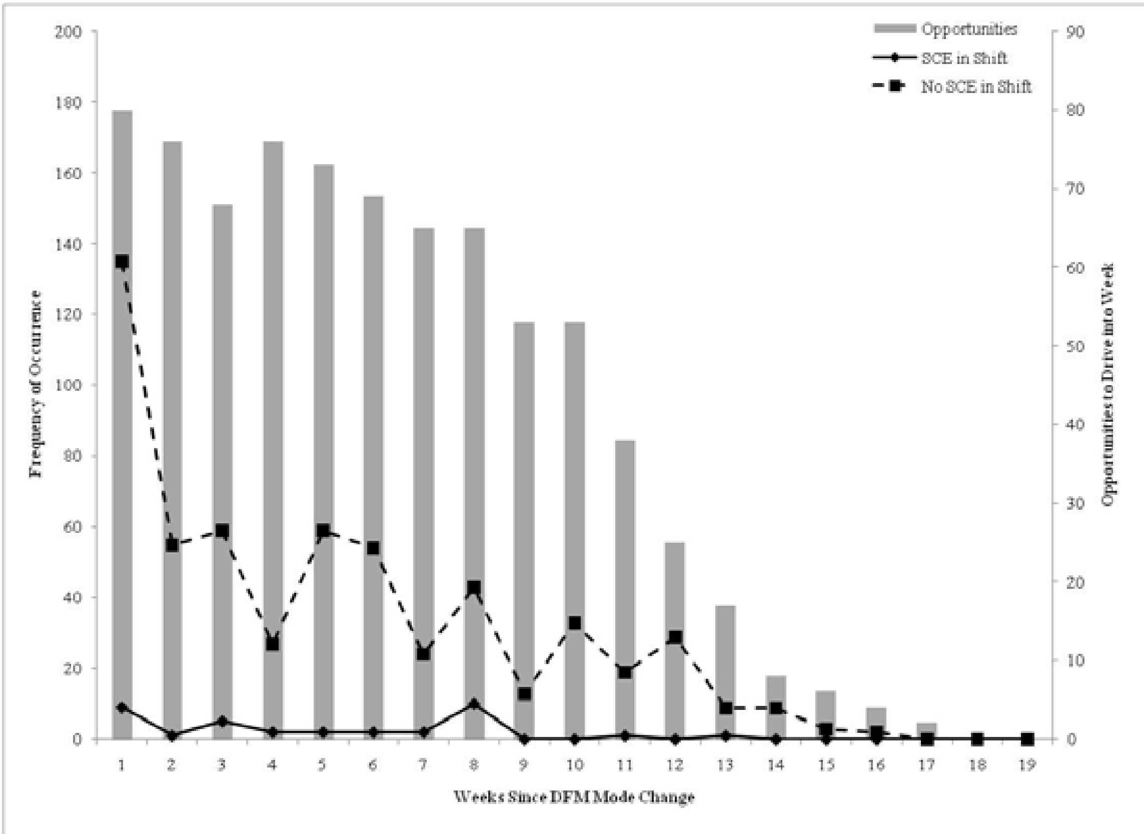


Figure 79. Frequency of Warning Sound Changes by SCE in Shift as a Function of Week

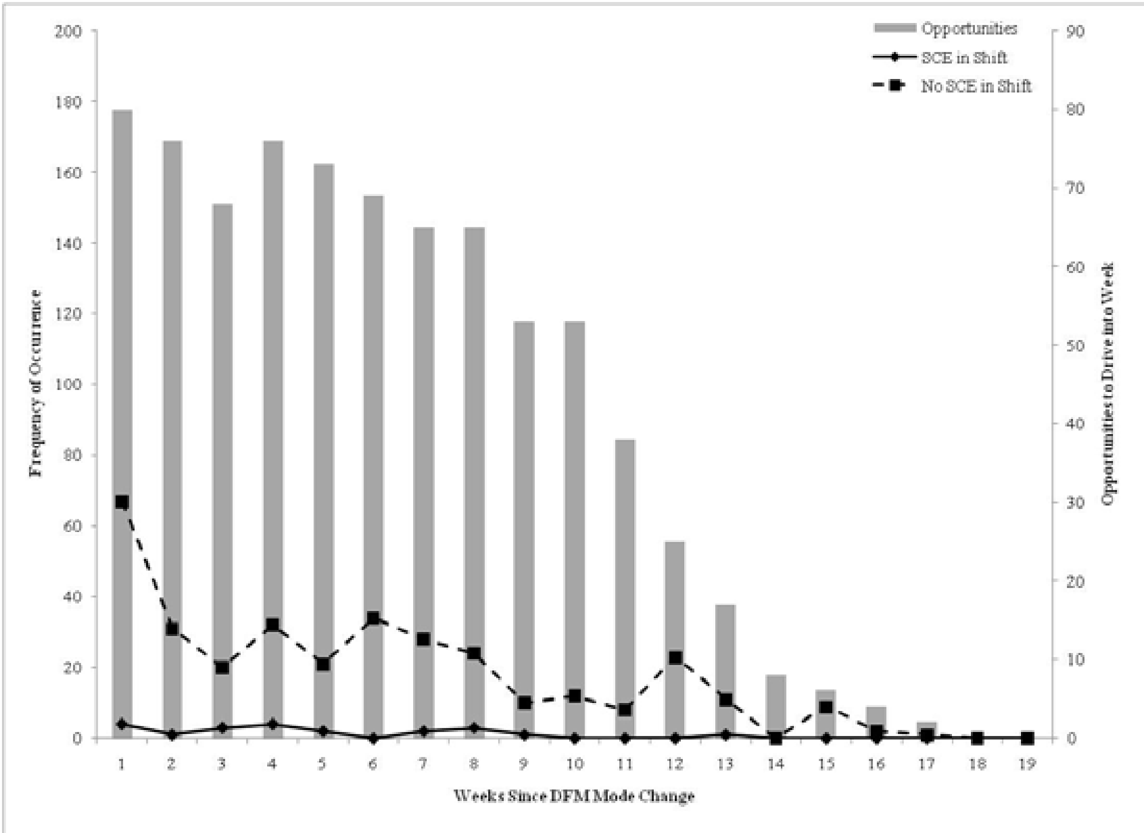


Figure 80. Frequency of Display Brightness Changes by SCE in Shift as a Function of Week

One marker of driver interaction with the DFM is the duration drivers allowed DFM alerts to continue before acknowledging them. This is calculated as a function of time of day, driving hour, and week in the study for drivers with both SCE and without SCE in their driving shift (Table 68).

Table 68. Mean Alert Duration by SCE in Shift as a Function of Time of Day

Time of Day	SCE in Shift (s)	No SCE in Shift (s)	SCE in Shift (N)	No SCE in Shift (N)	Opportunities
0:00 - 0:59	46.96	33.86	142	2,148	2,033
1:00 - 1:59	38.57	41.79	155	1,714	1,749
2:00 - 2:59	72.38	40.17	107	1,551	1,664
3:00 - 3:59	84.38	50.25	70	1,384	1,582
4:00 - 4:59	54.75	37.75	126	1,819	1,512
5:00 - 5:59	61.41	27.91	93	2,248	1,311
6:00 - 6:59	21.13	20.58	67	2,360	917
7:00 - 7:59	9.57	17.08	62	1,345	369
8:00 - 8:59	5.68	30.61	28	256	77
9:00 - 9:59	0.00	21.70	0	31	66
10:00 - 10:59	0.00	4.73	0	38	93
11:00 - 11:59	0.00	4.86	0	57	106
12:00 - 12:59	0.00	32.90	0	24	118
13:00 - 13:59	0.00	9.11	0	18	100
14:00 - 14:59	1.30	6.93	1	20	117
15:00 - 15:59	0.00	12.60	0	25	120
16:00 - 16:59	2.20	8.54	42	238	245
17:00 - 17:59	15.74	20.25	151	1,777	757
18:00 - 18:59	64.88	21.86	189	2,571	1,108
19:00 - 19:59	46.21	18.02	230	3,366	1,596
20:00 - 20:59	42.74	15.73	236	5,278	2,372
21:00 - 21:59	25.69	20.30	203	4,160	2,585
22:00 - 22:59	35.77	26.45	200	3,476	2,694
23:00 - 23:59	34.91	26.48	190	3,102	2,416

When viewed as a function of time of day, drivers with SCEs in their shift tended to allow the DFM alerts to continue the longest in the 2:00-2:59 and 3:00-3:59 hour blocks (Figure 81). Another peak in mean alert duration was observed at the 18:00-18:59 hour block. Drivers without SCEs occurring in their shift tended to acknowledge the alerts faster. However, they also displayed an increase in reaction time that followed the same temporal pattern as the SCE drivers. Similar reasons as mentioned previously for data points that seem to be outliers during the high illumination hours apply to this case as well. The data points in the middle of the day were checked; they usually occurred during adverse weather or in tunnels where the illumination level drops and the DFM becomes operational.

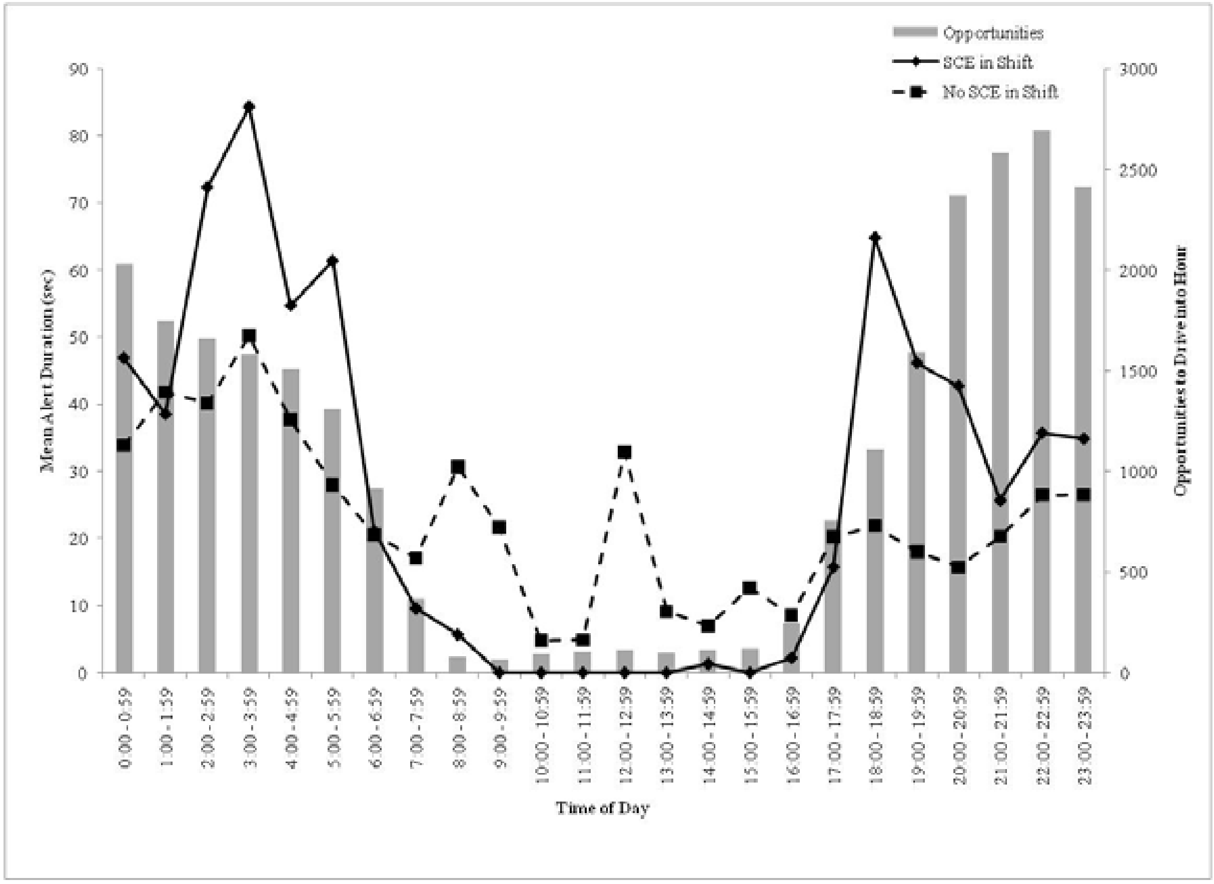


Figure 81. Mean Alert Duration by SCE in Shift as a Function of Time of Day

For the drivers with SCEs in their shifts, no clear pattern emerges when mean alert duration is calculated as a function of driving hour. These drivers, however, demonstrated globally higher reaction times (in the form of increased mean alert duration) as compared to the drivers without SCEs in their driving shifts (Table 69). Drivers without SCEs in their shifts tended to have a relatively unchanging mean alert duration for most of the driving shift, with the mean alert duration dropping toward the final driving hours (Figure 82).

Table 69. Mean Alert Duration by SCE in Shift as a Function of Driving Hour

Driving Hour	SCE in Shift (s)	No SCE in Shift (s)	SCE in Shift (N)	No SCE in Shift (N)	Opportunities
1	37.38	25.27	163	3,585	2,143
2	63.54	28.11	151	3,310	2,426
3	61.53	33.33	176	2,870	2,457
4	37.28	32.32	198	3,047	2,397
5	50.58	32.65	170	3,433	2,449
6	41.40	32.81	190	3,273	2,206
7	59.08	29.96	214	3,170	2,095
8	34.24	27.62	194	3,295	1,855
9	39.67	18.36	110	2,987	1,558
10	52.21	13.38	115	2,283	1,170
11	45.77	15.32	69	1,757	897

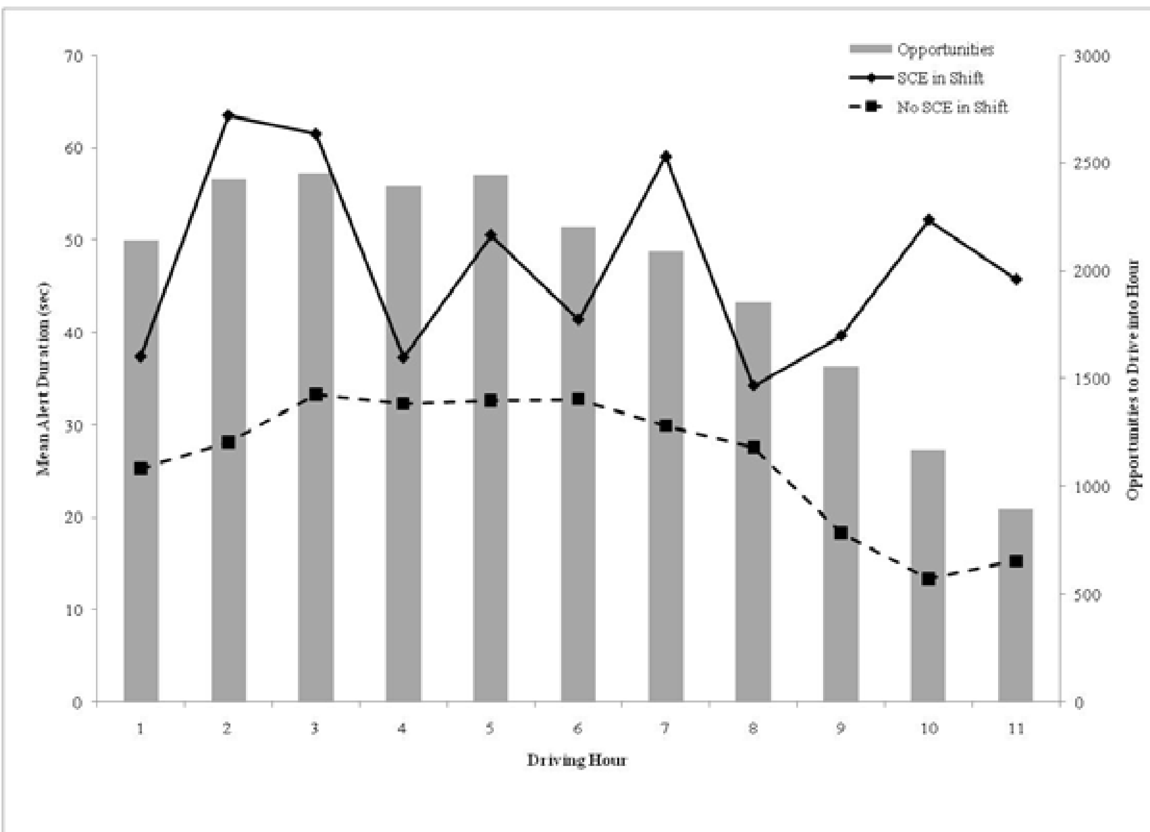


Figure 82. Mean Alert Duration by SCE in Shift as a Function of Driving Hour

Similarly, drivers with an SCE in their shifts displayed a higher global pattern of mean alert duration as compared to drivers without an SCE in their shifts (Figure 83) when viewed as a function of week in the study since the DFM mode was changed (Table 70). Drivers with an SCE in their shift not only had a higher mean alert duration but also had a greater amount of variability in their reaction time to these alerts. Drivers without an SCE in their shift did not display any specific pattern in mean alert duration. Note that a single (non-SCE-in-shift) driver caused a peak in mean alert duration between weeks 12 and 15. As mentioned earlier in this section, upon further examination, it was revealed this was the result of a single driver who chose to completely reduce the volume of DFM auditory alerts and to allow alerts to continue to sound. The behavior from this driver also contributed to the increase in mean alert duration in weeks 12 and 13; the driver did not drive past week 14.

Table 70. Mean Alert Duration by SCE in Shift as a Function of Weeks in Study Since DFM Mode Change

Weeks Since DFM Mode Change	SCE in Shift	No SCE in Shift	N SCE in Shift	N No SCE in Shift	Opportunities
1	8.19	51.42	68	3,331	80
2	22.37	18.57	139	4,040	76
3	47.02	9.68	342	3,702	68
4	62.13	12.51	250	3,087	76
5	10.44	24.08	346	3,813	73
6	19.39	21.22	416	4,363	69
7	39.46	17.44	186	3,224	65
8	89.26	22.36	321	3,921	65
9	11.57	23.76	38	2,515	53
10	9.25	27.51	2	2,873	53
11	80.24	8.95	86	1,959	38
12	31.35	48.12	66	975	25
13	104.27	84.99	30	501	17
14	0.0	231.87	0	342	8
15	10.05	36.99	2	293	6
16	0.0	15.66	0	65	4
17	0.0	1.90	0	1	2
18	0.0	0.0	0	0	0
19	0.0	3.50	0	1	1

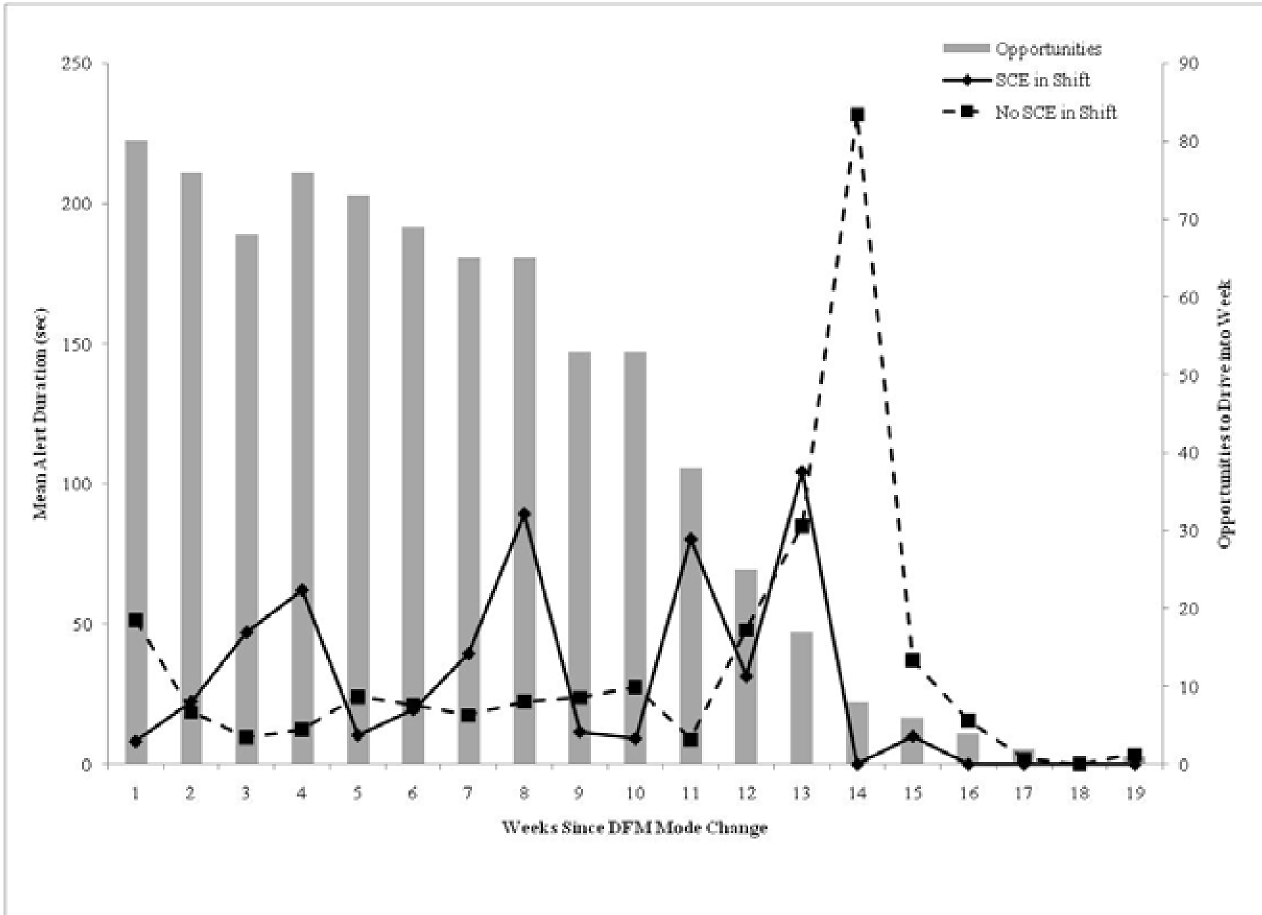


Figure 83. Mean Alert Duration by SCE in Shift as a Function of Week

Discussion

Overall, some differences in driver interaction with the DFM were present. The majority of drivers did not take advantage of the range of adjustments provided by the DFM's various sensitivity levels, instead choosing to leave the device set to the active mode default level. Similar patterns were present for driver interactions with the selection of a warning sound, with most choosing to leave the device set to the initial setting. Display brightness tended to be split between the two limits of adjustment, suggesting that most drivers were not taking advantage of the full range of adjustment or did not feel that it was needed. All of these device interactions, however, were different between drivers with and without SCEs within their shift. Drivers with an SCE occurring in their shift tended to have globally lower levels of interaction with the system in terms of the sensitivity, warning sound, and display brightness adjustments. These interactions decline in both SCE and non-SCE drivers as their participation/exposure to the system increased. Therefore, it is possible that there is an adjustment and learning period for the driver while he or she becomes accustomed to the operation of the DFM and use of a DDWS.

RESEARCH QUESTION 5: FAVORING DRIVERS - FOLLOW-UP

As part of this section the following research question (RQ) was answered:

> RQ 5.1: How did the DDWS operate for drivers who rated the system positively and negatively in the post-study survey?

Methods, analysis, measures of interest, results, and discussion related to this question are presented next.

Methods

Using the van der Laan composite scale (van der Laan, Heino, & De Waard, 1997) VNTSC identified extreme "favorability" attitudes. A rating of zero represents a neutral attitude, therefore only the drivers whose composite scores were greater than or equal to one for favoring attitudes and less than or equal to negative one for disfavoring attitudes were used (Figure 84). There were 11 drivers who had a favoring attitude towards the DDWS and 5 drivers who had a disfavoring attitude towards the DDWS. However, two of those drivers were not part of the original subset of 96 drivers considered for this study. This gives a total of 10 drivers with a favoring attitude towards the DDWS and 4 drivers with a disfavoring attitude towards the DDWS. Due to the small sample size (n = 14), only descriptive statistics were performed for this research question. Therefore, caution is advised when attempting to generalize from the results.

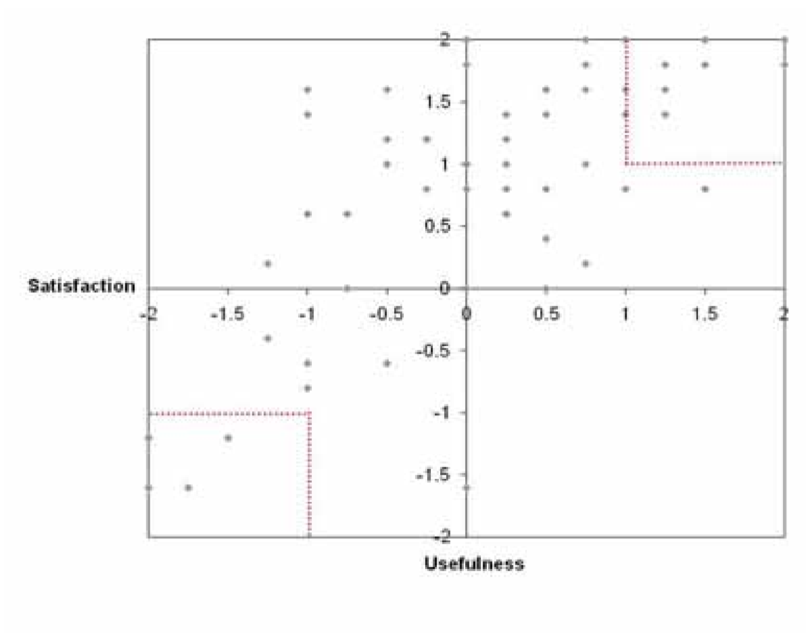


Figure 84. Results on van der Laan Composite Scale of Satisfaction and Usefulness. Based on All Drivers in Study (n = 96)

Data from the subsets of 14 drivers who had very favoring and very disfavoring ratings were compared with respect to main areas evaluated under Research Questions 1 through 4:

1. On-the-job drowsiness
2. Sleep hygiene
3. Involvement in SCEs
4. Human-machine interaction

In additions to the main topics above, the following areas were covered as well:

- Characteristics
    - Driver Characteristics
    - Operational Characteristics
- System Reliability

Analysis

Descriptive statistics, graphs, and tables were used to identify tendencies in the data following a format similar to the one used for the previous research questions. As mentioned previously, the small number of participants with extreme attitudes precludes a more thorough analysis.

Measures

Characteristics. In order to examine the driver characteristics, the drivers' age (in years) and experience (in years) was analyzed. Additionally, characteristics of the type of operation in which the driver worked were examined in order to identify any patterns related to the haul type (i.e., long- versus line-haul).

System Reliability. How well the system worked for a particular driver may have influenced the rating the driver gave the DFM. For this reason, the false alert rate for each group (favoring and disfavoring ratings) was compared.

Previous Research Question Comparison. The results from Research Questions 1, 2, 3, and 4 were re-examined for these two subsets of drivers (i.e., with favoring and disfavoring ratings). Difference in safety benefits between the two subsets were examined as well.

Results

Characteristics

Driver Characteristics. The years of driving experience for the two subsets of drivers ranged from 5-18 years for the drivers who gave the DFM a disfavoring rating (n = 4) and from 1-21 years for the ones who gave a favoring rating (n = 10). Figure 85 presents the mean age for each subset of drivers based on their DFM rating compared to the overall driver age experience for the 96 drivers who were considered for the previous analyses. On average, drivers with favoring opinions of the DFM had three fewer years of experience than drivers with a disfavoring rating of the DFM. Additionally, in the overall study data set (beyond the present analysis), drivers who gave favoring ratings of the DFM had two fewer years of experience than the average driver in the study. There were very few drivers in the two extreme subsets; therefore it is difficult to reach any conclusions regarding the relationship between a driver's years of experience and DFM rating.

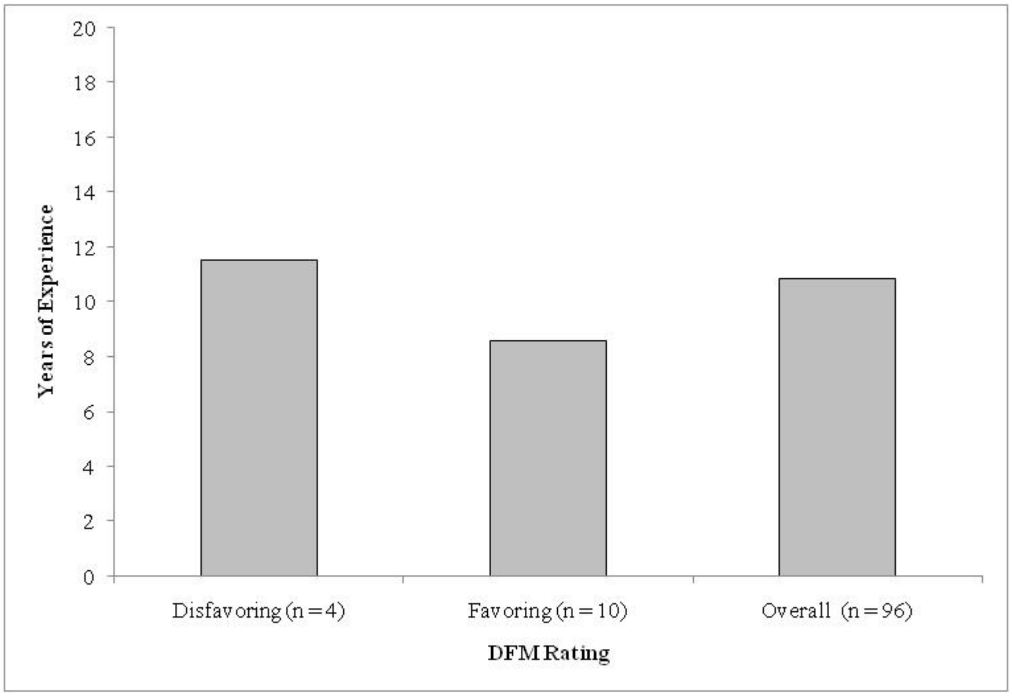


Figure 85. Mean Years of Experience by Driver DFM Rating. Based on All Drivers in Study (n = 96)

Figure 86 shows the mean age of drivers with favoring and disfavoring opinions of the system, as well as the overall average age of drivers in the study. On average, drivers with a favoring opinion of the DFM were 40 years old, and the drivers with a disfavoring opinion of the system were 41. Both subsets of drivers had a similar age, which was very similar to the overall study's average age.

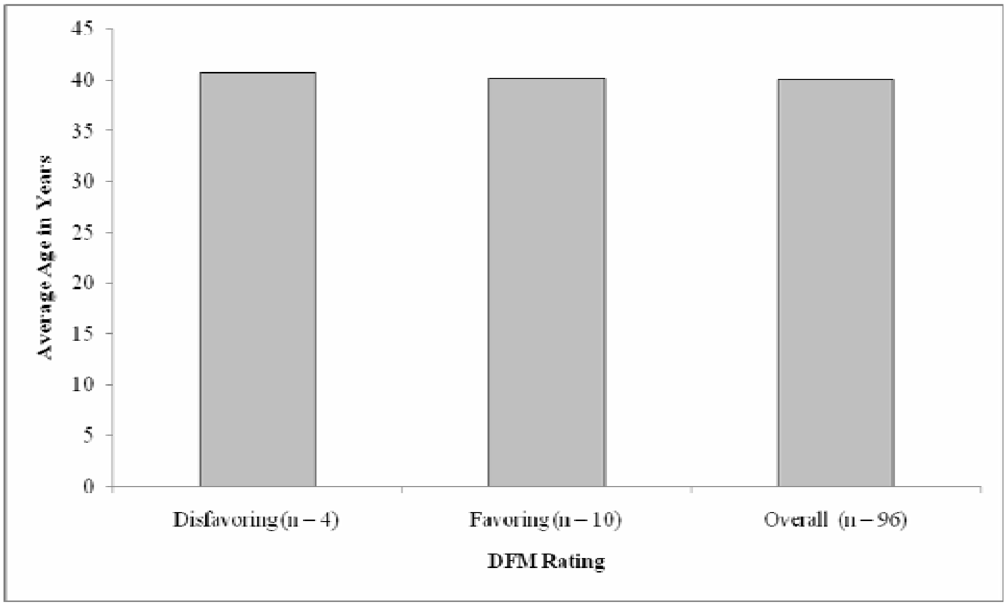


Figure 86. Mean Driver Age by Driver DFM Rating. Based on All Drivers in Study (n = 96)

Operational Characteristics. Among the drivers who gave the DDWS a favoring rating, all were long-haul drivers; however, among the drivers who gave the DDWS a disfavoring rating, half were line-haul while the remaining were long-haul. It should be noted that only four drivers gave the DDWS a disfavoring rating; therefore no conclusions should be drawn.

System Reliability

The valid alert rate was calculated as a function of driver rating of the DFM system (Table 71). A total of 15,386 alerts were reviewed, of which 721 were valid alerts. Among the drivers with a disfavoring rating of the system there were 700 alerts, of which 7 were valid. Drivers with a favoring rating of the system had 1,722 alerts, of which 104 were valid. A comparison of the 95 percent confidence intervals of the validity rates demonstrated that the validity rate for disfavoring drivers was significantly lower than both the validity rate of the drivers with a favoring rating of the DFM and the drivers overall. The validity rate for the drivers with a favoring rating of the system did not differ significantly from drivers in the overall study. However, drivers tended to have disfavoring ratings of the system when their validity rate was lower than the overall average for the study.

Table 71. Reviewed Alerts and Valid Alerts by DFM Rating

DFM Rating	Reviewed Alerts	Valid Alerts	Validity Rate	SE	Lower Confidence Level (LCL)*	Upper Confidence Level (UCL)*
Favoring (n = 10)	1,722	104	6.04	0.574	4.91	7.16
Disfavoring (n = 4)	700	7	1.00	0.376	0.26	1.74
Overall (n = 96)	15,386	721	4.69	0.170	4.35	5.02

* Observations were assumed to be independent, and the valid alerts were assumed to follow a binomial distribution. The confidence limits were constructed using a normal approximation of the binomial distribution.

On-The-Job Drowsiness

DFM PERCLOS Values. As in Research Question 1, both the DFM PERCLOS collected from the system and the manual PERCLOS values calculated from the NSTSCE baseline events were used for this analysis. Figure 87 presents an overview of the DFM PERCLOS values that characterize each experimental condition and DFM rating. The minimum DFM PERCLOS value for all combinations of experimental condition and DFM rating was zero. The maximum ranged from 29 to 63. For both the baseline Test and Test experimental condition, the drivers with a disfavoring rating of the DFM had much lower maximum values than those with a favoring rating of the DFM. As before, DFM PERCLOS values might be affected by scanning behaviors (i.e., eyes off of forward roadway). Therefore measures of central tendency give a more realistic and accurate representation. For both the mean and median, the drivers with a disfavoring rating of the DFM had higher DFM PERCLOS values than the drivers with a favoring rating of the system.

Evaluating the mean DFM PERCLOS values by driving hours showed that drivers in the test condition with a favoring rating of the DFM had, on average, lower DFM PERCLOS values throughout the hours driving in the study (Figure 88). These findings also show that drivers in the test condition with a disfavoring rating of the DFM had on average higher DFM PERCLOS values than the drivers in the same experimental condition with a favoring rating of the DFM. The data point in Figure 88 that represents the baseline test Favoring condition for hours 350-399 is composed of 2 observations from the same participant (Driver 129). This driver received only 11 alerts over the duration of the study, of which none were valid. As before, the DAS only collected the DFM output and not potential internal failure messages for software and hardware in the DFM. Therefore, it is not possible to determine with certainty that these two observations are outliers. The 95 percent confidence limits for the mean (Table 72) show that the difference between the baseline test favoring and disfavoring ratings was not significant. However, during the test condition, drivers with a disfavoring rating of the system had a higher mean manual PERCLOS than drivers with a favoring rating of the system. The difference between the favoring and disfavoring subsets was substantial.

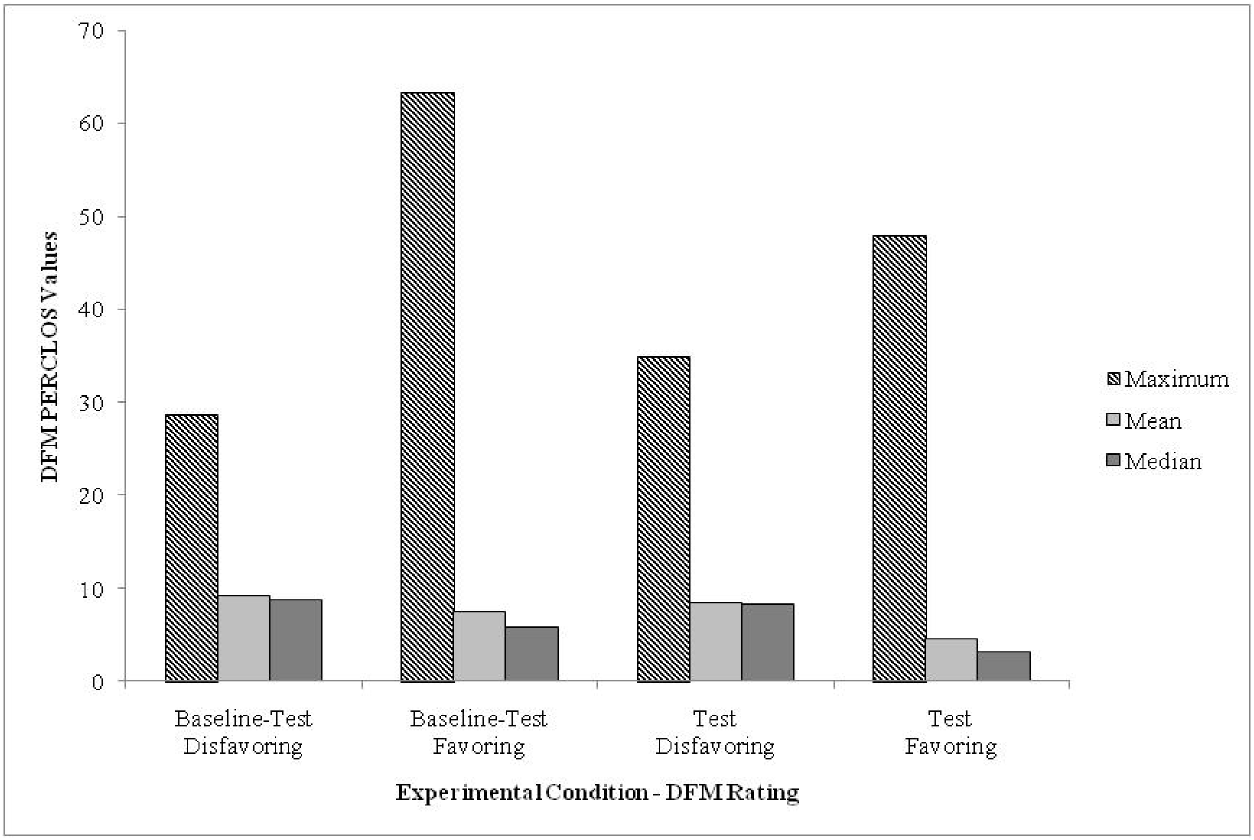


Figure 87. DFM PERCLOS Values by Experimental Group and DFM Rating. Based on a Subset of Drivers (n = 14)

Table 72. Mean DFM PERCLOS Values With 95 Percent Confidence Limits by DFM Rating

Experimental Condition	DFM Rating	N	Mean	SE	LCL*	UCL*
Baseline Test	Disfavoring	60	9.11	0.72	7.69	10.53
Baseline Test	Favoring	192	7.42	0.58	6.29	8.55
Test	Disfavoring	184	8.40	0.43	7.57	9.24
Test	Favoring	332	4.48	0.26	3.96	4.99

*Observations are taken from the same driver and are therefore not independent. Based on a subset of drivers, n = 14.

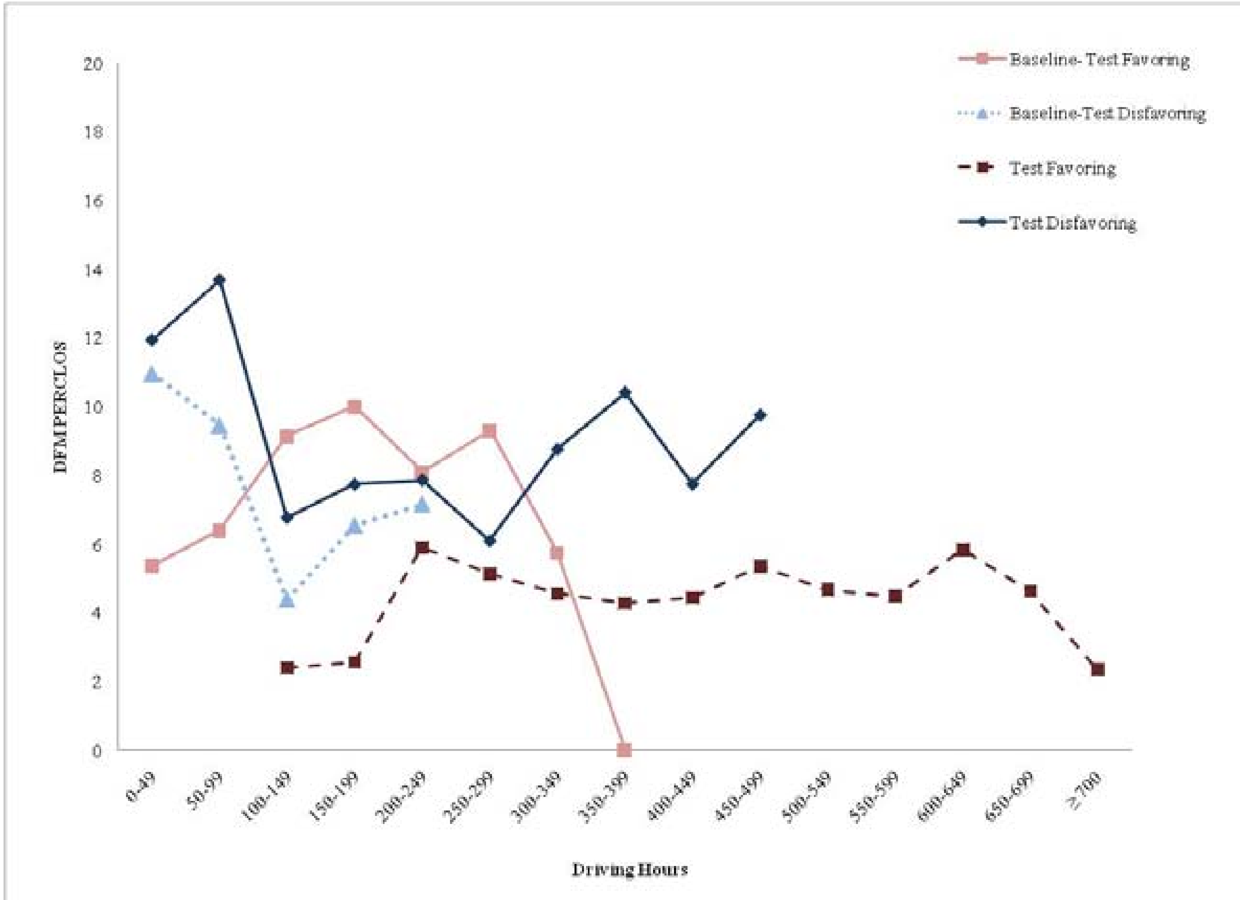


Figure 88. DFM PERCLOS Values by Driving Hours in Study and DFM Rating. Based on a Subset of Drivers (n = 14)

Figure 89 shows DFM PERCLOS values by days in the study and DFM rating. A similar pattern as the one discussed for Figure 88 above is presented for days in the study. On average, the drivers with a favoring rating of the DFM in the test condition had a lower DFM PERCLOS value than the drivers with a disfavoring rating. This low DFM PERCLOS pattern remains regardless of which day of the week the DFM value was obtained (Figure 90). When compared to the drivers who gave a disfavoring rating to the DFM, drivers with a favoring rating had a higher DFM PERCLOS value than the drivers with a favoring rating throughout the days in the study and regardless of day of the week. The baseline for the favoring rating show that the drivers who gave a favoring rating to the DFM had a higher DFM PERCLOS value in the beginning of the study and that it decreased over time. This was not the case for the drivers with a disfavoring rating of the DFM.

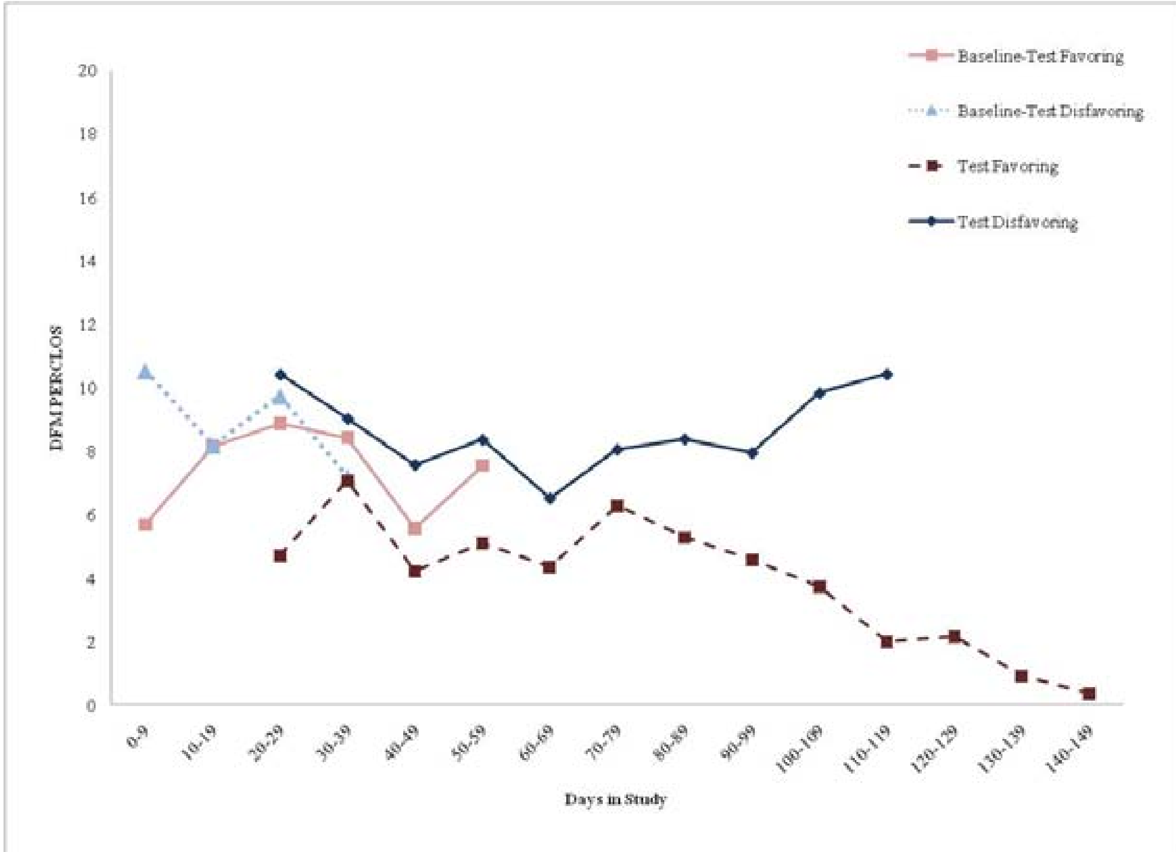


Figure 89. DFM PERCLOS Values by Days in Study and DFM Rating. Based on a Subset of Drivers (n = 14)

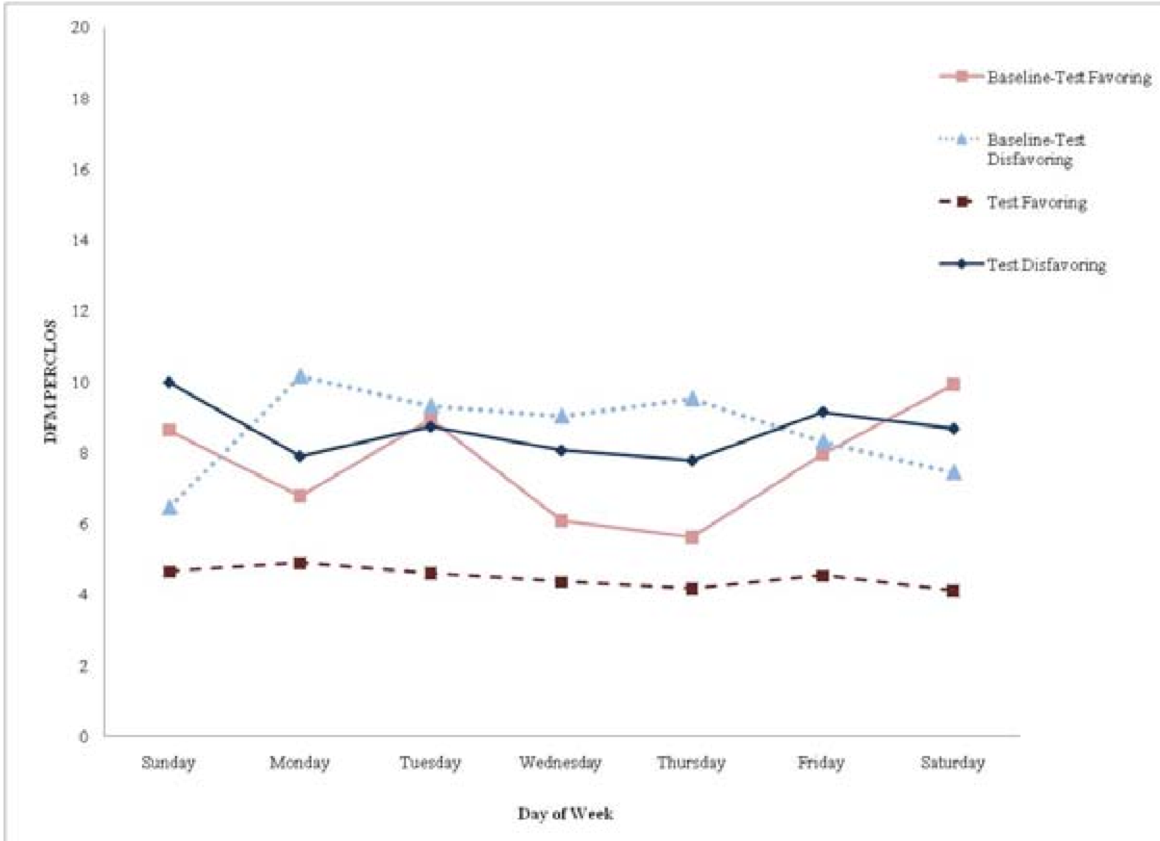


Figure 90. DFM PERCLOS Values by Day of Week and DFM Rating. Based on a Subset of Drivers (n = 14)

The time perspective obtained by evaluating the DFM PERCLOS values by weeks in the study (Figure 91) demonstrates the same pattern from previous analyses (see Figure 89). Drivers in the test condition who gave a favoring rating had a lower average DFM PERCLOS value. Additionally, drivers in the Test Experimental Condition with a disfavoring rating of the DFM had a higher average DFM PERCLOS value.

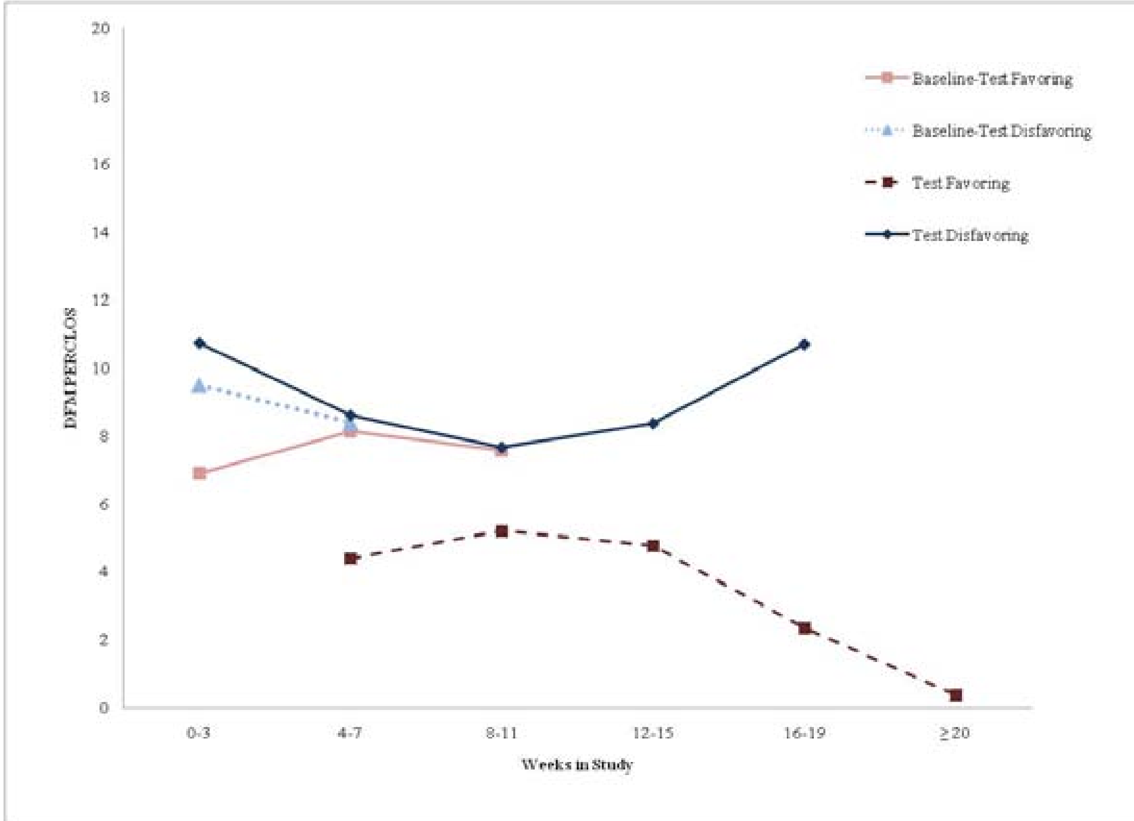


Figure 91. DFM PERCLOS Values by Weeks in Study and DFM Rating. Based on a Subset of Drivers (n = 14)

Manual PERCLOS Values. Figure 92 presents an overview of the manual PERCLOS values calculated for the NSTSCE baseline events. The maximum value obtained from the manual PERCLOS values are lower than those recorded from the DFM, with the maximum manual PERCLOS ranging from 24 to 31. The measures of central tendency show that, during the baseline test Condition, drivers with a disfavoring rating of the DFM had an average manual PERCLOS value similar to the test condition. However for the drivers with a favoring rating, the manual PERCLOS during their test condition was significantly lower than during their baseline. The 95 percent confidence limits for the mean (Table 73) showed that the difference between the baseline test favoring and disfavoring ratings were not significant. However, during the test condition, drivers with a disfavoring rating of the system had a higher mean PERCLOS than drivers with a favoring rating of the system. The difference between the favoring and disfavoring subsets was substantial.

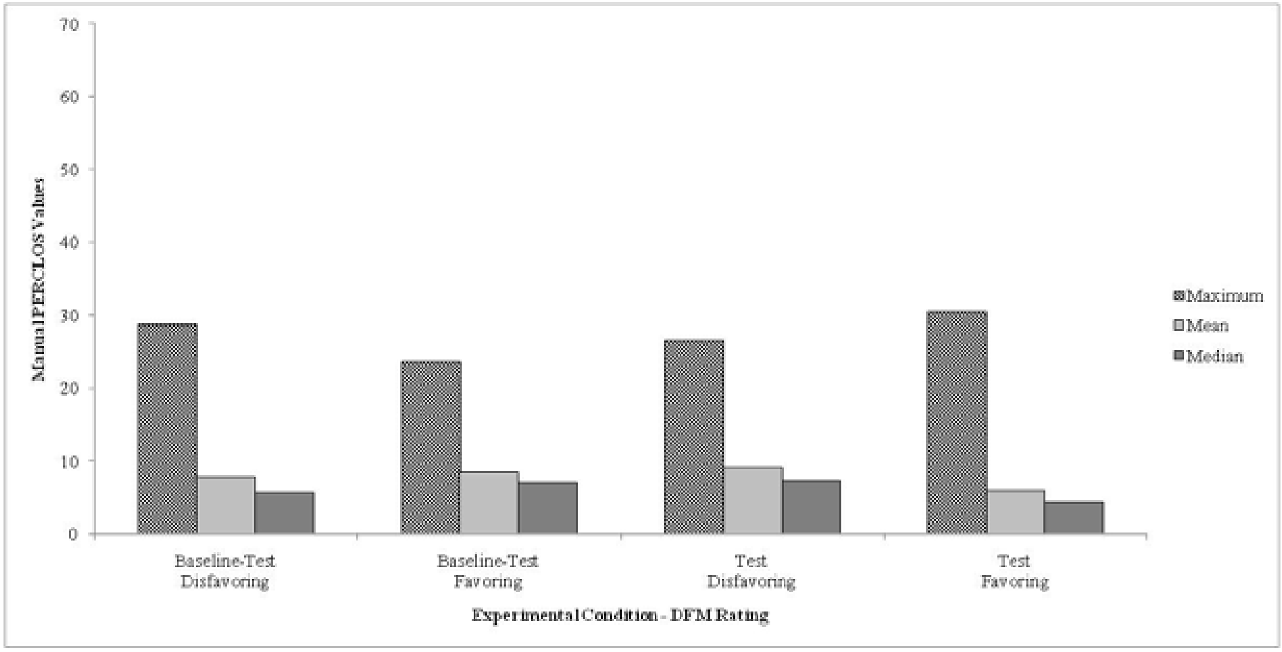


Figure 92. Manual PERCLOS Values by Experimental Group and DFM Rating.

Based on a Subset of Drivers (n = 14)

Table 73. Mean Manual PERCLOS Values With Lower and Upper Confidence Limits

Experimental Condition	DFM Rating	N	Mean	SE	LCL*	UCL*
Baseline Test	Disfavoring	22	7.80	1.50	4.86	10.74
Baseline Test	Favoring	63	8.43	0.71	7.05	9.82
Test	Disfavoring	47	9.25	1.10	7.10	11.41
Test	Favoring	114	5.97	0.54	4.91	7.03

*Observations are taken from the same driver and are therefore not independent. Based on a subset of drivers, n = 14.

Evaluating the mean manual PERCLOS as a function of driving hours demonstrates no discernable pattern (Figure 93). When viewing a time progression of manual PERCLOS values by days in the study, a difference in driver favorability is revealed. The drivers in the test condition who gave favoring ratings of the DFM had a lower mean manual PERCLOS value. The pattern seen in Figure 94 holds for every day of the week except for Saturday (Figure 95).

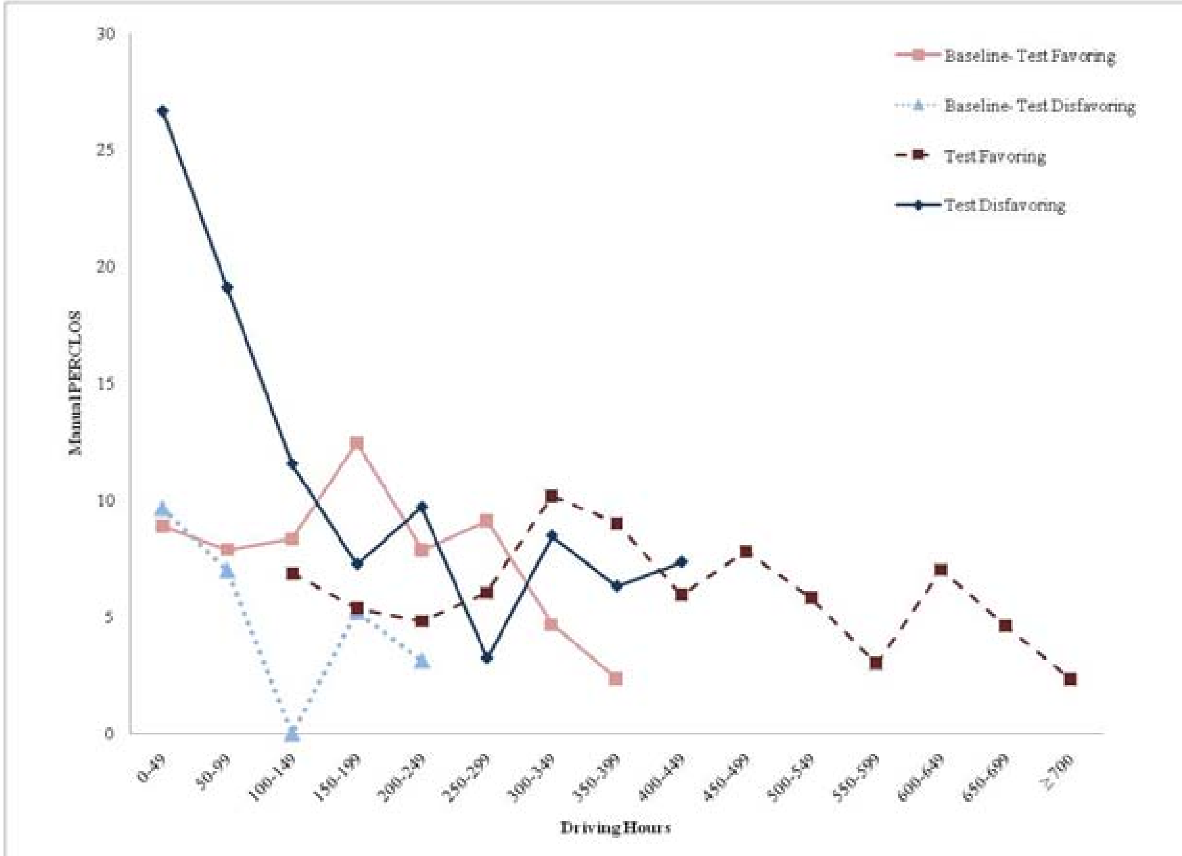


Figure 93. Manual PERCLOS Values by Driving Hours in Study and DFM Rating. Based on a Subset of Drivers (n = 14)

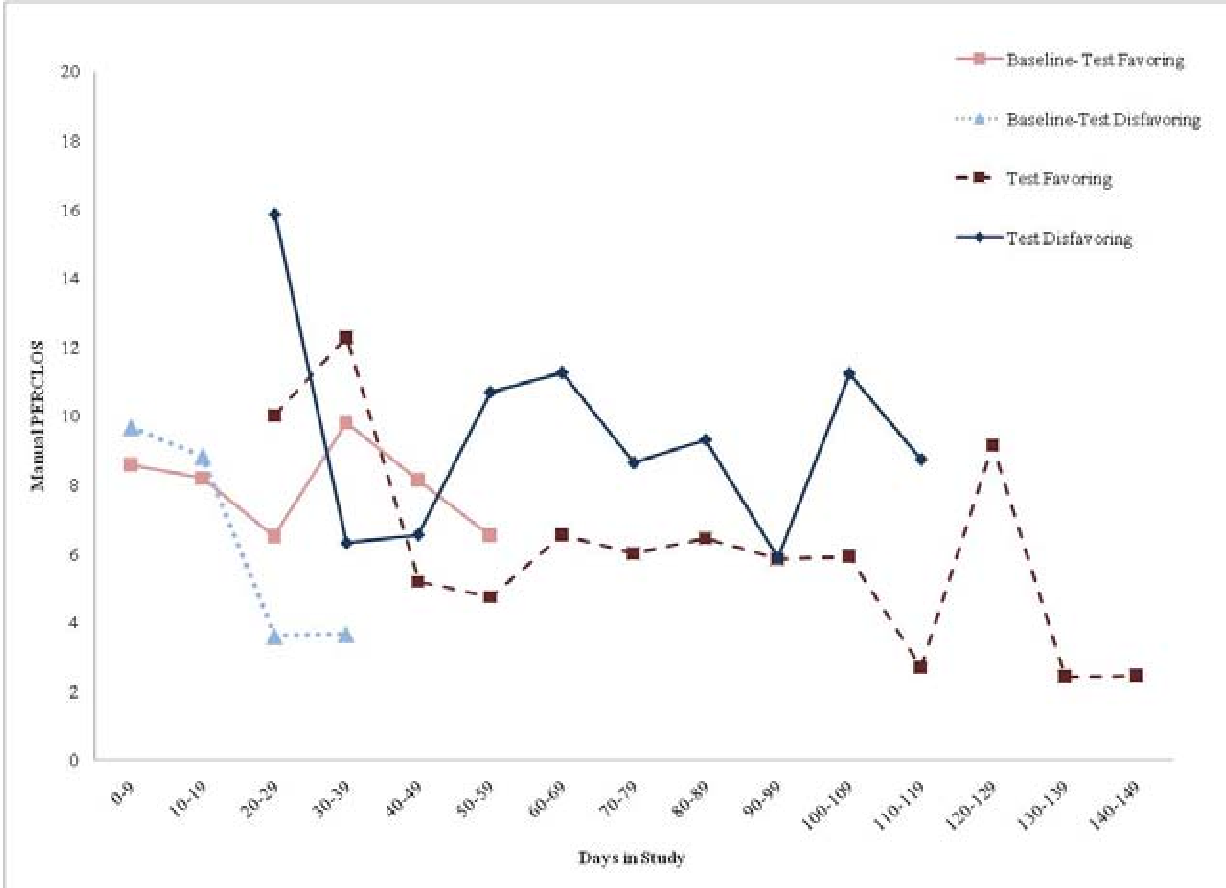


Figure 94. Manual PERCLOS Values by Days in Study and DFM Rating. Based on a Subset of Drivers (n = 14)

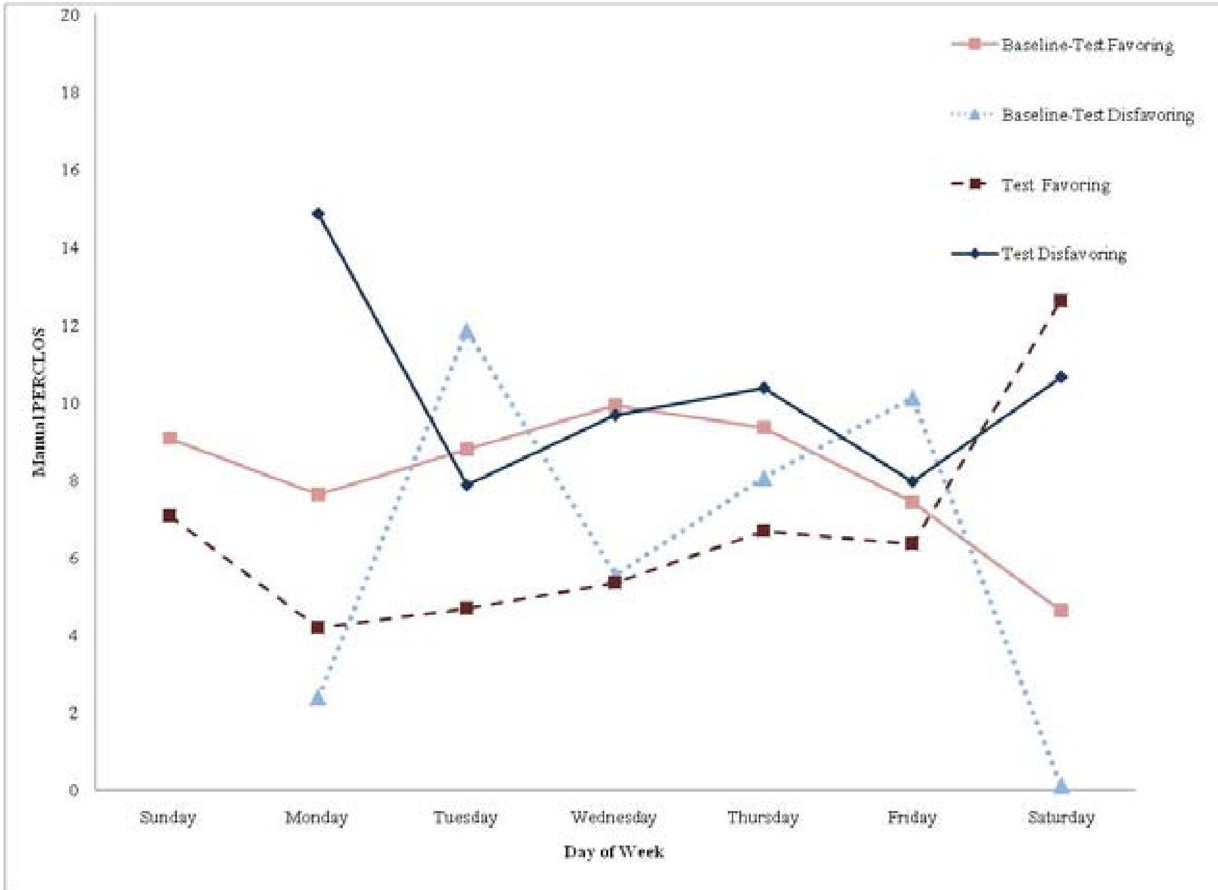


Figure 95. Manual PERCLOS Values by Day of the Week and DFM Rating. Based on a Subset of Drivers (n = 14)

Figure 96 presents the manual PERCLOS values as a function of weeks of participation in the study. This also presents a similar pattern as those described previously, where most drivers with a favoring rating for the DFM had a lower mean manual PERCLOS value when compared to the baseline and to the drivers with a disfavoring rating of the system during the test condition.

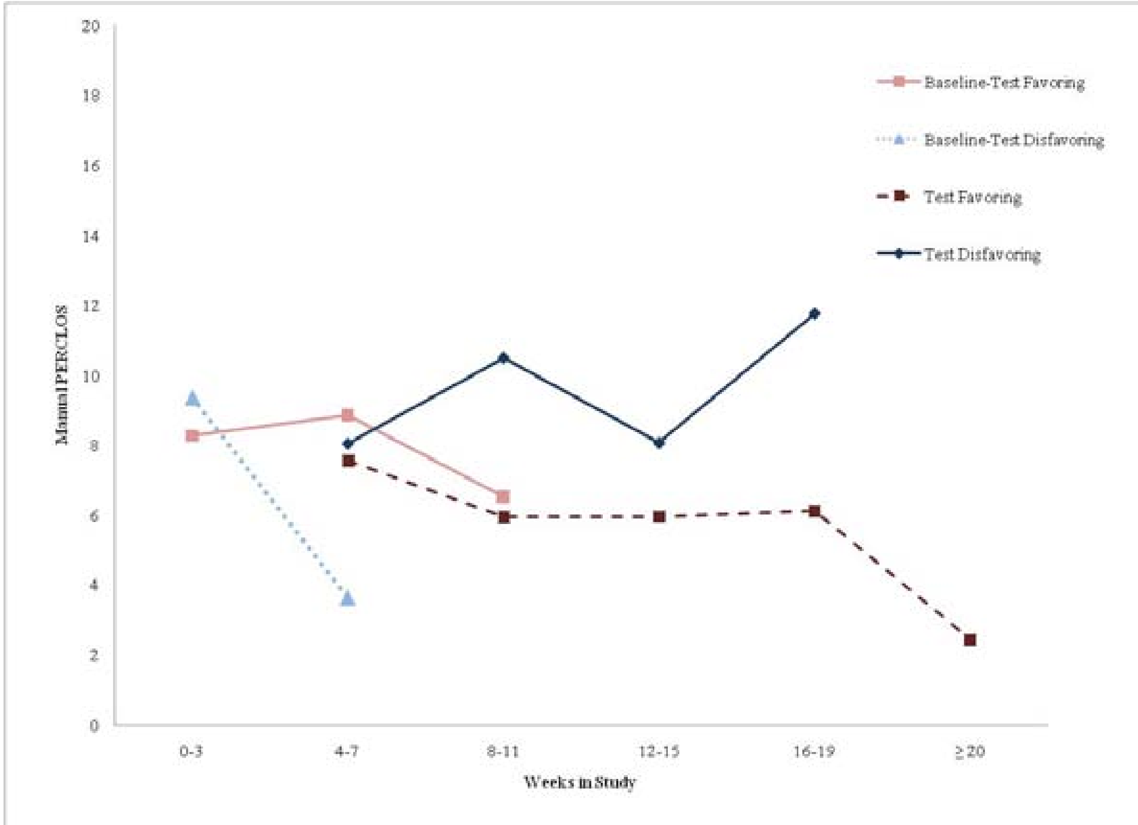


Figure 96. Manual PERCLOS Values by Weeks in Study and DFM Rating. Based on a Subset of Drivers (n = 14)

Valid Alerts. The results from the previous analysis provide an overview of on-the-job-drowsiness by examining DFM PERCLOS and manual PERCLOS. DFM alerts were based on the PERCLOS calculated by the system. If the PERCLOS values did not reach the predetermined threshold the system did not provide the driver with an alert. However, it cannot be assumed that because a driver received more alerts that this driver had higher level of on-the-job-drowsiness. Although valid alerts were identified, some of the alerts the drivers received may have been due to increased situational awareness (e.g., monitoring the mirrors) or other system limitations (e.g., not capturing the driver's eyes properly). The next analysis will focus only on the valid alerts that the drivers received during the course of the study.

Due to the small subset of participants who were part of the two DFM rating extremes (favoring, disfavoring), this analysis was also limited to descriptive statistics. Drivers who gave a favoring rating to the DFM had more valid alerts than their counterparts, even for the baseline (Table 74). The total number of valid alerts received from the subset of drivers was 111; the drivers who gave favoring ratings of the system had 104 (93.7%) valid alerts, and the drivers with a disfavoring rating of the system had 7 (6.3%) alerts. The number of valid alerts for drivers with a disfavoring rating of the system is insufficient to support any further analysis. However, the results suggest that even during baseline (no feedback to the driver provided) the drivers with a favoring opinion of the system were able to have the DFM work more reliably for them.

Table 74. Proportion of Valid Alerts by Experimental Condition and DFM Rating. Based on a Subset of Drivers (n = 14)

Experimental Condition	DFM Rating	Frequency	Percent
Baseline Test	Favoring	37	33.33
Baseline Test	Disfavoring	1	0.90
Test	Favoring	67	60.36
Test	Disfavoring	6	5.41
Total	---	111	100

Alert Response - Stopping Behavior. Figure 97 shows the elapsed time from when the valid DFM alert was presented to when the vehicle was stopped for a period of 10 min or more. The data is split by DFM rating (favoring and disfavoring) for each of the two experimental conditions of interest (baseline test and test conditions). Overall, drivers with a favoring opinion of the DFM took a longer time to stop than drivers with a disfavoring rating of the system. Drivers with a disfavoring rating of the DFM drove an average of 1 hour and 15 min before stopping the vehicle in the baseline test condition, while drivers in the test condition drove slightly less (1 hour and 13 min) before stopping. Drivers with a favoring rating of the DFM drove an average of 1 hour and 33 min before stopping the vehicle in the baseline test condition, while drivers in the test condition drove longer (1 hour and 47 min) before stopping.

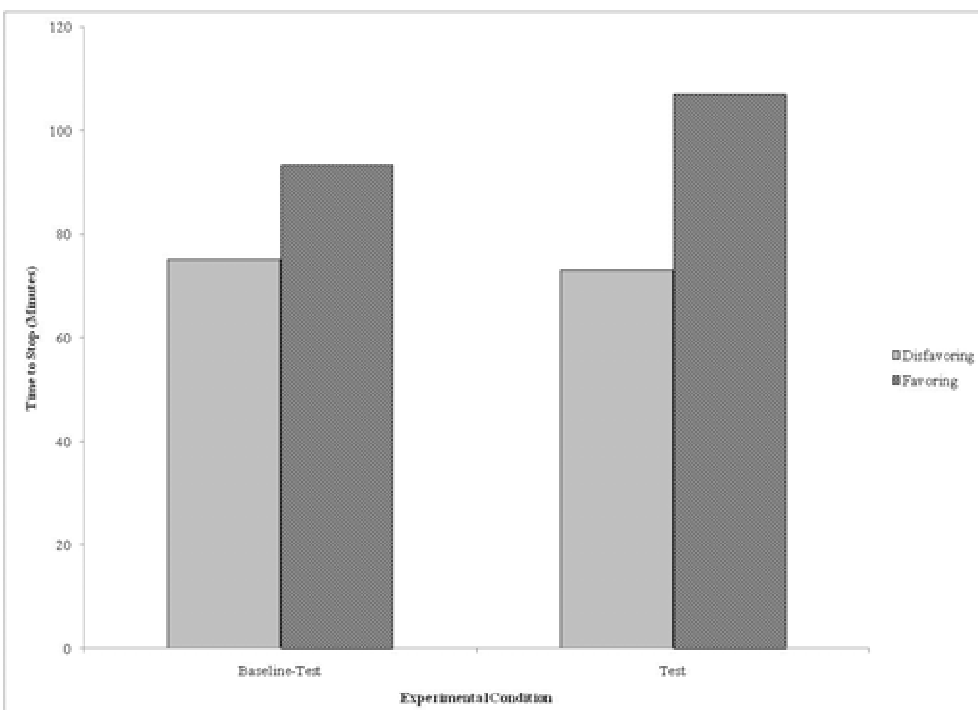


Figure 97. Elapsed Time From Valid Alert to Driver Stopping the Vehicle by DFM Rating. Based on a Subset of Drivers (n = 14)

Post-Alert Behaviors. Figure 98 presents the behaviors observed after a valid DFM alert was presented for the two extreme DFM rating subsets. Please remember that drivers with a disfavoring rating of the DFM represent one valid alert in the baseline test condition and six valid alerts in the test condition. These were drivers observed only to adjust their body, reach, or stretch after receiving a valid DFM alert. Drivers with a favoring rating of the DFM (37 valid alerts during baseline test and 67 valid alerts during test condition) exhibited a similar set of post-alert behavior where adjusting their body was the most common behavior followed by a few additional behaviors. However, the additional behaviors were likely due to the increased sample size and not necessarily to the drivers' favoring disposition toward the DFM.

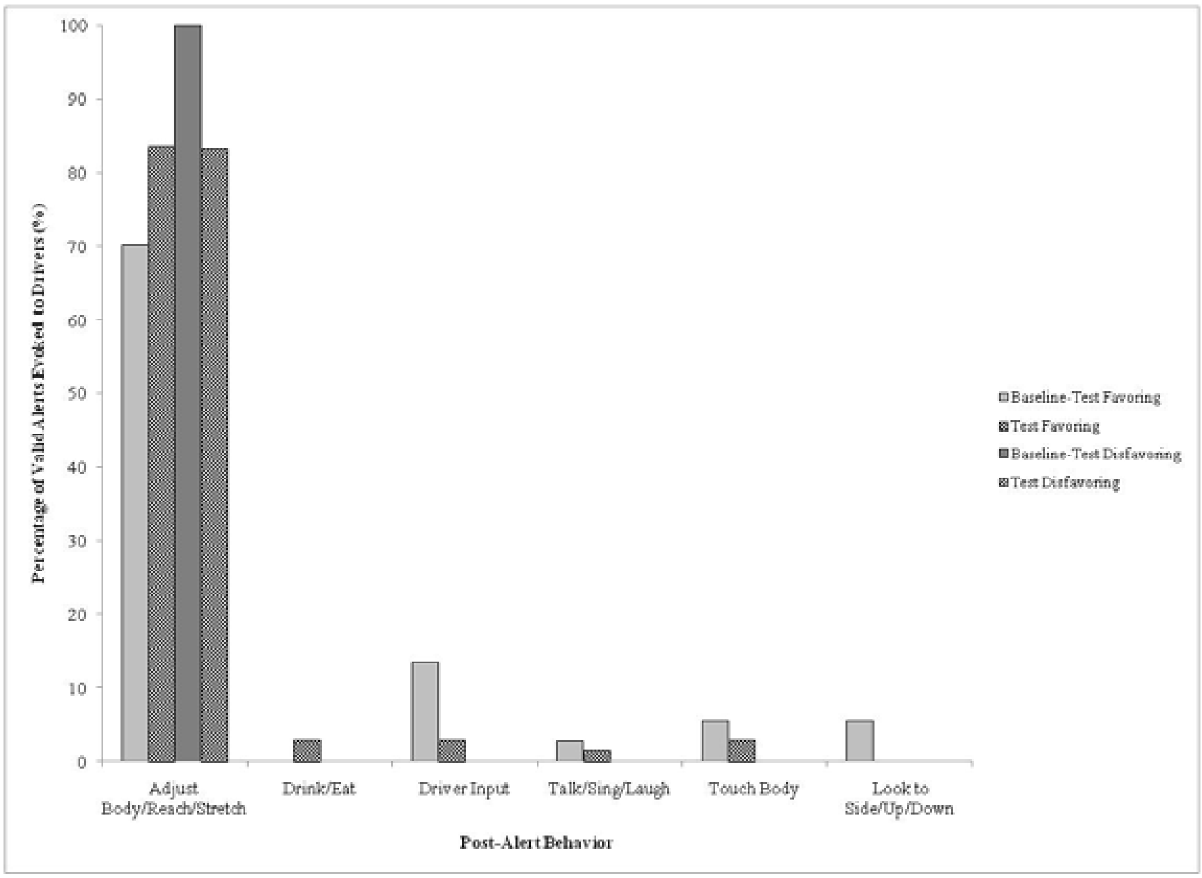


Figure 98. Percentage of Valid Alerts by Type of Behavior: DFM Rating.

Based on a Subset of Drivers (n = 14)

Figure 99 shows the average amount of time that elapsed from a valid DFM alert to the observed behavior for drivers with a favoring or disfavoring opinion of the DFM. The sparse amount of data makes it difficult to draw any general conclusions.

