Behavioral Indicators of Drowsy Driving: Active Search Mirror Checks

July 2022 | Final Report

SAFE-D
SAFETY THROUGH DISRUPTION

Transportation Institute, Virginia Tech

Texas A&M Transportation Institute

San Diego State University
Leadership Starts Here

Disclaimer

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. 05-084	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle Behavioral Indicators of Drowsy Driving: Active Search Mirror Checks	5. Report Date July 2022 6. Performing Organization Code:
7. Author(s) Jason E. Meyer Robert E. Llaneras	8. Performing Organization Report No. Report 05-084
9. Performing Organization Name and Address: Safe-D National UTC Virginia Polytechnic Institute and State University Virginia Tech Transportation Institute 3500 Transportation Research Plaza Blacksburg, Virginia 24061, USA	10. Work Unit No. 11. Contract or Grant No. 69A3551747115/05-084
12. Sponsoring Agency Name and Address Office of the Secretary of Transportation (OST) Office of the Assistant Secretary for Research and Technology University Transportation Centers Program Department of Transportation Washington, DC 20590 United States	13. Type of Report and Period Final Research Report 14. Sponsoring Agency Code

15. Supplementary Notes

This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation – Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program.

16. Abstract

Driver impairment due to drowsiness or fatigue has a significant impact on the safety of all road users. Assessing an impairment such as driver drowsiness through the use of vehicle-based technology continues to be an area of interest. Both the initial detection and continued monitoring of driver drowsiness have been the emphasis of vehicle-based driver monitoring systems (DMS). Particularly, in-vehicle eye tracking systems have been implemented as a way of determining driver state. Specifically, when hands-free driving assistance features are engaged, measures such as the driver's percentage of eye closure (PERCLOS) are being considered to determine driver drowsiness. However, one challenge of such a metric is its reliability, particularly with regard to false alarms (when a DMS indicates the driver is drowsy but in fact is not). Therefore, the use of more gross-level driver behavioral measures may serve as a way of cross-checking the assessments of a DMS. This work mined an available dataset in order to examine driver search behavior, with the goal of identifying relationships between driver vigilance and drowsy driving, to test the hypothesis that driver search behavior (e.g., mirror checks) degrades with increasing levels of drowsiness. Based on a statistical comparison of participant driving data encompassing instances of alert, moderately drowsy, and drowsy driving, no significant differences were observed among these three classifications.

17. Key Words Publication, guidelines, report, brochure, communication, marketing	18. Distribution Statement No restrictions. This document is available to the public through the Safe-D National UTC website, as well as the following repositories: VTechWorks, The National Transportation Library, The Transportation Library, Volpe National Transportation Systems Center, Federal Highway Administration Research Library, and the National Technical Reports Library.

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 10	22. Price $0

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Abstract

Driver impairment due to drowsiness or fatigue has a significant impact on the safety of all road users. Assessing an impairment such as driver drowsiness through the use of vehicle-based technology continues to be an area of interest. Both the initial detection and continued monitoring of driver drowsiness have been the emphasis of vehicle-based driver monitoring systems (DMS). Particularly, in-vehicle eye tracking systems have been implemented as a way of determining driver state. Specifically, when hands-free driving assistance features are engaged, measures such as the driver's percentage of eye closure (PERCLOS) are being considered to determine driver drowsiness. However, one challenge of such a metric is its reliability, particularly with regard to false alarms (when a DMS indicates the driver is drowsy but in fact is not). Therefore, the use of more gross-level driver behavioral measures may serve as a way of cross-checking the assessments of a DMS. This work mined an available dataset in order to examine driver search behavior, with the goal of identifying relationships between driver vigilance and drowsy driving, to test the hypothesis that driver search behavior (e.g., mirror checks) degrades with increasing levels of drowsiness. Based on a statistical comparison of participant driving data encompassing instances of alert, moderately drowsy, and drowsy driving, no significant differences were observed among these three classifications.

Acknowledgements

This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation – Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program.

