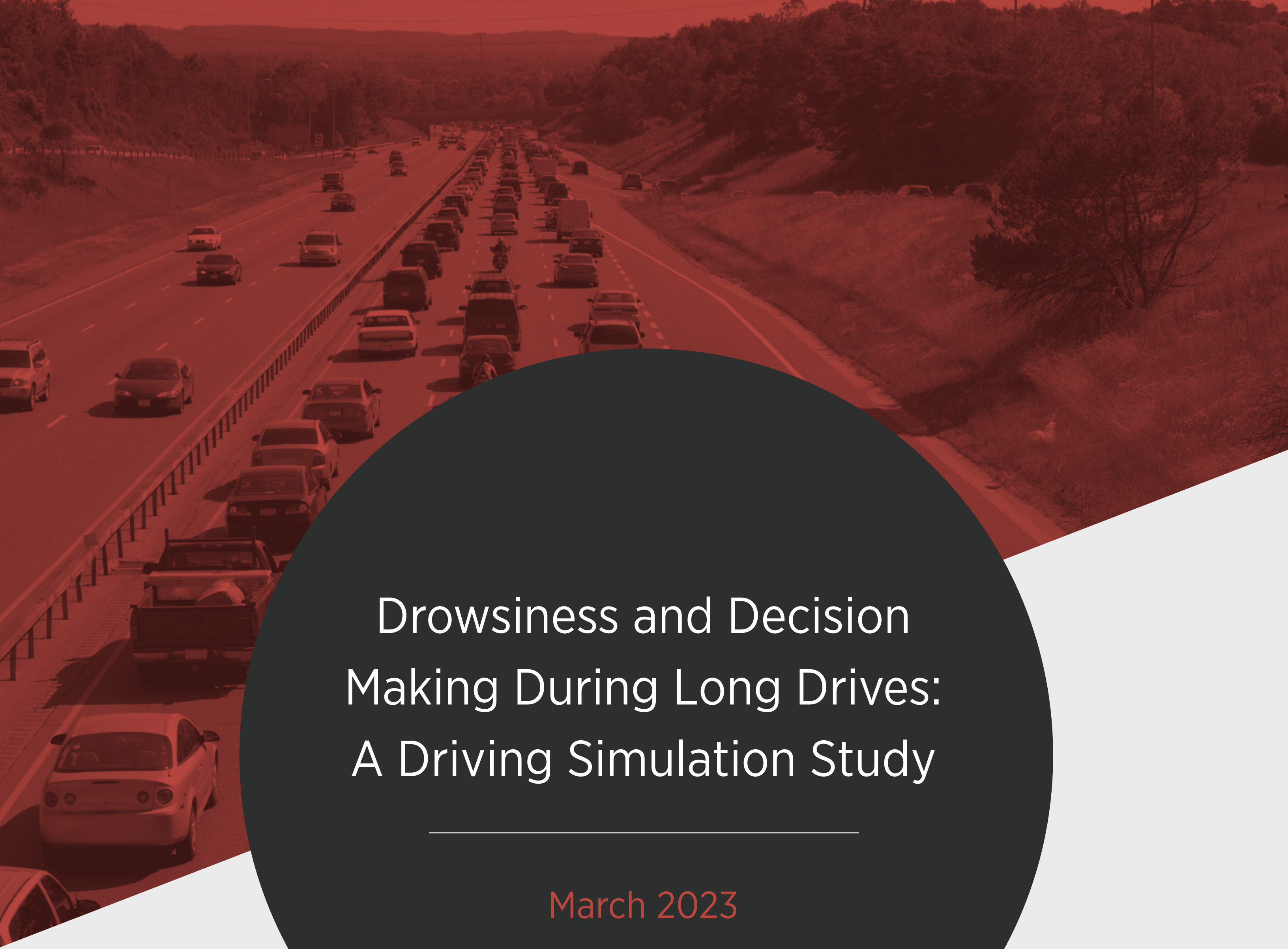

Drowsiness and Decision Making During Long Drives: A Driving Simulation Study

March 2023

607 14th Street, NW, Suite 201 | Washington, DC 20005 | 202-638-5944

FOUNDATION FOR TRAFFIC SAFETY

Title

Drowsiness and Decision Making During Long Drives: A Driving Simulation Study

March 2023

Authors

John Gaspar & Cher Carney

University of Iowa

Foreword

The AAA Foundation for Traffic Safety seeks to guide drivers toward safe decisions. When experiencing drowsiness behind the wheel, research shows that the safe decision is to stop driving and rest, because the only cure for drowsiness is sleep. Unfortunately, drivers often underestimate how drowsy they are, or overestimate their ability to fight it, sometimes with tragic consequences. Our research estimates that drowsy driving is a factor in as many as one in five deadly crashes.

This report describes the results of an experiment designed to examine drowsy drivers' awareness of how drowsy they are, and their decisions about when to stop to rest. The findings of this research offers insights that can be used to educate drivers about the dangers of drowsy driving and the importance of recognizing the early warning signs of drowsiness. This report should be of interest to researchers, practitioners such as healthcare providers, risk managers, and driving instructors, as well as other health and safety stakeholders.

C. Y. David Yang, Ph.D.

President and Executive Director
AAA Foundation for Traffic Safety

About the Sponsor

AAA Foundation for Traffic Safety
607 14th Street, NW, Suite 201
Washington, D.C. 20005
202-638-5944
www.aaafoundation.org

Founded in 1947, the AAA Foundation for Traffic Safety in Washington, D.C., is a nonprofit, publicly supported charitable research and educational organization dedicated to saving lives by preventing traffic crashes and reducing injuries when crashes occur. Funding for this report was provided by voluntary contributions from AAA/CAA and their affiliated motor clubs, individual members, AAA-affiliated insurance companies, and other organizations or sources.

Table of Contents

Table of Figures

Executive Summary

Background

Drowsiness plays a large and often underestimated role in traffic crashes, injuries, and deaths. While official statistics from the U.S. Department of Transportation indicate that driver drowsiness is a factor in only about 2%–3% of crashes, injuries, and deaths nationwide, most experts regard these statistics as a substantial underestimate. Research by the AAA Foundation for Traffic Safety has estimated (a) that as many as 6%–10% of all police-reported crashes and 16%–21% of fatal crashes may involve driver drowsiness, (Tefft, 2012; 2014; Owens et al., 2018), (b) that missing as little as 1 hour of sleep significantly increases a driver's risk of causing a crash, and (c) that much of the excess risk associated with sleep deprivation is attributable to increased risk of committing a variety of ordinary driving errors, not only to falling asleep at the wheel (Tefft, 2018).

In a review of recent literature and discussions with experts to document the state of knowledge regarding drowsy driving countermeasures, the AAA Foundation for Traffic Safety found that obtaining sufficient sleep, napping, and consuming caffeine are among the few evidence-based countermeasures drivers can employ to prevent or mitigate drowsy driving (Bayne et al., 2022). Unfortunately, previous research has shown that drivers may underestimate the likelihood that they fall asleep (Reyner & Horne, 1998). Thus, drivers might choose to begin driving or resist stopping to take a break when their driving performance is likely to become or is already impaired by drowsiness. There is a need to understand the relationship between self-perceived and objective measures drowsiness, as well as how perceived and objectively measured drowsiness relate to drivers' decisions regarding whether or when to stop driving or employ other countermeasures to attempt to mitigate their drowsiness and maintain safety.

Objective

This study investigated the relationship between subjective ratings of drowsiness, objective measures of drowsiness, and measures of driving performance among participants in a driving simulation study designed to induce drowsiness. The study also examined what factors influence drowsy drivers' decisions regarding whether to take breaks during long drives.

Method

The study utilized a novel driving simulator methodology to examine drowsy driving during a long overnight drive. Participants followed a protocol designed to induce partial sleep deprivation prior to the session, and then drove a route designed to induce drowsiness. Participants had the option to stop to rest at designated areas throughout the drive. At various points throughout the session, several measures of drowsiness, including self-ratings and objective measures derived from video-coded eyelid closures, were collected. To replicate the motivational tradeoffs of drowsy driving, the study utilized a novel incentive methodology to mimic the decision-making tradeoff between continuing to drive to reach their destination more quickly versus stopping to rest to maintain safety.

Results and Discussion

Results showed that self-assessments of drowsiness were often poorly calibrated with objective drowsiness based on eyelid closures. This indicates that drowsy drivers might over- or underestimate their level of drowsiness, which could lead to situations in which drivers might decide to continue driving despite high levels of objective drowsiness. Furthermore, the data showed that self-ratings of drowsiness were the key predictor of the likelihood that drivers would stop to take a break. Other factors, including objective drowsiness and driving performance, were not significantly associated with the likelihood of stopping to take a break. This suggests that despite the finding that self-assessments of drowsiness may be poorly aligned with objective measures, drowsy drivers rely on self-ratings when determining whether to stop to take a break. Importantly, many drivers continue driving even when they rate their drowsiness as very high.

Conclusion

These results can help inform efforts to educate the public about drowsy driving. Drivers should understand that their self-perceived drowsiness may not align with other established objective measures of drowsiness. Thus, drivers should be encouraged to consider stopping to rest before they feel severely drowsy. The results of this study can also help to inform drowsy drivers that their own self-evaluations of drowsiness may be inaccurate and focus on promoting the use of effective, evidence-based countermeasures in the challenging context of drowsy driving.

Introduction

While official statistics from the U.S. Department of Transportation indicate that driver drowsiness is a factor in only about 2%–3% of crashes, injuries, and deaths nationwide, most experts regard these statistics as reflecting only the tip of the iceberg. Drowsiness refers to the homeostatic drive for sleep, or sleep pressure (Shen et al., 2006). Research by the AAA Foundation for Traffic Safety has estimated (a) that as many as 10% of all police-reported crashes and 16%–21% of fatal crashes may involve driver drowsiness, (Tefft, 2012; 2016; Owens et al., 2018), (b) that missing as little as 1 hour of sleep significantly increases a driver's risk of causing a crash, and (c) that much of the excess risk associated with sleep deprivation is attributable to increased risk of committing a variety of ordinary driving errors, not only to falling asleep at the wheel (Tefft, 2018).

In a review of recent literature and discussions with experts to document the state of knowledge regarding drowsy driving countermeasures, the AAA Foundation for Traffic Safety found that obtaining sufficient sleep, napping, and consuming caffeine are among the few evidence-based countermeasures drivers can employ to prevent or mitigate drowsy driving (Bayne et al., 2022). Unfortunately, previous research has shown that although drowsy drivers are generally aware that they are drowsy, some drivers underestimate the likelihood that they will soon fall asleep (Reyner & Horne, 1998). In one study that interviewed drivers after their involvement in crashes in which they had fallen asleep, 60% reported that they had felt only moderately, slightly, or not at all drowsy before the crash (Stutts et al., 1999). Thus, drivers might choose to begin driving or resist stopping to take a break when their driving performance is likely to become or is already impaired by drowsiness. There is a need to understand the relationship between self-perceived and objective measures drowsiness, as well as how perceived and objectively measured drowsiness relate to drivers' decisions regarding whether or when to stop driving or employ other countermeasures to attempt to mitigate their drowsiness and maintain safety.

Drowsiness and drowsiness mitigation are difficult topics to study for several reasons. Part of the challenge in studying drowsy driving is the multitude of ways in which drowsiness can be assessed before, during, and after a drive. Self-ratings, such as the Karolinska Sleepiness Scale (KSS; Åkerstedt & Gillberg, 1990), ask drivers to rate their drowsiness on a scale from alert to sleepy. These subjective measures are often captured during breaks in the driving task. Objective measures include eyelid closure metrics such as percent eyelid closure (Wierwille et al., 1994), number of long (typically > 500ms) eyelid closures, and electroencephalography (EEG). Driving performance measures such as lane keeping performance and lane departure frequency are often used as indices of drowsy driving as well, as they have been shown to be strongly correlated with drowsiness (e.g., Knipling & Wierwille, 1994).