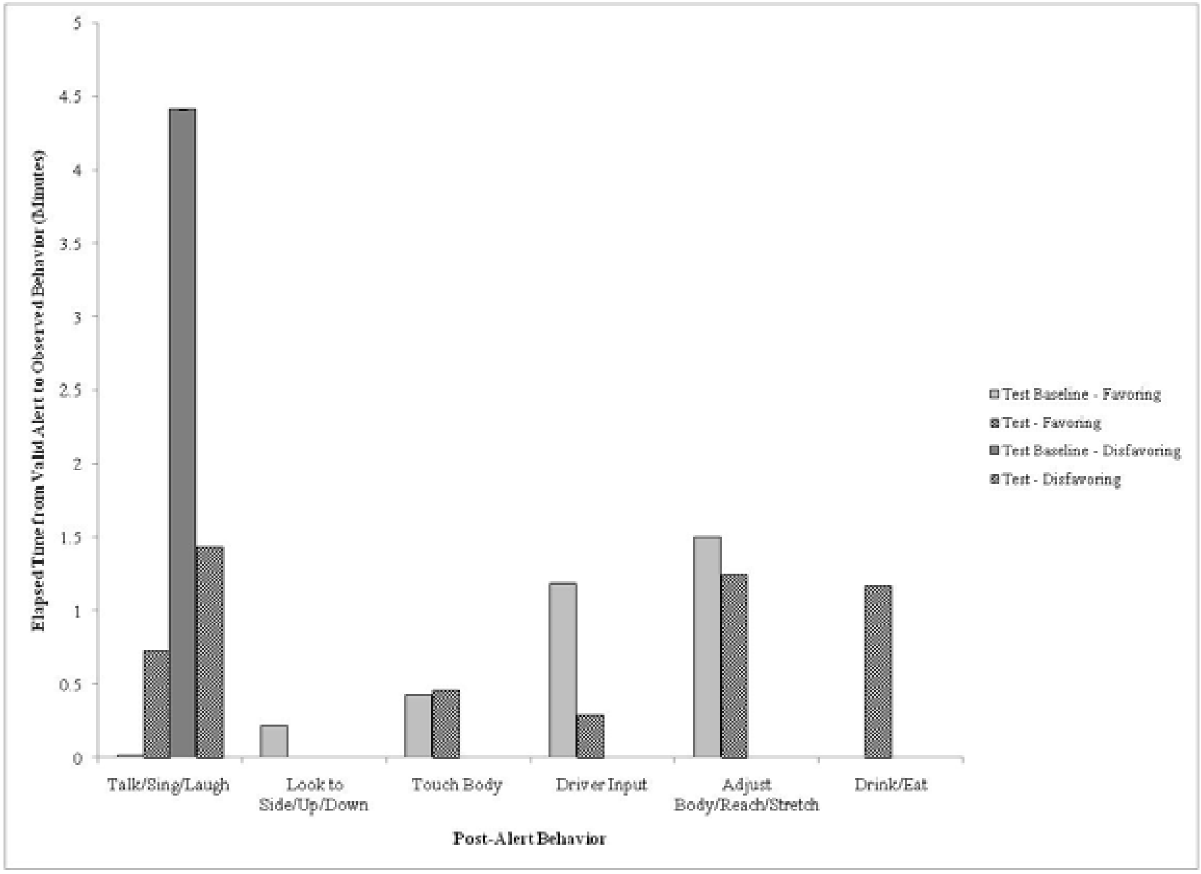


Figure 99. Elapsed Time from Valid Alert to Post-Alert Behavior: DFM Rating.

Based on a Subset of Drivers (n = 14)

Sleep Hygiene

Table 75 shows the mean quantity of sleep obtained by DFM rating (favoring or disfavoring). It appears that drivers with favoring opinions of the DFM obtained a greater amount of sleep than those drivers with a disfavoring rating of the DFM (Table 75). A comparison of the 95 percent confidence limits indicated that the difference between the average sleep for the favoring and disfavoring drivers was significant. However, there was no increase in the amount of sleep obtained from the baseline test to the test condition when the rating was favoring.

Table 75. Mean Sleep Quantity by DFM Rating

Experimental Condition	DFM Rating	Number of Days	Mean Sleep in 24 h[1]	SE (h)	LCL*	UCL*
Baseline Test	Disfavoring	27	3.52	0.31	2.91	4.13
Baseline Test	Favoring	305	5.67	0.14	5.40	5.94
Test	Disfavoring	241	4.18	0.20	3.79	4.57
Test	Favoring	531	5.65	0.11	5.43	5.87

*Observations are taken from the same driver and are therefore not independent. Based on a subset of drivers, n = 14.

Table 76 presents the mean sleep efficiency by DFM rating. Drivers with a favoring opinion of the DFM obtained more efficient sleep than drivers with a disfavoring rating of the DFM. However, there was no increase in the sleep efficiency for the baseline test to the test condition when the rating was favoring. Table 77 presents the mean number of awakenings by DFM rating. There was no statistical difference between the drivers with a favoring rating of the DFM and drivers giving disfavoring ratings in the test condition.

Table 76. Sleep Efficiency by DFM Rating

Experimental Condition	DFM Rating	Number of Days	Efficiency (Percent)	SE (Percent)	LCL*	UCL*
Baseline Test	Disfavoring	69	74	2	70	78
Baseline Test	Favoring	273	86	1	84	88
Test	Disfavoring	210	71	1	69	73
Test	Favoring	450	85	1	83	87

*Observations are taken from the same driver and are therefore not independent. Based on a subset of drivers, n = 14.

Table 77. Awakenings by DFM Rating

Experimental Condition	DFM Rating	Number of Days	Mean Awakenings	SE	LCL*	UCL*
Baseline Test	Disfavoring	69	11.35	1.18	9.04	13.66
Baseline Test	Favoring	273	9.23	0.44	8.37	10.09
Test	Disfavoring	210	12.86	0.73	11.43	14.29
Test	Favoring	450	10.66	0.41	9.86	11.46

*Observations are taken from the same driver and are therefore not independent. Based on a subset of drivers, n = 14.

[1] Note that the sleep quantity values cannot be used to evaluate overall sleep of CMV drivers, but only provide a means of making relative comparisons across groups. The method of eliminating nap data precludes making more general or absolute statements regarding overall sleep.

Table 78 presents the mean number of scored sleep periods (SSPs) of duration greater than 20 min by DFM rating. Based on this subset of data, drivers in the test condition with a favoring rating of the DFM obtained more SSPs longer than 20 min than drivers with a disfavoring rating of the DFM. This is also true for the mean longest SSP duration (Table 79).

Table 78. Mean Number of SSPs Over 20 Minutes by DFM Rating

Experimental Condition	DFM Rating	Number of Days	Mean #SSPs Longer than 20 min	SE	LCL*	UCL*
Baseline Test	Disfavoring	69	3.77	0.31	3.16	4.38
Baseline Test	Favoring	273	4.41	0.16	4.10	4.72
Test	Disfavoring	210	3.99	0.19	3.62	4.36
Test	Favoring	450	4.74	0.14	4.47	5.01

*Observations are taken from the same driver and are therefore not independent. Based on a subset of drivers, n = 14.

Table 79. Mean Longest SSP in an O-O Interval by DFM Rating

Experimental Condition	DFM Rating	Number of Days	Mean Longest SSP (min)	SE (min)	LCL*	UCL*
Baseline Test	Disfavoring	69	81.91	8.02	66.19	211.64
Baseline Test	Favoring	273	151.8	5.71	140.61	427.39
Test	Disfavoring	210	104.5	7.19	90.41	281.70
Test	Favoring	450	149.6	4.57	140.64	425.26

*Observations are taken from the same driver and are therefore not independent. Based on a subset of drivers, n = 14.

Taken as a whole, these data suggest that drivers with a favoring opinion of the DFM obtained a higher quality of sleep than drivers with a disfavoring rating of the DFM. However, there was no difference between the baseline and test conditions for the subset of drivers with a favoring rating of the DFM. This suggests that drivers favoring the DFM might have already had a better sleep hygiene than the disfavoring drivers.

Involvement in Safety-Critical Events

The total number of SCEs evaluated in this study is 1,124 events (28 crashes, 112 near-crashes, and 984 crash relevant conflicts). The subset of drivers for Research Question 5 represents 212 (19%) of these SCEs (i.e., 6 crashes, 21 near-crashes, and 185 crash relevant conflicts). Table 80 shows the frequency of SCE by DFM rating and Experimental Condition. Table 81 shows the number of hours driven using a similar split of the data.

Table 80. SCEs by Experimental Condition and DFM Rating.

Based on a Subset of Drivers (n = 14)

Experimental Condition	Favoring	Disfavoring	Total
Baseline Test	50	5	55
Test	141	16	157
Total	191	21	212

Table 81. Number of Hours Driven by Experimental Condition and DFM Rating.

Based on a Subset of Drivers (n = 14)

Experimental Condition	Favoring	Disfavoring	Total
Baseline Test	2144	471	2615
Test	3953	1269	5222
Total	6097	1740	7837

Safety Critical Events by DFM Rating. Table 82 shows the frequency of SCE based on the drivers' subjective DFM rating. Results indicate that both rating groups have a similar distribution (Fisher's $p = 0.40$). These results also indicate there is not a statistically significant relationship between the experimental condition and the type of SCE. In addition to the frequency of SCE, SCE rate was computed as the number of SCEs per 100 hours driven by DFM rating (Figure 100). The SCE rate for crashes and near-crashes was similar for drivers with opposing DFM ratings, but drivers with a disfavoring rating had a lower rate of crash relevant conflicts.

Table 82. Frequency of SCE by Type of SCE and DFM Rating.

Based on a Subset of Drivers (n = 14)

DFM Rating	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Favoring	5	18	168	191
Disfavoring	1	3	17	21
Total	6	21	185	212

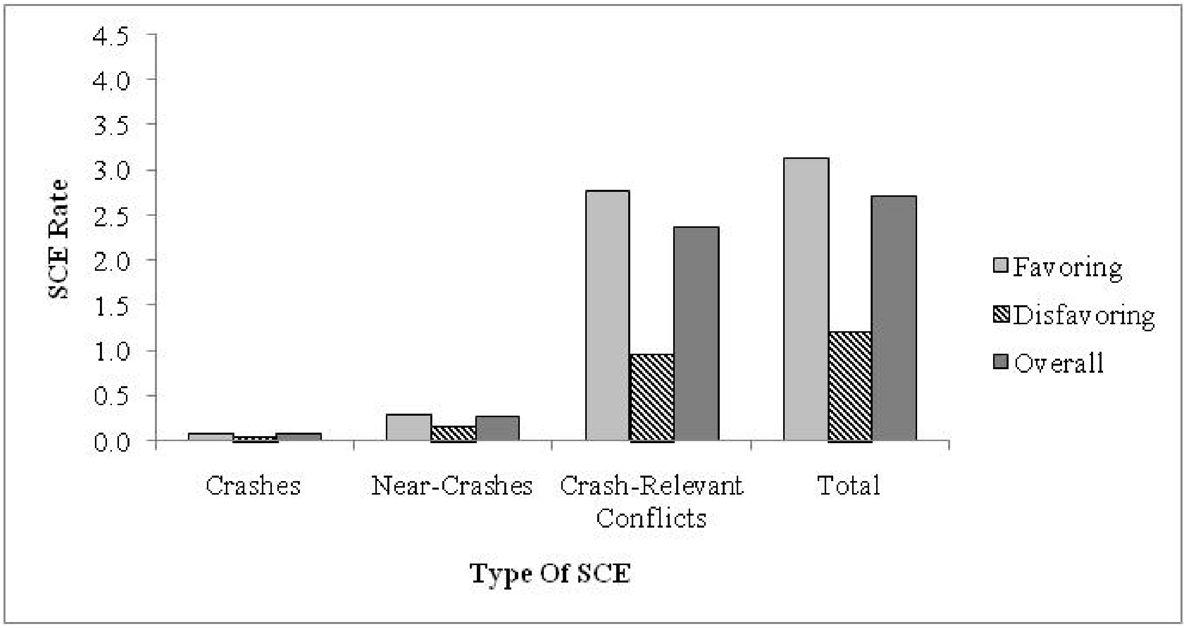


Figure 100. SCE Rates by Event Type and DFM Rating.

Based on a Subset of Drivers (n = 14)

Safety Critical Events by Experimental Condition. Figure 101 presents the SCEs by the experimental conditions of interest for this research question (baseline test and test conditions). Drivers with a favoring opinion of the DFM had 191 SCEs, with 50 of them occurring in the baseline test condition. Drivers with a disfavoring rating of the DFM had 21 SCEs, and 5 of those were in the baseline test condition.

Examination by the Fisher test indicates the two conditions and groups have the same distribution of SCEs ($p = 0.31$). These results also indicate that there is no statistically significant relationship between the experimental condition and the type of SCE.

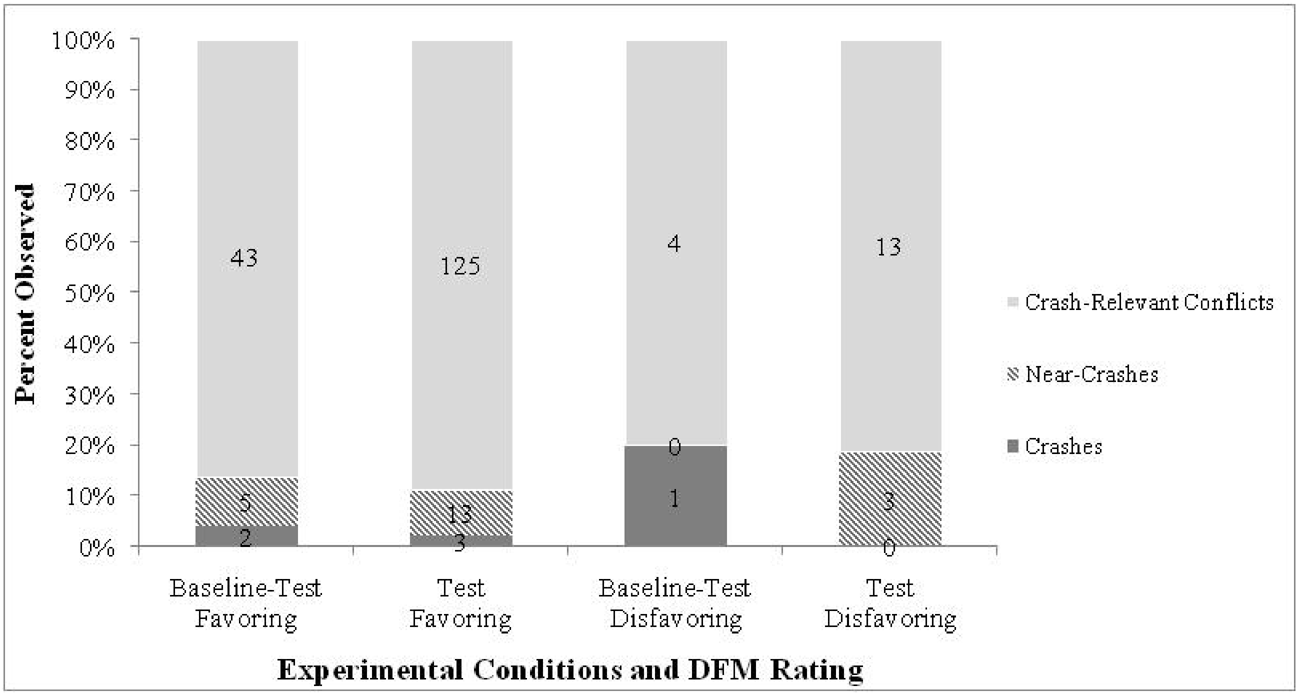


Figure 101. Frequencies and Proportion of SCEs by Experimental Condition and DFM Rating. Based on a Subset of Drivers (n = 14)

Manual PERCLOS for Safety Critical Event by DFM Rating. Two surrogate measures of driver drowsiness were employed to characterize level of drowsiness during SCEs. The first measure evaluated was the manual PERCLOS. For 212 SCEs, 58 SCEs were lacking sufficient video data to calculate manual PERCLOS (Wiegand, Hanowski, Olson, & Melvin, 2008). The mean manual PERCLOS value was 6 with a maximum observed value of 24. Table 83 shows the descriptive statistics for the manual PERCLOS values at the moment of the SCE by DFM rating and experimental condition. The average manual PERCLOS value increased from the baseline test to the test conditions for the drivers with a disfavoring rating of the DFM. However, this baseline had only two data points available, making it difficult to provide an accurate trend. The mean manual PERCLOS for the drivers with a favoring opinion of the DFM was lower for the test condition when compared to the baseline test, suggesting a positive impact of the DFM for this group.

Table 83. Manual PERCLOS Values by Experimental Condition and DFM Rating. Based on a Subset of Drivers (n = 14)

Experimental Condition	DFM Rating	Total SCE	SCE With Manual PERCLOS Available	Minimum	Maximum	Mean
Baseline Test	Disfavoring	5	2	1	4	3
Baseline Test	Favoring	50	37	1	24	8
Test	Disfavoring	16	11	0	22	10
Test	Favoring	141	104	0	22	5

The frequency of PERCLOS values by experimental condition is presented in Table 84. The table shows that 62 percent of SCEs have PERCLOS values below 12. Given that approximately 27 percent of the SCE did not have manual PERCLOS available, no further statistical analysis was performed.

Table 84. Manual PERCLOS Over 12 by Experimental Condition and DFM Rating. Based on a Subset of Drivers (n = 14)

Experimental Condition	DFM Rating	PERCLOS <12	PERCLOS ≥12	PERCLOS N\A	Total
Baseline Test	Disfavoring	2	0	3	5
Baseline Test	Favoring	26	11	13	50
Test	Disfavoring	6	5	5	16
Test	Favoring	98	6	37	141
Total	---	132	22	58	212

Drowsiness-Related Driving Behavior Evaluation for Safety Critical Event by DFM Rating. As mentioned earlier, all SCEs were evaluated and driver behavior was categorized if the driver appeared drowsy, sleepy, asleep, fatigued, or showed signs of reduced alertness that occurred during the period of time leading to the SCE. Table 85 shows the frequency of drowsiness-related SCEs by experimental condition and DFM Rating. Of the 212 SCEs, 28 (13.2%) are drowsiness related. For the drivers with a disfavoring rating of the DFM, drowsiness-related SCEs range from 40 to 50 percent, compared with less than 10 percent for the drivers with a favoring opinion of the DFM.

Table 85. Frequency of Drowsiness-Related SCE by Experimental Condition and DFM Rating. Based on a Subset of Drivers (n = 14)

Experimental Condition	DFM Rating	Not Drowsy	Drowsy	Total
Baseline Test	Disfavoring	3	2	5
Baseline Test	Favoring	46	4	50
Test	Disfavoring	8	8	16
Test	Favoring	127	14	141
Total	---	184	28	212

Table 86 shows the distribution of events by group, test condition, and whether or not they were drowsiness related. Table 87 shows the SCE rate by test condition and type of SCE. The crash rate of drowsiness-related events in the test condition for the group with favoring ratings of the DFM is 0.35, compared with 0.19 for the ones with a disfavoring rating of the DFM.

Table 86. Drowsiness-Related SCE by Experimental Condition, Event Type, and DFM Rating. Based on a Subset of Drivers (n = 14)

Experimental Condition	DFM Rating	Drowsiness Related	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Baseline Test	Favoring	No	2	4	40	46
Baseline Test	Favoring	Yes	0	1	3	4
Baseline Test	Favoring	Total	2	5	43	50
Test	Favoring	No	3	11	113	127
Test	Favoring	Yes	0	2	12	14
Test	Favoring	Total	3	13	125	141
Baseline Test	Disfavoring	No	1	0	2	3
Baseline Test	Disfavoring	Yes	0	0	2	2
Baseline Test	Disfavoring	Total	1	0	4	5
Test	Disfavoring	No	0	0	8	8
Test	Disfavoring	Yes	0	3	5	8
Test	Disfavoring	Total	0	3	13	16

Table 87. SCE Rate by 100 Hours Driven for Drowsiness-Related SCE by Experimental Condition, Event Type, and DFM Rating. Based on a Subset of Drivers (n = 14)

Experimental Condition	DFM Rating	Drowsiness Related	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Baseline Test	Favoring	No	0.09	0.19	1.87	2.15
Baseline Test	Favoring	Yes	0.00	0.05	0.14	0.19
Baseline Test	Favoring	Total	0.09	0.23	2.01	2.33
Test	Favoring	No	0.08	0.28	2.86	3.21
Test	Favoring	Yes	0.00	0.05	0.30	0.35
Test	Favoring	Total	0.08	0.33	3.16	3.57
Baseline Test	Disfavoring	No	0.21	0.00	0.42	0.64
Baseline Test	Disfavoring	Yes	0.00	0.00	0.42	0.42
Baseline Test	Disfavoring	Total	0.21	0.00	0.85	1.06
Test	Disfavoring	No	0.00	0.00	0.63	0.63
Test	Disfavoring	Yes	0.00	0.24	0.39	0.63
Test	Disfavoring	Total	0.00	0.24	1.02	1.26

Driver at Fault. The proportion of drivers at fault are displayed in Figure 102. The driver of the experimental vehicle (V1) was judged to be at fault between 66 and 87 percent of the time. Results show no statistically significant differences in the distribution of SCEs between test conditions or within the test conditions for each Experimental Group (Fischer $p = 0.072$).

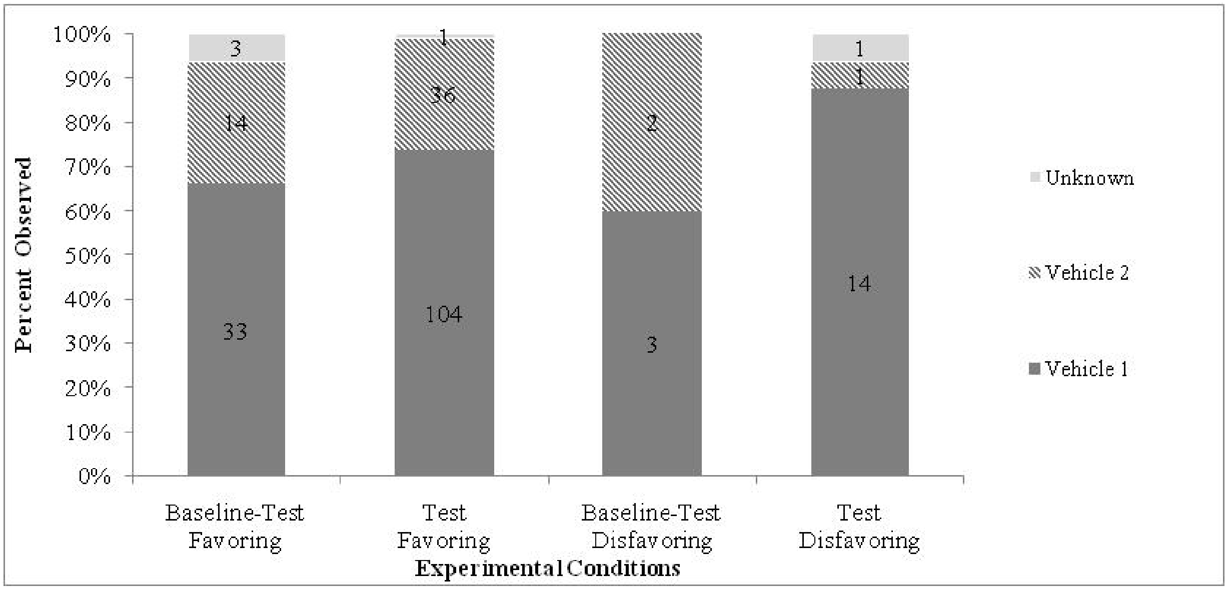


Figure 102. Vehicle at Fault by Experimental Condition and DFM Rating.

Based on a Subset of Drivers (n = 14)

Safety Critical Events Within the Operating Envelope that Received Alerts. Similar to the analysis for Research Question 3 (examining SCE involvement), the alerts presented for the SCEs in this subset of data were identified. Fifty-eight alerts were recorded in the previous hour for five of the SCEs. However, none of the alerts were valid. Four of the SCEs that received alerts were for Driver 2, who gave the DFM disfavoring ratings in the baseline. This driver received a total of 58 valid alerts and 57 full alerts. Only one SCE was associated with one initial alert and one full alert in the hour prior to the event for the drivers who gave favoring ratings to the system (Driver 125).

Human-Machine Interaction

The subset of drivers who gave the DFM disfavoring ratings contained only 4 drivers, while the subset of participants with favoring ratings of the DFM contained 10 drivers. In order to account for the unbalanced nature of the sample, the rate of change per driver was examined instead of the frequency of DFM changes.

Between drivers with favoring ratings of the DFM those with disfavoring ratings of the DFM, there was little difference in the rate per driver of DFM sensitivity changes by time of day (Figure 103). However, from 20:00 to 23:59 the disfavoring drivers had more DFM sensitivity changes per driver than did the favoring drivers.

Figure 104 and Figure 105 show the changes per driver in both the DFM warning sound and the DFM brightness level. Both of these graphs show similar patterns; specifically, the drivers with a disfavoring opinion of the DFM had, overall, more changes in both brightness and warning sounds per driver than the drivers favoring the system.

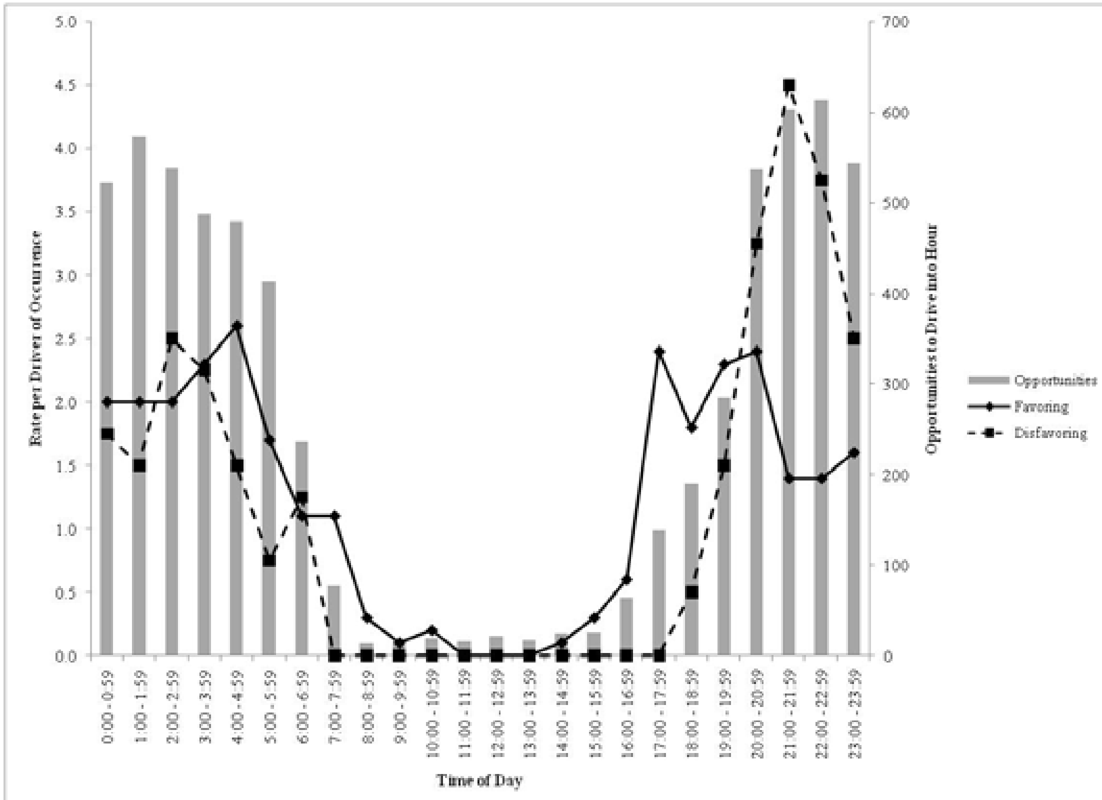


Figure 103. Rate of DFM Sensitivity Changes per Driver by Time of Day and DFM Rating. Based on a Subset of Drivers (n = 14)

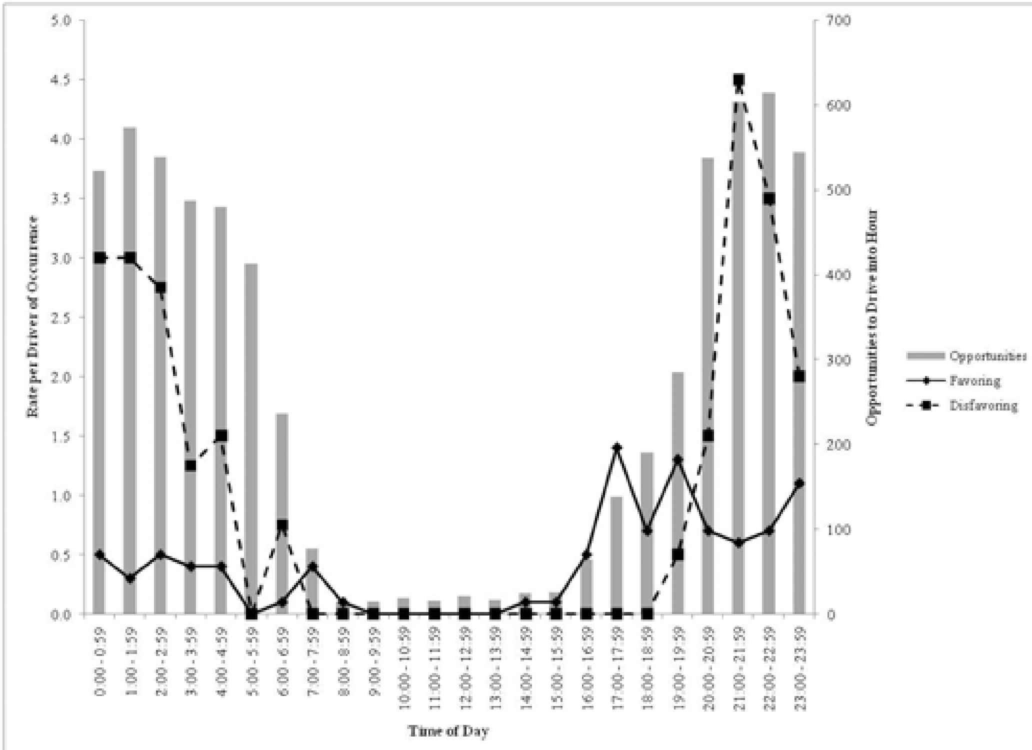


Figure 104. Rate of Warning Sound Changes per Driver by Time of Day and DFM Rating. Based on a Subset of Drivers (n = 14)

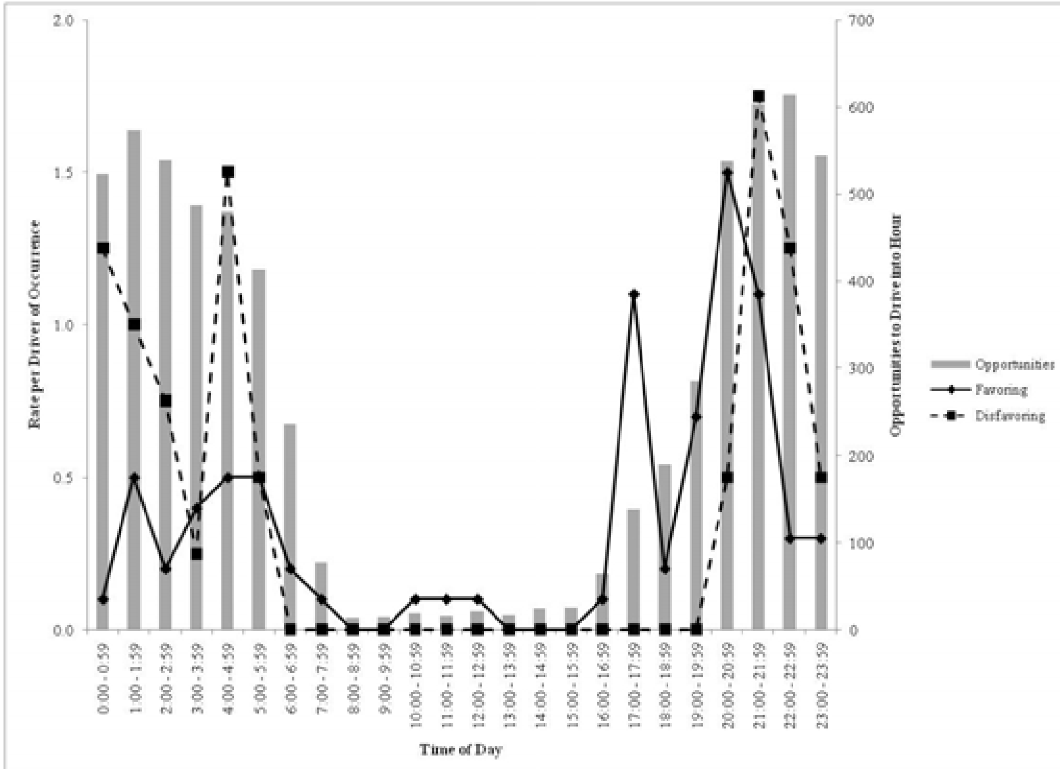


Figure 105. Rate of Brightness Changes per Driver by Time of Day and DFM Rating. Based on a Subset of Drivers (n = 14)

Figure 106 presents the rate per driver of the changes in sensitivity level as the driver progressed through their shift. Overall the number of opportunities decreased as the driver hours progressed. However, there does not appear to be a pattern that would differentiate between the drivers with favoring and disfavoring opinions of the system.

The drivers with a favoring rating of the system had a steadier rate of DFM warning sound changes per driver as the driving hours progressed (Figure 107). The drivers with a disfavoring rating of the system started with a high rate of DFM warning sound changes per driver; then, as the shift progressed the rate of changes dropped dramatically.

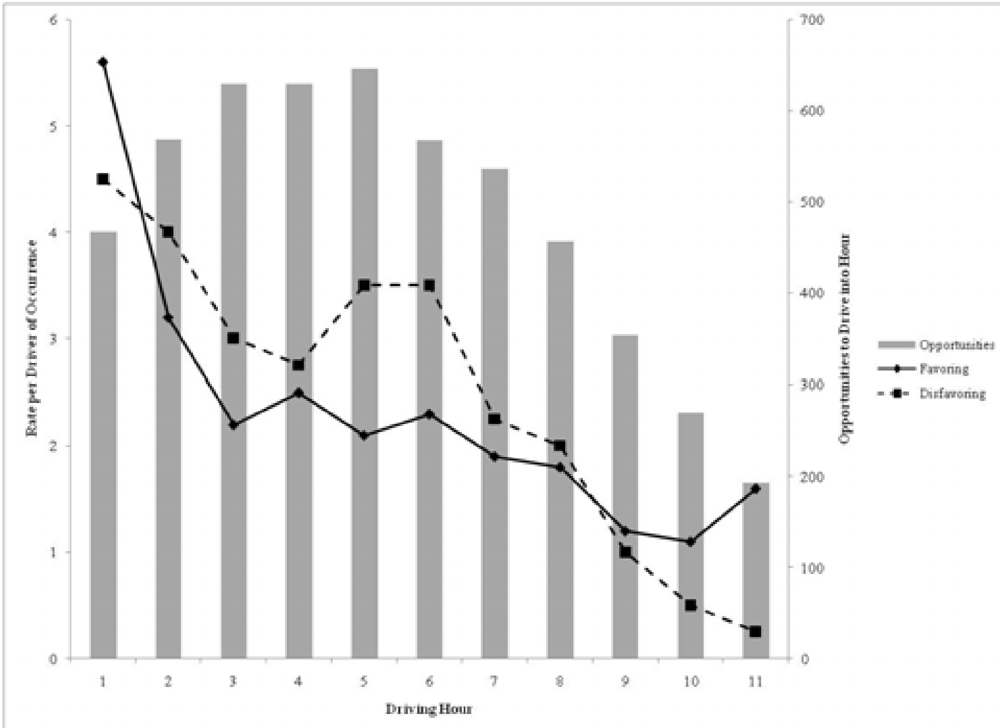


Figure 106. Rate of DFM Sensitivity Changes per Driver by Driving Hour and DFM Rating. Based on a Subset of Drivers (n = 14)

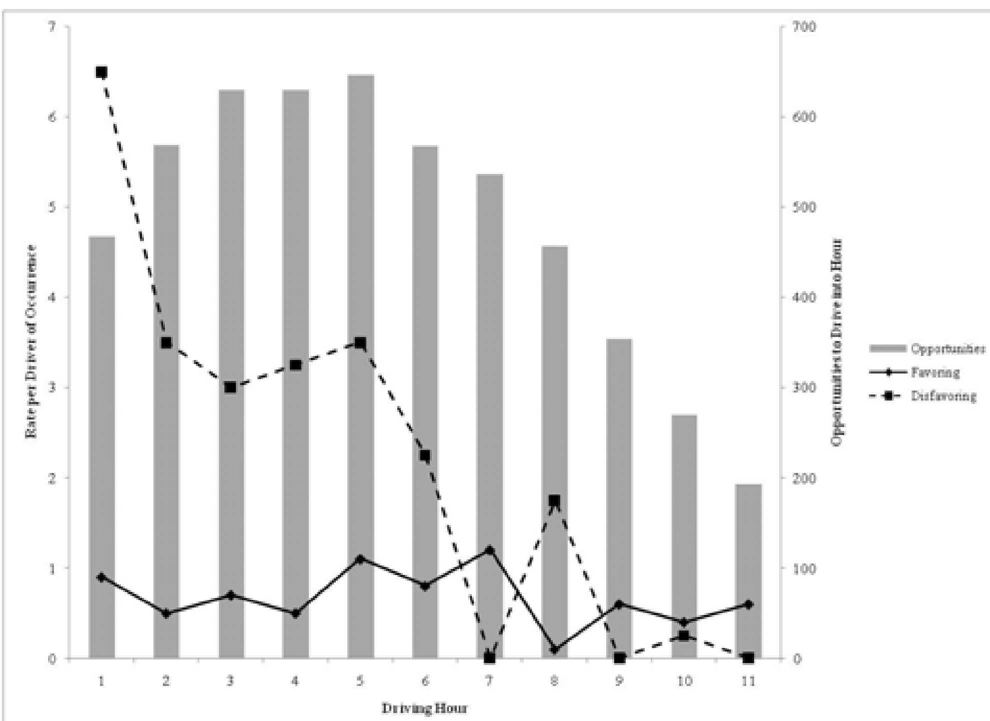


Figure 107. Rate of DFM Warning Sound Changes per Driver by Driving Hour and DFM Rating. Based on a Subset of Drivers (n = 14)

As with the DFM warning sound, the drivers with a disfavoring rating of the system started a shift with a higher rate of DFM brightness changes. Then, as the shift progressed the rate of changes dropped (Figure 108). However, the drivers with a favoring rating of the system did not have any discernable pattern in their rate of DFM brightness changes (Figure 108).

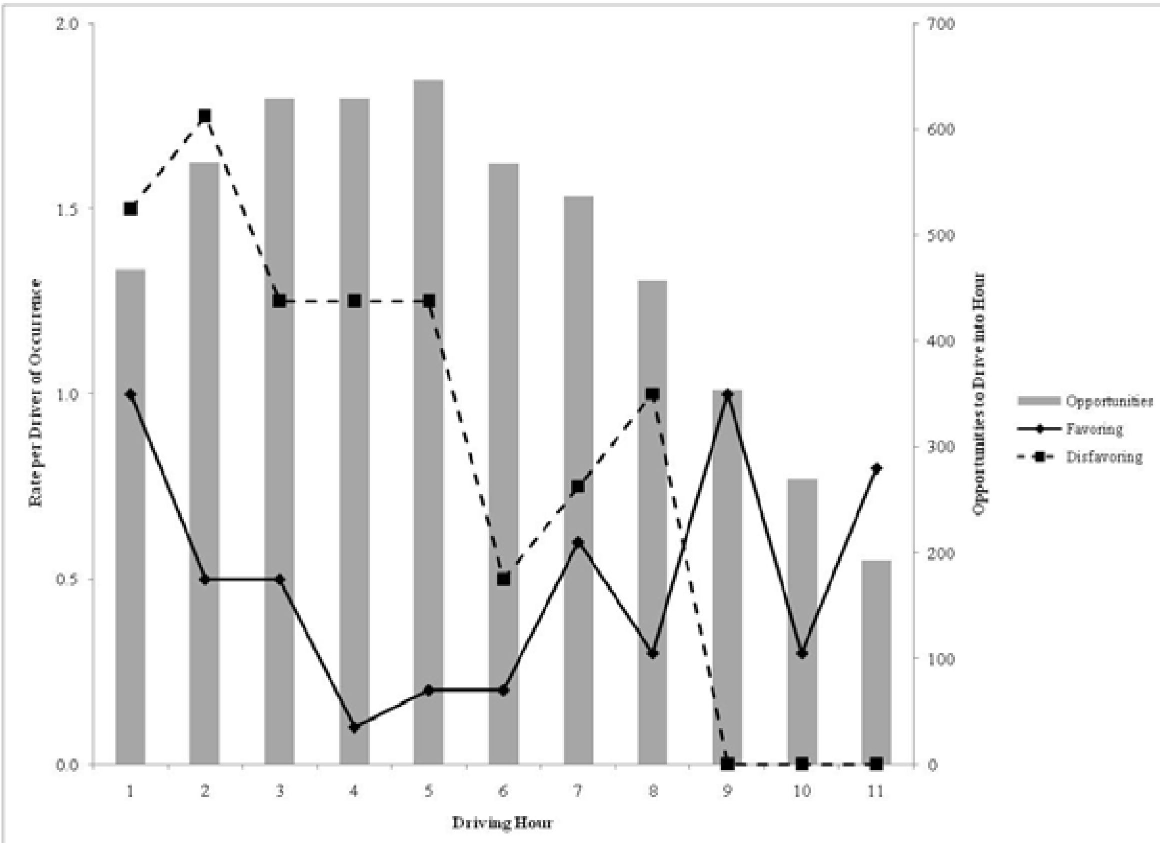


Figure 108. Rate of DFM Brightness Changes per Driver by Driving Hour and DFM Rating. Based on a Subset of Drivers (n = 14)

Figure 109 shows no discernable trend in the rate of DFM sensitivity level changes per driver as the study progressed. One striking feature of the graph is that the drivers with a disfavoring rating of the system had a similar rate of changes in DFM sensitivity level as the favoring drivers for the first week. The interaction with the system was very different between both groups from that point on.

The drivers with a favoring rating of the system had a steadier rate per driver of DFM warning sound changes than the disfavoring drivers (Figure 110). Once again the drivers with disfavoring ratings of the system had a higher rate of changes per driver during the first week. This was also followed by spikes throughout the study.

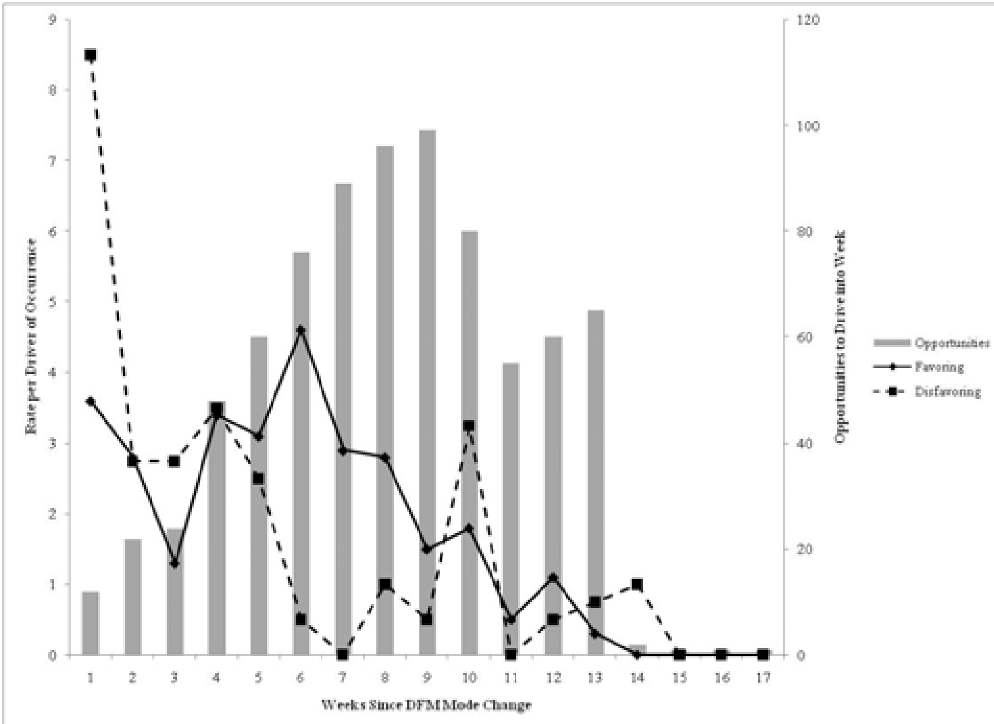


Figure 109. Rate of DFM Sensitivity Level Changes by Week Since DFM Mode Change and DFM Rating. Based on a Subset of Drivers (n = 14)

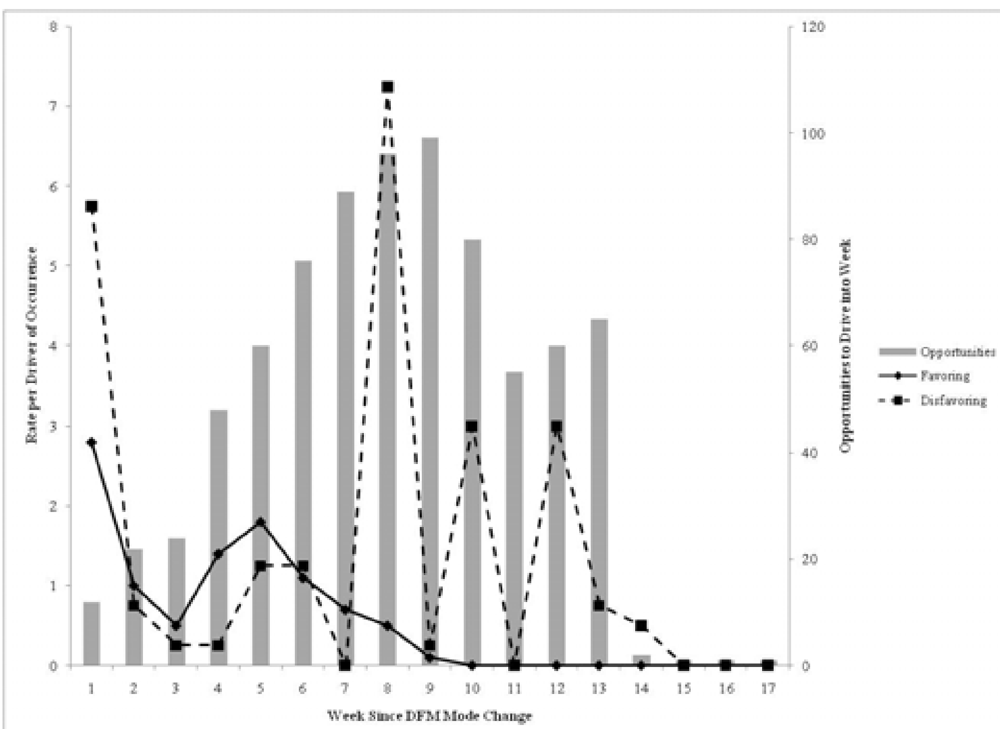


Figure 110. Rate of DFM Warning Sound Changes by Weeks Since DFM Mode Change and DFM Rating. Based on a Subset of Drivers (n = 14)

There appears to be no pattern in the rate of DFM brightness changes per driver as the drivers progressed in the study (Figure 111). As before, the drivers with a disfavoring rating of the system had a higher rate of DFM brightness changes during week one, similar to the drivers with favoring ratings of the system. The favoring drivers exhibited more changes during the earlier part of the study.

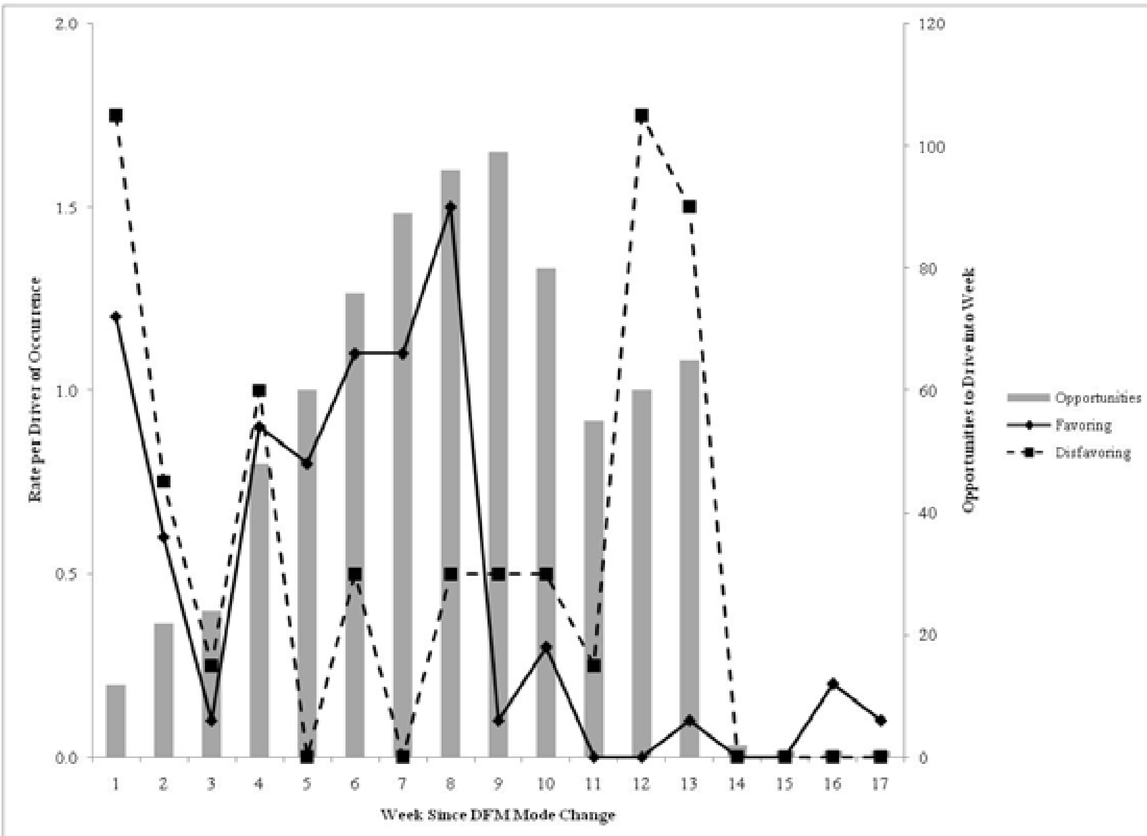


Figure 111. Rate of DFM Brightness Changes by Weeks Since DFM Mode Change and DFM Rating. Based on a Subset of Drivers (n = 14)

Overall the drivers with disfavoring ratings of the DFM had a higher mean alert duration. This is also true when it is evaluated by time of day (Figure 112). Note that Figure 112 shows one extreme observation at 8:00-8:59; this average was composed of only one observation. When comparing the mean alert duration, depending on the driving hours during a shift, once again the disfavoring drivers had a higher mean alert duration than the drivers with a favoring rating of the system (Figure 113).

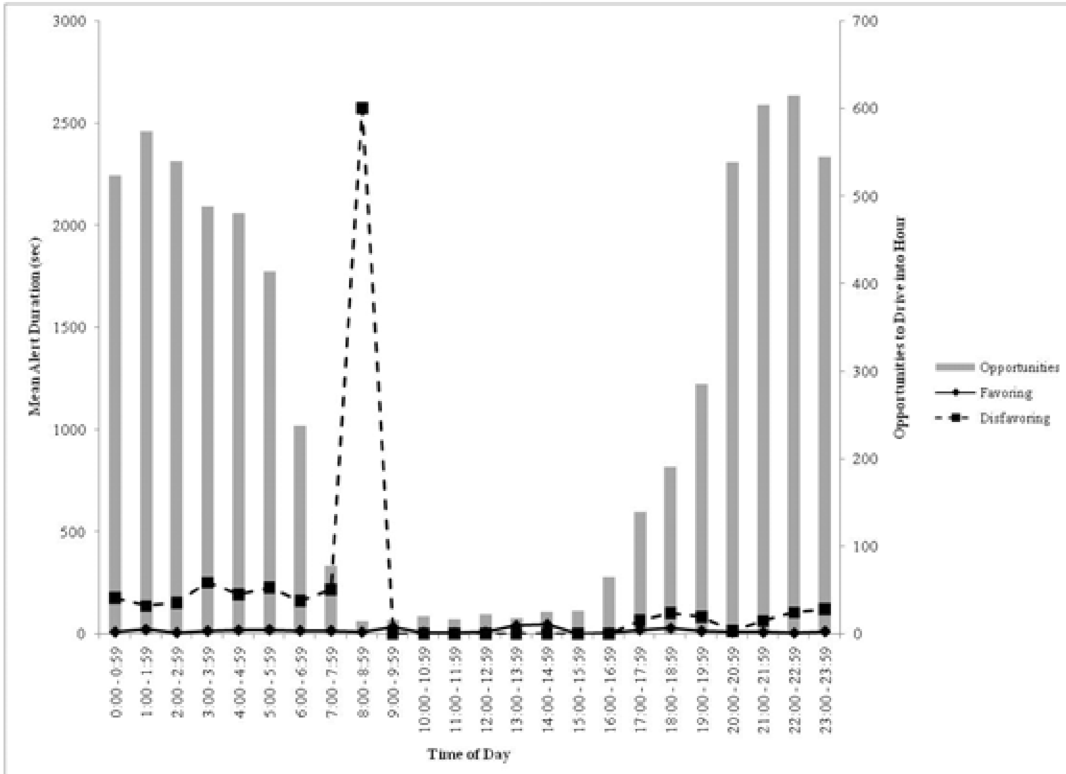


Figure 112. Mean Alert Duration by Time of Day and DFM Rating. Based on a Subset of Drivers (n = 14)

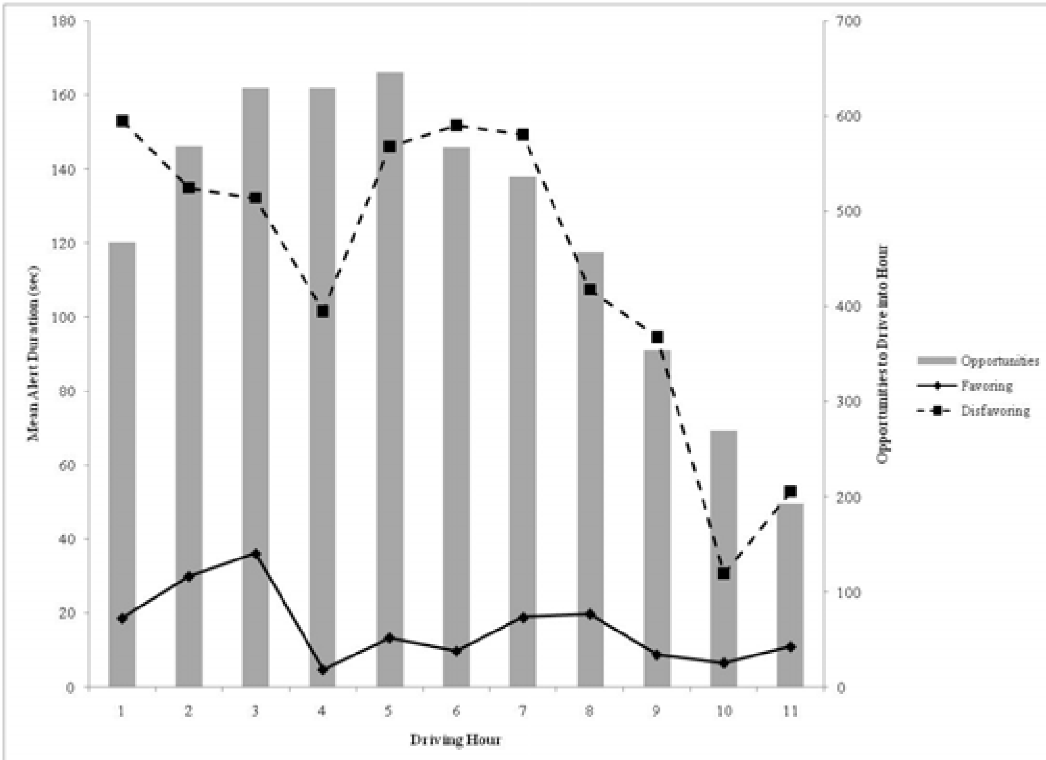


Figure 113. Mean Alert Duration by Driving Hour and DFM Rating. Based on a Subset of Drivers (n = 14)

Figure 114 does not show as strong of a pattern for either subset of drivers in terms of mean alert duration. However, drivers with a disfavoring rating of the DFM had a higher and less steady mean alert duration as the study progressed.

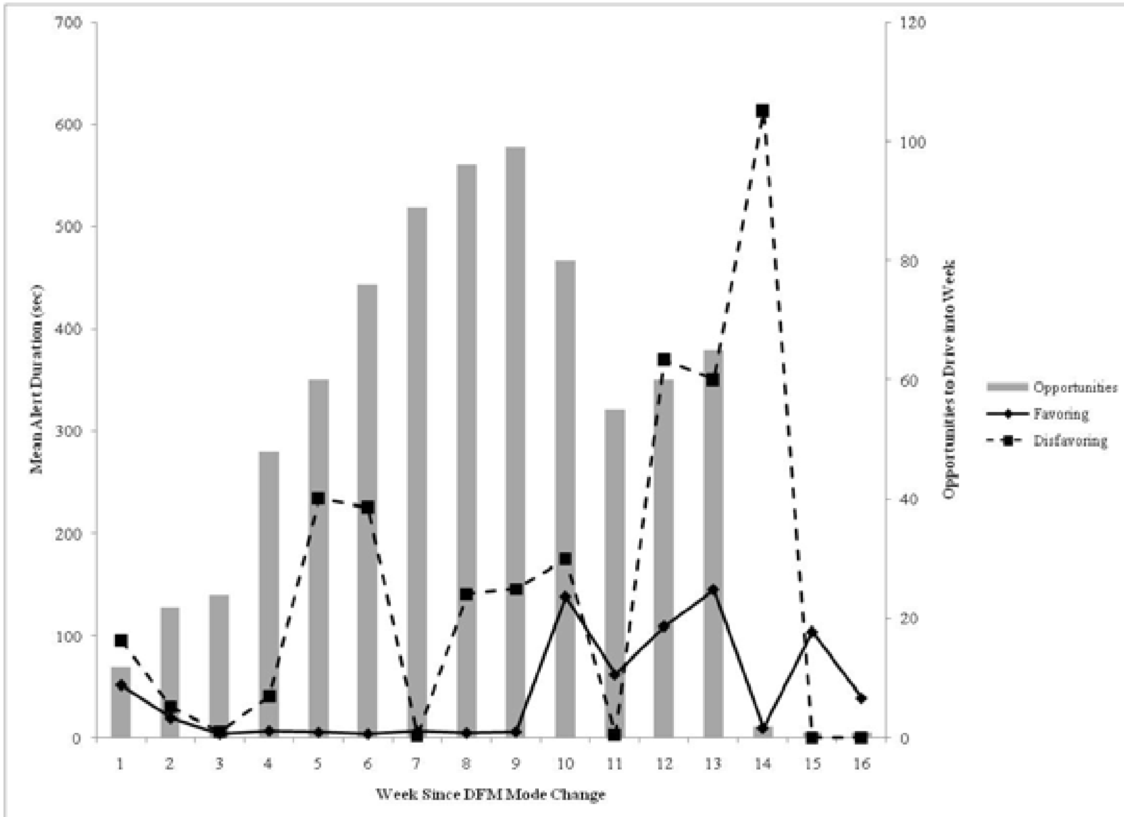


Figure 114. Mean Alert Duration by Weeks Since DFM Mode Change and DFM Rating. Based on a Subset of Drivers (n = 14)

Figure 115 shows the mean duration for each brightness setting. The most frequently used settings were the two extremes for both favoring and disfavoring drivers. Figure 116 shows the mean duration for each sensitivity level. The least frequently used setting was the high sensitivity level, and the most frequently used was the medium sensitivity level. Figure 117 shows the mean duration of each warning sound. Overwhelmingly, the most frequently used warning sound was 0. All of these findings are consistent with the findings in Research Question 4 (examining the human-machine interaction of the prototype DFM system). That is, it appears that the drivers did not take advantage of the full range of settings provided. However, this was not restricted to one single rating group.

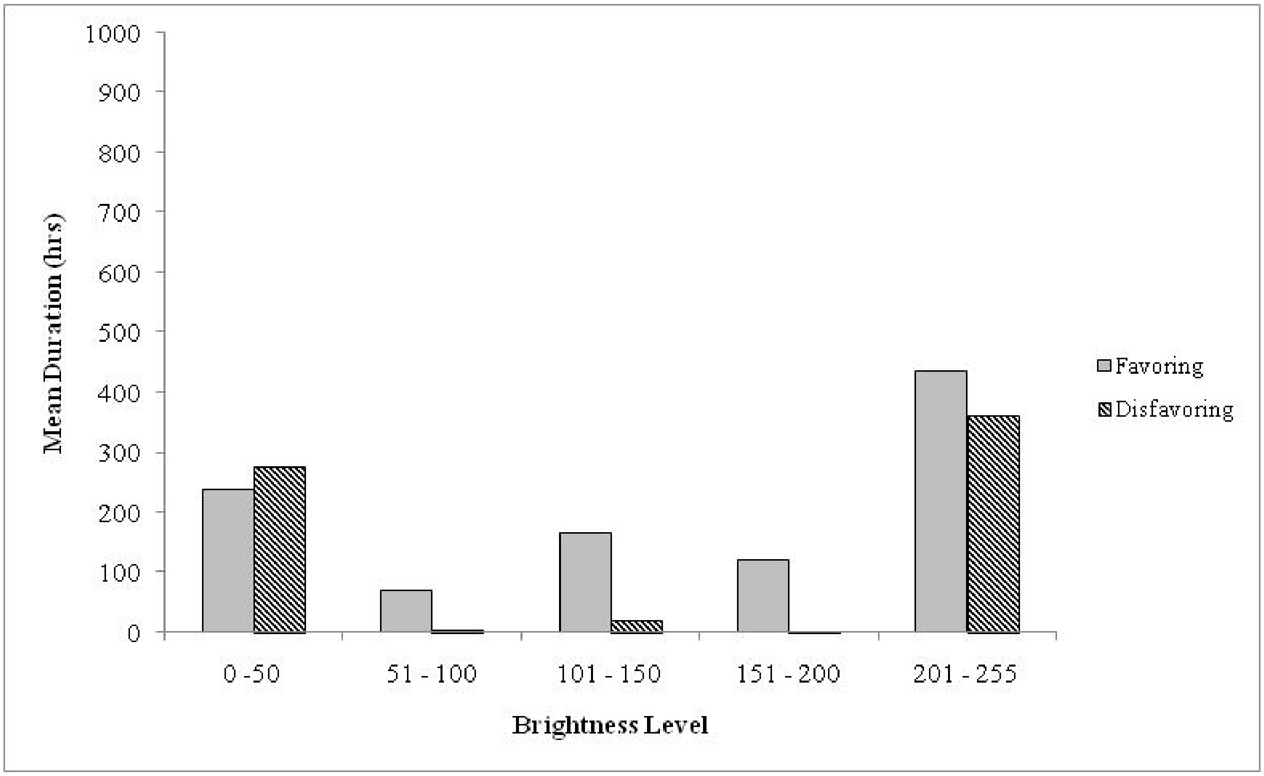


Figure 115. Mean Duration, in Hours, of Each Display Brightness Setting by DFM Rating. Based on a Subset of Drivers (n = 14)

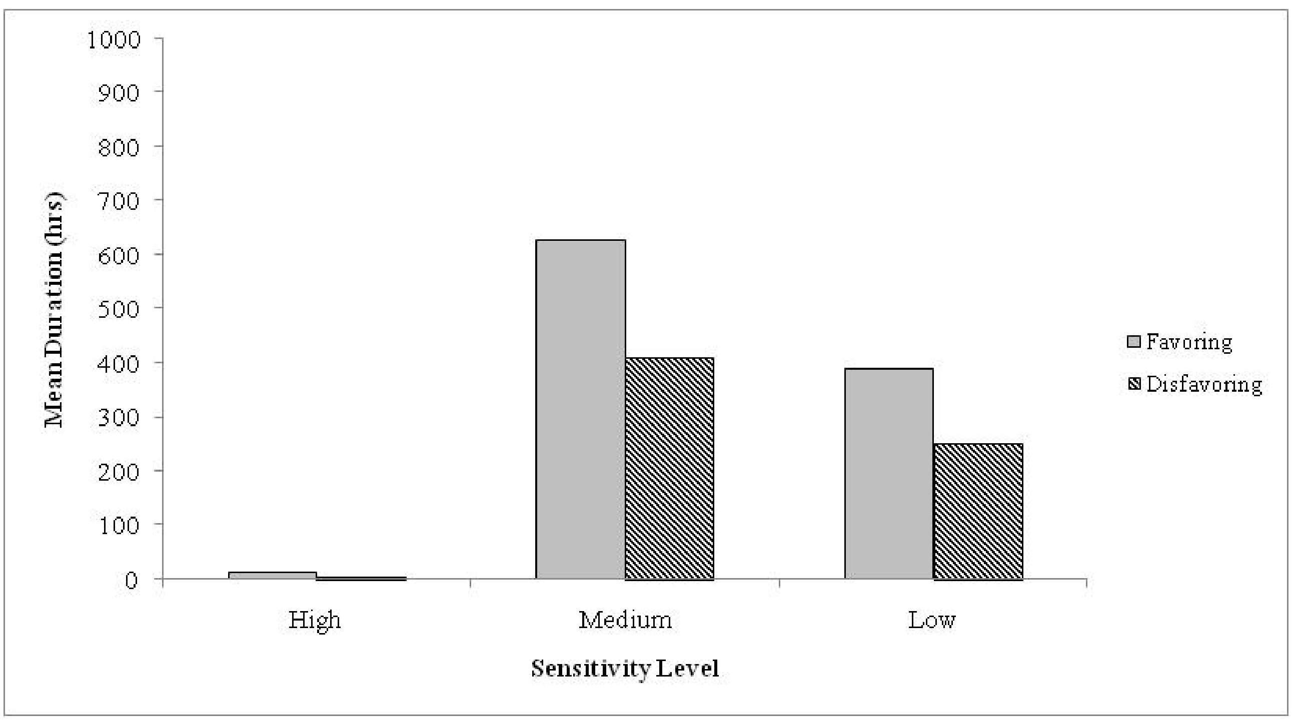


Figure 116. Mean Duration, in Hours, of Each Sensitivity Level by DFM Rating.

Based on a Subset of Drivers (n = 14)

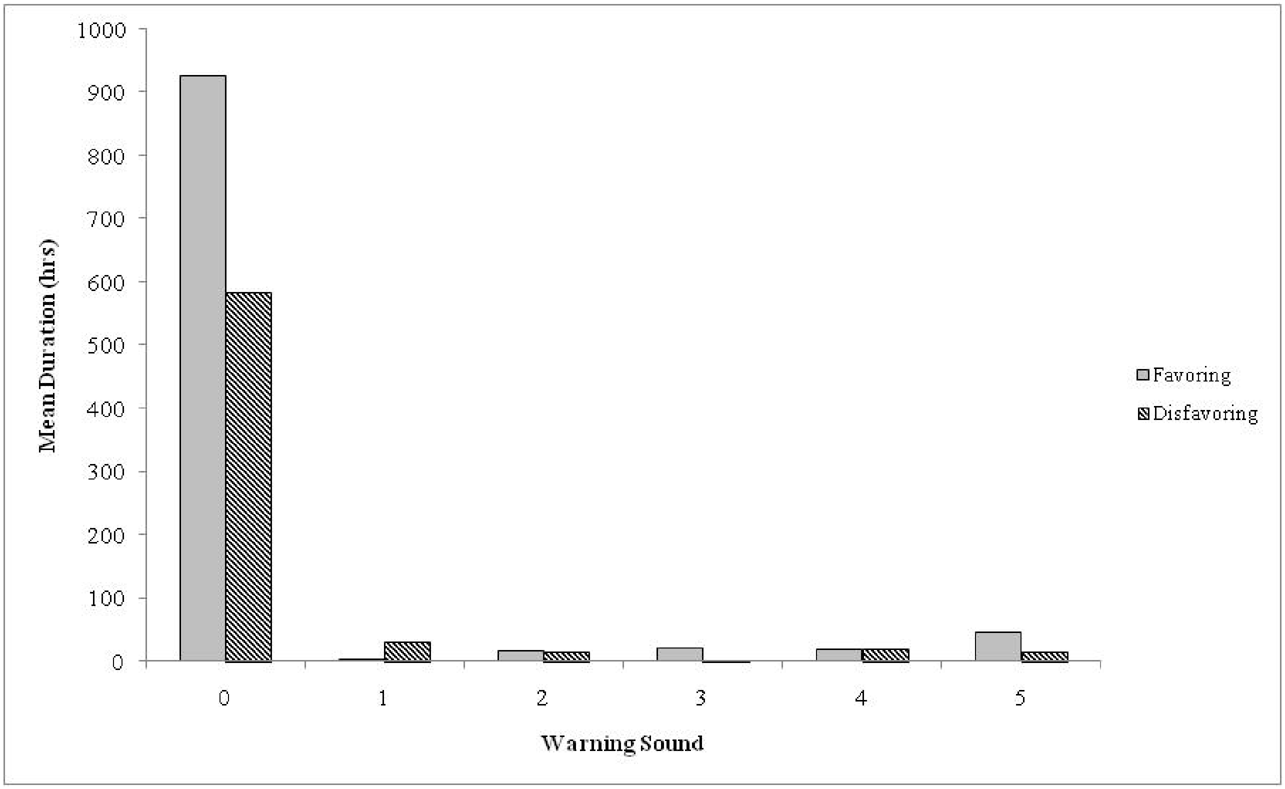


Figure 117. Mean Duration, in Hours, of Each Warning Sound by DFM Rating.

Based on a Subset of Drivers (n = 14)

Discussion

The drivers with a disfavoring rating of the system were on average older and had more experience; however, none of these differences were significant. Of the 4 drivers with a disfavoring rating of the system, 50 percent were line-haul drivers. All of the drivers with a favoring rating (n = 10) of the system were long-haul drivers. Drivers with disfavoring opinions of the system had a significantly lower rate of valid alerts as compared to the favoring drivers. Although the reason for this difference cannot be determined based on the sample size available (i.e., 10 favoring, 4 disfavoring drivers), there did appear to be an increase in the safety benefits between the drivers with favoring opinions and those with disfavoring opinions of the DFM. This can be seen in both the reduction of the DFM PERCLOS and the manual PERCLOS values, which would be indicative of a decrease in on-the-job-drowsiness. The favoring drivers, as a whole, had a higher sleep quality than the drivers with disfavoring opinions of the system. However, there was no change in the favoring drivers between the baseline test and Test. Therefore it is possible that the favoring raters had a predisposition to lower drowsiness than the disfavoring raters, or were more optimistic about the technology or participation in the research study. There appeared to be a decrease in the occurrence of drowsiness-related SCEs in the group with favoring ratings of the DFM. Due to the small sample size of drowsiness-related SCEs, this could only be examined as a general pattern not supportable by any rigorous statistical method. As expected, the drivers with a disfavoring rating of the DFM also made more adjustments to the system on average and also did not respond the DFM alerts as quickly as the drivers with favoring opinions of the system. Because the system worked less reliably for the disfavoring rating drivers, it is hypothesized that this is the reason the drivers made more adjustments to the DFM; once the drivers noticed the low reliability, they began to ignore the system. However, the drivers with a favoring rating of the DFM on average took longer to stop after receiving a valid alert. Though this difference was not significant, this finding may warrant further exploration. It is beyond the scope of this study to investigate if during that period of time there were any opportunities to stop or if that set of participants stopped at the earlier possible time while other drivers took less time to stop but had more potential places to do so.

It is worth noting once again that for this research question a small sample of drivers fell in the two extreme ratings of interest. Therefore, caution should be used in the interpretation of the results presented. Also due to the small sample size, any attempt to generalize these results should be done with caution. Given this, further exploration into the reason for the difference seen in these groups may be warranted.

RESEARCH QUESTION 6: AT-RISK DRIVER FOLLOW-UP

As part of this section the following research question (RQ) was answered:

> RQ 6.1: How did the DDWS operate for drivers who had significantly more alerts during the baseline period?

Methods, analysis, measures of interest, results, and discussion related to this question are presented next.

Methods

In order to assess potential system operation differences among drivers with more alerts in the baseline period, two groups of drivers were identified within the study drivers. The at-risk drivers were identified as the seven participants in the study with the highest number of valid alerts per hour driven in the baseline test experimental condition. The low-risk drivers were identified as the seven participants with the least number of valid alerts per hour driven in the baseline test experimental condition. This subset of participants was then analyzed using the data sets developed in the previous research questions.

Due to the small number of participants identified (14 drivers total); only descriptive statistics were used in the analysis.

Data from the subsets of at-risk and low-risk drivers were compared with respect to the main areas evaluated under Research Questions 1 through 4:

1. On-the-job drowsiness
2. Sleep hygiene
3. Involvement in SCEs
4. Human-machine interaction

In addition to the main topics above, the following areas were covered as well:

- Characteristics
  - Driver characteristics
  - Operational characteristics
- System reliability

It should be noted that due to the prototype DFM system's high false alert rate and the method used to identify at-risk and low-risk drivers, the system reliability portion of the analysis could be biased against certain drivers (i.e., the system may have simply worked in a more reliable fashion for drivers with the most alerts per driving hour, leading them to be identified as at-risk).

Analysis

Descriptive statistics, graphs, and tables will be used to identify tendencies in the data. As mentioned previously, the small number of participants with extreme attitudes precluded a more thorough analysis.

Measures

Characteristics. In order to examine the driver characteristics, the drivers' age (in years) and experience (in years) were analyzed. Additionally, operating characteristics of the system were examined in order to identify any patterns related to the haul type (long or line).

System Reliability. How well the system worked for a particular driver may have influenced the rating the driver gave the DFM, in addition to influencing the determination of their risk level. For this reason, the false alert rates for each group (at-risk and low-risk) were compared.

Previous Research Question Comparison. The results from Research Questions 1, 2, 3, and 4 were re-examined with two subsets of drivers, at-risk and low-risk, and then compared. In re-examining these questions, any differences in safety benefits between the two groups were presented.

Results

Characteristics

Driver Characteristics. Figure 118 shows the mean years of experience for drivers. On average the drivers considered at-risk had more years of experience than the drivers who were considered low-risk, even though both groups had more years of experience than the overall mean for all drivers in the study. As there were only seven drivers in each category, it is difficult to make any generalizations about the relationship between the years of experience and risk. Figure 119 shows the mean age of the participants who were considered at-risk and low-risk compared to that of all study drivers. Once again, the drivers considered to be at-risk were older on average than the drivers who were considered low-risk, and both groups had a higher mean age than the overall average age of the drivers in the study. There were no significant differences in any of the demographic data.

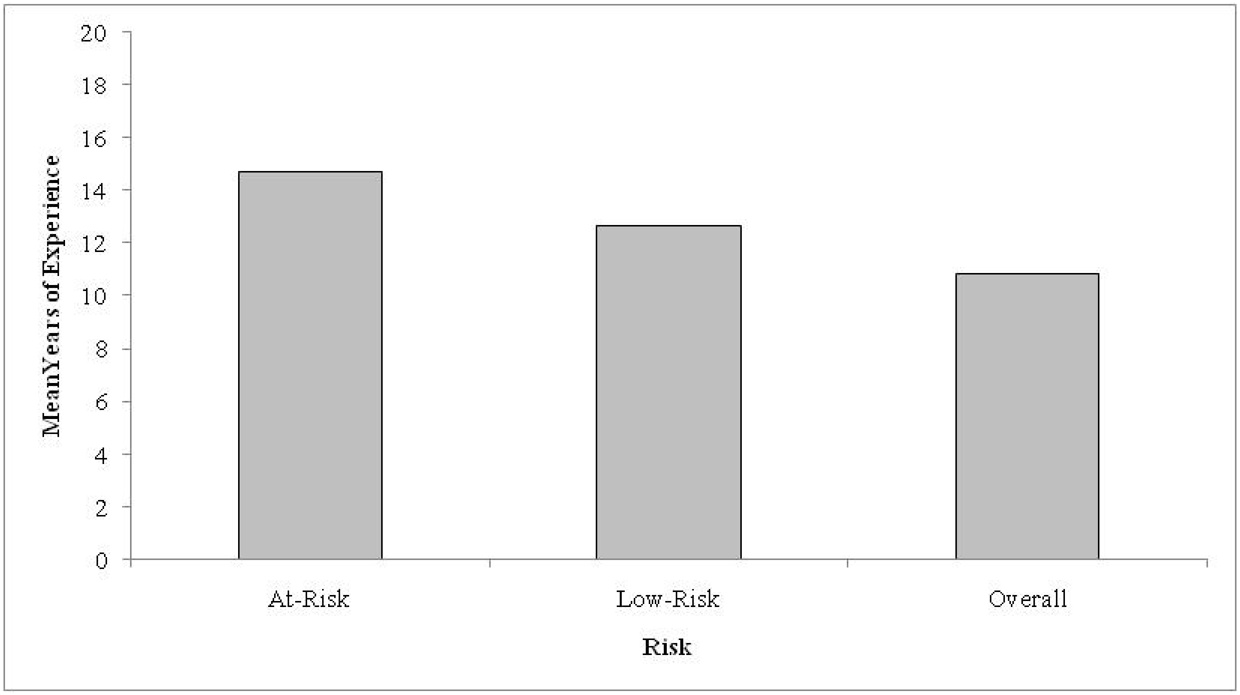


Figure 118. Mean Years of Experience by Risk Level. Based on a Subset of Drivers (n = 14)

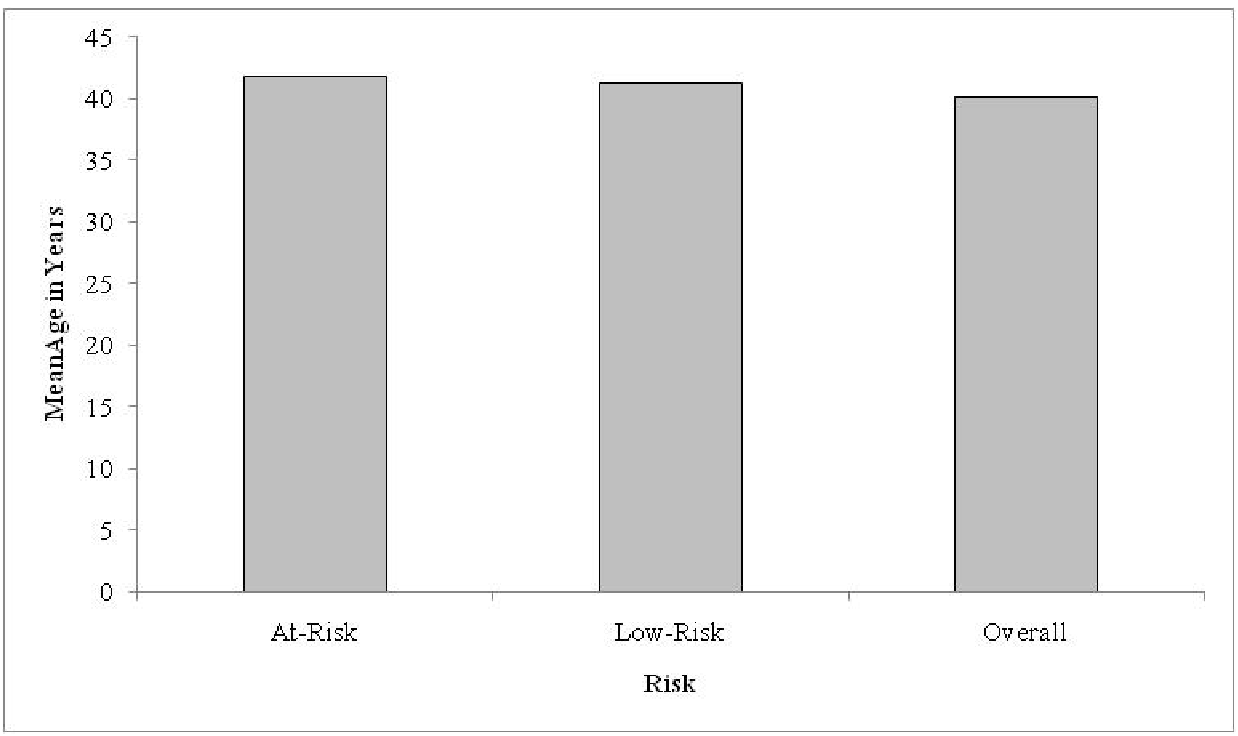


Figure 119. Mean Age by Risk Level. Based on a Subset of Drivers (n = 14)

Operational Characteristics. Among the at-risk drivers, five were long-haul drivers and two were line-haul drivers. Similarly, among the low-risk drivers, six were long-haul drivers and one was a line-haul driver. It should be noted that both groups only had seven drivers; therefore no conclusions should be drawn.

System Reliability

It should be reiterated that any conclusion based on the system reliability could be influenced by the method in which at-risk and low-risk drivers were identified. Given this, low-risk drivers had a significantly lower valid alert rate than drivers considered at-risk (Table 88). The low-risk drivers also had a significantly lower valid alert rate than the overall validity rate from all of the drivers in the study. Conversely, the drivers in the at-risk group had a significantly higher valid alert rate than the overall valid alert rate taken from all of the drivers in the study. Therefore, the system appeared to work more reliably for the drivers in the at-risk group.

Table 88. Validity Rate by Risk Level

Risk Level	Reviewed Alerts	Valid Alerts	Validity Rate (Percent)	SE	LCL*	UCL*
Low-Risk (n=7)	2,124	27	1.27	0.24	0.79	1.75
At-Risk (n=7)	1,418	149	10.51	0.81	8.91	12.10
Total (n=96)	15,386	721	4.69	0.17	4.35	5.02

* Observations were assumed to be independent, and the valid alerts were assumed to follow a binomial distribution. The confidence limits were constructed using a normal approximation of the binomial distribution.

On-The-Job Drowsiness

DFM PERCLOS Values. Both the DFM PERCLOS measure collected from the system and the manual PERCLOS values calculated from the NSTSCE baseline events were used for this analysis. Figure 120 shows the maximum, mean, and median for the baseline test and test conditions of both the at-risk and low-risk groups. During the baseline test condition there was no significant difference between the at-risk and low-risk groups. However, during the test condition the low-risk group had a significantly lower DFM PERCLOS average (Table 89).