Table of Contents

List of Figures

Introduction and Background

Driver impairment due to drowsiness has a significant impact on the safety of all road users (Knipling & Wang, 1994). Assessing drowsiness through the use of vehicle-based technology continues to be an area of interest, and one filled with challenges. Driver monitoring systems (DMS) have been employed to track drivers' attention to the forward roadway. Although not the original emphasis of the technology, DMS could also be applied to monitoring driver eye-glance behavior to determine drowsiness. Already, performance-based metrics (such as vehicle lateral control) are being studied as indices of driver drowsiness (Van Loon et al., 2015). Additionally, there may be benefits from incorporating the specific eye-glance-related data collected from a DMS in determining the driver's state of alertness, in terms of both the initial detection and the continued monitoring of drowsiness.

The method employed by some DMS is measuring the driver's percentage of eye closure (PERCLOS) (Wierwille et al., 1994) as a predictor of drowsiness (Jackson et al., 2016). Especially in cases when the task of driving becomes monotonous, the prevalence of drowsiness will occur (Vogelpohl et al., 2018). When hands-free partially automated driving assistance features are engaged, a measure such as PERCLOS could be utilized to determine driver alertness levels and drowsiness. A significant challenge in this area continues to be making reliable assessments of driver state since camera-based DMS technologies may not work under the full range of real-world driving environments. Of particular concern is reducing false alarms, cases where DMS technologies may incorrectly flag an alert driver as drowsy.

The use of more gross-level driver behavioral measures (e.g., mirror checks) may serve as a way of cross-checking and increasing the accuracy and reliability of DMS assessments. In the partially automated realm, this work supports the identification of driver drowsiness state to help avoid instances where drivers fall asleep behind the wheel of "self-driving" vehicles that still require drivers to bear some level of responsibility in taking control of the vehicle (Vogelpohl et al., 2018).

This work used an available dataset to examine driver search behavior, with the goal of identifying relationships between driver vigilance and drowsy driving. The dataset, from a proprietary study conducted by the Virginia Tech Transportation Institute (VTTI), consists of DMS data of multiple drivers who completed a prescribed route intended to induce periods of drowsy driving. The hypothesis was that driver search behavior degrades with increasing levels of drowsiness. Practical applications may be to incorporate driver search behavior into the "toolbox" of metrics for estimating driver drowsiness.

Method

In order to investigate relationships between driver vigilance and drowsy driving, an existing dataset was employed. A previous VTTI research study involved participants driving a prescribed route in order for researchers to capture instances of drowsy driving. Participants were instructed to drive an extended 5-hour trip, intended to induce periods of drowsy driving. The vehicle included a DMS unit with a single camera and two integrated LED units mounted on the vehicle steering column. The DMS was time-synchroninzed with the VTTI-installed data acquistion system (DAS) unit recording key driver measures, such as PERCLOS, blink duration, tracking status, gaze location (on-road or off-road), and other measures of driver eye state, and onboard DAS cameras. The DAS cameras provided multiple video feeds of the driving session, including driver face, forward roadway, and rear roadway views.

Two researchers accompanied each driver, monitoring the session and driver performance from the rear seat of the vehicle. The participants were instructed to drive the route, to minimize interaction with experimenters in the vehicle, to minimize lane changes, and to keep cruise control engaged as much as possible. This resulted in a 5-hour driving session consisting of multiple instances of recorded alert driving, as well as moments of driving "epochs" of drowsy driving. The classification of "alert" and "drowsy" driving epochs was chosen with "alert" being PERCLOS values < 5% and "drowsy" being PERCLOS values > 10%. A "moderately drowsy" classification (PERCLOS > 5% but < 10%) was also employed in order to provide more insight and expand the mere binary of "alert" and "drowsy" conditions.

Approval through Virginia Tech's Institutional Review Board (IRB) for conducting research with human subjects was needed in order to begin mining the existing database. Participants in that study had allowed their data to be used solely for the purpose of that project; and IRB approval was required in order to use existing participant video and images in conducting an investigation into driver vigilance and drowsy driving.

Reduction and Sampling Driving Epochs

The existing data from the previous project provided a view of the driver's face through a synchronized video feed. Based on this video data, in conjunction with the PERCLOS data collected, a protocol for classifying glances to mirrors and other pertinent locations was developed. Glance locations were determined by a video analyst assigning each individual glance by the driver to locations either inside or outside the vehicle. Thus, the objective was to manually have analysts code driver search behavior and record glances to the mirrors, as well as over-the-shoulder glance behavior, under different driver states, including alert and drowsy states.