Beyond measuring drowsiness, there are many challenges inherent in understanding the decision-making process that drivers use to determine whether to continue driving versus stop to take a break and rest. One challenge is that the various measures of drowsiness, from subjective ratings of drowsiness to objective measures such as eyelid closures, often do not overlap in test settings. Measures of eyelid closures have long been used as a gold standard for measuring objective drowsiness in naturalistic settings (Hanowski et al., 2008). However, little is known about the relationship between drivers' subjective self-assessments of their drowsiness and such objective measures. That is, are drivers good judges of their own drowsiness?

Another challenge is that studying break-taking behavior in a drowsy driving context is difficult in experimental settings. Although break taking can be observed in naturalistic and on-road data, it is difficult to quantify the many diverse factors that may influence drivers' decisions regarding whether to take a break or attempt to continue driving. Doing so would seemingly require controlled experimental research. However, studying drivers' decisions to take breaks in the context of a controlled experiment presents its own set of challenges. For example, the motivation to try to quickly arrive at one's destination (e.g., one's home or workplace) is typically absent when a study participant is essentially "captive" while they are in a laboratory participating in an experiment. Even if a research participant can be incentivized to attempt to complete an experimental drive quickly, their motivation to drive safely and avoid crashing is likely different in a laboratory setting—where risk of causing or sustaining physical injury or property damage is absent—than it would be in a real-world setting. Studying break taking in a controlled setting such as a driving simulator thus requires replicating the complex motivational tradeoffs present in a real-world drowsy-driving situation. However, it may be possible to manipulate incentive structures to recreate the motivational conditions of drowsy driving. This would provide the advantage of being able to study drowsiness in a large sample of drivers, across controlled drowsiness and driving conditions, and to sample an array of subjective and objective measures of drowsiness and driving performance.

Objectives

The goal of this research was to generate data to help inform messaging and countermeasures related to drowsy driving. The objectives of this study were to investigate two central research questions:

1. How do subjective measures of driver drowsiness relate to objective measures of drowsiness and driving performance?
2. When and why do drivers stop to rest versus continue to drive?

The study utilized a drowsy-driving protocol in which participants drove for approximately 3 hours in a driving simulator. An incentive scheme was used to attempt to mimic the real-world motivational tradeoffs that might lead a person to stop for a break or to continue driving while drowsy (e.g., a desire to reach one's destination safely, but also quickly). Recreating the motivational tradeoffs of drowsy driving is critical to accurately understanding the factors that influence decision making. Using a driving simulator allowed the research team to systematically generate moderate to high levels of drowsiness and to evaluate driver state and performance in an experimentally controlled (and safe) setting.

Research question 1 was addressed by collecting both subjective measures (i.e., self-ratings) and objective measures (eye closures) of drowsiness over the same set of time windows. This allowed the comparison of driver self-ratings with an objective assessment of drowsiness.

Research question 2 was evaluated by examining the relationship between a variety of factors including but not limited to those discussed above with respect to the likelihood of taking a break during the drive. The objective was to identify the factors that influenced whether a driver was likely to take a break at a given point in the drive.

Method

Recruiting and Eligibility Criteria

Study participants were recruited from the NADS subject registry and mass email solicitations to the University. All study procedures were approved by the University of Iowa Institutional Review Board and all participants provided written informed consent prior to participation. Participants were required to meet the following criteria to be eligible to participate in the study:

- Between ages 21 and 55
- Hold a valid U.S. driver's license with no restrictions for vision
- Drive at least 2,000 miles per year
- Drive without special equipment
- Live within 30 minutes of University of Iowa Research Park
- Have normal sleep patterns
- Not have obstructive sleep apnea
- Have not been diagnosed with a serious illness or taking any medication that impairs driving or induces drowsiness
- Report no history of motion sickness
- Report no history of neck and/or back pain
- Had not participated in previous simulator studies on drowsy driving at the National Advanced Driving Simulator (NADS)
- Be able to participate in an overnight study visit of up to nine hours and refrain from sleep after waking at a specified time on the day of the study visit
- Be willing to complete an activity log on the day of the study visit
- Refrain from consuming caffeine after 1 p.m. on the day of the study visit
- Pass a breath-alcohol test upon arrival for the study visit

Apparatus

Drives were completed on the NADS miniSim simulator. The simulator was fixed-base (i.e., no motion) with three forward channels; realistic steering wheel, brake, and accelerator pedals; and a simple dashboard interface. An entertainment system was added to allow for participants to listen to music (as a countermeasure one might seek to employ to combat drowsiness). An armrest with a cupholder was included in the simulator as a location to store drinks. Figure 1 shows an image of the miniSim.

Figure 1. NADS miniSim

Driving Task

The study drive consisted of a 150-mile drive along a highway interstate loop in dark, nighttime lighting conditions that were consistent throughout the drive (Figure 2). The interstate loop was 82 miles long, meaning that participants completed just under two full loops. Because the road geometry was similar, participants were unaware the study drive was along a loop. The divided highway road network consisted of three lanes of traffic in each direction with speed limit of 65 mph. A solid paved shoulder extended beyond the roadway. The surrounding visual environment was a sparse suburban environment. Occasional ambient traffic was present in the participant's direction of travel, with ambient vehicle speed varying such that participants occasionally passed and were passed by other vehicles. All participants had the option to engage standard cruise control via buttons on the steering wheel.

The simulated highway loop contained four identical rest areas, indicated by the curved sections in Figure 2. Rest areas were previewed by signs approximately one minute in advance (when traveling at 65 mph). The rest areas were spaced approximately 20 miles apart. The study drive began at the north interchange and participants completed the drive in a counterclockwise direction.

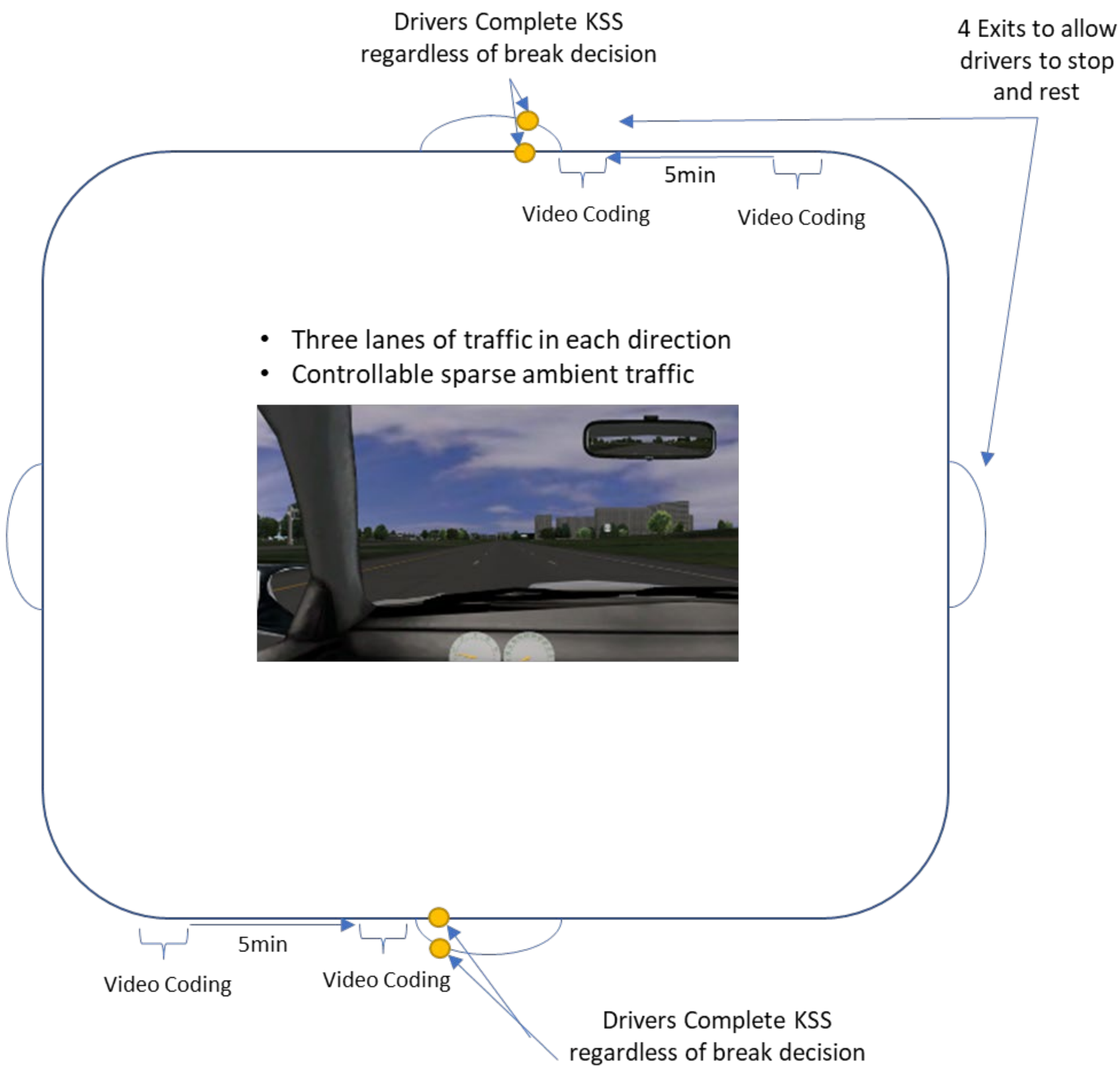

Figure 2. Study Database and Simulator Screenshot

Participants had a maximum of 3 hours to complete the 150-mile route, which would take slightly less than 2.5 hours at 65 mph without stopping, and were told that the goal was to drive 150 miles in under 3 hours. The drive was discontinued if participants reached 150 miles or the 3-hour time limit. Speed was managed by giving participants reminders about the speed limit when their speed deviated from the posted speed limit by more than ±5 mph. Road departures were defined as any time more than half the participant's vehicle departed the road surface onto the shoulder. (Unintentional drifts into other lanes were not categorized as road departures.) Crashes were defined as any collision with another vehicle or object in the roadway. During the drive, participants were free to decide if and when to take breaks.

A system of monetary incentives, adapted from a procedure developed by Gaspar and colleagues for a prior study (under review), was used to replicate the motivational tradeoffs present in a drowsy-driving situation (Table 1). Prior to the study drive, the research team informed participants they would start with a base compensation of $50 for their participation in the study. Participants were informed they would earn bonus compensation for reaching the destination in less than 3 hours, prorated at $1 per minute under the time limit, up to a maximum bonus of $50. However, they were also told that they would lose the entire amount and receive no compensation if they departed the road or collided with another vehicle. Participants were informed that the drive would be discontinued if a crash occurred. The incentive was designed to encourage participants to consider a tradeoff between (a) continuing to drive and earning more compensation, but risking safety (and losing the compensation), versus (b) stopping to rest to reduce risk of crashing but giving up at least some of the possible time bonus. In practice, all participants were paid $100 irrespective of time taken to complete the drive or whether they crashed (note: no participant actually crashed or departed the road during the study). The research team debriefed participants regarding the purpose of this deception at study completion.

Table 1. Incentive Structure

Starting Compensation	Rewards	Penalties
$50	+$1/minute under 3 hours, up to $50 max	–$50 for road departure or crash

*Note: all participants actually received $100 for completing the study.

