Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies

Jennifer F. May *, Carryl L. Baldwin

Department of Psychology, Old Dominion University, Norfolk, VA 23529, United States

ABSTRACT

Driver fatigue is an ill-defined term in the literature. It has been broadly used to refer to a wide range of driver states, each with different causal mechanisms. Technologies currently exist which enable detection of driver fatigue and interventions that have the potential to dramatically reduce crash probability. The successful implementation of these technologies depends on the cause and type of fatigue experienced. Sleep-related (SR) forms of driver fatigue result from accumulated sleep debt, prolonged wakefulness or troughs in the circadian rhythms. SR fatigue is resistant to most intervention strategies. Conversely, technologies for detecting and countering task-related (TR) fatigue (caused by mental overload or underload) are proving to be effective tools for improving transportation safety. Methods of detecting and counteracting the various forms of driver fatigue are discussed. Emphasis is placed on examining the effectiveness of existing and emerging technologies for combating TR forms of driver fatigue.

1. Introduction

Fatigue is a multidimensional construct that has been difficult for researchers to define (Brown, 1994; Desmond & Hancock, 2001). In fact, fatigue, sleepiness and drowsiness are often used synonymously (Johns, 1998). Fatigue and sleep are contributing factors in thousands of crashes, injuries and fatalities annually (Knippling & Wang, 1994; NHTSA, 2006). Several researchers have reviewed technological devices for the detection and countering of driver fatigue (Boivin, 2000; Hartley, Horberry, Mabbott, & Krueger, 2000; Mallis, 1999). To our knowledge, no attention has been given to differentiating technology by subcategorizing driver fatigue based on the causal factors of the fatigue. This distinction is important when evaluating the effectiveness of these technological devices. Driver fatigue can be subcategorized into sleep-related (SR) and task-related (TR) fatigue based on the causal factors contributing to the fatigued state. Sleep deprivation, extended duration of wakefulness and time of day (circadian rhythm effect) affect SR fatigue. Certain characteristics of driving, like task demand and duration, can produce TR fatigue in the absence of any sleep-related cause.

Certain detection technologies may be helpful in determining both types of fatigue, but technology aimed at increasing alertness may only be suitable for countering TR fatigue. Research has shown that among all the self-initiated strategies that drivers use (i.e. cranking up the AC, blasting the radio, rolling down the window), only naps and caffeine produce a reduction in driver fatigue, specifically in the context of SR fatigue (Horne & Reyner, 2001; Reyner & Horne, 1998, 2000, 2002). This review will focus on examining the effectiveness of driver fatigue technologies in the context of these different causal factors.

SR fatigue can be caused by circadian rhythms, sleep deprivation and sleep restriction. Sleep/wake patterns follow the body's natural circadian rhythm or internal clock, which drives humans to sleep during the night and be awake during...
Circadian effects have also been demonstrated during a driving simulator task (Lenne, Triggs, & Redman, 1997). Speed deviation (standard deviation of driver's average speed) varied significantly as a result of time of day, with the greatest variability occurring at 6:00 am, 2:00 pm and 2:00 am. SR fatigue is also influenced by homeostatic factors, such as the duration of wakefulness and sleep deprivation. Performance becomes worse the longer a person remains awake. Sleep restriction, or not obtaining adequate sleep will also result in increased sleepiness and a decline in performance.

In an effort to evaluate the progression of performance decline in response to various levels of sleep deprivation, Jewett, Dijk, Kronauer, and Dinges (1999) tested participants on the psychomotor vigilance task (PVT). The PVT is a reaction time task that requires participants press a button as soon as they see numbers appear on a display. Participants completed this task after 0, 2 (from 3 am to 5 am), 5 (from 1:30 am to 6:30 am) or 8 (from midnight to 8 am) hours of sleep. The PVT was performed at 10 am. All measures of the PVT, including number of lapses, slowest reaction time, fastest reaction time, time on task decrement and median reaction time improved as hours of sleep increased. Results from a 40-h sleep deprivation study showed that PVT performance decreased (lapses increased in frequency and reaction time increased) as the homeostatic pressure to sleep (or time awake) increased (Graw, Krauchi, Knoblauch, Wirz-Justice, & Cajochen, 2004).