The existing data consists of a sample of 40 drivers, all of whom completed the 5-hour trip. During these trips, each driver experienced multiple episodes of drowsy driving (based on PERCLOS data collected at the time). The sampling of drivers was limited to the 32 of the 40 drivers who completed their sessions in a daytime setting in order to eliminate any effect that a darkened environment may have on the results. Specific driving epochs from this collection were isolated for further analysis. Samples for each driver were drawn across "alert," "moderately drowsy," and "drowsy" driving. Several time epochs were used for reduction and analysis, including larger 5-minute epochs and a focused 1-minute epoch. At a minimum, three epochs were analyzed per driver (one epoch for each of the three PERCLOS classifications). This resulted in 135 cases or epochs varying in driver alertness levels (Table 1) (e.g., alert, moderately drowsy, drowsy).

Table 1. Total Number of Epochs by PERCLOS Classification

PERCLOS Classification	Number of Epochs
Alert	47
Moderately Drowsy	35
Drowsy	53
Total	**135**

In an effort to minimize variability, only epochs that were comparable with regard to traffic, speed and driving maneuvers were selected. For example, instances where drivers changed lanes or with incomparable traffic conditions in the environment were both eliminated from consideration as such conditions would influence driver glance behavior. This reduced the original 135 epochs under consideration to 83 epochs (Table 2).

Table 2. Total Number of Epochs – Limit to "No Lane Changes"

Lane Changes	PERCLOS Classification	Number of Epochs
Yes	Alert	24
Yes	Moderately Drowsy	13
Yes	Drowsy	15
Total Yes	**-**	**52**
No	Alert	23
No	Moderately Drowsy	22
No	Drowsy	38
Total No	-	**83**

The PERCLOS value generated from the DMS is the result of a 150-second rolling window where the system is measuring eyelid closure. In an investigation of glance behavior, the time leading up to the transition from "alert" to "drowsy" is critical. For purposes of analysis, two time windows were defined. A 5-minute epoch was used to provide a larger window of analysis. The specific duration was selected in order to capture a window of time (5 minutes prior to reaching a "drowsy" or "alert" or "moderately drowsy" threshold) that encompasses both the PERCLOS collection window and the events leading up to the transition. A smaller 1-minute time epoch was also defined to capture glance behavior just prior to the maximum observed drowsy levels. A comparable 1-minute alert window was also developed for each driver. Thus, two time-based epochs were used to explore search behavior: a broader 5-minute window and a more focused 1-minute window. These time windows are illustrated in Figure 1.