During the drives, participants had the option to use the radio in the vehicle to listen to music of their choosing. Participants could also bring beverages (e.g., coffee, soda) into the simulator once they had taken a break and then resumed driving. Participants had the option to stop to rest at any of the four rest areas in the loop (Figure 2). Participants were told during training they would have the option to stop at rest areas, but not given information about their location or spacing. During breaks, participants had the option to remain in the vehicle or exit the simulator. If they exited the simulator, they had access to a private room that included a comfortable chair. They could choose to take a nap, use the restroom, walk around the hall, step outside, and consume coffee (8oz), other caffeinated beverage, or other foods from the vending machine. There was no cap on the length of a break. A researcher recorded the length of each break and all activities engaged in. If participants cited symptoms of simulator sickness as a reason for taking a break, a researcher followed up with them and discontinued the session if symptoms were too severe. The participant could also decide to bring food or drink back into the vehicle to consume during the remaining drive, and these activities were recorded by a researcher.

At two points throughout the interstate loop, participants were asked to rate their drowsiness using the KSS on a tablet computer positioned as the vehicle infotainment screen. The KSS asks participants to rate their drowsiness on a scale from 1 (extremely alert) to 9 (very sleepy), as shown in Table 2. The location of these KSS probes coincided with the north and south rest areas, as shown in Figure 2. Participants completed the KSS at these locations regardless of whether they stopped for a break or continued driving. These probes were located based on the assumption that participants had just made a decision about whether to stop for a break or continue driving.

Table 2. Karolinska Sleepiness Scale (KSS)

Score	Description
1	extremely alert
2	very alert
3	alert
4	rather alert
5	neither alert nor sleepy
6	some signs of sleepiness
7	sleepy, but no effort to keep awake
8	sleepy, some effort to keep awake
9	very sleepy, great effort to keep awake, fighting sleep

Procedure

Prior to the study visit, participants completed a phone screening, provided informed consent, and were scheduled for a study visit. Participants were asked to complete a food and activity log to confirm that they were awake 18 hours before the start of their study drive. Participants were also asked to refrain from napping or consuming caffeine within the final 13 hours before their study drive. Two participants completed the study each night. The first participant each night was scheduled to begin their drive at 11 p.m. and the second participant was scheduled to begin at 2:30 a.m.

Participants arranged their own transportation to the facility and were told they could not transport themselves home after the session. Upon arrival, researchers reviewed the activity logs. Participants then completed a training presentation instructing them about the driving task, options for taking a break, and the incentive. Before entering the simulator, participants completed the KSS. The drive lasted until participants traveled 150 miles, reached the 3-hour time limit, or withdrew from the study. Following the drive, participants completed a post-drive questionnaire about their decision making and perceived level of drowsiness throughout the drive and were debriefed about the incentive. The research team provided transportation home following the study visit.

Measures

Video Coding

To provide an objective measure of drowsiness, eyelid closures were coded frame-by-frame for the four rest area sections where KSS was recorded. Figure 3 shows the drive layout for each of the two loops. Participants encountered 7 rest areas throughout the drive. Video coding and KSS scores were collected for rest areas 1, 3, 5, and 7, highlighted in Figure 3. For each of these four rest areas, eyelid closures were coded for two 1-minute windows. The first window was 5 minutes before the rest area. The second window was 1 minute before the rest area.

Video coding was performed using Boris open-source software (Friard & Gamba, 2016) using video from a driver-facing camera (30Hz). Gaze was coded as either eyes open or eyes closed when the participant's eyes were visibly closed. The start and stop time of each eye opening and closure were coded and used to compute durations for each individual eyelid closure. Video data were single-coded by a trained team of research assistants. Three trained raters completed the video coding, with spot checks performed by another member of the research team.

Using these coded eyelid closures, the research team computed a modified measure of percent eyelid closure score (PERCLOS; Wierwille et al., 1994). The standard PERCLOS measures the percentage of time for which the participant's eyelids are more than 80% closed and is usually coded over a period of several minutes. The modified PERCLOS employed in the current study instead computed the percentage of the 1-minute window in which the participant's eyes were completely closed. Previous research has reported that PERCLOS coded over a 1-minute window demonstrated good agreement with PERCLOS coded over longer windows (Owens et al., 2018). Full eyelid closures were selected for coding based on accuracy of coding and to increase overall coding efficiency. However, this modified PERCLOS necessarily produces lower values than PERCLOS as defined by Wierwille et al. (1994), as time when eyes are completely closed is a subset of time when eyes are more than 80% closed, meaning that any given value of this modified PERCLOS would correspond to a higher level of drowsiness than the same level of the standard PERCLOS measure using the 80% threshold for eyelid closure.

Driving Data

The main measure of driving performance was standard deviation in lane position (SDLP). SDLP provides a measure of lateral vehicle control, with larger values representing poorer lane-keeping performance. Previous studies have shown degradation in lane keeping because of drowsiness (e.g., Gaspar et al., 2017).

Driving data were reduced using custom Matlab scripts. The driving data analysis windows were the same one-minute windows preceding the rest areas over which eyelid closures were coded (Figure 3). Driving data were merged with gaze data using custom R scripts (R Core Team, 2016).

Break Measures and Non-Driving Behaviors

At each break, participants recorded the reason for taking a break and completed a KSS rating. Note that while video coding was only performed for rest areas 1, 3, 5, and 7, breaks could occur at any rest area during the drive and KSS was collected for each break. During breaks, a researcher recorded the activities a participant engaged in (e.g., consuming coffee). Break durations were determined based on the stop and start times from the simulator data.

Participants were also able to engage in non-driving behaviors while driving. Potential behaviors included consuming coffee or other beverages and using the radio. To record the frequency and timing of these behaviors, a second round of video coding was conducted by a research team member, who watched the entire drive and coded the start and end time of any behaviors (using Boris software). Behavior coding was also spot checked by another member of the research team.

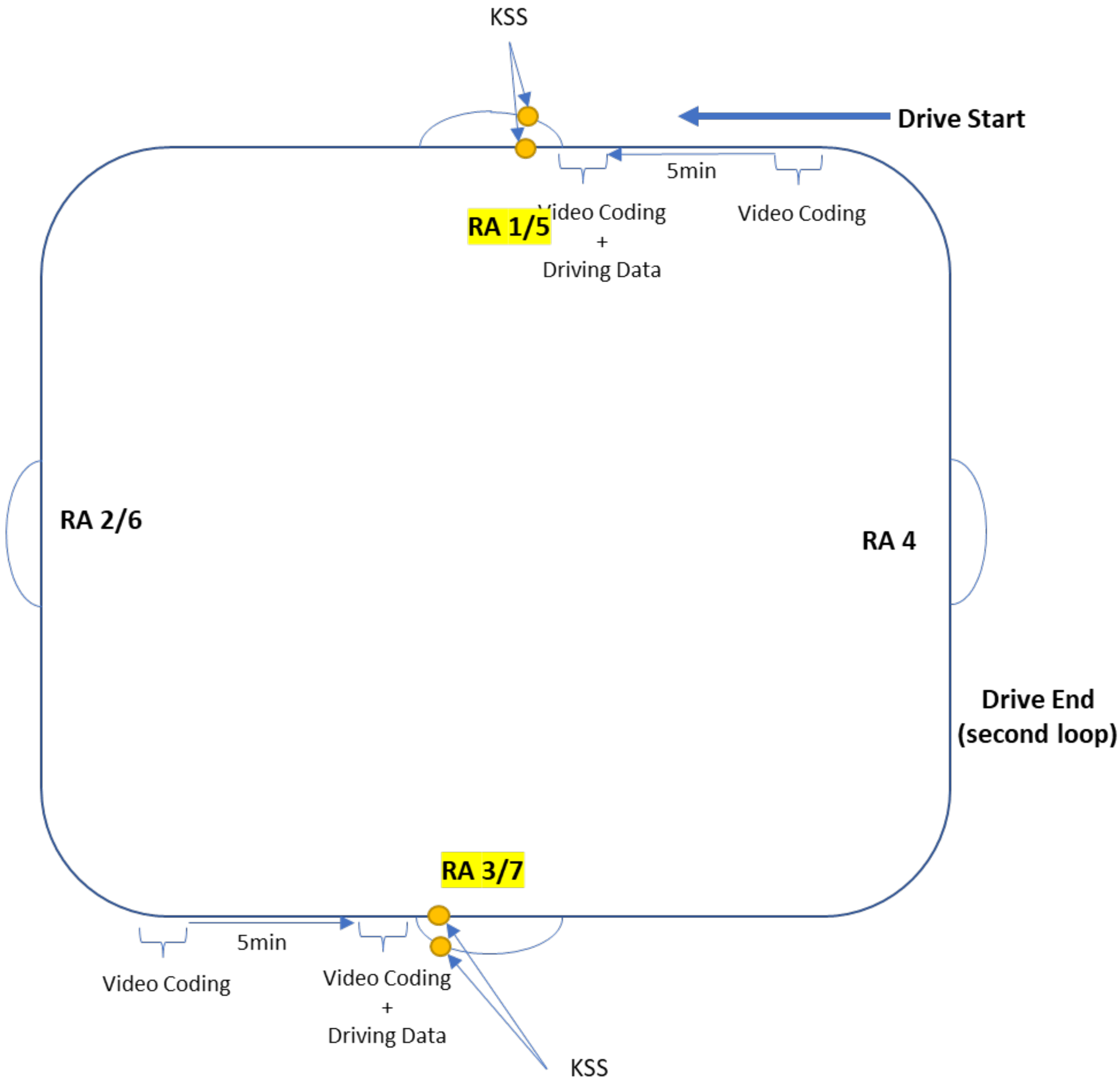

Figure 3. Drive Layout with Coded Rest Areas Highlighted

Results and Discussion

Participants

A total of 121 participants were enrolled in the study. Fourteen participants did not complete the drive due to simulator sickness and were excluded from the final sample. Seventeen additional participants did not complete the study drive session, either due to technical issues or withdrawing from the study prior to completion. The final sample included 90 drivers, aged 21–55. Table 3 shows demographic and pre-drive data from the sample. The final sample consisted of 50 females and 40 males.

Table 3. Sample Demographics and Pre-Drive Data

	Mean	SD	Min	Max
Age	31.39	10.29	21	55
Driving experience (years)	15.71	9.90	4	39
Annual miles driven	11,547	1601	2,000	15,000
Hours of sleep night before study session	7.65	1.24	4	10.5
Wake-up time day of study session	6:56 AM	NA	3:30 AM	8:00 AM
Pre-drive KSS	3.24	1.72	1	7

Drowsiness Measures

Evolution of Drowsiness Over the Drive

Measures of drowsiness were computed and compared at the four marked rest areas throughout the drive (Figure 3), referred to as rest areas 1, 3, 5, and 7. Self-rated drowsiness was measured using the KSS. Events where no KSS response was recorded (13 total) were removed from the analyses described in this section. No participant was missing more than one KSS response. Objective drowsiness was assessed using modified PERCLOS calculated from recorded eyelid closures during the one-minute window preceding the rest area.

The first analysis focused on whether measures of drowsiness changed over the course of the drive. Repeated measures ANOVAs were used to compare KSS ratings and PERCLOS measures across the four rest areas, which are used as a proxy for driving time. Paired comparisons were used to evaluate the difference between rest areas, representing the difference in drowsiness measures at different points in the drive.

KSS scores increased over the course of the drive (Figure 4). There was a significant main effect of rest area on KSS ratings ($F(3,321) = 8.10$, $p < 0.001$). Paired comparisons showed that rest KSS ratings in rest areas 5 and 7 were higher than rest area 1 or 3 ($p < 0.01$). The difference in ratings between rest areas 5 and 7 was not statistically significant ($p = 0.29$). KSS scores in rest area 3 were higher than rest area 1 ($p < 0.001$) and lower than rest area 5 ($p < 0.01$). These results suggest that KSS ratings generally increased over time, with the highest scores in the last two rest areas. These rest areas represent the last ~40 minutes of the three-hour drive. The lack of difference between rest areas 5 and 7 may evidence a ceiling effect of the KSS scale.