TR fatigue is caused by the driving task and driving environment. Desmond and Hancock (2001), as well as Gimeno, Cer-ezuela, and Montanes (2006) suggest that driver fatigue can be produced by active or passive TR fatigue. Active fatigue is the most common form of TR fatigue that drivers experience (Desmond & Hancock, 2001). Gimeno et al. (2006) relate active fatigue to mental overload (high demand) driving conditions and passive fatigue with underload conditions. Examples of high task demand situations include high density traffic, poor visibility, or the need to complete an auxiliary or secondary task (i.e. searching for an address) in addition to the driving task. Passive fatigue is produced when a driver is mainly monitoring the driving environment over an extended period of time when most or the entire actual driving task is automated. Passive fatigue may occur when the driving task is predictable. Drivers may start to rely on mental schemas of the driving task which results in a reduction in effort exerted on the task (Gimeno et al., 2006). Underload is likely to occur when the roadway is monotonous and there is little traffic.

Most studies of driver fatigue focus on sleep deprivation or circadian rhythm effects, but require drivers to perform driving tasks during monotonous, highway conditions. This confounds the effects of SR and TR fatigue. Regardless, it is clear that driver fatigue does produce performance decrements in driver simulation and on-road driving tasks. Fig. 1 illustrates the three types of fatigue, their causes, consequences and interactions.

Lenne, Triggs, and Redman (1998) conducted a driving simulator study where participants completed a 20 min drive every 3 h between 8:00 am and 8:00 pm after either an 8 h night of sleep or a night of complete sleep deprivation. The driving simulator task also included a secondary reaction time task. Results showed that lane position variability was greater in the sleep deprivation condition and increased across each trial. Like their previous study, sleep deprived drivers drove closer to the centerline of their lane than non-sleep deprived drivers. The standard deviation of speed was greater for sleep deprived drivers, and was worse at 8:00 am and 2:00 pm, showing a circadian effect on performance. Mean reaction time was greater in the sleep deprived condition and improved throughout the day regardless of condition.

Philip et al. (2005) tested the effects of sleep restriction on a real highway. Participants slept either 8.5 h or 2 h in the laboratory, and then drove on a straight highway with light traffic and fair weather in a car outfitted with additional passenger controls. During the on-road portion of the experiment, a researcher accompanied the driver, prepared to take over the driving task if necessary. Driving sessions were 105 min long and occurred approximately 2 h apart with a total of 5 sessions. After each driving session, participants completed subjective sleepiness scales and a 10 min reaction time test.
ipants in the sleep restricted condition exhibited significantly more inappropriate line crossings, increasing the risk of such an action by 8 times that of rested drivers. Mean reaction time and subjective sleepiness were both greater for the sleep restricted participants.

Thiffault and Bergeron (2003) tested the effects of monotonous environments on driving performance during a simulated driving task and revealed poorer performance in the monotonous condition, as indicated by a greater number of large steering wheel movements (over corrections). Increased automation of the driving task (e.g., using cruise control) during prolonged driving and the low task demands of monotonous roadway conditions result in passive fatigue (Desmond & Hancock, 2001).

Although the same types of performance decrements may be seen for both SR and TR fatigue, these causal factors must be distinguished when considering technology designed to detect or counter driver fatigue. Devices for detecting driver fatigue may be acceptable for both types of fatigue, but countermeasures will only be effective for combating fatigue. In addition, it is critical to distinguish between active and passive fatigue. Adding an additional task for the driver to perform like a reaction time task, could improve performance during passive fatigue. There are, in fact, commercially available devices that function this way. However, an additional task would be detrimental to a driver experiencing TR active fatigue. The following section addresses driver fatigue technology. Devices are grouped according to those intended for fatigue detection, crash prevention and fatigue reduction. These devices are evaluated for their effectiveness in detecting or countering SR and TR fatigue.

2. Technological aids for driver fatigue

2.1. Detection and warning technology

Detection and warning technologies use measures that are sensitive to physiological and performance changes in fatigue, such as eye movements, head-nodding and steering performance. The goal of these devices is to warn the driver of their fatigue so that they can stop driving and rest. Additional research must determine the effectiveness of these devices in the context of different causal factors, such as SR or TR fatigue. However, as SR fatigue may more likely lead to drivers actually falling asleep, these devices may be best suited for identifying SR fatigue.

2.1.1. Eye closures

The PERCLOS system calculates the amount of eyelid closure over the pupil based on video monitoring of the eyes. It measures this in 1–3 min intervals and derives an index of fatigue. Specifically, the algorithm calculates the proportion of time within 1 min blocks that the eyelid covers 80% of the pupil (Dinges & Grace, 1998). This system has been validated in on-road driving studies (Wierwille, Lewin, & Fairbanks, 1996) as well as with the PVT (Dinges, Mallis, Maislin, & Powell, 1998). PERCLOS values correlated with lane departures and subjective sleepiness in sleep deprived drivers. PERCLOS values also correlated with PVT lapses such that participants with higher PERCLOS values exhibited more lapses. The PERCLOS system has been tested against EEG algorithms, eye blink software, and head-nodding technology and was found to be superior compared to the other techniques in detecting fatigue and was highly correlated with performance decrements in the PVT (Dinges et al., 1998; Mallis, 1999).