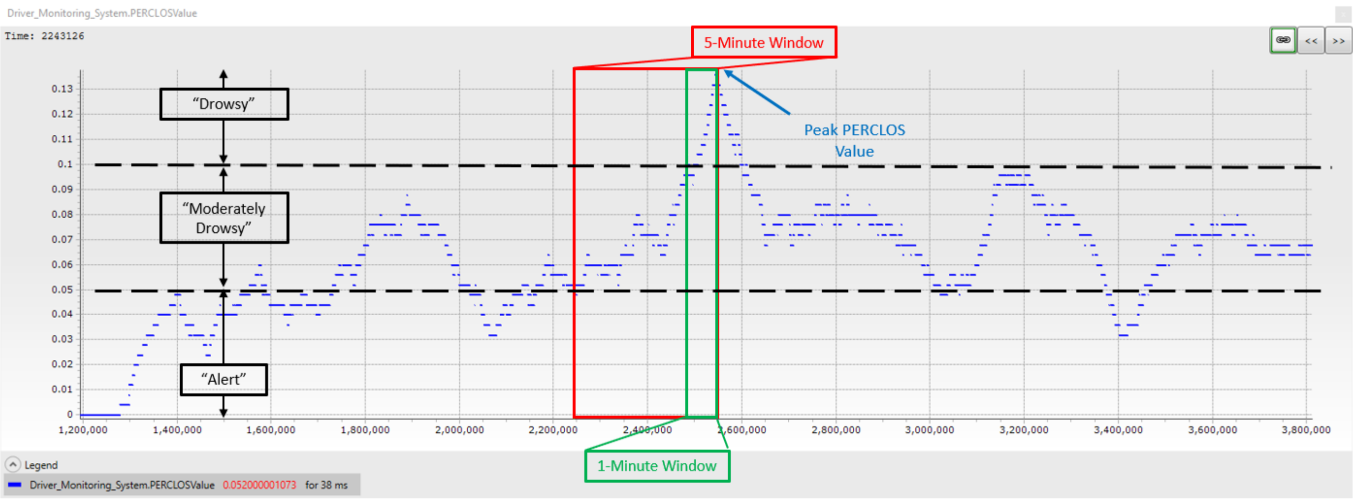

Figure 1. Diagram. 5-minute and 1-minute time windows.

Therefore, for this effort, video analysts viewed and coded participants’ eye-glance frequencies during a 5-minute window for each of the PERCLOS driving classifications of “alert,” “moderately drowsy,” and “drowsy.” Analysts were trained to identify specific locations of glances to include mirror locations of interest such as “Rearview,” “Left Mirror,” or “Right Mirror,” as well as glances “Over-The-Shoulder” for instances when drivers may be glancing to a blind spot location over their left shoulder.

Results

The number of glances to locations was normalized to define a glance rate—the number of glances per minute—in order to make comparisons with a common scale. The results below are based on the 32 daytime drivers who all experienced comparable epochs of “alert,” “moderately drowsy,” and “drowsy” driving. Data are presented using two alternate analysis frameworks or approaches. The first uses cases as the unit of analysis, where each observation (alert, moderately drowsy, drowsy) is treated independently in the analysis; in all 83 cases are included. The second uses the driver as the unit of analysis, where the data for each driver are aggregated to represent a single point estimate or mean for each driver; in all 32 drivers are represented, each weighted equally.

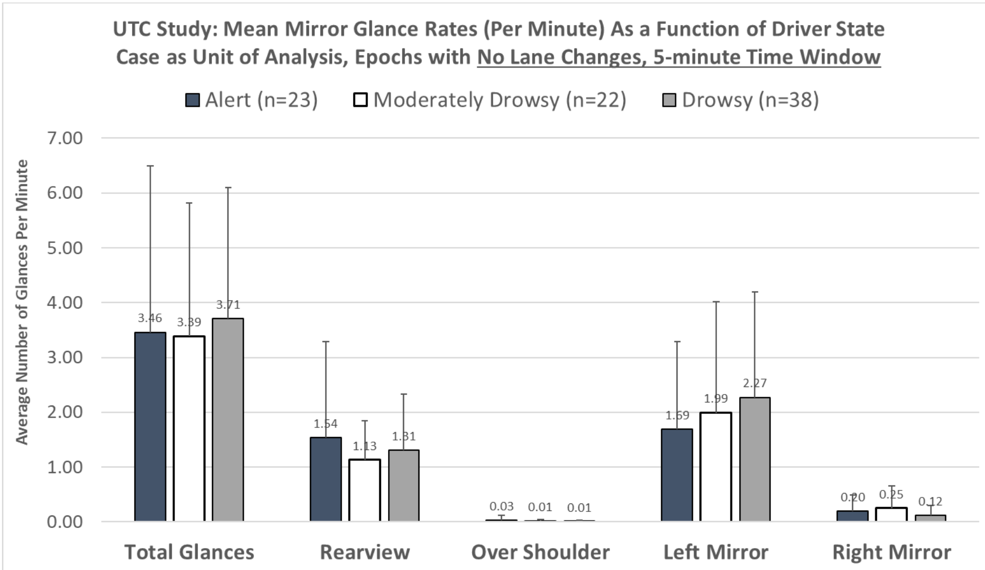

Figure 2. Chart. Mean mirror glance rates – case as unit of analysis (5-minute epoch).