One effect worth noting in these data is the timeline for the onset of changes to drowsiness. Within the first 40 minutes of the drive, there was a significant increase in perceived drowsiness. Then, within the next 40 minutes, perceived drowsiness appears to have plateaued, at least based on KSS ratings. This indicates that self-perception of drowsiness can increase quickly during the course of a long drive, and that it does not take long for drivers to reach the highest levels of perceived drowsiness.

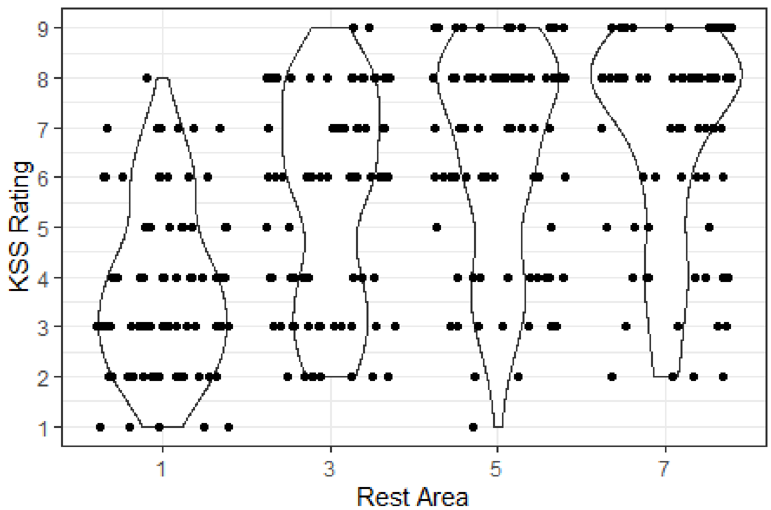

Figure 4. KSS ratings by rest area section.

Modified PERCLOS was calculated for the one-minute segments preceding each rest area, based on coding of eyelid closures. Higher PERCLOS values represent higher levels of drowsiness. PERCLOS was analyzed by comparing mean PERCLOS at each rest area location. A repeated measures ANOVA showed a significant main effect of rest area location on PERCLOS ($F(3,321) = 3.12$, $p = 0.03$). Pairwise comparisons showed that PERCLOS was higher for rest area 7 than for rest area 1 ($p < 0.001$). PERCLOS in rest area 7 was not significantly different than PERCLOS in rest area 3 ($p > 0.50$) or rest area 5 ($p > 0.48$). The difference between rest areas 3 and 5 was not significant ($p = 0.59$). These results indicate that, along with KSS ratings, objective measures of drowsiness, as defined by measures of long eyelid closures, also increased over the course of the drive, particularly between the first segment of the drive at later segments (i.e., after approximately 40 minutes).

To show the trends for other gaze-based measures, Figure 5 also presents other measures of eyelid closure, including the number of long eyelid closures (i.e., closures > 0.5 seconds), mean eyelid closure duration, and maximum eyelid closure duration.

Although the analysis intentionally focused on PERCLOS as the objective measure of drowsiness, it is worth noting a few results from inspection of the plots in Figure 5. First, it is clear that some participants experienced significant drowsiness events, likely in the form of microsleeps. At least three participants had maximum eyelid closures greater than three seconds long, occurring during the coding windows preceding rest areas 5 and 7. Furthermore, several participants had over ten long eyelid closures within the 60-second coding window. Importantly, these clusters of several long eyelid closures appeared as early as rest area 3, indicating the rapid onset of severe drowsiness (i.e., within the first 40 minutes of driving).

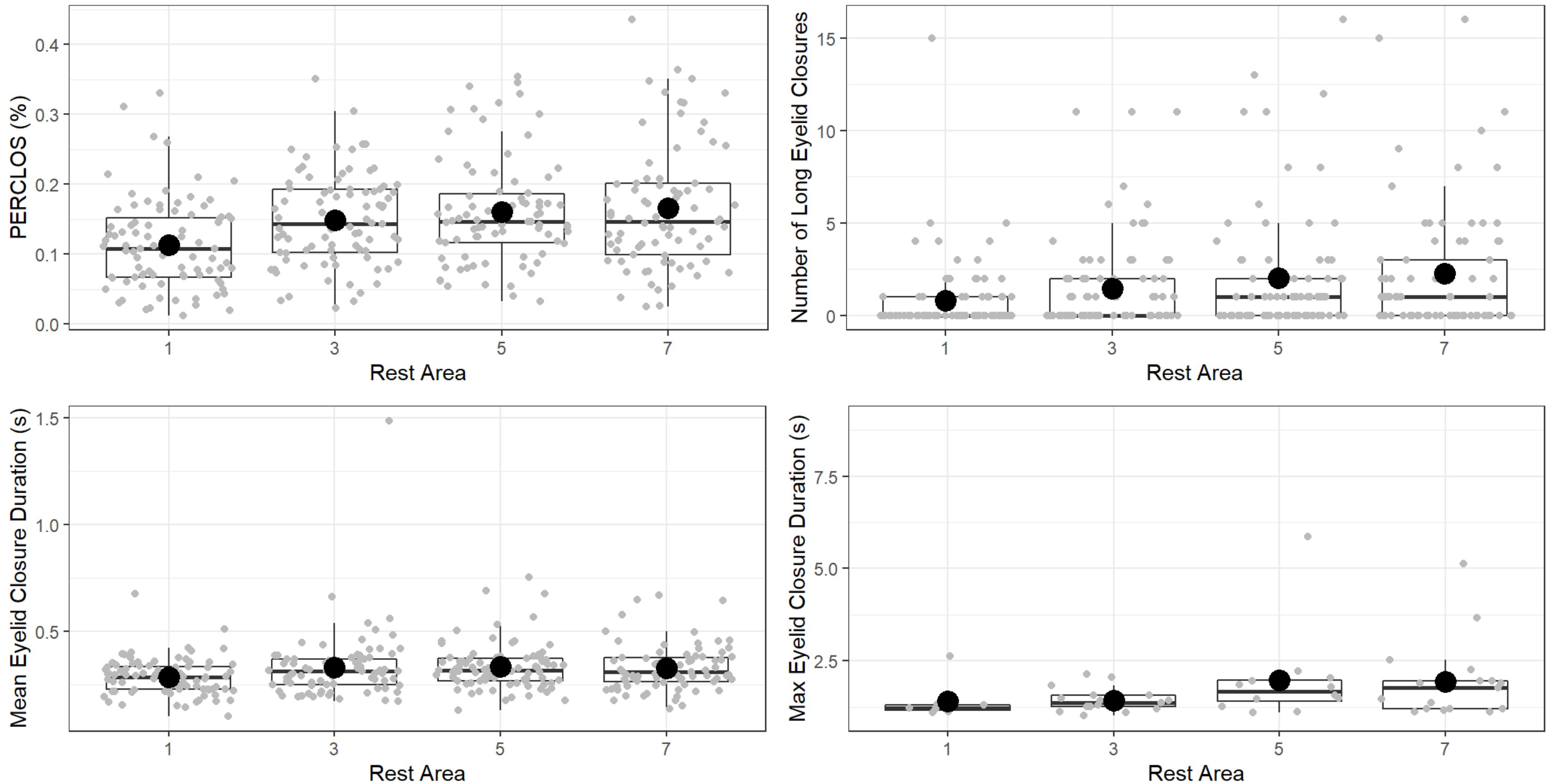

Figure 5. Gaze-based measures of drowsiness by rest area section

Relationship Between Drowsiness Measures

One of the main objectives of this project was to understand the relationship between the subjective feeling and objective measures of drowsiness. Previous research has suggested that one reason drivers may continue to drive versus stopping to rest when drowsy is a misunderstanding of how drowsy they are, leading them to overestimate their capability for safe driving and underestimate their risk of a crash.

The first step in this analysis was to examine the relationship between KSS ratings and PERCLOS, using the entire dataset from each of the four rest area locations. Linear mixed effects models with participant as a random effect and event as a fixed effect were used to examine the relationship between KSS and PERCLOS, and p values were obtained by likelihood ratio tests comparing the full mixed model to a partial model without the effect in question (Bates, Maechler, Bolker, & Walker, 2015). Figure 6 shows the relationship between KSS ratings and PERCLOS. KSS was related to PERCLOS ($\chi^2(1) = 7.77$, $p = .01$), indicating a statistically significant positive relationship between KSS and PERCLOS values.

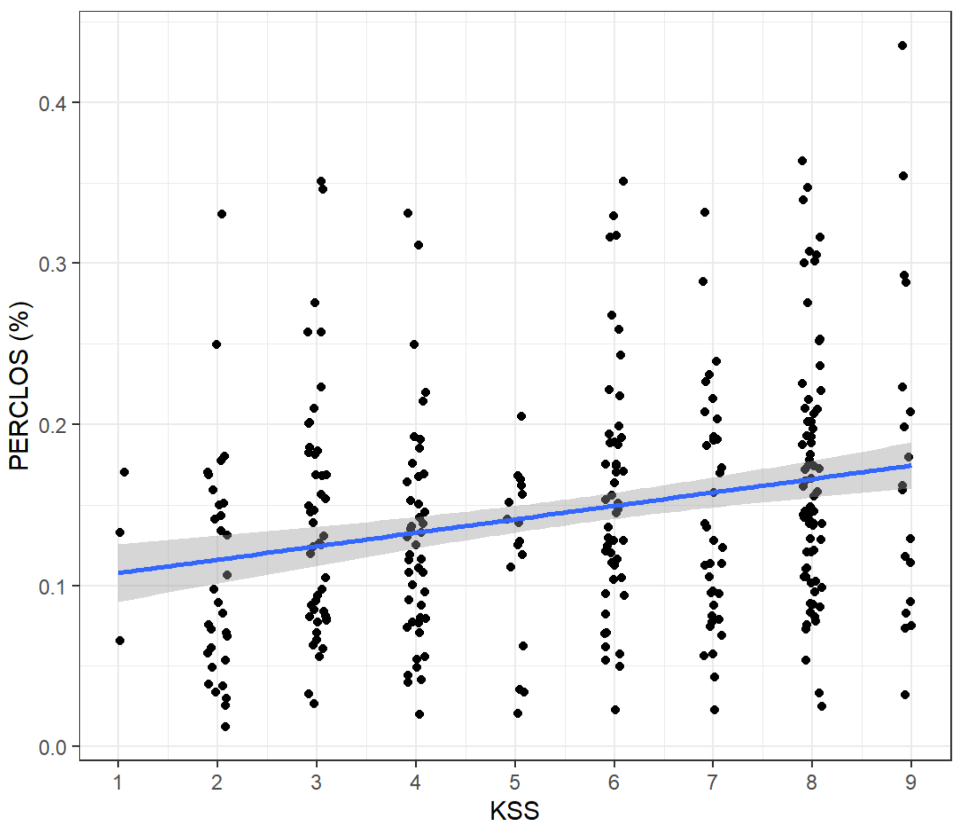

Figure 6. PERCLOS vs. KSS ratings

Beyond examining the overall correlation between subjective and objective drowsiness, the study also examined the extent to which participants were able to assess their own level of drowsiness at individual points throughout the drive. That is, are drivers more likely to underestimate their drowsiness than to overestimate it?

To perform this comparison, both KSS ratings and PERCLOS measures were categorized based on their perceived mapping to different levels of drowsiness (Table 4). For KSS, levels 1–4 were categorized as 1 or "low drowsiness" because they indicated some degree of perceived wakefulness or alertness; levels 5–7 were categorized as 2 or "moderate drowsiness;" and levels 8–9 were categorized as 3 or "high drowsiness."