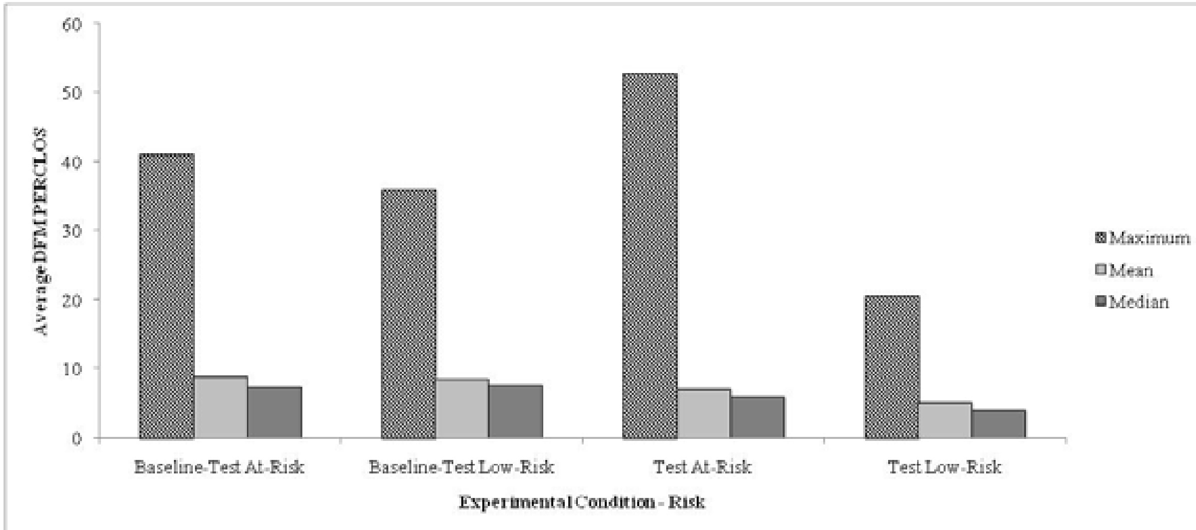


Figure 120. DFM PERCLOS Values by Experimental Group and Risk Level.

Based on a Subset of Drivers (n = 14)

Table 89. Mean DFM PERCLOS with 95 Percent Confidence Limits by Risk Level.

Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	N	Mean	SE	LCL*	UCL*
Baseline Test	Low-Risk	170	8.53	0.5	7.56	9.51
Baseline Test	At-Risk	65	8.76	0.88	7.04	10.48
Test	Low-Risk	285	5.02	0.21	4.61	5.44
Test	At-Risk	106	6.92	0.62	5.72	8.13

*Observations are taken from the same driver and are therefore not independent.

Evaluating the mean DFM PERCLOS values by taking driving hours into consideration, the study showed that drivers in the test condition considered low-risk had, on average, lower DFM PERCLOS values throughout the hours driving in the study (Figure 121).

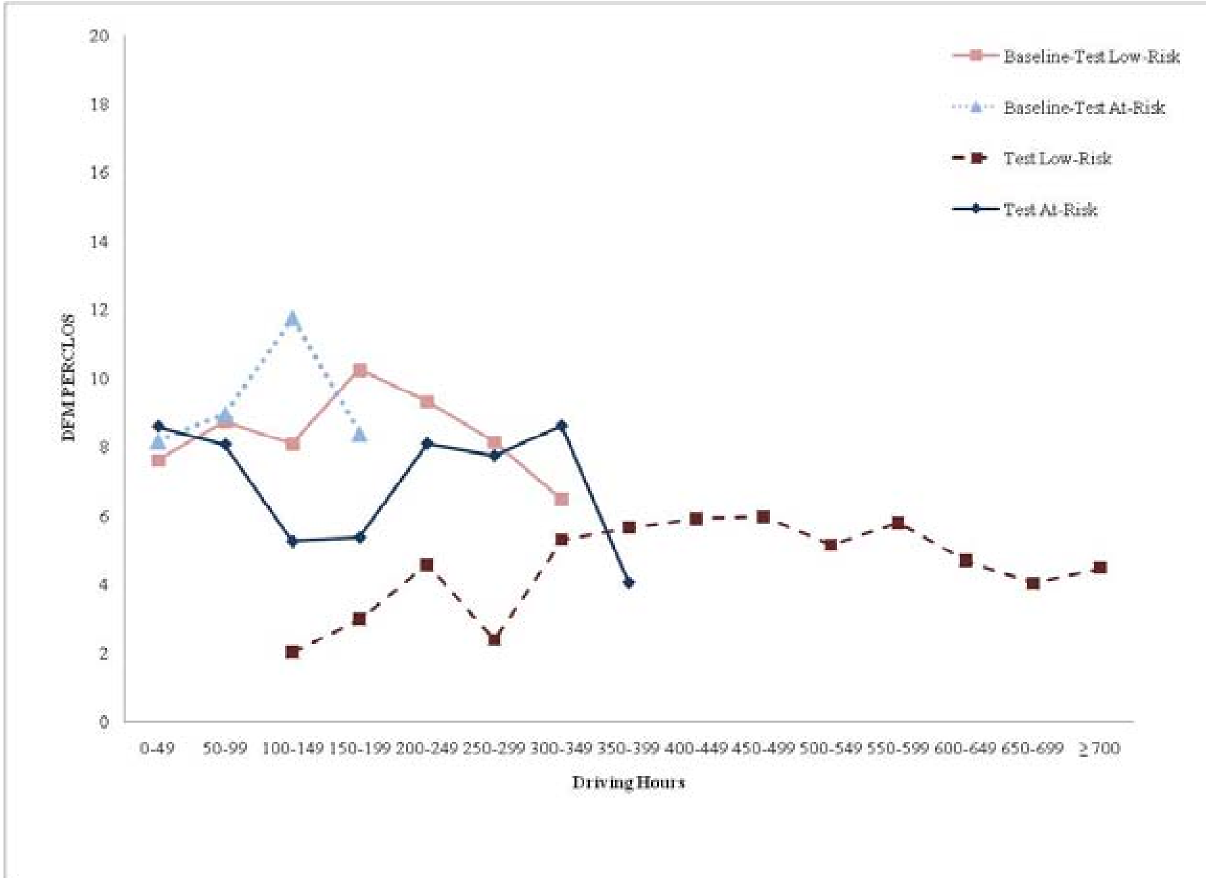


Figure 121. Mean DFM PERCLOS Values by Driving Hours and Risk Level.

Based on a Subset of Drivers (n = 14)

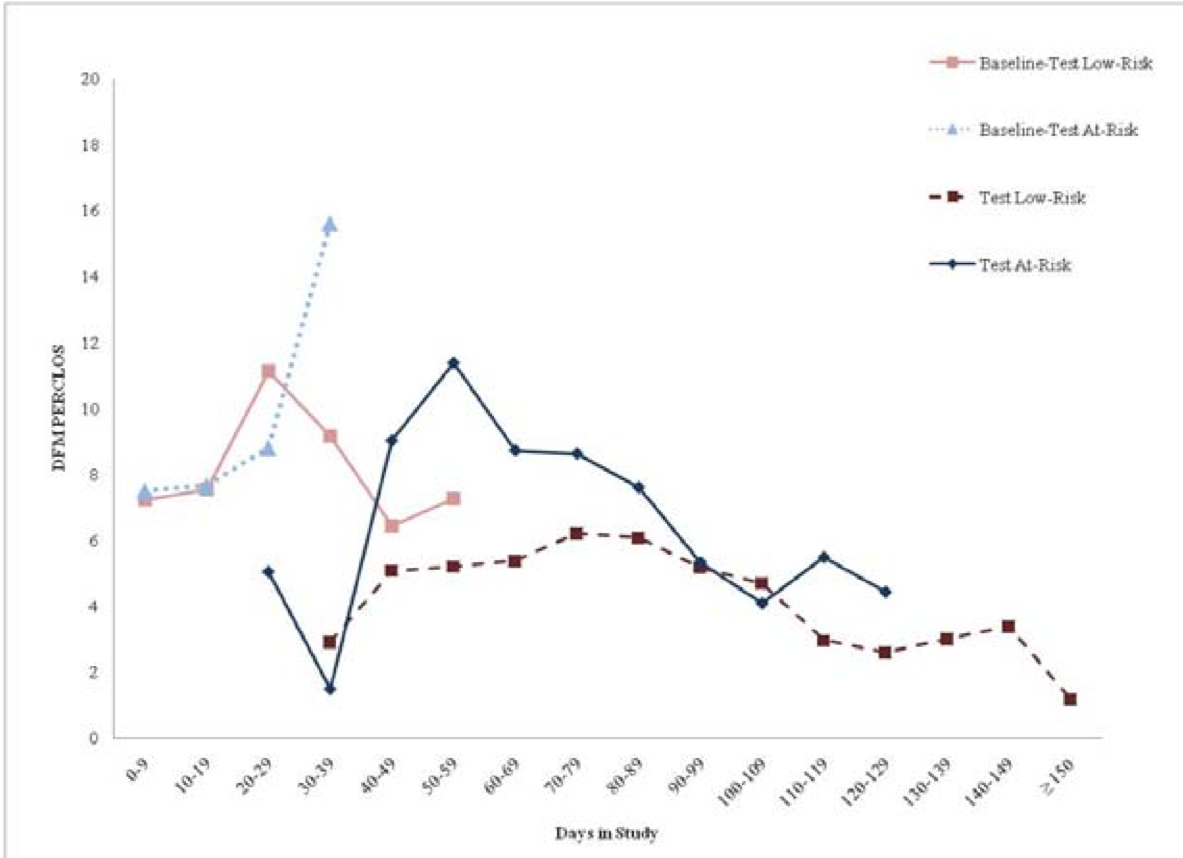


Figure 122. Mean DFM PRECLOS Values by Days in Study and Risk Level.

Based on a Subset of Drivers (n = 14)

Figure 122 shows the same overall pattern that Figure 121 showed. That is, on average, drivers considered low-risk had lower DFM PERCLOS values. As shown in Figure 123, this pattern extended to the day of the week. This pattern was still present when another measure of time was taken into consideration: weeks in the study (Figure 124).

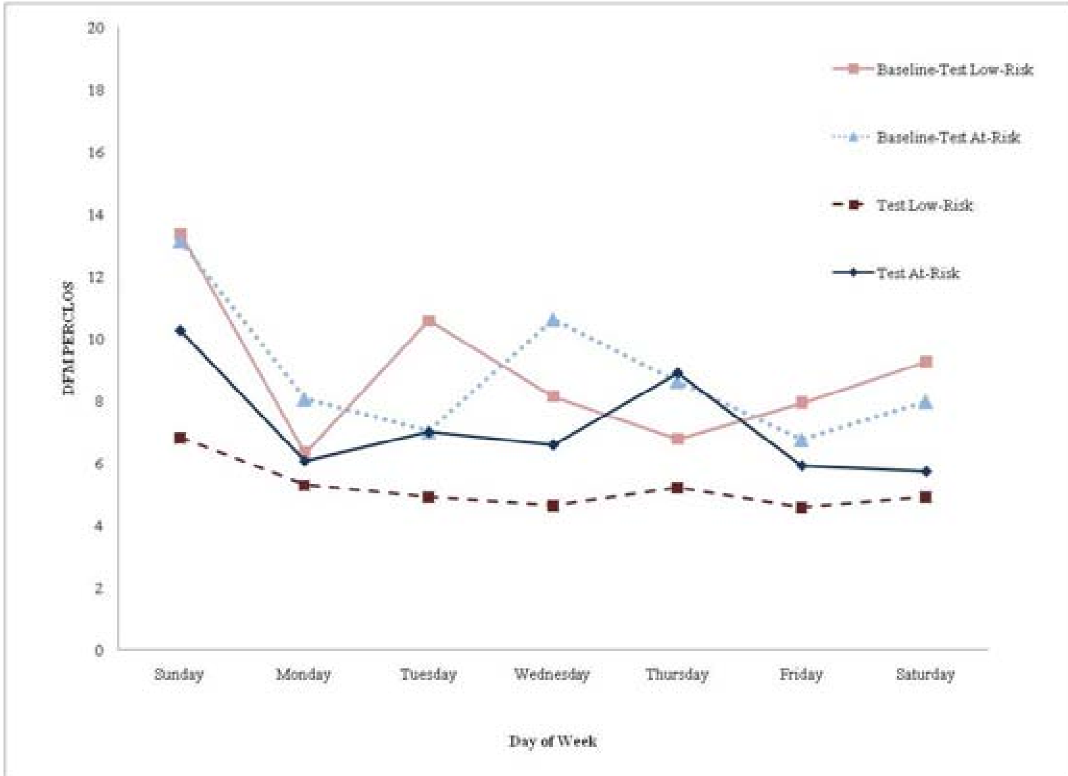


Figure 123. Mean DFM PERCLOS by Day of Week and Risk Level.

Based on a Subset of Drivers (n = 14)

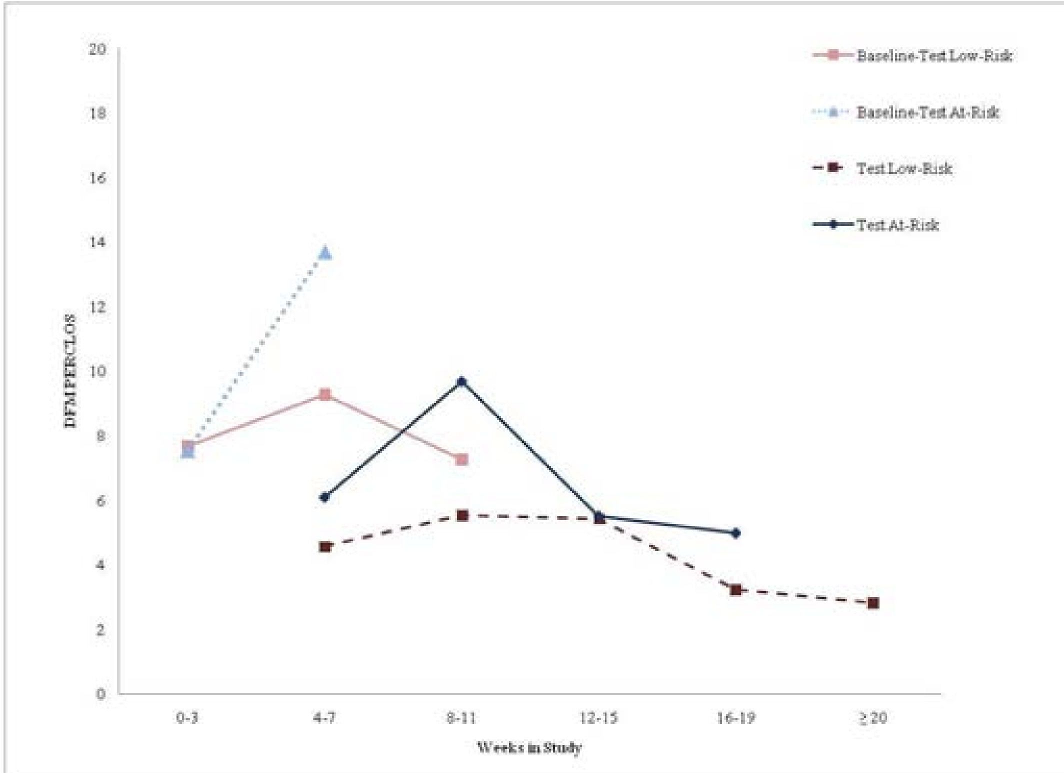


Figure 124. Mean DFM PERCLOS Values by Weeks in Study and Risk Level.

Based on a Subset of Drivers (n = 14)

Manual PERCLOS Values. Figure 125 presents an overview of the manual PERCLOS values calculated for NSTSCE baseline events. The maximums observed within the manual PERCLOS values (which range from 15.25 to 24.12) were lower than those recorded from the DFM. The measures of central tendency show that during the baseline period, the drivers considered low-risk had a lower mean and median manual PERCLOS value than those drivers considered at-risk. This pattern was also seen during the test condition phase of the study. From Table 90, the 95 percent confidence limits show a significant difference between at-risk and low-risk drivers during the baseline test condition; however, the difference seen during the test condition phase of the study was not significant.

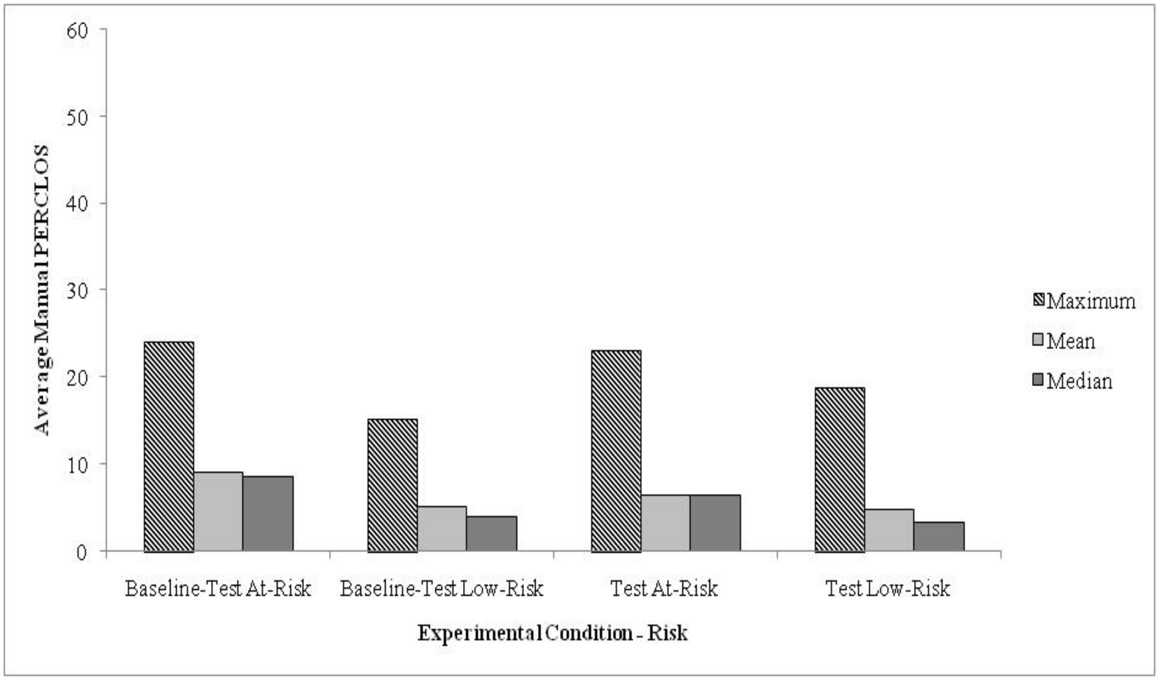


Figure 125. Maximum, Mean, and Median Manual PERCLOS Values by Risk Level.

Based on a Subset of Drivers (n = 14)

Table 90. Mean Manual PERCLOS Values and 95 Percent Confidence Levels by Risk Level. Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	N	Mean	SE	LCL*	UCL*
Baseline Test	Low-Risk	60	5.1	0.51	4.10	6.10
Baseline Test	At-Risk	40	9.09	0.92	7.29	10.89
Test	Low-Risk	87	4.76	0.49	3.80	5.71
Test	At-Risk	41	6.46	0.62	5.24	7.68

*Observations are taken from the same driver and are therefore not independent.

Evaluating the mean manual PERCLOS, in consideration of the hours of driving while in the study, showed low-risk drivers had lower average manual PERCLOS values. It also demonstrated that drivers in the low-risk group had a drop in their manual PERCLOS values between the baseline test condition and the test condition.

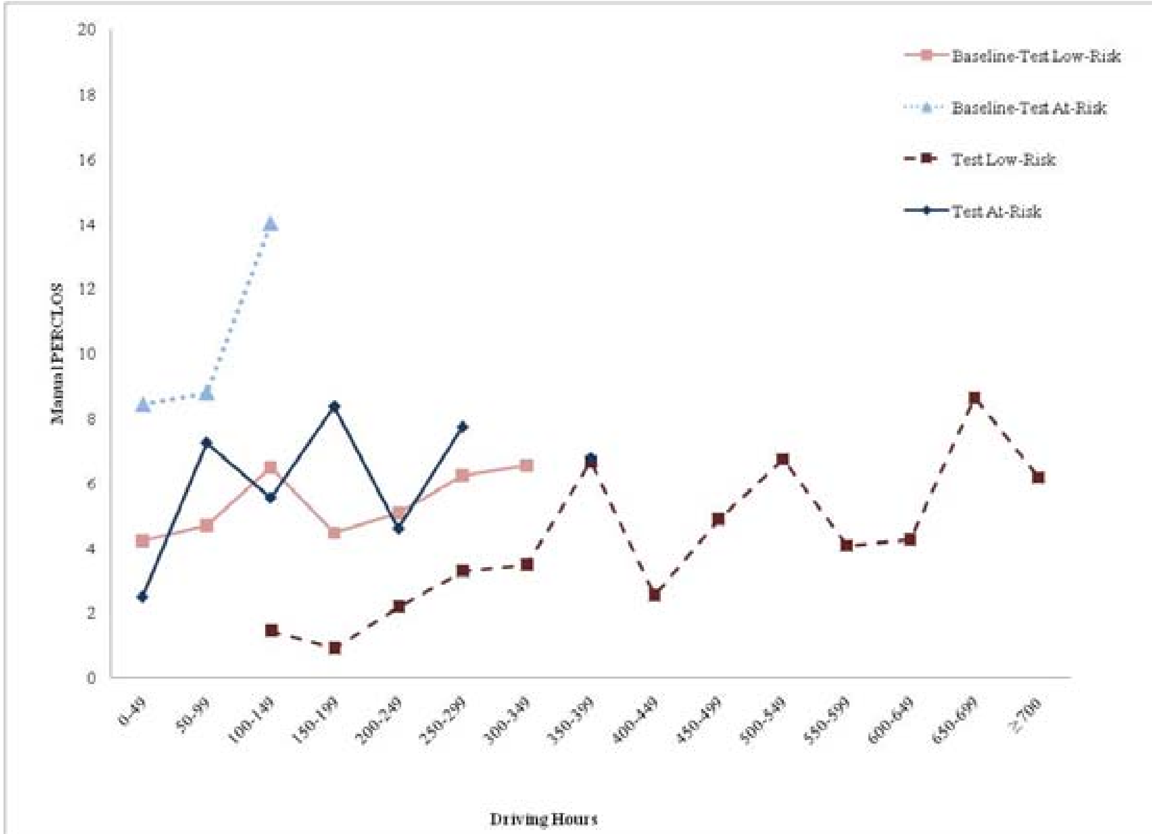


Figure 126. Mean Manual PERCLOS Values by Driving Hours and Risk Level.

Based on a Subset of Drivers (n = 14)

When manual PERCLOS values were examined as a function of days in study, the mean manual PERCLOS values do not have as strong of a pattern (Figure 127). The drivers considered low-risk overall have a lower mean manual PERCLOS than drivers considered at-risk. However, for low-risk drivers, no difference between baseline test and Test are present. low-risk drivers also have a lower average manual PERCLOS value for each day of the week when compared to the at-risk drivers (Figure 128).

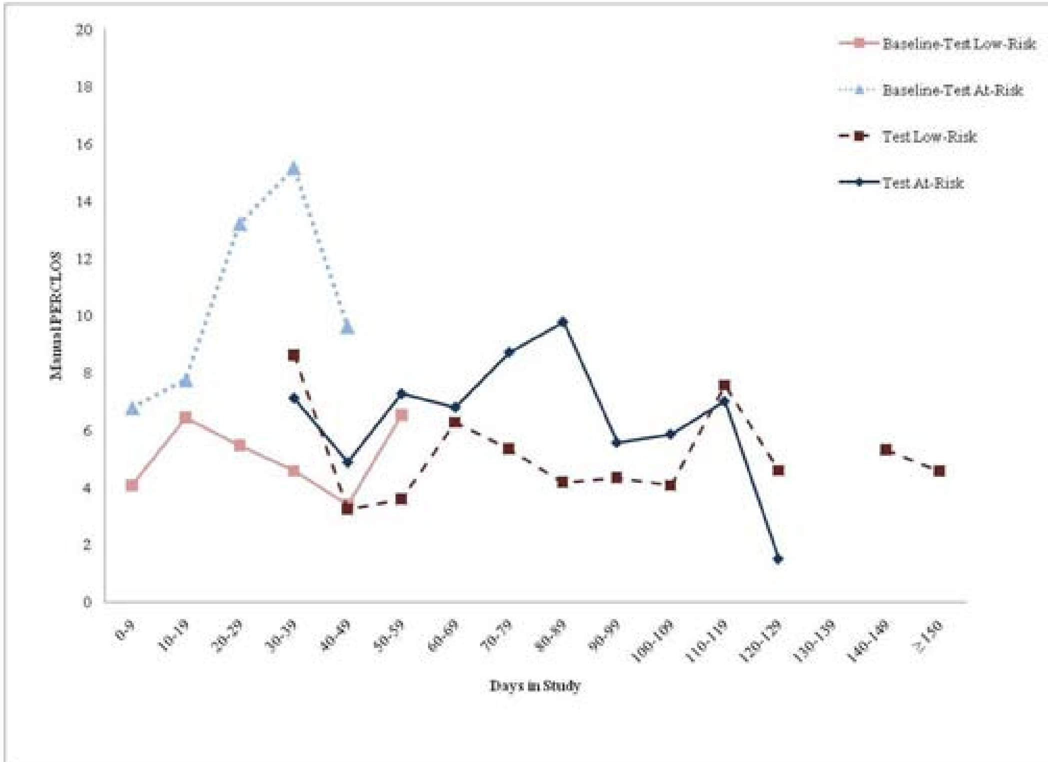


Figure 127. Mean Manual PERCLOS by Days in Study and Risk Level.

Based on a Subset of Drivers (n = 14)

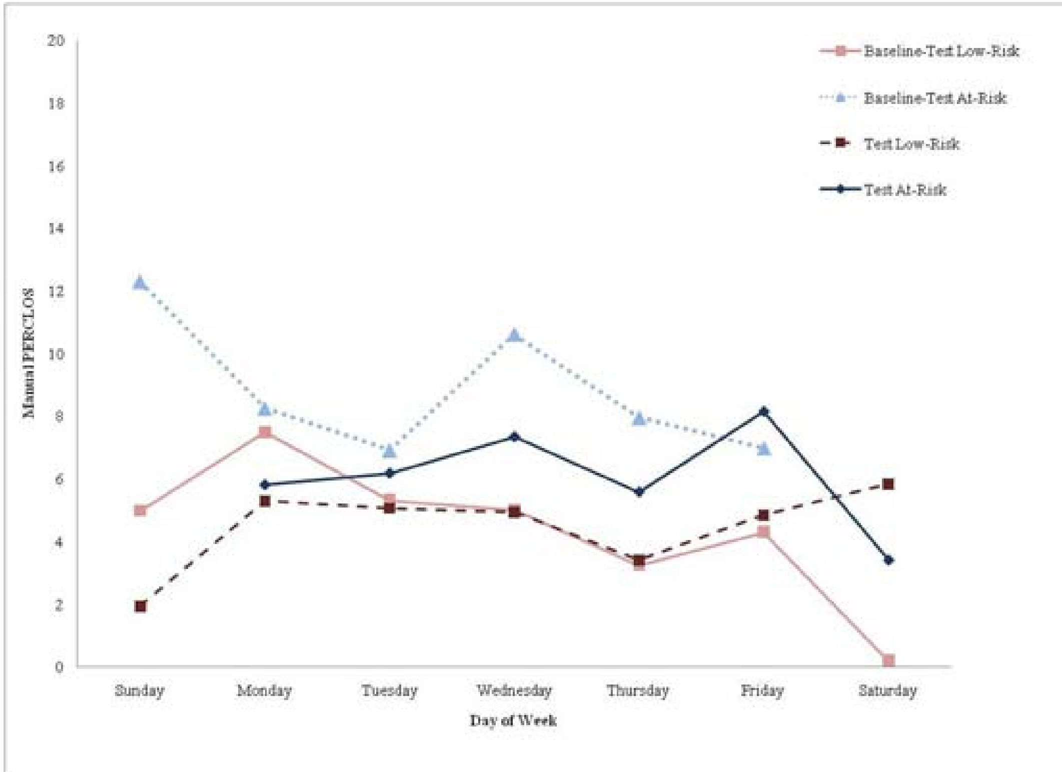


Figure 128. Mean Manual PERCLOS Values by Day of Week and Risk Level.

Based on a Subset of Drivers (n = 14)

When using weeks in the study as the time measure of interest, the manual PERCLOS values for low-risk drivers was once again lower on average (Figure 129). Low-Risk drivers on average had both a lower DFM PERCLOS and manual PERCLOS values.

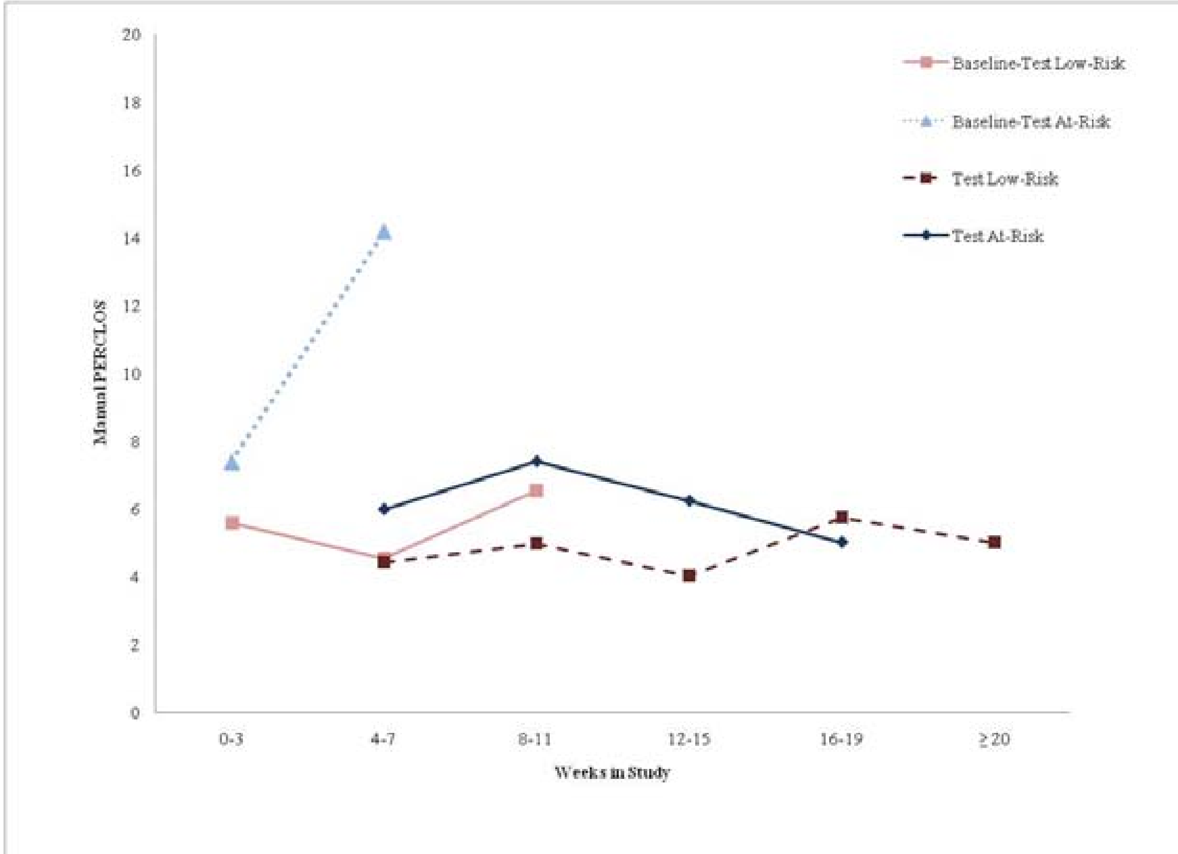


Figure 129. Mean Manual PERCLOS Values by Weeks in Study and Risk Level.

Based on a Subset of Drivers (n = 14)

Valid Alerts. The results from the previous analysis provide an overview of on-the-job-drowsiness by examining both DFM and manual PERCLOS. The DFM alerts were based on the PERCLOS calculated by the DFM system. If the PERCLOS values did not reach the predetermined threshold, the system did not provide the driver with an alert. However, it cannot be assumed that a driver who received more alerts experienced more on-the-job-drowsiness. The alerts the drivers received may be due to increased situational awareness (e.g., monitoring the mirrors) or other system limitations (e.g., not capturing the driver's eyes properly). Therefore the next analysis focuses only on the valid alerts that the drivers received during the course of the study.

Due to the small number of participants and the limited number of valid alerts for drivers in the at-risk and low-risk conditions, this analysis was also limited to descriptive statistics. Again, the at-risk drivers had more valid alerts than the low-risk drivers. The total number of valid alerts received from the subset of drivers was 176: at-risk drivers had 149 valid alerts (84.7%), and low-risk drivers had 27 valid alerts (15.3%). The number of valid alerts that low-risk drivers received is insufficient to support any further analysis.

Table 91. Proportion of Valid Alerts by Experimental Condition and Risk Level. Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	Frequency	Percentage
Baseline Test	Low-Risk	9	5.11
Baseline Test	At-Risk	94	53.41
Test	Low-Risk	18	10.23
Test	At-Risk	55	31.25
Total	---	176	100

Alert Response – Stopping Behavior. Figure 130 shows the elapsed time from the receipt of a valid DFM alert to the driver stopping the vehicle for a period of 10 min or more. Overall, drivers with a higher crash risk took a longer time to stop than drivers with a lower crash risk. Higher crash risk drivers drove an average of 1 hour and 2 min before stopping the vehicle in the baseline test condition, while higher crash risk drivers in the Test condition drove longer (1 hour and 15 min) before stopping. Drivers who had a lower crash risk drove an average of 1 hour and 43 min before stopping the vehicle in the baseline Test condition, while lower crash risk drivers in the Test condition drove 1 hour and 25 min before stopping.

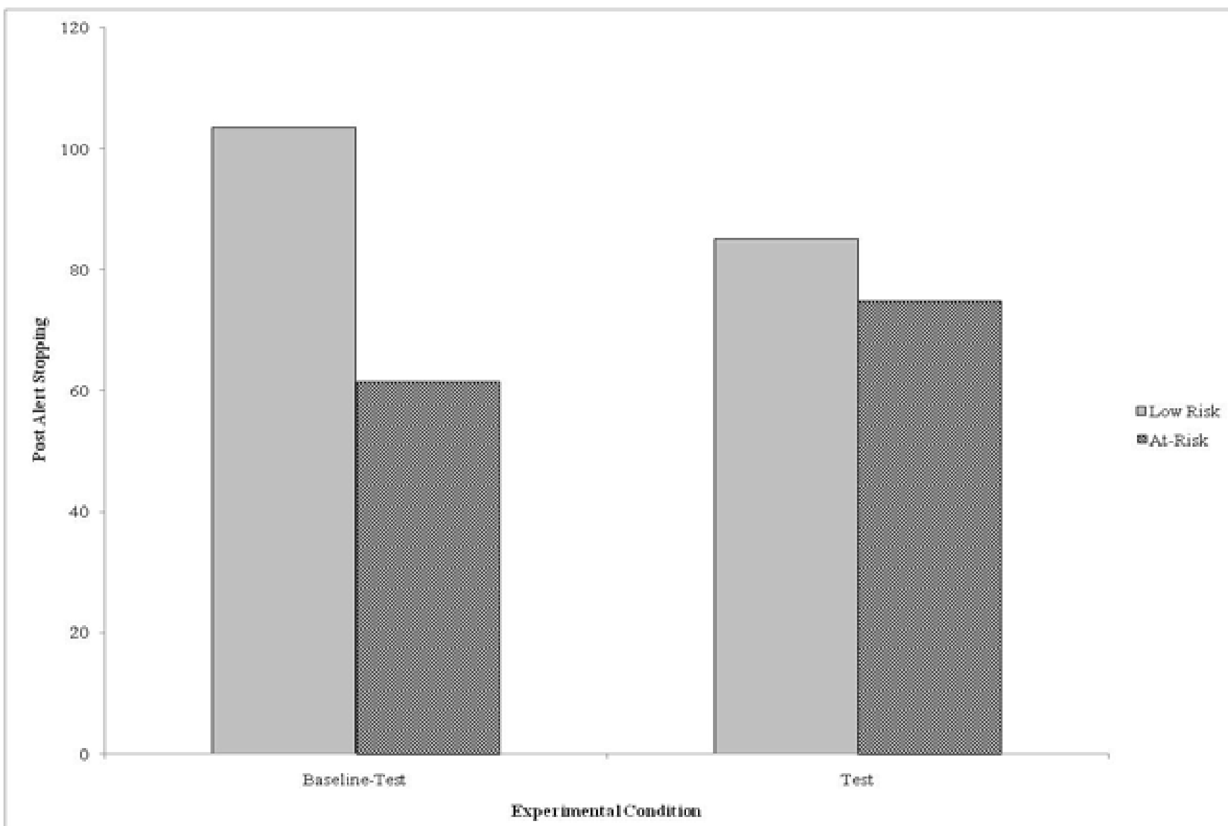


Figure 130. Elapsed Time (Minutes) from Valid Alert to Driver Stopping the Vehicle by Risk Level. Based on a Subset of Drivers (n = 14)

Post-Alert Behaviors. Figure 131 presents the behaviors that were observed after valid DFM alerts were presented to drivers considered to have lower and higher crash risks. With respect to those drivers with a lower crash risk, 9 valid alerts were generated in the baseline test Condition, while 18 valid alerts were generated in the test condition. These drivers were observed only to adjust their body, initiate a driving maneuver, or touch their body after receiving a valid DFM alert. With respect to those drivers considered to have a higher crash risk, 94 valid alerts were generated in the baseline test condition, while 55 valid alerts were generated in the test condition. In addition to the behaviors previously mentioned, these drivers were also observed to drink/eat and look to the side. However, this difference was likely due to the increased number of valid alerts in the baseline test condition and not necessarily to differences in the drivers' crash risk.

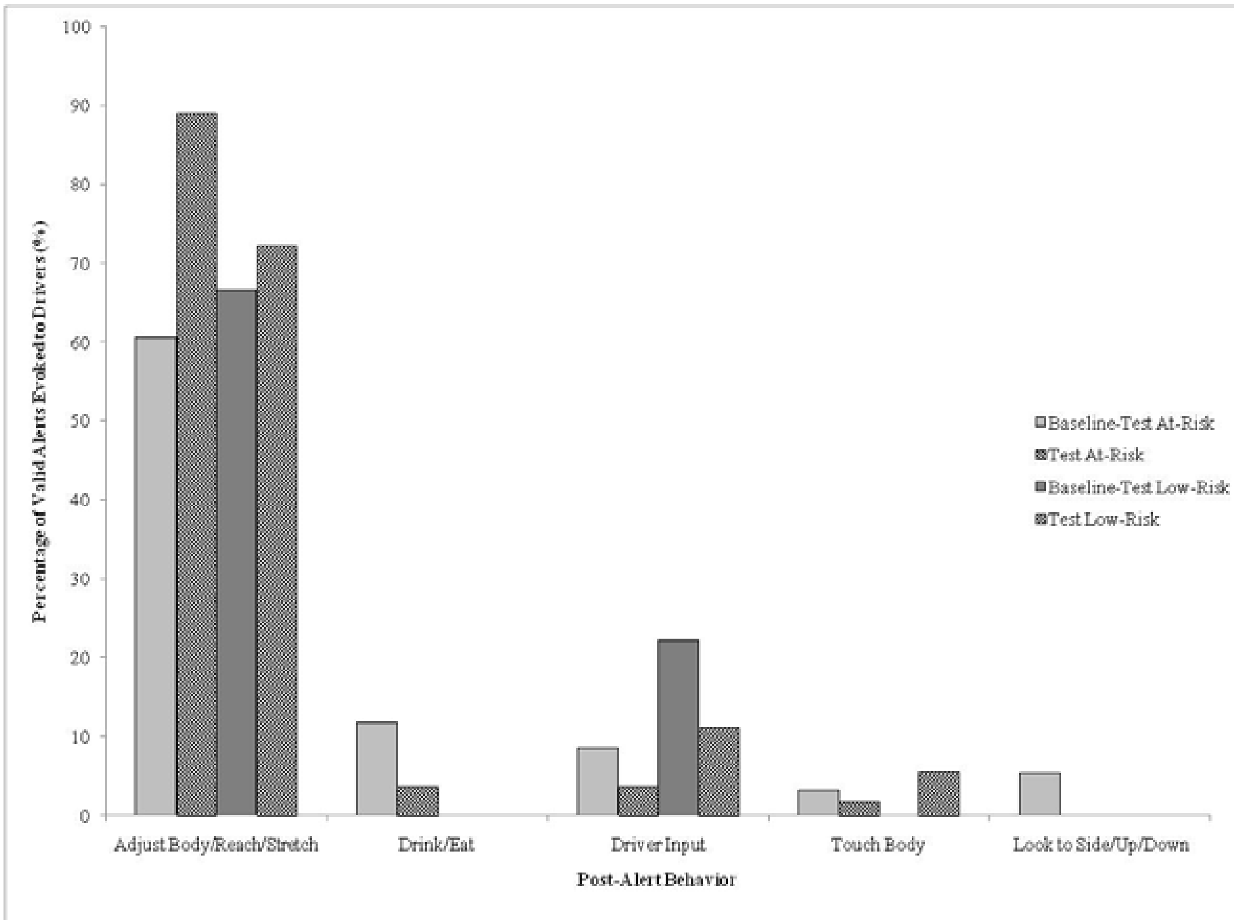


Figure 131. Percentage of Valid Alerts by Type of Behavior: Risk Level.

Based on a Subset of Drivers (n = 14)

Figure 132 shows the mean time that elapsed from a valid DFM alert to the observed behavior for drivers who had a lower or higher crash risk. No patterns are apparent.

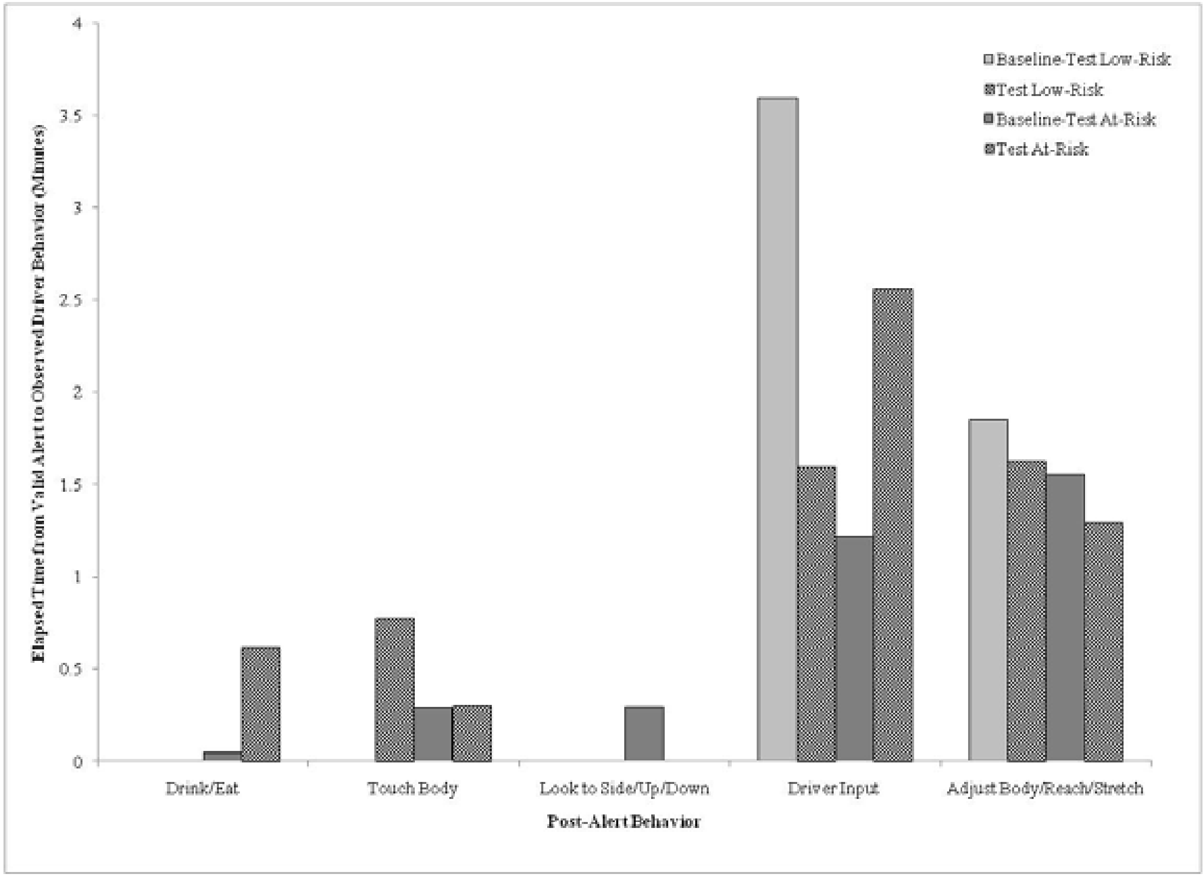


Figure 132. Elapsed Time from Valid Alert to Post-Alert Behavior: Risk Level.

Based on a Subset of Drivers (n = 14)

Sleep Hygiene

Table 92 presents the mean quantity of sleep obtained by drivers who were considered to have a higher or lower crash risk. There does not appear to be a difference in quantity of sleep between drivers who were considered to have a higher crash risk and those drivers considered to have a lower crash risk.

Table 92. Mean Sleep Quantity by Risk Level. Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	Number of Days	Mean Quantity of Sleep in 24h[2]	Standard Error (h)	LCL*	UCL*
Baseline Test	Low-Risk	266	6.02	0.15	5.73	6.32
Baseline Test	At-Risk	162	5.56	0.22	5.12	6
Test	Low-Risk	338	5.75	0.2	5.36	6.13
Test	At-Risk	223	6.33	0.17	6	6.67

*Observations are taken from the same driver and are therefore not independent.

Table 93 presents the mean sleep efficiency for drivers who were considered to have a higher or lower crash risk. No patterns are apparent.

Table 93. Sleep Efficiency by Risk Level. Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	Number of Days	Efficiency (Percent)	Standard Error (h)	LCL*	UCL*
Baseline Test	Low-Risk	238	0.9	0.01	0.88	0.92
Baseline Test	At-Risk	145	0.82	0.01	0.8	0.84
Test	Low-Risk	302	0.81	0.01	0.79	0.83
Test	At-Risk	199	0.84	0.01	0.82	0.86

*Observations are taken from the same driver and are therefore not independent.

Table 94 presents the mean number of awakenings for drivers who were considered to have a higher or lower crash risk. It appears that drivers considered to have a lower crash risk woke up more frequently during the O-O (sleep onset-sleep offset) interval than drivers who were considered to have a higher crash risk.

Table 94. Awakenings by Risk Level. Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	Number of Days	Mean Awakenings	Standard Error (h)	LCL*	UCL*
Baseline Test	Low-Risk	238	12.32	0.63	11.09	13.55
Baseline Test	At-Risk	145	9.8	0.65	8.53	11.07
Test	Low-Risk	302	12.32	0.63	11.09	13.55
Test	At-Risk	199	7.27	0.44	6.41	8.13

*Observations are taken from the same driver and are therefore not independent.

[2] Note that the sleep quantity values cannot be used to evaluate overall sleep of CMV drivers, but only provide a means of making relative comparisons across groups. The method of eliminating nap data precludes making more general or absolute statements regarding overall sleep.

Table 95 presents the mean number of SSPs (scored sleep periods) longer than 20 min for drivers who were considered to have a higher or lower crash risk. It appeared that drivers considered to have a lower crash risk obtained more SSPs longer than 20 min as compared to drivers considered to have a higher crash risk.

Table 95. Mean Number of SSP Over 20 Min by Risk Level.

Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	Number of Days	Mean #SSPs Longer than 20 min	SE (min)	LCL*	UCL*
Baseline Test	Low-Risk	238	5.29	0.21	4.88	5.7
Baseline Test	At-Risk	145	4.3	0.19	3.93	4.67
Test	Low-Risk	302	4.44	0.17	4.11	4.77
Test	At-Risk	199	3.83	0.15	3.54	4.12

*Observations are taken from the same driver and are therefore not independent.

Table 96 presents the mean longest SSP obtained by drivers who were considered to have a higher or lower crash risk. It appeared that drivers considered to have a higher crash risk obtained the longest SSP on average as compared to drivers considered to have a lower crash risk.

Table 96. Mean Longest SSP in an O-O Interval by Risk Level.

Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	Number of Days	Mean Longest SSP	SE (min)	LCL*	UCL*
Baseline Test	Low-Risk	238	129.6	5.16	119.5	139.7
Baseline Test	At-Risk	145	145	8.77	127.8	162.2
Test	Low-Risk	302	144.4	6.73	131.2	157.6
Test	At-Risk	199	200.2	7.48	185.5	214.9

*Observations are taken from the same driver and are therefore not independent.

Involvement in Safety-Critical Events

The total data set includes 1,124 SCEs. This data set consisted of 28 crashes, 112 near-crashes, and 984 crash relevant conflicts. The at-risk/low-risk drivers who are the object of this research question had 127 SCEs: 1 crash, 13 near-crashes, and 113 crash relevant conflicts. Table 97 shows the frequency of SCEs per driver for each risk level. Table 98 shows the number of hours driven by risk for the baseline test and the Test condition.

Table 97. Safety Critical Events by Experimental Condition and Risk Level.

Based on a Subset of Drivers (n = 14)

Experimental Condition	Low-Risk	At-Risk	Total
Baseline Test	39	23	62
Test	57	8	65
Total	96	31	127

Table 98. Number of Hours Driven by Experimental Condition and Risk Level.

Based on a Subset of Drivers (n = 14)

Experimental Condition	Low-Risk	At-Risk	Total
Baseline	1,849	518	2,367
Test	2,677	650	3,327
Total	4,526	1,168	5,694

Safety Critical Events by Risk Level. Table 99 shows the frequency of SCEs based on the driver's risk. Results indicate that both risk groups have the same distribution of SCEs (Fisher's $p = 0.19$).

Table 99. Frequency of SCEs by Type of SCE and Risk Level.

Based on a Subset of Drivers (n = 14)

Risk Level	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
At-Risk	1	4	26	31
Low-Risk	0	9	87	96
Total	1	13	113	127

In addition to the frequency, the SCE rate was computed as the number of SCEs per 100 hours driven by Risk Level, as shown in Figure 133. The SCE rates for At-Risk drivers were higher than for low-risk drivers, with the average frequency being 25 percent higher for the At-Risk group.

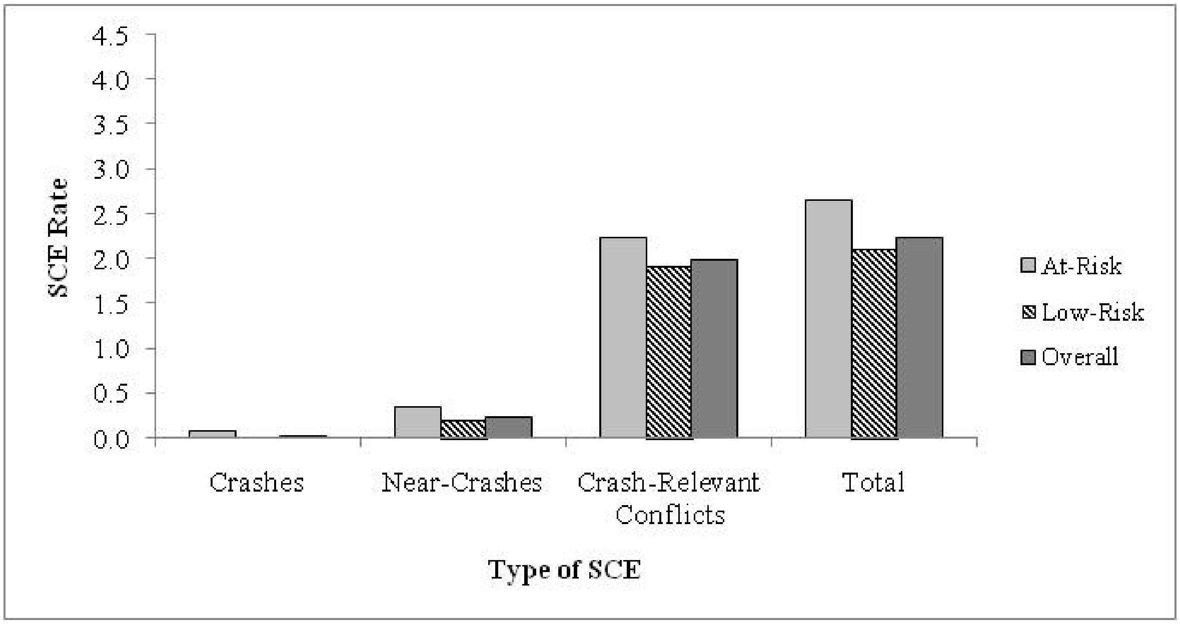


Figure 133. SCE Rates by Event Type and Risk Level.

Based on a Subset of Drivers (n = 14)

Safety Critical Events by Experimental Condition. Table 100 presents the SCEs by the experimental conditions of interest for this research question (i.e., baseline test and test conditions). Those drivers considered as At-Risk had 31 SCEs, with 23 of them occurring in the baseline period. The drivers considered low-risk had 96 SCEs, with 39 of them in the baseline period. Examination by a Fisher test indicated that there are statistically significant differences in the distribution of events for the four groups. ($p = 0.002$). However, the low-risk baseline test and test condition distributions are not statistically significant ($p = 0.34$).

Table 100. Frequencies of SCE by Experimental Condition and Risk Level. Based on a Subset of Drivers (n = 14)

Experimental Condition and Risk Level	Crash-Relevant Conflicts	Near-Crashes	Crashes
Baseline-Test At-Risk	22	0	1
Test At-Risk	4	4	0
Baseline-Test Low-Risk	34	5	0
Test Low-Risk	53	4	0

The SCE rates, as the number of SCEs per 100 hours driven, by each condition are compared in Figure 134. The crash rate for the at-risk/baseline test group was statistically higher than the rates observed in other groups. However, generalizations from this difference are difficult to reach due to the small sample size within this subset of the overall data.

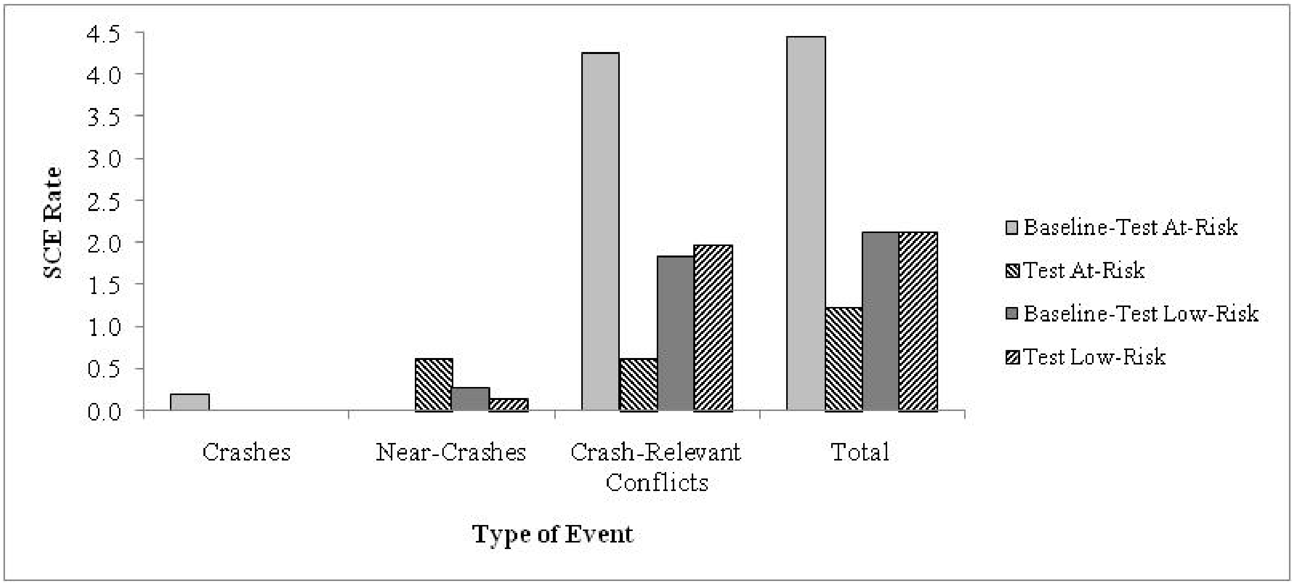


Figure 134. SCE Ratio by Experimental Condition and Risk Level.

Based on a Subset of Drivers (n = 14)

Manual PERCLOS For SCE by Risk Level. Two surrogate measures of driver drowsiness were employed to characterize level of drowsiness during SCEs. The first measure evaluated was manual PERCLOS. For the 127 SCEs, 21 SCEs were lacking sufficient video data to calculate manual PERCLOS (Wiegand, Hanowski, Olson, & Melvin, 2008). The mean manual PERCLOS value was 6, with a maximum observed value of 24.5. Table 101 shows the descriptive statistics for the manually calculated PERCLOS values at the moment of SCE occurrence by Risk and experimental condition. The average manual PERCLOS value was higher for the at-risk group than for the low-risk group; these differences were statistically significant (Fischer $p = 0.0171$).

Table 101. Manual PERCLOS Values by Experimental Condition and Risk Level

Experimental Condition	Risk Level	Total SCE	SCE With Manual PERCLOS Available	Minimum	Maximum	Mean
Baseline Test	At-Risk	23	18	2	25	11
Test	At-Risk	8	7	2	25	11
Baseline Test	Low-Risk	39	32	0	13	5
Test	Low-Risk	57	49	0	13	4

The frequencies of manual PERCLOS values by experimental condition and Risk Level are presented in Table 102. The table shows that 72 percent of SCEs have manual PERCLOS values below 12. Given that approximately 17 percent of the SCE did not have manual PERCLOS available, no further statistical analysis was performed.

Table 102. Manual PERCLOS Over 12 by Experimental Condition and Risk Level. Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	PERCLOS <12	PERCLOS ≥12	PERCLOS N\A	Total
Baseline Test	At-Risk	9	9	5	23
Test	At-Risk	4	3	1	8
Baseline Test	Low-Risk	30	2	7	39
Test	Low-Risk	48	1	8	57
Total	---	91	15	21	127

Drowsiness-Related Driving Behavior Evaluation for SCE by Risk Level. As mentioned earlier, all SCEs were evaluated and driver behavior was categorized if the driver appeared drowsy, sleepy, asleep, fatigued, or showed signs of reduced alertness that occurred during the period of time leading to the SCE. Table 103 shows the frequency of drowsiness-related SCEs by experimental condition and Risk Level. Of the 127 SCEs, 18 are drowsiness related. Results indicate that the differences in frequency of SCEs were not statistically significant among experimental conditions and Risk Level (p = 0.0575).

Table 103. Frequency of Drowsiness-Related SCE by Experimental Condition and Risk Level. Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	Not Drowsy	Drowsy	Total
Baseline Test	At-Risk	20	3	23
Test	At-Risk	4	4	8
Baseline Test	Low-Risk	34	5	39
Test	Low-Risk	51	6	57
Total	---	109	18	127

Table 104 shows the distribution of SCEs by group and test condition and if they are drowsiness-related. Table 105 shows the SCE rate by experimental condition, Event Type, and Driver's Risk Level. The SCE rate of drowsiness-related events for the test condition under Low Risk was 0.22, compared with 0.62 for the drivers in the at-risk level. However, during the baseline period the Low Risk rate of drownsiness-related events was 0.27 and for the at-risk level it was 0.58. This initially indicates that there is no reduction in the number of drownsiness-related events associated with the use of the DFM..

Table 104. Drowsiness-Related SCEs by Experimental Condition, Event Type, and Risk Level. Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	Drowsiness Related	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Baseline Test	At-Risk	No	1	0	19	20
Baseline Test	At-Risk	Yes	0	0	3	3
Baseline Test	At-Risk	Total	1	0	22	23
Test	At-Risk	No	0	2	2	4
Test	At-Risk	Yes	0	2	2	4
Test	At-Risk	Total	0	4	4	8
Baseline Test	Low-Risk	No	0	5	29	34
Baseline Test	Low-Risk	Yes	0	0	5	5
Baseline Test	Low-Risk	Total	0	5	34	39
Test	Low-Risk	No	0	4	47	51
Test	Low-Risk	Yes	0	0	6	6
Test	Low-Risk	Total	0	4	53	57

Table 105. SCE Rate by 100 Hours Driven for Drowsiness-Related SCEs by Experimental Condition, Event Type, and Risk Level. Based on a Subset of Drivers (n = 14)

Experimental Condition	Risk Level	Drowsiness Related	Crashes	Near-Crashes	Crash Relevant Conflicts	Total
Baseline Test	At-Risk	No	0.19	0.00	3.67	3.86
Baseline Test	At-Risk	Yes	0.00	0.00	0.58	0.58
Baseline Test	At-Risk	Total	0.19	0.00	4.24	4.44
Test	At-Risk	No	0.00	0.31	0.31	0.62
Test	At-Risk	Yes	0.00	0.31	0.31	0.62
Test	At-Risk	Total	0.00	0.62	0.62	1.23
Baseline Test	Low-Risk	No	0.00	0.27	1.57	1.84
Baseline Test	Low-Risk	Yes	0.00	0.00	0.27	0.27
Baseline Test	Low-Risk	Total	0.00	0.27	1.84	2.11
Test	Low-Risk	No	0.00	0.15	1.76	1.91
Test	Low-Risk	Yes	0.00	0.00	0.22	0.22
Test	Low-Risk	Total	0.00	0.15	1.98	2.13

Driver at Fault. The frequency of vehicle at fault is displayed in Table 106. The driver of the experimental vehicle (Vehicle 1) was judged to be at fault between 63 to 87 percent of the time. Results showed no statistically significant differences in the distribution of SCE by Risk or test conditions (Fischer's $p = 0.45$).

Table 106. Vehicle at Fault by Experimental Condition and Risk Level. Based on a Subset of Drivers (n = 14)

Experimental Condition and Risk Level	Unknown	Vehicle 2	Vehicle 1
Baseline-Test At-Risk	1	2	20
Test At-Risk	0	3	5
Baseline-Test Low-Risk	2	9	28
Test Low-Risk	2	8	47

Safety Critical Events Within the Operating Envelope That Received Alerts. Similar to the analysis for Research Question 3, the alerts presented for the SCEs in this subset of data were identified.

The DFM recorded 78 initial alerts in the hour prior to 9 of the 127 SCEs. Forty-nine of these alerts occurred for 5 SCEs in the at-risk group. The percentage of SCEs that received alerts was 13 and 25 percent for the at-risk group for the baseline and test condition, respectively. The percentage of SCEs that received alerts for the low-risk group was 3 percent for the baseline and 5 percent for the test condition. Six SCEs received 55 full alerts. The percentage of SCEs that received full alerts for the test condition was 25 percent for the at-risk group compared with 5 percent of the low-risk group. Valid initial alerts occurred in the at-risk group in only 3 of the SCEs, and full alerts occurred in only one SCE.

Human-Machine Interaction

As previously noted, this analysis consists of seven participants in each of the two risk categories. Therefore, examining the frequency of changes is equivalent to examining the rate per driver of changes made to the DFM. There was no discernable pattern found when examining the frequency of DFM sensitivity level by time of day (Figure 135). From 16:00 to 21:00 there was a sharp increase in frequency of changes made to the warning sound level of the DFM in the low-risk drivers. However, other than this singular event there was no significant difference between the two groups (Figure 136).

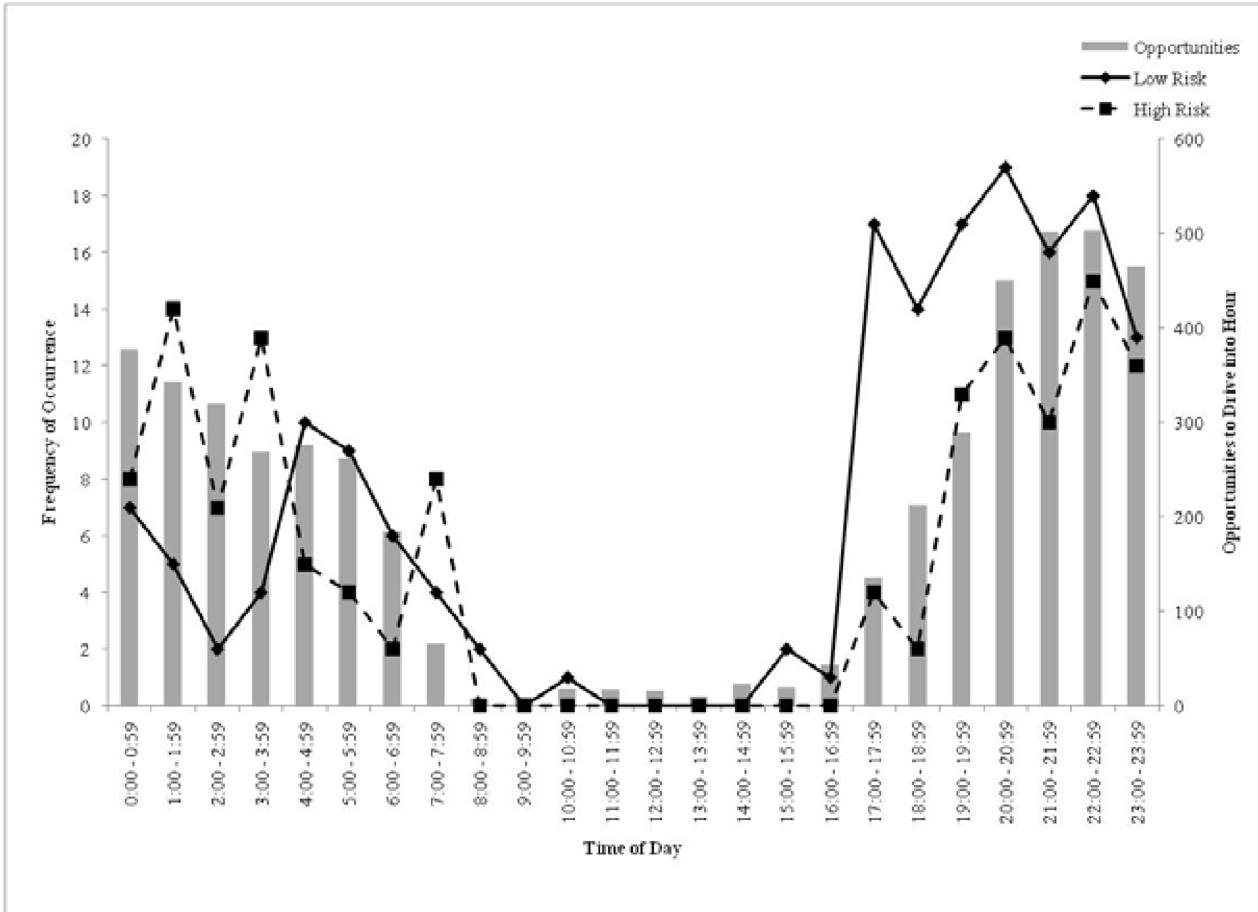


Figure 135. Frequency of DFM Sensitivity Level Changes by Time of Day and Risk Level. Based on a Subset of Drivers (n = 14)

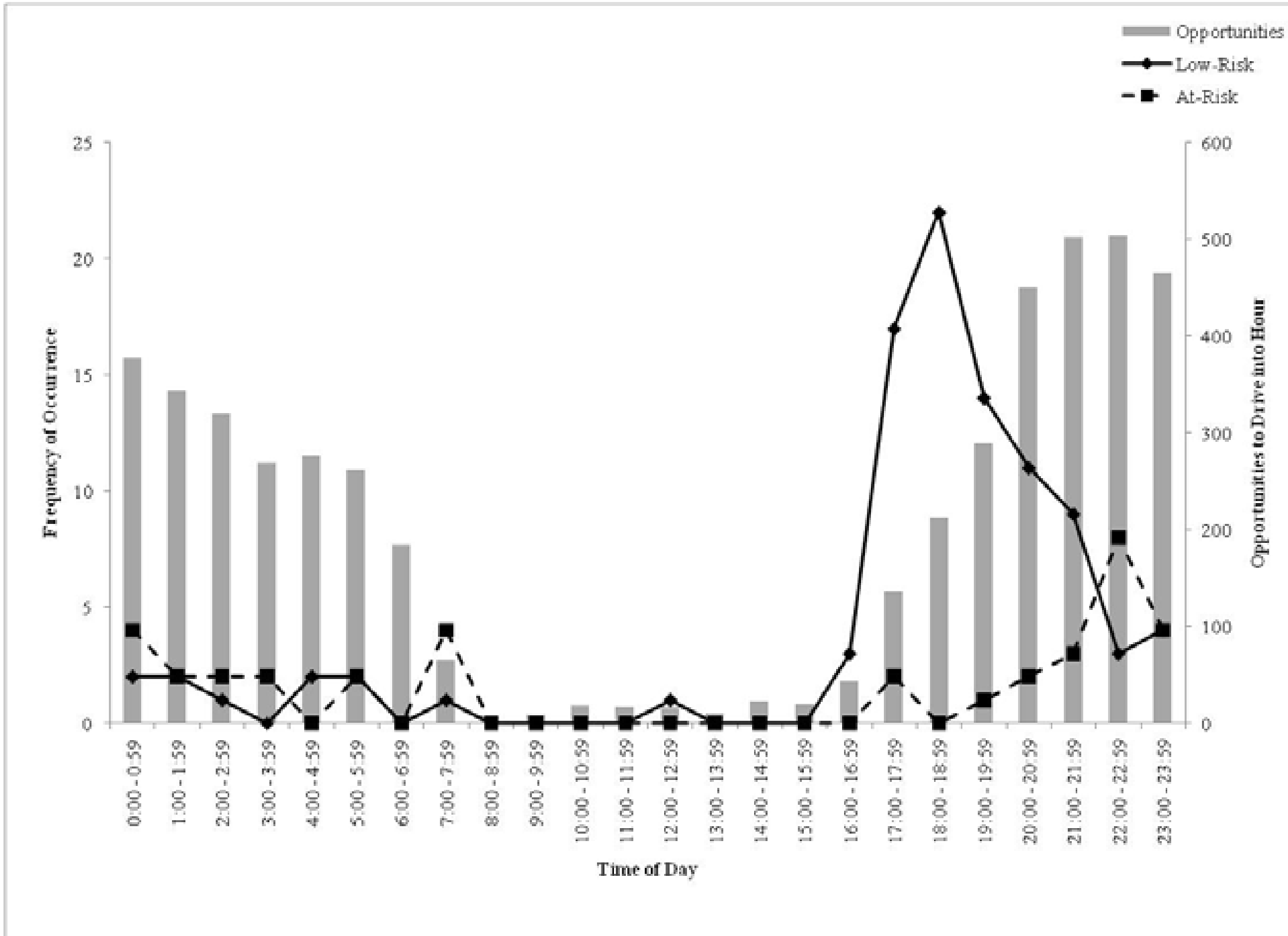


Figure 136. Frequency of DFM Warning Sound Changes by Time of Day and Risk Level. Based on a Subset of Drivers (n = 14)

The same sharp increase in frequency of changes made to the DFM warning sound level was observed in the frequency of changes made to the DFM brightness by the low-risk drivers.

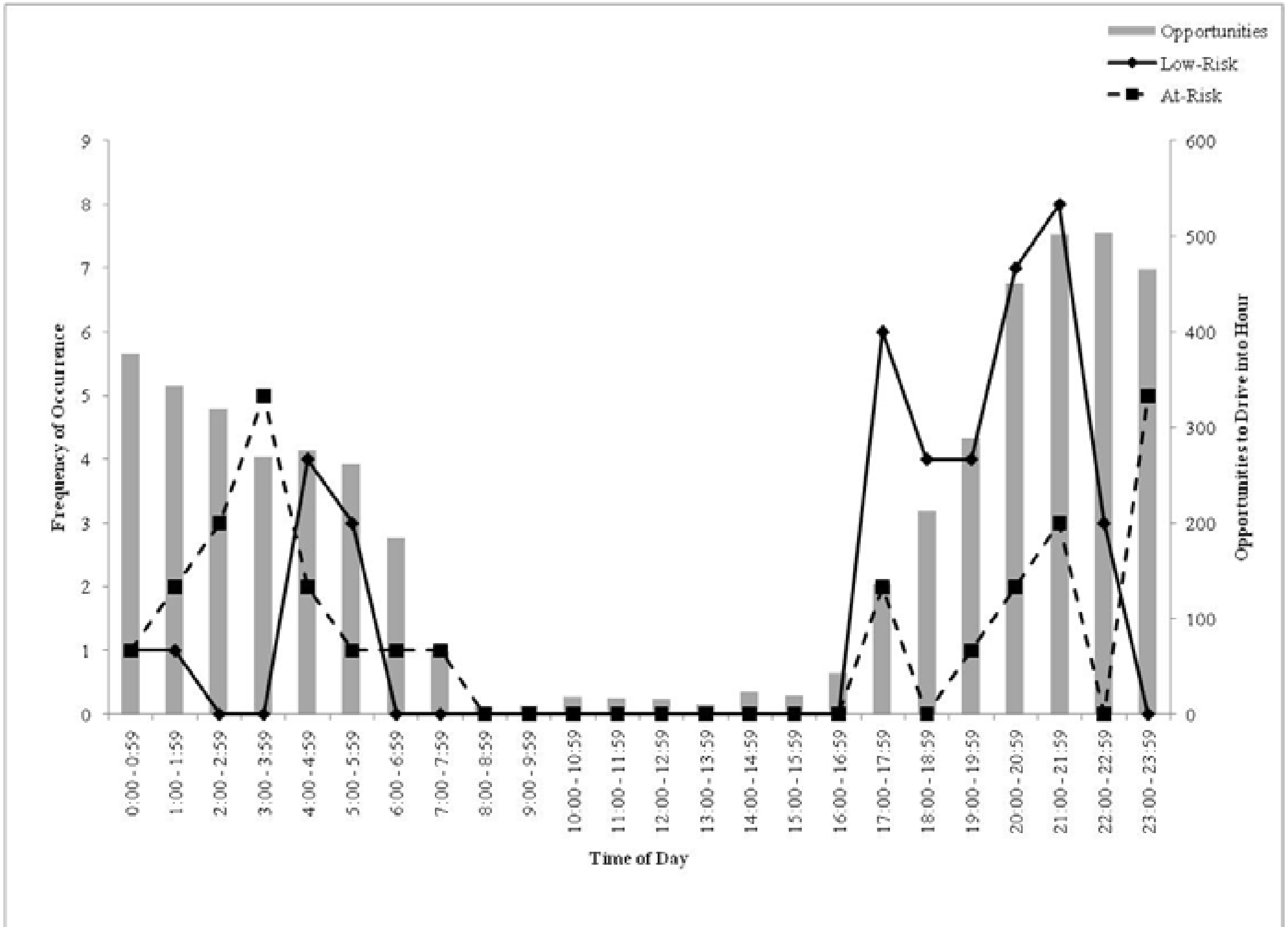


Figure 137. Frequency of DFM Brightness Changes by Time of Day and Risk Level.

Based on a Subset of Drivers (n = 14)

For those drivers considered at-risk, there was an apparent downward pattern in the frequency of changes made to the DFM sensitivity when viewed as a function of driving hours in a shift. However, this pattern was not present for low-risk drivers (Figure 138). The frequency of changes made to the DFM warning sounds has a downward pattern for the at-risk drivers (Figure 139). The low-risk drivers have a sharp increase in their frequency of DFM warning sound level changes from three to six hours into the shift. Then, from the sixth hour, the frequency of the DFM warning sound level changes begins to fall (Figure 139).

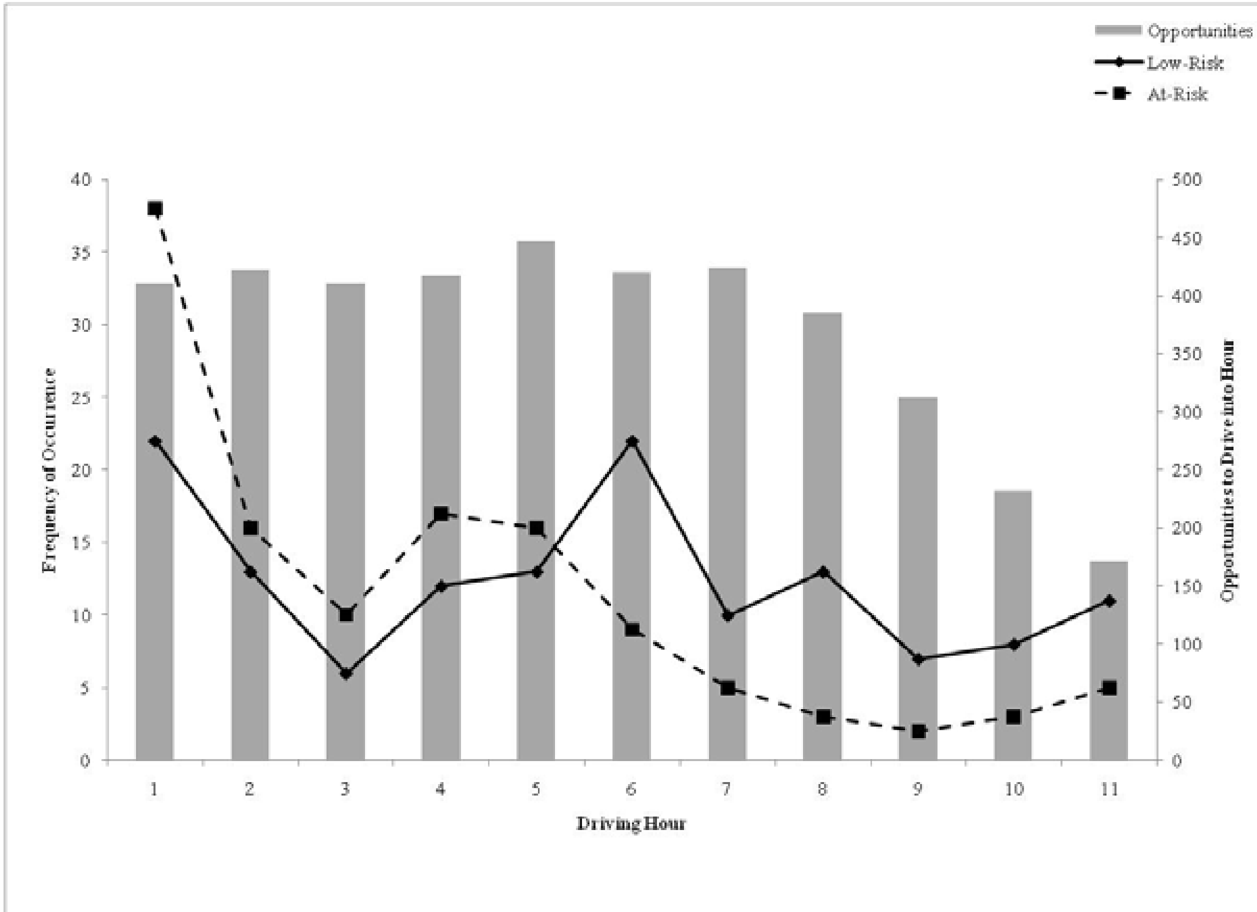


Figure 138. Frequency of DFM Sensitivity Level Changes by Driving Hour and Risk Level. Based on a Subset of Drivers (n = 14)

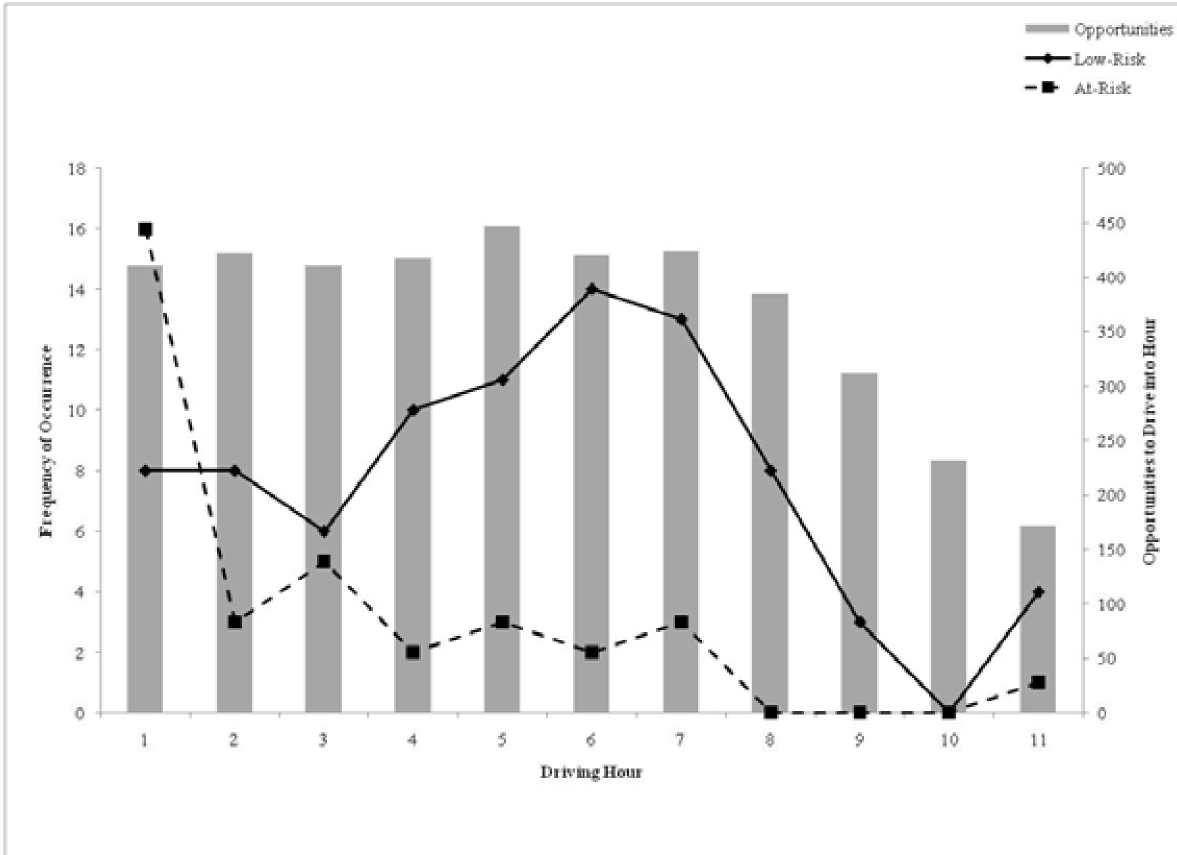


Figure 139. Frequency of DFM Warning Sound Changes by Driving Hour and Risk Level. Based on a Subset of Drivers (n = 14)

There is a downward pattern in the frequency of changes made to the DFM brightness level by at-risk drivers. For low-risk drivers there is a sharp increase in the number of changes made to the DFM brightness level during the eighth and ninth hours of driving in the shift. However, except for this singular event, the frequency of DFM brightness changes is relatively stable across groups.