Kozak et al. (2005) measured eye closure, reaction time and subjective sleepiness during a 3 h driving simulator task in completely sleep deprived and fully rested participants. Sleep deprived participants had a higher frequency of lapses and longer reaction times, higher PERCLOS values and greater subjective sleepiness. These measures also increased over the drive.

This system is currently used within the driver fatigue monitor DD850 (Attention Technologies, http://www.attention-technology.com) used for fatigue monitoring in commercial truckers. The driver fatigue monitor gives the driver feedback with audible alerts and visually presents information on the duration of the eye closures and distance driven during that time.

2.1.2. Head-nodding technology

As a driver starts falling asleep, his/her head starts to nod as muscles relax (Hartley et al., 2000). The No-NAP (Global Info-tech Pune, n.d.) fits over the driver’s ear and monitors head position. When the head nods down, the driver receives a buzz in the ear to wake him or her up. This type of technology would be most efficient in detecting the onset of sleep, as the head-nod phenomenon appears to be a last cue of sleepiness before a micro-sleep or sleep period occurs (Hartley et al., 2000). Other head-nodding technologies include the Dozer’s Alarm (Haworth & Vulcan, 1991) and the micro-nod detection system (Hartley et al., 2000). The Dozer’s alarm was tested during a simulated driving task at night. Results showed that the alarm did not improve lane tracking performance and did not allow drivers to maintain alertness, but it did detect drowsiness (Haworth & Vulcan, 1991). Although the head-nod technology appears efficient for detecting the onset of sleep, drivers are likely unfit to drive safely before this point of sleepiness. As such, the head-nod technology warns drivers too late as to their sleepiness. In addition, there may be a high potential for false alarms as the driver looks around in his/her environment. False alarms cause user mistrust and annoyance in a system. This can lead a user not to use the technology or ignore true alarms which may result in a crash.
2.1.3. Deadman switch technology
Deadman switches are designed for the user to continuously press a switch. When the switch is released, it is assumed that the user has become impaired and alarms sound. The release of the deadman switch has been related to longer reaction times, and presence of stages 1 and 2 of sleep (Ogilvie, Wilkinson, & Allison, 1989). Unfortunately, like the head-nod technology, this type of system, if developed for vehicle use, would be a late-alert system which identifies sleepy drivers only when sleep has set in.

2.2. Crash prevention technology
Crash prevention technologies were designed to prevent crashes regardless of the causal factors. However, because crashes which result from running off the road are characteristic of fatigue-related crashes, rumble strips and lane departure warnings can reduce the crash risk for drivers experiencing both TR and SR fatigue. Collision avoidance warning systems can help fatigued drivers (SR and TR) by compensating for their deteriorated reaction times.

2.2.1. Roadway designs
Rumble strips are a roadway design option for monitoring weaving and alerting drivers when they drive off the road. Rumble strips have a grooved design within the pavement that produces a loud noise and vibrations within the car when a driver crosses or drives along the strip. The advantage of rumble strips is that they are available for all drivers. The two types of rumble strips constructed today are milled and rolled rumble strips.

Milled rumble strips are constructed on existing asphalt shoulders or new ones, and are concave circular depressions approximately 180 mm wide, 400 mm long and 13 mm deep (Perrillo, 1998). Milled rumble strips are preferred because they produce more noise and vibrations than the other types (Perrillo, 1998). Rolled rumble strips can only be installed in new or reconstructed shoulders while the asphalt is hot. They are constructed by creating rolls of excess asphalt above the flat surface of the shoulder and are narrower than the milled rumble strips and result in a shallower tire drop thus producing less noise and vibration than milled rumble strips (Perrillo, 1998). Center line rumble strips have resulted in a reduction of crashes along rural two-lane roads (Persaud, Retting, & Lyon, 2004). Rumble strips have significantly reduced the amount of run off the road crashes in New York (Perrillo, 1998).