A similar categorization was performed for PERCLOS measures. This categorization was based on the categorization of PERCLOS by Hanowski and colleagues (2008). In previous work, PERCLOS less than or equal to 0.125 was categorized as 1 or "low drowsiness;" values between 0.125 and 0.25 were categorized as 2 or "moderate drowsiness;" and values above 0.25 were categorized as 3 or "high drowsiness" (Table 4). Note that the current study used calculated a modified PERCLOS score based on full eyelid closures rather than the >80% eyelid closure as implemented in previous work (Hanowski et al., 2008), thus the thresholds used in previous work would imply somewhat greater levels of drowsiness in the current study than they did in previous work (i.e., if the participant's eyes were fully closed for 25% of the time in the current study, they were necessarily >80% closed for more than 25% of the time).

Table 4. KSS and PERCLOS Categories

KSS Score Range	KSS Category
1–4	Low Drowsiness
5–7	Moderate Drowsiness
8–9	High Drowsiness
PERCLOS Range	**PERCLOS Category**
≤0.125	Low Drowsiness
0.125–0.250	Moderate Drowsiness
>0.25	High Drowsiness

Using this categorization, the agreement between drowsiness self-assessments and objectively measured PERCLOS scores was examined for each individual rating by comparing the KSS category (low, moderate, high) with the PERCLOS category for each event. Figure 7 visualizes the resulting confusion matrix (Table 5), color-coded for the accuracy of the drowsiness estimate.

Overall, participants were most accurate at classifying "moderate" levels of drowsiness, with 65% of self-ratings of moderate drowsiness also categorized as moderately drowsy based on PERCLOS. Drivers tended to overestimate moderate to high levels of drowsiness. In 75% of the windows where participants rated their drowsiness as low, drowsiness assessed using PERCLOS was actually moderate or high. Similarly, participants also appear to have overestimated low to moderate levels of drowsiness with 70% of self-ratings of high drowsiness actually categorized by PERCLOS as low or moderate drowsiness. These results suggest that KSS ratings were more congruent with objective levels of drowsiness at moderate levels of drowsiness than at either low or high levels of drowsiness.

Taken together with the previous results examining relationships between subjective and objective measures in aggregate, these results suggest that at the event level, participants were somewhat poorly calibrated in their evaluations of drowsiness, particular at high and low levels based on PERCLOS. This may have important implications for decision making in the context of drowsy driving. It is also worth noting that even at moderate drowsiness, where subjective ratings were most accurate, their subjective self-assessments disagreed with the more objective PERCLOS-based measure of drowsiness 35% of the time. Overall, these results suggest that drivers should be cautious about estimating their level of drowsiness.

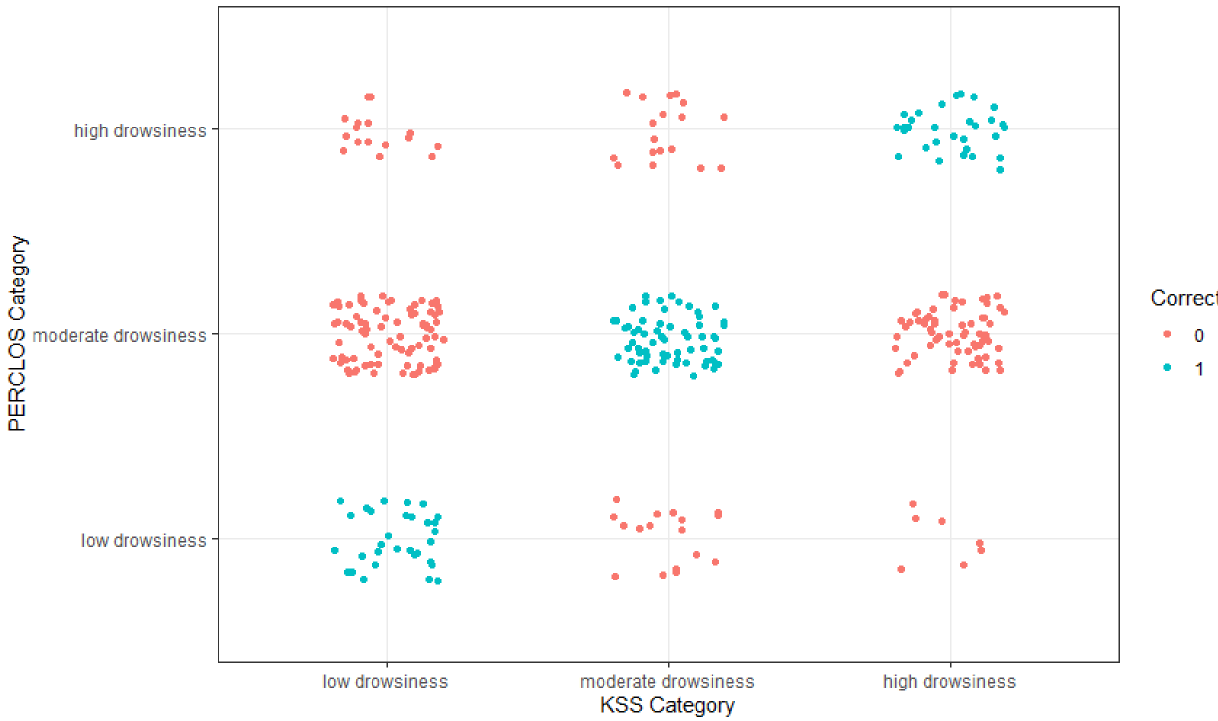

Figure 7. PERCLOS categorization vs. KSS categorization (Points represent individual ratings)

Table 5. Confusion matrix for drowsiness measures

PERCLOS Category	KSS Category		
	High 101 (25%)	Moderate 100 (24%)	Low 128 (31%)
High	30 (30%)	18 (18%)	16 (13%)
Moderate	64 (63%)	65 (65%)	80 (63%)
Low	7 (7%)	17 (17%)	32 (25%)

Note: percentages represent % of self-ratings.

Drowsiness and Driving Performance

This analysis examined how driving performance changed over time and the relationship between driving performance and measures of drowsiness. Figure 8 shows SDLP for the four rest area segments. There was a main effect of segment ($F(3,321) = 16.21$, $p < 0.001$).

Paired comparisons showed that lane keeping performance differed between all four rest area segments ($p < 0.02$). Performance was worst in rest area 7, followed by rest area 3, rest area 5, and rest area 1, respectively. Interestingly, performance was worse in rest area 3 than in rest area 5. As expected, lateral control was best in the earliest segment of the drive (rest area 1). For reference, lanes in the simulator database were 12 feet wide. In many cases, SDLP doubled or tripled in rest area 7 compared to rest area 1, signifying significantly more swerving in the lane and a greater likelihood of unintentionally departing the lane.

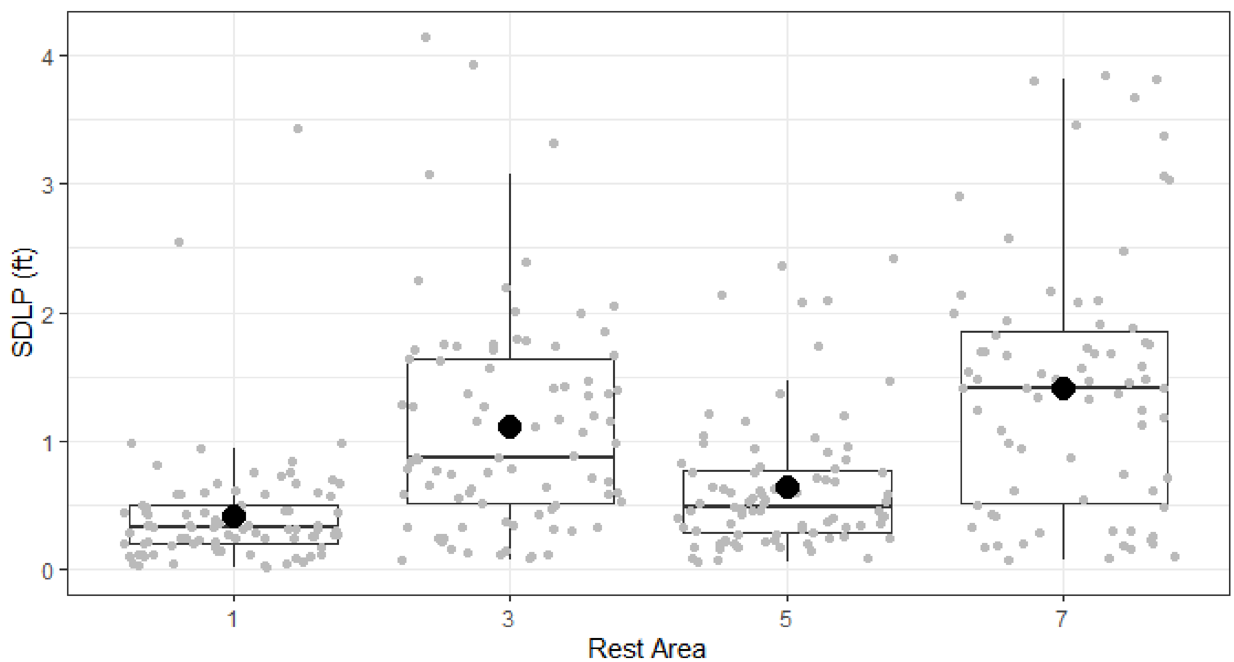

Figure 8. SDLP by rest area segment

Considering these results in combination with drowsiness measures across the event windows (Figure 5) suggests that some of the SDLP increase in rest area 3 compared to rest area 5 may have been due to some underlying physical feature(s) of one rest area compared to the other. Drowsiness was elevated at rest area 5 compared to rest area 3, but SDLP showed the opposite pattern between the two event windows. Although the rest areas were essentially copies of one another, it may be the case that some difference, such as differences in the ambient traffic present at the two locations, led to increased SDLP in event 3 relative to event 5.

Figure 9 shows the relationships between subjective drowsiness and SDLP (left) and between objective drowsiness and SLDP (right). Linear mixed effects models with participant as a random effect and event as a fixed effect were used to examine the relationship between SDLP and both KSS and PERCLOS, respectively. KSS showed a statistically significant relationship to SDLP, ($\chi^2(1) = 4.14$, $p = .04$). Interestingly, the relationship between SDLP and PERCLOS was not statistically significant, ($\chi^2(1) = 1.30$, $p = .25$). As perceived drowsiness increased, driving performance (in the form of lateral control) deteriorated. These findings suggest that drivers are able to perceive increased levels of drowsiness in association with deteriorating driving performance. However, measurable worsening in lateral control may not itself be a strong indicator of increasing objective drowsiness as measured using PERCLOS, at least within the range of drowsiness examined in the current study.