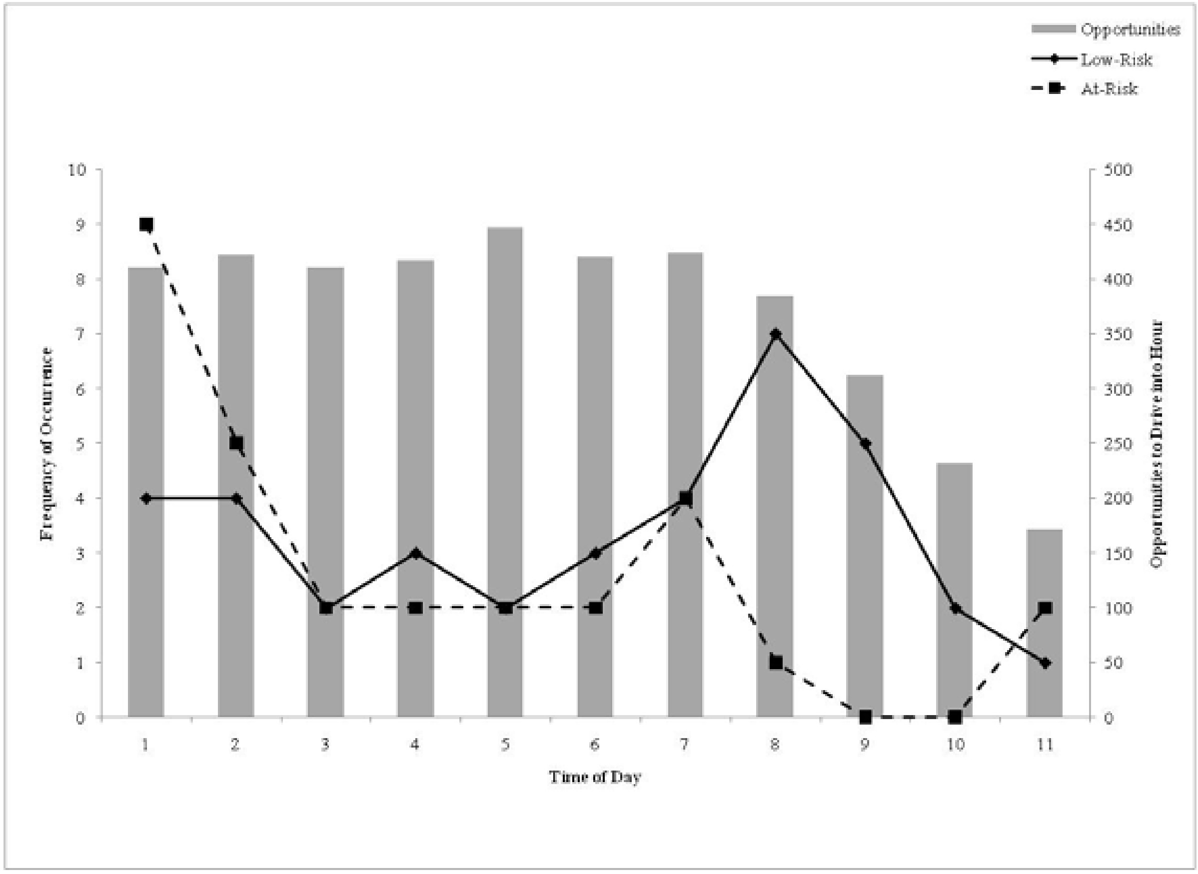


Figure 140. Frequencies of DFM Brightness Changes by Driving Hour and Risk Level. Based on a Subset of Drivers (n = 14)

For sensitivity level changes (Figure 141), warning sound changes (Figure 142), and brightness level changes (Figure 143), there does not appear to be any strong patterns in the frequency of changes made in the number of weeks since the time the DFM mode was switched to active. There are some slight downward patterns, but none are either significant or important.

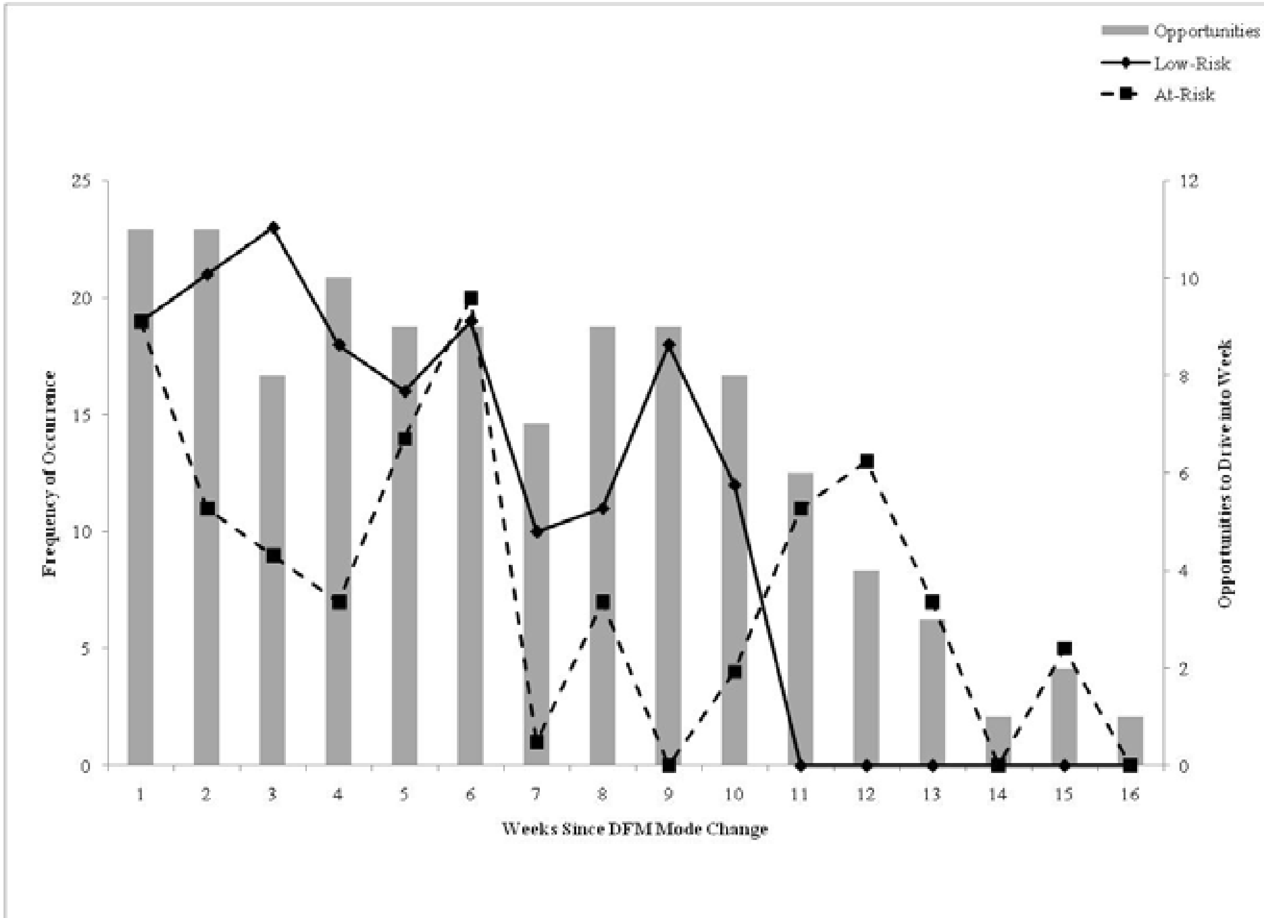


Figure 141. Frequency of DFM Sensitivity Level Changes by Weeks Since DFM Mode Change and Risk Level. Based on a Subset of Drivers (n = 14)

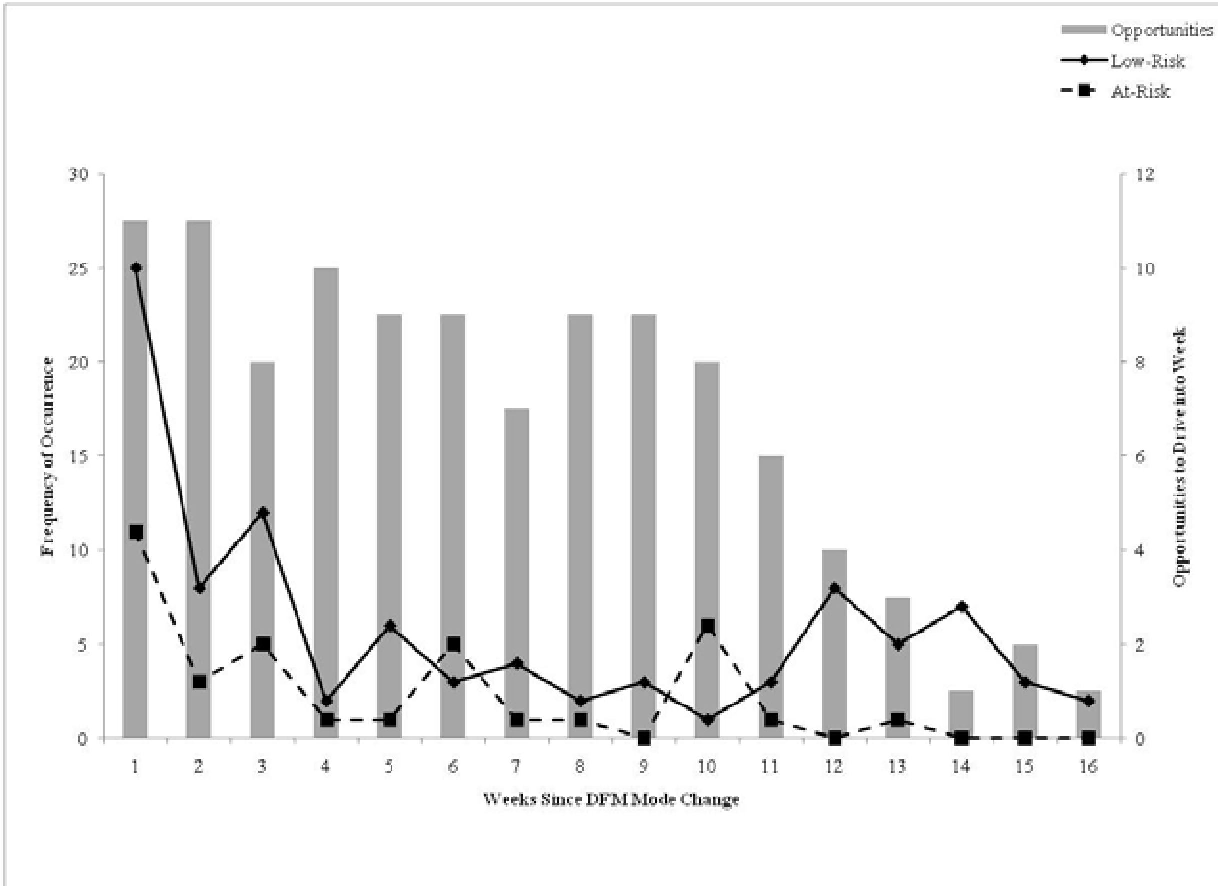


Figure 142. Frequency of DFM Warning Sound Changes by Weeks Since DFM Mode Change and Risk Level. Based on a Subset of Drivers (n = 14)

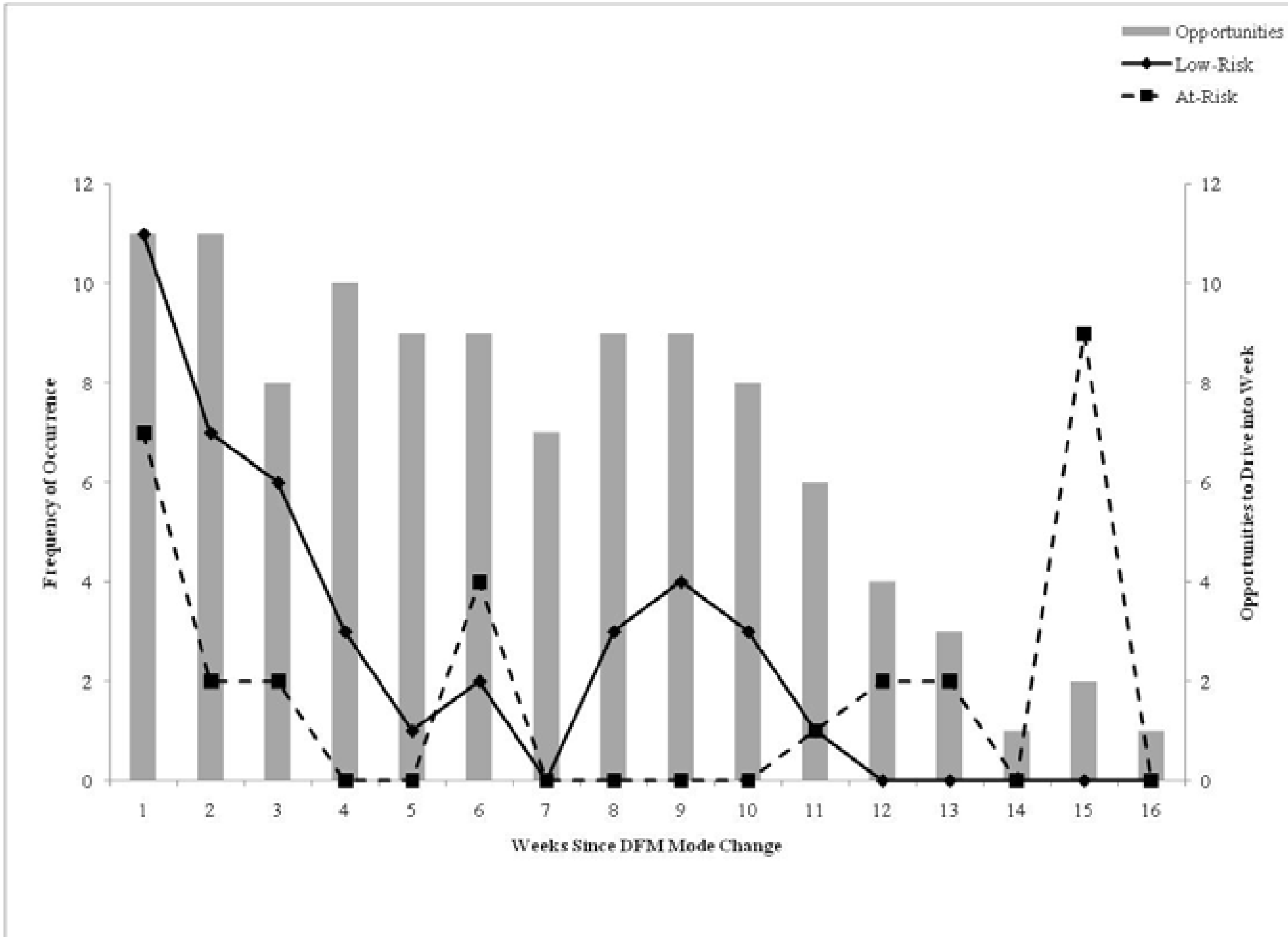


Figure 143. Frequency of DFM Brightness Level Changes by Weeks Since DFM Mode Change and Risk Level. Based on a Subset of Drivers (n = 14)

From noon to 12:59 there was a large increase in the mean duration of the DFM alerts for the low-risk drivers (Figure 144). This data point is composed of only one observation and could possibly be an outlier. For the mean alert duration for the driving hours in a shift, there was no pattern between the at-risk drivers and the low-risk drivers (Figure 145).

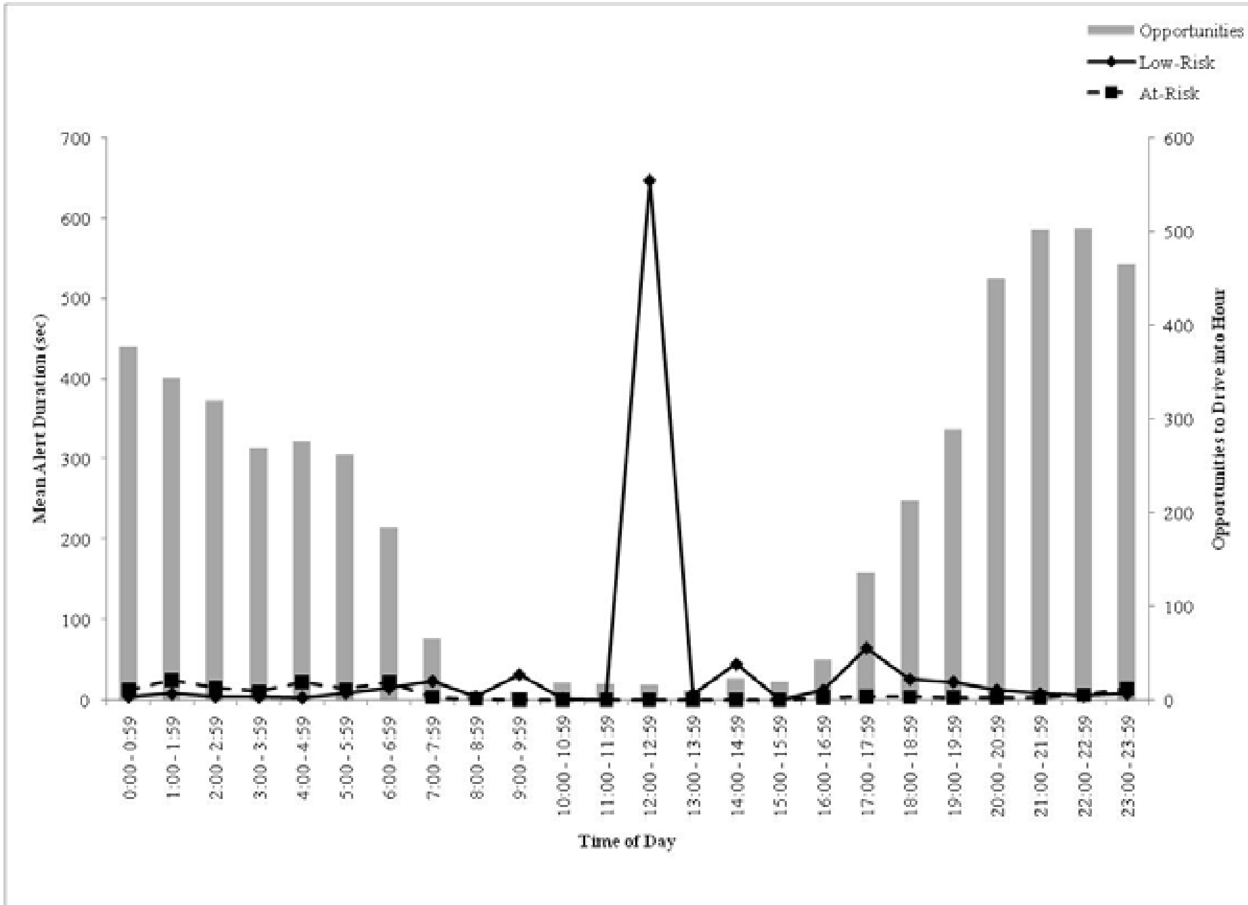


Figure 144. Mean Alert Duration by Time of Day and Risk Level.

Based on a Subset of Drivers (n = 14)

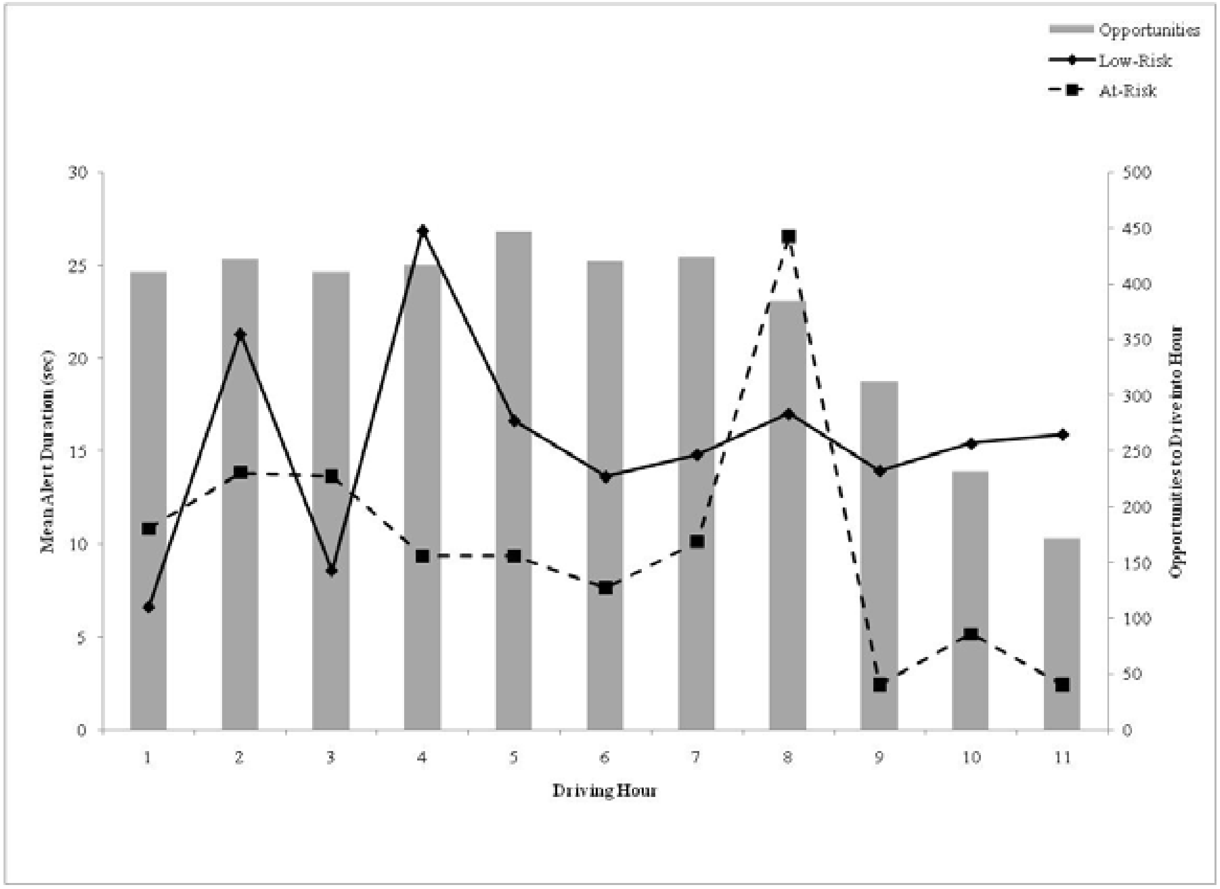


Figure 145. Mean Alert Duration by Driving Hour and Risk Level.

Based on a Subset of Drivers (n = 14)

Figure 146 shows the mean alert duration of both the low-risk and at-risk drivers in the study. For low-risk drivers in the study, there was a large increase in the mean alert duration observed during the first week after the DFM mode was switched to active. Outside of this event, the average alert duration the two groups was relatively similar.

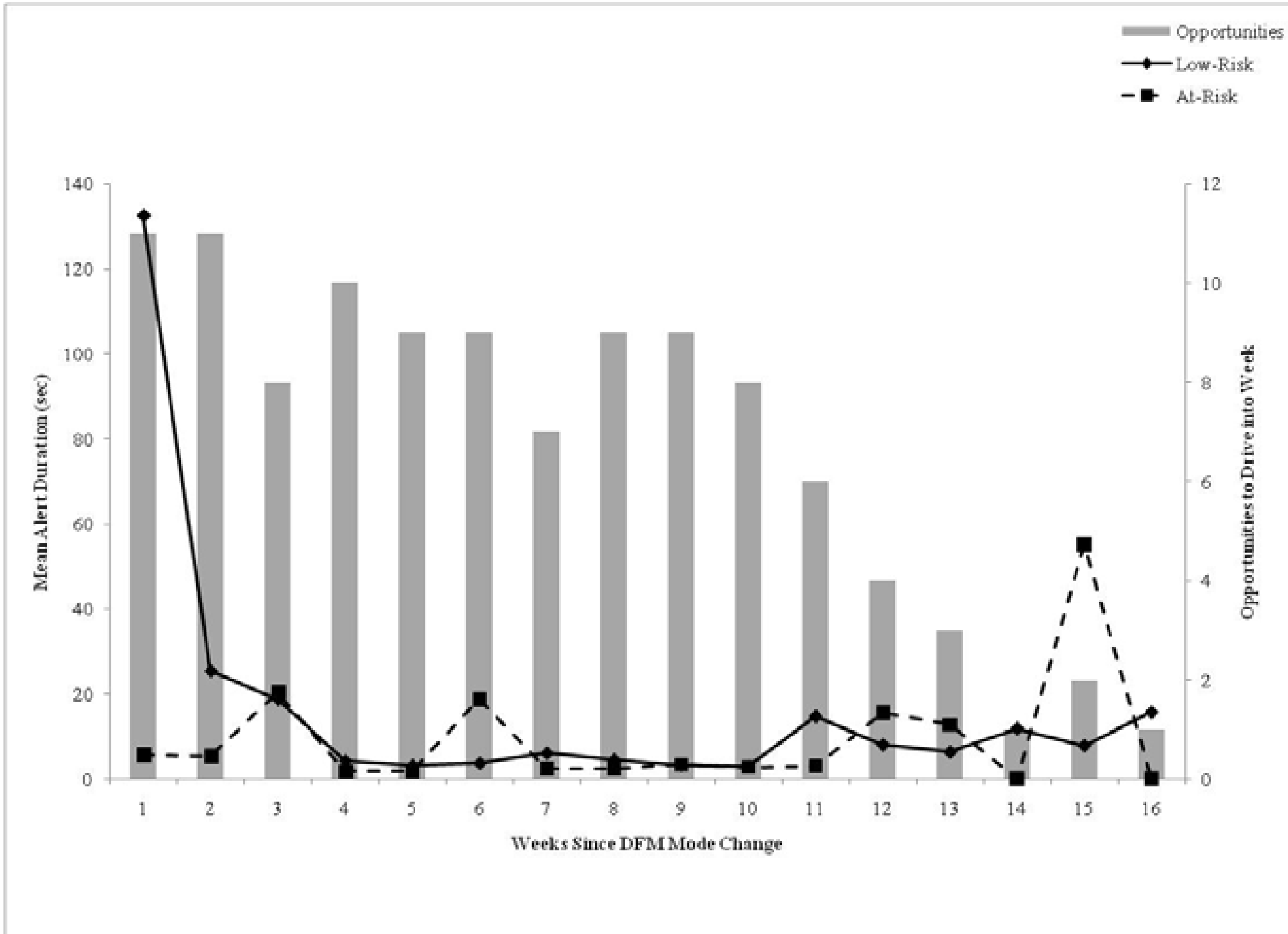


Figure 146. Mean Alert Duration by Weeks Since DFM Mode Change and Risk Level. Based on a Subset of Drivers (n = 14)

Figure 147 shows the mean duration for each brightness setting. The most frequently used settings were the two extremes for both the at-risk and low-risk drivers. Figure 148 shows the mean duration for each sensitivity level. The least frequently used setting was the high sensitivity level, and the most frequently used was the medium sensitivity level. Figure 149 shows the mean duration of each warning sound. Overwhelmingly, the most frequently used warning sound was 0. All of these findings are consistent with the findings in Research Question 4 (an examination of the human-machine interaction), where the drivers did not use the full range of settings provided. However, this was not restricted to one single risk group.

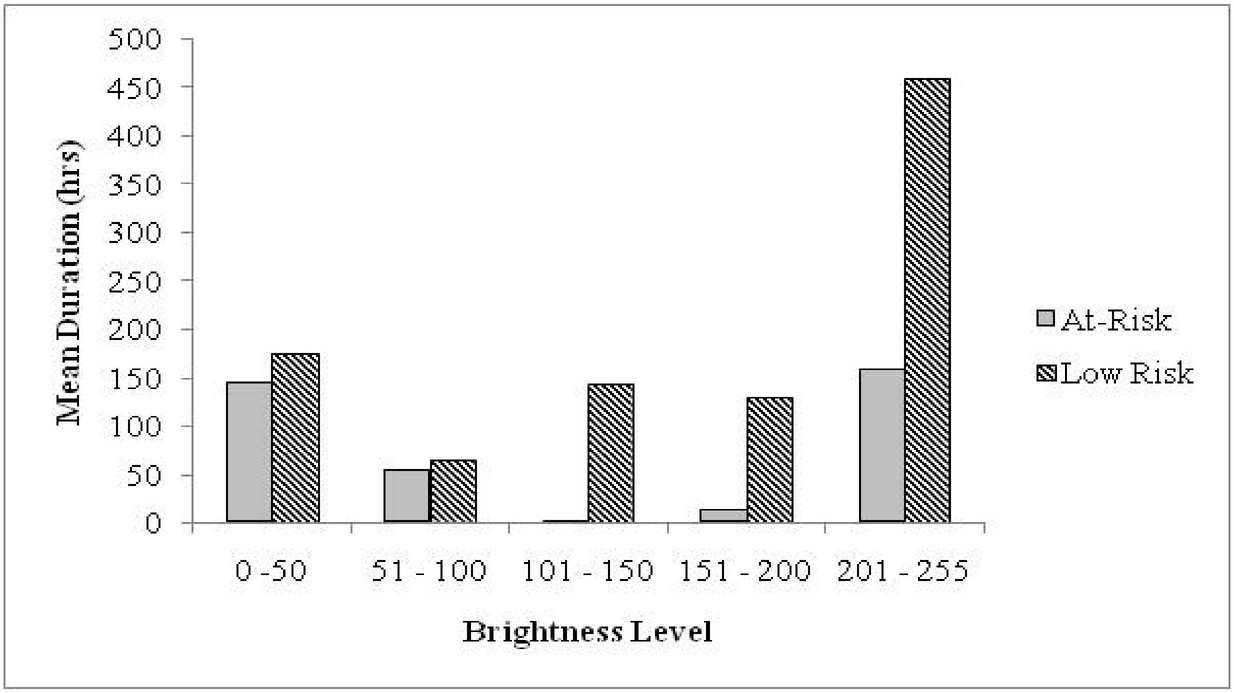


Figure 147. Mean Duration, in Hours, of Each Display Brightness Setting by Risk Level. Based on a Subset of Drivers (n = 14)

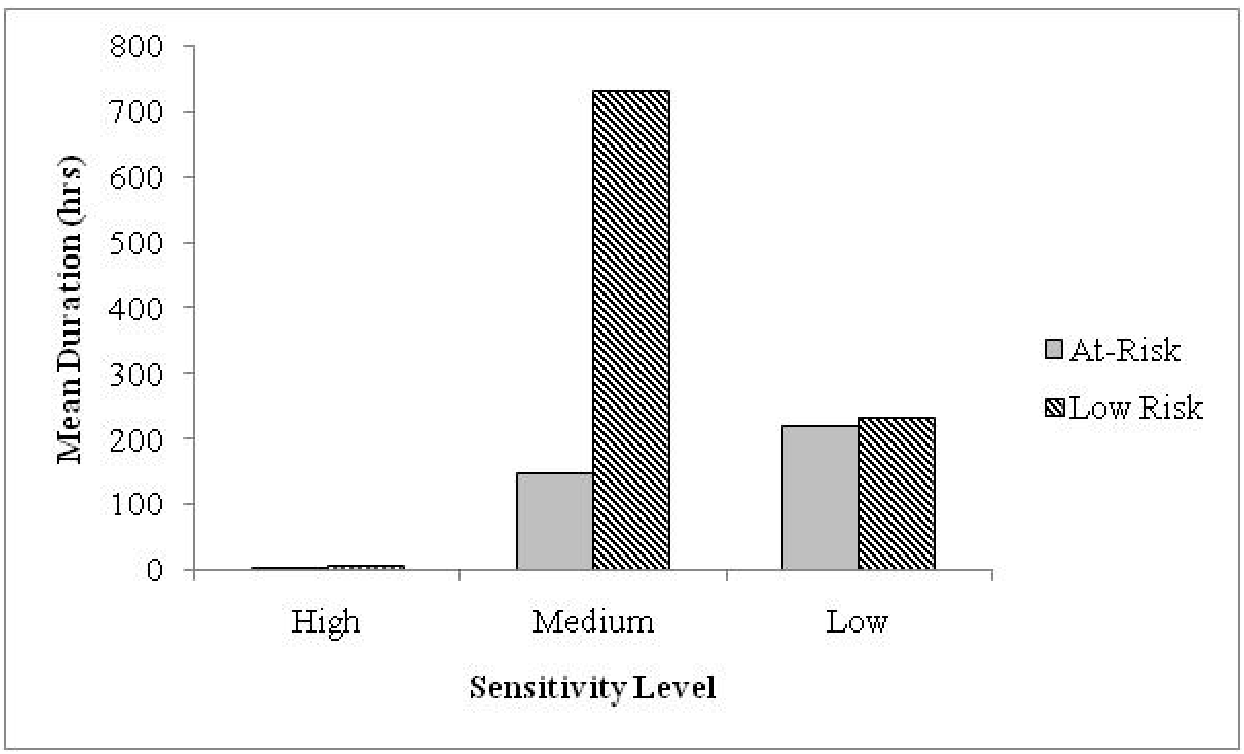


Figure 148. Mean Duration, in Hours, of Each Sensitivity Level by Risk Level. Based on a Subset of Drivers (n = 14)

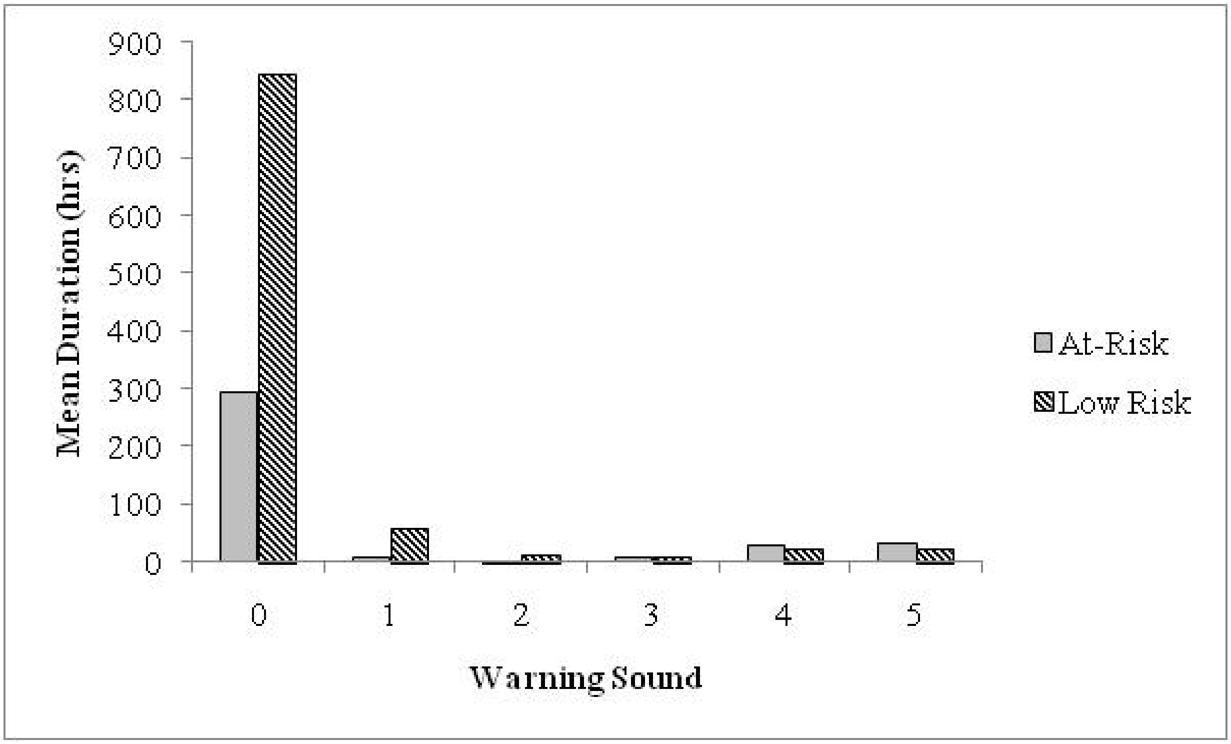


Figure 149. Mean Duration, in Hours, of Each Warning Sound by Risk Level. Based on a Subset of Drivers (n = 14)

Discussion

The at-risk drivers were on average slightly older and more experienced drivers than the low-risk drivers. However, the differences were not large, and the small number of drivers involved in the analysis precluded any estimates of significance. Haul type appears to have no effect on the risk group the drivers belonged to. The at-risk group had two line-haul drivers and five long-haul drivers, and the Low Risk group had one line-haul driver and six long-haul drivers. Those drivers considered to be at-risk did not have the same level of reduction of on-the-job-drowsiness as low-risk drivers. The low-risk drivers' valid alert rate was not as high. However, the low-risk drivers did have a lower DFM PERCLOS value and a lower manual PERCLOS value on average. When considering valid alerts, caution should be used in making any generalizations from the results due to the system's overall high false alert rate and the method used to identify at-risk and low-risk drivers. The system reliability may be biased and simply provide more reliable operation to the drivers with the most alerts per hour (i.e., with more alerts, there is a greater chance that one of them might be valid), leading them to be identified as at-risk.

Some of the sleep quality data (specifically, the number of awakenings and longest SSPs) indicated that drivers considered to have a higher risk obtained better sleep quality as compared to those drivers considered to have a lower risk. This pattern is counterintuitive as one would expect drivers with poor sleep quality to have a higher risk and be involved in SCEs due to drowsiness. However, this counterintuitive finding may be an artifact of the method used, which did not consider sleep periods under 20 min. On the contrary, drivers considered to be at lower risk obtained more SSPs longer than 20 min compared to drivers who had a higher risk. However, the small sample of drivers makes it difficult to generalize the findings to the heavy-vehicle driver population at large.

The small sample size also prevents generalizations about the at-risk drivers and the low-risk drivers with regard to SCEs. On average, there were no practical differences between the at-risk and the low-risk drivers. This also holds for the human-machine interaction. The at-risk drivers had a downward pattern in some of the graphs; however, due to the small sample size no strong conclusions can be made. It appeared as if the drivers also did not take advantage of the full range of options for the brightness level, sensitivity level, and warning sounds.

CHAPTER 5. DISCUSSION OF RESULTS

An overall view of the research objectives, measures, and results presented in earlier chapters is summarized in Table 107. Although each research question was addressed individually, Research Questions 5 and 6 required examination of a specific subset of data. Unfortunately, this approach precluded a more in-depth analysis due to small sample size and, at points, prevented statistical examinations of the data. In these cases, however, a descriptive examination was performed.

IMPACTS ON DRIVER DROWSINESS

Research Question 1 sought to examine potential impacts of DDWS device usage: fewer instances of drowsy driving, a potential decrease in the number of alerts over time, an impact on post-alert behaviors.

Evaluation of instances of driver drowsiness required an accurate assessment of DFM alert quality. This assessment was conducted through two separate, independent evaluations of driver drowsiness, which examined both the estimations of PERCLOS as determined by the DFM and PERCLOS as manually calculated by data reductionists. This approach was made necessary because of known limitations to the operation of the prototype DFM system. Specifically, the DFM's estimation of PERCLOS was artificially inflated due to instances of the DFM camera being unable to capture the driver's eyes properly or the camera going out of alignment. This was further exacerbated by instances where a large amount of driving-related scanning was required of the driver, or rapidly shifting light levels were observed. The present results provide an interesting addition to the findings of earlier studies, such as those in Wiegand, Hanowski, Olson, & Melvin (2008).

The evaluation of DFM-calculated PERCLOS did not show a dramatic reduction of drowsiness over the course of the study. However, within these results the test condition (drivers who received DFM feedback as to their drowsiness level) tended to have the lowest PERCLOS values. This was the trend for most of the days in the study, regardless of when the value was obtained. Similar findings for the test condition were present in the examination of driving hours. Drivers in the test condition tended to have a lower PERCLOS value. These findings were not directly echoed in the examination of manual PERCLOS values, which did not indicate the presence of a clear trend for the test condition. However, when examined considering the number of weeks in the study, manual PERCLOS for the test condition was lower than that of the the control condition from the eighth week on.

Overall, there is some evidence suggesting that the prototype DFM system was successful in reducing levels of driver drowsiness. However, these findings must be viewed in the strict terms of the prototype system's limited operating envelope. The present system was limited to a very small set of conditions, such as low illuminance, drivers not requiring eyeglasses, and other restrictions. Evaluation of the system in conditions falling outside of the operating envelope of the prototype system (performed using manually calculated PERCLOS) did not show a similar reduction in the level of drowsiness.

Table 107. Summary of Research Questions

Research Question	Measures	Results
1.1. Does the DDWS result in fewer episodes of drowsy driving?	DFM PERCLOS, Manual PERCLOS	A significant reduction in drowsiness over time is observed with DFM PERCLOS for the test condition (i.e., when the DDWS is providing feedback to the driver). However, based on manual PERCLOS, that was not the case. During a drowsiness level assessment of random normal driving instances (i.e., baseline events not limited by the DFM's operating envelope) there was no significant difference in the level of drowsiness over time.
1.2. Does the frequency of DDWS alerts decrease over time?	DFM Valid Alerts	No practical differences exist when the frequency of alerts is compared between the test and control conditions over time.
1.3. Do DDWS alerts have an impact on post-alert behavior?	DFM Valid Alerts, Post-Alert Behavior	No statistically significant differences were observed in post-alert behaviors following a valid DFM alert.
2.1. Does a DDWS influence drivers to get more sleep?	Actigraphy: Sleep Quantity	No statistical difference in sleep quantity was observed.
2.2. Do drivers using a DDWS achieve better quality sleep?	Actigraphy: Sleep Quality	No statistical difference in sleep quality was observed.
3.1. Does a DDWS affect involvement in safety critical events?	Event Type, Event Frequency Manual PERCLOS, Drowsiness Related Driving Behavior, SCE Rate per 100 Hours Driven	No statistical differences in SCE involvement were observed for the measures of interest between the different experimental conditions when all SCEs were considered in the analyses. The analyses were performed taking into consideration all SCEs and also only the SCEs within the DFM's operating envelope.
3.2. Does a DDWS affect involvement in at-fault safety critical events?	Event Type, Event Frequency Manual PERCLOS, Drowsiness-Related	For the SCE where the truck driver was at-fault, no statistical differences in SCE involvement were observed for the measures of interest between the different experimental conditions. The analyses were performed taking into consideration all SCEs and also only the

Research Question	Measures	Results
	Driving Behavior, SCE Rate per 100 Hours Driven	SCEs within the DFM's operating envelope.
4. How do drivers operate the DDWS in a real-world environment?	Changes in Device Settings	Most drivers did not take advantage of the full range of settings available to them, instead choosing to leave most controls at the default setting. There is a distinct difference in DFM operating patterns between those drivers with and without SCEs in their shifts. Drivers with SCEs in their shift tended to have lower levels of interaction with the system as compared to those without SCEs in their shift.
5. How did the DDWS operate for drivers who rated the system favorably in the post-study survey?	Same measures as RQ1.1-4, based on a small subset (n = 14) of drivers with strong opinions on the system	Drivers' subjective opinion of the DFM was parallel to the results of Research Questions 1 through 4, reflected in the valid alert counts, overall drowsiness, sleep quality, involvement in SCEs, and interactions with the device.
6. How did the DDWS operate for drivers who had significantly more alerts during the baseline?	Same measures as RQ1.1-4, based on a small subset (n = 14) of drivers with extremely high and low number of valid DFM alerts	Drivers considered at-risk of drowsiness related events did not show the same reduction in drowsy driving observed in low-risk drivers. The small sample size (i.e., two subset of seven drivers each) precluded further examination of SCE and device interaction data.

Alerts were only provided to drivers while the system was in active mode and after PERCLOS values reached a predetermined threshold. Although they were not provided to the driver, alerts were recorded during the dark mode by the DAS. These alerts were validated and those determined to be valid alerts were used to assess the frequency of alerts over time. Most valid alerts were in the test condition, with the highest proportion of valid alerts occurring weekdays and equally distributed across time of day. A Poisson regression was used to analyze this count data; results suggest statistically significant differences between the experimental conditions. However, the regression's estimates of the two groups differ by a fraction of an alert. Thus, there are no practical differences observed between the groups (i.e., with and without DFM feedback).

Additional caveats are necessary when examining the number of alerts provided to drivers. It is possible that any increase in valid alerts indicate that the drivers were accommodating their posture or scanning behavior in order to reduce false alarms. This is especially true if the increase is observed after DFM feedback was activated (the test condition), which is the pattern observed.

Before examining the specific post-alert behaviors, the nature of received alerts was examined. It was found that most valid alerts were received while the DFM was operating in the PERCLOS-3, or medium, sensitivity setting. Valid alerts were observed across all hours within the DFM's operating envelope. This yields further support for the validity of such alerts, as long distance driving is highly monotonous and conducive to a vigilance decrement.

A variety of behaviors were observed in drivers following the receipt of an alert from the DFM. These behaviors were organized into a taxonomy for further examination. However, no measurable impact of valid DFM alerts was present in terms of driver post-alert behavior. This was found regardless of the type of measure: the frequency or the temporal response of post-alert behaviors. Thus, it is possible that many of the post-alert behaviors observed and recorded were drivers' physical or cognitive responses to the feelings of drowsiness.

The overall findings of Research Question 1, looking for possible impacts of the prototype DFM device upon instances of on-the-job driver drowsiness, were generally inconclusive. This is most likely due to the rather large number of false alerts the DFM system was found to provide drivers. However, an examination of the false alerts determined that the majority of such alerts were provided in periods during which the driver's attention was not directed toward the road ahead. Even though the device was not always sensitive to drowsiness, it was sensitive to eyes-on-road attention. Future development of the system requires further refinement to the detection algorithms to reduce false alerts. Even with the large number of false alerts, some interesting findings and trends were present. In general, drivers in the Test Group had lower PERCLOS values overall as compared to the other experimental conditions. This would indicate that providing some type of feedback as to the driver's level of drowsiness would result in a reduction in drowsy driving.

IMPACTS ON DRIVER SLEEP HYGIENE

The hypothesized safety model (see Figure 1) indicated the possibility of an improvement in drivers' sleep hygiene. An analysis was conducted on the quantity and quality of observed sleep of drivers in the study. These analyses compared drivers' sleep quality across the various experimental conditions of the study. No significant differences were observed in either the sleep quantity or quality between experimental conditions.

Although no specific effects of the prototype DFM system on driver sleep hygiene were present, some interesting findings were revealed. Using a novel approach to the calculation of a sleep period (as compared to previous research using the same data set, see Hanowski, Hickman, Fumero, Olson, & Dingus, 2007), the typical sleep obtained by CMV drivers was found to be rather poor in general. Using the present calculation, drivers were producing some motor activity (scored as an awakening) approximately every 30 min. As further details, such as those available in a formal sleep study, were not within the scope of this study, broader conclusions about this finding are impossible. This frequency of awakening, however, indicates that a poorer quality of sleep was being obtained by the drivers. Further research into CMV driver sleep quality and other aspects affecting it is warranted. Furthermore, this study illustrates the overall difficulty in the measurement of sleep outside of laboratory settings. The ability to accurately and non-intrusively measure sleep, as well as the methods used to analyze such findings, is still in development.

SAFETY CRITICAL EVENT INVOLVEMENT

The possible safety benefits of the prototype DFM system were explored by examining driver involvement in safety critical events. For the purposes of this study, SCEs were defined as crashes, near-crashes, and crash relevant conflicts. Specific analyses of all SCEs, those SCEs occurring within the DFM operating envelope, and the SCEs that occurred when the truck driver was at-fault were examined. However, it is important to note that SCEs did not only occur within the operational envelope of the prototype DFM system (specifically the speed and illumination requirements of the system). Using DFM PERCLOS as the measure to identify drowsiness in SCEs would exclude a significant portion of all SCEs, thus skewing any subsequent analysis. Therefore, all SCEs were examined using the manual PERCLOS measures.

The resulting data set included a total of 1,124 SCEs. Of these, there were 28 crashes, 112 near-crashes, and 984 crash relevant conflicts. Of the total 1,124 SCEs, 221 occurred within the operating envelope of the DFM. Statistical tests considering all the SCEs showed no statistically significant difference between the Test and Control Groups. Moreover, no statistically significant differences in the SCE distribution were observed for drivers during the baseline and the test conditions.

Additional analyses were conducted to further examine connections between SCEs and drowsiness. It was found that over 60 percent of the SCEs had a manual PERCLOS value below 12, indicating that most SCEs were not occurring in the study's definition of a state of drowsiness. A second behavioral analysis was conducted to characterize each of the SCEs (see

Hickman et al., 2005). Drowsiness-related behaviors were observed in a total of 143 SCEs (13%). However, no statistical differences between the experimental conditions were present.

Although drowsy driving plays a large role in crashes on the highway system, in some cases up to 20 percent of all SCE (Hanowski, Wierwille, & Dingus, 2003), they do not represent the majority of these events. Additionally, crashes are relatively rare events as evidenced by the 28 observed crashes in this study. Therefore, the results of the present work are not entirely unexpected. These findings do not indicate any lessening of the magnitude of the drowsy driving problem, they only illustrate the difficulty in studying a serious but unpredictable event.

HUMAN-MACHINE INTERACTIONS

Several interesting findings were present when drivers' interaction with the prototype DFM was examined. The majority of drivers did not choose to adjust settings on the DFM system, instead choosing to leave the DFM set to the defaults of active mode. This pattern was observed in examinations of sensitivity levels, warning sounds, and display brightness.

When drivers did choose to adjust the settings of the DFM, binomial patterns were observed. This was especially salient in the display brightness settings, where brightness settings tended to be split between the two limits of adjustment. Overall, this suggests that drivers were not using the full range of adjustment, did not feel that it was needed, or were using the adjustment for a secondary (non-intended) purpose such as masking DFM alerts.

There were marked differences in driver-device interactions when drivers were compared based on the presence of an SCE within their shift. Those drivers with an SCE occurring in their shift tended to have a lower level of interaction with the system across sensitivity, warning sound, and display brightness adjustments. One possible explanation for this is that drivers who were operating under some existing level of stress were limiting interactions with the device in order to dedicate more resources to the driving task. Another possibility is that there is some link between drivers with an increased risk of SCEs and willingness to interact with the system; however, evaluating that possibility was not within the scope of the present study.

These results highlight the difficulty in designing user interfaces, especially for complex, novel systems in complex environments. The finding that drivers did not use the full range of adjustments available is evidence of this. In addition to further refinement of the detection algorithms to reduce false alerts, further refinement of the human-machine interface is needed. The challenges in refining this prototype device's interface are the same faced by any number of secondary systems which are increasingly common in the vehicle. A better understanding of the optimal interface is of great importance.

FAVORING DRIVERS

In order to determine if there were operational differences in the DDWS between drivers who rated the system positively and those who rated the system negatively, new subsets of the larger data sets were created. These subsets were created by identifying drivers who rated the prototype system in an extremely favoring (high satisfaction, high usefulness) or disfavoring (low satisfaction, low usefulness) manner. These identified drivers were then screened to ensure they were part of the original 96 drivers of the full study data set. The resultant 14 drivers' data from the data sets for Research Questions 1 through 4 were used to generate the subset for Research Question 5. The small sample size of this group precluded examination using traditional statistical testing. Instead, the analyses conducted were descriptive in nature.

On average, drivers with a disfavoring opinion of the system were older and more experienced drivers, and most were line-haul drivers. This is in sharp contrast to the drivers with favoring opinions of the system, all of whom were long-haul drivers. The disfavoring drivers tended to have fewer valid alerts when compared to the drivers with favoring opinions of the system. Although the limited sample size precludes further examination, there were some trends indicating an increase in safety benefits for drivers favoring the system over the ones with disfavoring opinions. Further differences between the drivers were observed in sleep quality. Drivers with favoring opinions tended to have better sleep quality than those with disfavoring ratings of the system. This serves as a possible explanation for the differences observed in valid alert frequency between the two groups, as it is possible that the favoring raters had a predisposition to lower drowsiness than the disfavoring raters.

There appeared to be a decrease in the occurrence of drowsiness-related SCEs for favoring rating drivers. However, due to the small sample size of drowsiness-related SCEs, this could only be examined as a general trend. Further analysis would require a larger and, ideally, balanced sample.

Interestingly, the drivers with disfavoring ratings of the DFM tended to make more adjustments to the system and, as would be expected, did not respond the DFM alerts as quickly as the favoring rating drivers. Several reasons for the greater levels of device interaction are possible. As the system worked less reliably for the disfavoring-rating drivers, these drivers may have made more adjustments to the DFM and, upon realizing the low reliability, began to ignore the warnings.

AT-RISK DRIVERS

Similar to the process of determining driver subjective favorability, a limited number of drivers were identified as being either at-risk or at low-risk of driver drowsiness. The two groups were created by selecting the seven drivers with the greatest number of valid alerts and the seven drivers with the least number of valid alerts. These drivers were compared with respect to the questions outlined in Research Questions 1 through 4, in addition to driver and operational characteristics and the overall reliability of the prototype DFM.

On average, at-risk drivers were slightly older and more experienced drivers compared to low-risk drivers. No difference in haul type was found between the two drowsiness risk groups. Interestingly, drivers considered to be at-risk did not show the reduction of on-the-job-drowsiness displayed by low-risk drivers. Likewise, low-risk drivers did not have as high a valid alert rate. They did, however, have a lower DFM PERCLOS value and a lower manual

PERCLOS value. Although this initially suggests that some distinct differences between the groups are present, restraint should be used in generalizing from these results due to the prototype system's overall high false alert rate and the method of identifying at-risk and low-risk drivers. It is possible that system reliability was biased and simply provided more reliable operation to the drivers with the most alerts per hour, which in turn led these drivers to be identified as at-risk.

Interesting results are present in the sleep quality data. Drivers with a higher risk of drowsy driving tended to have better sleep quality when compared with drivers at lower risk. Although this is an interesting and somewhat counterintuitive finding, it should be noted that these results could be an artifact of the sleep scoring method (which did not consider sleep periods under 20 min). This is further supported by the fact that drivers considered to be at lower risk obtained more SSPs longer than 20 min as compared to drivers at higher risk.

When differences between at-risk and low-risk drivers were examined in terms of SCE, no significant differences were found. Likewise, no significant differences were present for the two groups in terms of their interaction with the DFM prototype. Although at-risk drivers had a downward trend in some aspects of their interaction with the DFM, the small sample size precludes any broad conclusions.

PERFORMANCE AND CAPABILITIES

Calibration, Maintenance, Adjustment, and Reliability

One of the problems revealed by field testing the DFM was its handling of situations where the eyes were not visible to the system, yet the driver was not truly drowsy. This produces a false alarm, which could have a detrimental effect on the driver's acceptance of the system (Bliss & Acton, 2003). False alarms have been classified into three major subtypes: traditional, nuisance, and inopportune (Xiao & Seagull, 1999). Traditional false alarms are those in which the device or sensor is improperly calibrated or otherwise faulty in some way. Nuisance alarms are those in which the alarm activates in the incorrect context. The inopportune alarm is an otherwise valid alarm that activates before the danger, in effect turning into a warning.

During the alert validation process (see Appendix A), data reductionists were asked to code any apparent causes of the invalid alerts. There were many categories available to the reductionists; however, if no category clearly described the apparent cause of the invalid alert, the reductionist could enter a description of the apparent cause. These were entered as free text (string variables) by the reductionists and saved to the data set for that driver.

Of the 14,665 invalid alerts recorded for 90 drivers, the vast majority were not attributed to one of the predefined categories. In order to determine if any discernable pattern was present, the reductionist-entered text was converted to categories based on keywords in the description (e.g., eyes not found, driver moving). These were then examined in an attempt to reduce the overall number of unexplained invalid alerts.

The existing categories of invalid alert causal factors were placed into an initial taxonomy. This consisted of Driver Distraction and Outside of the Device Specifications. The category of Driver Distraction was based primarily on current knowledge of driver distractions. Therefore, any action that removed the driver's attention from the task of driving was included in this category.

The second major category, Outside of the Device Specifications, was created to capture driver actions that were not specifically accounted for by the prototype DFM system, yet would cause an alert to register (not related to Driver Distraction). Because the prototype DFM system depended on a restricted set of circumstances (i.e., no eyeglasses, specific levels of illumination) and on the driver's visual fixations being in the forward direction, any other driver action not accounted for by driver distraction was included in this category. The resulting process left a significant number of invalid alerts unattributable to a specific category. Based on the processing of the reductionist comments in the data files, the additional categories Device Failure, Multiple Actions, and Undetermined were added.

The category Multiple Actions was created to capture those situations where multiple apparent causes for the failure of the DFM were present. These were commonly situations such as driver scanning combined with using a CB radio. Invalid alerts where the driver's eyes remained in a forward fixation and no other apparent cause was present were classified as Device Failures. At the conclusion of this process, 32 of the initial 14,665 alerts remained unclassified as no clear reason for the invalid alert was observed. No further determination was possible for these 32 invalid alerts.

Sixteen major factors causing invalid alerts were indentified in the final analysis. Their frequency of observation and contributing percentages are presented in Table 108. The most frequently observed cause of an invalid alert was a device failure. This situation was most often indentified by data reductionists as the device having a clear and unobstructed view of the driver's eyes and face yet not being able to locate the driver's eyes. The next two most frequently observed causes were necessary driver actions, such as mirror and roadway scanning. Fluctuations in illuminance level was the fourth major contributor.

Table 108. Type and Frequency of Invalid Alert Causes

Type	Description	Frequency	Percent
Device Failure	Device not working	5,780	39.41
Driver Distraction	Cell phone	349	2.38
Driver Distraction	Reading	267	1.82
Driver Distraction	Eating/drinking	176	1.20
Driver Distraction	Tobacco use	159	1.08
Driver Distraction	CB radio	116	0.79
Driver Distraction	Personal hygiene	112	0.76
Driver Distraction	Other electronic device	40	0.27
Multiple Actions	Multiple actions	358	2.44
Outside Specifications	Looking away	3,015	20.56
Outside Specifications	Driving-related scanning	2,137	14.57
Outside Specifications	Outside light level	1,275	8.60
Outside Specifications	Eyes obstructed	324	2.21
Outside Specifications	Driver shifting position	229	1.56
Outside Specifications	Driver movement	209	1.43
Outside Specifications	Glasses/sunglasses	87	0.59
Undetermined	Undetermined	32	0.22
Total		**14,665**	**100.00**

An interesting pattern emerges when the frequency of invalid alert causal factor types are examined (Table 109). Nearly half of invalid alert causes may be classified as events outside of the prototype DFM's specifications.

Table 109. Type and Frequency of Invalid Alert Causes

Type	Frequency	Percent
Outside Device Specifications	7,276	49.61
Device Failure	5,780	39.41
Driver Distraction	1,219	8.31
Multiple Actions	358	2.44
Undetermined	32	0.22
Total	**14,665**	**100.00**

As mentioned in the description of the device above, the DFM specfically monitors for and depends on the driver's face being in a forward-oriented position. Violations of this assumption often resulted in an invalid alert being recorded by the prototype DFM. The majority of the invalid alert types may be classified as either situations outside of the prototype DFM's specifications or as a simple device failure (Figure 150).

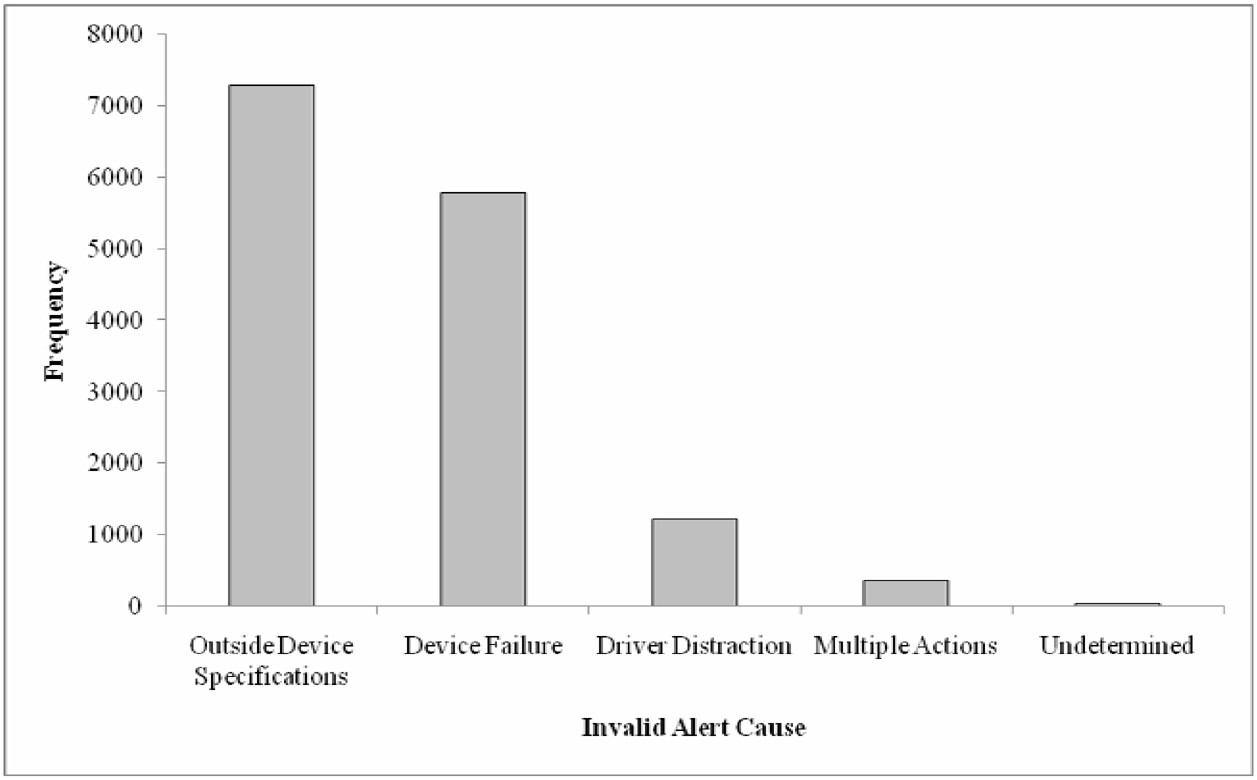


Figure 150. Types of Invalid Alert Causes

Between the two common barriers to any system (i.e., technological and theoretical barriers), the majority of invalid alerts observed in the current study may be considered technological in nature. The possibility of refining the device's eye detection algorithm and modifying the manner in which driver-related scanning is handled should greatly reduce the number of invalid alerts produced by the DFM and greatly increase the utility of future DDWS implementations.

Although many of the false alarms observed in the present experiment are likely due to sensor or algorithm calibration issues (such as lighting issues or drivers' head movement causing traditional false alarms), it is entirely possible that some of the false alarms observed during the experiment's duration were actually of the inopportune variety. That is, they were alerting the driver to a real scenario of likely drowsiness, but the temporal presentation of the alarm was suboptimal. Whether this is an issue of alert algorithm calibration or some other form of adjustment necessary to the DFM is unclear. Any future DDWS must contend with the need for an optimized algorithm for the detection of drowsiness onset in order to minimize the risks to the driver, such as distraction and habituation, that arise from the various forms of false alarms.

Overall, minimal real-world calibration of the system was required following the initial installation. Any adjustments that were necessary were due to repair operations for the device. Problems occurred such as buttons or screws coming loose, parts coming unglued, and switches failing. These are not entirely unexpected in a prototype system, and with continued development they should be eliminated from any production system.

Technical Challenges and Future Designs

The prototype DFM used in the present study was based on measuring PERCLOS, a measure of slow eye-closure. The system operates by monitoring the driver's eyes and determining those periods where the eyelids are between 80 and 100 percent closed. This is done through the use of a near-infrared camera positioned within approximately 20° of the centerline of the driver. Several adjustments are available to the driver so that he or she may adjust the PERCLOS threshold for initial alerts, the display brightness, and the alert sound and volume.

In the current trial, the DFM was frequently observed to respond to situations where the driver was not exhibiting true symptoms of drowsiness (i.e., false alarms). These situations present a problem for driver acceptance of the system, even though most drivers seemed to believe the system was overall beneficial and provided a useful and needed source of information for their safety. When the invalid alerts produced by the DFM were examined, the majority of instances were outside of the true operating envelope of the prototype system. Another major portion of the invalid alerts were generated by driver actions that may be considered distracted, meaning the system was alerting the driver to a loss of visual roadway information (as would occur in driver drowsiness). Thus, many of the invalid alerts may be considered nuisance alarms; these alerts were valid for a different context than the one intended.

One possibility which must be noted is the act of responding to the alert produced by the DFM acting as a psychophysiological arousal mechanism. Some evidence (Ayoob, Grace, & Steinfeld, 2005) of a DDWS acting as a driver arousal mechanism exists. If the DDWS is acting as a mechanism for awakening a drowsy driver, then it may be considered a countermeasure as well as an alert. Two complementary mechanisms for a DDWS countermeasure to drowsiness may exist. The act of detecting the alert could possibly provide some degree of cognitive arousal, and the act of reaching forward to respond to the alert using the warning response button could provide an increase in the level of physiological arousal. These concepts are not mutually exclusive; instead they are complementary toward raising the driver's level of arousal among the physical and cognitive dimensions. Further investigation into the nature of alerting a drowsy driver, and what impacts this process has on their level of arousal, would likely prove to be of some benefit to designers of such systems.

REFERENCES

Advocates for Highway and Auto Safety. (2001). *Fact Sheet: Truck Driver Fatigue*. Retrieved July 21, 2008, from http://saferoads.org/issues/fs-truckdriverfatigue.htm

Amit, P., Boyle, L.N., Boer, E.R., Tippin, J., & Rizzo, M. (2005). *Steering Entropy Changes as a Function of Microsleeps*. Driving Assessment 2005: 3rd International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Rockport, ME.

Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W., & Pollak, C.P. (2003). The role of actigraphy in the study of sleep and circadian rhythms. *Sleep, 26*(3), 51.

Andreassi, J. (2000). *Psychophysiology: Human Behavior and Physiological Response*. London: Lawrence Erlbaum Associates.

Artaud, T., Planque, S., Lavergne, C., Cara, H., de Lepine, P., Tarrière, C., & Gueguen, B.M. (1994). *An On-Board System for Detecting Lapses of Alertness in Car Driving*. 14th E.S.V. Conference, Session 2-Intelligent Vehicle Highway System and Human Factors: Munich, Germany.

Ayoob, E.M., Grace, R., & Steinfeld, A. (2005). *Driver-Vehicle-Interface (DVI) Development of a Drowsy Driver Detection and Warning System for Commercial Vehicles.* Technical Report No. CMU-RI-TR-05-46. Robotics Institute, Carnegie Mellon University.

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

Blanco, M., Hickman, J.S. Olson, R.L., Bocanegra, J.L., Hanowski, R.J., Nakata, A., Greening, M., Madison, P., Holbrook, G.T., & Bowman, D. (2008). *Investigating Critical Incidents, Driver Restart Period, Sleep Quantity, and Crash Countermeasures in Commercial Operations using Naturalistic Data Collection: Final Report*. (Contract No. DTFH61-01-C-00049, Task Order # 23). Blacksburg, VA: Virginia Tech Transportation Institute.

Bliss, J.P., & Acton, S.A. (2003). Alarm mistrust in automobiles: How collision alarm reliability affects driving. *Applied Ergonomics, 34*, 499-509.

Boer, E.R. (2000). Behavioral entropy as an index of workload. *Proceedings of the IEA 2000/HFES 2000 Congress.*

Boer, E.R., Rakauskas, M.E., Ward, N.J., & Goodrich, M.A. (2005). *Steering Entropy Revisited*. Driving Assessment 2005: 3rd International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Rockport, ME.

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

Boyle, L., Hill, J., Tippin, J., Faber, K., & Rizzo, M. (2007). The effects of sleep apnea on heart rate variability in a simulator. *Proceedings of the Fourth International Driving*

Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, 306-313.

Brookhuis, K. (1995). Driver impairment monitoring by physiological measures. In L. Hartley (Ed.), *Driver Impairment, Driver Fatigue and Driver Simulation*. London: Taylor & Francis, 181-188.

Brookhuis, K.A., & de Waard, D., (1993). The use of psychophysiology to assess driver status. *Ergonomics, 36*, 1,099-1,110.

Brookhuis, K.A., Louwerens, J.W. & O'Hanlon, J.F. (1986). EEG energy-density spectra and driving performance under the influence of some antidepressant drugs. In J.F. O'Hanlon & J.J. de Gier (Eds.), *Drugs and Driving*, London, Philadelphia: Taylor & Francis, 213-221.

Brown, I.D. (1994). Driver fatigue. *Human Factors, 36*(2), 298-314.

Buxton, S. & Hartley, L. (2001, August). Napping to prevent fatigue: A policy. Paper presented at the *Insurance Commission of Western Australia: 2001 Conference on Road Safety*, Perth, Australia.

Byeon, M.K., Han, S.W., Min, H.K., Wo, Y.S., Park, Y.B., & Huh, W. (2006). A study of HRV analysis to detect drowsiness states of drivers. *Proceedings of the Fourth International Conference on Biomedical Engineering*, Innsbruck, Austria.

Cameron, A.C., & Trivedi, P.K. (1986). Econometric models based on count data: Comparisons and applications of some estimators and tests. *Journal of Applied Economics, 1*, 29-53.

Cole, R.J., Kripke, D.F., Gruen, W., Mullaney, D.J., & Gillin, J.C. (1992). Automatic sleep/wake identification from wrist activity. *Sleep, 15*(5), 461-469.

Comsis Corporation (1996). *Preliminary Human Factors Guidelines for Crash Avoidance Warning Devices*. NHTSA Project No. DTNH22-91-C-07004.

Desmond, P.A., Matthews, G., & Hancock, P.A. (1997). Dimensions of subjective fatigue states in driving. In C. Mercier-Guyon (Ed.), *Proceedings of the 14th International Conference on Alcohol, Drugs, and Traffic Safety*.

de'Sperati, C. (2005). Smooth pursuit-like eye movements during mental extrapolation of motion: The facilitatory effect of drowsiness. *Cognitive Brain Research, 25*(1), 328-338.

Dinges, D.F. (1995). An overview of sleepiness and accidents. *Journal of Sleep Research* (4), 4-14.

Dinges, D.F., & Grace, R. (1998). *PERCLOS: A Valid Physiological Measure of Alertness as Assessed by Psychomotor Vigilance*. Federal Highway Administration Publication No. FHWA-MCRT-98-006.

Dinges, D.F., Maislin, G., Brewster, R.M., Krueger, G.P., & Carroll, R.J. (2005). Pilot test of fatigue management technologies. *Transportation Research Record, 1922*, 175-182.

Dingus, T.A., Hardee, H.L., & Wierwille, W.W. (1987). Development of models for on-board detection of driver impairment. *Accident Analysis and Prevention, 19*(4), 271-283.

Dinges, D.F., & Mallis, M. (1998). Managing fatigue by drowsiness detection: Can technological promises be realized? In L. Hartley (Ed.), *Managing Fatigue in Transportation*, London: Pergamon Press, 209-229.

Duley, A. (2006). *Affective Information Processing and Anxiety: Attentional Bias and Short Lead Interval Startle Modification*. Unpublished doctoral dissertation. Gainesvill, FL: University of Florida.

Erwin, C.W. (1976). *Studies of Drowsiness: Final Report*. Durham, NC: Duke University.

Erwin, C.W. (1973). *Psychophysiologic Indices of Drowsiness*. Detroit, MI: International Automotive Engineering Congress.

Erwin, C.W., Weiner, E.L., Hartwell, J.W., Truscott, T.R., & Linnoila, M.I. (1975). *Alcohol Effect on Vigilance Performance*. Society of Automotive Engineers Paper No. 750880, presented at the Automotive Engineering Meeting, Detroit, MI.

Evans, L. (2004). *Traffic Safety*. Bloomfield Hills, MI: Science Serving Society.

Fairclough, S.H. (1997). Monitoring driver drowsiness via driving performance. In I. Noy (Ed.), *Ergonomics and Safety of Intelligent Driver Interfaces*. New Jersey: Lawrence Erlbaum Associates, 363-379.

Federal Motor Carrier Safety Administration (2005). The Revised Hours of Service Regulations – 2005. Retrieved July 22, 2008 from http://www.fmcsa.dot.gov/rules-regulations/truck/driver/hos/brochure2005.htm

Filiatrault, D.D., Cooper, P.J., King, D.J., Siegmund, G.P., & Wong, P.K.H. (1996). Efficiency of vehicle-based data to predict lane departure arising from the loss of alertness due to fatigue. *Proceedings of the 40th Annual Association for the Advancement of Automotive Medicine*, Vancouver, BC.

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

Grace, R., & Benjamin, A.L. (1999). Application of a Heavy Vehicle Drowsy Driver Detection System. *Proceedings of the SAE International Truck & Bus Meeting and Exposition*, November 15-17, 1999.

Grace, R., Byrne, V.E., Bierman, D.M., Legrand, J-M., Gricourt, D., France, E., Davis, R.K., Staszewski, J.J., & Carnahan, B. (1999). A drowsy driver detection system for heavy vehicles. *Technical Conference Proceedings, Ocular Measures of Driver Alertness*, Herndon, VA.

Grace, R., & Stewart, S. (2001). *Drowsy Driver Monitor and Warning System*. Paper presented at the International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Aspen, CO.

Greco, K.E., Deaton, C., Kutner, M., Schnelle, J.F., & Ouslander, J.G. (2004). Psychoactive medications and actigraphically scored sleep quality in frail nursing home patients. *Journal of the American Medical Directors Association, 5*(4), 223-227.

Hanowski, R.J., Blanco, M., Nakata, A., Hickman, J.S., Schaudt, W.A., Fumero, M.C., Olson, R.L., Jermeland, J., Greening, M., Holbrook, G.T., Knipling, R.R., & Madison, P. (2008). *The Drowsy Driver Warning System Field Operational Test: Data Collection Final Report*. Contract No. DTNH22-00-C-07007, Task Order 14. Blacksburg, VA: Virginia Tech Transportation Institute.

Hanowski, R.J., Hickman, J., Fumero, M.C., Olson, R.L., & Dingus, T.A. (2007). The sleep of commercial vehicle drivers under the 2003 hours-of-service regulations. *Accident Analysis and Prevention, 39*(6), 1140-1145.

Hanowski, R.J., Olson, R.L., Bocanegra, J., & Hickman, J.S. (2007). *Analysis of Risk as a Function of Driving-Hour: Assessment of Driving Hours 1 Through 11*. Unpublished Research Report. Blacksburg, VA: Virginia Tech Transportation Institute.

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

Heitmann, A., & Guttkuhn, R. (2001). Technologies for the monitoring and prevention of driver fatigue. *Proceedings of the International Driving Symposium of Human Factors in Driver Assessment, Training and Vehicle Design*, 81-86.

Hickman, J.S., Knipling, R.R., Olson, R.L., Fumero, M.C., Blanco, M., & Hanowski, R.J. (2005). *Preliminary Analysis of Data Collected in the Drowsy Driver Warning System Field Operational Test: Task 5, Preliminary Analysis of Drowsy Driver Warning System Field Operational Test Data*. Report DTNH22-00-C-07007, Task Order 21. Blacksburg, VA: VTTI.

Hollander, M., & Wolfe, D.A. (1999). *Nonparametric Statistical Analyses* (2nd Ed.). New York: John Wiley & Sons.

Hyoki, K., Shigeta, M., Tsuno, N., Kawamuro, Y., & Kinoshita, T. (1998). Quantitative electro-oculography and electroencephalography as indices of alertness. *Electroencephalography and Clinical Neurophysiology*, 106, 213-219.

Ingre, M., Åkerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2006). Subjective sleepiness, simulated driving performance and blink duration: examining individual differences. *Journal of Sleep Research, 15*(1), 47-53.

Jean-Louis, G., Kripke, D.F., Mason, W.J., Elliott, J.A., & Youngstedt, S.D. (2001). Sleep estimation from wrist movement quantified by different actigraphic modalities. *Journal of Neuroscience Methods, 105*(2), 185-191.

Johns, M., Tucker, A., & Chapman, R. (2005). *A New Method for Monitoring the Drowsiness of Drivers*. Paper presented at the 2005 International Conference on Fatigue in Transportation Operations, Seattle, WA.

Kloss, J.D., Szuba, M.P., & Dinges, D.F. (2003). Sleep loss and sleepiness: Physiological and neurobehavioral effects. In K.L. Davis, D. Charney, J.T. Coyle, & C. Nemeroff (Eds.), *Neuropsychopharmacology: The Fifth Generation of Progress*, Baltimore: Lippincott, Williams and Wilkins, 1,895-1,905.

Knipling, R.R. (2005). Evidence and dimensions of commercial driver differential crash risk. In *Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, 2-8. Iowa City, IA: Public Policy Center

Knipling, R.R., & Wierwille, W.W. (1994). Vehicle-based drowsy driver detection: Current status and future prospects. *Proceedings of the IVHS America 1994 Annual Meeting*, pp 245-256, Atlanta, GA.

Lal, S.K.L., & Craig, A. (2000). Psychophysiological effects associated with drowsiness: Driver fatigue and electroencephalography. *International Journal of Psychophysiology*, 35, 39.

Lal, S.K.L., Craig, A., Boord, P., Kirkup, L., & Nguyen, H. (2003). Development of an algorithm for an EEG-based driver fatigue countermeasure. *Journal of Safety Research, 34*(3), 321-328.

Lancaster, R., & Ward, R. (2002). *The Contribution of Individual Factors to Driving Behaviour: Implications for Managing Work-Related Road Safety* (Report No. 020). Entec UK Limited, Health and Safety Executive: United Kingdom.

Lee, L.F. (1986). Specification test for Poisson regression models. *International Economic Review, 27*, 689-706.

Lenné, M.G., Triggs, T.J., & Redman, J.R. (1997). Time of day variations in driving performance. *Accident Analysis and Prevention, 29*(4), 431-437.

Mackie, R.R. and Miller, J.C., 1978. *Effects of Hours of Service, Regularity of Schedules, and Cargo Loading on Truck and Bus Driver Fatigue*. Technical Report 1765-F NHTSA. Goleta, CA: Human Factors Research, Inc.

Mallis, M., Maislin, G., Konowal, N., Byrne, V., Bierman, D., Davis, R., et. al. (2000). *Biobehavioral Responses to Drowsy Driving Alarms and Alerting Stimuli*. Final report to develop, test and evaluate a drowsy driver detection and warning system for commercial motor vehicle drivers sponsored by the National Highway Traffic Safety Administration. Federal Highway Administration, Office of Motor Carriers.

Mitler, M.M., Miller, J.C., Lipsitz, J.J., Walsh, J.K., & Wylie, C.D. (1997). The sleep of long-haul truck drivers. *The New England Journal of Medicine, 337*(11).

Moller, H.J., Kayumov, L., Bulmash, E.L., Nhan, J., & Shapiro, C.M. (2006). Simulator performance, microsleep episodes, and subjective sleepiness: Normative data using convergent methodologies to assess driver drowsiness. *Journal of Psychosomatic Research, 61*(3), 335-342.

Moore-Ede, M.C., Sulzman, F.M., & Fuller, C.A. (1982). *The Clocks that Time Us*. Cambridge, MA: Harvard University Press.

Mulder, L. (1992). Measurement and analysis methods of heart rate and respiration for use in applied environments. *Biological Psychology, 34*, 205-336.

NHTSA. (2008). *Fatality Analysis Reporting System* [database]. Washington, DC: National Highway Traffic Safety Administration.

Neale, V.L., Dingus, T.A., Klauer, S.G., Sudweeks, J., & Goodman, M. (2005). *An Overview of the 100-Car Naturalistic Study and Findings*. Paper No. 05-0400. Blacksburg, VA: Virginia Tech Transportation Institute.

Neale, V.L., Klauer, S.G., Dingus, T.A., Holbrook, G.T., & Peterson, A.D. (2001). *Task 1 Report, Final Measure Selection: Set of Potential Variables for Naturalistic Driving Study*. Washington, DC: National Highway Traffic Safety Administration.

Nishimura, C., & Nagumo, J. (1985). Feedback control of arousal using skin potential level as an index. *Ergonomics, 28*(6), 905-913.

Oron-Gilad, T., & Ronen, A. (2007). Road characteristics and driver fatigue: A simulator study. *Traffic Injury Prevention, 8*(3), 281-289.

Paz-Frankel, E. (2006). Truck driver turnover reaches record level. *Memphis Business Journal*, March 31, 2006. Retrieved on July 21, 2008 from http://www.bizjournals.com/memphis/stories/2006/04/03/story3.html?jst=s_cn_hl

Petit, C., Chaput, D., Tarriere, C., Le Coz, J.Y., & Planque, S. (1990). Research to prevent the driver from falling asleep behind the wheel. *Proceedings of the 34th AAAM Conference*, American Association of Automotive Medicine, 505-523.

Pizza, F., Contardi, S., Mostacci, B., & Cirignotta, F. (2004). A driving simulation task: Correlations with Multiple Sleep Latency Test. *Brain Research Bulletin, 63*(5), 423-426.

Porcu, S., Ferrara, M., Urbani, L., Bellatreccia, A., & Casagrande, M. (1998). Smooth pursuit and saccadic eye movements as possible indicators of nighttime sleepiness. *Physiology and Behavior, 65*(3), 437-443.

Rechtschaffen, A., & Siegel, J. (2000). Sleep and dreaming. In E.R. Kandel, J.H. Schwartz, and T.M. Jessell (Eds.), *Principles of Neural Science* (4^{th} Ed.), New York: McGraw-Hill Medical, 937-947.