Figure 2 illustrates the average glance rate to each of the specified locations over the 5-minute sample period. Among both the total number of glances, as well as the specific mirrors/locations under investigation, glance rates were highest for the left mirror and the rearview mirrors. Glances to the right mirror (outside passenger mirror) and over-the-shoulder glances were relatively low-frequency events; this is likely due to the fact that the vehicle was positioned in the right lane and lane changes were excluded or minimized during the sampled 5-minute epochs. Statistical analyses were performed using the SAS analysis software with effects evaluated using the PROC GLM procedure (general linear model) to perform analysis of variance (ANOVA) tests. Alpha levels were set to 0.05, and experiment-wise error was controlled using Duncan's Multiple Range post hoc test to follow-up significant main effects. Results found that average number of glances per minute were not significantly different when considering the PERCLOS classifications of the drivers. Some drivers experienced multiple drowsy epochs that would lead to an inaccurate weight given to those drowsy epochs when making a comparison. In order to eliminate any single driver from inaccurately influencing the data by the number of usable cases extracted from their data, each driver was equally weighted (in other words, analyses used the driver as the unit of analysis).

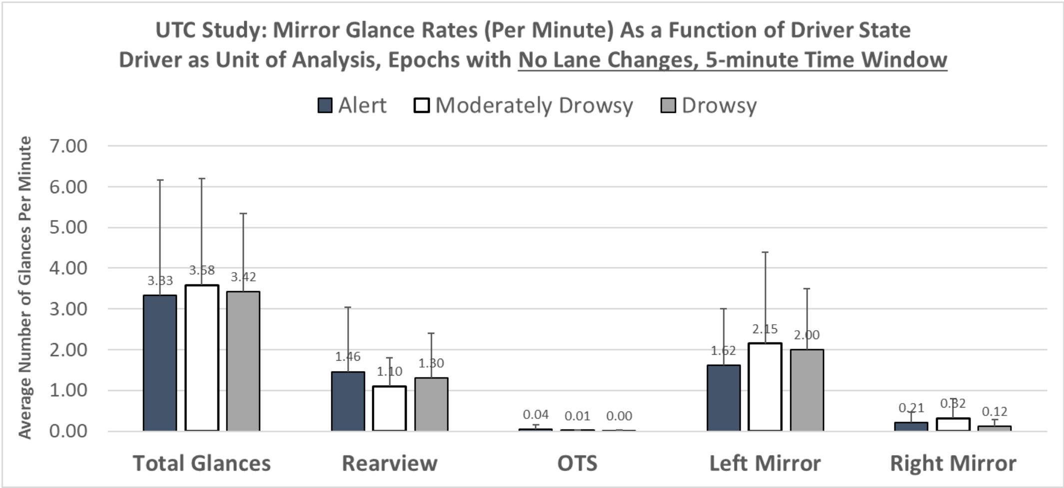

Figure 3. Chart. Mirror glance rates – driver as unit of analysis (5-minute epoch).

Figure 3 illustrates the average glance rate to each of the specified locations over the sample period using the driver as the unit of analysis. As shown in the figure, drivers appeared to make less frequent (fewer) glances to the rearview mirror when “moderately drowsy” or “drowsy” relative to an “alert” state. However, this trend was small and not statistically significant, with substantial variability among samples (large standard deviations). Data with regard to the left mirror (driver mirror) show an unexpected pattern, with elevated glance rates under the drowsy conditions. This may be a result of other uncontrolled factors such as passing traffic, for example. Overall, no strong evidence was found to indicate that glance behavior (measured by glance rates) was significantly impacted by driver state as examined here.

Based on these results, a reanalysis of the data was conducted, limiting the duration of analysis to a 1-minute window of time within the peak PERCLOS value (for the “moderately drowsy” and “drowsy” conditions) and a 1-window of time when the participant was classified as “alert” (based on the PERCLOS value). It was suspected that perhaps a 5-minute window of time allowed for too much variability in the changes reflected by PERCLOS, and that 60 seconds of glances may provide insight into any differences in search behavior as related to PERCLOS classification. Results of a more limited, 1-minute time window are shown below in Figure 4.