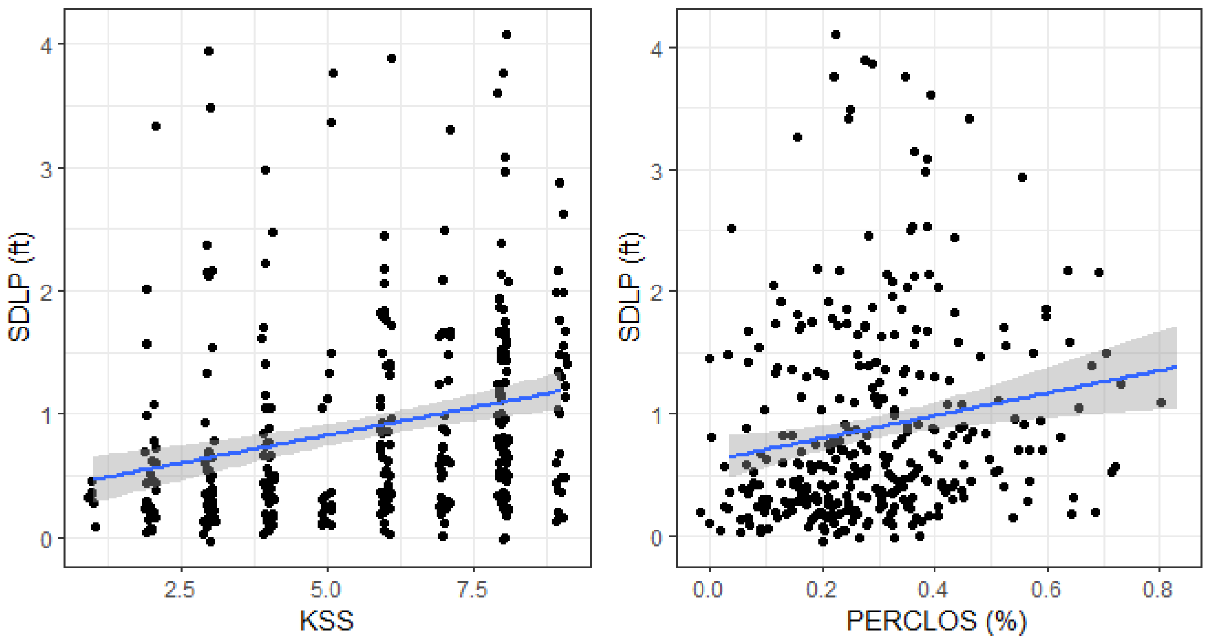

Figure 9. Drowsiness measures vs. SDLP

Break Taking

Break Frequency, Duration, and Activities

The second objective of the study was to understand break-taking behavior and what factors led to drivers stopping to take a break. Data for the following analyses were captured from all breaks, irrespective of whether the breaks fell within one of the video coding windows described above. Table 6 shows the number of participants who took 0, 1, or 2 breaks during the study drive. A total of 57 breaks occurred during the study, with 35 drivers taking one break and 11 drivers taking two breaks. Roughly half of the sample completed the drive without taking a break.

Table 6 also presents mean and range of break duration for both the first and second breaks. Breaks ranged from approximately 5 minutes in duration at the low end to slightly over 30 minutes. It is worth noting that the minimum break duration may reflect the minimum time necessary to complete study procedures such as the break survey and to restart the simulation. On average, the survey took 1.5 minutes to complete.

Table 6. Number and Duration of Breaks Taken

Total Breaks	Number of Participants	Duration of Break (in min) Mean	Min	Max
0	44 (49%)	NA	NA	NA
1	35 (39%)	7.98	5.04	31.10
2	11 (11%)	9.54	5.43	23.40

In addition to the frequency of breaks, the analysis examined when breaks occurred in the drive. Figure 10 visualizes where breaks occurred within the study drive. Table 7 shows corresponding descriptive statistics for the timing of break taking. The timing of the first break ranged from 22–135 minutes into the drive. On average, the first break occurred a little over an hour into the drive (77 minutes), although there was a cluster of breaks that occurred about 100 minutes into the drive. For drivers who took a second break, on average that break occurred 44 minutes after the first break and approximately 100 minutes into the total drive.

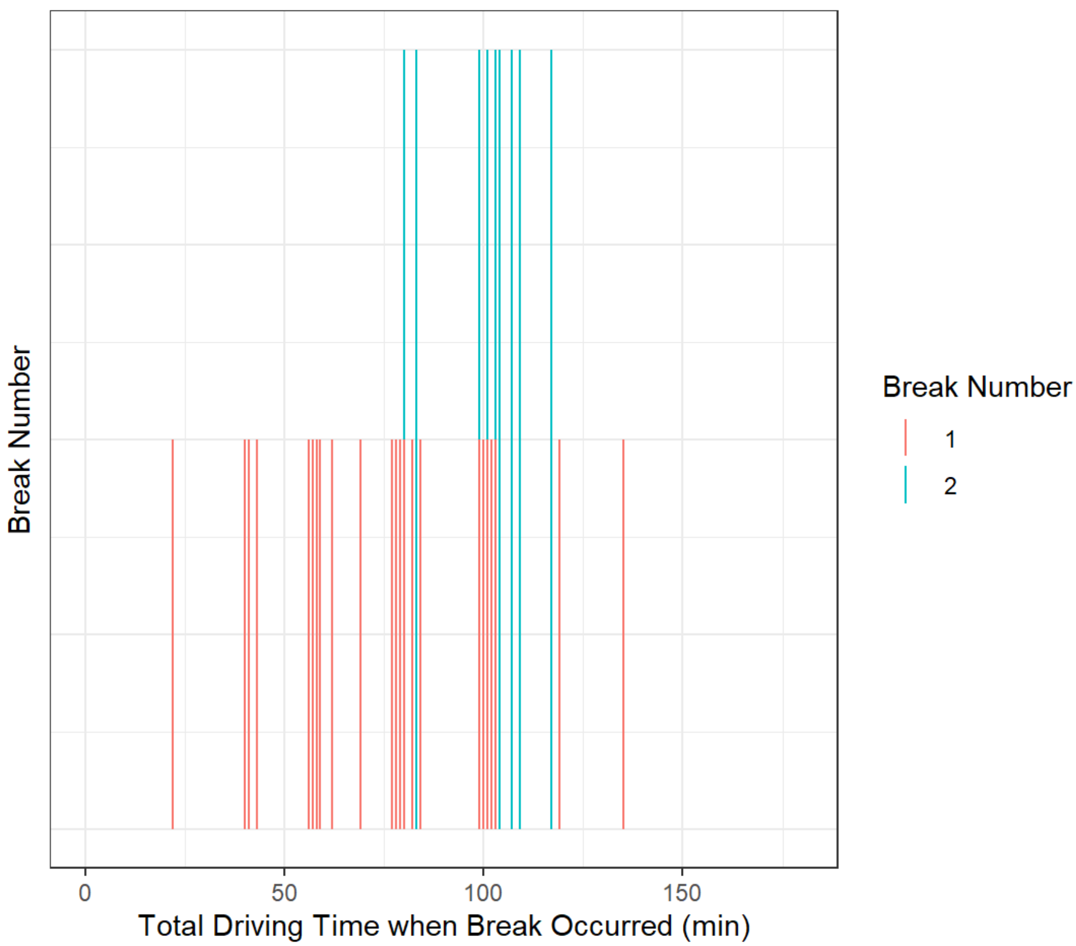

Figure 10. Break location in drive (all participants)

Table 7. Descriptive statistics for timing of breaks

Break Number	Time Into Drive when Break Occurred (in min)				
	Mean	Median	SD	Min	Max
1	73.27	77	25.30	22	135
2	98.73	103	12.35	80	117

Reasons for Breaks and Break Activities

At each break, researchers recorded both the participant's stated reason for the break and the activities performed during the break. Figure 11 shows the reasons recorded for breaks (left) and the activities recorded during the breaks (right).

Reasons for breaks were coded by categories and included the following:

- Drowsiness: descriptions such as tired, drowsy, sleepy
- Fatigue: reasons such as exhaustion, physical strain
- Driving Performance: descriptions included driving in lane, departing lane
- Boredom
- Eye Strain
- Needed Restroom

Break activities included consuming coffee, using their phone, resting, using the restroom, stretching or walking, changing clothes, eating, and applying lip balm.

Drowsiness was the most cited reason for taking a break. Nearly 40% of the breaks were prompted by feelings of drowsiness. Using the restroom was also a common reason for taking a break. Somewhat surprisingly, deterioration of driving performance was not frequently cited as a reason for taking a break, only noted as a reason for two breaks. The most common activities engaged in during breaks were using the restroom (~35% of breaks), resting (defined as taking a nap or sitting quietly in a comfortable chair without perform other activities, ~30% of breaks), and consuming a cup of coffee (17% of breaks).

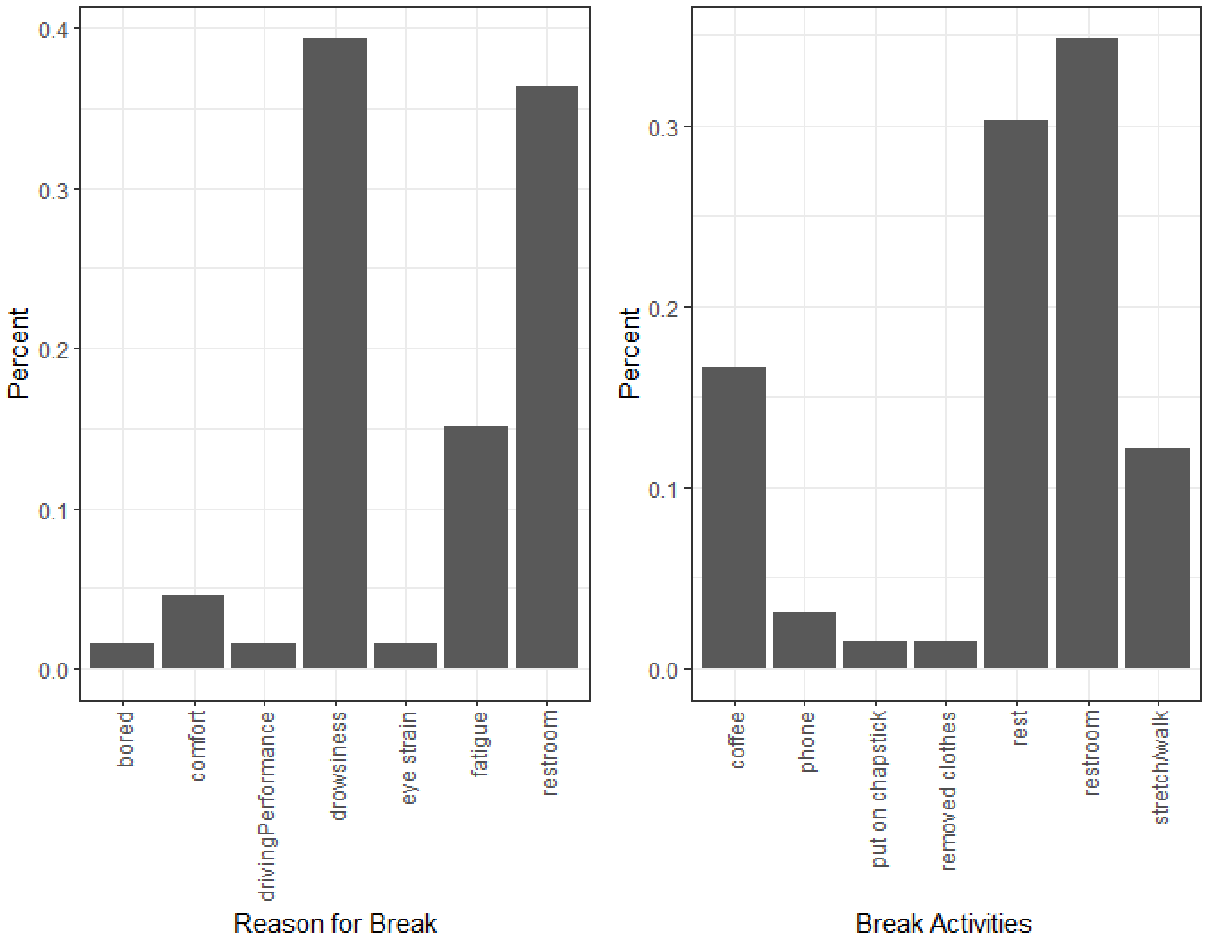

Figure 11. Reasons for breaks and proportion of breaks

Impact of Breaks on Driving Performance

To determine whether break taking had an impact on performance, we compared driving performance as measured by SDLP in the one-minute window preceding the break to four other windows around the break. One comparison window was five minutes before the break. The other three windows were post-break, at 5, 10, and 20 minutes, respectively. The 5-minute post-break window was selected to provide a buffer for participants to start the vehicle and merge back onto the interstate. Each of these windows was also 60 seconds in duration. We calculated the difference in SDLP between the one-minute window before the break with these four other windows, with lower scores representing a reduction in SDLP, indicative of better lane keeping performance.