Rimini-Doering, M., Manstetten, D., Altmueller, T., Ladstaetter, U., & Mahler, M. (2001). Monitoring driver drowsiness and stress in a driving simulator. *Proceedings of the International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design*, Aspen, CO.

Rimini-Doering, M., Altmueller, T., Ladstaetter, U., & Rossmeier, M. (2005). Effects of lane departure warning on drowsy drivers' performance and state in a simulator. *Proceedings of the International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Rockport, ME.

RSSB. (2002). *Rail Safety & Standards Board Research Program Human Performance, Driver Vigilance Devices: Systems Review*. Unpublished technical report. Rail Safety & Standards Board, London, UK.

Schmidt, C., Collette, F., Cajochen, C., & Peigneux, P. (2007). A time to think: Circadian rhythms in human cognition. *Cognitive Neuropsychology, 24*(7), 755-789.

Singer, J.D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, *24*(4), 323-335.

Singer, J.D. & Willett, J.B. (2003). *Applied Longitudinal Data Analysis*. Oxford University Press.

Sirevaag, E.J., & Stern, J.A. (2000). Ocular measures of fatigue and cognitive factors. *Proceedings of the 11th Symposium on Human Interface*, Kyoto, Japan: Human Interface Society, 225-230.

Sivak, M. (1996). The information that drivers use: Is it indeed 90% visual? *Perception, 25*(9), 1081-1090.

Skipper, J.H., & Wierwille, W.W. (1986a). An investigation of low-level stimulus-induced measures of driver drowsiness. In A.G. Gale et al. (Eds.), *Vision in Vehicles, Proceedings of a Conference on Vision in Vehicles*, Nottingham, UK: North Holland Elsevier Science Publishers, 139-148.

Skipper, J.H. & Wierwille, W.W. (1986b). Drowsy driver detection using discriminant analysis. *Human Factors, 28*(5), 527-540.

Smiley, A. (2002). Fatigue and driving. In R.E. Dewar & P.L. Olson (Eds.), *Human Factors in Traffic Safety*, Tucson, AZ: Lawyers & Judges, 143-176.

Spiegel, K., Tasali, E., Penev, P., & Van Cauter, E. (2007). Brief Communication: Sleep Curtailment in Healthy Young Men Is Associated with Decreased Leptin Levels, Elevated Ghrelin Levels, and Increased Hunger and Appetite. *Annuals of Internal Medicine, 141*(11), 847.

Stern, J.A., Boyer, D.J., & Schroeder, D. (1994). Blink rate: A possible measure of fatigue. *Human Factors, 36*(2), 285-297.

Stutts, J.C., Wilkins, J.W. & Vaughn, B.V. (1999). *Why Do People Have Drowsy Driving Crashes? Input from Drivers Who Just Did*. Washington, DC: AAA Foundation for Traffic Safety.

Sugarman, R.C., & Cozad, C.P. (1972). *Road Tests of Alertness Variables*. Calspan Report No. ZM-5019-B-1. Buffalo, NY: Calspan.

Tani, P., Lindberg, N., Wendt, T. N.-V., Wendt, L. V., Alanko, L., Appelberg, B., et al. (2005). Actigraphic assessment of sleep in young adults with Asperger syndrome. *Psychiatry and Clinical Neurosciences, 59*, 206–208.

Tasali, E., Leproult, R., Ehrmann, D. A., & Van Cauter, E. (2008). Slow-wave sleep and the risk of type 2 diabetes in humans. *Proceedings of the National Academy of Sciences*, 0706446105.

Thiffault, P., Bergeron, J. (2003). Monotony of road environment and driver fatigue: A simulator study. *Accident Analysis and Prevention, 35*, 321-331.

van der Berg, J., Neely, G., Wiklund, U., & Landstöm, U. (2005). Heart rate variability during sedentary work and sleep in normal and sleep-deprived states. *Clinical Physiology and Functional Imaging, 25*(1), 51-57.

van der Laan, J.D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research: Part C, 5*(1), 1-10.

Van Dongen, H.P., Maislin, G., Mullington, J.M., & Dinges, D.F. (2003). The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep, 26*(20), 117-126.

Volow, M.R., & Erwin, C.W. (1973). *The Heart Rate Variability Correlates of Spontaneous Drowsiness Onset*. Society of Automotive Engineers Paper No. 730124, presented at the International Automotive Engineering Congress, Detroit, MI.

Wang, J.S., Knipling, R.R., and Blincoe, L.J. (1999). The dimensions of motor vehicle crash risk. *Journal of Transportation and Statistics, 2*(1), 19-43

Webb, W.B. (Ed.). (1982). *Biological Rhythms, Sleep, and Performance*. New York: Wiley.

Wiegand, D. M., Hanowski, R. J., Olson, R., & Melvin, W. (2008). *Fatigue Analyses from 16 Months of Naturalistic Commercial Motor Vehicle Driving Data*. Blacksburg, VA: National Surface Transportation Safety Center for Excellence.

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

Wierwille, W.W. & Ellsworth, L.A. (1994). Evaluation of driver drowsiness by trained observers. *Accident Analysis and Prevention, 26*(5), 571-581.

Wierwille, W.W., Ellsworth, L.A., Wreggit, S.S., Fairbanks, R.J., Kirn, C.L. (1994). *Research on Vehicle-Based Driver Status/Performance Monitoring: Development, Validation, and Refinement of Algorithms for Detection of Driver Drowsiness*. Report No. DOT-HS-808-247. Washington, DC: National Highway Traffic Safety Administration.

Wierwille, W. W., Hanowski, R.J., Olson, R.L., Dinges, D.L., Price, N.J., Maislin, G., et al. (2003). *NHTSA Drowsy Driver Detection and Interface Project*. Report No. DTNH 22-00-D-07007. Blacksburg, VA: VTTI.

Wilson, B., Popkin, S., Barr, L., & Hitz, J. (2002). *Evaluation Plan for the Drowsy Driver Warning System Field Operation Test*. Unpublished Report. Cambridge, MA: Volpe Center.

Williamson, A., Feyer, A., & Friswell, R. (1996). The impact of work practices on fatigue in long distance truck drivers. *Accident Analysis and Prevention, 28*(6), 709-719.

Wylie, C. D., Schultz, T., Miller, J. C., Mitler, M. M., & Mackie, R. R. (1996). *Commercial Motor Vehicle Driver Fatigue and Alertness Study: Technical Summary* (Rep. No. FHWA-MC-97-001). Washington, DC: U.S. Department of Transportation, Federal Highway Administration.

Xiao, Y., & Seagull, F.J. (1999). An analysis of problems with auditory alarms: Defining the roles of alarms in process monitoring tasks. *Proceedings of the Human Factors & Ergonomics Society 43rd Annual Meeting*, 256-260.

APPENDIX A. ALERT VALIDATION PROTOCOL

The driver fatigue monitor is a device that is used to determine if a driver is drowsy. The DFM uses image recognition to measure PERCLOS and alerts the driver when PERCLOS values reach a specified threshold; however, the reliability of these alarms is dependent upon the DFM's ability to continually locate the driver's eyes and make an accurate determination as to whether they are open or closed. The DFM may give false alarms when it either incorrectly locates the driver's eyes or incorrectly records them as closed when they are open. The data reductionist will need to determine: (1) if the alert presented by the DFM is valid (i.e., if there are slow eye-closures at the time of the alert), and (2) if invalid, the cause for the invalid alert.

This document specifies the protocol for assessing DFM's alert validity and describing the reason for the invalid alerts. A drowsiness assessment is required to determine whether a DFM alert is valid.

Sign Out Events:

1. Go to My Network Places
2. Browse to \\Bigfoot\425803\Reduction\MasterLists
3. Open "ADDWStriggersbydriver.xls"
4. Initial a Driver to indicate that you've signed it out to work on.
5. Go to that driver's spreadsheet and begin working on Week 1.
    a. Complete the columns for each trigger as your review it to indicate whether it is valid or invalid.
    b. When you have reached 7 valid triggers for each week, move on to the next week.
    c. If you review 20 triggers for a given week and none are valid, move on to the next week.
    d. If you review 50 triggers for a given week and are unable to reach 7 valids, move on to the next week.
    e. Once you complete a week's worth of triggers and move on to the next week, it may be helpful to somehow mark the remainder of that week so that you do not start working on it again later. (e.g., use shading, etc.)
6. Save

Loading Dart:

1. Load DART
2. Open the ThirtyFourTruck data collection (Figure 151)
3. Logon (username is your VTTI userid, password is VTTI)
4. Please change your password at first use under User > View Profile
5. Switch to the ADDWS reduction (User → Change Reduction → ADDWS)

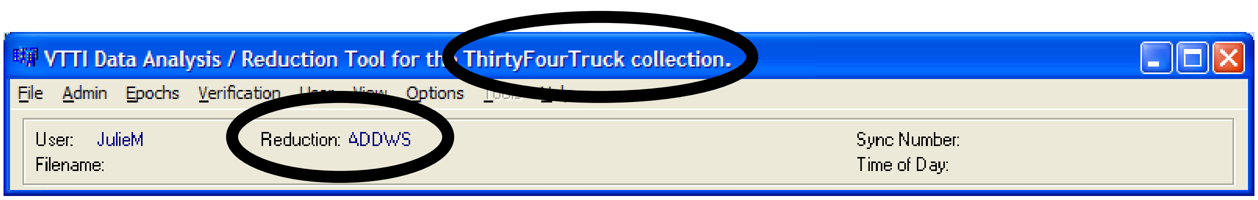


Figure 151. Opening the Proper Collection/Reduction

Open Trigger for Analysis:

1. In DART, load Query Tool.
2. Load the "Triggers by ID" query.
3. Copy the Trigger ID you signed out from the Excel log and paste it into the Condition List in DART.
4. Click on Go.
5. Open the Trigger listed in the results.

Set up Views:

You will need the following views open (Figure 152).

1. View → Video and Play Controller
2. View → Triggers
3. View → Network Speed
4. View→PERCLOS→PERCLOS_3
5. View → PERCLOS → Status → Sensitivity Level
6. View → PERCLOS → Status → Number of Eyes Found

Once you've arranged your windows the way you want them, save your View Setup for use the next time you log in (User → View Setups → Add). A sample View Setup is below.

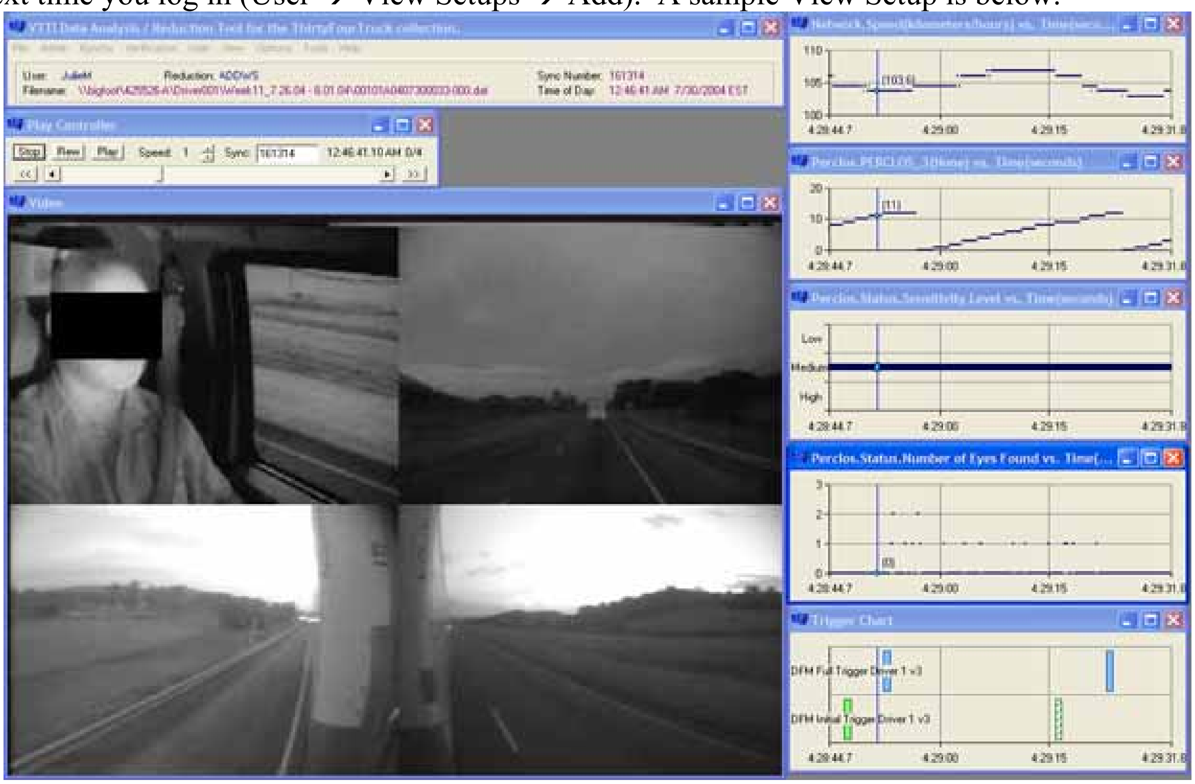


Figure 152. View Arrangement Within DART

Analyze Trigger:

1. Pull the trigger down into a new Event row.

2. Open the Question Reduction window by right clicking on the event and selecting the "Alert Validation" option. The question screen will appear (Figure 153).

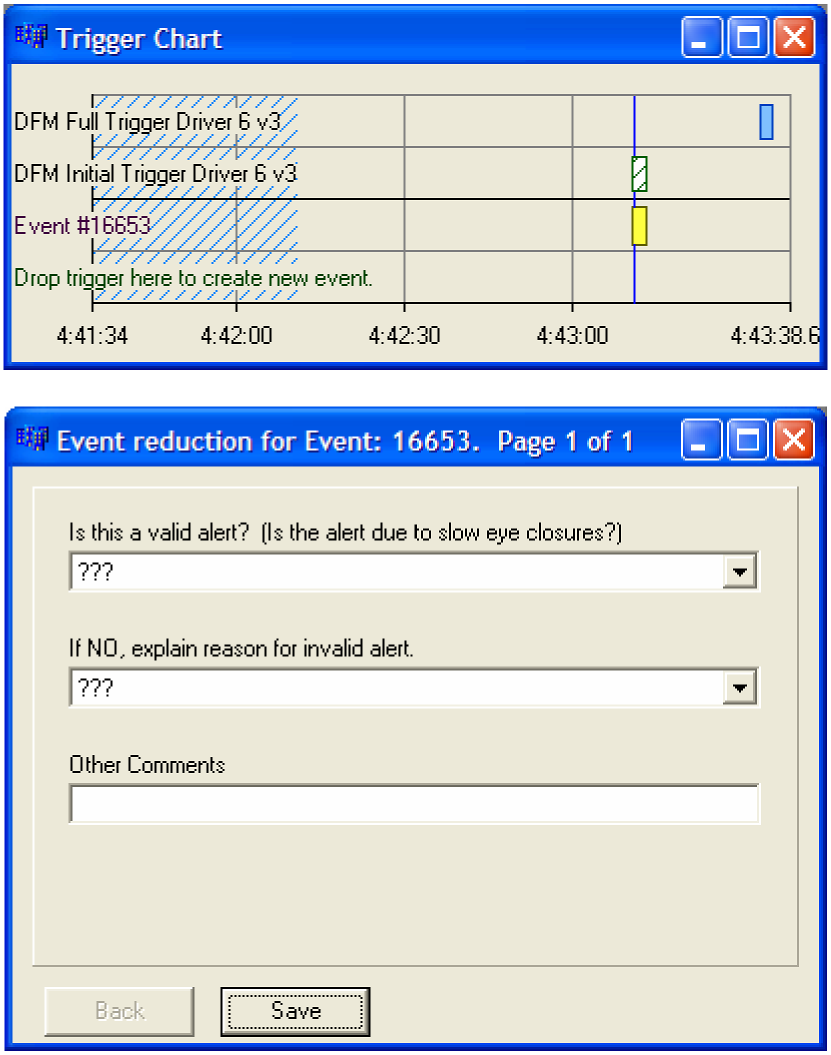


Figure 153. Trigger Analysis

3. Review the 1 minute of video prior to the alert and answer the questions using the pull down menus. The questions and the available responses are detailed below. The one minute mark can be easily found by subtracting 600 from the first frame of the trigger.
4. When complete, click on Save.
5. Return to the Excel log, mark that trigger as complete, and move to the next trigger, repeating steps 1 through 5.

Alert Validation Questions:

1. Is this a valid alert? (Is the alert due to slow eye-closures?)

   a. Yes (alert due to slow eye-closures) – See below for further explanation

b. No (alert NOT due to slow eye-closures)

Description of Slow Eye-Closures:
The DFM was designed to detect slow eye-closures as a measurement of driver drowsiness. An alert is considered to be valid if the alert was caused by slow eye-closures. These slow eye-closures may be accompanied by slowly dropping the head toward the chest or rolling the head up toward the head rest. If the alert is caused by anything other than a slow eye-closure (e.g., head turning, looking down), the alert is invalid. To determine which activities caused invalid DFM alerts, the reductionist should examine the increase in PERCLOS value with the video. If it is determined that the upward trend in PERCLOS values was generally caused by eye-closures, the alert should be scored as valid. The Number of Eyes Found variable can be used to determine if the increase in PERCLOS was caused by eye-closures or by the DFM's inability to find the drivers eyes. The DFM assumes if no eyes are found the driver is closing their eyes. If no eyes are found and the driver is looking forward and their eyes are open it can be assumed that the DFM was not locating the driver's eyes correctly.

Below are listed the alert parameters for three settings of the DFM. The driver selects the sensitivity level, and then receives alerts as the DFM-calculated PERCLOS scores reach the listed values. The numbers in parentheses represent the cumulative length of time that the DFM needs to record the eyes as closed over the specified interval in order for the driver to receive an alert. This is also the minimum amount of time you should have before either an initial or full alert (as indicated). For example, once a full alert is reached (which resets the PERCLOS back to zero), at least 30 seconds will elapse before a new initial alert is given when the sensitivity level is set to Low. If an alert is actually given after only 30 seconds, then the DFM monitor has recorded the eyes as Closed for that entire 30 second interval. Use this information as a guideline in determining if alerts are valid. If the eyes were in fact NOT closed for that interval, then the alert is invalid.

P1 - High Sensitivity

PERCLOS calculation period	1 min (PERCLOS 1)
Initial advisory tone PERCLOS	8 percent (4.8 s)
Full warning PERCLOS	12 percent (7.2 s)

P3 - Medium Sensitivity

PERCLOS calculation period	3 min (PERCLOS 3)
Initial advisory tone PERCLOS	9 percent (16.2 s)
Full warning PERCLOS	12 percent (21.6 s)

P5 - Low Sensitivity

PERCLOS calculation period	5 min (PERCLOS 5)
Initial advisory tone PERCLOS	10 percent (30 s)
Full warning PERCLOS	12 percent (36 s)

Please note that if you feel the above descriptions overlooked something important or do not properly describe what you are viewing, then supplement the description with your own best judgment in making your rating.

2. If NO, explain reason for invalid alert.

If the alert is determined to be invalid in step #1 above, please indicate the reason for the invalid alert based on the 60 seconds of video viewed. In the case that 60 full seconds are not available, however

30 to 60 seconds of video is available, perform the analysis as usual. If less than 30 seconds is available mark as "Not Enough Video". If any of the following behavior is found, select the corresponding option in the drop down menu:

Please select first alternative if the alert is a valid alert:

- Not Applicable – Alert is valid

For invalid alerts:

- Eyes Obstructed (hat, hands, etc.)
- Glasses/Sunglasses
- Change in lighting (overhead, glare, etc.)
- Outside light level to bright
- Driving-related scanning (mirrors, etc.)
- CB Radio usage
- Cell phone usage
- Other electronic device (PDA, iPod, etc.)
- Reading
- Eating/Drinking
- Looking down to the front or side (other or unknown reason)
- Smoking/Tobacco related
- Personal hygiene
- Other (explain below)

3. Other Comments

If the reason for the invalid alert is not covered in the above list, or further explanation is necessary, please include an explanation here.

Quality Assurance:

For the 1st 100 alerts there will be a 100 percent inspection by a CTBS member(s) attached to the ADDWS project. After this initial period, QA inspectors will only need to review those triggers that have been marked as valid. If the number of valid triggers for a given driver is less than 10 percent of all the triggers, then invalid triggers will be randomly reviewed until the 10 percent quota has been attained.

APPENDIX B. ALERT VALIDATION QUALITY ASSURANCE PROTOCOL

BACKGROUND

The driver fatigue monitor is a device that is used to determine if a driver is drowsy. The DFM uses image recognition to measure PERCLOS and alerts the driver when PERCLOS values reach a specified threshold; however, the reliability of these alerts is dependent on the DFM's ability to continually locate the driver's eyes and make an accurate determination as to whether they are open or closed. The DFM may give false alarms when it either incorrectly locates the driver's eyes or incorrectly records them as closed when they are open. The data reductionist determines: (1) if the alert presented by the DFM is valid (i.e., if there are slow eye-closures at the time of the alert), and (2) if invalid, the cause for the invalid alert. The role of the QA inspector is to verify the judgment of the data reductionist.

PURPOSE

This document specifies the protocol for performing quality control of the data reductionists' assessment of DFM alert validity. For the 1st 100 triggers in a reductionist's effort, there will be 100 percent inspection by the CTBS member(s) attached to the ADDWS project. After this initial period, 10 percent will be spot checked by reductionists with a third person reviewing all discrepancies on a weekly basis.

PROTOCOL

Loading Dart:

1. Load DART
2. Open the ThirtyFourTruck data collection (Figure 154)
3. Logon (username is your VTTI userid, password is VTTI)
4. Please change your password at first use under User > View Profile
5. Switch to the ADDWS reduction (User → Change Reduction → ADDWS)

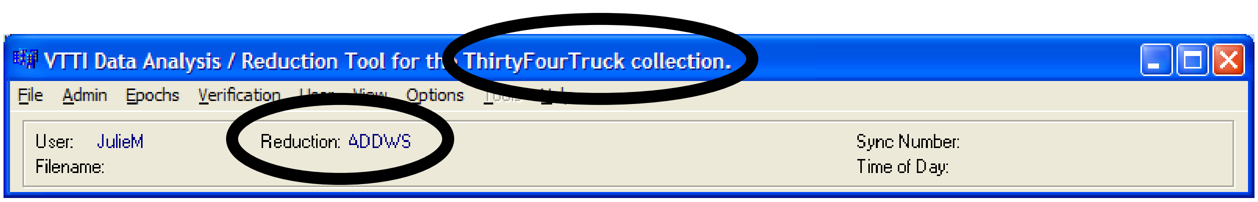


Figure 154. Opening the Proper Collection/Reduction

Open Trigger for Analysis:

1. In DART, load Query Tool.
2. Load the "Triggers by ID" query.
3. Copy the Trigger ID you signed out from the Excel log and paste it into the Condition List in DART.
4. Click on Go.
5. Open the Trigger listed in the results.

Set up Views:

You will need the following views open (Figure 155).

1. View → Video and Play Controller
2. View → Triggers
3. View → Network Speed
4. View→PERCLOS→PERCLOS_3
5. View → PERCLOS → Status → Sensitivity Level
6. View → PERCLOS → Status → Number of Eyes Found

Once you've arranged your windows the way you want them, save your View Setup for use the next time you log in (User → View Setups → Add). A sample View Setup is below.

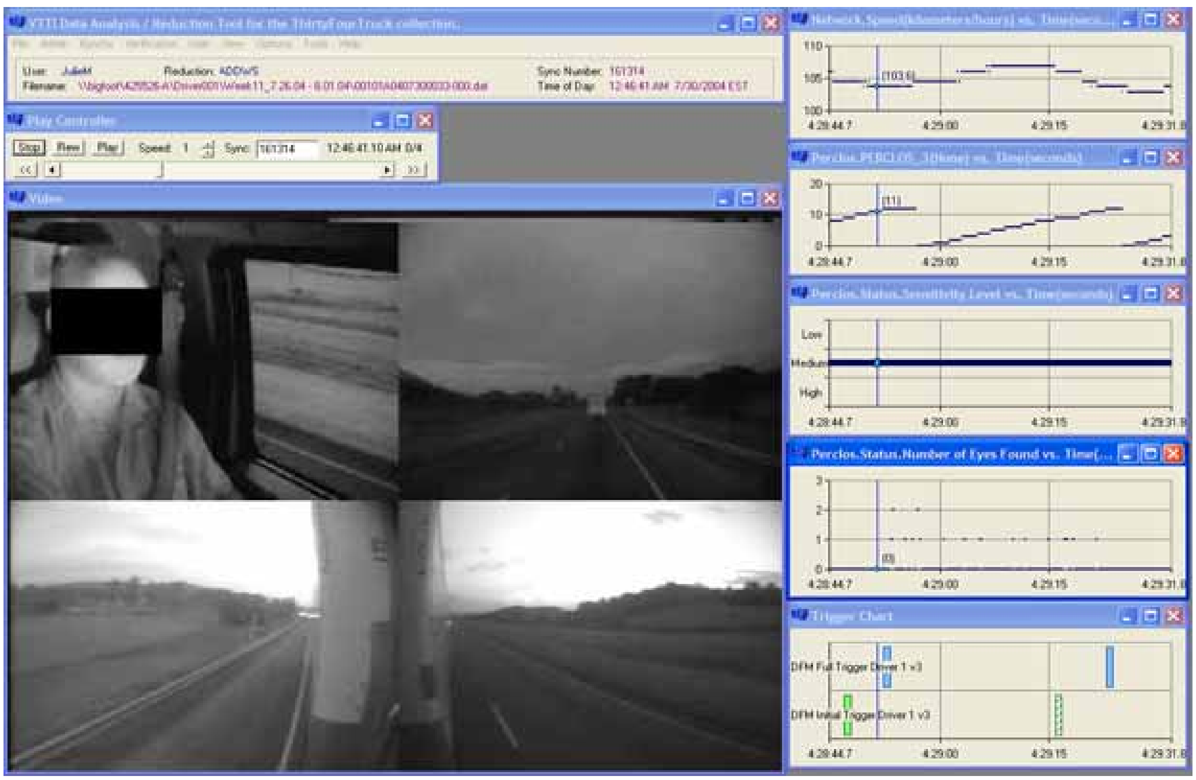


Figure 155. View Arrangement Within DART

Sign Out Events:

1. Go to My Network Places
2. Browse to \\Bigfoot\425803\Reduction\MasterLists
3. Open "Driver Sign Out & QA.xls"
4. Go to the "All Drivers" Tab
5. Identify a reductionist that hasn't been reviewed for inspection and sign them out.

6. Specify the QA inspection type (100 percent or 10%), QA inspector, and Status. (Look for any reductionist that hasn't had the first 100 triggers reviewed since they are high priority).

100 Percent Inspection

1. As a QA inspector, you are to review the first 100 triggers a reductionist reviews.
2. For a given reductionist, locate the driver number they have worked on (column A).
3. Locate the driver spreadsheet for that reductionist. For example, Jen, who worked on driver 111, would have worked on the Drivers 111-121 Jen & Sarah.xls spreadsheet.
4. Review 1 minute of video prior to the alert. The one minute mark can be easily found by subtracting 600 from the first frame of the trigger.
5. Look for slow eye-closures and whether the PERCLOS is inflated due to the eyes being closed or the driver's head being out of range from the DFM.
6. Complete the Quality Assurance (QA) columns for each trigger (both valid and invalid triggers), marking whether the trigger is valid or invalid in your opinion.
7. If you disagree with the reductionist's assessment, please highlight this trigger in yellow so it can be easily identified by a 3rd QA inspector. Please also initial the trigger so we know who reviewed it.
8. Save the spreadsheet.
9. After you have reviewed the triggers, you are to notify Rebecca Olson. Rebecca will inspect the discrepancies and provide a third and final assessment of the trigger validity.
10. Based on the results of the final reviewer, you are to meet with the initial reviewer and go over the reasons why their assessment was incorrect.

10 Percent Inspection

1. After a QA inspector has reviewed the first 100 triggers for a reductionist, QA inspectors will only need to review those triggers that have been marked as valid. If the number of valid triggers for a given driver is less than 10 percent of all the triggers, then invalid triggers will be randomly reviewed until the 10 percent quota has been attained. To see if this case exists for your driver, go to the "10 percent QA Inspection" tab and look up your driver number. If your driver number is there, inspect the triggers that have been listed.
2. Next to the valid triggers in the driver spreadsheet, complete the Quality Assurance (QA) columns for each valid trigger, marking whether the trigger is valid or invalid.
3. Review 1 minute of video prior to the alert. The one minute mark can be easily found by subtracting 600 from the first frame of the trigger.
4. Look for slow eye-closures and whether the PERCLOS is inflated due to the eyes being closed or the driver's head being out of range from the DFM.
5. If you determine that a trigger marked as valid is indeed invalid, please take note of this in the QA columns in the driver worksheet. Please highlight this trigger in yellow so it can be easily identified by a 3rd QA inspector. Please also initial the trigger so we know who reviewed it.
6. Save the spreadsheet.
7. After you have reviewed all the triggers, you are to notify Rebecca Olson. Rebecca will inspect the discrepancies and provide a third and final assessment of the trigger validity.

8. Based on the results of the final reviewer, you are to meet with the initial reviewer and go over the reasons why their assessment was incorrect.

APPENDIX C. POST-ALERT BEHAVIOR IDENTIFICATION PROTOCOL

PURPOSE

The purpose of this reduction is to document the driver behavior that occurs in the five minutes following a valid DFM alert. This document specifies the protocol for assessing DFM post-alert behavior.

PROTOCOL

Sign Out Events:

1. Go to My Network Places
2. Browse to \\bigfoot\425803\Reduction\DFM Post-Alert Behavior\
3. Open "PostAlertBehavior_SignOut&Drivers1-10"
4. Initial a driver to indicate that you've signed it out to work on.
5. Open the Excel file that contains the spreadsheet for that driver. Initial an event under your driver to indicate that you've signed it out to work on it. Also specify the date you took it out.
6. Save

Loading Dart:

1. Load DART
2. Open the ThirtyFourTruck data collection (File → Open Collection → Select Thirty Four Truck from the drop down list)
3. User → Login → Logon (username is your VTTI userid, password is VTTI)
    a. Please change your password at first use under User > View Profile
4. Switch to the ADDWS reduction (User → Change Reduction → ADDWS, Figure 156)

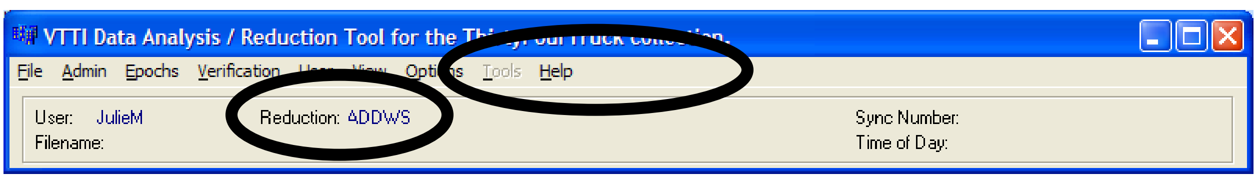


Figure 156. Opening the Proper Collection/Reduction

Open Trigger for Analysis:

1. In DART, load Query Tool (File → Query Tool).
2. Load the 'EventbyID' query (File → Saved Query List)
3. Copy the Event ID you signed out from the Excel log and paste it into the Condition List in DART.
4. Click on Go.
5. Open the Event listed in the results by double clicking on it.

Set up Views:

You will need the following views open (Figure 157).

1. View → Video and Play Controller
2. View → Triggers
3. View → Network Status
4. View → Network Speed
5. Tools → Event Reduction Manager

Once you've arranged your windows the way you want them, save your View Setup for use the next time you log in (User → View Setups → Add). A sample View Setup is below.

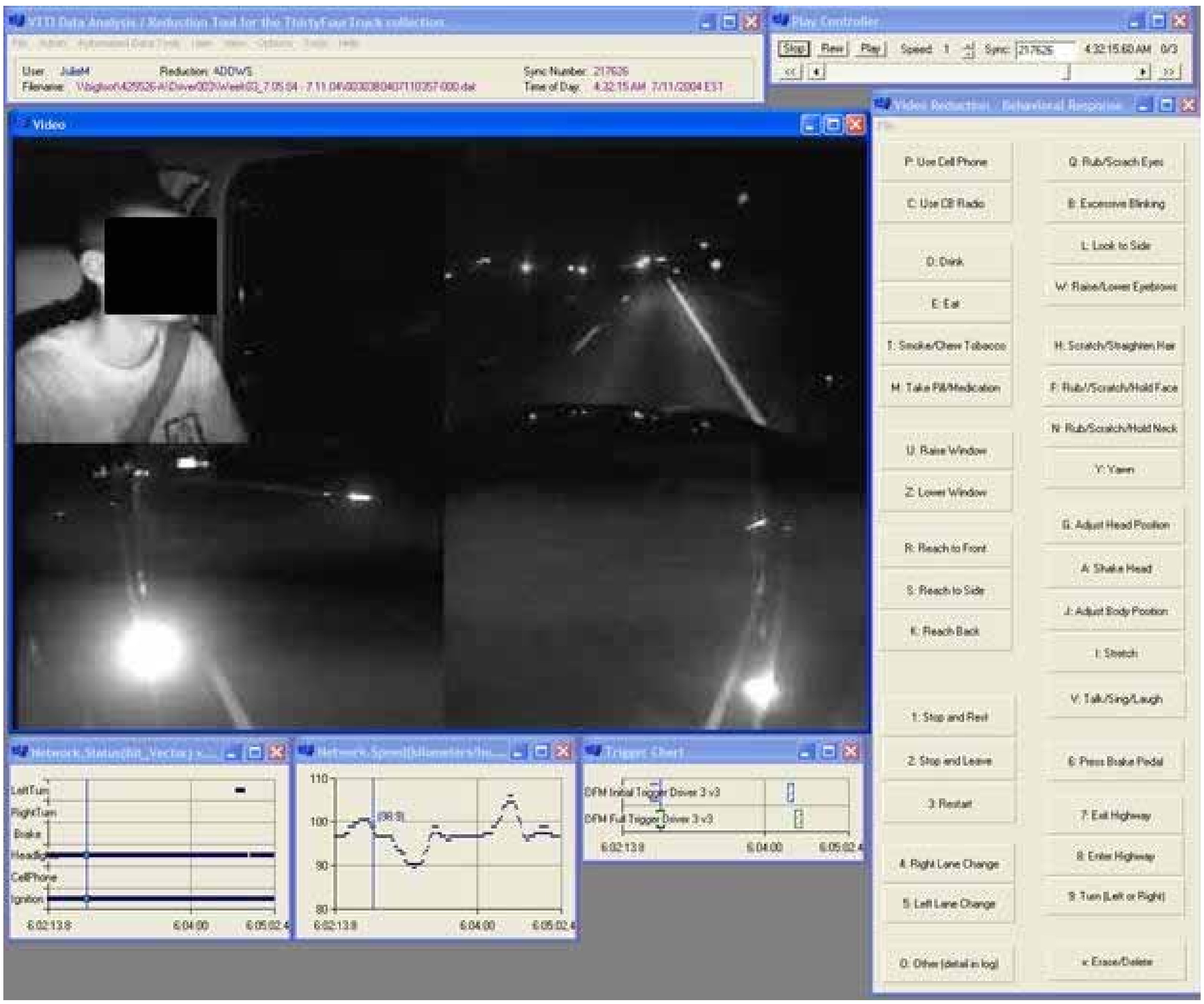


Figure 157. View Arrangement Within DART

Document Environmental Conditions:

- For a given trigger, right click in the trigger window and select validation. This shows the event number for the triggers.
- Right click on the event associated with the valid trigger.
- Select the environmental conditions question reduction from the popup menu.
- You are to mark the environmental conditions listed below, that exist at the sync point of the valid DFM alert. You may watch the video for a few seconds if necessary to obtain all of the needed information.

Weather

- 01 = No adverse conditions (If difficult to tell what the condition is, err on the side of no adverse conditions. You must be able to say with certainty that an adverse condition, such as snow, exists in order to mark one down.)
- 02 = Rain
- 03 = Sleet (Sleet must be on the road, and not just on the side of the road)
- 04 = Snow (Snow must be on the road, and not just on the side of the road)
- 05 = Fog
- 06 = Rain & fog
- 07 = Sleet & fog
- 08 = Other (smog, smoke, sand/dust, crosswind, hail)
- 09 = Unknown

Roadway Surface Condition

- 01 = Dry
- 02 = Wet (doesn't have to be raining)
- 03 = Snow or slush
- 04 = Ice (are they slipping around)
- 05 = Sand, oil, dirt
- 08 = Other
- 09 = Unknown

Traffic-way Flow

- 00 = Not physically divided (center 2-way left turn lane): This is when you have traffic going in opposite directions. In between both lanes is a center turn lane that is accessible to vehicles going in either direction of travel.
- 01 = Not physically divided (2-way trafficway): This is when there is a double yellow line separating traffic going in opposite directions.
- 02 = Divided (median strip or barrier): Typical example would be the interstate. This is when you have something physically separating the lanes that head in opposite directions.
- 03 = One-way trafficway: This would be exit ramps, exit ramps, or alleyways
- 09 = Unknown

Number of Travel Lanes

Note: Per GES, if road is divided, only lanes in travel direction are counted, if undivided, all lanes are counted; code in relation to subject vehicle. Count all contiguous lanes at the time of the trigger (e.g., include entrance/exit lanes or turn lanes if contiguous). Do not include lanes if blocked by cones or barrels.

- 01 = 1
- 02 = 2
- 03 = 3
- 04 = 4
- 05 = 5
- 06 = 6
- 07 = 7+
- 09 = Unknown

Traffic Density

Note: Code the traffic density at the time of the trigger.

- 01 = LOS A: Free flow – Individual users are virtually unaffected by the presence of others in the traffic stream. Freedom to select desired speeds and to maneuver within the traffic stream is extremely high. The general level of comfort and convenience provided to the motorist, passenger, or pedestrian is excellent.
- 02 = LOS B: Flow with some restrictions – In the range of stable traffic flow, but the presence of other users in the traffic stream begins to be noticeable. Freedom to select desired speeds is relatively unaffected, but there is a slight decline in the freedom to maneuver within the traffic stream from LOS A, because the presence of others in the traffic stream begins to affect individual behavior.
- 03 = LOS C: Stable flow, maneuverability and speed are more restricted – In the range of stable traffic flow, but marks the beginning of the range of flow in which the operation of individual uses becomes significantly affected by the interactions with others in the traffic stream. The selection of speed is now affected by the presence of others, and maneuvering within the traffic stream requires substantial vigilance on the part of the user. The general level of comfort and convenience declines noticeably at this level.
- 04 = LOS D: Unstable flow: temporary restrictions substantially slow driver – Represents high-density, but stable traffic flow. Speed and freedom to maneuver are severely restricted, and the driver or pedestrian experiences a generally poor level of comfort and convenience. Small increases in traffic flow will generally cause operational problems at this level.
- 05 = LOS E: Flow is unstable; vehicles are unable to pass, temporary stoppages, etc. – Represents operating conditions at or near the capacity level. All speeds are reduced to a low, but relatively uniform value. Freedom to maneuver within the traffic stream is extremely difficult, and it is generally accomplished by forcing a vehicle or pedestrian to “give way” to accommodate such maneuvers. Comfort and convenience levels are extremely poor, and driver or pedestrian frustration is generally high. Operations at this level are usually unstable, because small increases in flow or minor perturbations within the traffic stream will cause breakdowns.

- 06 = LOS F: Forced traffic flow condition with low speeds and traffic volumes that are below capacity. Queues' forming in particular locations – This condition exists whenever the amount of traffic approaching a point exceeds the amount which can traverse the point. Queues form behind such locations. Operations within the queue are characterized by stop-and-go waves, and they are extremely unstable. Vehicles may progress at reasonable speeds for several hundred feet or more, and then be required to stop in a cyclic fashion. LOS F is used to describe the operating conditions within the queue, as well as the point of the breakdown. It should be noted, however, that in many cases operating conditions of vehicles or pedestrians discharged from the queue may be quite good. Nevertheless, it is the point at which arrival flow exceeds discharge slow which causes the queue to form, and LOS F is an appropriate designation for such points.
- Unknown/unable to determine

Comments

- For any condition that is not made clear by the above choices, please make a comment.

Document Post-Alert Behavior:

1. Go to Tools → Event Reduction Manager
   a. In the Trigger chart (View → Triggers) right click on the trigger and then click on validation
   b. The events associated with the triggers are then shown in the rows below the triggers.
   c. Locate the Event# associated with the trigger
   d. In the event manager tool, look up this event number and click on it.
   e. This event is then shown in the status pane on the right.
   f. Click event start (This locates the start of the trigger. This can be done at any point).
   g. Keep this tool open throughout the video reduction.

2. From the sync_begin point of the DFM alert trigger, you are to analyze up to 5 minutes of video (3,000 frames) and record the post-alert driving behaviors.
   a. If the video file ends before the 5 minute mark (events highlighted in red), you are to open the next trip file to see what occurs in the remaining time. To do this, look up the '2nd Event ID' in the spreadsheet, and open that event through the Query Tool (saved Query – ByEventID). You do not need to re-enter the environmental conditions for the trigger in the second file.

3. You need to record the first occurrence of each behavior in the 5-minute interval. For example, if the driver takes two separate sips from a can of soda in the 5-minute period, you only need to record the first time the can of soda touches his/her mouth. **Please note**, if the 5-minute period is split between two files, you still need to only record the first occurrence between the entire 5 minutes (does not restart marking behaviors with second file).

4. It is suggested that you go through the video first at double-speed to get an idea of what behaviors occur.

5. Next, go back to the trigger sync_begin and play the video at normal speed. When you get to a behavior, step through the video footage frame by frame to find the exact begin_sync for that behavior.

6. A list of post-alert behaviors is provided in the video reduction tool. When a post-alert behavior is observed, such as “Yawn”, you are to press the “y” key at the frame the yawn begins. The table below specifies the operational definitions for the initiation of each behavior.
7. If you make a mistake, go back to the frame the “y” key was pressed and press the “x” key to erase your entry.
8. If a behavior is observed that is not listed, press the “0” key for ‘Other’. Make a note of the sync number where the other behavior began, and also note what the other behavior was.

Table 110. Operational Definitions for the Post-Alert Behaviors

Code	Post-Alert Behavior	When to Record Sync
P	Use Cell Phone	Mark the frame the cell phone first appears in the video image
C	Use CB Radio	Mark the frame the driver picks up the CB radio
D	Drink	Mark the frame the driver puts drink to mouth.
E	Eat	Mark the frame the driver puts food in his mouth.
T	Smoke/Chew Tobacco	Mark the frame the driver touches the cigarette or chewing tobacco to his/her mouth.
M	Take Pill/Medication	Mark the frame the driver puts the medication in his/her mouth.
Z	Raise/Lower Window	Mark the frame the window begins to move.
R	Reach to Front	Mark the frame the driver begins to move his arm.
S	Reach to Side/Reach Up	Mark the frame the driver begins to move his arm.
K	Reach Back	Mark the frame the driver begins to move his arm.
U	Veer Off Road	Mark the frame the truck unintentionally leaves the lane, either to the left or right side. Look for the paint markings to disappear under the truck as you would be lane changes.
V	Talk/Sing/Laugh	Mark the frame the driver begins to move his mouth.
Q	Rub/Scratch Eyes	Mark the frame the driver touches his eyes.
B	Excessive Blinking	Mark the frame the driver begins to blink at the start of the excessive blinking behavior.
L	Look to Side/Up/Down	Mark the frame the driver’s eyes/head initiate their fixation to the side (when the eyes land on where they’re looking).
W	Raise/Lower Eyebrows	Mark the frame the driver begins to raise or lower his eyebrows.
H	Scratch / Straighten Hair	Mark the frame the driver begins to touch his/her hair.
F	Rub / Scratch / Hold Face	Mark the frame the driver begins to touch his/her face.
N	Rub/Scratch/ Hold Neck	Mark the frame the driver begins to touch his/her neck.

Code	Post-Alert Behavior	When to Record Sync
Y	Yawn	Mark the frame the driver begins to open his/her mouth.
G	Adjust Head Position	Mark the frame the driver begins to move his/her head in the adjustment.
A	Shake Head	Mark the frame the driver begins to move his/her head when shaking it.
J	Adjust Body Position	Mark the frame the driver begins to move his/her body in the adjustment.
I	Stretch	Mark the frame the driver begins to move his/her body when stretching.
1	Stop and Rest	Mark the frame the truck comes to a complete stop based on the forward roadway video image, and driver remains in the driver seat to rest.
2	Stop and Leave	Mark the frame the truck comes to a complete stop based on the forward roadway video image and the driver subsequently leaves the driver seat.
3	Restart	Mark the frame the truck begins to move after the driver has taken a break. Stopping for reasons other than rest, such as at a red light, do not count.
4	Right Lane Change	Mark the frame the truck crosses the lane markings when making a right lane change. Only the first right lane change should be marked.
5	Left Lane Change	Mark the frame the truck crosses the lane markings when making a left lane change. Only the first left lane change should be marked.
6	Press Brake Pedal	Mark the frame the brake pedal is pressed using View → Network → Status. Only the first brake pedal press should be marked.
7	Exit Highway	If on a highway, mark the frame the truck enters the exit ramp using the forward roadway image.
8	Enter Highway	If entering a highway, mark the frame the truck enters the entrance ramp using the forward roadway image.
9	Turn (Left or Right)	Mark the frame the truck initiates the right or left hand turn based on the forward roadway camera view. Only the first turn should be marked.
O (letter)	Other	Any other behavior you see (e.g., dancing). Also record in the Excel log what the behavior is along with the sync number.
X	Erase (Delete)	To delete or erase a behavior mistakenly coded.

Examples of behaviors that should be marked as other:

- Eyes closed for extended period of time (to indicate that the driver *may* be falling asleep. Look for eyes closed more than 2 seconds.)
- Reading

Comments

In the excel log file, please note when the driver is excessively drowsy and you think it is a good example to show people during the final presentation. This will help us track down good examples when we need to show them.

APPENDIX D. POST-ALERT BEHAVIOR IDENTIFICATION QUALITY ASSURANCE PROTOCOL

The purpose of this document is to present the protocol for reviewing the reductionists' assessment of post-alert behaviors.

Sign Out Events:

1. Go to My Network Places
2. Browse to \\bigfoot\425803\Reduction\DFM Post-Alert Behavior\
3. Open "PostAlertBehavior_SignOut&Drivers1-10"
4. Initial a driver to indicate that you've signed it out to work on.
5. Open the Excel file that contains the spreadsheet for that driver. Initial an event under your driver to indicate that you've signed it out to work on it. Also specify the date you took it out.
6. Save

Loading Dart:

1. Load DART
2. Open the ThirtyFourTruck data collection (File → Open Collection → Select Thirty Four Truck from the drop down list)
3. User → Login → Logon (username is your VTTI userid, password is VTTI)
    a. Please change your password at first use under User > View Profile
4. Switch to the ADDWS reduction (User → Change Reduction → ADDWS, see Figure 158)

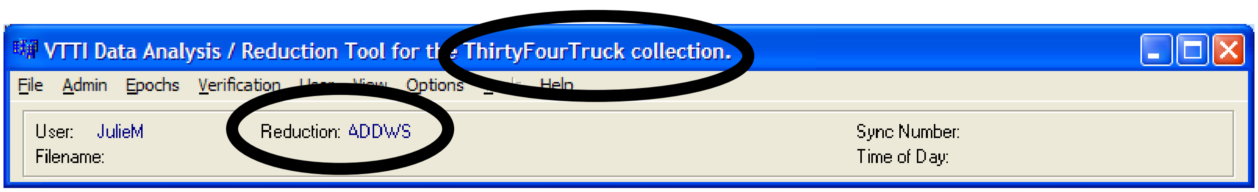


Figure 158. Opening the Proper Collection/Reduction

Open Trigger for Analysis:

1. In DART, load Query Tool (File → Query Tool).
2. Load the 'EventbyID' query (File → Saved Query List)
3. Copy the Event ID you signed out from the Excel log and paste it into the Condition List in DART.
4. Click on Go.
5. Open the Event listed in the results by double clicking on it.

Set up Views:

You will need the following views open.

1. View → Video and Play Controller
2. View → Triggers
3. View → Network Status
4. View → Network Speed

5. Tools → Event Reduction Manager

Once you've arranged your windows the way you want them, save your View Setup for use the next time you log in (User → View Setups → Add). A sample View Setup is below (Figure 159).

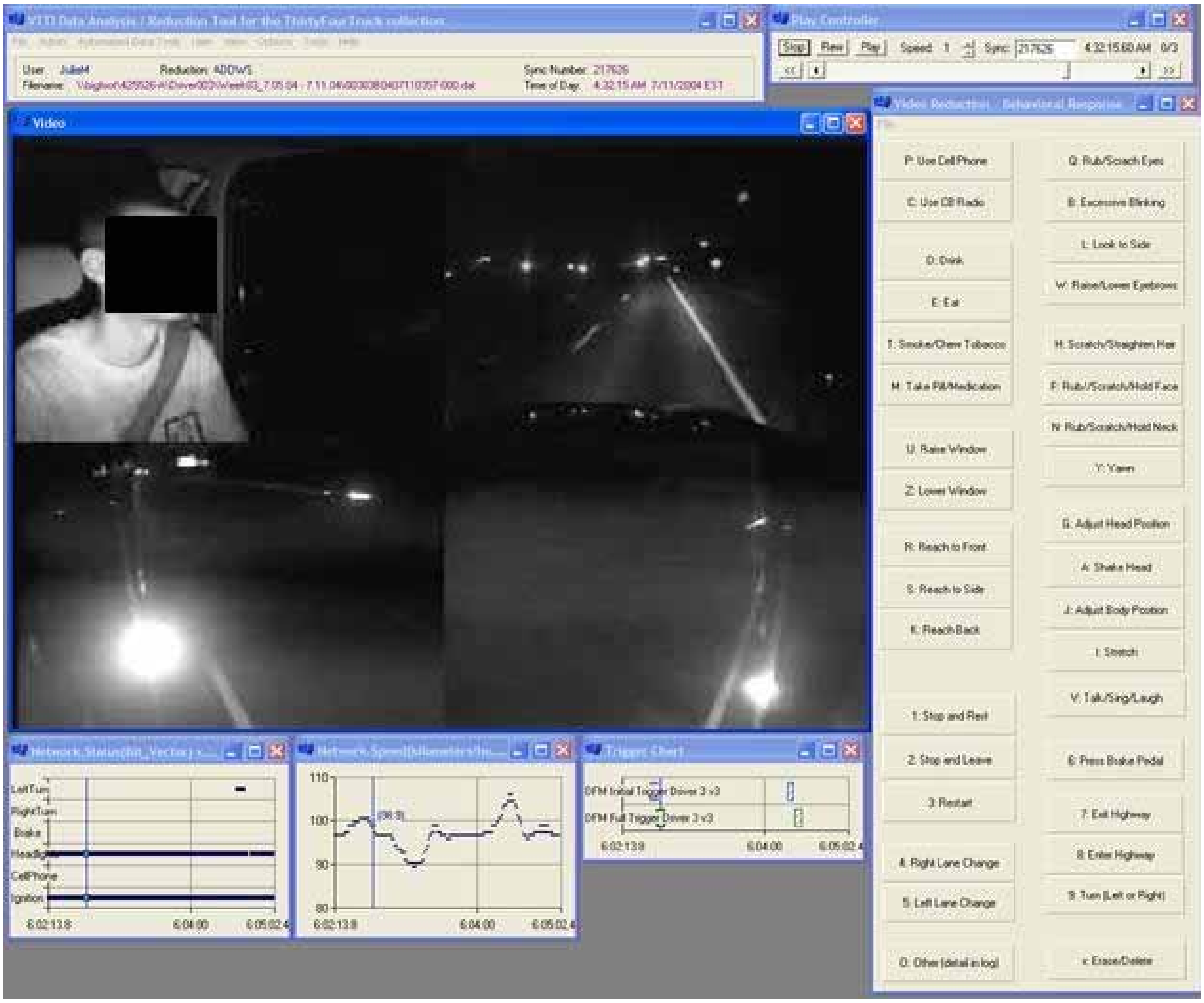


Figure 159. View Arrangement in DART

Inspect Environmental Conditions:

- For a given trigger, right click in the trigger window and select validation. This shows the event number for the triggers.
- Right click on the event associated with the valid trigger.
- Select the environmental conditions question reduction from the popup menu.
- Review the environmental conditions that were specified by the reductionist. Watch the video for a few seconds if necessary to obtain all of the needed information.

Weather

- 01 = No adverse conditions (If difficult to tell what the condition is, err on the side of no adverse conditions. You must be able to say with certainty that an adverse condition, such as snow, exists in order to mark one down.)
- 02 = Rain
- 03 = Sleet (Sleet must be on the road, and not just on the side of the road)
- 04 = Snow (Snow must be on the road, and not just on the side of the road)
- 05 = Fog
- 06 = Rain & fog
- 07 = Sleet & fog
- 08 = Other (smog, smoke, sand/dust, crosswind, hail)
- 09 = Unknown

1. Roadway Surface Condition

- 01 = Dry
- 02 = Wet (doesn't have to be raining)
- 03 = Snow or slush
- 04 = Ice (are they slipping around)
- 05 = Sand, oil, dirt
- 08 = Other
- 09 = Unknown

2. Traffic-way Flow

- 00 = Not physically divided (center 2-way left turn lane): This is when you have traffic going in opposite directions. In between both lanes is a center turn lane that is accessible to vehicles going in either direction of travel.
- 01 = Not physically divided (2-way trafficway): This is when there is a double yellow line separating traffic going in opposite directions.
- 02 = Divided (median strip or barrier): Typical example would be the interstate. This is when you have something physically separating the lanes that head in opposite directions.
- 03 = One-way trafficway: This would be exit ramps, exit ramps, or alleyways
- 09 = Unknown

3. Number of Travel Lanes

<u>Note:</u> Per GES, if road is divided, only lanes in travel direction are counted, if undivided, all lanes are counted; code in relation to subject vehicle. Count all contiguous lanes at the time of the trigger (e.g., include entrance/exit lanes or turn lanes if contiguous). Do not include lanes if blocked by cones or barrels.

- 01 = 1
- 02 = 2
- 03 = 3
- 04 = 4
- 05 = 5
- 06 = 6
- 07 = 7+
- 09 = Unknown

4. Traffic Density

Note: Code the traffic density at the time of the trigger.

- 01 = LOS A: Free flow – Individual users are virtually unaffected by the presence of others in the traffic stream. Freedom to select desired speeds and to maneuver within the traffic stream is extremely high. The general level of comfort and convenience provided to the motorist, passenger, or pedestrian is excellent.
- 02 = LOS B: Flow with some restrictions – In the range of stable traffic flow, but the presence of other users in the traffic stream begins to be noticeable. Freedom to select desired speeds is relatively unaffected, but there is a slight decline in the freedom to maneuver within the traffic stream from LOS A, because the presence of others in the traffic stream begins to affect individual behavior.
- 03 = LOS C: Stable flow, maneuverability and speed are more restricted – In the range of stable traffic flow, but marks the beginning of the range of flow in which the operation of individual uses becomes significantly affected by the interactions with others in the traffic stream. The selection of speed is now affected by the presence of others, and maneuvering within the traffic stream requires substantial vigilance on the part of the user. The general level of comfort and convenience declines noticeably at this level.
- 04 = LOS D: Unstable flow: temporary restrictions substantially slow driver – Represents high-density, but stable traffic flow. Speed and freedom to maneuver are severely restricted, and the driver or pedestrian experiences a generally poor level of comfort and convenience. Small increases in traffic flow will generally cause operational problems at this level.
- 05 = LOS E: Flow is unstable; vehicles are unable to pass, temporary stoppages, etc. – Represents operating conditions at or near the capacity level. All speeds are reduced to a low, but relatively uniform value. Freedom to maneuver within the traffic stream is extremely difficult, and it is generally accomplished by forcing a vehicle or pedestrian to "give way" to accommodate such maneuvers. Comfort and convenience levels are extremely poor, and driver or pedestrian frustration is generally high. Operations at this level are usually unstable, because small increases in flow or minor perturbations within the traffic stream will cause breakdowns.
- 06 = LOS F: Forced traffic flow condition with low speeds and traffic volumes that are below capacity. Queues' forming in particular locations – This condition exists whenever the amount of traffic approaching a point exceeds the amount which can traverse the point. Queues form behind such locations. Operations within the queue are characterized by stop-and-go waves, and they are extremely unstable. Vehicles may progress at reasonable speeds for several hundred feet or more, and then be required to stop in a cyclic fashion. LOS F is used to describe the operating conditions within the queue, as well as the point of the breakdown. It should be noted, however, that in many cases operating conditions of vehicles or pedestrians discharged from the queue may be quite good. Nevertheless, it is the point at which arrival flow exceeds discharge slow which causes the queue to form, and LOS F is an appropriate designation for such points.
- Unknown/unable to determine

5. **Comments**
   - For any condition that is not made clear by the above choices, please make a comment.

Review Post-Alert Behavior:

9. Go to Tools → Event Reduction Manager
   a. In the Trigger chart (View → Triggers) right click on the trigger and then click on validation
   b. The events associated with the triggers are then shown in the rows below the triggers.
   c. Locate the Event# associated with the trigger
   d. In the event manager tool, look up this event number and click on it.
   e. This event is then shown in the status pane on the right.
   f. Click event start (This locates the start of the trigger. This can be done at any point).
   g. Keep this tool open throughout the video reduction.

10. From the sync_begin point of the DFM alert trigger, review the 5 minutes of video (3,000 frames) that occur after the trigger and document any behaviors the reductionist missed.
    a. If the video file ends before the 5 minute mark (events highlighted in red), you are to open the next trip file to see what occurs in the remaining time. To do this, look up the '2nd Event ID' in the spreadsheet, and open that event through the Query Tool (saved Query – ByEventID). You do not need to re-enter the environmental conditions for the trigger in the second file.

11. Go through the video first at double-speed to get an idea of what behaviors occur.

12. Next, go back to the trigger sync_begin and play the video at normal speed. When you get to a behavior, step through the video footage frame by frame to find the exact begin_sync for that behavior.

13. A list of post-alert behaviors is provided in the video reduction tool. When a post-alert behavior is observed, such as "Yawn", you are to press the "y" key at the frame the yawn begins. The table below specifies the operational definitions for the initiation of each behavior.

14. If you make a mistake, go back to the frame the "y" key was pressed and press the "x" key to erase your entry.

15. If a behavior is observed that is not listed, press the "0" key for 'Other'. Make a note of the sync number where the other behavior began, and also note what the other behavior was.

Table 111. Operational Definitions for the Post-Alert Behaviors

Code	Post-Alert Behavior	When to Record Sync
P	Use Cell Phone	Mark the frame the cell phone first appears in the video image
C	Use CB Radio	Mark the frame the driver picks up the CB radio
D	Drink	Mark the frame the driver puts drink to mouth.
E	Eat	Mark the frame the driver puts food in his mouth.
T	Smoke/Chew Tobacco	Mark the frame the driver touches the cigarette or chewing tobacco to his/her mouth.
M	Take Pill/Medication	Mark the frame the driver puts the medication in his/her mouth.
Z	Raise/Lower Window	Mark the frame the window begins to move.
R	Reach to Front	Mark the frame the driver begins to move his arm.
S	Reach to Side/Reach Up	Mark the frame the driver begins to move his arm.
K	Reach Back	Mark the frame the driver begins to move his arm.
U	Veer Off Road	Mark the frame the truck unintentionally leaves the lane, either to the left or right side. Look for the paint markings to disappear under the truck as you would be lane changes.
V	Talk/Sing/Laugh	Mark the frame the driver begins to move his mouth.
Q	Rub/Scratch Eyes	Mark the frame the driver touches his eyes.
B	Excessive Blinking	Mark the frame the driver begins to blink at the start of the excessive blinking behavior.
L	Look to Side/Up/Down	Mark the frame the driver's eyes/head initiate their fixation to the side (when the eyes land on where they're looking).
W	Raise/Lower Eyebrows	Mark the frame the driver begins to raise or lower his eyebrows.
H	Scratch / Straighten Hair	Mark the frame the driver begins to touch his/her hair.
F	Rub / Scratch / Hold Face	Mark the frame the driver begins to touch his/her face.
N	Rub/Scratch/ Hold Neck	Mark the frame the driver begins to touch his/her neck.
Y	Yawn	Mark the frame the driver begins to open his/her mouth.
G	Adjust Head Position	Mark the frame the driver begins to move his/her head in the adjustment.
A	Shake Head	Mark the frame the driver begins to move his/her head when shaking it.
J	Adjust Body Position	Mark the frame the driver begins to move his/her body in the adjustment.

Code	Post-Alert Behavior	When to Record Sync
I	Stretch	Mark the frame the driver begins to move his/her body when stretching.
1	Stop and Rest	Mark the frame the truck comes to a complete stop based on the forward roadway video image, and driver remains in the driver seat to rest.
2	Stop and Leave	Mark the frame the truck comes to a complete stop based on the forward roadway video image and the driver subsequently leaves the driver seat.
3	Restart	Mark the frame the truck begins to move after the driver has taken a break. Stopping for reasons other than rest, such as at a red light, do not count.
4	Right Lane Change	Mark the frame the truck crosses the lane markings when making a right lane change. Only the first right lane change should be marked.
5	Left Lane Change	Mark the frame the truck crosses the lane markings when making a left lane change. Only the first left lane change should be marked.
6	Press Brake Pedal	Mark the frame the brake pedal is pressed using View → Network → Status. Only the first brake pedal press should be marked.
7	Exit Highway	If on a highway, mark the frame the truck enters the exit ramp using the forward roadway image.
8	Enter Highway	If entering a highway, mark the frame the truck enters the entrance ramp using the forward roadway image.
9	Turn (Left or Right)	Mark the frame the truck initiates the right or left hand turn based on the forward roadway camera view. Only the first turn should be marked.
O (letter)	Other	Any other behavior you see (e.g., dancing). Also record in the Excel log what the behavior is along with the sync number.
X	Erase (Delete)	To delete or erase a behavior mistakenly coded.

Examples of behaviors that should be marked as other:

- Eyes closed for extended period of time (to indicate that the driver *may* be falling asleep. Look for eyes closed more than 2 seconds.)
- Reading

Review Process

- Review the first ten triggers that reduced by a reductionist.
- Provide feedback to reductionist on their post-alert behavior reduction
- When all reductionists have had their first 10 triggers reviewed, review 10 percent of the triggers for each driver

APPENDIX E. BIBLIOGRAPHY OF RELEVANT RESEARCH AND LITERATURE

REPORTS ON THE DETECTION OF DROWSY DRIVERS

Ayoob, E.M., Grace, R., & Steinfeld, A. (2005). ***Driver-Vehicle-Interface (DVI) Development of a Drowsy Driver Detection and Warning System for Commercial Vehicles.*** **Technical Report No. CMU-RI-TR-05-46. Robotics Institute, Carnegie Mellon University.**

This report details the creation of the interface for the prototype DFM used in this experiment. The report includes a general background on driver drowsiness, the considerations of designing warnings and alerting systems for the commercial vehicle driver, and results of focus groups examining specific characteristics of a potential DDWS interface.

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

This report describes the results of a laboratory and field study examining the effects of a lack of sleep on driver performance. The researchers used actigraphy to measure sleep obtained during the field study.

Carroll, R.J. (Ed.). (1999). ***Ocular Measures of Driver Alertness: Technical Conference Proceedings*****. Report Number FHWA-MC-99-136. Washington, DC: Office of Motor Carrier and Highway Safety/Federal Highway Administration.**

This is the technical proceedings from a conference held in April 1999 in Herdon, Virginia. The conference was co-sponsored by the Federal Highway Administration Office of Motor Carrier and Highway Safety and the National Highway Traffic Safety Administration Office of Vehicle Safety Research. The conference sought to document current FHWA/NHTSA research on the validity of using ocular measurements of driver alertness, disseminate information on recent technological advances in the domain, and review the feasibility and future of actively monitoring levels of driver arousal.

Dinges, D.F., & Grace, R. (1998). ***PERCLOS: A Valid Psychophysiological Measure of Alertness as Assessed by Psychomotor Vigilance*****. Federal Highway Administration Publication Number FHWA-MCRT-98-006.**

This study attempted to evaluate the reliability and sensitivity of various drowsiness detection methods (including three measures of PERCLOS). In addition, this study examined the effect of auditory and tactile stimuli on alertness and the vigilance state. PERCLOS was demonstrated to be a highly reliable method of assessing driver drowsiness. Additionally, the researchers found that neither auditory nor tactile stimulation was able to positively affect the drowsy driver's level of arousal.

Dinges, D.F., Maislin, G., Krueger, G.P., Brewster, R., & Carroll, R.J. (2005). ***Pilot Test of Fatigue Management Technologies*****. Report No. FMCSA-RT-05-002. Smyrna, GA: American Transportation Research Institute.**

Describes the initial development work and pilot/field test of driver acceptance of and potential benefits from fatigue management technologies. This work examined the use of actigraphy, PERCLOS eye monitoring, lane-tracking, and a steering assistance device. Their findings indicated that drivers were receptive to fatigue management technologies and, if improvements in their performance could be made, would benefit from their use.

Dinges, D.F., & Mallis, M. (1998). Managing fatigue by drowsiness detection: Can technological promises be realized? In L. Hartley (Ed.), ***Managing Fatigue in Transportation*****, London: Pergamon Press, 209-229.**

This is a book chapter describing various methods of drowsiness detection. It is notable for the description of driving as an extended vigilance task. The authors posit that the successful drowsy driver detection system is attempting to detect the onset of the vigilance decrement, which should be one of the most robust effects of sleepiness.

Eskandarian, A., Sayed, R., Delaigue, P., Blum, J., & Mortazavi, A. (2007). ***Advanced Driver Fatigue Research*** **(Technical Report No. FMCSA-RRR-07-001). Ashburn, VA: George Washington University.**

Describes research performed investigating a drowsy driver detection system at George Washington University under contracts from the FMCSA. The system developed only monitors driver's steering wheel movements and was initially developed for automobile drivers. The report concludes that although there were issues with false alarms this approach to detecting driver drowsiness is valid for use in commercial motor vehicles.

Hartley, L., Horberry, T., Mabbot, N., & Krueger, G.P. (2000). ***Review of Fatigue Detection and Prediction Technologies: Final Report*****. (Technical report). Melbourne, Australia: National Road Transport Commission.**

This report reviews the different technologies and systems available for detecting drowsy driving available at the time. This report also examines such systems in the larger context of regulation and enforcement. One notable conclusion from the report is that technological methods of detecting drowsy driving are a "last ditch safety device," and that companies should seek rational fatigue management programs.

Knipling, R. R. (February, 1998). Three fatigue management revolutions for the 21st century. ***Proceedings of the Third International Conference on Fatigue and Transportation*****, Fremantle, Australia.**

A paper discussing the interrelated issues of Hours of Service regulations, outcome-based fatigue management, and driver-performance-based fatigue management. Potential benefits and recommendations are presented.

O'Neill, T.R., Krueger, G.P., Van Hemel, S.B., & McGowan, A.L. (1999). *Effects of Operating Practices on Commercial Driver Alertness*. (Rep. No. FHWA-MC-99-140). Washington, DC: Federal Highway Administration Office of Motor Carrier and Highway Safety.

This report describes an effort to describe fatigue within the context of commercial motor vehicle operators. This study examines the physical requirements of the task and what impacts they have on operator fatigue. In addition, results of a simulator study investigating the impact of on-duty, non-driving activities is presented. The overall findings suggest that performance deteriorates as a function of time-of-day, with slight increases in driver performance after breaks.

Williamson, A., & Chamberlain, T. (2005). *Review of On-Road Driver Fatigue Monitoring Devices*. Unpublished Report. New South Wales, Australia: NSW Injury Risk Management Research Centre, University of New South Wales.

A paper assessing various fatigue detection technologies, their performance, and the issues surrounding their possibilities for implementation. This work divides detection methods into those focusing on the driver's state (physiological measures) and on driver performance (lateral positioning). A summary of currently available devices is included in the report.

Wright, N.A., Stone, B.M., Horberry, T.J., & Reed, N. (2007). *A Review of In-Vehicle Sleepiness Detection Devices*. (Technical Report No. PPR157). Wokingham, UK: Transportation Research Laboratory.

This report details several sleepiness detection devices (both commercial and in development). Devices are grouped into categories of physiological detection (using methods such as EEG, GSR, and eye activity), physical activity detection (using methods such as actigraphy and head movements), behavior model detection, sleep/active cycle models, and combination approaches.

VIRGINIA TECH TRANSPORTATION INSTITUTE REPORTS ON DDWS/DFM/PERCLOS RESEARCH

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., Jermeland, J., & Knipling, R.R. (2006). *The 100-Car Naturalistic Driving Study; Phase II- Results of the 100-Car Field Experiment*. Contract No. DTNH22-00-C-07007 (Task Order No. 06). Washington, DC: National Highway Traffic Safety Administration.

The 100-Car naturalistic driving study was one of the first large-scale attempts to gather naturalistic data from drivers, thereby removing the common complaint of generalizability from laboratory studies. This report details some of the findings of this study and discusses the methods of data collection. These methods served as some of the initial data collection methods later to be adapted for use in the current study.

Dingus, T., Neale, V., Garness, S., Hanowski, R., Keisler, A., Lee, S., Perez, M., Robinson, G., Belz, S., Casali, J., Pace-Schott, E., Stickgold, R., & Hobson, J. (2002). *Impact of Sleeper*

Berth Usage on Driver Fatigue: Final Project Report. Contract no. DTFH61-96-C-00068. Washington, DC: Federal Motor Carrier Safety Administration.

Hanowski, R.J., Nakata, A., Olson, R.L., Holbrook, G.T., Knipling, R.K., Wierwille, W.W., & Koepfle, B.J. (2003). *Drowsy Driver Warning System (DDWS) Field Operational Test (FOT) Task 13 Report: Phase II Test Workplan – Draft*. Contract No. DTNH22-00-C-07007, Task Order No.14. Blacksburg, VA: Virginia Tech Transportation Institute.

Hanowski, R.J., Wierwille, W.W., Garness, S.A., & Dingus, T.A. (September, 2000). *Impact of Local/Short Haul Operations on Driver Fatigue: Final Report*. Report No. DOT-MC-00-203. Washington, DC: Federal Motor Carrier Safety Administration.

Hickman, J.S., Knipling, R.R., Olson, R.L., Fumero, M.C., Blanco, M., & Hanowski, R.J. (2005). *Heavy Vehicle-Light Vehicle Interaction Data Collection and Countermeasure Research Project: Phase 1 – Preliminary Analysis of Data Collected in the Drowsy Driver Warning System Field Operational Test: Task 5, Preliminary Analysis of Drowsy Driver Warning System Field Operational Test Data*. Report Number DTNH22-00-C-07007. Washington, DC: National Highway Traffic Safety Administration.

Knipling, R.R., Olson, R.L., Hanowski, R.J., Hickman, J.S., & Holbrook, T.G. (2004). *Phase I – Preliminary Analysis of Data Collected in the Drowsy Driver Warning System Field Operational Test: Task 2, Analysis Specification Report*. Contract No. DTNH22-00-C-07007 (Task Order No. 21). Washington, DC: National Highway Traffic Safety Administration.

Wierwille, W.W., Hanowski, R.J., Olson, R.L., Dinges, D.F., Price, N.J., Maislin, G., Powell, J.W. IV, Ecker, A.J., Mallis, M.M., Szuba, M.P., Ayoob, A., Grace, R. & Steinfeld, A. (2003). *NHTSA Drowsy Driver Detection and Interface Project – Final Report*. Contract No. DTNH22-D- 00-07007, Task Order 1. Washington, DC: National Highway Traffic Safety Administration.

Wierwille, W.W. (1999a). Historical perspective on slow eyelid closure: Whence PERCLOS? Report No. FHWA-MC-990136, *Ocular Measures of Driver Alertness*, 31-52. Washington, DC: National Highway Traffic Safety Administration.

Wierwille, W.W. (1999b). Desired attributes of a drowsy driver monitoring system and candidate technologies for implementation. Report No. FHWA-MC-136, *Ocular Measures of Driver Alertness*, 130-143. Washington, DC: Federal Highway Administration.

Wierwille, W.W., & Ellsworth, L.A. (1994). Evaluation of driver drowsiness by trained raters. *Accident Analysis and Prevention, 26*(5), 571-581.

Hanowski, R.J., Blanco, M., Nakata, A., Hickman, J.S., Schaudt, W.A., Fumero, M.C., Olson, R.L., Jermeland, J., Greening, M., Holbrook, G.T., Knipling, R.R., & Madison, P. (2008). *The Drowsy Driver Warning System Field Operational Test: Data Collection Methods Final Report*. Unpublished Research Report. Blacksburg, VA: Virginia Tech Transportation Institute.

This series of reports details work performed at the Virginia Tech Transportation Institute (VTTI) developing a practical driver fatigue monitor /Drowsy Driver Warning System (DDWS). VTTI has, over the past decade, been heavily involved with the development of such systems, performing both laboratory and field testing. Note that initial research, development, and testing on such systems occurred at Virginia Tech before the existence of VTTI (see Wierwille & Ellsworth, 1994).

Part II: Driver Acceptance, Fleet Management Acceptance, and Deployment

Heidi D. Howarth
Tara E. Di Domenico
Lawrence C. Barr
Stephen M. Popkin
Kevin Green

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This report was sponsored by the National Highway Traffic Safety Administration under Project Plan Agreement HS99. Dr. Paul Rau of NHTSA's Office of Vehicle Safety Research, Heavy Vehicle/Driver Interaction Division provided oversight, technical advice, and feedback over the duration of the project.

The data used in this report were collected by Virginia Tech Transportation Institute under NHTSA contract DTNH22-05-D-01019, Task Order #18 and was analyzed by the Volpe Center under Project Plan Agreement HW1Q. Dr. Paul Rau was the NHTSA program manager for this project. Bruce Wilson, Andy Lam, Wassim Najm, and Kevin Green from the Volpe Center and Kate Hartman, of the Federal Highway Administration's Joint Program Office, supported this effort.

CHAPTER 1. INTRODUCTION

Successful deployment of vehicle technologies such as a drowsy driver warning system generally require reliable and valid system performance, a projected safety benefit, and also the knowledge that users accept the device and will heed its feedback (in this case, alerts provided by the prototype driver fatigue monitor). Part II of this report addresses DDWS acceptance from the perspective of both drivers and managers, and additionally considers aspects of projected technology deployment.

DRIVER ACCEPTANCE

Past work in the area of user acceptance includes that of Dinges and Mallis (1998), Stearns and Boyle (2002), Whitlock (2002), and Bekiaris, Nikolaaou, and Mousadakou (2004). These efforts specified a number of criteria that should be considered from the point of view of the technology user when evaluating such systems. Building on this information where possible and practical, U.S. Department of Transportation research staff conceptualized a methodology to systematically assess user acceptance when evaluating various new and emerging vehicle technologies. This methodology is largely based on the National Highway Traffic Safety Administration Intelligent Transportation Systems Joint Program Office's report to Congress (1997). It has evolved and been iteratively expanded for various DOT projects involving FOTs.

Using this approach, acceptance depends upon the degree to which a driver perceives the benefits derived from a system as greater than the costs. For example, if a system's safety potential is not perceived to outweigh its potential costs in other areas (e.g., user annoyance regarding system feedback), it is likely that the system will not be purchased, or that it will be purchased, but not utilized. Or, if there is a benefit whereby device use is perceived to enhance safety and driving skill, then there is also the potential danger that users will over-rely on the technology and feel comfortable engaging in riskier driving behavior. It is important that outcomes such as these are considered as part of any comprehensive evaluation of the safety and user acceptance of vehicle technologies.

In our assessment of a DDWS, we sought to determine the extent to which drivers were able to easily use the DFM and understand its functioning, find that it operated as intended and endorse it, and heed the feedback it provided. Ultimately, if drivers do not accept a technology such as this, they may be inclined to ignore its warnings, or even generate additional risk by using the device in an inappropriate manner. Conceptualized in this way, it is critical to understand user acceptance on a number of levels. As such, we have structured our approach to address five broad elements, or objectives: *ease of use, ease of learning, perceived value, advocacy,* and *driver changes*. Evaluation of these elements includes a systematic assessment based on human factors principles as applied from the perspective of the technology user.

FLEET MANAGEMENT ACCEPTANCE AND DEPLOYMENT

Considering management-level perspectives is additionally important to the process of successful deployment of a DDWS. Not only must company management support the use of the technology, but they are also responsible for understanding its potential impact on their employees and company operations, as well as anticipated safety and economic benefits. It is similarly important to take into account the views and goals of the various developers of these devices and those who publicly represent the interests of commercial motor vehicle operators.

RESEARCH QUESTIONS

Based on the objectives of this research effort, 12 areas of interest were deemed necessary to examine driver acceptance, fleet management acceptance, and deployment of a DDWS. Each of these areas was then operationalized as a research question, with individual components serving as subquestions. They are delineated as follows:

Driver Acceptance

Research Question 7 – Ease of Use

The below research questions address DDWS usability, use patterns, and the degree of understanding and tolerance reported by those in the Test Group who experienced device feedback.

> <u>RQ 7.1</u>: Does use of the DDWS create extra demands on the driver, such as added stress or increased fatigue?
> <u>RQ 7.2</u>: For various degrees of fatigue, how often and what duration do drivers require to observe the device in order to understand its output? Additionally, how do assessments of device accuracy change under varying degrees of fatigue? *(Not addressed due to data constraints.)*
> <u>RQ 7.3</u>: To what degree are drivers willing to tolerate false alarms? Also, what is their degree of reliance on the system, and their perception of correct alarms?
> <u>RQ 7.4</u>: To what degree were drivers able to recognize DDWS alerts?
> <u>RQ 7.5</u>: Do drivers understand the DDWS' operational limitations?
> <u>RQ 7.6</u>: To what extent are the DDWS controls easy and intuitive for drivers to use?
> <u>RQ 7.7</u>: What actions do drivers take to improve their alertness based on the warnings they received? Under what circumstances do they take such actions or not take action?

Research Question 8 – Ease of Learning

The below research questions address the degree to which participants believed that the training they received enhanced their understanding of fatigue management, as well as the DDWS and its application.

> <u>RQ 8.1</u>: Were drivers able to retain information about device operation and the meanings and uses of its output?
> <u>RQ 8.2</u>: How much time does it take for drivers to feel proficient with the DDWS and its output? How much time does it take for driver to learn both the capabilities and limitations of the device? *(Addressed using available focus group data.)*

RQ 8.3: Was the training drivers received on the DDWS complete, and was it understandable? Was the fatigue management training drivers received complete, and was it understandable?

Research Question 9 – Perceived Value

The below research questions address the utility of the DDWS in terms of its perceived ability to measure alertness state and details participants' perception of safety, health-related, and data confidentiality concerns as pertaining to the device.

RQ 9.1: How frequently did drivers indicate they received appropriate warnings based on an accurate alertness assessment?
RQ 9.2: What is the degree to which drivers' felt that the DDWS enhanced the effectiveness of their fatigue management program and practices?
RQ 9.3: Do drivers view the DDWS as a liability or invasion of privacy?
RQ 9.4: Did drivers feel that the DDWS effectively decreased instances of fatigued driving, thus keeping driving skills at an appropriate level for safety
RQ 9.5: Do drivers feel that use of the DDWS will have an adverse effect on their health?
RQ 9.6: How do drivers evaluate their driving safety based on use of the DDWS?

Research Question 10 – Advocacy

The below research questions address the Test Group's degree of reported satisfaction with the DDWS in the context of its usefulness, participants' willingness to endorse it, and potential future device use.

RQ 10.1: How satisfied were drivers with the DDWS? How useful did drivers find the DDWS to be?
RQ 10.2: Are drivers interested in having the DDWS purchased for their entire fleet?
RQ 10.3: Are drivers interested in purchasing the DDWS for their truck, or sharing the cost of device with their employer?
RQ 10.4: Are drivers willing to endorse the DDWS to drivers within and outside their own company?

Research Question 11 – Driver Changes

The below research questions address pre- and post-study levels of perceived fatigue, DDWS usage outcomes and the degree to which Test Group participants used device output to alter their behavior during driving and non-driving periods.

RQ 11.1: What are drivers' perceived levels of fatigue/alertness while driving with and without the aid of the DDWS?

RQ 11.2: Did drivers initiate behaviors as a result of exposure to the DDWS, including unexpected uses for the device? If so, what are they?
RQ 11.3: How much time do drivers spend monitoring the DDWS under various degrees of fatigue? Where do drivers reallocate time spent monitoring the DDWS? *(Not addressed due to data constraints.)*
RQ 11.4: Was use of the DDWS feedback, potentially to adjust driving style, associated with health improvements (e.g., altered work/rest cycles)?
RQ 11.5: What behavioral changes may have been brought about as a result of extended exposure to the DDWS?

Fleet Management Acceptance

Research Question 12 – Perceived Driver Acceptance

The below research questions address the degree to which trucking company interviewees expected their drivers to support the use of such an alertness-monitoring device, in addition to their views regarding its various potential advantages and disadvantages.

RQ 12.1: What are managers' personal opinions regarding driver acceptance of the device?
RQ 12.2: What are fleet managers' opinions regarding driver acceptance of the device based on feedback from drivers?
RQ 12.3: Based on information provided at the briefing, what are fleet managers' perceptions of the DDWS' capabilities, advantages, and disadvantages? Does fleet management believe that drivers will approve or disapprove of the DDWS?

Research Question 13 – Safety and Economic Benefit

The below research questions address trucking company interviewees' perceptions of the potential economic benefits of alertness-monitoring technologies, as attributed to increased safety, and how much they would be willing to spend for such a device. Respondents also offered insight pertaining to what types of insurance and federal incentives would make them more likely to recommend the purchase of this technology for their fleets.

RQ 13.1: What are fleet managers' perceptions of the potential economic benefits, as attributed to increased safety, of the DDWS?
RQ 13.2: What federal incentives would fleet management like to have associated with the adoption of DDWS?
RQ 13.3: What insurance incentives would fleet management like to have associated with the adoption of DDWS?
RQ 13.4: How much would fleet management be willing to pay for a DDWS?