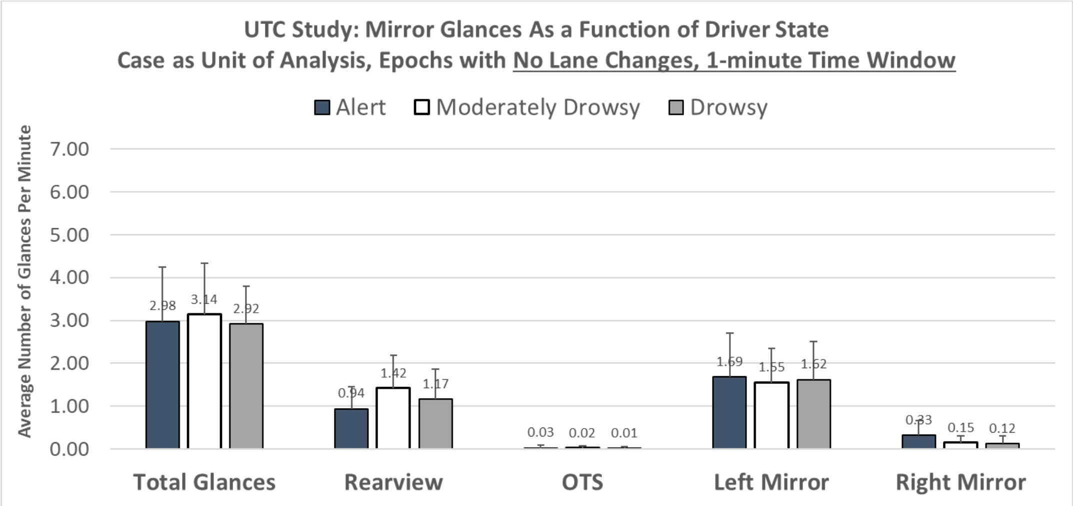

Figure 4. Chart. Mirror glance rates – Case as unit of analysis (1-minute epoch).

Again, despite a narrower 1-minute window of observation, no strong evidence was found to indicate that glance behavior was significantly impacted by driver drowsiness level. Figure 4 shows that drivers glanced more often to the rearview mirror when alert compared to periods of higher drowsiness. However, this difference was not found to be statistically significant. Again, there is an unexpected elevated number of glances to the left mirror under drowsy conditions compared to when the driver was "alert."

Similar to the approach to observing the 5-minute window using individual drivers as the unit of analysis, Figure 5 shows the results of each of the 32 drivers being equally weighted in the analysis with the narrower 1-minute time window.

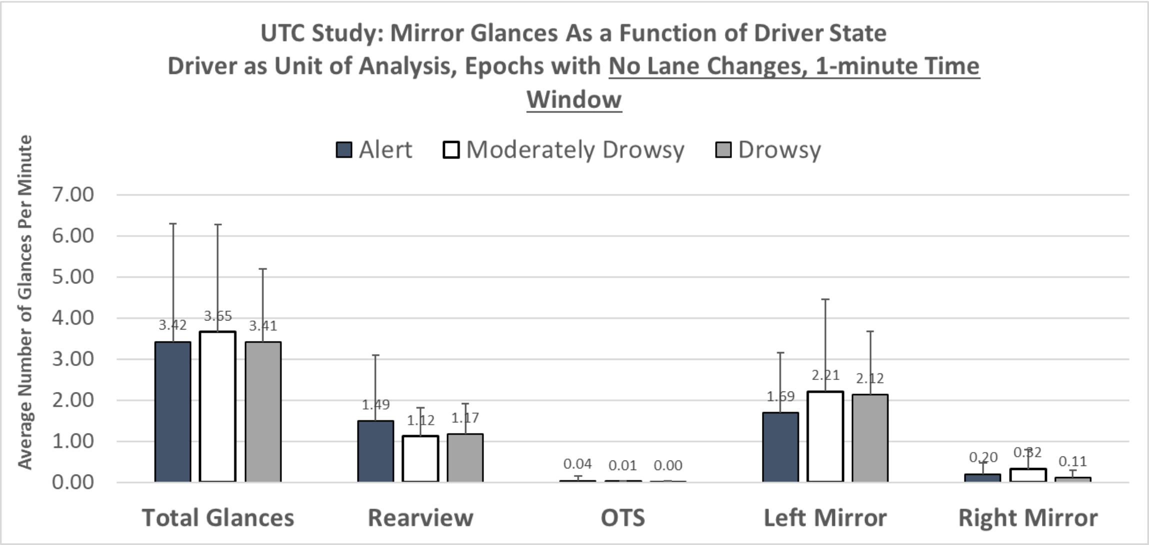

Figure 5. Chart. Mirror glance rates – driver as unit of analysis (1-minute epoch).