Figure 12 shows the change scores for each of the comparison windows. Note that these windows were calculated only for instances where a break occurred. To show trends across participants, lines represent a unique break and its corresponding comparison windows. A change of 0 indicates there was no change in SDLP between the one-minute window before the break and a given comparison window.

A repeated-measures ANOVA was used to compare SDLP at each of the five windows, including the one-minute window leading up to the break. The main effect of time window was not significant, ($F(4,275) = 1.48$, $p = 0.21$). These results suggest that breaks had a minimal impact on driving performance, at least in the twenty minutes immediately following the break.

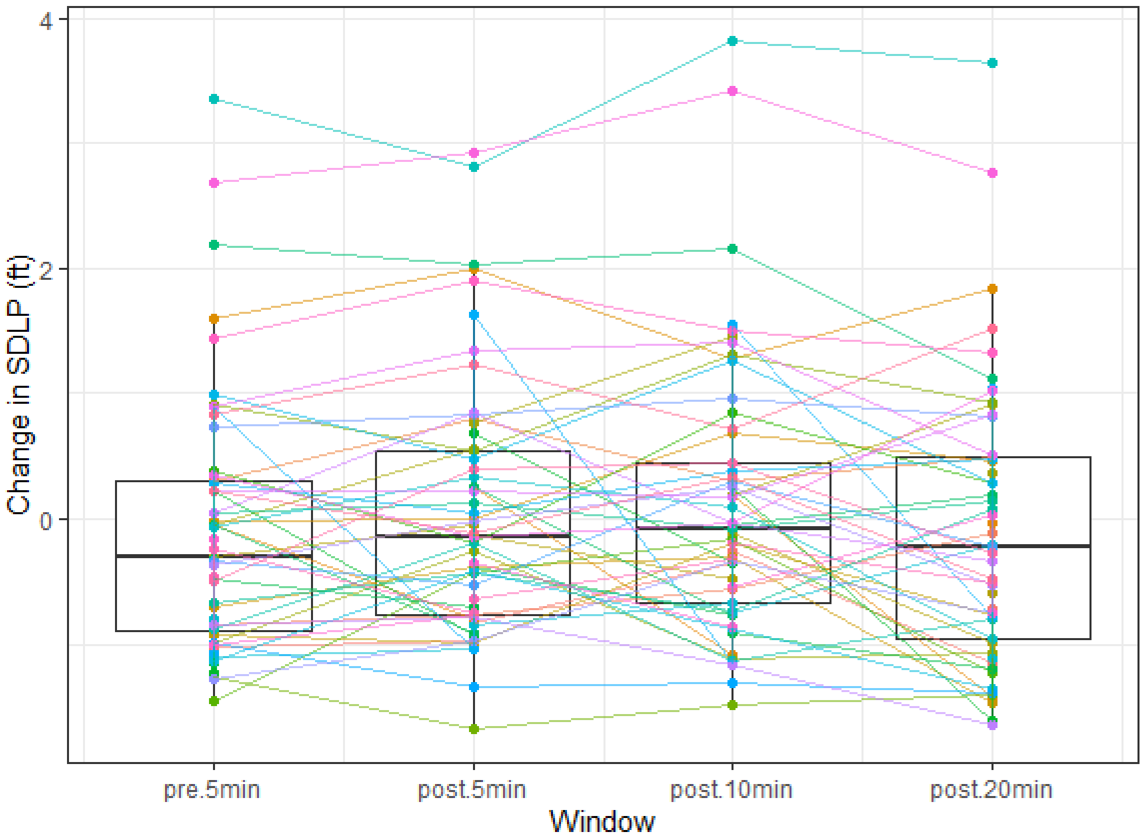

Figure 12. Change in SDLP compared to the 60-second window preceding break (Lines represent individual breaks)

Likelihood of Taking a Break

The second main focus of the study was an evaluation of the factors that contribute to the likelihood of taking a break during drowsy driving. This analysis used a subset of rest area events (i.e., windows) from the drive where glances were coded (i.e., rest areas 1, 3, 5, and 7). This constituted a total of 330 event windows. The events were coded based on whether a break was taken or not. Thirty (9%) of the windows contained a break and the remaining 300 (91%) did not include a break.

The analysis was focused on identify which factors predicted the likelihood of taking a break. Factors considered as predictors included the following:

- Event (i.e., rest area segment)
- PERCLOS in the 1-minute window before the rest area
- PERCLOS in the 1-minute window 5 minutes before the rest area
- SDLP in the 1-minute window before the rest area
- SDLP in a 1-minute window 5 minutes before the rest area
- KSS category: low, moderate, or high drowsiness
- Cumulative drive time: total time since the start of the study drive
- Drive time: time since starting to drive, either following a break or the start of the drive

Figure 13 shows a correlation matrix with the different break predictors and the break outcome, in the boxed region. Overall, there were not strong correlations between most of the predictor variables and the likelihood of taking a break, with the exception of a relationship between break taking and KSS self-ratings ($p = 0.01$).

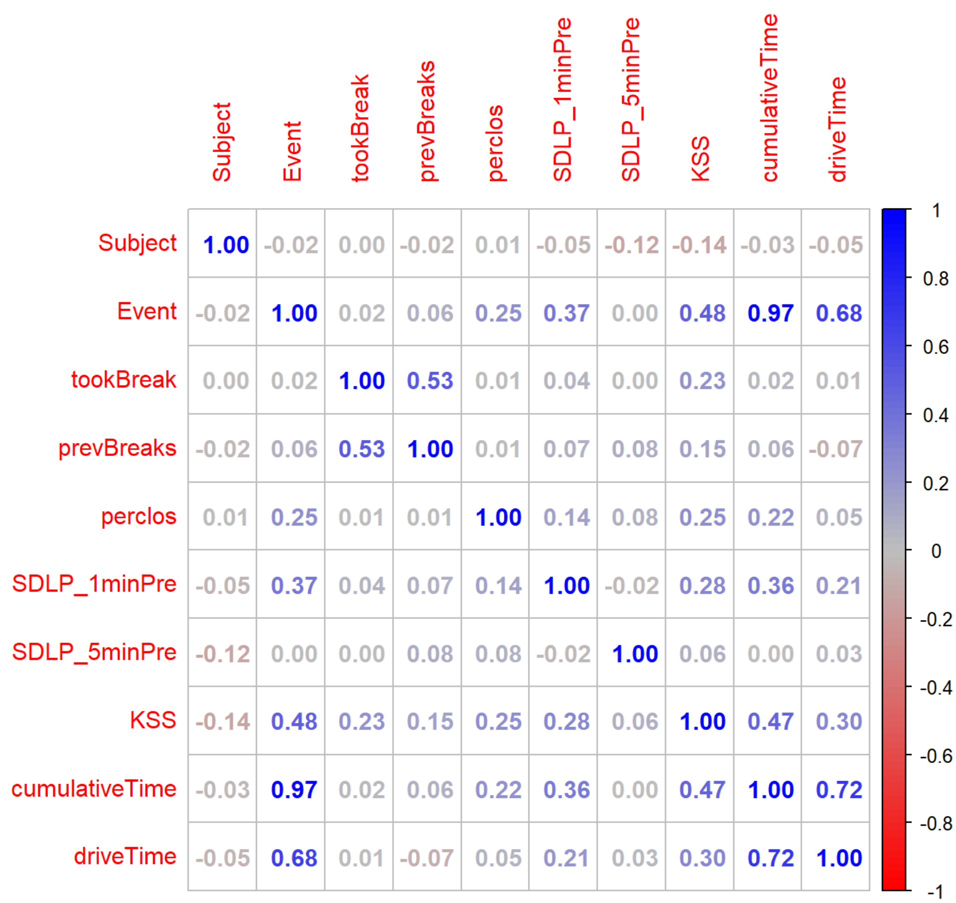

Figure 13. Correlation plot of break predictors and break outcome

A logistic regression was performed to determine the effects of the above predictor variables on likelihood of taking a break. The logistic regression model was statistically significant ($p < 0.001$, McFadden $R^2 = 0.14$). KSS score was the only predictor variable significantly associated with taking a break: a one-unit increase in KSS was associated with an 84% increase in the odds of taking a break ($p < 0.001$, OR = 1.84, CI = [1.40, 2.56]). No other predictor variables, including PERCLOS and SDLP measures, were significantly associated with the likelihood of taking a break.

This finding suggests that, at least in the context of the incentive structure in this study, drowsy drivers rely on self-perceptions of their drowsiness when deciding whether to take a break or not. Measures of lateral control, both just before a break opportunity and five minutes earlier, appear to have played a limited role in drivers' decisions regarding whether to take a break. Somewhat surprisingly, drive time did not have a significant impact on predicting whether drivers would stop to take a break, despite the finding that drowsiness increased over the course of the drive.

The shift in decision making with respect to self-evaluations of drowsiness can be visualized by plotting break likelihood by KSS rating. Figure 14 shows a histogram of break taking by KSS rating, color coded to represent the distribution of KSS ratings where a break did or did not occur. Table 8 shows the total number of rest area events alongside the number of breaks for each KSS rating. The likelihood of taking a break increased as KSS ratings increased. The data in Table 8 suggests there may be two transition points in the KSS scale in relation to break-taking behavior. One shift happens from KSS rating 5 to rating 6, with a 10% increase in the likelihood of a break for rating 6–8.

It is worth noting that most drivers chose to continue driving despite rating themselves as drowsy. Even at the highest KSS rating (9), more than 75% of the time participants decided to continue driving rather than taking a break. This is in line with the finding that KSS ratings were somewhat poorly calibrated with objective PERCLOS measurements. At high levels of drowsiness, participants tended to underestimate their own drowsiness. The finding that breaks were infrequent, even when drivers rated their own drowsiness as high, suggest that drowsy drivers may also overestimate their ability to continue driving despite impairment. Drivers appear to have discounted the potential penalty of losing their monetary incentive in favor of reaching the destination within the time deadline.

