Boring but Demanding: Using Secondary Tasks to Counter the Driver Vigilance Decrement for Partially Automated Driving

Scott Mishler[1] & Jing Chen[2]

[1]Old Dominion University, Norfolk, VA, USA

[2]Rice University, Houston, TX, USA

Précis: Driving-related and non-driving related secondary tasks were added to a simulated, prolonged partially driving task. Non-driving related secondary task mitigated the vigilance decrement. Resource depletion and disengagement were both shown to be causes of the vigilance decrement.

Topic choice: Surface Transportation

Acknowledgement

Jing Chen https://orcid.org/0000-0003-0394-0375

The authors have no known conflicts of interest to disclose.

This work was supported by the National Science Foundation award #2007386 and #2245055. The authors thank Alexa Quesnel, Kyle Canady, Manuela Urbanske Donnelly, and Kayla Wall-Green for their assistance in data collection for this study. Part of the data was presented at the 2022 Annual Meeting of Human Factors and Ergonomics Society and included as an extended abstract in the conference program.

*Correspondence concerning this article should be addressed to Jing Chen, jingchen@rice.edu, Department of Psychological Sciences-MS25, Rice University, 6100 Main Street, Houston, TX 77005.

Abstract

Objective. We investigated secondary-task-based countermeasures to the vigilance decrement during a simulated partially automated driving (PAD) task, with the goal of understanding the underlying mechanism of the vigilance decrement and maintaining driver vigilance in PAD.

Background. Partial driving automation requires a human driver to monitor the roadway, but humans are notoriously bad at monitoring tasks over long periods of time, demonstrating the vigilance decrement in such tasks. The overload explanations of the vigilance decrement predicted the decrement to be worse with added secondary tasks due to increased task demands and depleted attentional resources, whereas the underload explanations predicted the vigilance decrement to be alleviated with secondary tasks due to increased task engagement.

Method. Participants watched a driving video simulating PAD and were required to identify hazardous vehicles throughout the 45-min drive. A total of 117 participants were assigned to three different vigilance-intervention conditions including a driving-related secondary task (DR) condition, a non-driving-related secondary task (NDR) condition, and a control condition with no secondary tasks.

Results. Overall, the vigilance decrement was shown overtime, reflected in increased response times, reduced hazard detection rates, reduced response sensitivity, shifted response criterion, and subjective reports on task-induced stress. Compared to the DR and the control conditions, the NDR displayed mitigated vigilance decrement.

Conclusion. This study provided convergent evidence for both resource depletion and disengagement as sources of the vigilance decrement.

Application. The practical implication is that infrequent and intermittent breaks using a non-driving related task may help alleviate the vigilance decrement in PAD systems.

Keywords: sustained attention, driver vigilance, resource depletion, task engagement, driving automation systems

Boring and Demanding: Using Secondary Tasks to Counter the Driver Vigilance Decrement for Partially Automated Driving

The current partially automated driving (PAD) systems still require the human driver to monitor the driving scene to look out for unexpected problems in the roadway. Drivers must maintain sustained attention, or vigilance, to watch for any potential issues, which is a challenging task for humans (Greenlee et al., 2018; Körber et al., 2015). The act of monitoring the roadway while the automation does much of the primary driving task can quickly become monotonous and boring over time. Drivers tend to respond more slowly and make fewer safe responses to dangerous events as time goes on, resulting in what is commonly referred to as the vigilance decrement (Parasuraman, 1986). PAD still requires the human driver to remain vigilant, yet the driving task becomes more monotonous than manual driving, which imposes greater challenges for driving safety. Our study examined this vigilance decrement issue in the context of PAD to develop strategies to mitigate the decrement through various secondary tasks and to further understand the theoretical underpinnings of the vigilance decrement.

Vigilance

Vigilance refers to the ability to pay attention and maintain focus on a task while responding to infrequent, unpredictable target stimuli over prolonged periods of time (Parasuraman, 1986; Warm et al., 2008). Vigilance is required for many common tasks including manual driving, and the vigilance decrement has been shown to be robust in this task (Larue et al., 2010; Thiffault & Bergeron, 2003). The vigilance decrement has also been shown in PAD, for which the human driver is required to monitor the environment for occasional although critical hazardous signals.