Research Question 14 – Impact on Operation

The below research questions address interviewees' perceptions regarding the potential for trucking companies to monitor and use the data collected by such a device, the degree to which company access to device data might influence driver utilization and acceptance, and potential policies concerning required driver behavior following device warnings.

> RQ 14.1: How much training do fleet managers believe is required for their drivers to make the best use of the DDWS? And, are they willing to provide it? *(Not addressed due to data constraints; fleet management interviewees did not directly experience device.)*
> RQ 14.2: Does fleet management plan to monitor DDWS alerts? If so, how do they plan to use this information? What might the influence of this be on acceptance by drivers?
> RQ 14.3: What sort of policy, if any, does fleet management plan to implement regarding required driver behavior following DDWS alerts?
> RQ 14.4: To what extent is fleet management willing to modify their FMPs based on FOT findings? *(Not addressed due to data constraints; fleet management interviewees did not directly experience device.)*

Research Question 15 – Improvements

The below research questions address interviewees' perceptions regarding additional features or other improvements they would recommend for the concept device as it was described to them. Additionally, they were asked to comment on any concerns they could anticipate regarding device performance.

> RQ 15.1: What are fleet managers' concerns regarding performance of the DDWS?
> RQ 15.2: What features does fleet management desire in the DDWS? How much will fleet management pay for these features?
> RQ 15.3: Does fleet management seek other improvements in the DDWS? *(Not addressed due to data constraints; fleet management interviewees did not directly experience device.)*

Deployment

Research Question 16 – Introduction and Price Range

The below research questions address interviewees' perceptions regarding the market for alertness-monitoring devices, and including the general availability of their product, pricing strategies, and views on long-term deployment.

> RQ 16.1: Is there a distinct introductory period anticipated for the DDWS? If so, how long will it last? What will be the introductory price for DDWS? What will be the maximum penetration level of DDWS in trucks (heavy vehicles) during the introductory period? When will suppliers and truck manufacturers consider deployment of the system?

RQ 16.2: What will be the price range of DDWS deployments over the next 15 years? What is the expected rate of DDWS deployment over the next 15 years? Will the DDWS be available as an option for all trucks, or only more expensive models?

Research Question 17 - Perspectives of Trucking-related Organizations

The below research question addresses the interviewee's perceptions regarding advantages and disadvantages of this type of technology, organized labor's expected views, and liability concerns. Additionally, attitudes pertaining to expected utilization, acceptance, and deployment of this technology were assessed.

RQ 17.1: What are organized labor's concerns regarding the DDWS?

Research Question 18 – Additional Activities

The below research questions address perceptions regarding activities that interviewees believed would help promote the deployment of an alertness-monitoring technology, as well as research undertakings they would recommend in the area of drowsy-driving as related to crash avoidance. Additionally, the extent to which trucking fleets have already purchased or considered the use of alertness-monitoring devices was explored.

RQ 18.1: Would there be a need for non-research activities beyond the FOT in order to expedite DDWS deployment, assuming it is desirable to expedite deployment? What else, if anything, should be done in drowsy driver related crash avoidance research?

CHAPTER 2. METHODS

This chapter outlines the methods and procedures used to gather DDWS FOT data pertaining to the driver acceptance elements introduced in Chapter 1. Sections detail participant recruitment, selection, compensation and training, as well as survey development and administration, data handling, and statistical and descriptive analyses. Information is also provided regarding post-FOT focus groups and phone survey efforts.

FOT PARTICIPATION

Study participation was voluntary and occurred from March 2004 through September 2005. Naturalistic data collection took place during the routine driving of a sample of line- and long-haul CMV operators.

Recruitment, Selection and Compensation

Participants were recruited within three trucking companies at several terminal locations throughout Virginia and North Carolina. FOT conductor staff and company management personally recruited drivers for potential study participation. The FOT conductor also distributed recruitment flyers in an effort to obtain additional participants.

Drivers who expressed interest in the study were individually screened by FOT conductor staff members to determine their typical driving schedule, as well as to ensure a suitable level of auditory and visual (at least 20/40 corrected vision) acuity. As part of this process, the FOT conductor administered a pre-participation survey in order to confirm that drivers operated their trucks at night and did not require eyeglasses to do so, as the DFM would not function properly if these two conditions were not met. All drivers selected for FOT participation tested at acceptable levels on the vision and hearing tests, with the exception of one participant who required a hearing aid. The final step in the selection process required the FOT conductor to couple potential participants with the DFM itself to test for proper slow-ramp eye closure detection. Each person sat in front of a static, operational device and opened and closed their eyes while FOT conductor staff monitored its output. Using this procedure, it was determined that the DFM was able to correctly measure eye closure for all drivers who were assessed.

Selected FOT participants were randomly assigned to either the Control or Test Group. The final experimental design prescribed that Test Group participants would experience two weeks of baseline driving, with the DFM monitoring their behavior (i.e., collecting data), but not providing feedback. This was to be followed by nine weeks in the treatment period, where the device was fully functional provided visual and auditory feedback. Those in the Control Group were to drive under baseline conditions for their entire nine-week participation in the FOT (see Part I, Chapter 4 for a full description of the FOT conditions). FOT conductor staff placed the participant with a hearing aid who did not pass the audiometer test in the Control Group so that attention to DFM-generated auditory alerts would not be necessary.

Upon completion of the required FOT screening and training sessions (see Pre-FOT Training, below) participants received $50. Additional compensation for study participation included $75 per week of FOT data collection. Participants were rewarded with a $250 bonus for successful completion of the FOT, including participation in the entire FOT data collection period, completing and returning all paperwork and surveys, and returning all study-related items provided by the FOT conductor. Payment for participants who terminated data collection in advance of the specified time was pro-rated. Participation in focus group sessions following the FOT was voluntary and compensation was $100.

Pre-FOT Training

A research scientist affiliated with the FOT conductor provided drivers selected for study participation with a two-hour Microsoft PowerPoint training session on fatigue management prior to the onset of the FOT. The session addressed topics such as the effects of fatigue on driver alertness and performance; basic information about sleep, circadian rhythms, and shift work; as well as suggestions for improving sleep and alertness levels, including the use of various fatigue countermeasures. The FOT conductor made an effort to provide each participant with the fatigue management training no more than two weeks before beginning the FOT; however, this was not always possible due to time and hours-of-service (HOS) constraints. For this reason, sessions were frequently offered at the same time as the participant screening, in groups of two to six participants. One-on-one training was also available for participants who could not make it to a group session.

At the start of each training session, the FOT conductor provided participants with a hard copy of the PowerPoint fatigue management presentation for the purposes of note taking. After the presentation, participants were given the opportunity to ask questions and discuss topics of particular interest to them. Once discussions and questioning concluded, those in the Control Group were excused and participants in the Test Group received an additional half-hour of training on the DFM. This training explained the operation of the device, its purpose and components, and how to use the system. Informal final instruction was also provided to Test Group participants as they began the FOT in order to refresh their understanding of the DFM's purpose and operation, in conjunction with a two-page "quick reference" guide that covered basic device operation.

FOT SURVEYS

Four surveys were administered to participants (see Appendix C). These included pre-participation and pre-study surveys (provided prior to/at the start of participation), and post-study and debriefing surveys (provided following participation). Survey items included a combination of Likert-type scales, checklists, yes/no items, visual analog scales (VAS), and open-ended responses.

Survey Development, Administration and Screening

The pre-participation survey was used to obtain general demographic information, as well as confirm that drivers operated their trucks at night and did not require eyeglasses to do so. The pre-study survey solicited self-report information regarding additional participant demographics, attitudes toward the DDWS (Test Group only), fatigue, health and well-being, driver work schedules, and the driving environment. The post-study survey reframed many pre-study survey items in light of FOT participation (and exposure to the Active device; Test Group only), thus allowing for the possibility of comparing responses prior to and following the study. The debriefing survey afforded an opportunity for participants to respond regarding their study participation experience and to document their reactions to the DDWS.

The FOT conductor pilot-tested draft versions of each FOT driver acceptance survey with CMV operators who underwent identical screening and training procedures as future FOT participants and then completed a 10-hour data collection run in order to experience the DFM. Pilot study participants volunteered verbal feedback, suggesting that portions of the draft surveys were somewhat difficult to understand because the wording was “too academic” and “not driver-friendly.” Based on this information, the researchers contracted to assist in the design and analysis of the FOT surveys iteratively revised the draft versions to make them more amenable to the FOT target audience. Calculations of scores on the Flesch-Kincaid readability measure, as provided by Microsoft Word, were adjusted to a high school reading level ($\leq 12^{th}$ grade) for each of the finalized surveys.

For the official study, the FOT conductor administered pre-participation surveys per the participant recruitment and selection procedures described above. Once selected for participation, at the onset of data collection, FOT conductor staff provided pre-study surveys to Control and Test Group participants for completion. Most participants finished this activity on site; however, some required additional time to do so and were therefore instructed to return the survey to FOT conductor staff prior to beginning the FOT. Upon completion of the FOT data collection period, the FOT conductor administered a post-study survey to participants. Thereafter, at the conclusion of their study participation, participants received the debriefing survey.

FOT conductor staff visually scanned participant surveys upon receipt to help ensure that all questions had been answered. If a missing response was noted, it was requested that the participant complete the items or otherwise indicate a non-response. Additionally, the FOT conductor performed data verification procedures on post-study surveys that included screening for fixed-responses and verification of consistent attitudes for certain, related VAS items. If a discrepancy was detected, FOT conductor staff requested that the participant verify or otherwise explain the intended response.

Data Receipt and Safeguards

The FOT conductor maintained the confidentiality of all survey data, identifying participants using numerical codes. As they became available, photocopies of completed, original surveys and VHS recordings of focus group sessions were mailed to the research staff responsible for their analysis. The receipt of each survey copy from the FOT conductor was logged upon arrival. Participant survey responses were coded using data reduction forms and subsequently entered into SPSS for statistical analysis. At each stage of these data management and data entry procedures, verification was performed by an additional member of the research staff to ensure accuracy.

Data Analysis

The FOT survey data analyses were structured using variables specified in an a priori data analysis plan that mapped relevant items to the various driver acceptance elements. The data analysis strategy included procedures for descriptive investigation of quantitative and qualitative data. As appropriate, summary statistics (e.g., response distribution, mean, median, standard deviation) were calculated for quantitative survey data. Results were plotted graphically in various formats and used to present information when this would add value to the clarity of the results. In addition, where possible, and where meaningful and informative comparisons existed, independent-samples nonparametric analyses between Control and Test Groups were performed, as well as paired samples, nonparametric comparisons of responses to congruent pre- and post-survey items (see Appendix A).

The qualitative information contained in responses to open-ended survey items and anecdotal information obtained from focus group sessions were summarized (see Focus Groups section, below) to supplement the outcomes of the quantitative DDWS acceptance data. Particular anecdotes and anonymous quotations offered by participants were incorporated into report sections to the extent that they were representative and appropriate.

FOCUS GROUPS

At the culmination of the FOT, focus group session participation was offered to those who were part of the Test Group and did not terminate their study participation early, to help ensure similar periods of exposure to the DFM. The FOT conductor performed focus group recruitment by phone. Those who agreed to participate were scheduled for one of two sessions and called by the FOT conductor a few days prior as a reminder to help ensure full participation. In total, 14 males participated (one session contained 8 participants, the other contained 6).

Focus Group Materials and Sessions

In a manner similar to that used to develop the FOT surveys, research staff members created a discussion guide and corresponding PowerPoint slides to facilitate the focus group process. Developmental iteration included inputs from the FOT conductor staff and the focus group facilitator. The materials were designed to obtain additional information pertaining to each of the five driver acceptance elements (see Chapter 1). A semi-structured format allowed the facilitator to selectively probe for information in areas that researchers required more information, or that were not fully addressed by the FOT surveys.

Focus group sessions were hosted in conference rooms at two different FOT-participating fleet terminals. The first session contained long-haul drivers; line-haul drivers made up the second session. Prior to beginning discussions, participants read and signed informed consent forms permitting audio/video recording for the purposes of ensuring accurate information. FOT conductor staff and a driver acceptance research staff member (via telephone) supported the focus group facilitation process. The FOT conductor staff member present at the session ensured that all participants had equal opportunity to express their opinions. Each focus group session lasted approximately two hours.

PHONE SURVEY: FLEET MANAGEMENT ACCEPTANCE

A research staff member conducted telephone interviews to obtain attitudes and perspectives pertaining to fleet management acceptance of a DDWS. Because the interviewees were not familiar with the particular DDWS used for this FOT, the interviewer provided them a description of the device, framed as a “conceptual” vehicle-based alertness-monitoring technology.

Survey Methods

A list supplied by the FOT conductor provided contact information for management-level staff at 20 for-hire and 20 less-than-a-truckload trucking companies. Nine of 24 randomly selected companies agreed to participate in the phone survey, and consisted of 5 long-haul, 3 short-haul, and 1 leasing supplier of trucks for both long- and short-haul operations. Interviewees included management staff in the following positions: four safety managers/directors, two fleet managers, one company president, one co-owner, and a maintenance director.

Prior to the interview, the interviewer provided a brief conceptual description of a DDWS (note that this intentionally described the DFM) through e-mail or over the phone. The following summarizes the description participants received:

The device mounts on the truck’s dashboard, to the right of the driver, so as not to block the forward view or interfere with any other equipment in the cab. A stationary camera at the top of the device uses technology that directs minimal levels of harmless infrared light towards the driver’s face and eyes. To the extent that the driver’s eyes are “in view” of the device camera, it processes data in real time to estimate “percent eye closure” during periods when the truck is traveling above 35 miles-per-hour and the cab is dark. If percent eye closure estimates are greater than a predetermined level, this indicates that the driver’s eyes may be closed, or nearly closed, thus suggesting a potential lapse in alertness. If such an event occurs, the device issues an audible warning tone to the driver. After this alert, if percent eye closure measures continue to rise, the tone sounds repeatedly and red warning lights illuminate and flash until the driver responds by pressing a button on top of the unit. Once this button is activated and the warning silenced, the device presents a bar-graph-like display showing the duration of the longest eye closure, as well as a numeric display indicating the total number and the rate-per-hour for warnings received (see Part I, Chapter 2 for a detailed description of the prototype DFM system).

A 16-item, semi-structured phone interview questionnaire (see Appendix D) developed by the research staff assessed trucking company management attitudes regarding expected approval or disapproval of the conceptual DDWS by drivers, its potential for use at their company, and possible economic benefits related to insurance and other incentives. The interview process lasted between 20 and 35 minutes. Participants were ensured of the confidentiality of all information provided and that they would not be individually identified in any publication of survey outcomes.

PHONE SURVEY: DEPLOYMENT

A member of the research staff conducted telephone interviews (see Appendix D) and in one case, solicited information via email, from technology developers and a trucking advocacy organization. Responses provided insight regarding the introduction and deployment of existing alertness-monitoring systems, as well as more general views regarding such technologies.

Survey Methods

Representatives from two companies who are actively involved in the development and deployment of unobtrusive, vision-based, alertness-monitoring systems were interviewed over the phone to solicit their views on current and future deployment efforts and pricing strategies for their products. Both companies have been developing alertness-monitoring technologies for nearly 10 years. A representative from the Safety and Operations Office of the American Trucking Association was also interviewed to obtain the perspective of a U.S. trucking advocacy organization on alertness-monitoring devices and their potential for use in commercial vehicles. Questions were e-mailed to the ATA representative due to phone interview scheduling difficulties, and requested similar information as was asked of company managers (see Chapter 4 and Chapter 5).

CHAPTER 3. DRIVER ACCEPTANCE RESULTS

Results are reported using the available survey data for each analysis. Characteristics of the final driver acceptance survey sample, as influenced by its small overall size and further split by subgroups of unequal size, in most instances made descriptive reporting the most appropriate statistical approach. Additionally, survey item response distributions were generally non-normal (either skewed or bimodal). For this reason, findings are most often described in terms of the response distribution, as it is more meaningful and representative than providing means or other measures of central tendency. In cases where the data permitted between-group comparisons or pre- and post-FOT paired-comparisons, we conducted nonparametric statistical analyses. However, as only one comparison yielded a significant result (see Ease of Use section), the non-significant outcomes of the remainder of the tests are provided in Appendix A.

PARTICIPANTS

Participant recruitment and screening procedures, as outlined in Chapter 2, resulted in an initial FOT sample of 102 drivers (101 male, 1 female). Additionally, one driver who wore glasses was used to test the DFM under conditions outside of its intended operational parameters. As an outcome of the inherent challenges to naturalistic data collection and a field test of this magnitude, a subsequent set of participant selection criteria were required to help ensure the integrity and meaningfulness of the planned driver acceptance analyses and outcomes.

Conducting the planned driver acceptance pre- and post-FOT attitudinal comparisons necessitated a uniform participant sample. As such, requirements for inclusion in the driver acceptance survey data sample encompassed the following:

1. Acceptable screening outcomes based on the pre-participation survey (and DFM screening tests)
2. Available pre-study survey data
3. Available post-study survey data
4. No exposure to DFM Active mode prior to completion of baseline data
5. No documented pre-study survey completion date after exposure to DFM Active mode
6. No documented pre-study survey completion date more than three days after FOT start date.

These screening criteria yielded a final driver acceptance participant sample of 48. The Control Group consisted of 15 participants and the Test Group consisted of the remaining 33 participants.

Demographics

A summary of self-reported demographic information for the driver acceptance participant sample is reported by group in Table 112. Focus group participants were not identified beyond the information provided in Chapter 2 and are therefore not included in the table.

Table 112. Participant Summary Demographic Information

Demographic Variable	% Response		
	Full Sample (*n* = 48)	Control Group (*n* = 15)	Test Group (*n* = 33)
Gender			
Male	97.9	100	97.0
Female	2.1	--	3.0
Marital Status			
Married or living with partner	72.9	73.3	72.7
Single or widowed	14.6	13.3	6.1
Divorced or separated	8.3	13.3	15.2
Missing	4.2	--	6.1
Ethnic Background			
Caucasian	64.6	66.7	63.6
African American	27.1	20	30.3
Native American	4.2	6.7	3.0
Asian	2.1	--	3.0
Missing	2.1	6.7	--
Highest Achieved Education Level			
Did not complete high school	10.4	6.7	18.2
High school graduate	50.0	46.7	42.4
Beyond high school	37.5	46.7	33.3
Missing	2.1	--	6.1
Haul Type			
Long	75.0	86.7	69.7
Line	25.0	13.3	30.3
Trucking Company			
J.B. Hunt	41.7	46.7	39.4
Howell's	37.5	46.7	33.3
Pitt-Ohio Express	20.8	6.7	27.3
	Mean (*SD*)		
Age (years)	41.1 (8.5)	38.7 (8.7)	42.2 (8.3)
CMV driving experience (years)	11.67 (8.2)	9.4 (6.9)	12.7 (8.7)

EASE OF USE

This section addresses DDWS usability, use patterns, and the degree of understanding and tolerance reported by those in the Test Group who experienced device feedback.

Survey Results

Usability of the DDWS was assessed using several items in the post-study survey. When asked whether the location of the DDWS within the cab was acceptable, 90.9 percent of the sample responded "yes." The mean rating for a VAS item regarding how easy it was to "read the DDWS display while driving" (0 = not at all; 100 = extremely) was 76.9 (*SD* = 23.9). Another VAS item measured the reported degree of driving ease or difficulty when using the DDWS (0 = very difficult; 100 = very easy) and resulted in a mean rating of 80.5 (*SD* = 23.2).

A number of post-study survey items investigated device use patterns. A checklist of driving scenarios, as depicted in Figure 160, indicates that participants were most likely to utilize DDWS warnings during times when they felt very tired (52.1%) and, to a lesser extent, when they were experiencing very little traffic or poor driving conditions.

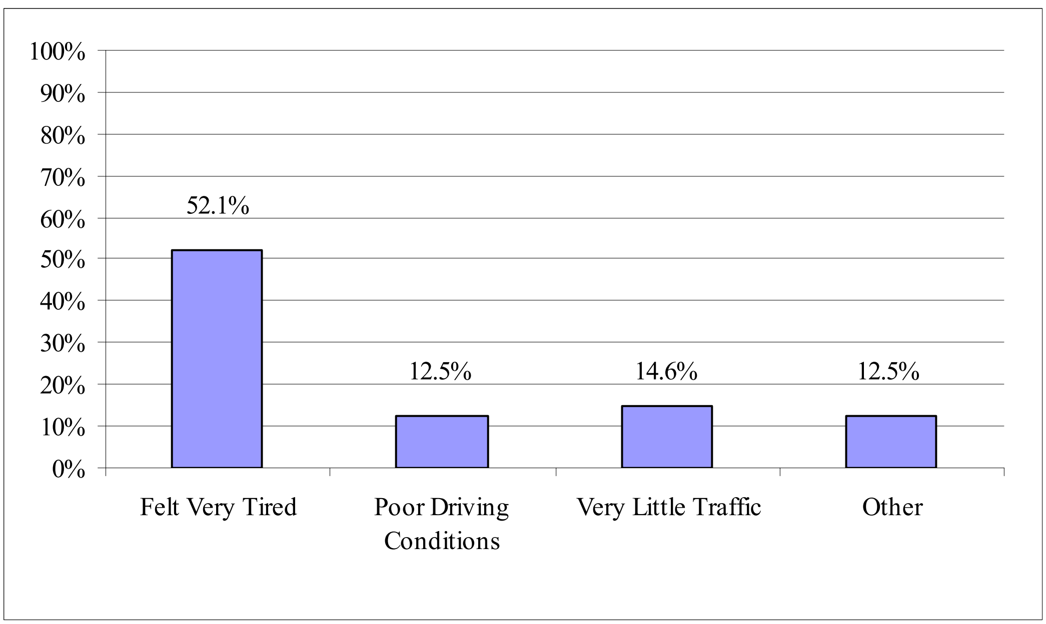


Figure 160. Response Distribution for Scenarios When Participants Were Most Likely to Use DDWS Warning Information (Check All That Apply)

A VAS item regarding device usage habits asked about the degree to which participants relied on the DDWS for warnings about their level of fatigue (0 = not at all; 100 = extremely). The mean response for this item was below the scale midpoint at 39.6 (*SD* = 29.9). Figure 161 shows the distribution of participant ratings, where responses from 45.5 percent of the sample fell within the bottom third of the scale range. A related VAS item assessed device over-reliance. Participants responded to the statement: "Overall, I found myself relying too much on the

DDWS device" (0 = not at all; 100 = extremely). The mean rating for this item was 22.8 (*SD* = 26.0), with 78.8 percent of all participants responding within the bottom third of the scale range.[1]

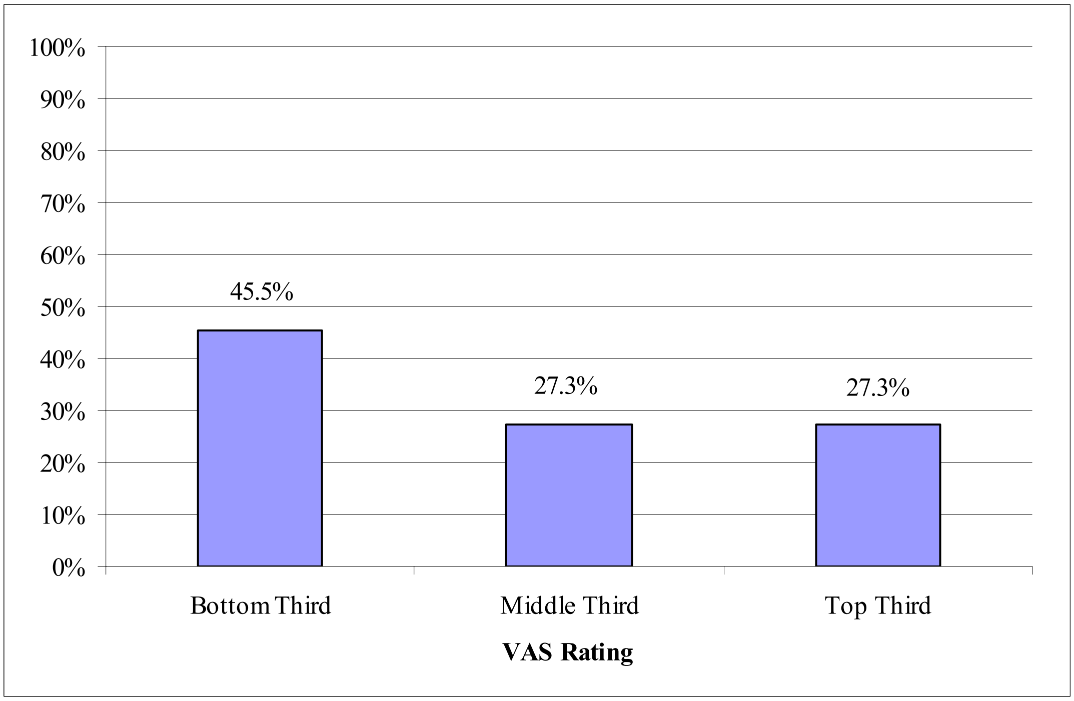


Figure 161. Distribution of VAS Ratings Regarding Degree of Reliance on DDWS to Warn About Fatigue Level (0 = Not at All; 100 = Extremely)

Survey items also assessed participant responses to device warnings. Upon receiving an alert, 39.4 percent of the sample reported "almost never" ignoring it, and another 42.4 percent indicated "occasionally" ignoring the warning. Separate, parallel survey items asked how often drivers were "unwilling" or "unable" to take actions to improve their alertness once the DDWS provided a warning. Approximately half of the participants in each case indicated "almost never" being unwilling (51.5%) or unable (45.5%) to take action. Smaller proportions of the sample reported being "occasionally" unwilling and unable to act, at 30.3 percent and 39.4 percent respectively. Situations when participants most frequently did not take action to improve their alertness after receiving a warning were provided in survey responses using an open-ended format. Generally, these encompassed instances where the device was perceived to be malfunctioning (e.g., false alarms, dawn, dusk) and when participants were already close to, or late to arrive at, their destination or stopping point.

The response distribution rating usefulness of the DDWS for the purpose of fatigue management while at work was assessed prior to and after device exposure using survey items with parallel construction (see Figure 162). As depicted, there was a shift in attitudes; the highest frequency of responses prior to experiencing the DDWS indicated that participants felt the device would be

[1] As a caveat pertaining to this type of data, it should be noted that humans do not tend to perform well in detecting their own levels of drowsiness (Dinges, 1989).

“quite useful,” while after exposure to the device, they most often reported that it was only “a little useful.” A Wilcoxon signed-ranks analysis resulted in a significant difference between the assessments of device usefulness made prior to exposure to the device and thereafter ($z = -3.11$, $p = 0.002$; see Appendix A).

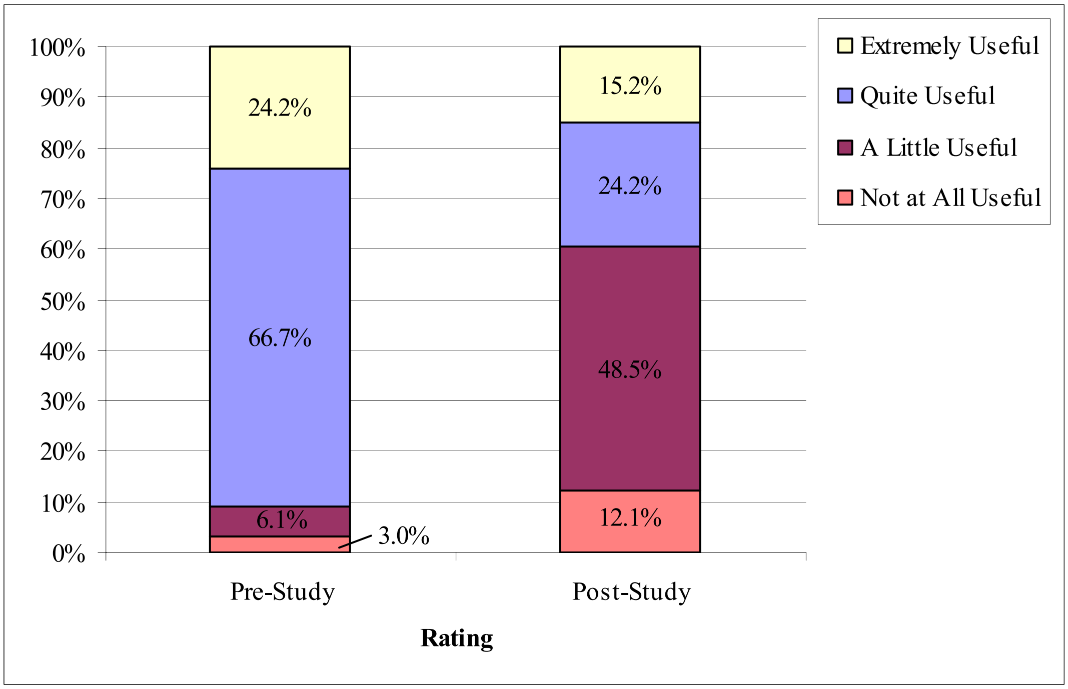


Figure 162. Response Distribution for Usefulness of DDWS for Fatigue Management Purposes While at Work (1 = Not at All Useful; 4 = Extremely Useful)

Two related post-study survey items measured the degree to which participants felt that the DDWS was operating properly when it provided warnings. When asked how often the device correctly measured alertness, the highest response frequency was in the category of “occasionally,” while for incorrect fatigue warnings, most responses fell into the categories of “occasionally” and “frequently” (see Figure 163).

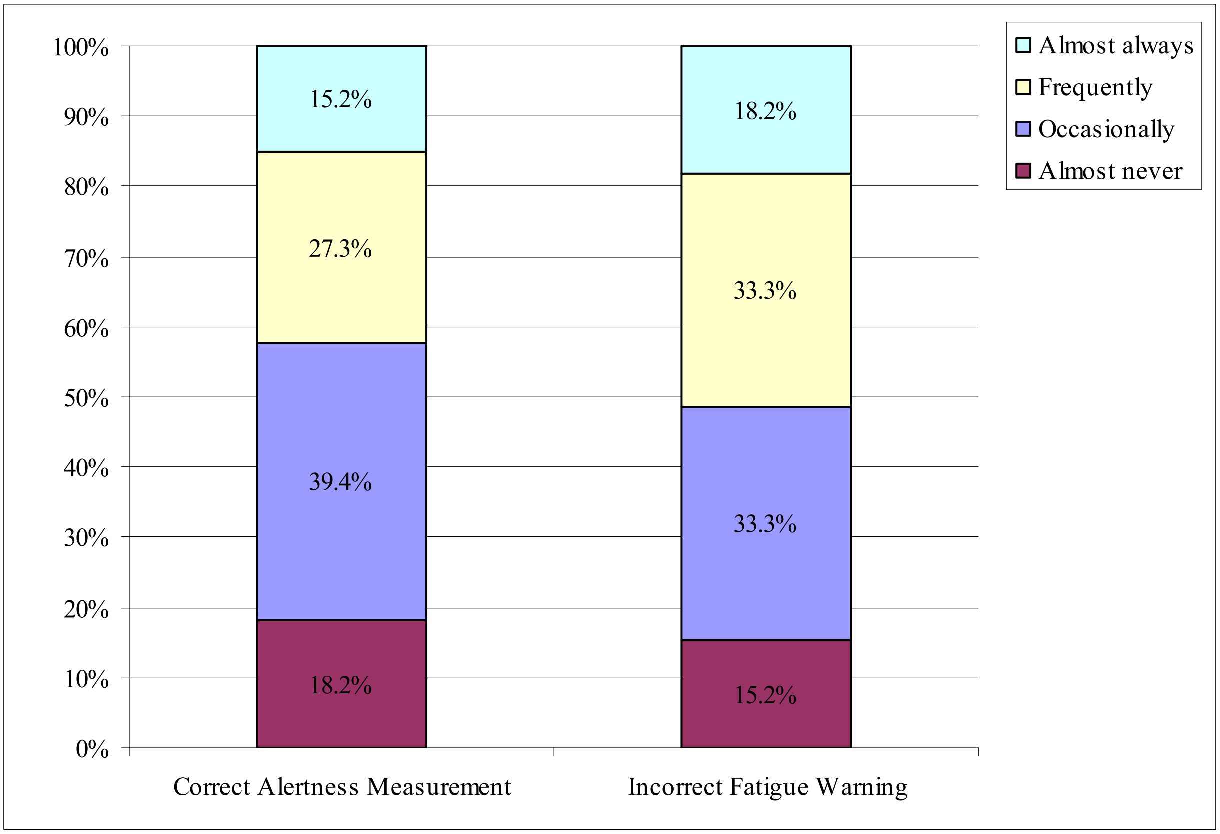


Figure 163. Response Distribution for How Often DDWS Correctly Measured Alertness and Incorrectly Provided Fatigue Warnings (1 = Almost Never; 4 = Almost Always)

In addition to measuring the frequency of device false alarms, survey items also assessed participants' tolerance of these warnings. A total of 81.8 percent of the sample indicated that they were aware of situations during the study where the DDWS did not operate as it should have. Additionally, the number of incorrect warnings provided by the device per duty period was requested in an open-ended survey item. This resulted in a reported mean of 12.4 false warnings per duty period and a range of responses from 0 to 100. The mean reported annoyance with such incorrect device warnings was 62.6 (*SD* = 33.1), as measured using a VAS (0 = not at all; 100 = extremely). Nevertheless, for another survey item, two-thirds of the sample indicated that they did not stop relying on the DDWS as a result of false alarms.

Items in the post-study survey gauged participants' understanding of device warnings and their ability to utilize the device (note that the brightness of the display, and the type and volume of the alert received, were adjustable and may have influenced the effectiveness of warnings. However, information regarding specific user-adjustments was not captured in the surveys (see Part I, Chapter 4, Research Question 4 for specific findings on device adjustment). A VAS survey item assessed the degree to which participants were able to easily recognize device alerts and resulted in a mean of 82.9 (*SD* = 21.9; 0 = not easily; 100 = very easily). A similar VAS item measured the extent to which the DDWS caught participants' attention quickly when it provided an alert (0 = strongly disagree; 100 = strongly agree). The mean level of agreement with this statement was 80.7 (*SD* = 23.9).

As shown in Figure 164, the main issue with DDWS use, when provided a checklist of responses, was reported as "hearing the warning." The second most frequently reported category was "other," which included written-in responses regarding issues with false warnings, especially during dusk and dawn. Results for a related survey item indicated that understanding device functioning in various situations was not a problem for 72.7 percent of the sample. For those who did report instances of not understanding device operation, situations provided in an open-ended response were similar to the above and included actual mechanical malfunction, as well as difficulties with false alarms.

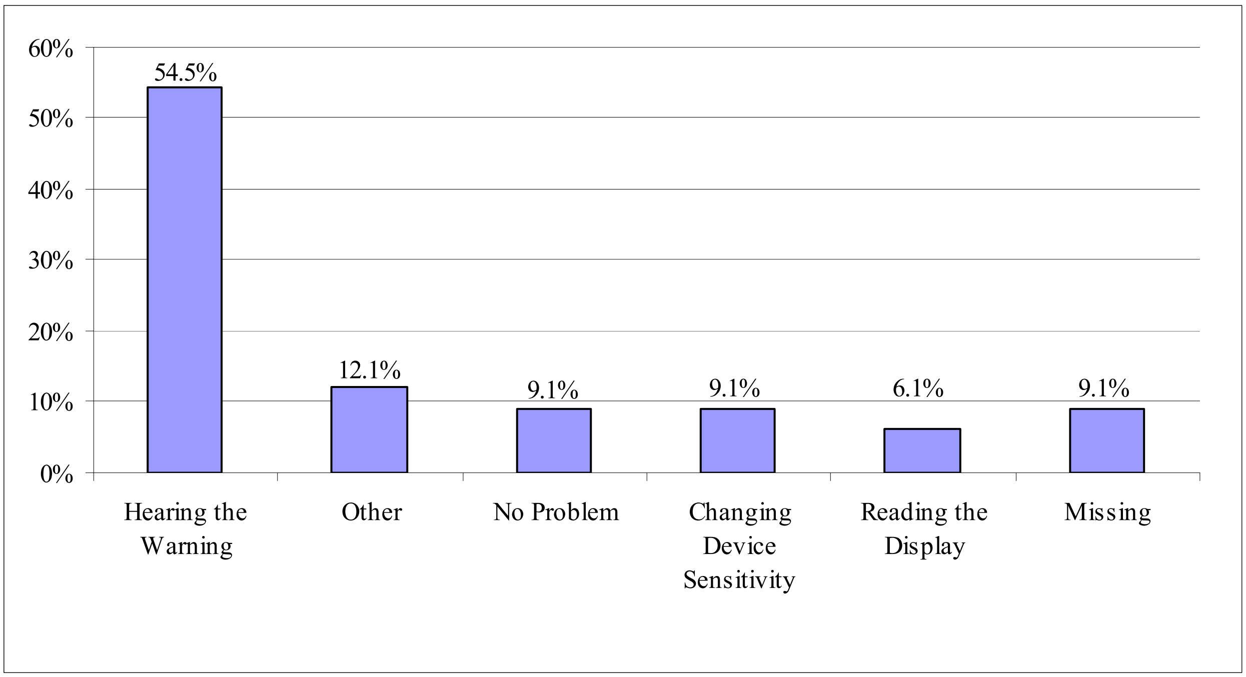


Figure 164. Reported "Biggest Problem" When Using DDWS (Check One)

Focus Group Results

Focus group participants did not report major usability issues with the DDWS itself, to the extent that its limitations had been explained prior to the onset of the FOT. Suggestions for possible additional improvement of system usability included the following: lighting the dials during night driving to make user adjustments easier; creating a wider range of camera operation to minimize losing track of a driver's eyes as a result of in-cab movement, especially when checking side mirrors; and moving the location or reducing the size of the device to make it less obtrusive and avoid occluding mirrors. Pertaining to the auditory alerts, one participant commented that he particularly liked the sound because it was very alerting, while another mentioned that he was sometimes unable to hear the warning due to the CB or other noises in the cab. Contextually, it was not clear whether these participants were aware of, or had adjusted, the DDWS auditory warning sound or volume controls.

Participants reacted to the device in different ways when receiving warnings that they perceived to be legitimate. Most focus group members reported that they sometimes found it difficult to comply with warnings due to limited truck stop locations and problems with parking availability. However, at least one focus group member indicated that he always found a place to stop. False alarms were found to be distracting by some, especially as two focus group participants reported receiving them about once every ten minutes, necessitating numerous device resets. Participants handled warnings that were not believed to be legitimate, especially during dusk and dawn when the device provided the greatest number of false alarms, by employing a variety of techniques. One method was to turn down the volume so that alerts were barely or, according to one focus group member, not at all audible. Additionally, in one case, a participant recounted putting something over the device's camera; however, he also admitted that this technique did not appear to solve the problem. Despite reported attempts to decrease the number of false alarms in these ways, some participants were quick to agree that there should not be a way to lower the volume of the warnings at all. They acknowledged that the purpose of the device is to alert the driver, and when the volume is turned down, warnings cannot be heard well enough to do so.

Discussion

With regard to DDWS usability findings, Test Group participants largely reported that the device was easy and intuitive to use while driving and did not indicate any major problems, in particular as certain device limitations were explained to them prior to use. Suggestions for improvement offered during the focus group sessions were related to some of these limitations, including increasing the device camera's field of view to help prevent it from losing track of drivers' eyes. In addition, it was suggested that device dials light up to allow for easier adjustments during night driving. Somewhat varied statements made during the focus group sessions regarding auditory warnings were mirrored in survey responses. Comments offered by focus group members indicated that some participants found warnings to be very alerting, while others thought they were not loud enough. In comparison, survey findings indicated that, generally, participants easily recognized alerts and that the alerts quickly attracted their attention. However, at the same time, "hearing the warning" was reported as the main usability issue with the DDWS. These contrasting findings suggest that the warning adjustment features on the device may deserve further investigation, as the degree to which the driver acceptance survey sample adjusted the volume or changed the sound of the warning itself is unclear.

Survey findings indicated that nearly half of participants were not very likely to rely on the DDWS to warn them about their level of fatigue. More specifically, participants reported being most likely to utilize device warnings when they felt "very tired," as opposed to situations related to environmental conditions, such as traffic. The vast majority of participants reported "almost never" or "occasionally" ignoring DDWS feedback and, related to this, approximately half of the sample reported "almost never" being either unable or unwilling to take action based on warnings provided by the DDWS. Examples provided in the surveys regarding situations where participants did not act upon warnings they received largely referred to false alarms and instances when a destination was near, or they were running late. Additionally, focus group members reported that they sometimes found it difficult to take action based on warnings due to limited truck stop locations and parking availability.

Attitudes regarding incorporating DDWS feedback into fatigue management work practices evidenced a statistically significant, negative shift after exposure to device operation during the FOT. Prior to experiencing the DDWS, two-thirds of the participant sample predicted that the device would prove itself "quite useful" for this purpose; however, subsequent to device exposure, nearly half of the participants reported that it was only "a little useful." This shift may be due to reports by over 80 percent of participants who were aware of situations during the FOT where the DDWS did not operate properly. Additionally, over half of the sample indicated that they felt the device was "almost never" or "occasionally" correct regarding their alertness level. Accordingly, false warnings, especially during dusk and dawn, came to be expected by participants and reported annoyance with them was moderate. Nevertheless, two-thirds of the sample indicated that perceived device false warnings did not cause them to stop relying on the DDWS.

EASE OF LEARNING

This section documents the degree to which participants believed that the training they received enhanced their understanding of fatigue management, as well as the DDWS and its application.

Survey Results

Responses to two VAS survey items addressed the perceived completeness of the DDWS and fatigue management training received by participants. The items used parallel construction in their wording and identical scale anchors, ranging from 0 (not at all) to 100 (extremely). Results indicated that the mean completeness rating for the DDWS training received by the Test Group was 81.8 (SD = 15.5). The mean completeness rating for fatigue management training was 67.6 (SD = 25.2) for the Test Group and 69.0 (SD = 20.8) for the Control Group. Figure 165 shows the percent of participant responses above and below the scale midpoint for these two survey items.

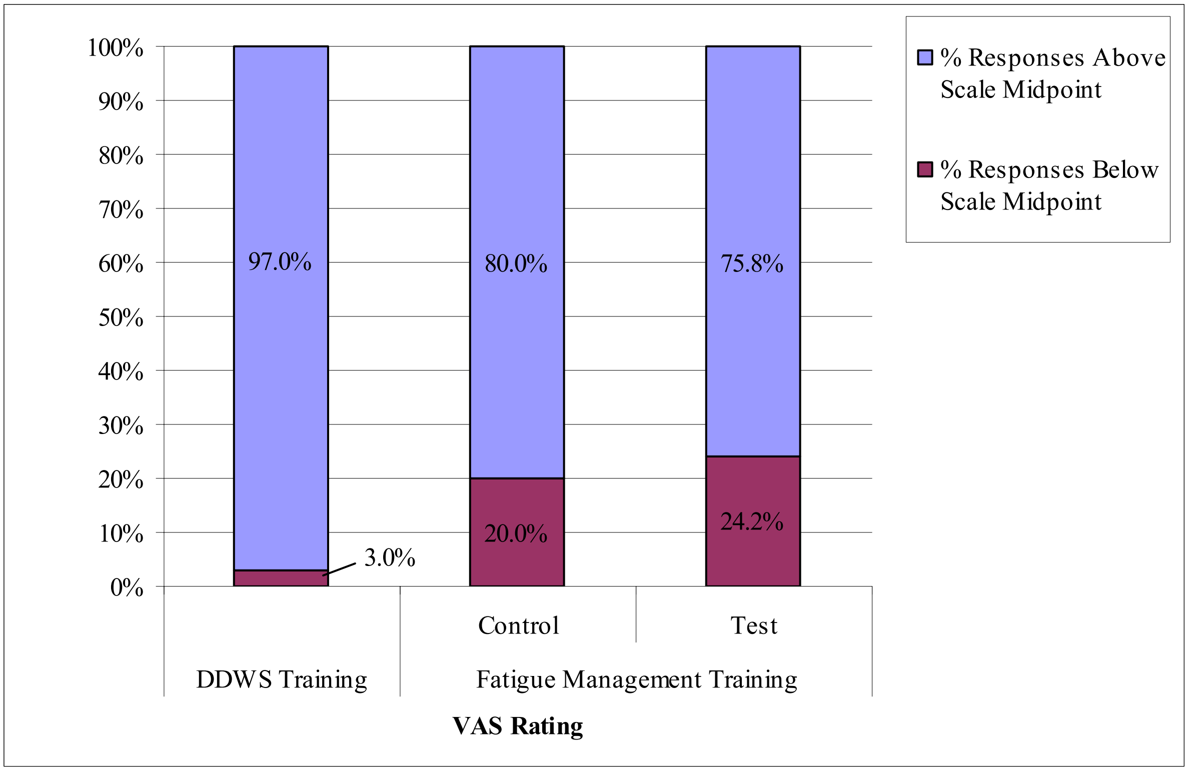


Figure 165. Distribution of VAS Completeness Ratings for Training on DDWS and Fatigue Management (0 = Not at All; 100 = Extremely)

Another survey item asked whether participants could "easily understand" the written information they received regarding the DDWS. Results indicated that none of the participants had difficulty with the materials (i.e., 100% of the sample responded "yes"). In a related item, a single participant would have preferred training information written in Spanish.

Focus Group Results

Focus group participants indicated that the DDWS training they received was straightforward and useful. In particular, participants found the two-page quick-reference guide helpful during initial device exposure at the onset of the FOT. Furthermore, for those who experienced a lag between training and FOT data collection, the guide proved to be particularly useful for the purposes of reviewing DDWS features and usage instructions. One participant likened his experience to "…turning your VCR on. After 10 minutes you've figured out what it's [going to] do and what it's not [going to] do." As an area for improvement, some reported that the training was too brief and suggested providing "hands-on" learning in the form of live demonstration or a video in order to further enhance understanding of the device and its operation.

Discussion

Results suggest that the device training and related fatigue management training were effective and well received, overall. Nearly the entire Test Group found information about the DDWS easy to comprehend. Moreover, as an outcome of focus group discussions, there was a particular emphasis on the utility of the quick-reference guide for understanding device operation, with additional suggestions to incorporate this information into a live or video-based demonstration.

For both types of training, completeness-ratings for over 75 percent of all groups fell above the scale midpoint (higher scores indicating greater completeness), though mean ratings for the fatigue management training were lower than those for the device training. A review of the content of both sessions suggests the possibility that scores for fatigue management training completeness might have been higher had Test Group course materials been tailored to the FOT and in particular, linked with DDWS output, as opposed to stand-alone information about sleep, fatigue, and fatigue countermeasures. An example of such a linkage would be the following statement: “When you receive [insert description of warning type; e.g., auditory, or auditory plus visual], the device has detected [insert fatigue severity prediction based on type of warning] and you should consider [insert fatigue management countermeasure action] as soon as you are safely able to do so.” Providing this level of prescriptive information may, in turn, have aided drivers to a greater degree when making fatigue management decisions based on device feedback, especially as pertaining to which countermeasures are most appropriate and effective in various situations. However, to some degree, tailoring the fatigue management training towards the prototype device’s output would have complicated comparisons between the Test and Control Groups, as they no longer would have received the same fatigue management training.

PERCEIVED VALUE

This section assesses the utility of the DDWS in terms of its perceived ability to measure alertness state and details participants’ perception of safety, health-related, and data confidentiality concerns as pertaining to the device.

Survey Results

A series of survey items assessed the prevalence of fatigue and fatigue-related safety outcomes in the current sample, prior to and after the FOT (and exposure to DDWS feedback in the case of the Test Group). In the pre-study survey, a total of 42.4 percent and 53.3 percent of the Test and Control Group participants, respectively, answered “yes” when asked whether they had “ever, even for a moment, fallen asleep behind the wheel.” In a parallel post-study survey item, a similar percentage resulted for the Test Group (39.4%) and a somewhat lower percentage responded affirmatively in the Control Group (40.0%). Regarding collisions, prior to exposure to the DDWS, 13.3 percent of the Test Group participant sample and 15.2 percent of the Control sample indicated having been involved in an accident or incident while working that they felt was related to sleepiness or fatigue. In a related post-study survey item asking whether the DDWS helped to avoid an accident or close call when it was turned on during the FOT, 27.3 percent of the Test Group replied “yes.”

A post-study survey item measured perceived device accuracy in terms of how often the DDWS correctly measured alertness by providing appropriate warnings. As shown in Figure 166, combined responses for the two most utilized scale options indicate that 66.7 percent of the Test participant sample felt that the DDWS “occasionally” or “frequently” correctly measured alertness by issuing a warning.

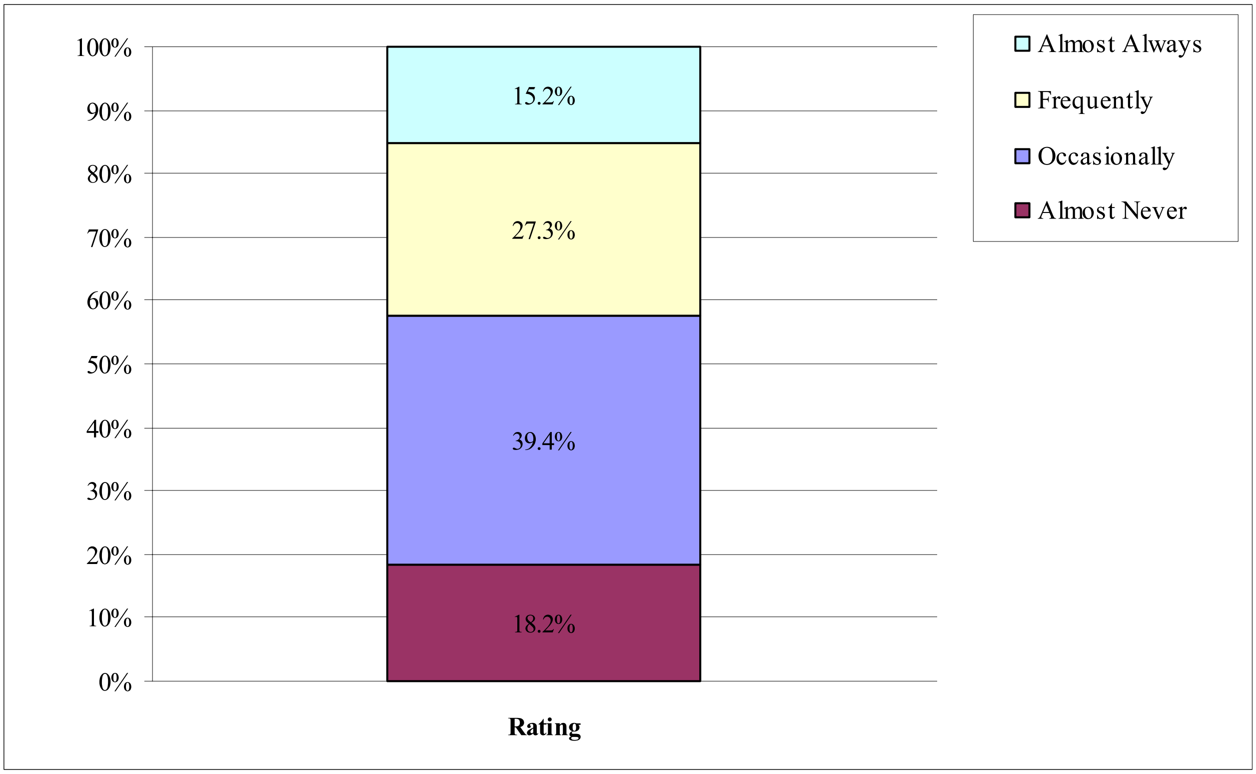


Figure 166. Response Distribution for Perceived Accuracy of DDWS Alertness Measurement (1 = Almost Never; 4 = Almost Always)

Two VAS survey items used parallel construction in their wording to assess device utility with regard to predicted likelihood of being involved in a fatigue-related accident when using the DDWS compared with not using it. As depicted in Figure 167, the majority of responses in both cases fell within the bottom third of the scale range, where lower scores indicate lower predicted accident likelihood. Comparatively, fewer Test Group participants rated their perceived likelihood of accident involvement as high (i.e., using the top third of the scale) when using the DDWS (6.1%) than without the aid of the device (15.2%).

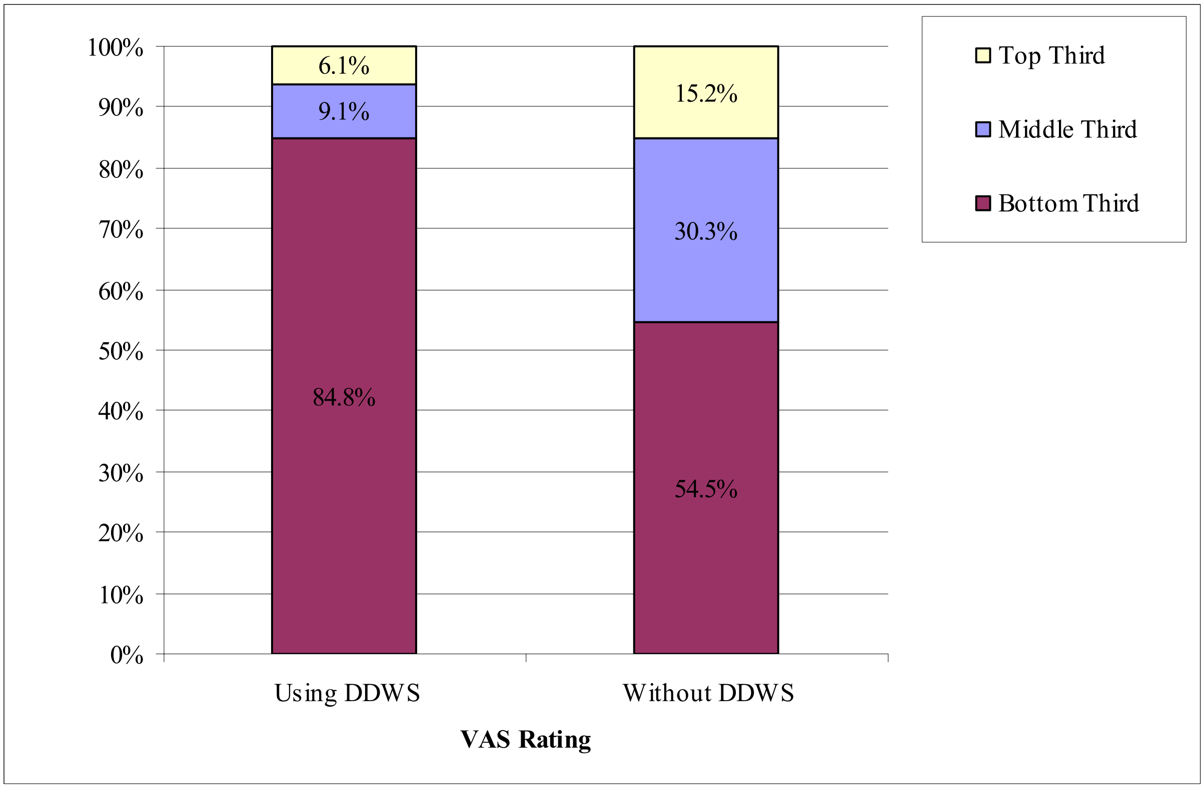


Figure 167. Distribution of VAS Likelihood Ratings for Predicted Involvement in a Fatigue-Related Accident Using the DDWS Compared to Without the DDWS (0 = Not at All; 100 = Extremely)

A number of survey items addressed the utility of employing DDWS driver state information for fatigue management purposes. For one post-study survey item, 81.8 percent of Test Group participants affirmatively indicated that the information provided by the device was "useful for managing [their] fatigue." Results of a similar post-study VAS survey item assessing the degree to which the DDWS was useful, in addition to other steps taken to manage fatigue, are shown in Figure 168. The mean rating was close to the scale midpoint, at 52.4 (*SD* = 30.3), however the response distribution also deserves consideration. As depicted, over 75 percent of all responses fell within the top and bottom thirds of scale.

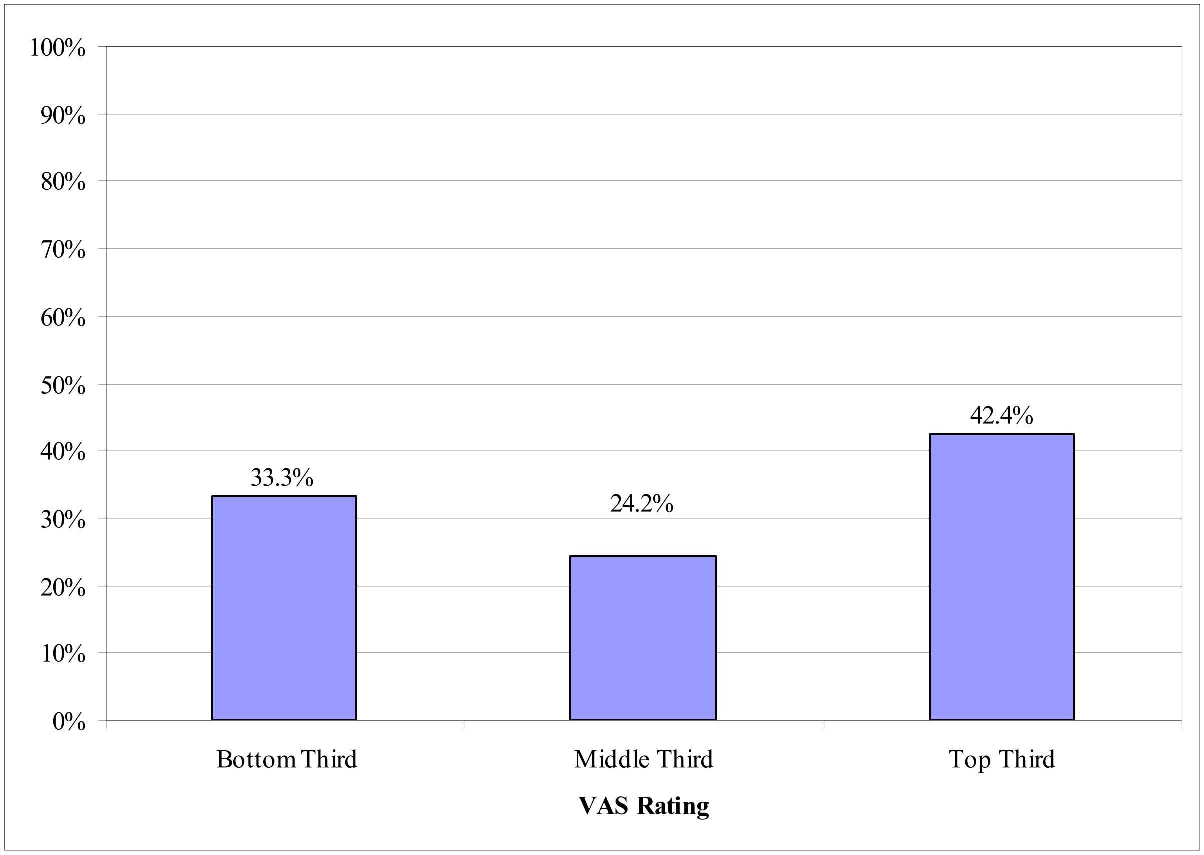


Figure 168. Distribution of VAS Ratings for Usefulness of DDWS in Addition to Other Fatigue Management Steps (0 = Not at All; 100 = Extremely)

A comparison of similarly constructed pre- and post-study survey items indicated that, prior to device exposure, 69.7 percent of the Test Group participant sample believed that the DDWS feedback would "very much" help them to follow their company's fatigue management program, whereas only 18.2 percent of participants responded this way after device exposure. In a related post-study item, 75.8 percent responded "no" when asked whether they used the DDWS to help with fatigue management activities, such as altering sleep times; nor did the majority of participants (81.8%) find other applications for the DDWS. Most of the 18.2 percent who indicated they had found alternate uses for the device did not respond to an open-ended follow-up item requesting examples. Those who did reply indicated uses ranging from "Improve my driving posture and attentiveness to traffic situations" to "Nice place to park my sunglasses."

Parallel items in the pre- and post-study surveys assessed perceived benefits regarding driving safety related to DDWS use. Visual comparison of the response distribution for each item in Figure 169 illustrates that the greatest shift in responses from before to after exposure to device feedback occurred for the category corresponding to the belief that the DDWS did "not at all" improve driving safety. Participant responses using this rating increased from 3.0 percent to 27.3 percent after exposure to the device.

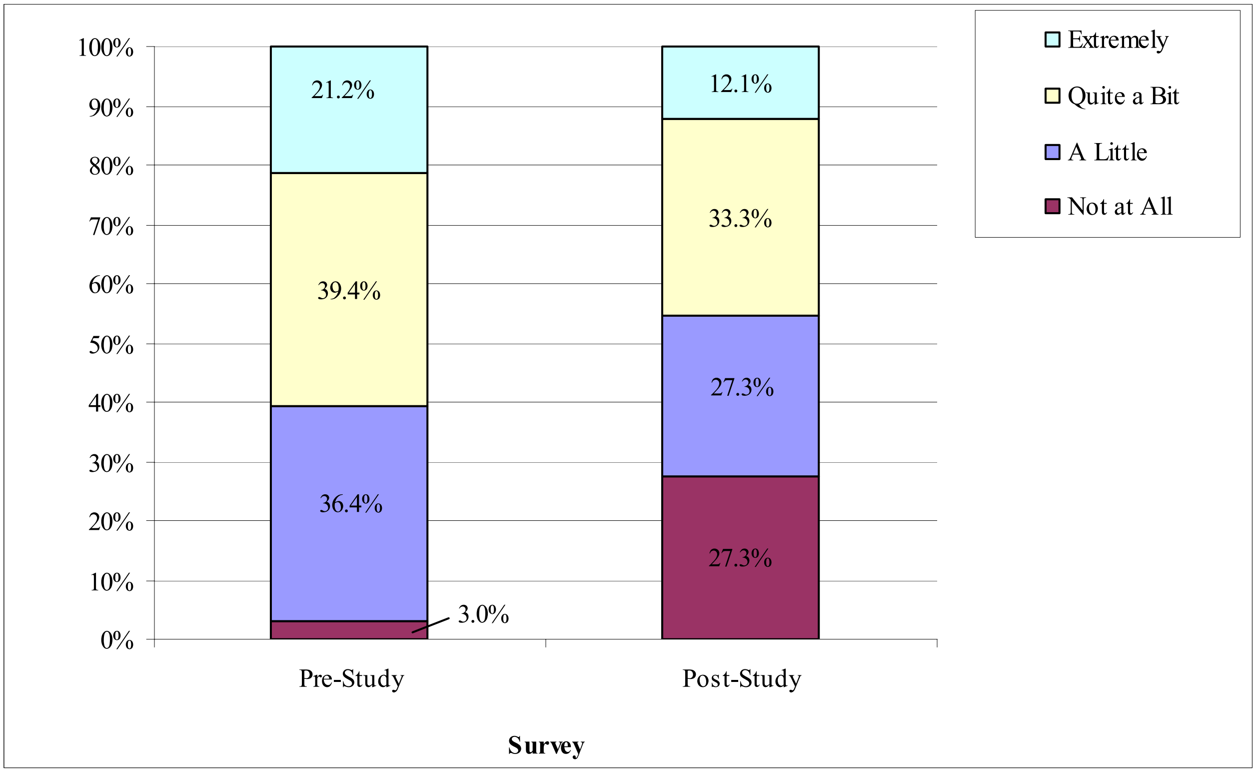


Figure 169. Response Distribution for Ratings of How Much Participants Believed DDWS Would/Did Improve Driving Safety (1 = Not at All; 4 = Extremely)

A similar VAS item regarding the degree to which Test Group participants agreed that the DDWS increased their driving safety resulted in a mean of 52.9 (*SD* = 31.0), just above the scale midpoint (0 = strongly disagree; 100 = strongly agree). The response distribution for this item depicted in Figure 170 shows that participants most frequently provided ratings in the neutral, middle third of the scale, although responses were fairly well-dispersed across the top and bottom thirds, as well.

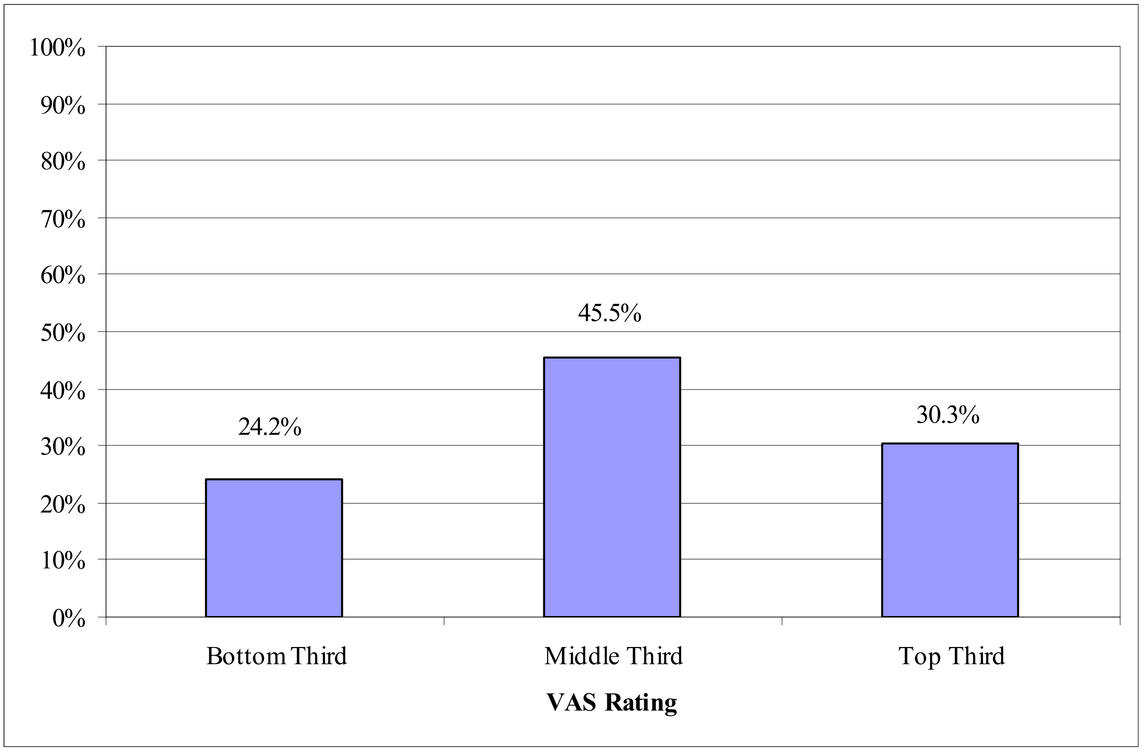


Figure 170. Distribution of VAS Agreement Ratings Regarding Degree to which DDWS Increased Driving Safety (0 = Strongly Disagree; 100 = Strongly Agree)

Additional survey items assessed the perceived value of the device with respect to health and data confidentiality concerns. Prior to the FOT, 100 percent of Test Group participants replied "no" to an item in the pre-study survey asking if they felt that use of the device would harm their health. After exposure to the device, 93.9 percent of participants replied "no" to a post-study survey item asking if they felt that using the DDWS, for any reason, made them feel more fatigued. Additionally, with regard to data collected by the device, 100 percent of post-study survey affirmative responses indicated that participants believed the DDWS information was kept confidential.

Nevertheless, 36.4 percent of the group replied "no" regarding whether they would be "willing to continue using the DDWS device after the study…if the information it gives [them] is not recorded or available to [their] company and others." A follow-on question requested that participants report "Why or why not?" Responses reflected this split opinion and included the following: "It can only improve safety," "It really did change how I drove or how I felt," "If the device is working like it should," and "It is too much of a distraction."

Focus Group Results

Overall, participants in the focus group sessions believed that the DDWS was able to accurately measure a fatigued state during nighttime driving; however, many were not satisfied with its inability to operate during daylight hours, especially after lunch. Additionally, several focus group members viewed the lack of accurate DDWS feedback around dawn and dusk as a major drawback of the device, given biologically established alertness decrements during this time (Wehr, Aeschbach, & Duncan, 2001). As for using DDWS alertness feedback for fatigue management, one participant noted: "I'm watching the bar, and once that bar is all the way over, I'm taking a nap."

Although focus group participants did not register concern about DDWS data confidentiality during the FOT itself, some worried about future device use and the potential for law enforcement and/or management to obtain access to their information. In line with this possibility, participants agreed that they would consequently like the DDWS a "tremendous amount less," as many choose to operate CMVs for the "freedom" and lack of constant managerial supervision that trucking provides. As one participant noted, because "sometimes you have to break the rules a bit," there is a resulting fear of the DDWS becoming a "black box" for the trucking industry. Participants did acknowledge a potential exception to external data access for health professionals. Several focus group members expressed that it may be useful to allow those working in this field access to their DDWS data for assistance with fatigue management, including sleep-wake scheduling. However, participants stressed that this would only be acceptable if the information provided by the device did not yield any negative repercussions. As a final note, focus group members did not report feelings of adverse health or safety effects resulting from the infrared light source utilized for operation by the DDWS.

Discussion

Approximately half each of the Control and Test Groups acknowledged fatigue as at least a one-time problem on the job, both prior to their participation in the FOT, and to a similar though slightly lesser extent when reflecting over the course of the study. Although less than 20 percent of both groups reported past accident or incident involvement that they felt was fatigue-related, closer to 30 percent of the Test Group sample indicated that the DDWS helped them avoid an accident or close call during study participation.

Part of the value of the DDWS as a safety-enhancing device rests with how accurately participants felt that it assessed alertness, as warnings that are perceived to be valid are logically more likely to be attended to. Nearly 40 percent of the Test Group indicated that the device provided accurate and appropriate warnings only "occasionally." As elucidated in the focus group sessions, the prevalence of this response choice may be related to concerns regarding the inability of the device to operate accurately during the day, and particularly in the early morning hours, when the risk of fatigue is great. Nevertheless, a smaller percentage of participants felt that they were highly likely to be involved in a fatigue-related collision using DDWS compared to without it; however, this difference was accounted for by a small number of actual cases.

Opinions varied regarding the utility of the DDWS for managing fatigue. Responding to one survey item, over 80 percent of participants agreed that the DDWS was "useful" for this purpose. However, attitudes were much more widely distributed with increased flexibility in response options (using a 0 – 100 scale). Furthermore, approximately three-quarters of the sample reported not employing the device to assist with particular fatigue management activities, such as altering sleep times; however, some did suggest alternate "off label" uses. Responses reflecting that participants "very much" expected device feedback would help them follow their company's fatigue management program dropped by over 50 percent after device exposure, relative to the pre-exposure responses in this category.

The perceived value of the device in terms of a safety benefit also decreased to some degree after participants experienced driving with the DDWS. Responses corresponding to the device improving driving safety “quite a bit” or “extremely” dropped by approximately 15 percent after exposure. Further, participants expressed neutral opinions on average when asked about the degree to which they agreed that the DDWS increased driving safety. Overall, the sample expressed some doubt regarding the utility of the device as a safety technology, which may be related to ease of use findings indicating the most participants were aware of instances where the device did not work properly and that it was “almost never” or “occasionally” correct regarding their alertness level.

Although nearly the entire Test Group sample agreed that the DDWS did not impact their health by making them feel more fatigued, and all responses indicated that participants believed that the data the device collected was kept confidential, some hesitation was expressed regarding future use of the DDWS. Over 35 percent of the participants indicated that they would not be willing to use the device after the study if the data collected were not recorded or available to an outside party. However, for those who expressed that they would continue to use the DDWS, opinions corresponded with the belief that it could only help their driving. At the same time, participants felt a strong need for assurances that their data would not be made available to outside parties – except perhaps to health professionals who could help with fatigue management – and that available information would not be used punitively by employers. Indeed, these concerns appear warranted. All trucking company management representatives who were interviewed as part of a phone survey (see Chapter 4) indicated that they would desire access to device data as a means of identifying fatigued and unsafe drivers, though for educational/training purposes only. For those in the focus groups who expressed misgivings regarding continued use of the DDWS, much of the reluctance stemmed from the failure of the device to operate reliably and accurately over the course of a day.

ADVOCACY

This section assesses the Test Group’s degree of reported satisfaction with the DDWS in the context of its usefulness, participants’ willingness to endorse it, and potential future device use.

Survey Results

Post-study survey responses to a multidimensional summated rating scale eliciting attitudes regarding new transport technologies were evaluated (van der Laan, Heino, & de Waard, 1997). Referred to as the “Driver Acceptance Scale” within the current effort, this scale has evidenced reliability and appropriateness for use in prior field and simulator research and consists of nine, five-point response, item-pairs that generate composite ratings along dimensions of “usefulness” and “satisfaction” (dimensional subscales were initially determined through factor analysis).

The usefulness subscale is comprised of responses to five item-pairs: useful/useless; good/bad; effective/superfluous (ineffective); assisting/worthless; and raising alertness/sleep-inducing. The satisfaction subscale is made up of responses to the remaining four item-pairs:

pleasant/unpleasant; nice/annoying; likeable/irritating; and desirable/undesirable. Response choices consist of five checkboxes. The center checkbox represents neutral attitudes and is coded as “0.” Positive and negative responses are scored from -2 to +2 to correspond with appropriate scale anchors, employing reverse-item coding where appropriate. Calculations of Cronbach’s alpha to assess scale reliability for this sample was sufficiently large (> 0.8) for both subscales.

Mean composite scores were calculated by participant for each of the subscales. The overall mean satisfaction subscale score was 0.15 (*SD* = 0.99), and the overall mean usefulness subscale score was 0.78 (*SD* = 0.95). Figure 171 provides a plot of subscale scores as x-y coordinate pairs (satisfaction, usefulness) as a way to graphically represent user attitudes towards the DDWS, where data points shown in grey signify the overlap of two participants. Positive satisfaction and positive usefulness attitudes are depicted in the upper right quadrant (*n* = 18); negative satisfaction and negative usefulness attitudes are depicted in the lower left quadrant (*n* = 5; note that further investigation of a subset of these two driver groups is discussed in Appendix B). Mixed responses indicating positive usefulness scores, paired with negative satisfaction scores are shown in the upper left quadrant (*n* = 5). The five remaining coordinate pairs fell directly on either or both the “x” and “y” axes, indicating a neutral usefulness and/or satisfaction score as part of the pairing.

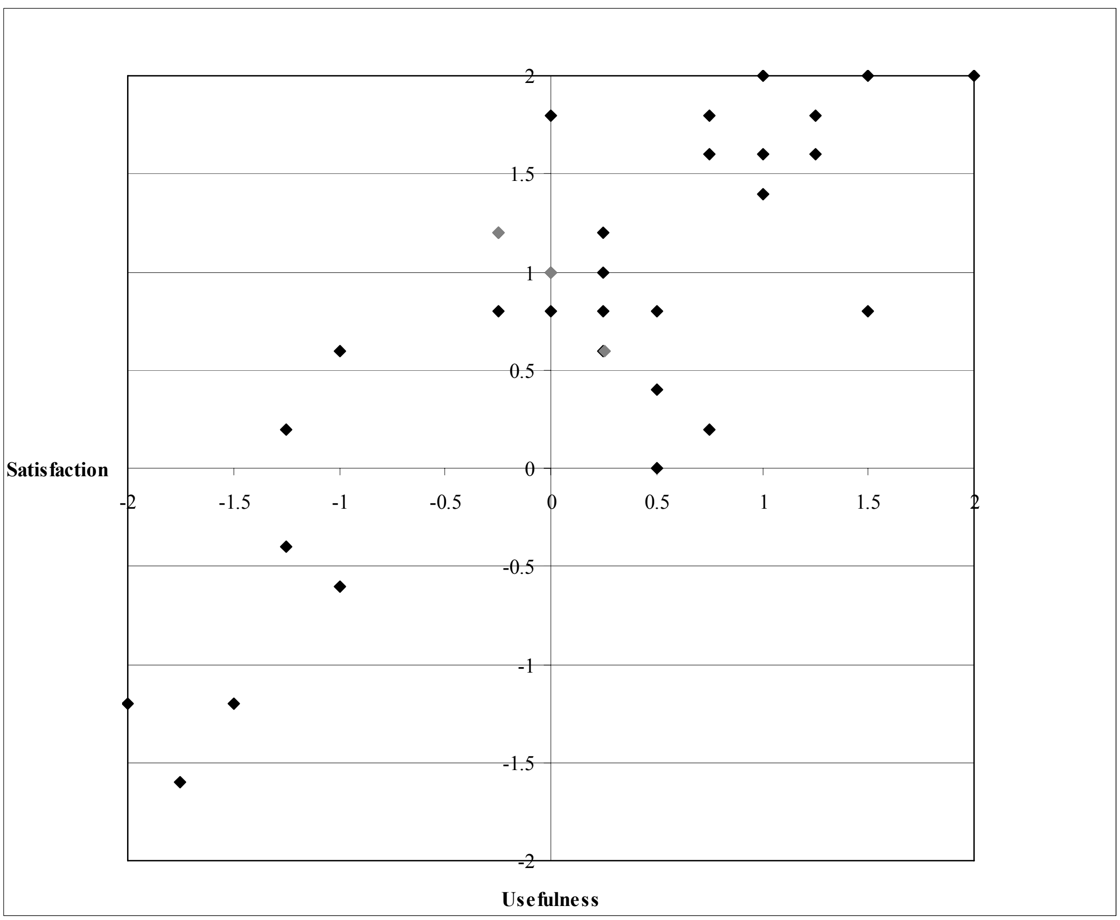


Figure 171. Participant-Paired Ratings of Overall DDWS Acceptance: Satisfaction (x) and Usefulness (y) Subscale Scores

In a related survey item, a VAS asked how satisfied participants “usually were with the DDWS system” (0 = very unsatisfied; 100 = very satisfied). Results evidenced a nearly balanced distribution of ratings above and below the scale midpoint (48.5 percent and 51.5 percent, respectively), as participant satisfaction levels were variable. Visual inspection of the response distribution reflected that the highest number of participant ratings (18.1%) fell in the range of 71-80.

Two survey items addressed advocacy in terms of device endorsement. Test Group participants were asked whether they had recommended use of the DDWS to other drivers in their company and, in a separate item, if they had recommended it to professional drivers outside the company. In both cases, a total of 33.3 percent reported recommending device use to fellow drivers.

A number of survey items assessed general attitudes towards device purchase (note that no specific price range was provided). A total of one-third of the participants indicated that they would be willing to ask their respective employers to purchase the DDWS for the entire fleet, and a lesser proportion (20.8%) were willing to ask their employer to purchase the device for their truck. Results pertaining to individual DDWS purchases were lower; 12.5 percent indicated that they would be willing to buy the device themselves, while only 9.1 percent of the participants indicated that they would be willing to share the cost of the device with their employer.

Focus Group Results

Focus group attitudes regarding advocacy were influenced by perceived false alerts, data-confidentiality issues, and device purchase price. Some participants indicated that false alerts were a “major concern” and that they would not use such a device if the amount of false alerts were not reduced: “The potential is there if it’s working properly to be very effective. As [it] is now, it’s not effective.” Others saw more promise: “Super great idea. I wish more companies had these things. If they could just clear up the small little thing with the false alerts I think they would be great.” When asked directly about device purchase, most focus group members indicated that they would not consider a price that exceeded $400-$500. Additionally, it was clear that participants would not be willing to purchase the device unless guaranteed that the data it collects were not accessible to outside parties. With regard to DDWS usefulness for the purposes of fatigue management, one participant stated: “I would recommend it. I think it’s a good device. It helped me manage my day better. If you work out the little kinks with it, I think it would be a great product.”

Discussion

In particular for those who participated in the focus groups, advocacy of the DDWS appears heavily influenced by the prevalence of false alerts, as the device often provided warnings at times when it was known to function unreliably (e.g., dusk and dawn). In some cases, drivers appeared to be more bothered by this than were others; with the hope for an improved technology, certain focus group participants were in fact very positive. Nevertheless, concern existed that the data collected by the device could be obtained by employers or outside parties and used punitively, although for the FOT, it was stressed that all data were confidential.

Results of the Driver Acceptance Scale analyses indicate that aggregate opinions regarding overall device usefulness were more positive than those regarding device satisfaction, which were closer to neutral. This may be interpreted to suggest that participants acknowledged the utility and effectiveness of the system as a means of notifying them to a compromised alertness state, more so than they were pleased with device feedback outcomes, which, in some instances, included false alarms. Findings from a related survey item regarding "usual" levels of satisfaction illustrated somewhat mixed attitudes towards the device, as ratings were nearly equally distributed between the top and bottom half of the 100-point response scale. An explanation for these results may be found in the focus group discussions, where opinions were less positive regarding actual device performance, and more positive as pertaining to device potential if the "kinks" were worked out.

Attitudes concerning device endorsement and purchase also appear to reflect limited support for the DDWS. The majority of participants had not recommended device use to other drivers and, while one-third of the sample indicated that they would be willing to ask their employer to purchase the device for their entire fleet, just over 10 percent reported being willing to buy the DDWS on their own or share the cost with their employer.

DRIVER CHANGES

This section reports on pre- and post-study levels of perceived fatigue, DDWS usage outcomes and the degree to which Test Group participants used device output to alter their behavior during driving and non-driving periods.

Survey Results

Survey items assessed the extent to which exposure to the DDWS in the Test Group matched with expected outcomes of reduced driving fatigue. Prior to the FOT and again after exposure to the DDWS, participants used a VAS to report their usual degree of fatigue when driving in the early morning hours. Participants provided ratings for periods of driving through rural interstates, and also city/highway areas. Figure 172 and Figure 173 combine responses to these four survey items for the purposes of comparing the distribution of pre- and post-study survey ratings across groups. Generally, the profiles of responses were similar across the Test and Control Group participant samples. For all conditions, both before and after participation in the FOT (and exposure to DDWS feedback in the case of the Test Group), at least half of all responses fell within the bottom third of the scale range. Overall, this indicates that perceived fatigue levels were not particularly high, whether driving with or without the device. Increases in reported moderate-to-severe fatigue, as found in comparisons of the top two-thirds of the pre-post item distributions, suggest greater perceived fatigue levels after exposure to the DDWS. The increase for Test Group participants when reporting on rural driving was 12.1 percent; for city/highway driving, the increase was 6.1 percent. Similarly, reported fatigue increased by 4.8 percent under rural driving conditions for the Control Group drivers, and by 2.5 percent for city/highway driving. The same pattern holds true across groups and driving conditions for reported severe fatigue, using the top third of the scale. However, it should be noted that given a small sample, where one case accounts for a few percentage points (or several percentage points in the case of the Control Group), caution should be exerted when interpreting these results.