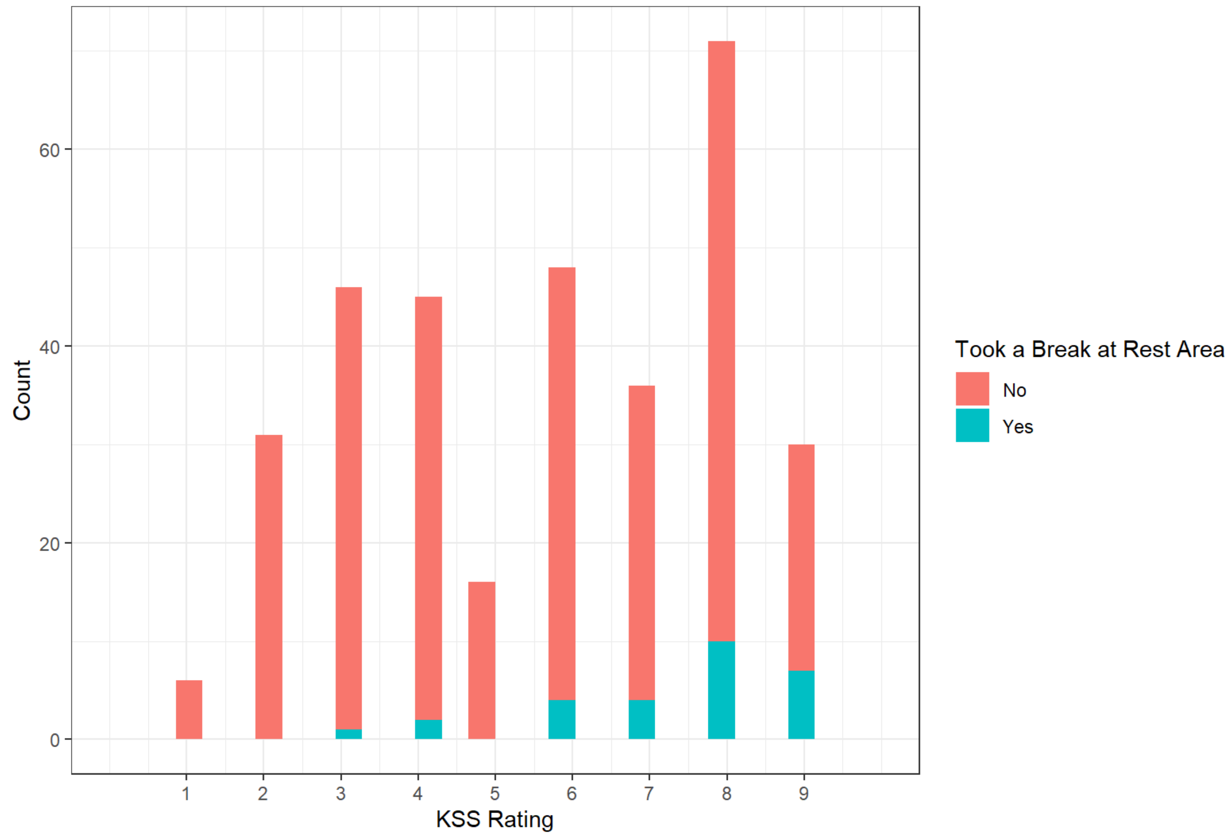

Figure 14. Histogram of breaks by KSS rating

Table 8. Rest area windows and breaks by KSS rating

KSS Rating	N	Number of Breaks	Percent of Events with Break	Median PERCLOS
1	6	0	0%	0.18
2	31	0	0%	0.16
3	46	1	2.2%	0.27
4	43	2	4.4%	0.23
5	16	0	0%	0.25
6	49	5	10.2%	0.30
7	36	4	11.1%	0.23
8	71	10	14.1%	0.31
9	30	7	23.3%	0.30

Conclusion

The objective of this study was to understand the relationship between different measures of drowsiness across a long drowsy-driving situation and to examine the factors that predict the likelihood that a drowsy driver will stop to take a break. The study employed a novel driving simulator experiment that combined several features necessary for answering these two research questions:

- A long, three-hour nighttime drive on an interstate loop
- An incentive method designed to mimic the motivational tradeoffs of drowsy driving. Specifically, the approach was intended to replicate the tradeoff between continuing to drive to reach a destination and stopping to preserve safety.
- Participants had the opportunity to stop to rest at rest areas, and to engage in countermeasures both within the drive and during breaks.
- Collection of objective and subjective drowsiness metrics over the same sampling window, which also corresponded to when drivers were making decisions about whether to stop or continue driving

The results of the study showed that both objective and subjective drowsiness increased over the drive. While there was an overall correlation between objective and subjective measures, comparison of individual ratings suggests that drivers often misclassified their level of drowsiness, particularly at both the low and high levels of drowsiness. PERCLOS is considered a “gold standard” for objectively assessing drowsiness in the context of driving. Therefore, it would be expected that a well-calibrated driver would be able to self-rate drowsiness in a manner that aligns with PERCLOS measures. These results suggest self-ratings are often poorly calibrated. Drowsy drivers overestimated low levels of drowsiness and underestimated high levels of drowsiness.

One conclusion from this finding is that messaging around drowsy driving could focus on the relatively poor correlation between self-ratings and objective drowsiness. Drivers could be made aware that they are poor judges of their own drowsiness, particularly as they become more drowsy. Other potential countermeasures might help calibrate drivers to their drowsy state through the integration of real-time feedback. Additional research is needed to understand the impact of a range of countermeasures on the relationship between subjective and objective drowsiness. Future research should also seek to understand whether benefits observed from particular countermeasures (e.g., in-vehicle alerts; Gaspar et al., 2017) may be the product of improved self-assessment calibration at certain levels of drowsiness.

The second research question centered on the factors that predict whether a drowsy driver will stop to rest. The analysis compared the effects of several factors including driving performance, objective drowsiness, subjective drowsiness, and time-on-task. Only subjective drowsiness had a significant effect on the likelihood that drivers would stop to rest. This finding aligns with the data showing somewhat poor calibration of self-ratings of drowsiness (e.g., Smith et al., 2016; Cochran et al., 2021). Drowsy drivers often misperceive their own state, in some cases underestimating high levels of drowsiness. Importantly, drivers’ self-assessments of their own drowsiness appear key to their decisions about whether to stop to rest. Nonetheless, even when drivers rated their level of drowsiness as very high (KSS scores of 9), they still bypassed rest areas without stopping on more than three of every four occasions. It is possible that on some occasions, drivers experiencing high levels of drowsiness may have failed to notice the rest areas.

One interesting finding is that driving performance was not a significant predictor of the likelihood of a break. Although the study did not ask drivers to rate their driving performance, performance declined over the course of the drive, suggesting that drivers might either fail to detect changes in driving performance or neglect to incorporate declining driving performance into their decision-making process about whether to take a break. It is also worth noting that driving performance was correlated with self-ratings of subjective drowsiness. If drivers decide to take breaks based on subjective drowsiness, it could be the case that this leaves limited variance left to be explained by driving performance. It is also important to point out that the penalty component of the incentive method was based on driving performance in that roadway departures would result in loss of compensation. Despite this, most drivers continued driving even when they rated their own drowsiness as high. This finding suggests that it may be difficult to shift drowsy drivers' decision making, particularly if that decision making relies heavily on self-assessments.

Several limitations from the current study are worth mentioning. First, the study used a driving simulator with an incentive method designed to replicate the motivational tradeoffs of drowsy driving. Although survey data from previous studies suggest that decision-making processes were similar between the experiment and similar on-road situations, more extensive effort should be devoted to validating the decision-making processes resulting from the incentive approach. As with any simulator study, the perceived cost associated with crashes is likely to be lower than in on-road environments, despite the fact that the incentive method heavily penalized errors in driving performance. Along similar lines, in the real world, drivers may encounter situations where options for taking a break are less appealing due to unfamiliarity or safety concerns, which were not present in the simulator study. However, it should be pointed out that this would likely only further serve to decrease the likelihood of taking a break in real-world driving conditions. The study also examined a subset of the drives corresponding to rest area windows. More research is needed to better understand when drivers begin to make decisions about stopping to take a break, and what other factors might contribute to those decisions that were not included in the present study. Factors such as familiarity with the driving route, ability to predict the specific location of upcoming rest areas, and options for different activities at different rest areas could all impact decision making and should be explored in future research.

This study provides insight into how drowsiness impacts decision making during long, nighttime drives. The results demonstrate a need for countermeasures to both improve self-assessment of drowsiness and to shift decision making in drowsy driving situations. Although much progress has been made to understand driver drowsiness, its contribution to crash risk, and potential methods for detection and mitigation, much work is still needed to understand how to shift decision making toward safer behaviors.

Acknowledgements

The authors would like to thank AAA Foundation for Traffic Safety for funding this research and for their collaboration in completing this research. The authors would also like to thank NORC, specifically Alycia Bayne, Neha Trivedi, and Madeline Liotta, for their collaboration and feedback on the study and report. Several individuals at the University of Iowa contributed to this research. The authors would like to thank Kasey Hagedorn for her leading role in helping organize and complete the study. The authors would like to thank Michael Neuhaus and Zachary Coleman, who served as overnight data collection specialists. Finally, the authors wish to acknowledge Mary Windsor, Katarzyna Wolniak-Bujakowska, and Caitlyn Nyugen for their support in video coding.

References

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, *52*(1–2), 29–37. https://doi.org/10.3109/00207459008994241

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bayne, A., Trivedi, N., Liotta, M., Siegfried, A., Gaspar, J. & Carney, C. (2022). Countermeasures to Reduce Drowsy Driving: Results of a Literature Review and Discussions with Experts (Technical Report). Washington, D.C.: AAA Foundation for Traffic Safety.

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution, 7*(11), 1325–1330. https://doi.org/10.1111/2041-210X.12584

Cochran, K. R., Letvak, S. A., McCoy, T. P., Bacon, C., & Karper, W. B. (2021). Can Nurses Self-Assess Fatigue?: Implications for Nurse Leaders. *JONA: The Journal of Nursing Administration, 51*(10), 507–512. https://doi.org/10.1097/NNA.0000000000001055

Gaspar, J.G., Schwarz, C.W., Marshall, D., Jenness, J., De Leonardis, D., & Blenner, J. A. (In Review). In-Vehicle Drowsy Driving Detection and Alerting: Final Report. National Highway Traffic Safety Administration.

Gaspar, J. G., Brown, T. L., Schwarz, C. W., Lee, J. D., Kang, J., & Higgins, J. S. (2017). Evaluating driver drowsiness countermeasures. *Traffic Injury Prevention, 18*(sup1), S58–S63. https://doi.org/10.1080/15389588.2017.1303140

Hanowski, R. J., Bowman, D., Alden, A., Wierwille, W. W., & Carroll, R. (2008). PERCLOS+: Development of a robust field measure of driver drowsiness. In *15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual Meeting*, New York, NY.

Knipling, R. R., & Wierwille, W. W. (1994). Vehicle-based drowsy driver detection: Current status and future prospects.

Owens, J.M., Dingus, T.A., Guo, F., Fang, Y., Perez, M., McClafferty, J. & Tefft, B.C. (2018). Prevalence of Drowsy Driving Crashes: Estimates from a Large-Scale Naturalistic Driving Study (Research Brief). Washington, D.C.: AAA Foundation for Traffic Safety.

R Core Team. (2016). R*: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available from https://www.R-project.org/

Reyner, L. A., & Horne, J. A. (1998). Falling asleep whilst driving: are drivers aware of prior sleepiness? *International Journal of Legal Medicine,* 111: 120–123. https://doi.org/10.1007/s004140050131

Shen, J., Barbera, J., & Shapiro, C. M. (2006). Distinguishing sleepiness and fatigue: focus on definition and measurement. *Sleep Medicine Reviews, 10*(1), 63-76. https://doi.org/10.1016/j.smrv.2005.05.004

Smith, B. P., Browne, M., Armstrong, T. A., & Ferguson, S. A. (2016). The accuracy of subjective measures for assessing fatigue related decrements in multi-stressor environments. *Safety Science, 86*, 238–244. https://doi.org/10.1016/j.ssci.2016.03.006

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.

Tefft, B. C. (2018). Acute sleep deprivation and culpable motor vehicle crash involvement. *Sleep, 41*(10): zsy144. https://doi.org/10.1093/sleep/zsy144

Tefft, B. C. (2014). *Prevalence of Motor Vehicle Crashes Involving Drowsy Drivers, United States, 2009–2013*. Washington, D.C.: AAA Foundation for Traffic Safety.

Tefft, B. C. (2012). Prevalence of motor vehicle crashes involving drowsy drivers, United States, 1999–2008. *Accident Analysis & Prevention, 45*(1): 180–186. https://doi.org/10.1016/j.aap.2011.05.028.

Wierwille, W.W., Ellsworth, L.A., Wreggit, S.S., et al. (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.