2.2.2. Lane departure warning systems
Lane drifting or departure technology receives information from a video camera that monitors the road ahead and graphs the boundaries of the lane. An alarm sounds when a driver veers out of the lane. SafeTRAC (AssistWare Technology, 2005) is such a device. SafeTRAC warns drivers if they begin to drift out of their lane without using their turn signal. This device also calculates a continuous ‘score’ of performance, and if the score drops, SafeTRAC generates an alert recommending rest (AssistWare Technology, 2005). In this way, SafeTRAC not only detects lane departures, but lane position variability or weaving such that performance can be evaluated.

AutoVue is another lane departure warning system manufactured by Iteris (http://www.iteris.com/av/passenger.html) which incorporates the car’s speed, turn signal, and road markings in order to identify unintentional lane departures. If the car drifts out of the lane, a virtual rumble strip sound is emitted to warn the driver. This system is currently installed in the Infiniti M45 and Infiniti FX passenger vehicles. Both the SafeTRAC and the AutoVUE advertise that they have low false alarm rates, but neither company provide research to support this claim.

2.2.3. Collision avoidance system (CAS) warnings
CAS warnings are currently integrated in the production of some vehicles and are intended to alert drivers of potential rear-end or from the side crashes. This type of technology is useful in vigilance conditions where the ability to detect these critical events is reduced. CAS warnings can measure time to collision or temporal headway (Ben-Yaacov, Maltz, & Shinar, 2002). Temporal headway is defined as the time needed for a car to reach the position of the car ahead of it (Ben-Yaacov et al., 2002). Other sensors can detect when another vehicle or obstacle comes within a certain distance from the sides of the car. Lee, McGeehee, Brown, and Reyes (2002) demonstrated with a driving simulator task that in both distracted and non-distracted drivers, early collision avoidance warnings have been shown to reduce collisions by 80% as compared to drivers who did not receive a warning. CAS warnings in this study also reduced collision velocity and minimum time to collision, demonstrating a reduction in crash severity. The collision avoidance warning consisted on both an auditory tone and a visual icon displayed above the instrument panel. May, Baldwin, and Parasuraman (2006) found that crashes were reduced in drivers exhibiting task-related fatigue when presented with an auditory warning prior to a potential head-on crash in a driving simulator task.

2.3. Fatigue countermeasures
Technology designed to reduce fatigue or improve performance is suitable for drivers experiencing TR fatigue. Technological countermeasures include the use of automation in vehicles to reduce task load as well as interactive technology designed to increase engagement of the driver to prevent passive TR fatigue.
2.3.1. Automation

Automation refers to transferring control of a system or task (in this case, a component of driving such as speed control) from the operator to the vehicle. A recent study examining TR fatigue indicated that lane keeping performance was improved when drivers had cruise control activated (Funke, Matthews, Warm, Emo, & Fellner, 2005). Adaptive automation can link driver fatigue and task load in an effort to optimize driving performance (Hancock & Verwey, 1997). The purpose of adaptive automation is to modify the task load of the driver depending on the driver's state.

Hancock and Verwey (1997) described two adaptive automation systems. The generic intelligent driver support system (see Michon, 1993) utilizes a computerized scheduling system which analyzes the mental and visual workload of the driver, prioritizes information messages to send to the driver and decides how to present the information. If the driver is under high visual workload, the system may present messages verbally or by tactile sensation. Although this system does not specifically address fatigue, it does attempt to reduce driver workload. By reducing driver workload, active TR fatigue may be alleviated.

The SAVE system (Hancock & Verwey, 1997) integrates a monitoring unit, an automatic control device and a warning system. The monitoring unit detects fatigue, intoxication and unresponsive drivers. The automatic control device stops the car on the side of the road if the driver is not responding to warnings. The warning system informs the driver of his or her unsafe condition and also informs surrounding traffic and a traffic control center.

Active or adaptive cruise control is an advancement of cruise control which automatically adjusts the car's speed in response to the speed of the vehicle ahead of the driver (BMW World, n.d.; Marsden, McDonald, & Brackstone, 2001). This type of adaptive automation helps drivers in heavier traffic, but may not be beneficial on straight roadways with little traffic (Young & Stanton, 2004).

The S.A.M.G-3, steering attention monitor (Barton, 2003; Electronic Safety Products, n.d.) monitors normal corrective steering movements, defined by the micro-corrections that drivers must constantly make to maintain performance. An alarm sounds when normal steering movements cease. In a form of adaptive automation, there is an option to allow the system to turn off cruise control when steering wheel movements cease. Other steering monitors, such as the ZzzzAlert Driver Fatigue Warning System and the TravAlert early warning system also monitor corrective steering movements and produce audible alerts if the driver stops steering, but they do not attempt to modify the driver's task load or take over any driving systems (Hartley et al., 2000).