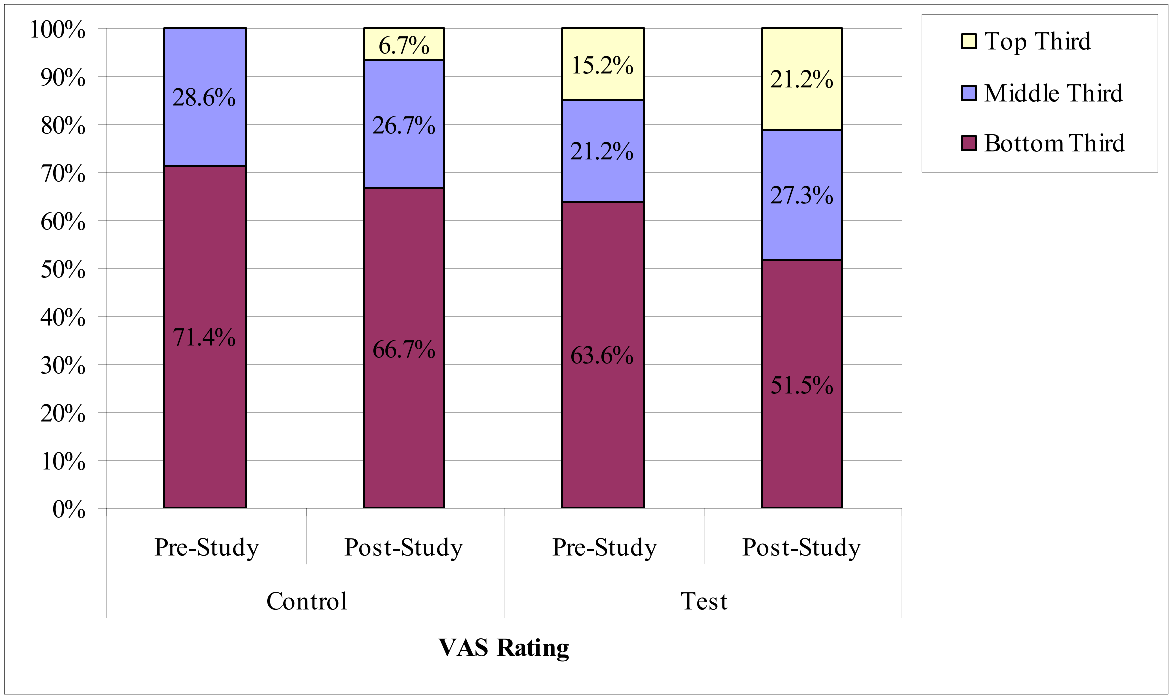


Figure 172. Distribution of VAS Ratings Regarding Usual Degree of Fatigue While Driving During Early Morning Hours on Rural Interstates (0 = Not at All; 100 = Extremely)

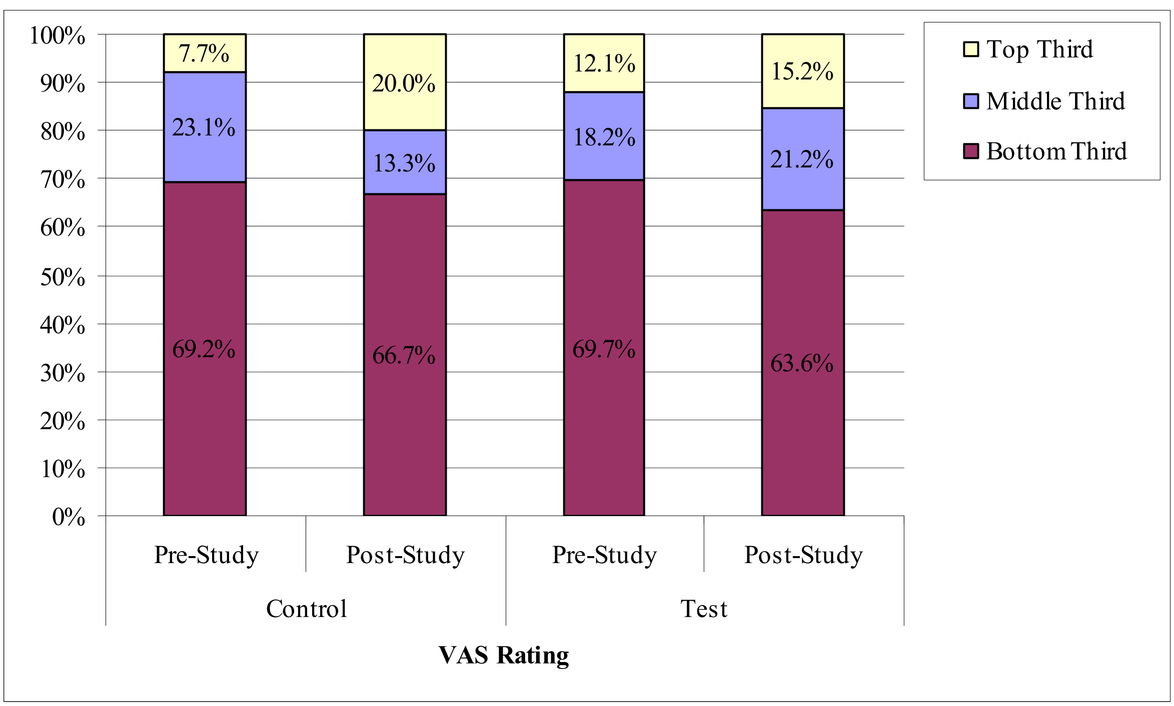


Figure 173. Distribution of VAS Ratings Regarding Usual Degree of Fatigue While Driving During Early Morning Hours Through Cities and Highways (0 = Not at All; 100 = Extremely)

The debriefing survey asked if Test Group participants felt "more alert throughout the night during the study, in general, than [they] would have otherwise," and also provided for an open-ended explanation as to "why?" Thirty-five percent of the sample indicated that they experienced greater alertness during night driving and cited reasons such as the device making them "more aware" of their alertness level and also that the system's flashing lights were a "constant reminder that someone's watching." Fifty-one percent of participants reported that they did not feel more alert, however. Explanations included the sentiment that "if I was getting tired, I would get tired with or without the system in the same way" and that "the study didn't really change my sleeping style." One participant cited job-related pressures: "I drove pretty tired even with the box in there…I mean, [if you've] got to get a load there, [you've] got to get a load there." The remaining 14 percent of the sample replied with ambiguous responses that could not be classified.

The surveys also assessed the ways in which Test Group participants behaved in relation to the DDWS and the feedback it provided. A checklist item in both the pre- and post-study surveys allowed for a comparison of the manner in which participants anticipated interacting with the device, with actual behavior after device exposure. Responses, as shown in Figure 174, indicated that participants' intentions to stop driving and pull over to sleep fell by 30.3 percent after experience with the DDWS. In contrast, planned caffeine consumption increased by 21.2 percent. Finally, a 27.3 percent increase in "other" responses was evidenced. These included actions such as resetting or ignoring the device and continuing to drive, opening the window, listening to the radio, and driver's self-assessing their level of fatigue or distraction.

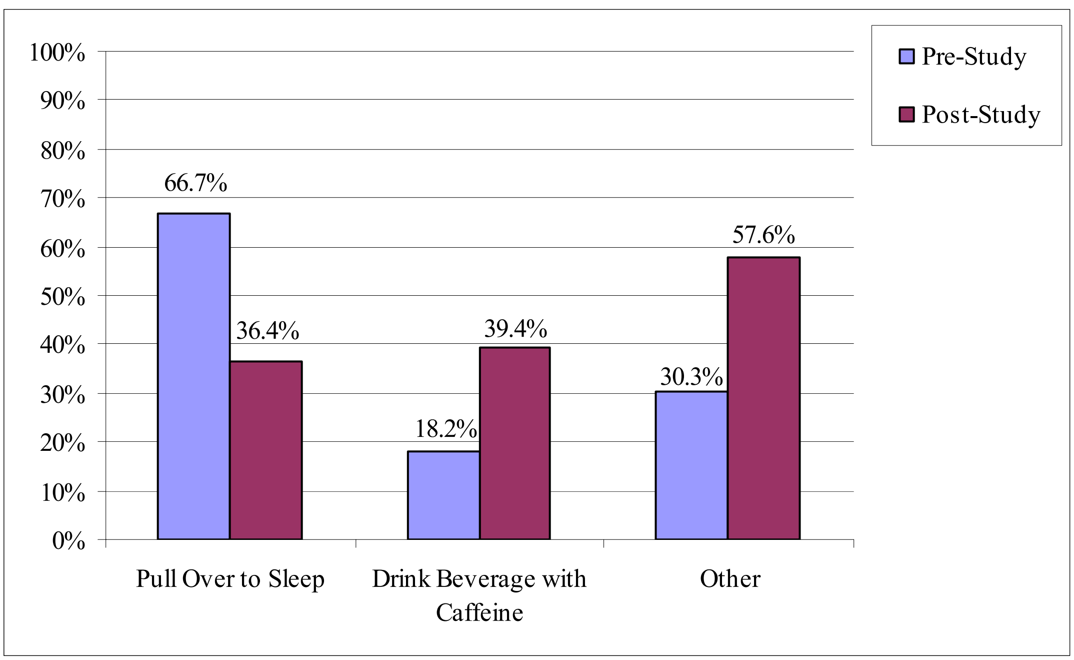


Figure 174. Planned and Reported Immediate Use of DDWS Warnings During Driving (Check All That Apply)

In a related post-study survey item, participants listed "fatigue-fighting actions or activities" that they engaged in throughout the FOT. We grouped and categorized similar responses as presented in Figure 175. In-cab activities encompassed the largest percent response and included actions such as talking on a cell phone or CB, listening to the radio, and opening a window.

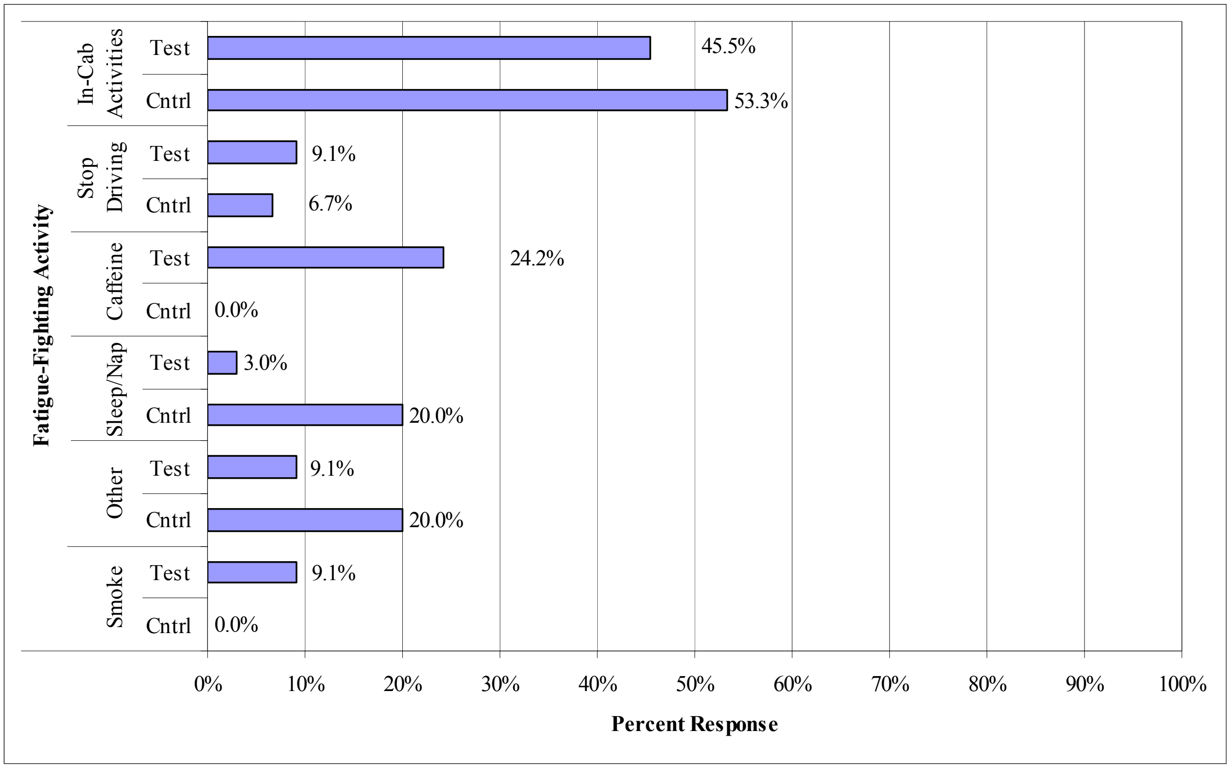


Figure 175. Post-Study Survey Item Responses Listing Fatigue-fighting Actions or Activities Performed When Driving

The additional possibility that Test Group participants interacted with the DDWS in an effort to "trick" it into not providing warnings was addressed in the post-study survey. A vast majority of responses (87.9%) specified that participants "almost never" attempted to foil the operation of the device. "Occasional" attempts at doing so were reported at 12.1 percent.

Ability to integrate the DDWS into personal fatigue management habits, as well as how the device may have impacted non-driving time, was assessed using a number of survey items. The mean response to a post-study survey VAS item asking how well Test Group participants were able to include the DDWS into their fatigue management habits fell just above the midpoint of the scale at 54.6. As shown in Figure 176, the responses to this item were more heavily weighted within the top two-thirds of the scale range, totaling 78.8 percent.

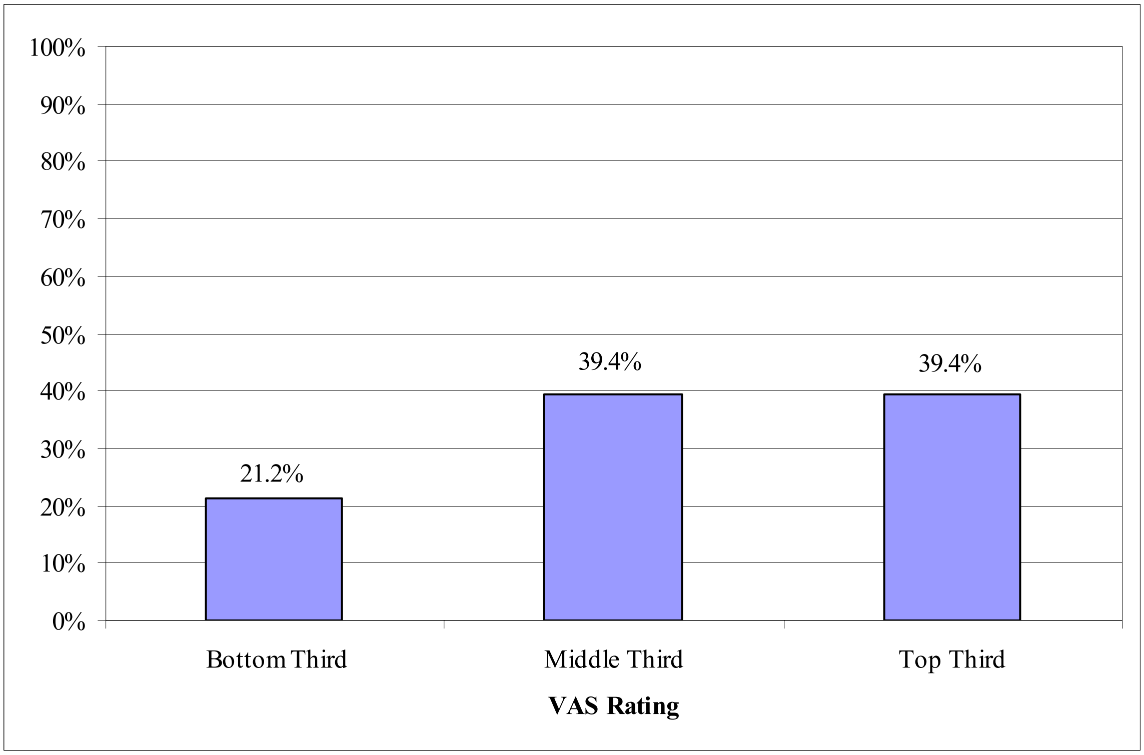


Figure 176. Distribution of VAS Ratings Regarding How Well Participants Felt Able to Include DDWS Into Fatigue Management Habits (0 = Not at All; 100 = Extremely)

Similar items in the pre- and post-study surveys asked Test Group participants to check off planned and actual uses of the DDWS as applied to their non-driving time. As provided in Figure 177, the largest shift in responses was evidenced in a 51.5 percent decrease in utilizing the device outside of driving for the purpose of fatigue management. A decrease of 42.5 percent was found in responses indicating use of the DDWS for the purposes of sleep/wake scheduling.

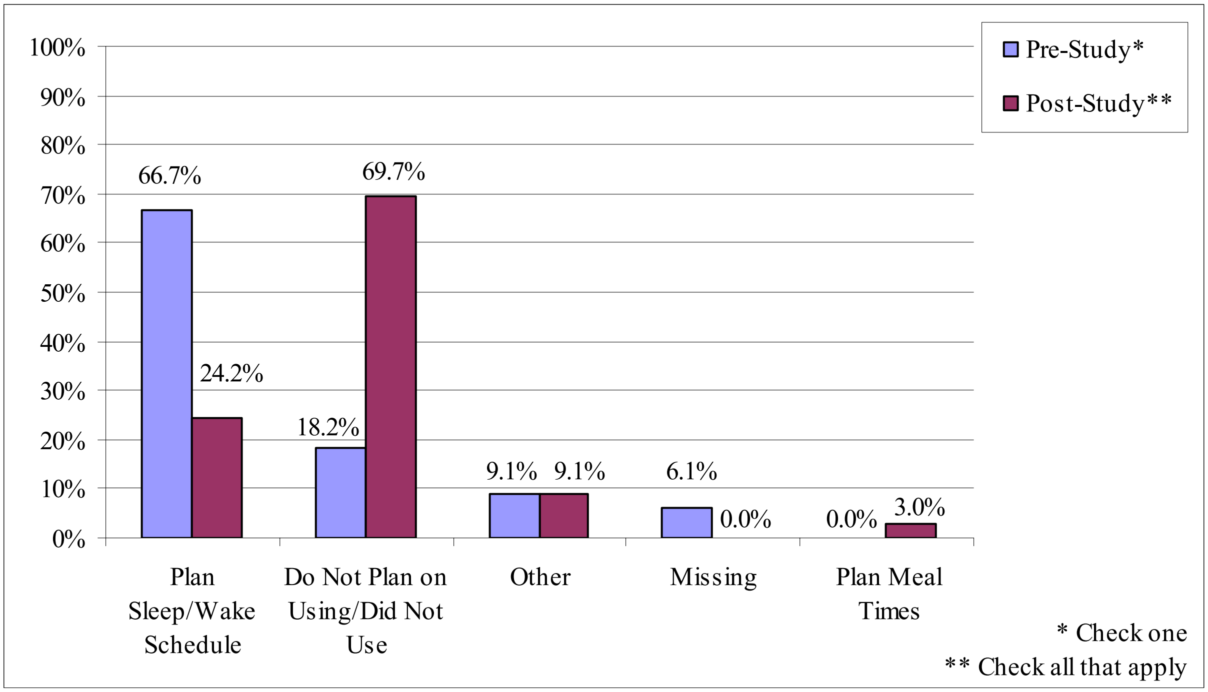


Figure 177. Comparison of Planned and Actual Reported Use of the DDWS During Non-Driving Time (Checklist)

A post-study survey item assessed perceived changes in driver health as a function of the DDWS. Responses specified that 75.8 percent of participants felt that the device had "no effect" on their health, while a smaller percentage (21.2%) felt that the device "improved" their health in some way, while one respondent indicated that the device "impaired" his health.

Focus Group Results

Focus group discussion provided insight into the degree to which participants changed their behavior to incorporate the DDWS into work-related fatigue management habits. One focus group member stated that he had not found any uses for DDWS output other than as a "conversation [piece] at truck stops." Similarly, another explained that, "it's so annoying and so inaccurate you get completely complacent with it and you don't even care if it's there [because] you just get sick of hearing it." Although these views appeared to some degree to be shared among focus group participants, a utility to the device output when the DDWS operated as intended and expected was nevertheless acknowledged: "If I knew I had to drive late, I know this thing [is going to] work. I can push it a little further and it'll catch me, and it always did," but he added, "daytime hours [were] another problem."

Focus group participants also discussed using device feedback during non-driving time. Nearly all focus group members confirmed that they had not used the DDWS to plan off-duty activities. As one person offered, "as soon as you got out of the truck, you stopped thinking about it." Nevertheless, one participant claimed to have increased the amount of time he spent sleeping while off-duty as a result of DDWS output.

Discussion

Survey results from both the Test and Control Groups indicated that, overall, driving fatigue was not considered an extreme problem, either prior to or after participation in the FOT, or across various conditions of rural and city/highway driving. Nevertheless, consideration of post-FOT ratings made at a level suggesting moderate-to-severe fatigue during early morning hours evidenced a comparative increase, raising the possibility that participants actually felt more fatigued driving with the device, and with the addition of device feedback in the case of the Test Group. Although the calculated percent shift was actually not very remarkable, especially in light of the small sample size, this finding is nevertheless further supported with the outcome that the majority of participants did not report feeling more alert during night driving while participating in the FOT. An alternate explanation for these results is that the fatigue management training that was provided, coupled with exposure to device warnings, enabled participants to better recognize (and consequently mitigate) their fatigue. Nevertheless, over 75 percent of the sample reported that use of the DDWS neither improved nor impaired their health (i.e., "did not affect"), making it difficult to pinpoint the nature of the perceived impact, if any, the device had on driver state and overall well being.

Outcomes of survey data analyses also explored the degree to which Test Group participants changed their driving habits as a result of the device. Prior to exposure to the DDWS, participants most often expected to stop driving and pull over to sleep upon receiving warnings. However, after engaging in fatigue management training and experiencing driving with the DDWS over the course of the FOT, most participants altered their response to report "other" means of interacting with the device, including resetting the DDWS and/or ignoring it, in addition to operational strategies such as opening a window, as was suggested as part of the FOT fatigue management training participants received. Reports of caffeine intake also increased over planned use at the onset of the FOT. Despite noted instances of some unintended uses for the device (as provided in Perceived Value), nearly 90 percent of the survey sample indicated that they almost never attempted to foil the DDWS in order to avoid receiving warnings while driving. Nevertheless, survey responses suggested neutral attitudes, on average, regarding how well participants felt they were able to include the DDWS into their fatigue management routine. Although one focus group member was more positive in this regard, his interaction with the device could be viewed as an undesirable outcome, in that he used it to push himself further when he was fatigued, trusting that the warnings would "catch" him. Finally, participants did not appear to incorporate DDWS feedback into their non-driving lives. Most responded to a survey item asking about use of the DDWS by indicating that they did not utilize device feedback in this manner. In comparison, prior to DDWS exposure, over half the participants intended to use the device to plan sleep/wake schedules.

CHAPTER 4. FLEET MANAGEMENT ACCEPTANCE

This chapter provides a management-level perspective with regard to perceptions of a DDWS. It addresses anticipated driver acceptance of the device, its potential safety and economic benefits, as well as the perceived impact of such a technology on operations. Additionally, feedback is offered regarding potential device improvements.

RESULTS

Phone interview results are summarized by theme as related to the fleet management acceptance objective. We incorporated various anecdotes and anonymous quotations offered by nine trucking company management-level interviewees to the extent that they were representative of the overall sample and helped to elucidate a point. As further detailed in Chapter 2, interviewees represented five long-haul and three short-haul operations, as well as a leasing supplier of trucks for both long- and short-haul operations. Interviewee positions within these organizations included four safety managers/directors, two fleet managers, one company president, one co-owner, and a maintenance director.

Perceived Driver Acceptance

The concept of "perceived driver acceptance" explored the degree to which trucking company interviewees expected their drivers to support the use of such an alertness-monitoring device, in addition to their views regarding its various potential advantages and disadvantages.

Six of the nine trucking company interviewees were enthusiastic supporters of alertness-monitoring technologies for use in commercial vehicles as a result of the potential for these devices to reduce accident rates. Interviewees from both long-haul and short-haul operations reported that the technology would be a welcome, added tool to help drivers stay alert and reduce the likelihood of fatigue-related accidents. The alerting feature of the system was cited by several participants as having the potential to be of particular use for the drivers. As one interviewee stated, "Based on my [past] personal experience as a driver who has been drowsy, this technology could save lives. You don't have to be sleep deprived to fall asleep. It can happen quickly."

Despite general management support for the concept device, attitudes regarding its utility varied across situations. For example, one interviewee who herself was supportive of the technology pointed out that where she worked, the decision to purchase such a device would be entirely up to the driver, as they are an owner-operator company. Her opinion was that employees may not perceive a need for the device and would be therefore less likely to spend the money to purchase it stating, "Our drivers have excellent driving records and a lot of experience, so they know when they need to rest and stop driving." A similar view was expressed by an owner (and driver) of a long-haul operation who, along with the company's co-owner, each possess over 30 years of job experience. He commented that, "We don't push ourselves too hard, and we rest when we need to." Employing an operations-based rationale, an interviewee from a short-haul company also did not deem the technology necessary, suggesting that his drivers return home every night, their hours are closely monitored, and they call the office every couple of hours while on duty. However, he did acknowledge a potential benefit for long-haul drivers, who typically drive for long periods with no interruptions and may routinely sleep away from home. Conditional support for the technology was also provided by one participant who would advocate its use on a limited, test-basis only, reporting that his company does not have a problem with fatigue-related or run-off-road incidents.

Survey responses regarding anticipated driver attitudes toward the concept device were split. A single interviewee expressed unqualified support, stating that drivers at his long-haul company would welcome the technology, since it does not impede driving and would provide a warning to help bolster alertness and prevent drivers from operating in the “tunnel zone,” a state in which the driver has lost situational awareness. Four company representatives offered conditional support, suggesting that drivers would likely eventually accept the technology, even if its use was mandatory, if the safety benefits were clearly explained to them. As one interviewee clarified, “if you appeal to their egos and make them part of the safety process, they [will] be more willing to accept the device.” Additionally, two interviewees offered the opinion that company drivers would be more amenable to accepting the technology than owner-operators because (being their own boss) owner-operators have a higher level of autonomy and independence than do company drivers. As such, they tend to be confident in their ability to self-monitor for fatigue. Further, it was pointed out that owner-operators, unlike company drivers, would have to purchase the device themselves; yet, if they did not perceive a need for the technology, they would be unlikely to spend the money for it. The remaining four company representatives expected drivers to oppose the technology, stating that they were not aware of any fatigue-related incidents at their organization, and that employees are sensitive to the potential for company-monitoring of driving activity through equipment in the cab that may be viewed as “intrusive.” It was added that drivers may fear job loss or other punitive action as a result of the information collected from this device.

Respondents cited the safety benefit associated with potential reductions in accident rates as the main advantage of a technology such as the concept device. One interviewee stated that companies can ill-afford the downtime resulting from an accident that takes a truck and/or driver out of service. Another asserted that the most important thing is to “stay out of the ditch” (i.e., to avoid run-off-road crashes caused by falling asleep at the wheel). In addition, interviewees suggested that the device could help determine whether drivers are getting the proper amount of rest, if they suffer from sleep disorders such as sleep apnea, or if they are falsifying their log books. Potential disadvantages associated with the technology included concerns regarding driver acceptance of a device that monitors activity, implementation and start-up issues (“getting the bugs out”), system cost and maintenance, and the possibility of legal action against the company if device data were available to attorneys.

When asked their opinions regarding training, interviewees were unanimous in their agreement that supplying education and training to drivers about the capabilities, advantages, and disadvantages of this type of technology would make employees more likely to support and approve of such a device. Respondents offered that training and education would demonstrate a company commitment to driver safety, concern for driver well-being, and would provide the platform to illustrate that such technology is more than a potential means to monitor work hours and driving behaviors.

Safety and Economic Benefit

Interviewees were surveyed regarding their perceptions of the potential economic benefits of alertness-monitoring technologies, as attributed to increased safety, and how much they would be willing to spend for such a device. Respondents also offered insight pertaining to what types of insurance and federal incentives would make them more likely to recommend the purchase of this technology for their fleets.

In all cases, respondents suggested that the cost savings associated with potential accident rate reductions would be the primary economic benefit of this technology. Company costs associated with accidents include loss of life, loss of equipment, driver injuries and associated medical costs, vehicle property damage, and workman's compensation claims. Three interviewees also noted the economic benefit of potentially lower insurance rates as related to improved safety records, and fewer lane-change and run-off-road accidents.

Interviewees suggested two federal incentives in particular that would make them more likely to recommend the purchase of an alertness-monitoring device to their companies: tax credits and federal grants. In citing government grants as an incentive, one respondent stated that his home state of New Jersey provides money to trucking companies for the installation of heaters and air conditioners in their vehicles; a similar federal program could provide an incentive for the use of alertness-monitoring devices. Regarding insurance incentives, all participants expressed that the most influential would be offering rate reductions in conjunction with the purchase of alertness-monitoring technologies. Additionally, one interviewee noted, "If our insurance carrier told us they would not cancel our policy if we had this device on board our trucks, it would make us more likely to consider purchasing it." However, it was also stressed that drivers would have to be protected from insurance companies obtaining and using device information in the event of a lawsuit.

Four interviewees did not offer a specific price point at which they would be willing to purchase the device as described to them, but expressed that keeping the price "low" would be critical, especially in light of the recent high fuel costs. The remaining respondents suggested a range of $200 to $1,500 per unit; one participant quoted a unit price of $500 for small fleets (4 or 5 trucks), but indicated that there should be a discounted unit price for large fleets of 100 or more.

Impact on Operation

In order to assess the expected operational impact of this technology, participants were questioned regarding the potential for trucking companies to monitor and use the data collected by such a device, the degree to which company access to device data might influence driver utilization and acceptance, and potential policies concerning required driver behavior following device warnings.

All interviewees expressed an interest in collecting and monitoring device data for the immediate purpose of identifying instances of fatigued driving, and with the ultimate goal of mitigating such situations and avoiding associated accidents. Respondents further indicated that the eye closure data would be a valuable tool when educating drivers about getting proper rest. In addition, it was acknowledged that, although the availability of this data would permit the tracking of employee work hours, it would not be used punitively or as a basis for dismissal. One participant expressed a desire to monitor device data only if it were transmitted in real time to the company, with the rationale that the value of such data diminishes over time. Another interviewee who had reported that he would want to collect and monitor device data offered the caveat that the company would be more reluctant to do so if it were admissible in court.

In most cases, participants shared the opinion that, assuming drivers are guaranteed that device data is employed strictly for safety managing purposes, the potential for management to collect and use data from an alertness-monitoring system would positively influence driver acceptance. As one respondent stated, "The warnings are a message to the driver that if he doesn't get off the road, he can kill himself or someone else. It is about safety, not big brother." Two interviewees provided the neutral assessment that collecting and using data from the device would not affect driver acceptance and utilization. For example, one participant reasoned that if drivers do not exceed hours of service regulations and do not drive at times when they should not (i.e., while sleepy), they have nothing to fear from being monitored. However, drivers who habitually drive beyond legal limits and do not get enough rest would be much less likely to endorse the use of this device.

With regard to how likely their company would be to implement policies prescribing mandatory actions in response to device warnings, a common sentiment was that management would need to solicit driver opinions and work collaboratively with employees in order for such an undertaking to succeed. Several participants strongly cautioned against using any such policy as a basis to penalize or dismiss drivers. Of the suggested driver actions in response to device warnings that were mentioned, pulling off the road, taking a break, and/or obtaining rest were cited most frequently. One respondent also proposed that drivers who received repeated warnings could have their route altered or perhaps be switched from long- to short-haul operations, if possible, within their company.

Improvements

Interviewees were questioned regarding additional features or other improvements they would recommend for the concept device as it was described to them. Additionally, they were asked to comment on any concerns they could anticipate regarding device performance.

Reliability, accuracy, and maintainability of the alertness-monitoring device were the most critical performance concerns reported by the interview participants. Several respondents stressed that getting the "bugs worked out" prior to operational use would be extremely important. For instance, it was suggested that too many false alarms would likely antagonize drivers, perhaps forcing them to disengage the device. One participant also commented that truck cabs are often already outfitted with other accessories that could interfere with the effectiveness of this type of alertness-monitoring device. For example, a CB radio dangling in front of the driver might interfere with the infrared light source.

When questioned, four of the interviewees provided suggestions for potential additional device features including monitoring driver physical status, detecting alcohol on the driver's breath, monitoring driver distraction as well as alertness, and combining measurements of eye closure with measurements of vehicle parameters, such as lane movement/lane departure and following distance. One participant commented that, in his view, the additional vehicle parameter measurements could increase the device's worth from between $200 to $400 to $800 to $1,000. The remaining five respondents had no suggestions for potential additional features, but two remarked that the audible warning alert feature would be the most meaningful element of such a device.

Discussion

Overall, interviewees expressed support for the potential of an alertness-monitoring device as a tool to improve safety, and consequently save money by reducing accidents and injuries. However, some respondents indicated that they would not recommend the use of this technology at their company in particular because they do not believe the need exists. In one case, the company was a short-haul operation, and it was explained that drivers took frequent breaks during the day to load and unload trucks and then returned home to sleep every night. In the second instance, it was offered that the company has an excellent safety record and no perceived problem with driver fatigue. Although the potential benefits of such a device were apparent to most of the interviewees given situations of drowsy driving, they often expressed that they did not find it necessary for their own operations. To the extent that the results of this interview are representative of the industry as a whole, it may then be the case that operator fatigue is not openly considered by trucking company management to be a universal problem.

Responses pertaining to driver acceptance of the proposed technology suggested overall concern that employees would consider it a "big brother" tool for management to monitor their actions and use it as a basis for punishment or dismissal. However, all respondents reported that providing training and education to drivers about the various capabilities, advantages, and disadvantages of this type of technology would likely garner more approval and support for the device. Furthermore, it was anticipated that driver acceptance could be bolstered through clear discussion of management's intended use for the data, including informing drivers of potential safety hazards, as long as there were not punitive outcomes. Although such an arrangement may, in theory, be possible, it is nevertheless unlikely that labor unions would easily approve, and there may be other privacy issues pertaining to health information access to consider. Management interviewees deemed the potential for data admissibility in court as problematic, and anticipated the same would be true for their employees (though quite possibly for different reasons).

Despite several respondents indicating that they did not believe that use of such a device was necessary at their company, reduced insurance rates and government tax credits were cited as incentives that might make them more likely to purchase the technology, especially in light of rapidly-rising fuel costs. The most frequently cited economic benefit of alertness-monitoring technologies was also cost-related, as interviewees anticipated that a safety benefit for the device would translate into reduced accident, injury, legal, and property damage expenditures. Although nearly half of the interviewees did not offer a price at which they would be willing to purchase the device, those who did provided a wide range of $200 to $1,500 per unit. Concerns with regard to reliability, accuracy (in particular, false alerts), and maintainability of the device were noted as potentially critical system performance issues.

CHAPTER 5. DEPLOYMENT

This chapter addresses short- and long-term deployment issues for alertness-monitoring technologies from the perspective of companies currently involved with the development and deployment of such devices. Additionally, it provides the views of a trucking-related organization representative regarding potential advantages and disadvantages of such technologies, driver acceptance and device deployment, as well as anticipated concerns of management and organized labor.

RESULTS

Interview results are summarized by theme as related to the deployment objective. We incorporated various anecdotes and anonymous quotations offered by two respondents from alertness-monitoring technology development companies and one representative of the ATA to the extent that they were representative of the interviewee sample and helped to elucidate a point.

Introduction and Price Range

In order to better understand the market for alertness-monitoring devices, interviewees were asked about the general availability of their product. One indicated that his company's system has been available for use by the general public for approximately six months and is currently being used for road transport applications in light commercial vehicles, heavy trucks, and off-road equipment. He also pointed out that the technology would be equally suited to rail, maritime, and aviation environments. Prior to deployment, his company tested the device during an introductory period of "over one year," using approximately a dozen vehicles operating in several different environments and with a number of partners. Additionally, the system was used during this period by several customers for driver behavior research.

The respondent from the second company indicated that, at the time of the interview, their device was not under high-volume production. The technology had been developed to a pre-production level, where the company focus was on providing robust eye and face-forward detection performance, and its functionality was demonstrated to both original equipment manufacturers and public audiences. He further explained, "Our company believes the sensing components of such a system are ready, but the intended countermeasure for warnings differ[s] among OEMs." Continuing efforts lie with the development of production-intent hardware in conjunction with a committed OEM production program. He added that OEMs are currently considering deployment of this technology as early as 2010 and that interest has been primarily expressed by the commercial vehicle segment. A deployment introductory period is not planned, rather the company intends to initially roll out the product across select platforms and then expand to higher volume platforms.

With regard to pricing strategies, one interviewee offered that, "Pricing information is generally treated confidentially and depends on volume. Pricing for mass-market deployment of the technology scales to levels that are typical of the automotive industry. We do not see cost as a barrier to mass-market uptake of the technology." The other company representative commented that there is a direct, inverse relationship between price and volume, and that his firm believes the market will expect gradual price reductions, or performance enhancements (in lieu of price reductions), over time. When asked about an introductory price, he replied that this would depend primarily on OEM requirements/functions, volumes, and roll-out plans.

A series of questions were used to determine the developers' views on the long-term deployment potential of alertness-monitoring technologies. Regarding an estimate of the technology deployment rate over the next 5 to 10 and 10 to 15 years, one respondent indicated that his company expects alertness-monitoring systems to become available in all upper-market passenger vehicles and in commercial heavy vehicles within 5 to 10 years. In 10 to 15 years, his company expects that these systems will be available in most vehicles on the market, and additionally offered that "as more vehicle systems become automated, the need for driver alertness warning systems will only become more critical. With increased automation, the driving task becomes increasingly boring, increasing the dangers of fatigue and distraction." The other interviewee suggested that projected deployment depends on the initial adoption dates, volumes, and prices, but the expectation is that the rate will mirror deployment curves for similar advanced safety technologies. He said that one difference which will impact penetration rates is the integration of other functions (e.g., driver recognition, security,) and technology (e.g., vehicle-infrastructure integration, wireless tracking) into transport vehicles. Respondents were not able to provide pricing estimates over the requested time periods. One interviewee commented that the timeframes were too far in the future to make a reasonable estimate of the pricing, but reiterated that his company does not view cost as being a barrier to broad OEM adoption of their technology. The other respondent offered that price depends on initial adoption dates and volumes, and that the price range for their device is expected to follow price curves for similar technologies.

Pertaining to long-term deployment projections for various motor vehicle classes, one interviewee stated that his company expects that the deployment of their technology by OEMs will be similar to other driver assistance systems that have come about over the past several years, such as anti-lock braking systems and air bags. He indicated that these systems "have always been introduced first in expensive vehicles and have worked their way into more affordable vehicles," typically over a 10- to 15-year period. The second participant commented that if the advantages of an alertness-monitoring system can be measured and quantified, monetary benefits will encourage OEMs to increase penetration rates in order to generate cost savings and help justify the additional expense for such equipment.

Perspectives of Trucking-related Organizations

In terms of the advantages and disadvantages of this type of technology, the ATA representative indicated that he did not have detailed knowledge of any alertness-monitoring systems, as they had not observed any being utilized in actual field operations. Nevertheless, he expressed general support for any technology that would assist a driver in monitoring his or her alertness while driving. The respondent did add, however, that according to statistics from the Federal Motor Carrier Safety Administration, truck driver fatigue contributes very little to actual crash occurrences. Consequently, his belief was that motor carriers will focus their resources more so on engineered systems that reduce high crash-likelihood scenarios such as rollovers, collisions, and lane departures.

When questioned regarding organized labor's expected views regarding alertness-monitoring technologies, the ATA representative indicated that, "This will depend on how drivers perceive the device will actually function as designed and marketed. For example, does it produce an alert when a driver is actually having unexpected drowsiness? If so, the better it will be received." In terms of the types of incentives that would be desired by organized labor in return for its members utilizing an alertness-monitoring device, ATA declined to offer a position, suggesting instead that it would be best to direct this question to representatives of organized labor and their labor-relations counterparts in management. He did provide, however, that any discussions between management and labor would likely take place in keeping with safety policy requirements.

In response to a question about anticipated liability concerns as related to use of this technology, the interviewee suggested that liability issues are always a possibility, especially when there is awareness of a problem (e.g., the driver is drowsy or nodding off) and no corrective action is taken. With regard to trucking company liability, the respondent stated that there are potentially significant data discovery issues if the data are stored, since liability laws in most states "allow trucking to be more financially liable than the requisite degree of negligence." He also conveyed a liability concern for drivers that false or misleading device data could be used against them unless legislative/regulatory safeguards are provided.

In questioning whether the ATA representative expected that the utilization of alertness-monitoring devices would raise issues pertaining to potential management data collection and data usage policies, he offered the opinion that motor carriers will likely want to have real-time, over-the-air data on drivers' alertness to assure that at-risk conditions can be immediately corrected. Further, he suggested that management will presumably want a degree of protection from the unauthorized release of data and third-party inquiries during discovery proceedings. In terms of policy issues that would prescribe a driver's response to warnings, the respondent suggested that "management will likely want to implement internal controls to assure the device is used properly."

The only potential issue identified by the representative from ATA with regard to acceptance of this technology was that false readings and frequent alerts that require an operator response will likely lead to negative reactions by drivers and an overall lack of driver acceptance. When asked to comment on issues related to technology developments and/or device capabilities, the ATA representative expressed that any audible or visual signals from the device should not produce distractions that could conflict with safe operation of the truck. The signals should be coordinated with other on-board safety technologies (e.g., lane departure systems, collision avoidance systems, and seat belt usage alarms) or other truck operational-alerts. He added that "only the highest of all priority alerts should be given to the driver, whether [it be] collision potential, tire pressure loss, electrical fire, head-nodding, brake loss, etc."

Finally, the representative from ATA offered a general comment about the prospective deployment and marketability of alertness-monitoring technologies by emphasizing the need for real-time output and information from the device, while also acknowledging that this capability would likely increase the cost of the system. The respondent also indicated that a device with the sole functionality of detecting driver fatigue may not be enough enticement to trigger motor carrier investment. The ability to detect driver vigilance, such as scanning to detect conditions in mirrors and other general "road conditions," and also to monitor the level of driver distraction, may have added appeal for carriers.

Additional Activities

As part of the phone survey of trucking company management staff, responses were sought regarding activities that interviewees believed would help promote the deployment of an alertness-monitoring technology, as well as research undertakings they would recommend in the area of drowsy-driving as related to crash avoidance. Additional survey questions explored the extent to which trucking fleets have already purchased or considered the use of alertness-monitoring devices.

Participant opinions varied regarding how to effectively promote and successfully deploy this type of device for use by trucking companies. Four interviewees suggested that carriers use the technology on a trial basis and publicly report the results: a documented safety benefit, positive driver testimonials, and backing from insurance companies would all lend support for further deployment and utilization. One respondent specified that the government should sponsor and fund such a test program. In addition, educational seminars were suggested as a way to "sell" the technology to trucking company safety directors and promote management buy-in. Two participants also recommended economic incentives, such as reduced insurance costs, as an effective promotional strategy. Alternately, it was noted that for some owner-operators and trucking companies with excellent safety records and no perceived issues with driver fatigue, device purchase and use would be unlikely, regardless of attempts at promotion.

Pertaining to scientific research in the area of drowsy-driving as related to crash avoidance, one participant commented that drivers in particular and companies in general are different from one another, and as such, broad research that could be universally applied to all segments of the industry would be difficult to design. The only other response to this query provided that "enough research pertaining to hours of service has already been done, and the industry should review and learn from it." The interviewee continued with the suggestion that perhaps more research into the effects of sleep apnea on driver behavior would be helpful, since it is an issue of particular concern in the industry.

At the time of interview, no companies had purchased or considered the use of an alertness monitoring system. Related to this, it should be noted that three of the companies initially contacted declined to participate in the phone survey, citing that they were "small, low-tech" operators with no interest in advanced technology safety systems. However, two respondents working for long-haul companies offered that their companies purchased and are currently utilizing other advanced safety technologies in their fleets. One company uses a forward collision warning system; the other carrier installed a lane departure warning system, as well as a front/side collision warning system on several of its units. Demonstrating a proactive approach and commitment to safety, the representatives of both of these companies expressed enthusiastic support for alertness-monitoring technologies during their interviews.

Discussion

Both technology company interviewees reported that, as part of development, their respective systems had undergone extensive research, design, laboratory, and on-road testing. At interview time, one resulting technology had already been deployed and the other was engaged in further development in conjunction with OEMs. Although neither respondent provided specific price points for their company's technology or these technologies in general, one suggested that cost would not be expected to impede mass-market use, and similarly, the other indicated that as usage volume increases, prices decrease. To that end, it was anticipated that within 10 to 15 years, such systems will have been introduced even in some lower-end vehicle models, similar to how other driver assistance systems, such as ABS, have become widely available.

Although not having been directly exposed to alertness-monitoring technologies in the field, the ATA representative who was interviewed expressed support for the premise of such devices, despite believing that motor carriers would be more likely to focus available resources at engineered systems geared toward reducing high crash-likelihood scenarios. His views regarding liability included both driver and company concerns, having suggested that available data could serve to implicate each for different reasons. At the same time, the interviewee offered that trucking company management would nevertheless likely be interested in data that would allow for the opportunity to mitigate at-risk (i.e., fatigued) driving conditions. Regarding his views on driver acceptance, the ATA representative suggested that false alarms and frequent alerts requiring a response would likely negatively impact acceptance, and that distractions resulting from information coming from too many technology sources could be problematic. He offered that, ultimately, organized labor's expected stance would be driven by driver attitudes to the device. In addition to real-time alertness-monitoring, he believed that technologies with additional features, such as the ability to detect driver distraction, would increase a device's likelihood for successful deployment in the CMV industry.

Trucking company management staff provided additional ideas for promotional and deployment strategies, including documenting driver acceptance and device safety benefits through trial field testing at companies, educational seminars, and economic incentives, such as reduced insurance costs. In cases where companies or owner-operators did not perceive that driver fatigue was an issue, it was offered that no amount of promotion or examples of successful industry deployment would be likely to lead to immediate purchases of such technology. As possibly related to this, at interview time, no companies had purchased or considered the utilization of an alertness-monitoring system, although various engineered systems (e.g., forward collision warning, lane departure) were in use.

CHAPTER 6. DISCUSSION OF RESULTS

Driver acceptance of a DDWS as assessed through the analysis of survey and focus group data was largely conditional (see Table 113 at chapter end for a summary of findings by RQ). From a usability perspective, participants generally found the device easy and intuitive to operate and understand, and were satisfied with the training they received in this regard. However, apart from the goal of the device to improve safety, participants could not entirely overlook the functional limitations of this technology in its current state. Some would have preferred a smaller device footprint or improved nighttime dial visibility. Yet, by far the most cited issue stemmed from the perceived inaccuracy of DFM alertness warnings. Despite explanations provided to participants of device operational shortcomings during daylight, dawn and dusk, many found it difficult to disregard these instances when evaluating the device. Nevertheless, there existed a small subset of drivers (7 of 33) who found the device to be particularly useful, and who were more satisfied with its performance than the typical FOT participant. Uniformly, this "extreme-favoring" sample subset comprised long-haul drivers, while an opposing group of "extreme-disfavoring" participants contained one long-haul and two line-haul drivers. Although beyond the scope of this effort, further investigation into the possibility that certain trucking operations or driver groups are more well-suited to successful use of a DDWS is warranted.

Results suggesting that the DDWS often did not meet user expectations pertaining to the anticipated operational abilities of a drowsiness detection technology contrasted with those indicating that participants only infrequently ignored the device, and that most did not stop relying on it despite false alarms. This is not entirely counterintuitive, however, as survey findings revealed that participants were most likely to use device warnings when they believed they were "very tired," perhaps choosing to disregard its feedback in other situations. Additionally, it may realistically have been difficult for participants to physically ignore warnings as they were disruptive by design and could not be silenced independently from turning off the unit.

Despite acknowledging on-the-job fatigue as an issue, participants reported that they did not over-rely on the DDWS, which is in line with findings that most participants questioned not only the accuracy of warnings (see discussion of false alerts in Part I, Chapter 5), but also the practical application of the information the device provided for the purpose of fatigue management. Some of the reluctance in using device output in this way may have resulted from the design of the fatigue management training provided as part of the FOT, which did not link the device's visual or auditory alerts with recommendations for fatigue-fighting countermeasures and actions during driving or non-driving time. In contrast to an orthogonal and non-prescriptive fatigue management training approach such as that used in the FOT, company managers who were interviewed anticipated a likely need for comprehensive training and open discussion among stakeholders regarding aspects of policy, data collection, and safety benefits in order to facilitate the driver acceptance of such a device.

Whether influenced by unmet device operational expectations, or because overall, most participants simply did not perceive excessive fatigue while driving during early morning hours,[1]

[1] As a caveat pertaining to this type of data, it should be noted that humans do not tend to perform well in detecting their own levels of drowsiness (Dinges, 1989).

exposure to device feedback over the course of the FOT appeared to impact views regarding its use as a safety-enhancing technology. Attitudes prior to experiencing the device were more optimistic regarding its ability to increase driving safety, whereas after device exposure, participant opinions were generally neutral regarding the extent to which the DDWS had improved their driving safety. Related to this, company managers were also somewhat disinclined to acknowledge driver fatigue as a problem for their organization, but saw the utility of such a device for other agencies. In any case, there was consistent support among managers for a technology that could provide a safety benefit, especially to the extent that this would generate cost savings and might be paired with insurance incentives.

Survey data and focus group session outcomes revealed that participant overall satisfaction levels were less positive regarding device performance and more positive about device potential if its "kinks" were worked out. Nevertheless, as tested during the FOT, participants were hesitant to endorse this device, or to indicate that they would purchase it on their own or in conjunction with their employer. This is not surprising, as interviewed representatives from technology companies indicated that their alertness-monitoring systems had undergone extensive research, design, laboratory, and on-road testing before being deployed. Therefore, initial driver acceptance hesitance and a few remaining bugs are not unexpected as part of the technology-development process. Some degree of driver reluctance in this area may also have been related to concerns over the confidentiality of device-collected information when considering future use of the DDWS. Supporting this view, management at surveyed trucking companies indeed noted that they would be interested in utilizing device alert information, if available, for the purpose of identifying fatigued and unsafe drivers. However, they stressed that this information would not be used punitively.

Regarding the usage of device output during non-driving time, prior to FOT participation over half of the sample intended to use this feedback to plan their sleep/wake schedule; however, after exposure, participants overwhelmingly indicated that they had not employed this information at all during their off-duty hours. In comparison, while on-duty, participants did attend to device feedback, and most often reported responding to warnings by engaging in activities such as opening a window, listening to the radio, and talking on a cell phone or CB. This is in contrast to pre-study predictions made by participants that they would most likely pull off the road to sleep in response to device alerts. Such findings point to the possibility that, rather than using the DDWS as a means to guide sleep times, upon receiving a warning participants may actually have pushed themselves to further extend driving by engaging in activities with limited efficacy for fatigue mitigation. Alternately, there may not have been any opportunity for them to immediately react by pulling off the road to rest, or engage in some other alertness-enhancing strategy. For these reasons, the authors suggested that participants may have benefited from training that included some amount of prescriptive information regarding appropriate and effective fatigue countermeasures in response to device warnings for various driving situations.

Overall, in line with improving driver alertness, a positive change in fatigue-management behavior was the desired outcome of DDWS use. Though most participants did not report feeling more alert when driving at night over the course of the FOT, it may be possible that they were better able to better identify their state of fatigue, given the information provided during fatigue management training in conjunction with device warnings. However, as the majority of the sample stated that use of the DDWS neither improved nor impaired their health, it is not entirely clear what impact, if any, the device had on participants' health and well-being.

The DDWS concept appears to have merit and value both to truck drivers, especially long-haul drivers, and with trucking companies. Results of this study strongly point to the need for significant user involvement in system interface development, including false-alarm suppression, and proper integration into a viable fatigue management program. Providing alerts, but not supplying drivers with an appropriate level of prescriptive information regarding countermeasures, or situations where drivers find themselves unable to employ a preferred countermeasure (e.g., no safe area to pull off the road) will likely lead to frustration, and perhaps non-use, in particular if it is believed that the intent of the device is to allow blame-shifting to drivers in the case of fatigue-related accidents. Additionally, if the driver subpopulation of owner-operators is to benefit, the system will need to be packaged with a basic fatigue management protocol. The fact that there are several manufacturers developing this technology is good news; if done well, there appears to be a waiting market to embrace this safety net device.

Table 113. Summary of Research Questions

Research Question	Measures	Results
RQ 7.1: Does use of the DDWS create extra demands on the driver, such as added stress or increased fatigue?	Survey	• Usual degree of early AM driving stress (test group; VAS: 0 = Not at All; 100 = Extremely) – Rural interstates • Pre-FOT: 63.6% responses = bottom 1/3 of scale • Post-FOT: 66.7% responses = bottom 1/3 of scale – City streets/highways • Pre-FOT: 63.6% responses = bottom 1/3 of scale • Post-FOT: 60.6% responses = bottom 1/3 of scale • FG members reported false alarms as distracting – 2 members reported receiving them every 10 minutes, necessitating numerous device resets • Also see Appendix A
RQ 7.2: For various degrees of fatigue, how often and what duration do drivers require to observe the device in order to understand its output? Additionally, how do assessments of device accuracy change under varying degrees of fatigue?		• No comment button data to assess perceived accuracy during driving or determine degree to which drivers understood device output – Not addressed due to data constraints
RQ 7.3: To what degree are drivers willing to tolerate false alarms? Also, what is their degree of reliance on the system, and their perception of correct alarms?	Survey	• Mean annoyance with false alarms = 62.6 (*SD* = 33.1) • 2/3 of participants reported they did not stop relying on the DDWS due to false alerts – No comment button or log data to address tolerance of false alarms, reliance on system, or perception of correct alarms while driving

Research Question	Measures	Results
RQ 7.4: To what degree were drivers able to recognize DDWS alerts?	Survey	• Degree to which DDWS alerts were easily recognized – Mean = 82.9 (*SD* = 21.9) • Degree to which DDWS got participants' attention quickly – Mean = 80.7 (*SD* = 23.9)
RQ 7.5: Do drivers understand the DDWS' operational limitations?	Survey	• 90.9% of sample found DDWS cab location acceptable • 81.8% of participants aware of situations where DDWS did not operate properly • Reported false warnings per duty period – Mean = 12.4 – Response range = 0 – 100
RQ 7.6: To what extent are the DDWS controls easy and intuitive for drivers to use?	Survey	• Ease/difficulty reading display – Mean rating = 76.9 (*SD* = 23.9) • Ease/difficulty driving using DDWS – Mean rating = 80.5 (*SD* = 23.2)
RQ 7.7: What actions do drivers take to improve their alertness based on the warnings they received? Under what circumstances do they take such actions or not take action?	Survey	• Responses to DDWS warnings – 39.4% reported "almost never" ignoring an alert – 42.4% "occasionally" ignored alerts • How often "unwilling" or "unable" to take actions to improve alertness once DDWS provided a warning – 51.5% "almost never" unwilling – 45.5% "almost never" unable – 30.3% "occasionally" unwilling – 39.4% "occasionally" unable
RQ 8.1: Were drivers able to retain information about device operation and the meanings and uses of its output?	Focus Group	• 2-page quick reference guide noted as especially helpful – In particular for those who experienced a lag between training and actual device use
RQ 8.2: How much time does it take for drivers to feel proficient with the DDWS and its output? How much time does it take for driver to learn both the capabilities and limitations of the device?	Focus Group	• Device use likened to "…turning your VCR on. After 10 minutes you've figured out what it's [going to] do and what it's not [going to] do"

Research Question	Measures	Results
RQ 8.3: Was the training drivers received on the DDWS complete, and was it understandable? Was the fatigue management training drivers received complete, and was it understandable?	Survey	• All participants could easily understand the written materials • 1 participant would have preferred the training materials to be written in Spanish
RQ 9.1: How frequently did drivers indicate they received appropriate warnings based on an accurate alertness assessment?	Survey	• Frequency drivers received an appropriate warning – 66.7% “occasionally” or “frequently”
RQ 9.2: What is the degree to which drivers’ felt that the DDWS enhanced the effectiveness of their fatigue management program and practices?	Survey	• 81.8% found the information provided by DDWS “useful for managing fatigue” • DDWS was/will be helpful for following company’s fatigue management program? – Pre-study: “very much” = 69.7% – Post-study: “very much” = 18.2% • 75.8% did not use DDWS to help with fatigue management activities • 81.8% did not find other applications for DDWS
RQ 9.3: Do drivers view the DDWS as a liability or invasion of privacy?	Survey, Focus Group	• 100% believed DDWS data was kept confidential • 36.4% willing to continue using DDWS after FOT if the information not recorded or available to company/others – Why/why not willing to continue use? ▪ “It can only improve safety” ▪ “It really did change how I drove or how I felt” ▪ “If the device is working like it should” ▪ “It sometimes is a nuisance”

Research Question	Measures	Results
RQ 9.4: Did drivers feel that the DDWS effectively decreased instances of fatigued driving, thus keeping driving skills at an appropriate level for safety?	Survey	• Fallen asleep behind the wheel – Pre-study "yes" ("ever"): TG = 42.4%; C = 53.3% – Post-study "yes" ("during FOT"): TG = 39.4%; C = 40.0% • Involved in accident or incident while working related to sleepiness/fatigue – Pre-study "yes" ("ever"): TG = 13.3%; C = 15.2% – Post-study "yes" (DDWS help avoid accident or close call while turned on): TG = 27.3%
RQ 9.5: Do drivers feel that use of the DDWS will have an adverse effect on their health?	Survey	• 100% of participants who responded believed DDWS did not harm their health
RQ 9.6: How do drivers evaluate their driving safety based on use of the DDWS?	Survey, Focus Group	• DDWS make you more fatigued? No = 93.9% • FG members did not report any adverse safety effects stemming from device use
RQ 10.1: How satisfied were drivers with the DDWS? How useful did drivers find the DDWS to be?	Survey	• Satisfaction with DDWS – Mean rating = 0.15 (*SD* = .99) • Usefulness of DDWS – Mean rating = .78 (*SD* = .95)
RQ 10.2: Are drivers interested in having the DDWS purchased for their entire fleet?	Survey	• 1/3 willing to ask employer to purchase DDWS for fleet • 20.8% willing to ask employer to purchase DDWS for truck
RQ 10.3: Are drivers interested in purchasing the DDWS for their truck, or sharing the cost of device with their employer?	Survey	• 12.5% willing to purchase DDWS themselves • 9.1% willing to share the cost with employer
RQ 10.4: Are drivers willing to endorse the DDWS to drivers within and outside their own company?	Survey	• 1/3 recommended DDWS use to other drivers within company • 1/3 recommended DDWS use to drivers outside company

Research Question	Measures	Results

RQ 11.1: What are drivers' perceived levels of fatigue/alertness while driving with and without the aid of the DDWS?	Survey	• Perceived fatigue levels were not particularly high, whether driving with or without the device
		  • Increases in reported moderate-to-severe fatigue suggest greater perceived fatigue levels after exposure to the DDWS. However, humans are characteristically not very good at self-detecting fatigue and increases were small
		    – TG increase: rural driving = 12.1%; city/highway driving = 6.1 %
		    – CG increase: 4.8% rural driving; city/highway = 2.5%

RQ 11.2: Did drivers initiate behaviors as a result of exposure to the DDWS, including unexpected uses for the device? If so, what are they?	Survey	• 87.9% "almost never" attempted to foil DDWS operation • 12.1% attempted to foil the DDWS "occasionally"
RQ 11.3: How much time do drivers spend monitoring the DDWS under various degrees of fatigue? Where do drivers reallocate time spent monitoring the DDWS?		• Video analysis would be required to analyze changes in driver scan patterns and time spent monitoring the device – Not addressed due to data constraints
RQ 11.4: Was use of the DDWS feedback, potentially to adjust driving style, associated with health improvements (e.g., altered work/rest cycles)?	Survey	• 75.8% felt that the DDWS did not affect their health • 21.2% felt that the DDWS improved their health • 3% felt that the DDWS impaired their health

Research Question	Measures	Results
RQ 11.5: What behavioral changes may have been brought about as a result of extended exposure to the DDWS?	Survey	• How well able to include (0 = not at all; 100 = extremely) the DDWS into fatigue management habits – TG mean: 54.6. Responses more heavily weighted within the top two-thirds of the scale range (78.8 %) • Check off planned and actual uses of the DDWS as applied to non-driving time – TG largest shift in responses was evidenced in a 42.5 % decrease in utilizing the device outside of driving for the purpose of fatigue management – Decrease of 51.5 % was found in responses indicating use of the DDWS for the purposes of sleep/wake scheduling
RQ 12.1: What are managers' personal opinions regarding driver acceptance of the device?	Fleet Management Interview	• General management support for the concept device • Company drivers likely more amenable to accepting this technology than owner-operators – Owner-operators have higher level of autonomy/independence – Owner-operators would have to purchase device on their own; company drivers would have it provided for them
RQ 12.2: What are fleet managers' opinions regarding driver acceptance of the device based on feedback from drivers?	Fleet Management Interview	• 4 of 9 interviewees stated that drivers would likely eventually accept device if its safety benefits are clearly explained – "If you appeal to their egos and make them part of the safety process, they [will] be more willing to accept the device" • Unanimous that education and training on device's advantages/disadvantages would make drivers more likely top accept it

Research Question	Measures	Results
RQ 12.3: Based on information provided at the briefing, what are fleet managers' perceptions of the DDWS' capabilities, advantages, and disadvantages? Does fleet management believe that drivers will approve or disapprove of the DDWS?	Fleet Management Interview	• Interviewees based responses on a hypothetical alertness-monitoring technology – Were not briefed on the actual DDWS in conjunction with the FOT • Main advantage of such a technology was safety benefit associated with reduced accident rates • Device could help determine if drivers are getting adequate sleep or falsifying log books
RQ 13.1: What are fleet managers' perceptions of the potential economic benefits, as attributed to increased safety, of the DDWS?	Fleet Management Interview	• Primary economic benefit would be cost savings associated with accident rate reductions, including: – Loss of life – Loss of equipment – Driver injuries and associated medical costs – Vehicle property damage – Workman's compensation claims
RQ 13.2: What federal incentives would fleet management like to have associated with the adoption of DDWS?	Fleet Management Interview	• 2 federal incentives in particular that would make recommendation to purchase more likely – Tax-credits – Federal grants
RQ 13.3: What insurance incentives would fleet management like to have associated with the adoption of DDWS?	Fleet Management Interview	• Lowered insurance rates for improved safety record • "If our insurance carrier told us they would not cancel our policy if we had this device on board our trucks, it would make us more likely to consider purchasing it" • Drivers would have to be protected from insurance companies obtaining and using device information for lawsuit
RQ 13.4: How much would fleet management be willing to pay for a DDWS?	Fleet Management Interview	• 4 of 9 did not offer specific price point – Keeping the price "low" would be critical • Other 5 interviewees responses ranged from \$200-\$1500 per unit – 1 participant quoted a unit price of \$500 for small fleets (4 – 5 trucks) with a discount for large fleets of 100+

Research Question	Measures	Results
RQ 14.1: How much training do fleet managers believe is required for their drivers to make the best use of the DDWS? And, are they willing to provide it?		• Difficult to obtain meaningful information to answer this question. Fleet management interviewees did not directly experience device and would likely defer to the training department/experts or to whatever vendor recommends/provides – Not addressed due to data constraints

RQ 14.2: Does fleet management plan to monitor DDWS alerts? If so, how do they plan to use this information? What might the influence of this be on acceptance by drivers?	Fleet Management Interview	• All interviewees expressed interest in collecting and monitoring device data
		  – Use to identify instances of fatigued driving and ultimately try to mitigate those situations to avoid associated accidents
		  – 1 participant would only want the data if available in real time (value diminishes over time)
		• Eye closure data useful for educating drivers about proper rest
		• Stressed that data would not be used to punish or dismiss drivers

RQ 14.3: What sort of policy, if any, does fleet management plan to implement regarding required driver behavior following DDWS alerts?	Fleet Management Interview	• Management needs to work collaboratively with employees to develop policies regarding mandatory behavior after a warning
		• Commonly suggested policies included:
		  – Pull off the road
		  – Take a break
		  – Obtain rest
		• 1 respondent proposed changing their route if a driver is getting repeated warnings, or switching them from long- to short-haul (if possible within the company)

RQ 14.4: To what extent is fleet management willing to modify their FMPs based on FOT findings?		• Difficult to obtain meaningful information to answer this question. Fleet management interviewees did not directly experience device and would have had difficulty absorbing FOT results over the phone – Not addressed due to data constraints

Research Question	Measures	Results
RQ 15.1: What are fleet managers' concerns regarding performance of the DDWS?	Fleet Management Interview	• Most critical performance concerns: – Reliability – Accuracy – Maintainability • Getting the "bugs worked out" prior to operational use is extremely important – Too many false alarms would antagonize drivers, possibly leading them to disengage the device
RQ 15.2: What features does fleet management desire in the DDWS? How much will fleet management pay for these features?	Fleet Management Interview	• 4 of 9 interviewees gave suggestions for additional device features including: – Monitor driver physical status – Detect alcohol on the driver's breath – Monitor driver distraction (in addition to alertness) – Combine measurements of eye closure with measurements of vehicle parameters (i.e., lane movement/lane departure and following distance)
RQ 15.3: Does fleet management seek other improvements in the DDWS?		• Fleet management interviewees did not directly experience device. Question is device-specific, and therefore would not yield useful information – Not addressed due to data constraints

Research Question	Measures	Results

RQ 16.1: Is there a distinct introductory period anticipated for the DDWS? If so, how long will it last? What will be the introductory price for DDWS? What will be the maximum penetration level of DDWS in trucks (heavy vehicles) during the introductory period? When will suppliers and truck manufacturers consider deployment of the system?	Deployment Interview	• 2 representatives from development companies discussed deployment of their respective products
		• 1 system had been available for use for ~6 months
		  – Currently used for road transport applications in light CMVs, heavy trucks, and off-road equipment
		  – Device was tested during introductory period of “over 1 year”
		    • Used ~12 vehicles operating in different environments with different partners
		    • Also used during this period by several customers for driver behavior research
		• Other system was at pre-production level
		  – OEM currently considering deployment of technology as early as 2010
		  – Interest primarily expressed by CMV segment
		  – Deployment introductory period not planned – intend to roll out product across select platforms and then expand to higher volume platforms

Research Question	Measures	Results

RQ 16.2: What will be the price range of DDWS deployments over the next 15 years? What is the expected rate of DDWS deployment over the next 15 years? Will the DDWS be available as an option for all trucks, or only more expensive models?	Deployment Interview	• 2 representatives from development companies discussed long-term deployment and expected future pricing of their respective products
		• "As more vehicle systems become automated, the need for driver alertness warning systems will only become more critical. With increased automation, the driving task becomes increasingly boring, increasing the dangers of fatigue and distraction."
		• 1 interviewee stated that "pricing information is generally treated confidentially and depends on volume" but that they "do not see cost as a barrier to mass-market uptake of the technology"
		• These systems "have always been introduced first in expensive vehicles and have worked their way into more affordable vehicles"
		    – Expects availability in all upper-market passenger/commercial heavy vehicles within 5-10 years
		    – Availability in most vehicles on the market within 10-15 years
		• Integration of other functions/technologies (e.g., driver recognition, wireless tracking) will impact penetration

Research Question	Measures	Results
RQ 17.1: What are organized labor's concerns regarding the DDWS?	Deployment Interview	• Liability issues are a possibility, especially when there is awareness of a problem (e.g., the driver is drowsy or nodding off) and no corrective action was taken • False readings and frequent alerts will likely lead to negative reactions by drivers and an overall lack of acceptance • Audible or visual signals from device should not distract or interfere with the safe operation of the truck – Signals should be coordinated with other on-board safety technologies (e.g., lane departure systems, collision avoidance systems) – "Only the highest of all priority alerts should be given to the driver, whether [it be] collision potential, tire pressure loss, electrical fire, head-nodding, brake loss, etc." • A device that only detects driver fatigue may not be enough to gain motor carrier investment – Ability to detect vigilance, distraction, and other general "road conditions" may add appeal for carriers

Research Question	Measures	Results

RQ 18.1: Would there be a need for non-research activities beyond the FOT in order to expedite DDWS deployment, assuming it is desirable to expedite deployment? What else, if anything, should be done in drowsy driver related crash avoidance research?	Deployment Interview	• Research question included as part of the phone survey of fleet management staff
		• Opinions varied regarding how to effectively promote and successfully deploy this type of device
		    – 4 interviewees suggested that carriers use the technology on a trial basis and publicly report the results (1 respondent felt this should be government funded)
		    – Educational seminars were suggested as a way to "sell" the technology and promote management buy-in
		    – 2 participants recommended economic incentives (e.g., reduced insurance costs)
		    – It may be impossible to sell the use of this device to some companies with excellent safety records and no perceived issues with driver fatigue, regardless of type of promotion
		• Drivers in particular and companies in general are very different from one another – broad research that could be universally applied to all segments of the industry would be difficult to design
		• "Enough research pertaining to HOS has already been done, and the industry should review and learn from it"
		    – More research into the effects of sleep apnea on driver behavior would be helpful, since it is an issue of particular concern to the industry
		• At the time of interview, no companies had purchased or considered the use of an alertness monitoring system
		    – 1 long-haul company had purchased and were using a forward collision warning system
		    – 1 long-haul company had installed a lane departure warning system, as well as a front/side collision warning system on several of its units

REFERENCES

Bekiaris, E., Nikolaaou, S., & Mousadakou, A. (2004). *System for effective Assessment of driver vigilance and Warning According to traffic risK Estimation (AWAKE): Design guidelines for driver drowsiness detection and avoidance* (Deliverable 9_1). Thessaloniki, Greece: Helenic Institute of Transport.

Dinges, D. F., & Mallis, M. M. (1998). Managing fatigue by drowsiness detection: Can technological promises be realized? In L. Hartley (Ed.), *Managing Fatigue in Transportation. Proceedings of the Third International Conference on Fatigue in Transportation* (pp. 209-229). Oxford: Pergamon.

NHTSA, ITS Joint Program Office. (1997). *Report to Congress on the National Highway Safety Administration ITS Program, Program progress during 1992-1996 and strategic plan for 1997-2002*. Washington, DC: National Highway Traffic Safety Administration.

Stearns, M. D., & Boyle, L. N. (2002). *Evaluation of GM/Delphi Rear-end Collision Countermeasures: Revised driver acceptance framework and survey data needs* (Technical Information Exchange). Cambridge, MA: Volpe National Transportation Systems Center.

Wehr, T.A., Aeschbach, D., & Duncan, W.C., Jr. (2001). Evidence for a biological dawn and dusk in the human circadian timing system. *Journal of Physiology, 535* (3), 937-951.

Whitlock, A. (2002). *Driver vigilance devices: Systems review final report*. Report No. 03 T024 QUIN 22. London: Rail Safety Standards Board.

APPENDIX F. NONPARAMETRIC ANALYSES

The small size of the survey sample and resulting non-normal data distributions necessitated the use of non-parametric statistical analysis procedures where meaningful and informative comparisons existed within the data. Mann-Whitney *U* was used for independent-samples comparisons between Control and Test Groups. Paired-samples comparisons of responses to congruent pre- and post-survey items were assessed using the Wilcoxon Signed-Ranks test. For both sets of analyses, the Sidak method was used to correct for the increased probability of Type I errors when conducting multiple comparisons. This correction resulted in an adjusted alpha of 0.004 in each case.

Table 114. Ease of Use Nonparametric Comparisons

Independent Samples Comparisons						
		Median		*Mann-Whitney*		
Survey Item Summary	Survey	Control	Test	*U*	*z*	*p* < 0.004*
VAS ratings regarding usual degree of stress while driving during the early morning hours through rural interstates (0 = Not at all; 100 = Extremely)	*Pre*	5.0	26.0	173.0	-1.66	0.097
	Post	20.0	27.0	200.0	-1.06	0.290
VAS ratings regarding usual degree of stress while driving during the early morning hours through city streets and highways (0 = Not at all; 100 = Extremely)	*Pre*	16.0	25.0	220.0	-0.61	0.541
	Post	47.0	18.0	203.0	-0.99	0.322

Paired Samples Comparisons					
		Median		*Wilcoxon Signed-Ranks*	
Survey Item Summary	Group	Pre-Survey	Post-Survey	*Z*	*p* < 0.004*
VAS ratings regarding usual degree of stress while driving during the early morning hours through rural interstates (0 = Not at all; 100 = Extremely)	*Control*	5.0	20.0	-0.57	0.570
	Test	26.0	27.0	-0.47	0.636
VAS ratings regarding usual degree of stress while driving during the early morning hours through city streets and highways (0 = Not at all; 100 = Extremely)	*Control*	16.0	47.0	-2.73	0.006
	Test	25.0	18.0	-0.25	0.802
Usefulness of the DDWS for the purpose of fatigue management while at work (1 = Not at all useful; 4 = Extremely useful)	*Control*	--	--	--	--
	Test	3.0	2.0	-3.11	0.002*

* Alpha adjusted using Sidak method to correct for Type I error when conducting multiple comparisons. Statistically significant findings indicated by $p < 0.004$.

Table 115. Ease of Learning Nonparametric Comparisons

Independent Samples Comparisons						
		Median		*Mann-Whitney*		
Survey Item Summary	Survey	Control	Test	*U*	*z*	*p* < 0.004*
Perceived completeness of the fatigue management training (0 = Not at all; 100 = Extremely)	*Pre*	--	--	--	--	--
	Post	73.0	75.0	244.0	-0.78	0.938

* Alpha adjusted using Sidak method to correct for Type I error when conducting multiple comparisons. Statistically significant findings indicated by $p < 0.004$.

Table 116. Perceived Value Nonparametric Comparisons

Independent Samples Comparisons						
		Median		*Mann-Whitney*		
Survey Item Summary	Survey	Control	Test	*U*	*z*	*p* < 0.004*
How well do you now think you will be able to follow your company's fatigue management program? (1 = Not at all; 4 = Very much)	*Pre*	--	--	--	--	--
	Post	3.0	3.0	161.0	-2.03	0.042

Paired Samples Comparisons					
		Median		*Wilcoxon Signed-Ranks*	
Survey Item Summary	Group	Pre-Survey	Post-Survey	*Z*	*p* < 0.004*
Perceived benefits regarding driving safety related to DDWS use (1 = Not at all; 4 = Extremely)	*Control*	--	--	--	--
	Test	3.0	2.0	-2.49	0.013

* Alpha adjusted using Sidak method to correct for Type I error when conducting multiple comparisons. Statistically significant findings indicated by $p < 0.004$.

Table 117. Driver Changes Nonparametric Comparisons

Independent Samples Comparisons						
		Median		*Mann-Whitney*		
Survey Item Summary	Survey	Control	Test	*U*	*z*	*p* < 0.004*
VAS ratings regarding usual degree of fatigue while driving during early morning hours through rural interstates (0 = Not at all; 100 = Extremely)	*Pre*	13.3	23.0	179.5	-1.20	0.231
	Post	20.5	30.0	200.0	-1.06	0.291
VAS ratings regarding usual degree of fatigue while driving during early morning hours through city streets and highways (0 = Not at all; 100 = Extremely)	*Pre*	7.0	19.0	151.5	-1.54	0.124
	Post	22.0	22.0	246.0	-0.03	0.973
VAS ratings regarding usual degree of alertness while driving during early morning hours through rural interstates (0 = Not at all; 100 = Extremely)	*Pre*	58.3	63.0	201.5	-0.54	0.591
	Post	85.0	71.5	180.5	-1.49	0.136
VAS ratings regarding usual degree of alertness while driving during early morning hours through city streets and highways (0 = Not at all; 100 = Extremely)	*Pre*	59.3	75.0	180.5	-1.18	0.240
	Post	74.0	76.5	242.0	-0.12	0.903

Paired Samples Comparisons					
		Median		*Wilcoxon Signed-Ranks*	
Survey Item Summary	Group	Pre-Survey	Post-Survey	*z*	$p < 0.004$*
VAS ratings regarding usual degree of fatigue while driving during early morning hours through rural interstates (0 = Not at all; 100 = Extremely)	*Control*	13.3	20.5	-1.29	0.198
	Test	23.0	30.0	-1.22	0.221
VAS ratings regarding usual degree of fatigue while driving during early morning hours through city streets and highways (0 = Not at all; 100 = Extremely)	*Control*	7.0	22.0	-1.43	0.152
	Test	19.0	22.0	-0.40	0.688
VAS ratings regarding usual degree of alertness while driving during early morning hours through rural interstates (0 = Not at all; 100 = Extremely)	*Control*	58.3	85.0	-1.26	0.209
	Test	63.0	71.5	-0.80	0.421
VAS ratings regarding usual degree of alertness while driving during early morning hours through city streets and highways (0 = Not at all; 100 = Extremely)	*Control*	59.3	74.0	-1.88	0.060
	Test	75.0	76.5	-0.79	0.432

* Alpha adjusted using Sidak method to correct for Type I error when conducting multiple comparisons. Statistically significant findings indicated by $p < 0.004$.

APPENDIX G. "EXTREME ATTITUDES" PARTICIPANT SUBSET

Targeted data analysis resulted in a small subset of participants who viewed the DDWS as particularly useful and were especially satisfied with the device, and a parallel, but opposing, subset of drivers who reported extreme dissatisfaction, and that they did not find the device very useful. In this appendix, we compare these groups to one another and the full Test Group.

METHODS

We selected a subset of Test Group participants (n = 33) who reported "extreme attitudes" regarding the device based on computed composite scores for the Driver Acceptance Scale (see Advocacy section). Those participants with scores that were greater than or equal to 1 on both the usefulness and satisfaction subscales comprised the "extreme-favoring" group (Favoring Group), while those with scores less than or equal to -1 on both the usefulness and satisfaction subscales formed the "extreme-disfavoring" group (Disfavoring Group). As illustrated in the top-right, outlined region of Figure 178, the Favoring Group consists of seven participants. Similarly, the bottom-left, outlined region contains three participants as part of the Disfavoring Group.

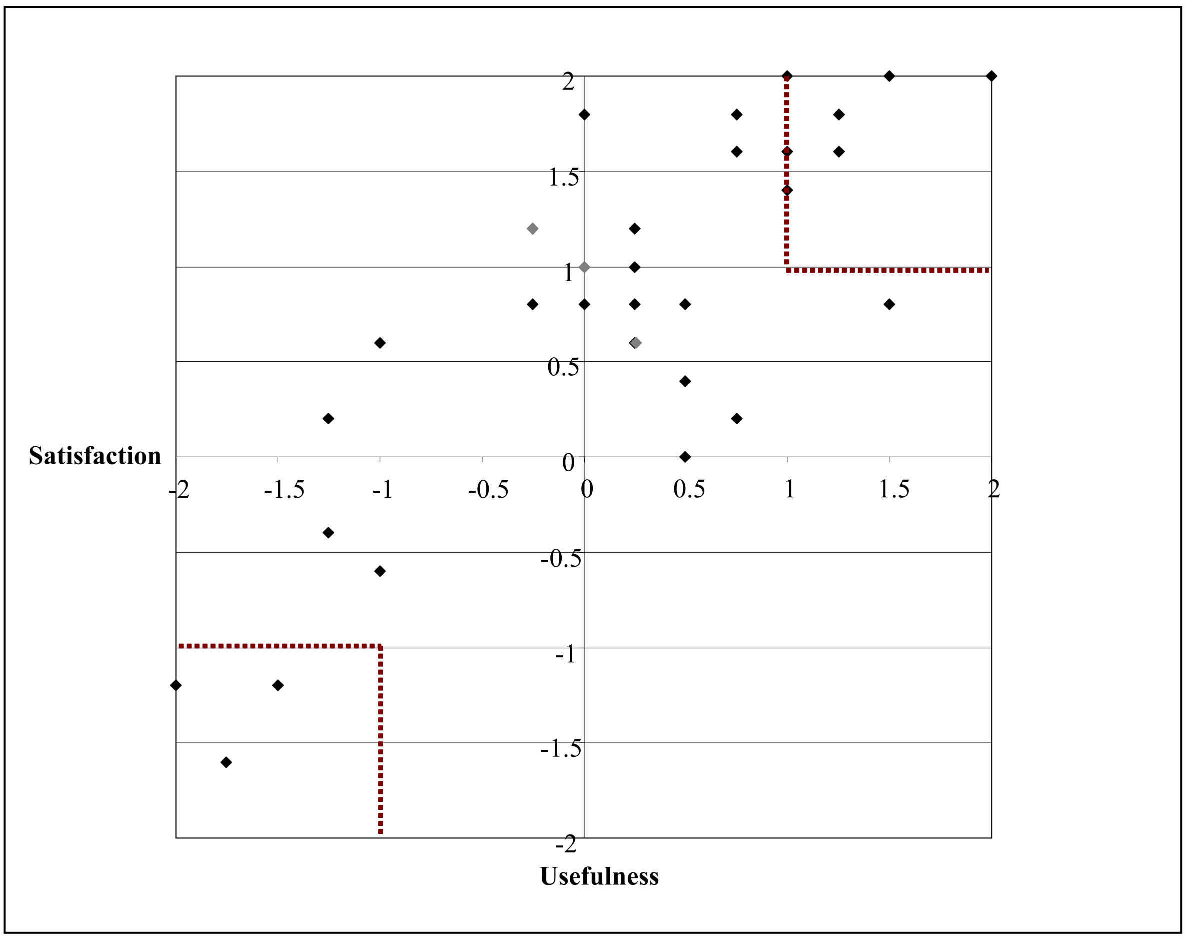


Figure 178. Plot of Paired Satisfaction (x) by Usefulness (y) Scores on Driver Acceptance Scale (Outline Denotes Extreme Attitudes Subset)

In this appendix, we report demographic information for these ten participants in comparison to the Test Group, as well as their responses to 10 post-study survey items as collaboratively selected with input from the FOT conductor. Comparisons between the extreme-attitudes groups and the full Test Group on these items assess attitudes towards the DDWS pertaining to safety outcomes, warning appropriateness, usability, and satisfaction levels. Findings are reported descriptively, given the small number of participants who comprised this subset.