Discussion

Based on the analysis, there was no statistically significant difference in the mean glance rate among the different PERCLOS classifications. It was expected that epochs of participants driving while "alert" would result in a significantly higher number of glances (glance rate) than when drowsy.

While it is possible there is no difference in drivers' levels of vigilance as it pertains to drivers' corresponding level of drowsiness, there are a number of factors that may have contributed to the findings. The route selected, while successful at inducing drowsy episodes in drivers, did consist of signalized intersections. This limited the instances of uninterrupted portions of the drive when drivers would be performing typical "scanning" of the environment as a driver decelerating upon approaching an amber or red traffic light will engage in different gaze behavior compared to maintaining a constant speed. Glance rates may also be sensitive to the traffic environment and location (mirror), suggesting that future studies or analyses should control for traffic volume or events such as cars approaching or passing.

Conclusions and Recommendations

While the findings of this exploratory study did not find significant differences among the PERCLOS classifications, this does not suggest that such differences do not exist. Considerations such as a larger sample size, as well as a more uniform sampling of data epochs, may provide more insight into any differences in mirror search behavior. Future efforts may focus on analyzing a window of time containing minimal traffic leading up to the classification of driver PERCLOS state. The protocol for the study, from which the data were mined, required the participants to drive at speeds lower than the posted speed limits. Therefore, the majority of driving epochs consisted of a queue of vehicles passing the slower participant vehicle in the neighboring left lane, thereby drawing attention to mirrors across all PERCLOS classification conditions. While a 5-minute window of time was chosen in order to increase reliability that drivers would be considered "drowsy" or "alert," a smaller 1-minute window of time was also considered in order to limit variability that may be found in the driving environment (such as the presence of traffic passing in the left lane). With this exploratory study, a future consideration would be to conduct a study in an environment with minimal instances of surrounding traffic or control to decrease the likelihood of a queue of vehicles passing the participant vehicle.

Additional Products

The Education and Workforce Development (EWD) and Technology Transfer (T2) products created as part of this project can be downloaded from the project page located on the Safe-D website. As this project did not produce any data, no datasets are available from this project on the Safe-D Collection of the VTTI Dataverse.

Education and Workforce Development Products

It was expected that the results of this effort would be used in the development of detection algorithms for the classification of driver states. Based on the results of this project, however, further data mining efforts, a refinement of the criteria for selection of epochs, a more controlled environment of traffic volume, or an increased sample of corresponding epochs would be necessary before developing algorithms for driver state classification.

Technology Transfer Products

An algorithm for classification of driver states was expected to be developed as a Technology Transfer Product. However, with a lack of statistically significant results, a classification model based on these data would not be reliable.

Data Products

The dataset used for analysis has been posted to the Safe-D repository, and can be accessed here: https://doi.org/10.15787/VTT1/R8WY4F

References

1. Jackson, M. L., S. Raj, R. J. Croft, A. C. Hayley, L. A. Downey, G. A. Kennedy, and M. E. Howard. (2016). Slow Eyelid Closure as a Measure of Driver Drowsiness and its Relationship to Performance. *Traffic Injury Prevention*, Vol. 17, No. 3, 2016, pp. 251-257.

2. Knipling, R., and J. Wang. Crashes and Fatalities Related to Driver Drowsiness/Fatigue. Washington, DC: National Highway Traffic Safety Administration, 1994.

3. Van Loon, R., R. F. T. Brouwer, and M. H. Martens. Drowsy Drivers' Under-performance in Lateral Control: How Much Is too Much? Using an Integrated Measure of Lateral Control to Quantify Safe Lateral Driving. *Accident Analysis & Prevention*, 2015.

4. Vogelpohl, T., M. Kühn, T. Hummel, and M. Vollrath. Asleep at the Automated Wheel Sleepiness and Fatigue During Highly Automated Driving. *Accident Analysis & Prevention*, 2018.

5. Wierwille, W. W., S. S. Wreggit, C. L. Kirn, L. A. Ellworth, L. and R. J. Fairbanks. Research on Vehicle-Based Driver Status/Performance Monitoring; Development, Validation, and Refinement of Algorithms for Detection of Driver Drowsiness. Report DOT HS 808 247. Washington, D.C.: National Highway Traffic Safety Administration, 1994.