2.3.2. Interactive technology

Verwey and Zaidel (1999) believed that giving drivers challenging secondary tasks to perform when they became tired would increase their alertness. They used a game box called “Car Mate” that was an auditory and verbal interface in the car designed to allow the driver to play 12 different games. Results showed that driving with the game box produced fewer incidents and accidents, as well as a lesser decline in vehicle control. Participants reported less drowsiness and fewer episodes of sleep in the game box condition.

The Interactive Knight–Warrior Sleep Alarm (CARRS-Q, 2002) is activated by the driver, presumably when the driver feels fatigued. The system will periodically emit a minor auditory alarm at a specified time interval. The driver has 1–3 s to turn off the alarm; otherwise a secondary wake-up siren will sound in an effort to wake the driver up.

Of all of these technologies, the crash prevention devices may be the most feasible to implement in terms of incorporating them into vehicle construction or roadway design. Lane departure warnings and rumble strips are currently in use to reduce off the road crashes. Collision avoidance warning systems are also installed in some vehicles to assist in reducing rear-end and side crashes. All three types of technology are potentially beneficial for both SR and TR fatigue.

3. Conclusions

Classifying driver fatigue into SR and TR (active and passive) categories based on the causal factors involved will lead to improved development and implementation of fatigue countermeasures. TR active fatigue stemming from high task load driving condition requiring sustained attention and prolonged driving will benefit from increased automation and in-vehicle technologies that offset driver workload. Conversely, further automation of the driving task will exacerbate TR passive fatigue caused by monotonous conditions and highly familiar roadways. When a driver is susceptible to TR passive fatigue, increasing the novelty and demand of the driving task through interactive technologies such as the CAR MATE can be of benefit. SR fatigue, on the other hand, comes about from dips in the circadian rhythm (time of day effects), sleep deprivation and prolonged wakefulness and may therefore be rather impervious to either form of countermeasure. However, either form of TR fatigue can exacerbate the performance decrements associated with SR fatigue. SR and both forms of TR fatigue may be measured similarly. Each form of fatigue has been shown to cause decrements in driving performance and reductions in visual scanning patterns and eye movements. However, performance patterns associated with each form are frequently confounded in the existing literature by examining them in combination rather than isolation. Future research should further examine the potential performance differences between SR and TR fatigue forms in isolation as well as in combination. For example, SR fatigue may result in a different signature of impairment in isolation versus in combination with either active or passive TR fatigue. If such a fatigue profile can be reliably detected, then adaptive forms of fatigue countermeasures might be implanted that target the specific interaction with the specific type of fatigue.
Table 1

The effectiveness of driver fatigue technologies in detecting or combating TR and SR fatigue.

<table>
<thead>
<tr>
<th>Detection technologies</th>
<th>TR-active fatigue</th>
<th>TR-passive fatigue</th>
<th>SR fatigue</th>
</tr>
</thead>
<tbody>
<tr>
<td>PerCLOS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Head-Nod</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deadman switch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crash prevention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane departure warning systems</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Collision avoidance warning systems</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Countermeasures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adaptive automation</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interactive technology</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

being experienced. Table 1 presents a summary of the countermeasures proposed to be beneficial for each type of fatigue. Several studies have shown that a nap and/or caffeine are the only effective countermeasures to SR fatigue (Maycock, 1996; Philip et al., 2006; Reyner & Horne, 2000).

Crash prevention technologies, such as lane departure warnings and collision avoidance warning systems, are likely to be helpful in both forms of TR and SR fatigue. In both types of fatigue, lane position variability and the potential for lane departures increase. Off the road crashes are the most common type of crashes from falling asleep at the wheel (George, 2005). Rumble strips are helpful in alerting drivers to off the road incidents, and are being installed more frequently on major highways.

In conclusion, future research testing measures of fatigue and fatigue-related technology must distinguish between different forms of fatigue. Countermeasure technologies for driver fatigue will be more or less effective depending on the type of fatigue the driver is experiencing. These technologies may be useful especially in the truck driving industry to help detect SR fatigue to indicate a need to take a break, or to help manipulate task load to better engage the driver. Future technologies that integrate several measures of fatigue (such as PERCLOS and lane departure warnings) might be more reliable and effective at detecting driver fatigue and reducing crashes associated with fatigue.

References


Lenné, M. G., Triggs, T. J., & Redman, J. R. (1998). Interactive effects of sleep deprivation, time of day and driving experience on a driving task. Sleep, 21(1), 38–44.