RESULTS

Table 118 presents comparative demographic information for the extreme-attitudes subset and the full sample of Test Group participants. The sample subset was similar in composition to the broader sample, though reported an average of 2.6 fewer years of CMV driving experience, and resulted in less response variability regarding marital status and ethnic composition.

Table 118. Demographic Comparison Information

Demographic Variable	% Response	
	“Extreme-Attitudes” Subset (n = 10)	Test Participant Group (n = 33)
Gender		
Male	100	97.0
Female	--	3.0
Marital Status		
Married or living with partner	80.0	72.7
Single or widowed	20.0	6.1
Divorced or separated	--	15.2
Missing	--	6.1
Ethnic Background		
Caucasian	50.0	63.6
African American	40.0	30.3
Native American	10.0	3.0
Asian American	--	3.0
Missing	--	--
Highest Achieved Education Level		
Did not complete high school	10.0	18.2
High school graduate	30.0	42.4
Beyond high school	40.0	33.3
Missing	20.0	6.1
Haul Type		
Long	80.0	69.7
Line	20.0	30.3
Trucking Company		
J.B. Hunt	50.0	39.4
Howell’s	30.0	33.3
Pitt-Ohio Express	20.0	27.3
	Mean (*SD*)	
Age (years)	40.8 (8.7)	42.2 (8.3)
CMV driving experience (years)	10.1 (6.0)	12.7 (8.7)

Further, participants who comprised the extreme-attitudes groups evidenced demographic profiles that were similar to one another, with one exception: the Favoring Group was entirely made up of long-haul drivers, whereas two of the three in the Disfavoring Group were line-haul drivers. A comparison of Driver Acceptance Scale scores for the full Test Group revealed a similar pattern. The mean satisfaction composite score for long-haul drivers was 0.25 compared to -0.20 for line-haul drivers, and the mean usefulness composite score was 1.20 compared to 0.16 for long- and line-haul drivers, respectively.

Comparison of Attitudes Towards the DDWS

Comparison of the Favoring and Disfavoring Groups with one another and in contrast to the full Test Group sample was undertaken with an interest in determining the degree to which their extreme attitudes towards the device were consistently strong across a number of survey items, as compared with the larger sample.

As indicated in the main body of the report, responses regarding general satisfaction with the DDWS using a 100-point VAS were nearly evenly distributed between halves of the scale. The same pattern existed for the extreme-attitudes participant subset, as the range of scores for the Favoring Group fell completely within the top half of the scale, and the scores for the Disfavoring Group were uniformly low (all ratings under 10). Related to satisfaction, device recommendation to drivers within and outside one's company resulted in two-thirds of the Test Group in each case indicating that they had not, at the time of the survey, recommended the DDWS to others. Responses for both items were similar within the favoring subsample; in the case of the Disfavoring Group, the device had uniformly not been recommended.

With regard to device interactions, the full Test Group sample responded most frequently having "almost never" or "occasionally" ignored the DDWS. These categories were also utilized by the Favoring Group for all responses, and not at all by the Disfavoring Group, whose members indicated "frequently" and "almost always" ignoring warnings. Reported ease or difficulty with driving using the device evidenced a similar split, whereby the Favoring Group responded more similarly on a VAS (0 = very difficult; 100 = very easy) to the overall Test Group sample, and respective means of 80.5 (*SD* = 23.2) and 90.0 (*SD* = 12.0) indicated positive attitudes towards driving with the DDWS. The Disfavoring Group evidenced much greater response variability, providing ratings ranging from 20-97 and a lower mean of 55.0 (*SD* = 39.0). In evaluating safety outcomes, the mean perceived agreement (100 = strongly agree) regarding increased driving safety as related to the device was neutral for the Test Group; however, the responses for the extreme-attitudes groups were not. Ratings for the Disfavoring Group ranged from 2-15, while ratings for the Favoring Group ranged from 61-100.

Responses reflecting attitudes towards DDWS warnings among those in the extreme-attitudes groups were uniformly negative for the Disfavoring Group, having provided a range of scores between 80-100 (100 = extremely annoyed) regarding receiving incorrect warnings. For Likert-type scale items pertaining to the frequency of perceived correct alertness measurement and incorrect fatigue warnings, in all cases, the Disfavoring Group responded uniformly using the negative scale anchor. The Favoring Group fell towards the opposing side of both response scales, and was positive to a degree beyond the attitudes offered by the Test Group sample as a whole. Finally, in an item assessing device usefulness for fatigue management while at work, it was again the case that the Disfavoring Group uniformly responded at the low end of the scale (i.e., “not at all useful”), while the Favoring Group responded exclusively using the higher end of the scale. In comparison, the full Test Group sample provided most responses at “a little useful,” an anchor towards the middle of the scale that was utilized by neither of the extreme-attitudes groups.

Overall, favoring and disfavoring attitudes were reliably positive or negative, depending on group membership, across a variety of survey items assessing a number of aspects of driver acceptance. It was uniformly the case that the Favoring Group responded considerably more positively towards the device than the Disfavoring Group, which was particularly small, but remarkably consistent (i.e., low response variability). Additionally, it was most often true that the extreme-attitudes groups responded to survey items more strongly in their respective attitudinal directions than was evidenced by the Test Group sample as a whole. Taken together, these findings suggest that the members of the favoring and Disfavoring Groups may be characterized as unvaryingly polarized in their views regarding the device.

Of additional interest, while demographically similar in most ways, the single bifurcation within this subgroup (and larger Test Group) that holds reasonable explanatory potential with regard to predicting driver reactions towards the device lies with operation type. Differences between long-haul and line-haul operations include a most obvious example of the possibility for marked variation in work schedules for over-the-road versus out-and-back deliveries. In the case of long-haul drivers, this difference can easily manifest itself as many hours behind the wheel with minimal rest, and the greater potential for experiencing fatigue in comparison to an out-and-back operation. While beyond the scope of the current effort to fully address or attempt to resolve, further investigation of this proposition is warranted.

APPENDIX H. FOT SURVEYS

Pre-Participation Survey

Instructions: Thank you for taking the time to fill in this survey. It should take you about one minute to complete. Please leave blank any information you are not sure of or do not feel comfortable giving. We will use this information to help us determine which drivers are best suited to participate in the study. The information you give will be kept confidential and will not be shared with any of your managers or other drivers.

Name: ___________________________________ Date:___________

1. What is your age?: _____ years
2. Gender *(Check one)*: ☐ Male ☐ Female
3. What is your eye color? ______________
4. Height: _____ feet _____ inches
5. Weight: _____ pounds
6. Build *(check one)*: ☐ Small frame ☐ Medium frame ☐ Large frame
7. Do you wear contact lenses? *(Check one)* ☐ No ☐ Yes (Lens color:______________)
8. Do you wear glasses at *night* when driving? ☐ No ☐ Yes – if *yes*: ☐ Metal frame ☐ Plastic frame
9. Do you use a hearing aid? ☐ No ☐ Yes, – if *yes*: in which ear(s)? ____________
10. Which of the following groups is most representative of your background? *(Check one)*

    ☐ African/American ☐ Asian/American ☐ Caucasian/American
    ☐ Hispanic/American ☐ Native American

11. Is English your language of preference for: reading? *(check one)* ☐ No ☐ Yes
    speaking? *(check one)* ☐ No ☐ Yes
12. How long have you been driving commercial vehicles? _____ years _____ months
13. How long have you been working for this company? _____ years _____ months
14. How long did you work for your previous employer (your job before this one)? _____ years _____ months
15. Are you a member of a union *(Check one)*? ☐ No ☐ Yes, which one?________________
16. Type of license and endorsements held: License:________________ Endorsements: _______________
17. I would like to participate in this experiment *(Check one)*: ☐ No ☐ Yes

If yes: What is your phone number? __________ What is the best time to reach you? ____________

The following three questions refer to your behavior while sleeping, trying to sleep, or while feeling sleepy. Please check one box for each question.

During the last month have you done or been told that you do any of the following?

	Never	Rarely (< 1 time per week)	1-2 times per week	3-4 times per week	5-7 times per week	Don't know
Loud Snoring						
Snorting or gasping						
Your breathing stops, or you struggle for breath						

Pre-Study Survey

Thank you for taking the time to complete this survey. It should take you about 30 minutes to fill out. The information that you give will not be shared with any of your managers or other operators, and your responses will be kept strictly confidential. This survey asks you questions about you, your health and well being, use of the drowsy driver warning system (DDWS), and your current work schedule and environment. You are not required to give answers to any survey items that make you uncomfortable. However, the more data you give, the better we will be able to understand the potential benefits of the device for drivers like yourself.

Operator Information:
The questions below about yourself and your background will help us to learn about what types of people are participating in this study.

1. Some people feel older or younger than their actual age. How old do you feel? ______ years

2. Please write down about how many years of each you have finished. Then check the correct box under whether or not you graduated.

	Number of years	*Did You Graduate?*	
High School	_________	☐ Yes	☐ No
Technical School	_________	☐ Yes	☐ No
Subject: __________________			
College	_________	☐ Yes	☐ No

3. What is your present marital status? *(Check one)*

☐ Single ☐ Divorced
☐ Married ☐ Separated
☐ Living with Partner ☐ Widowed

4. How many people depend on your income, including yourself? _________

5. How many other people in your household depend on you for a large part of their personal care? ______________

Health and Well Being:
The questions below will help us to better understand the status of your health. Please carefully follow any additional instructions.

1. Draw a line through the bar to show how healthy you usually feel.

Very healthy |————————————————| Not healthy at all

2. *Instructions:* The questions in this scale ask you about your feelings and thoughts <u>during the last month</u>. In each case, please check how often you felt or thought a certain way.

   a. In the last month, how often have you felt that you were unable to control the important things in your life?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   b. In the last month, how often have you felt confident about your ability to handle your personal problems?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   c. In the last month, how often have you felt that things were going your way?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   d. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

3. The questions below concern some of your body functions. Please try to answer each question by checking the appropriate box that shows how often you experienced each of the items *within the past year*. Check <u>one space</u> for <u>each item</u> below.

Items:	*Never*	*Less than once a month*	*Once or twice a month*	*Once a week*	*2 or 3 times a week*	*About every day*
Back Pain						
Acid indigestion, heartburn, or acid stomach						
Gastrointestinal problems						
Constipation						
Tight feeling in stomach						
Blurred vision						
Trouble falling asleep						
Feeling tired upon awaking						
Trouble staying asleep						
Other:						
Other:						

4. What health problems or chronic conditions have you been diagnosed with? Are they currently being treated? For example, diabetes, sleep apnea, heart condition, cancer, restless legs, etc.

Health Issue	Being Treated? (check one)	
______________________	☐ Yes	☐ No
______________________	☐ Yes	☐ No
______________________	☐ Yes	☐ No

5. How often do you usually have a drink of …? *(Check only one space for each drink item listed below). Then, place a ✓ in the shaded column if you use that beverage to stay awake while driving.*

Items:	More than 5 times a day	3 to 5 times a day	1 to 2 times a day	Less than once a day	Never or almost never	*Used to stay alert while driving?*
Beer						
Wine						
Other Liquor						
Coffee						
Tea						
Cola/ Soft Drink						

6. Do you smoke? ☐ Yes ☐ No

If yes, how much do you smoke per day, on average? *(Enter a number)*

How many packs of cigarettes per day __________
How many cigars per day __________
How many pipes of tobacco per day __________

Do you use smoking as a way to remain alert? ☐ Yes ☐ No

7. Please list the medications you take regularly and for what reason you take them. Also, please check the box for those that make you feel sleepy. For example, medications including those for high blood pressure, cholesterol, allergies, etc.

Medication:	*What reason?*	*Makes me sleepy:*
__________	______________________	☐
__________	______________________	☐
__________	______________________	☐
__________	______________________	☐
__________	______________________	☐

8. *Instructions:* Read each question carefully. Check the box (✓) underneath the most appropriate time frame.

a. If you were entirely free to plan your evening and had no commitments the next day, at what time would you choose to go to bed?

8 – 9 PM	9 – 10:15 PM	10:15 PM – 12:30 AM	12:30 – 1:45 AM	1:45 – 3 AM

b. You have to do 2 hours physically hard work. If you were entirely free to plan your day, in which of the following periods would you choose to do the work?

8 – 10 AM	11 AM – 1 PM	3 – 5 PM	7 – 9 PM

c. For some reason you have gone to bed several hours later than normal, but there is no need to get up at a particular time the next morning. Which of the following is most likely to occur?

- ☐ Will wake up at the usual time and not fall asleep again
- ☐ Will wake up at the usual time and doze thereafter
- ☐ Will wake up at the usual time but will fall asleep again
- ☐ Will not wake up until later than usual

d. You have a 2-hour test to take, which you know will be mentally exhausting. If you were entirely free to choose, in which of the following periods would you choose to sit the test?

8 – 10 AM	11 AM – 1 PM	3 – 5 PM	7 – 9 PM

e. If you had no commitments the next day and were entirely free to plan your own day, what time would you get up?

5 – 6:30 AM	6:30 – 7:45 AM	7:45 – 9:45 AM	9:45 – 11:00 AM	11:00 AM – 12:00 PM	12 PM or later

f. A friend has asked you to join him twice a week for a work-out in the gym. The best time for him is between 10pm - 11pm. Bearing nothing else in mind other than how you normally feel in the evening, how do you think you would perform?

- ☐ Very well
- ☐ Reasonably well
- ☐ Poorly
- ☐ Very poorly

g. One hears about 'morning' and 'evening' types of people. Which of these types do you consider yourself to be?

- ☐ Definitely morning type
- ☐ More a morning than an evening type
- ☐ More an evening than a morning type
- ☐ Definitely an evening type

9. *Instructions:* Below are eighteen statements that people sometimes make about themselves. Please indicate whether or not you believe each statement applies to you by marking whether you:

1 = Strongly Disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Strongly Agree

_____1. Are you inclined to keep in the background on social occasions?
_____2. Do you like to mix socially with people?
_____3. Do you sometimes feel happy, sometimes depressed, without any apparent reason?

_____4. Are you inclined to limit your acquaintances to a select few?
_____5. Do you like to have many social engagements?
_____6. Do you have frequent ups and downs in mood, either with or without apparent cause?

_____7. Would you rate yourself as a happy-go-lucky individual?
_____8. Can you usually let yourself go and have a good time at a party?
_____9. Are you inclined to be moody?
_____10. Would you be very unhappy if you were prevented from making numerous social contacts?
_____11. Do you usually take the initiative in making new friends?
_____12. Does your mind often wander while you are trying to concentrate?
_____13. Do you like to play pranks upon others?
_____14. Are you usually a "good mixer?"
_____15. Are you sometimes bubbling over with energy and sometimes very sluggish?
_____16. Do you often "have the time of your life" at social affairs?
_____17. Are you frequently "lost in thought" even when you should be taking part in a conversation?
_____18. Do you derive more satisfaction from social activities than from anything else?

10. Have any major stressful events taken place in your life within the last year? ☐ Yes ☐ No
If yes, what was the event(s)? ____

DDWS Device:
The questions below will help us to better understand what you think about the DDWS device before you use it.

1. Please describe in your own words how the DDWS may help you handle your fatigue while driving:____
____
____

2. How will you know if the device is not working right? *(Check any that apply)*

☐ Device gives no warnings
☐ Device gives constant warnings
☐ No visual display present
☐ Other: ____

3. How do you plan on using DDWS warnings <u>immediately</u> while you are driving? *(Check any that apply)*

   ☐ I will stop driving and pull over to sleep
   ☐ I will drink a beverage that has caffeine in it
   ☐ Other: ______________________________

4. <u>Along with</u> a fatigue management program, how useful do you think the DDWS will be in helping you manage your fatigue while at work?

   ☐ Not at all useful ☐ A little useful ☐ Quite useful ☐ Extremely useful

5. <u>Without</u> a fatigue management program, how useful do you think the DDWS will be in helping you manage your fatigue at work?

   ☐ Not at all useful ☐ A little useful ☐ Quite useful ☐ Extremely useful

6. The DDWS is designed for use by many different types of drivers. It may not always correctly measure the alertness level of each person. How will you know if this device is giving you correct information about your level of alertness? *(Check any that apply)*

   ☐ It will only sound warnings when I am fatigued
   ☐ It will never sound warnings when I am alert
   ☐ Other: ______________________________

7. Do you feel that using this device will <u>harm</u> your health? ☐ Yes ☐ No
   If <u>yes</u>, how? ______________________________

8. How much do you believe the DDWS will improve the <u>safety</u> of your driving?

   ☐ Not at all ☐ A little ☐ Quite a bit ☐ Extremely
   Why? ______________________________

9. Draw a line through the bar to show how <u>likely</u> it is for you to be involved in a fatigue-related accident or incident when <u>not</u> using the DDWS.

   Not at all |————————————————| Extremely

10. Draw a line through the bar to show how <u>likely</u> it is for you to be involved in a fatigue-related accident or incident when <u>using</u> the DDWS.

    Not at all |————————————————| Extremely

11. Under each clock time, use the scale below to write down how alert you expect to feel when driving at that time through rural interstates using the DDWS.

1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

12. Under each clock time, use the scale below to write down how alert you expect to feel when driving at that time through city streets and highways using the DDWS.

1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

13. Currently, how often do you look at the instrument panel in the cab?

- ☐ Once or more a minute
- ☐ Once every 5 to 10 minutes
- ☐ Once every 15 to 30 minutes
- ☐ Once every 30 to 60 minutes

14. Before you get a warning, how often do you expect to look at the DDWS?

- ☐ Once or more a minute
- ☐ Once every 5 to 10 minutes
- ☐ Once every 15 to 30 minutes
- ☐ Once every 30 to 60 minutes
- ☐ Only when an alarm is sounded

15. Once it has given a warning, how often do you expect to look at the DDWS?

- ☐ Once or more a minute
- ☐ Once every 5 to 10 minutes
- ☐ Once every 15 to 30 minutes
- ☐ Once every 30 to 60 minutes
- ☐ Only when an alarm is sounded

16. How often would the DDWS have to give you an incorrect warning for you to stop using it?

- ☐ More than once an hour
- ☐ Every couple of hours
- ☐ Several times a night
- ☐ Once a week
- ☐ Less than once a week

Current Work Schedule and Environment:
The questions below will help us to better understand the impact of your work schedule and work environment.

Instructions: The questions below ask about your stress level when driving.

1. Draw a line through the bar to show how stressed you usually feel when driving in the early morning hours through rural interstates:

Not at all |————————————————| Extremely

2. Under each clock time, use the scale below to write down how stressed you usually feel when driving at that time through rural interstates.

1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

3. Draw a line through the bar to indicate how stressed you usually feel when driving in the early morning hours through city streets and highways:

Not at all |————————————————| Extremely

4. Under each clock time, use the scale below to write down how stressed you usually feel when driving at that time through city streets and highways.

1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

Instructions: The questions below ask about your fatigue *level when driving.*

5. While operating a commercial motor vehicle, have you ever, even for a moment, fallen asleep behind the wheel? ☐ Yes ☐ No

6. While operating a commercial motor vehicle, have you ever been driving and found yourself somewhere, not remembering how you got there? ☐ Yes ☐ No

7. How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired? This refers to your usual way of life in recent times. Even if you have not done some of these things recently, try to work out how they would have affected you. Use the following scale to choose the most appropriate number for each situation:

0 = no chance of dozing
1 = slight chance of dozing
2 = moderate chance of dozing
3 = high chance of dozing

SITUATION CHANCE OF DOZING :

Sitting and reading ____________
Watching TV ____________
Sitting inactive in a public place (e.g., a theater or a meeting) ____________
As a passenger in a car for an hour without a break ____________
Lying down to rest in the afternoon when circumstances permit ____________
Sitting and talking to someone ____________
Sitting quietly after a lunch without alcohol ____________
In a car, while stopped for a few minutes in traffic ____________

8. Draw a line through the bar to show how fatigued you usually feel when driving in the early morning hours through rural interstates:

Not at all |————————————————| Extremely

9. Under each clock time, use the scale below to write down how fatigued you usually feel when driving at that time through rural interstates.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

10. Draw a line through the bar to show how fatigued you usually feel when driving in the early morning hours through city streets and highways:

Not at all |————————————————| Extremely

11. Under each clock time, use the scale below to write down how fatigued you usually feel when driving at that time through city streets and highways.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

12. How well are you currently able to follow your company's fatigue management program?

☐ Not at all ☐ A little ☐ Quite a bit ☐ Very much ☐ Non-applicable

13. How well do you think you will be able to follow your company's fatigue management program, using DDWS feedback?

☐ Not at all ☐ A little ☐ Quite a bit ☐ Very much ☐ Non-applicable

Instructions: The questions below ask about your alertness level when driving.

14. Draw a line through the bar to show how alert you usually feel when driving in the early morning hours through rural interstate areas:

Not at all |————————————————| Extremely

15. Under each clock time, use the scale below to write down how alert you usually feel when driving at that time through rural interstate areas.

1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

16. Draw a line through the bar to show how alert you usually feel when driving in the early morning hours through city streets and highways:

Not at all |——————————————————| Extremely

17. Under each clock time, use the scale below to write down how alert you usually feel when driving at that time through city streets and highways.

1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

18. During your non-driving time, how do you plan on using the DDWS feedback you received while driving?
    - ☐ Planning a sleep/wake schedule
    - ☐ Planning meal times
    - ☐ I do not plan on using the DDWS feedback I received during non-driving time
    - ☐ Other: ____________________

19. What time(s) do you usually take a meal break while on duty? *(Please write down hour(s) and AM/PM)*
    ____

20. Have you ever had an accident or incident while working that you feel was related to your sleepiness or fatigue level? ☐ yes ☐ no If yes, how many? _______

21. Have you been involved with other fatigue management programs within the past year?
    ☐ yes ☐ no If yes, please briefly explain what they contained: ____________________
    ____

22. Please indicate how familiar you are with fatigue management:

    ☐ Very familiar/expert ☐ Familiar ☐ Somewhat familiar ☐ Novice ☐ Not at all familiar

23. How many total hours do you work in a typical week? _________. How many of these hours do you spend driving? _________

24. *Please write down your <u>typical</u> work start and end times for each day in the grid below.*
    - Write *"off"* for those days you are <u>not</u> usually scheduled to work.
    - If the days you work change from week to week, fill in the *typical* time you might work for each day and check the '<u>variable work schedule</u>' box below the grid.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Start time							
End time							

☐ *variable work schedule*

25. What percent of your <u>work time</u> do you drive on:
    city streets and highways: __________%
    rural interstates: __________%
    Other: ____________________________%

26. How often during a <u>typical</u> work shift do you load and or unload your own cargo?

    ☐ Never
    ☐ Occasional shifts
    ☐ Sometimes (about half of the time)
    ☐ Most shifts
    ☐ Every shift

27. For <u>typical</u> loads, how long does it take you to load/unload the truck during deliveries? ____ minutes

    Do you find loading/unloading the truck to be more <u>alerting</u> or <u>fatiguing</u>?
    ☐ Alerting ☐ Fatiguing ☐ Neither Why: ____________________________________

    ☐ This question does not apply; I do not usually load/unload my truck.

28. What other work-related duties do you usually do, if any, <u>besides</u> driving and loading/unloading?
    ____

29. What type of <u>tractor</u> do you usually drive? ____

THANK YOU for completing this survey!

Do you have any <u>vacations</u>, <u>holiday</u> or other <u>planned time off</u> scheduled over the 17 weeks of this study?

If <u>yes</u>, please give the dates that you will be out of work:

____

Pre-Study Survey

Thank you for taking the time to complete this survey. It should take you about 15 minutes to fill out. The information that you give will not be shared with any of your managers or other operators, and your responses will be kept strictly confidential. This survey asks you questions about you, your health and well being, fatigue management, and your current work schedule and environment. You are not required to give answers to any survey items that make you uncomfortable. However, the more data you give, the better we will be able to understand the potential benefits of the device for drivers like yourself.

Operator Information:
The questions below about yourself and your background will help us to learn about what types of people are participating in this study.

1. Some people feel older or younger than their actual age. How old do you feel? ______ years

2. Please write down about how many years of each you have finished. Then check the correct box under whether or not you graduated.

	Number of years	*Did You Graduate?*	
High School	_________	☐ Yes	☐ No
Technical School	_________	☐ Yes	☐ No
Subject: ___________________			
College	_________	☐ Yes	☐ No

3. What is your present marital status? *(Check one)*

☐ Single ☐ Divorced
☐ Married ☐ Separated
☐ Living with Partner ☐ Widowed

4. How many people depend on your income, including yourself? _________

5. How many other people in your household depend on you for a large part of their personal care? ______________

Health and Well Being:
The questions below will help us to better understand the status of your health. Please carefully follow any additional instructions.

1. Draw a line through the bar to show how healthy you usually feel.

Very healthy |————————————————| Not healthy at all

2. *Instructions:* The questions in this scale ask you about your feelings and thoughts <u>during the last month</u>. In each case, please check how often you felt or thought a certain way.

   a. In the last month, how often have you felt that you were unable to control the important things in your life?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   b. In the last month, how often have you felt confident about your ability to handle your personal problems?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   c. In the last month, how often have you felt that things were going your way?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   d. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

3. The questions below concern some of your body functions. Please try to answer each question by checking the appropriate box that shows how often you experienced each of the items *within the past year.* Check <u>one space</u> for <u>each item</u> below.

Items:	*Never*	*Less than once a month*	*Once or twice a month*	*Once a week*	*2 or 3 times a week*	*About every day*
Back Pain						
Acid indigestion, heartburn, or acid stomach						
Gastrointestinal problems						
Constipation						
Tight feeling in stomach						
Blurred vision						
Trouble falling asleep						
Feeling tired upon awaking						
Trouble staying asleep						
Other:						
Other:						

4. What health problems or chronic conditions have you been diagnosed with? Are they currently being treated? For example, diabetes, sleep apnea, heart condition, cancer, restless legs, etc.

Health Issue	Being Treated? (check one)	
____________________	☐ Yes	☐ No
____________________	☐ Yes	☐ No
____________________	☐ Yes	☐ No

5. How often do you usually have a drink of …? *(Check only one space for each drink item listed below). Then, place a ✓ in the shaded column if you use that beverage to stay awake while driving.*

Items:	More than 5 times a day	3 to 5 times a day	1 to 2 times a day	Less than once a day	Never or almost never	*Used to stay alert while driving?*
Beer						
Wine						
Other Liquor						
Coffee						
Tea						
Cola/ Soft Drink						

6. Do you smoke? ☐ Yes ☐ No

   If yes, how much do you smoke per day, on average? *(Enter a number)*
   How many packs of cigarettes per day ________
   How many cigars per day ________
   How many pipes of tobacco per day ________

   Do you use smoking as a way to remain alert? ☐ Yes ☐ No

7. Please list the medications you take regularly and for what reason you take them. Also, please check the box for those that make you feel sleepy. For example, medications including those for high blood pressure, cholesterol, allergies, etc.

Medication:	*What reason?*	*Makes me sleepy:*
____________	____________________	☐
____________	____________________	☐
____________	____________________	☐
____________	____________________	☐
____________	____________________	☐

8. *Instructions:* Read each question carefully. Check the box (✓) underneath the most appropriate time frame.

a. If you were entirely free to plan your evening and had no commitments the next day, at what time would you choose to go to bed?

8 – 9 PM	9 – 10:15 PM	10:15 PM – 12:30 AM	12:30 – 1:45 AM	1:45 – 3 AM

b. You have to do 2 hours physically hard work. If you were entirely free to plan your day, in which of the following periods would you choose to do the work?

8 – 10 AM	11 AM – 1 PM	3 – 5 PM	7 – 9 PM

c. For some reason you have gone to bed several hours later than normal, but there is no need to get up at a particular time the next morning. Which of the following is most likely to occur?

- ☐ Will wake up at the usual time and not fall asleep again
- ☐ Will wake up at the usual time and doze thereafter
- ☐ Will wake up at the usual time but will fall asleep again
- ☐ Will not wake up until later than usual

d. You have a 2-hour test to take, which you know will be mentally exhausting. If you were entirely free to choose, in which of the following periods would you choose to sit the test?

8 – 10 AM	11 AM – 1 PM	3 – 5 PM	7 – 9 PM

e. If you had no commitments the next day and were entirely free to plan your own day, what time would you get up?

5 – 6:30 AM	6:30 – 7:45 AM	7:45 – 9:45 AM	9:45 – 11:00 AM	11:00 AM – 12:00 PM	12 PM or later

f. A friend has asked you to join him twice a week for a work-out in the gym. The best time for him is between 10pm - 11pm. Bearing nothing else in mind other than how you normally feel in the evening, how do you think you would perform?

- ☐ Very well
- ☐ Reasonably well
- ☐ Poorly
- ☐ Very poorly

g. One hears about 'morning' and 'evening' types of people. Which of these types do you consider yourself to be?

- ☐ Definitely morning type
- ☐ More a morning than an evening type
- ☐ More an evening than a morning type
- ☐ Definitely an evening type

9. *Instructions:* Below are eighteen statements that people sometimes make about themselves. Please indicate whether or not you believe each statement applies to you by marking whether you:

1 = Strongly Disagree   2 = Disagree   3 = Neutral   4 = Agree   5 = Strongly Agree

_____1. Are you inclined to keep in the background on social occasions?
_____2. Do you like to mix socially with people?
_____3. Do you sometimes feel happy, sometimes depressed, without any apparent reason?

_____4. Are you inclined to limit your acquaintances to a select few?
_____5. Do you like to have many social engagements?
_____6. Do you have frequent ups and downs in mood, either with or without apparent cause?

_____7. Would you rate yourself as a happy-go-lucky individual?
_____8. Can you usually let yourself go and have a good time at a party?
_____9. Are you inclined to be moody?
_____10. Would you be very unhappy if you were prevented from making numerous social contacts?
_____11. Do you usually take the initiative in making new friends?
_____12. Does your mind often wander while you are trying to concentrate?
_____13. Do you like to play pranks upon others?
_____14. Are you usually a "good mixer?"
_____15. Are you sometimes bubbling over with energy and sometimes very sluggish?
_____16. Do you often "have the time of your life" at social affairs?
_____17. Are you frequently "lost in thought" even when you should be taking part in a conversation?
_____18. Do you derive more satisfaction from social activities than from anything else?

10. Have any major stressful events taken place in your life within the last year? ☐ Yes ☐ No
If yes, what was the event(s)? ____

Current Work Schedule and Environment:
The questions below will help us to better understand the impact of your work schedule and work environment.

Instructions: The questions below ask about your stress level when driving.

1. Draw a line through the bar to show how stressed you usually feel when driving in the early morning hours through rural interstates:

Not at all |————————————————| Extremely

2. Under each clock time, use the scale below to write down how stressed you usually feel when driving at that time through rural interstates.

1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

3. Draw a line through the bar to indicate how stressed you usually feel when driving in the early morning hours through city streets and highways:

Not at all |———————————————| Extremely

4. Under each clock time, use the scale below to write down how stressed you usually feel when driving at that time through city streets and highways.

1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

Instructions: The questions below ask about your fatigue *level when driving.*

5. While operating a commercial motor vehicle, have you ever, even for a moment, fallen asleep behind the wheel? ☐ Yes ☐ No

6. While operating a commercial motor vehicle, have you ever been driving and found yourself somewhere, not remembering how you got there? ☐ Yes ☐ No

7. How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired? This refers to your usual way of life in recent times. Even if you have not done some of these things recently, try to work out how they would have affected you. Use the following scale to choose the most appropriate number for each situation:

   0 = no chance of dozing
   1 = slight chance of dozing
   2 = moderate chance of dozing
   3 = high chance of dozing

SITUATION CHANCE OF DOZING :

Sitting and reading ____________
Watching TV ____________
Sitting inactive in a public place (e.g., a theater or a meeting) ____________
As a passenger in a car for an hour without a break ____________
Lying down to rest in the afternoon when circumstances permit ____________
Sitting and talking to someone ____________
Sitting quietly after a lunch without alcohol ____________
In a car, while stopped for a few minutes in traffic ____________

8. Draw a line through the bar to show how likely it is for you to be involved in a fatigue-related accident or incident.

Not at all |———————————————| Extremely

9. Draw a line through the bar to show how fatigued you usually feel when driving in the early morning hours through rural interstates:

Not at all |————————————————| Extremely

10. Under each clock time, use the scale below to write down how fatigued you usually feel when driving at that time through rural interstates.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

11. Draw a line through the bar to show how fatigued you usually feel when driving in the early morning hours through city streets and highways:

Not at all |————————————————| Extremely

12. Under each clock time, use the scale below to write down how fatigued you usually feel when driving at that time through city streets and highways.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

13. How well are you currently able to follow your company's fatigue management program?

☐ Not at all ☐ A little ☐ Quite a bit ☐ Very much ☐ Non-applicable

Instructions: The questions below ask about your alertness *level when driving.*

14. Draw a line through the bar to show how alert you usually feel when driving in the early morning hours through rural interstate areas:

Not at all |————————————————| Extremely

15. Under each clock time, use the scale below to write down how alert you usually feel when driving at that time through rural interstate areas.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

16. Draw a line through the bar to show how alert you usually feel when driving in the early morning hours through city streets and highways:

Not at all |————————————————————| Extremely

17. Under each clock time, use the scale below to write down how alert you usually feel when driving at that time through city streets and highways.

1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

What time(s) do you usually take a meal break while on duty? *(Please write down hour(s) and AM/PM)*
____

18. Have you ever had an accident or incident while working that you feel was related to your sleepiness or fatigue level? ☐ yes ☐ no If yes, how many? _______

19. Have you been involved with other fatigue management programs within the past year?
☐ yes ☐ no If yes, please briefly explain what they contained:____________________
____

20. Please indicate how familiar you are with fatigue management:

☐ Very familiar/expert ☐ Familiar ☐ Somewhat familiar ☐ Novice ☐ Not at all familiar

21. How many total hours do you work in a typical week? _________. How many of these hours do you spend driving? _________

22. *Please write down your typical work start and end times for each day in the grid below.*
   - Write *"off"* for those days you are not usually scheduled to work.
   - If the days you work change from week to week, fill in the *typical* time you might work for each day and check the 'variable work schedule' box below the grid.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Start time							
End time							

☐ *variable work schedule*

23. What percent of your work time do you drive on:
   city streets and highways: __________%
   rural interstates: __________%
   Other: ______________________________%

24. How often during a typical work shift do you load and or unload your own cargo?

    ☐ Never
    ☐ Occasional shifts
    ☐ Sometimes (about half of the time)
    ☐ Most shifts
    ☐ Every shift

25. For typical loads, how long does it take you to load/unload the truck during deliveries? ____ minutes

    Do you find loading/unloading the truck to be more alerting or fatiguing?
    ☐ Alerting ☐ Fatiguing ☐ Neither Why: ______________________________

    ☐ This question does not apply; I do not usually load/unload my truck.

26. What other work-related duties do you usually do, if any, besides driving and loading/unloading?
    ____

27. What type of tractor do you usually drive? ______________________________

28. Currently, how often do you look at the instrument panel in the cab?

    ☐ Once or more a minute
    ☐ Once every 5 to 10 minutes
    ☐ Once every 15 to 30 minutes
    ☐ Once every 30 to 60 minutes

THANK YOU for completing this survey!

Do you have any vacations, holiday or other planned time off scheduled over the 17 weeks of this study?

If yes, please give the dates that you will be out of work:
____

Post-Study Survey

Thank you for taking the time to complete this survey. It should take you about 30 minutes to fill out. The information that you give will not be shared with any of your managers or other operators, and your responses will be kept strictly confidential. This survey asks you questions about your health and well being, use of the drowsy driver warning system (DDWS), and your current work schedule and environment. You are not required to give answers to any survey items that make you uncomfortable. However, the more data you give, the better we will be able to understand the potential benefits of the device for drivers like yourself.

Health and Well Being:

The questions below will allow us to better understand your health. Please carefully follow any additional instructions.

1. Draw a line through the bar to show how healthy you usually feel.

   Very healthy |————————————————————| Not healthy at all

2. The questions below concern some of your body functions. Please try to answer each question by checking the correct box that shows how often you experienced each of the items *during the course of the study.* Check one space for each item below.

Items:	*Never*	*Less than once a month*	*Once or twice a month*	*Once a week*	*2 or 3 times a week*	*About every day*
Back Pain						
Acid indigestion, heartburn, or acid stomach						
Gastrointestinal problems						
Constipation						
Tight feeling in stomach						
Blurred vision						
Trouble falling asleep						
Feeling tired upon awaking						
Trouble staying asleep						
Other:						
Other:						

3. Did any major stressful events take place in your life during the study? ☐ No ☐ Yes

   If "yes," what was the event(s)? ______________________________

4. *Instructions:* The questions in this scale ask you about your feelings and thoughts during the last month. In each case, please check how often you felt or thought a certain way.

   a. In the last month, how often have you felt that you were unable to control the important things in your life?
      ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   b. In the last month, how often have you felt confident about your ability to handle your personal problems?
      ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   c. In the last month, how often have you felt that things were going your way?
      ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   d. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?
      ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

DDWS Device:
The questions below will help us to better understand what you think about the DDWS device now that you have used it.

1. Please describe…
   a. how you think that the DDWS works (functions): ______________________________________

      ____

   b. how the DDWS may have helped you handle your fatigue while driving: __________________

      ____

2. Do you know of any situations where the DDWS device did not work right during the study?
   ☐ No ☐ Yes

   If "yes," how did you know that the device was not working right? *(Check all that apply)*

   ☐ Device gave no warnings
   ☐ Device gave constant warnings
   ☐ No visual display present
   ☐ Other: ____

3. How did you use DDWS warnings immediately while you were driving? *(Check all that apply)*

   ☐ I stopped driving and pulled over to sleep
   ☐ I drank a beverage that had caffeine in it
   ☐ Other: ____

4. Along with a fatigue management program, how useful do you think the DDWS was in helping you manage your fatigue while at work?

   ☐ Not at all useful ☐ A little useful ☐ Quite useful ☐ Extremely useful

5. Not considering what you know about fatigue management, how useful do you think the DDWS was in helping you manage your fatigue at work?

   ☐ Not at all useful ☐ A little useful ☐ Quite useful ☐ Extremely useful

6. The DDWS was designed for use by many different types of drivers. It may not always have correctly measured the alertness level of each person. How did you know whether this device was giving you correct information about your level of alertness? *(Check all that apply)*

   ☐ It only sounded warnings when I was fatigued
   ☐ It never sounded warnings when I was alert
   ☐ Other: ______________________________

7. How much do you believe the DDWS improved the safety of your driving?

   ☐ Not at all ☐ A little ☐ Quite a bit ☐ Extremely

   Why?______________________________

8. Before getting a warning, how often did you usually look at the DDWS device?

   ☐ Once or more a minute
   ☐ Once every 5 to 10 minutes
   ☐ Once every 15 to 30 minutes
   ☐ Once every 30 to 60 minutes
   ☐ Never

9. Once the DDWS gave a warning, how often did you usually look at the device?

   ☐ Once or more a minute
   ☐ Once every 5 to 10 minutes
   ☐ Once every 15 to 30 minutes
   ☐ Once every 30 to 60 minutes
   ☐ Only when an additional alarm was sounded

10. Draw a line through the bar to show how much effort you gave in checking the DDWS.

    None |————————————————| Extreme

11. Draw a line through the bar to show how much effort you gave in reacting to the DDWS warnings.

    None |————————————————| Extreme

12. How often did you feel that the DDWS correctly measured your level of alertness by giving you appropriate warnings?

    ☐ Almost never ☐ Occasionally ☐ Frequently ☐ Almost always

13. Draw a line through the bar to show how annoyed you were with incorrect DDWS warnings.

Not at all |————————————————| Extremely

14. How did you decide whether or not the DDWS was giving you correct warnings? *(Check all that apply)*

- ☐ I almost always got warnings when I was tired
- ☐ I knew that I was alert while driving and did not expect to get any warnings
- ☐ Other:______________________________

15. What did you do when the device gave you incorrect warnings? *(Check all that apply)*

- ☐ Ignored the device ☐ Reset the device ☐ Covered the device
- ☐ Other:______________________________

16. How often did you feel that the device gave you an incorrect warning about your level of fatigue?

☐ Almost never ☐ Occasionally ☐ Frequently ☐ Almost always

17. Under each clock time, use the scale below to write down how difficult it was to quickly understand and react to DDWS warnings.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3AM	3 AM – 6AM

18. Draw a line through the bar to show how easy it was to read the DDWS display while driving.

Not at all |————————————————| Extremely

19. Draw a line through the bar to show how easy it was to hear the DDWS warnings.

Not at all |————————————————| Extremely

20. List the steps you usually took when the DDWS gave you a warning:______________________________

____

21. Under each clock time, use the scale below to write down how often, in general, the DDWS gave you an incorrect warning.

1 = Almost never 2 = Occasionally 3 = Frequently 4 = Almost always

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

22. Did the DDWS give you all of the information that you needed to make good decisions about managing your fatigue while driving? ☐ Yes ☐ No
    What more information would you have liked? ____________________________

23. Was the location of the DDWS acceptable? *(Check one)* ☐Yes ☐No
    If "no," what was the problem?____________________________

24. When were you most likely to use the DDWS warning information? *(Check all that apply)*
    - ☐ When I felt very tired
    - ☐ During poor driving conditions
    - ☐ When there was very little traffic
    - ☐ Other:____________________________

25. Draw a line through the bar to show how much you felt that you relied on the DDWS device to warn you about your level of fatigue.

    Not at all |————————————————| Extremely

26. How often were you unable to take actions to improve your alertness once the DDWS gave a warning?

    ☐ Almost never ☐ Occasionally ☐ Frequently ☐ Almost always

27. How often were you unwilling to take actions to improve your alertness once the DDWS gave a warning?

    ☐ Almost never ☐ Occasionally ☐ Frequently ☐ Almost always

28. In what situations did you most often not do something to improve your alertness after the DDWS gave a warning?____________________________

29. What was your biggest problem when using the DDWS? *(Check one)*
    - ☐ Reading the display
    - ☐ Turning off the alarm
    - ☐ Changing device sensitivity
    - ☐ Hearing the warning
    - ☐ No problem
    - ☐ Other:____________________________

30. During your non-driving time, how did you use the DDWS feedback that you received while you were driving? *(Check all that apply)*
    - ☐ Planned my sleep/wake schedule
    - ☐ Planned meal times
    - ☐ I did not think about the DDWS feedback I received during non-driving time
    - ☐ Other: ____________________________

31. Did you use DDWS information to help with your other fatigue management activities (for example, changing the time you sleep, etc.)? ☐ No ☐ Yes

    If "yes," how did you use the DDWS information? ______________________________

32. Draw a line through the bar to show how comfortable you were using DDWS information to help with your other fatigue management activities (for example, changing the time you sleep, etc.).

    Not at all |————————————————| Extremely

33. How long did it take you to understand how you could use DDWS information in your other fatigue management activities?

    ☐ One day ☐ Less than 1 week ☐ Less than 1 month ☐ Still learning

34. Could you easily understand the written materials that you received about the DDWS? ☐ No ☐ Yes

35. Would it have helped you if DDWS training materials were written in a language *other* than English?

    ☐ No ☐ Yes If "yes," which language? ______________________

36. Draw a line through the bar to show how useful the DDWS was, in addition to the other steps you take to manage your fatigue.

    Not at all |————————————————| Extremely

37. Draw a line through the bar to show how complete the DDWS training that you got was.

    Not at all |————————————————| Extremely

38. Draw a line through the bar to show how complete your company's fatigue management training was.

    Not at all |————————————————| Extremely

39. Do you believe that the DDWS information was kept confidential as discussed at the beginning of the study? Yes ☐ No ☐ If "no," why not? ______________________________

40. How many warnings did the DDWS need to give you before you tried to do something to improve your alertness? ☐ 1 ☐ 2 ☐ 3 ☐ 4 or more

41. Do you find the information given to you by the DDWS useful for managing your fatigue?

    ☐ Yes ☐ No Why or why not? ______________________________

42. Have you found uses for the DDWS other than as a tool to manage your fatigue? ☐ No ☐ Yes

    If "yes," what are the uses? ______________________________

43. How do you feel that the DDWS has affected your health? *(Check one)*

☐ Improved my health
☐ Impaired my health
☐ No effect on my health
Why?________________________________

44. Draw a line through the bar to show how likely it is for you to be involved in a fatigue-related accident or incident when not using the DDWS.

Not at all |————————————————| Extremely

45. Draw a line through the bar to show how likely it is for you to be involved in a fatigue-related accident or incident when using the DDWS.

Not at all |————————————————| Extremely

46. Do you feel that, for whatever reason, using the DDWS made you feel more fatigued when you drove your usual hours? ☐No ☐ Yes Why? ____

47. Would you be willing to continue using the DDWS device after the study is over if the information it gives you is not recorded or available to your company and others? ☐ No ☐ Yes

Why or why not? ____

48. I would be willing to: *(Check all that apply)*

☐ Purchase a DDWS myself, if I could afford it
☐ Share the cost of the purchase of a DDWS for my truck cab with my employer
☐ Ask that my employer buy a DDWS device for my truck
☐ Ask that my employer buy DDWS devices for the fleet

49. Have you recommended the use of this device to your management? ☐ No ☐ Yes

50. Have you recommended the use of the DDWS to other drivers in your company? ☐ No ☐ Yes

51. Have you recommended the use of the DDWS device to professional drivers outside your company? ☐ No ☐ Yes

52. How often did you ignore DDWS warnings?

☐ Almost never ☐ Occasionally ☐ Frequently ☐ Almost always

53. How often did you trick the DDWS device so that it would no longer give you warnings?

☐ Almost never ☐ Occasionally ☐ Frequently ☐ Almost always

54. Draw a line through the bar to show how well you feel that you were able to include the DDWS into your fatigue management habits.

Not at all |————————————| Extremely

55. In the boxes below, write down any other fatigue-fighting actions or activities you do when driving. Please also fill in the spaces for how frequently you use each with this scale:

1 = Nightly 2 = Weekly 3 = Monthly 4 = Only rarely

Fatigue-fighting Action or Activity	*Frequency*

56. Please describe what you do personally to try to control your on-duty fatigue:____________________

____

57. Do you feel that the DDWS helped you to avoid an accident or a close call when it was turned on during the study? ☐ No ☐ Yes If "yes," how often? ______________________________

58. Did you take any fatigue or safety management classes, other than what was given for the DDWS, during the time when you were in this study?

☐ No ☐ Yes If "yes," what class? ______________________________

59. How annoying were the DDWS auditory alerts?

- ☐ Unacceptably annoying
- ☐ Somewhat annoying
- ☐ Tolerable
- ☐ Only slightly annoying
- ☐ Not at all annoying

60. Draw a line through the bar to show how much you agree that the DDWS increased your driving safety.

Strongly disagree |————————————| Strongly agree

61. Draw a line through the bar to show how satisfied you usually were with the DDWS system.

Very unsatisfied |————————————| Very satisfied

62. On average, how often per duty period do you feel that the DDWS device gave you an incorrect warning about your level of fatigue? _______ times per duty period.

63. Draw a line through the bar to show how comfortable you felt using the DDWS device.

Very uncomfortable |————————————| Very comfortable

64. Draw a line through the bar to show how easy or difficult you found it to drive using the DDWS device.

    Very difficult ⊢———————————————————⊣ Very easy

65. Draw a line through the bar to show how easily you were able to recognize alerts from the DDWS device.

    Not easily ⊢———————————————————⊣ Very easily

66. Draw a line through the bar to show how much you agree that the DDWS device got your attention quickly when it gave an alert.

    Strongly disagree ⊢———————————————————⊣ Strongly agree

67. Were there ever situations when the DDWS device worked in a way that you did not understand?
    ☐ No ☐ Yes If "yes", please explain: ____

    ____

68. On average, how often per duty period do you feel that the DDWS device did not give you a warning about your level of fatigue when you felt that one was needed? _______ times per duty period

69. Draw a line through the bar to show how easy it was to use the DDWS device while driving.

    Not at all ⊢———————————————————⊣ Extremely

70. Overall, I found myself relying too much on the DDWS device.

    Not at all ⊢———————————————————⊣ Extremely

71. Did you stop relying on the DDWS because of incorrect warnings? ☐ No ☐ Yes

72. While operating a commercial motor vehicle during the study, did you ever, even for a moment, fall asleep behind the wheel? ☐ No ☐ Yes

73. While operating a commercial motor vehicle during the study, did you ever find yourself somewhere and not remembering how you got there? ☐ No ☐ Yes

74. Please indicate your overall acceptance rating of the DDWS system.

    For each choice you will find 5 possible answers. When a term is completely appropriate, please put a check (√) in the square next to that term. When a term is appropriate to a certain extent, please put a check to the left or right of the middle at the side of the term. When you have no specific opinion, please put a check in the middle.

Test Group Post-study Survey

The DDWS system was:

useful						useless
pleasant						unpleasant
bad						good
nice						annoying
effective						superfluous (ineffective)
irritating						likeable
assisting						worthless
undesirable						desirable
raising alertness						sleep-inducing

Current Work Schedule and Environment:
The questions below will help us to better understand the impact of your work schedule and work environment.

Instructions: The questions below ask about your stress *level while driving using the DDWS.*

1. Draw a line through the bar to show how stressed you usually felt when driving in the early morning hours through rural interstates using the DDWS:

Not at all |———————————————| Extremely

2. Under each clock time, use the scale below to write down how stressed you usually felt when driving at that time through rural interstates using the DDWS.

   1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

3. Draw a line through the bar to show how stressed you usually felt when driving in the early morning hours through city streets and highways using the DDWS:

   Not at all |————————————| Extremely

4. Under each clock time, use the scale below to write down how stressed you usually felt when driving at that time through city streets and highways using the DDWS.

   1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

Instructions: The questions below ask about your fatigue level while driving using the DDWS.

5. Draw a line through the bar to show how fatigued you usually felt when driving in the early morning hours through rural interstates using the DDWS:

   Not at all |————————————| Extremely

6. Under each clock time, use the scale below to write down how fatigued you usually felt when driving at that time through rural interstates using the DDWS.

   1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

7. Draw a line through the bar to show how fatigued you usually felt when driving in the early morning hours through city streets and highways using the DDWS:

   Not at all |————————————————| Extremely

8. Under each clock time, use the scale below to write down how fatigued you usually felt when driving at that time through city streets and highways using the DDWS.

   1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

9. How well do you now think that you will be able to follow your company's fatigue management program, using DDWS feedback?

   ☐ Not at all ☐ A little ☐ Quite a bit ☐ Very much

10. How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired? This refers to your usual way of life in recent times. Even if you have not done some of these things recently, try to work out how they would have affected you. Use the following scale to choose the most appropriate number for each situation:

    0 = no chance of dozing
    1 = slight chance of dozing
    2 = moderate chance of dozing
    3 = high chance of dozing

SITUATION CHANCE OF DOZING :

Sitting and reading ____________
Watching TV ____________
Sitting inactive in a public place (e.g., a theater or a meeting) ____________
As a passenger in a car for an hour without a break ____________
Lying down to rest in the afternoon when circumstances permit ____________
Sitting and talking to someone ____________
Sitting quietly after a lunch without alcohol ____________
In a car, while stopped for a few minutes in traffic ____________

Instructions: The questions below ask about your alertness *level while driving using the DDWS.*

11. Draw a line through the bar to show how alert you usually felt when driving in the early morning hours through rural interstate areas using the DDWS:

Not at all |————————————————| Extremely

12. Under each clock time, use the scale below to write down how alert you usually felt when driving at that time through rural interstate areas using the DDWS.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

13. Draw a line through the bar to show how alert you usually felt when driving in the early morning hours through city streets and highways using the DDWS:

Not at all |————————————————| Extremely

14. Under each clock time, use the scale below to write down how alert you usually felt when driving at that time through city streets and highways using the DDWS.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

15. Did you drive any new or unfamiliar runs over the time of the study?
☐ No ☐ Yes If "yes," please list the runs below.

Run
Example: Pittsburgh to Boston via Routes 80, 84, 90

THANK YOU for completing this survey!

Post-Study Survey

Thank you for taking the time to complete this survey! It should take you about 15 minutes to fill out. The information that you give will not be shared with any of your managers or other operators, and your responses will be kept strictly confidential. This survey asks you questions about your health and well being, fatigue management, and your current work schedule and environment. You are not required to give answers to any survey items that make you uncomfortable. However, the more data you give, the better we will be able to understand the potential benefits of the device for drivers like yourself.

Health and Well Being:
The questions below will allow us to better understand your health. Please carefully follow any additional instructions.

1. Draw a line through the bar to show how healthy you usually feel.

   Very healthy |————————————————| Not healthy at all

2. The questions below concern some of your body functions. Please try to answer each question by checking the correct box that shows how often you experienced each of the items *during the course of the study*. Check one space for each item below.

Items:	*Never*	*Less than once a month*	*Once or twice a month*	*Once a week*	*2 or 3 times a week*	*About every day*
Back Pain						
Acid indigestion, heartburn, or acid stomach						
Gastrointestinal problems						
Constipation						
Tight feeling in stomach						
Blurred vision						
Trouble falling asleep						
Feeling tired upon awaking						
Trouble staying asleep						
Other:						
Other:						

3. Did any major stressful events take place in your life during the study? ☐ No ☐ Yes
   If "yes," what was the event(s)? ____

4. *Instructions:* The questions in this scale ask you about your feelings and thoughts during the last month. In each case, please check how often you felt or thought a certain way.

   a. In the last month, how often have you felt that you were unable to control the important things in your life?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   b. In the last month, how often have you felt confident about your ability to handle your personal problems?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   c. In the last month, how often have you felt that things were going your way?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

   d. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?
   ☐ Never ☐ Almost never ☐ Sometimes ☐ Fairly often ☐ Very often

Fatigue Management:
The questions below will allow us to better understand your views on fatigue management.

1. How useful do you think your company's fatigue management program was in helping you manage your fatigue at work?

   ☐ Not at all useful ☐ A little useful ☐ Quite useful ☐ Extremely useful

2. How well do you now think that you will be able to follow your company's fatigue management program?

   ☐ Not at all ☐ A little ☐ Quite a bit ☐ Very much

3. Draw a line through the bar to show how good you felt your company's fatigue management training program was.

   Not at all |————————————————| Extremely

4. Draw a line through the bar to show how likely it is for you to be involved in a fatigue-related accident or incident after you finished your company's fatigue management training.

   Not at all |————————————————| Extremely

5. In the boxes below, write down any fatigue-fighting actions or activities you do when driving. Please also fill in the spaces for how frequently you use each with this scale:

1 = Nightly 2 = Weekly 3 = Monthly 4 = Only rarely

Fatigue-fighting Action or Activity	*Frequency*

6. Please describe what you do personally to try to control your on-duty fatigue: ____________________

____

____

7. Did you take any fatigue or safety management classes, other than that given at the start of this study, during the time when you were a participant?
☐ No ☐ Yes If "yes," what class? ____

Current Work Schedule and Environment:
The questions below will help us to better understand the impact of your work schedule and work environment.

Instructions: The questions below ask about your stress *level while driving.*

1. Draw a line through the bar to show how stressed you usually felt when driving in the early morning hours through rural interstates:

Not at all |————————————————————| Extremely

2. Under each clock time, use the scale below to write down how stressed you usually felt when driving at that time through rural interstates.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

3. Draw a line through the bar to show how stressed you usually felt when driving in the early morning hours through city streets and highways:

Not at all |————————————————————| Extremely

4. Under each clock time, use the scale below to write down how stressed you usually felt when driving at that time through city streets and highways.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

Instructions: The questions below ask about your fatigue *level while driving.*

5. Draw a line through the bar to show how fatigued you usually felt when driving in the early morning hours through rural interstates:

Not at all |————————————————| Extremely

6. Under each clock time, use the scale below to write down how fatigued you usually felt when driving at that time through rural interstates.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

7. Draw a line through the bar to show how fatigued you usually felt when driving in the early morning hours through city streets and highways:

Not at all |————————————————| Extremely

8. Under each clock time, use the scale below to write down how fatigued you usually felt when driving at that time through city streets and highways.

1 = Not at all 2 = A little 3 = Quite a bit 4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

9. While operating a commercial motor vehicle during the study, did you ever, even for a moment, fall asleep behind the wheel? ☐ No ☐ Yes

10. While operating a commercial motor vehicle during the study, did you ever find yourself somewhere and not remembering how you got there? ☐ No ☐ Yes

11. How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired? This refers to your usual way of life in recent times. Even if you have not done some of these things recently, try to work out how they would have affected you. Use the following scale to choose the most appropriate number for each situation:

    0 = no chance of dozing
    1 = slight chance of dozing
    2 = moderate chance of dozing
    3 = high chance of dozing

SITUATION CHANCE OF DOZING :

Sitting and reading ____________
Watching TV ____________
Sitting inactive in a public place (e.g., a theater or a meeting) ____________
As a passenger in a car for an hour without a break ____________
Lying down to rest in the afternoon when circumstances permit ____________
Sitting and talking to someone ____________
Sitting quietly after a lunch without alcohol ____________
In a car, while stopped for a few minutes in traffic ____________

Instructions: The questions below ask about your alertness level while driving.

12. Draw a line through the bar to show how alert you usually felt when driving in the early morning hours through rural interstate areas:

    Not at all |————————————————| Extremely

13. Under each clock time, use the scale below to write down how alert you usually felt when driving at that time through rural interstate areas.

    1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

14. Draw a line through the bar to show how alert you usually felt when driving in the early morning hours through city streets and highways:

    Not at all |————————————————| Extremely

15. Under each clock time, use the scale below to write down how alert you usually felt when driving at that time through city streets and highways.

    1 = Not at all   2 = A little   3 = Quite a bit   4 = Extremely

9 PM – Midnight	Midnight – 3 AM	3 AM – 6 AM

16. Did you drive any new or unfamiliar runs over the time of the study? ☐ No ☐ Yes

If "yes," please list the runs below.

Run
Example: Pittsburgh to Boston via Routes 80, 84, 90

THANK YOU for completing this survey!

Debriefing Survey

The entire study staff would like to thank you for your participation in this experiment. We realize that this was not an easy or a short study and very much appreciate your willingness to complete all of the requested tasks. The data collected from this experiment will be used to evaluate and further develop drowsy driver warning systems for the American trucking industry. You have made a major contribution to our understanding of the safety benefits of this device and how we may improve upon it.

Instructions: Please write your answers under each question. Use the back of the page if you need additional space.

1. What part(s) of the experiment did you object to, or caused you hardship?

2. What did you enjoy or find interesting about the study?

3. How do you think the *training* on using the DDWS can be improved to be more helpful?

4. Was the amount of money you received for participating in the study a fair amount for the work that you had to do? (Check one) ☐ No ☐ Yes

   If you answered "No", what amount would have been fair?

5. Do you feel that a *reward* (for example, receiving an amount of money for reducing the number of warnings you received over the duration of the study) would have changed the way you used the DDWS feedback? (Check one) ☐ No ☐ Yes

   If "Yes", how?

6. Did you find any part of the study (for example, the training, explanation, focus groups, etc.) to be too rushed or poorly timed? (Check one) ☐ No ☐ Yes

   If "Yes", what part of the study was rushed and how might it be fixed?

7. What problems, if any, did you notice in how the study was conducted?

8. How aware were you of the DDWS and being in a study when the device was *not turned on*?

9. Did you feel *more alert* throughout the night *during the study*, in general, than you would have otherwise? ☐ No ☐ Yes

   If "yes", why do you think that you felt this way?

10. Overall, how did you feel that you were treated during the experiment? How could we improve the experience of future participants?

11. General comments and observations:

Debriefing Survey

The entire study staff would like to thank you for your participation in this experiment. We realize that this was not an easy or a short study and very much appreciate your willingness to complete all of the requested tasks. The data collected from this experiment will be used to evaluate and further develop technology for the American trucking industry. Thank you again for your contribution and for your efforts.

Instructions: Please write your answers under each question. Use the back of the page if you need additional space.

1. What part(s) of the experiment did you object to, or caused you hardship?

2. What did you enjoy or find interesting about the study?

3. Was the amount of money you received for participating in the study a fair amount for the work that you had to do? (Check one) ☐ No ☐ Yes

If you answered "No", what amount would have been fair?

4. Did you find any part of the study (for example, the fatigue management training, explanation, focus groups, etc.) to be too rushed or poorly timed? (Check one)
   ☐ No ☐ Yes
   If "Yes", what part of the study was rushed and how might it be fixed?

5. What problems, if any, did you notice in how the study was conducted?

6. Did you feel *more alert* throughout the night *during the study*, in general, than you would have otherwise? ☐ No ☐ Yes
   If "yes", why do you think that you felt this way?

7. Overall, how did you feel that you were treated during the experiment? How could we improve the experience of future participants?

8. General comments and observations:

APPENDIX I. FLEET MANAGEMENT ACCEPTANCE AND DEPLOYMENT INTERVIEW MATERIALS

I. Survey Trucking Company Management:

A maximum of nine managers will be surveyed in order to avoid conflict with OMB regulations. Participants will be ensured of the confidentiality of all information that they provide and that they will not be individually identified in any publication of survey outcomes.

The concept device will be secured to a truck's dashboard, to the right of the driver, so as not to block the forward view, or interfere with the steering wheel or existing CBs. A stationary camera at the top of the device will use technology that directs minimal levels of harmless infrared light towards the driver's face and eyes. *To the extent that the driver's eyes are "in view" of the device camera*, it processes data on the fly to estimate "percent eye closure" over time, during periods when the truck is traveling above 35 MPH and the cab is dark.

If device-estimates of percent eye closure are greater than a predetermined level, it is an indication that the driver's eyes are potentially closed (or nearly closed) over a specific amount of time, and that he or she may therefore be experiencing a lapse in alertness.

- When a lapse in alertness is detected, first a single, audible "warning" tone will sound.
- After the single warning tone, *if* device estimates of percent eye closure continue to rise, the tone will sound repeatedly, and red warning lights will illuminate and flash until the driver responds by pressing a button on top of the unit.
- Once this button is pressed and the warning silenced, a bar-graph-like display shows the duration of the longest eye closure (during the period over which the device was estimating percent eye closure). In addition, one numeric display shows the total number of warnings received from the time the truck was turned "on", and another shows the rate per hour for visual + auditory warnings.

Management Interview Questions
Mapped to Fleet Management Acceptance Objectives in SOW

Is your company primarily a long- or short-haul operation (or other)?

(1.1) Would you support or oppose this type of technology for use by your company's drivers?
- Why or why not?

(1.2) Do you think that your company's drivers would support or oppose using this type of technology?
- Why or why not?

(1.3) What do you see as the advantages and disadvantages of a device like this?

(2.4) How much would you be willing to pay for a device like this for use by your fleet?

(4.2) Can you think of additional features you would want in a device like this?
- How much would you be willing to pay for these features? (Try to obtain $-by-feature breakdown.)

(1.3) Do you think that providing education and training to your company's drivers about the capabilities, advantages, and disadvantages of this type of technology would affect their opinions?
- How so? That is, would they be more likely to approve or disapprove of it?

(4.1) As I described the device, are there any major issues you can foresee regarding its performance?
- If so, what are they? Please describe.

(3.2) If your company installed an alertness monitor similar to that previously described in its trucks, do you expect that the warnings drivers received would be monitored – that is, would you collect percent eye closure data as available from the device?
- If yes, how would you use this information in your company?

(3.2) Do you think that the potential for management to collect and use this data would influence driver utilization and acceptance of the device?
- If yes, how so?

(3.3) If an alertness-monitoring device such as the one I described were installed in your fleet's vehicles, what sort of policy, if any, might your company implement regarding required driver behavior following a warning about a lapse in alertness? (For example, would you allow drivers to dictate their own response to warnings or would a response be prescribed by management?)

(2.1) If an alertness-monitoring device such as the one I described were shown to improve driver safety, what do you think would be the economic benefits for your company?

(2.2) What federal incentives would make you more likely to recommend the purchase of a device such as this for your fleet?

(2.3) What insurance incentives would make you more likely to recommend the purchase of a device like this for your fleet?

Mapped to Deployment Objective 3 in SOW

1. In your view, what would need to be done in order to effectively promote and successfully deploy this type of device for use by trucking companies?

2. Can you think of any scientific research that you feel should be conducted in the area of drowsy-driving as related to crash avoidance?
    - If yes, please explain (try to elicit what sort of research/information interviewee seeks).

- When do you think this research should take place with reference to device deployment? (e.g., prior to, in parallel?)

3. Has your company purchased or considered the use of any alertness monitoring technologies?
    - If yes, please describe.
        - What technology?
        - Was it purchased for use? If so, how many trucks/employees experienced it?
        - Please provide any additional information pertaining to advantages/ disadvantages/lessons learned as pertaining to the technology.

II. Conduct Market Research/Survey Technology Developers:

Using the internet first to conduct market research, follow up with a phone survey of vendors/manufacturers/developers of alertness monitoring systems, as available (maximum of nine) to determine the following:

Technology Developer Interview Questions
(Mapped to Deployment Objectives 1.1 and 1.2 in the 6/1/07 SOW)

1. Name of technology

2. Background
    - How does it work?
    - Any research findings?

(1.1) Is the technology/product available to professional drivers (or the general public)?
If available,
- What mode/who is using?
- How long has the technology been available for such use?
- Was there an introductory period prior to deployment?
    - If yes, how long?
- How many vehicles were equipped (what types)?
- When will manufacturers and suppliers consider deployment of the system?
- What was the introductory price for the technology?
    - Will the price stay the same for full deployment?
    - If no, what will the new price be?

If not available,
- Is an introductory period planned prior to deployment?
    - If yes, how long?
- How many vehicles will be equipped (what types)?
- When will manufacturers and suppliers consider deployment of the system?
- What will be the introductory price for the technology?
    - Will the price stay the same for full deployment?
    - If no, what will the new price be?

(1.2) Long-term deployment:

- Can you estimate a price range for this technology's deployments over the next 5-10 and 10-15 years?
- Can you estimate the rate of deployment for this technology over the next 5-10 and 10-15 years?
- Do you expect that this technology will exist as option for all vehicles, or only more expensive models (if an integrated technology)?
    - When do you anticipate this happening?

III. <u>Survey Trucking-related Organizations</u>

Survey ATA (http://www.truckline.com/index) and other relevant organizations as available and necessary to determine the following:

[Provide description of drowsy driver warning system as above]

Trucking-related Organization Interview Questions
(Mapped to Deployment Objective 2 in the 6/1/07 SOW)

(2.1) What is your view of this technology in terms of advantages and disadvantages?

- What do you anticipate to be organized labor's view regarding this technology?

(2.2) Do you anticipate any liability concerns? If so, please explain:

- For trucking companies?
- For drivers?

(2.3) In your opinion, do you anticipate issues regarding potential management policies?

- For example, policies that would prescribe a driver's response to warnings?
- Or regarding data collection and how data may be used by management?

(2.6) In your view, what types of incentives would be desired by organized labor in return for its members utilizing such a device?

(2.4) Do you anticipate any other issues with acceptance of this type of device?

- In terms of technology developments and/or device capabilities? (2.5)
- Legal developments? (2.6)

CONCLUSIONS

In an effort to decrease the crashes and fatalities related to drowsiness that involve commercial motor vehicle drivers, NHTSA has been actively seeking new methods and technologies to enhance driver alertness under the Drowsy Driver Technology Program since 1996. The current research effort, funded by NHTSA, presents a complete assessment of the DDWS concept. The prototype system used to evaluate this concept (i.e., DFM) provided feedback to drivers when their level of alertness (as evaluated by their PERCLOS level) exceeded an acceptable threshold. The assessment of the DDWS was performed using objective data (e.g., PERCLOS, sleep quantity and quality, safety-critical events) and subjective data (e.g., questionnaires, focus groups, interviews). The following is a compilation of the main findings for this research effort under each of the five main areas evaluated: (1) safety benefits, (2) performance and capabilities, (3) driver acceptance, (4) fleet management acceptance, and (5) deployment.

A data set composed of naturalistic parametric data and subjective data was originally collected by VTTI for NHTSA and FMCSA in a previous research effort, which allowed for a comprehensive examination of drowsiness in commercial motor vehicle drivers. At its time, this was the largest data set ever collected by the U.S. Department of Transportation. VTTI was tasked with analyzing the parametric data to determine if driver performance differed based on the assistance of a DDWS in conflict situations and evaluate the prototype DFM system's performance and capabilities (Part I of this report). VNTSC led the acceptance and deployment aspects of the assessment (Part II of this report).

SAFETY BENEFITS

Driver Drowsiness

There were inconclusive results as to the impact of the prototype DFM system on driver performance in conflict situations. Although precautions during data reduction were taken in an attempt to eliminate periods when the DFM would likely operate in an unreliable fashion, a high number of false alerts (not related to drowsiness) were still registered. A sampling of alerts revealed a discrepancy between the prototype DFM system's estimate of PERCLOS and the DDWS concept requirement. For example, certain driver actions (e.g., scanning mirrors) were included the DFM's PERCLOS calculation. These make it difficult to draw any directly generalizable conclusions regarding the findings for the DDWS concept, but some general trends were observed.

With those precautions in mind, the analysis revealed that drivers in the Test Group had lower PERCLOS values overall as compared to other experimental conditions. Most of the valid DFM alerts were registered on dry two-lane highways with no adverse weather conditions present. These results suggest that providing the driver with feedback as to his or her level of arousal would lead to an overall reduction of instances of drowsy driving. Lastly, the prototype DFM system did not appear to impact driver's normal sleep pattern or sleep hygiene, whether measured by Actigraph data in terms of sleep quantity or sleep quality. These objective findings mirror the subjective evaluation results obtained; drivers originally mentioned that they would sleep if they received an alert, but after the end of the study they suggested that they actually engaged in behaviors other than stopping to obtain some restorative sleep.

A number of safety critical events were recorded over the course of the study. As expected, the majority of these were crash relevant conflicts and near-crashes. Regardless of the type of safety critical event, no statistically significant performance differences were present between those drivers with DFM feedback and those without.

The DDWS concept, if implemented properly, should be able to reduce the level of on-the-job drowsiness. The prototype evaluated during this assessment showed a positive trend that may be enhanced by implementing the system improvements suggested herein.

Driver Subjective Ratings and At-Risk Follow-Ups

Two follow-up analyses were conducted. The first compared drivers who had a favoring opinion of the prototype system to those with a disfavoring opinion of the system. Those drivers with disfavoring opinions of the system tended to have a lower rate of valid alerts. Drivers who gave favoring ratings of the system tended to have an increase in safety benefits. Therefore the system was not helpful for all. Although the DFM feedback may have been the influencing factor driving this difference, the very small sample size involved with this analysis precludes any sweeping conclusions. However, the more comprehensive driver acceptance analysis performed suggests that even though drivers agreed with the purpose of the concept, they could not overlook the functional limitations of the prototype.

The second follow-up analysis examined differences between drivers identified to be at risk of drowsiness-related events and those identified as having a lower risk of that type of event. Mixed results were found. Although the at-risk drivers had a reduction in drowsiness events, sleep quality data seemed to indicate that drivers with higher sleep quality (considering only sleep periods over 20 min) had a higher risk. As in the other follow-up analyses, this may be an artifact of the extremely low sample size identified for these analyses. However, the trend presented by the data is one of interest. If drivers who sleep more attempt driving longer hours, the benefit of improved sleep hygiene may not reduce drowsiness-related events. Therefore, it is important that a synergistic plan to coach drivers on the importance of sleep and safe driving conditions is implemented. If not, the benefit of longer sleep periods with higher sleep quality may not positively impact drivers who would like to take advantage of a DDWS. Moreover, if areas to safely park and rest are not available to drivers during their delivery runs, aspects such as improved sleep hygiene, accurate DDWS alerts, and comprehensive drowsiness management trainings might not yield the expected results.

Driver Interactions with the DFM

There were significant differences in the ways drivers interacted with the DFM. As would be expected with many in-vehicle devices, once initial driver preferences were determined the DFM tended to remain at those settings. That is, initial preferred warning sound and sensitivity level were frequently maintained. Some interesting patterns of use were apparent. One example is that drivers tended to use display brightness as a binary adjustment, using either the minimum or maximum of the full range of adjustments available to them.

Two user types became apparent. These two types were manifested as an interaction between the frequency of adjustments in DFM settings and the presence or absence of safety critical events within the driver's shift. Those drivers with safety critical events in their shift tended to have an overall lower level of interaction with the sound, brightness, and sensitivity settings of the DFM as compared to drivers without a safety critical event in their shift. Although the frequency of interactions with the system declined for drivers with and without a safety critical event in their shift across the duration of the study, a statistical difference between the two groups remains. This could be viewed as a difference in driver engagement with the DFM, and suggests that a factor such as driver engagement with the DDWS may also influence driver performance.

PERFORMANCE AND CAPABILITIES

In support of the second major objective of this research effort, VTTI characterized the performance and capabilities of the DDWS. This was done in terms of DDWS operation under naturalistic conditions, system integrity in real-world operations, and the specific human-machine interaction.

The prototype DFM system used in the present study was the result of a long series of experiments by the NHTSA Drowsy Driver Technology Program since 1996. Subsequent versions of this system are currently available to commercial vehicle drivers and operators. The system is based on PERCLOS (Wierwille, 1999), a relatively commonly used metric for determining driver drowsiness. Information about driver slow eye-closures, used to generate the prototype DFM system's estimate of PERCLOS, was captured via near-infrared cameras mounted on the vehicle dashboard.

Although this system allowed for a relatively clear view of the driver's eyes, it was not an optimal placement for the device. This may have contributed to the relatively large number of alerts not related to drowsiness generated by the prototype DFM system. It is recommended that refinements be made to the device algorithms used to determine what information is included and excluded from driver slow eye-closures, along with a more optimal placement of the device (which would require vehicle manufacturer support). These should yield great benefits in the operation of a future DDWS. Some of the aspects highlighted by the driver acceptance analysis suggest that from a usability perspective, participants generally found the device easy and intuitive to operate and understand. However, some would have preferred a smaller device footprint or improved nighttime dial visibility. Yet, by far the most cited issue stemmed from the perceived inaccuracy of the alerts.

DRIVER ACCEPTANCE

Driver acceptance of the DDWS was largely conditional. Generally, participants were satisfied with the training they received and found the device easy and intuitive to operate and understand. In fact, as noted in results presented for the larger participant sample used in Part I, once device sound and sensitivity levels were initially selected, these tended not to be altered. Reports indicated that participants were most likely to utilize device feedback when they felt very tired, although the majority did not report feeling more alert during night driving when using the device. In response to drowsiness alerts, participants often talked on the phone or rolled down a window, as opposed to stopping driving, which was the desired outcome.

Despite explanations provided to participants of the technology's operational shortcomings during daylight, dawn and dusk (indeed as reported in Part I, a high number of false alerts were registered), many found it difficult to disregard such instances when evaluating the device. Attitudes prior to experiencing the DDWS were more optimistic regarding its ability to increase driving safety, whereas after device exposure, participant opinions were generally neutral regarding the extent to which their driving safety had improved. While reluctant to integrate DDWS output into a daily fatigue management routine, participants did indicate only infrequently ignoring the device and most did not stop relying on it despite false alarms. Part I findings, using a larger participant sample, additionally demonstrated that the device did not appear to impact sleep hygiene, whether assessed in terms of sleep quantity or sleep quality.

Survey data and focus group session outcomes revealed that overall satisfaction levels were less positive regarding device performance and more positive about device potential if its "kinks" were worked out. As such, survey findings evidenced that participants were hesitant to endorse this device for use by others or for their own continued use. Nevertheless, there existed a small subset of drivers who found the device to be particularly useful, and who were more satisfied with its performance than the typical FOT participant. Uniformly, this "extreme-favoring" sample subset comprised long-haul drivers, while an opposing group of "extreme-disfavoring" participants contained one long-haul and two line-haul drivers. Although beyond the scope of this effort, further investigation into the possibility that certain trucking operations or driver subgroups in the CMV-driving population are more well-suited to successful use of a device such as the DDWS is warranted.

FLEET MANAGEMENT ACCEPTANCE

In an additional, separate data collection effort, we conducted telephone interviews with trucking company management in order to understand aspects of fleet management acceptance of a conceptual DDWS. Overall, as consistent with reported driver acceptance findings, interviewees expressed support for an alertness-monitoring device as a tool with promise to possibly improve safety, and further, were encouraged by its potential to save money by reducing accidents and injuries. However, there was shared concern that employees would consider it a means for management to monitor their actions and that it would be utilized as basis for punishment or dismissal; this view was corroborated during the focus group sessions where participants indicated a lack of willingness to purchase the device unless guaranteed that the data it collects would not be accessible to outside parties. Nevertheless, interviewees anticipated that driver acceptance could be bolstered through clear discussion of management's intended use for the data, including informing drivers of potential safety hazards, as long as there were not punitive outcomes.

Despite several respondents indicating they did not believe that use of such a device was necessary at their company, reduced insurance rates and government tax credits were cited as incentives that might make them more likely to purchase the technology, especially in light of the rapidly rising fuel costs. The most frequently cited economic benefit of alertness-monitoring technologies was also cost-related, as interviewees anticipated that a safety benefit for the device would translate into reduced accident, injury, legal, and property damage expenditures. Concerns with regard to reliability, accuracy (in particular, false alerts), and maintainability of the device were noted as potentially critical system performance issues.

DEPLOYMENT

In an additional, separate data collection effort, we conducted telephone interviews and in one case, solicited information via email, from technology developers and a trucking advocacy organization in order to understand aspects of deployment for alertness-monitoring technologies. Both technology company interviewees reported that, as part of development, their respective systems had undergone extensive research, design, laboratory, and on-road testing. At interview time, one resulting technology had already been deployed and the other was engaged in further development in conjunction with OEMs. Although not having been directly exposed to alertness-monitoring technologies in the field, the trucking advocacy organization representative who was interviewed expressed support for the premise of such devices, despite believing that motor carriers would be more likely to focus available resources at engineered systems geared toward reducing high crash-likelihood scenarios. As such, trucking company management staff provided additional ideas for promotional and deployment strategies for alertness-monitoring technologies, including documenting driver acceptance and device safety benefits through trial field testing at companies, educational seminars, and various economic incentives. In the case of organizations or among owner-operators where driver fatigue is not perceived as an issue, however, it was offered that no amount of promotion or examples of successful industry deployment would be likely to lead to immediate purchases of such technology.

CONCLUDING REMARKS

The assessment revealed interesting findings and challenges in data reduction and analysis. For example, the results demonstrated that the prototype device was attempting to provide alerts outside of the optimal operating envelope, providing for a large number of alarms not related to drowsiness, and for scenarios in which the device was not meant to operate, including driver distraction. Despite the prototype pitfalls, the DDWS concept seemed to flourish. The results of the assessment suggest that the even a prototype device does have an overall positive impact on driver safety. Therefore, further refinements in the algorithms, operational envelope, and driver-system interaction may promote stronger acceptance and impact in aspects such as on-the-job drowsiness. Drivers who favored the system seemed to receive a benefit from it; therefore, increasing the aspects that were favored and improving the aspects that were not will potentially help the drivers who need it the most (i.e., drivers at-risk of a higher number of drowsiness episodes). The DDWS concept appears to have merit and value to truck drivers, especially long-haul drivers, and trucking companies. Results of the driver acceptance portion of this study strongly point to the need for significant user involvement in system interface development, including “false alarm” suppression and proper integration into a viable fatigue management program. Providing alerts without providing drivers an appropriate level of prescriptive information regarding countermeasures, or situations where drivers find themselves unable to employ a referred countermeasure (e.g., no safe area to pull off the road), will likely lead to frustration and perhaps non-use, in particular if it is believed that the intent of the device is to allow blame-shifting to the drivers in the case of drowsiness-related safety critical events. Additionally, if the driver subpopulation of owner-operators is to benefit, the system will need to be packaged with a basic fatigue management protocol for this driver subpopulation. The fact that there are several manufacturers developing this technology is good news; if done well, there appears to be a waiting market to embrace this safety net device.

FUTURE RESEARCH

The large amount of naturalistic data VTTI collected during the original study makes numerous forms of re-examination possible. Besides being an efficient use of the data available, data mining allows for the originally collected data to extend well beyond its original purposes. A number of future research topics may be explored with the current data set.

Drowsiness and Health

Research has shown that sleep duration may affect the hormones that regulate hunger. For instance, it was found that two days with four hours of time in bed each night were associated with an 18 percent decrease in levels of leptin, the anorexigenic hormone responsible for informing the body of satiation, according to Spiegel, Tasali, Penev, and Van Cauter (2007). The authors report that leptin level decrements were also found to increase appetite for calorie-dense nutrients with high carbohydrate content, including sweets, salty snacks, and starchy foods by 33 to 45 percent. They go on to suggest that sleep curtailment can promote excessive eating.

The data currently available would allow a comparison of the body-mass index of drivers who obtained a greater than average amount of sleep to those drivers who obtained a less than average amount of sleep to investigate whether sleep hygiene is related to their weight. It is expected that drivers obtaining a less than average amount of sleep would possess a greater body-mass index than drivers who obtained a greater than average amount of sleep.

False Alarm Rate Acceptance for a DDWS

One of the more promising avenues of future research (in terms of overall increases in the safety of our transportation system) is using manual PERCLOS values, driving performance data, and other contributing factors pointing toward drowsiness on random samples from the existent naturalistic data sets (i.e., over 200 drivers) to develop guidelines which may be used to identify periods of high drowsiness. Besides providing a much-needed influx of information on real-world occurrences of drowsy driving, this effort would assist in the development of new and future DDWS technologies. By identifying taxonomies of specific driver behaviors, movements, etc., the amount of false alarms in a DDWS could be reduced and in turn increase the effectiveness of the concept. In addition, the current DDWS FOT data set could help identify the false alarm rate that a driver is willing to tolerate before he or she becomes frustrated with the system (e.g., lets the alert continue to sound, becomes angry, hits DFM). It is expected that the more false alarms that the driver receives, the more frustrated he or she will become with the DDWS and the less he or she will comply, but identifying that threshold was outside the scope of the current research effort.

DOT HS 811 117
April 2009

U.S. Department
of Transportation

**National Highway
Traffic Safety
Administration**

NHTSA

www.nhtsa.gov