To understand the vigilance tasks, a few terms need to be defined. There have been two distinct task paradigms used to study vigilance: discrimination and identification. In discrimination tasks, there are events that occur at relatively regular intervals and a small portion of those events are defined as the critical signals. Participants are required to discriminate the signal from the other, noise events. In identification tasks, there are no defined noise events, and participants are to respond to the signal events that occur at random times, thus requiring operators to hold extra information in working memory during the task. Event rate is how often the events, including both the signal and noise events, if any, occur over a period of time, so it applies to both the identification and discrimination paradigms. Signal probability is defined as the percentage of the signal events among all events, and thus only applies to the discrimination paradigm. The type of task and vigilance events can influence how the operator performs in a vigilance task. In contrast to some vigilance tasks that require operators to hold specific information in working further and consequently increase workload, driving tasks mostly have responses that come naturally to licensed drivers, so they already know what to look out for.

For the discrimination paradigm, signal detection theory (SDT) can be used to depict participant's performance, showing shifts in operators' responses which can explain potential underlying attentional shifts (Green & Swets, 1966). Response sensitivity refers to the ability to correctly distinguish the signal from the noise, and response bias is the tendency of the individual to report the presence of the signal (See et al., 1995). When an individual says "yes" (i.e., target present) to a signal it is considered a *hit* and saying "no" (i.e., target absent) to a signal is a *miss*. Alternatively, saying "yes" to noise is a *false alarm* and saying "no" to noise is a *correct rejection*.

Theories of the vigilance decrement are comprised of two general categories that describe the cause of the decrement: underload and overload. Among the underload explanations, the arousal-based theory has been the key explanation of the vigilance decrement (Welford, 1968). The arousal theory postulates that performance is low when arousal is at low levels. During a vigilance task, because the individual is attending to rare signals in a monotonous task, they become under stimulated over time and their arousal level decreases. An important feature of the vigilance task is that signal rates are low, leading to long gaps between events that require attention. The mindlessness theory (Manly et al., 1999) proposes that sustained attention is determined by an internal supervisory ability to control attention, and the attentional control wanes when signal rates are low due to a lack of external stimulation. This leads to an absent-minded approach where the signal is lost in the monotony of the task due to attentional drift. Research looking at heart rate and respiration during vigilance tasks demonstrated decrements in each along with decreased performance over time, providing evidence for under-arousal and disengagement (Pattyn et al., 2008). Expanding on the mindlessness theory, the mind-wandering hypothesis (Robertson et al., 1997; Smallwood et al., 2004; Smallwood & Schooler, 2006) postulates that attention is shifted away from the central task toward internal, task-unrelated thoughts during the vigilance task, thus resulting in a decrement of vigilance. Both the arousal and mindlessness theories of the vigilance decrement can be considered as underload theories that attribute the vigilance decrement to the operator being underloaded.

In contrast to the underload theories, there are theories that attribute the vigilance decrement to the operator being overloaded. The earlier sustained demand theory focuses on the information processing demands of the vigilance task, and posits that arousal and stress increase due to the high demands of maintaining attention (Parasuraman, 1979). The continual effortful demand of information processing leads to high mental workload, particularly in vigilance tasks that require more difficult discriminations, resulting in higher stress and decreased vigilance over time (Deaton & Parasuraman, 1993; Warm et al., 1996, 2008). The resource theory of vigilance, based on attentional resource models such as the unitary resource model (Kahneman, 1973) and the multiple resource model (Wickens, 2002), expands this idea of stress and mental demand based on evidence of the depletion of attentional resources in vigilance tasks (Grier et al., 2003; Warm et al., 2008). The resource theory has been supported by findings showing that vigilance tasks are demanding, involve high workload, and induce stress (Johnson & Proctor, 2004; Szalma et al., 2004; Warm et al., 1996). When vigilance tasks require discriminations based on cognitive information (e.g., memory), the decrement is likely to be larger because of the higher mental demand than when discriminations are based on sensory information (e.g., color; See et al., 1995; Warm et al., 2008). Further evidence for the resource theory showed that resources were depleted as time progressed in a vigilance task, and that more demanding tasks show larger vigilance decrements (Greenlee et al., 2019; Helton & Russell, 2011; Helton & Warm, 2008).

The underload and overload vigilance theories have important implications for the design of intervention strategies to mitigate the driver vigilance decrement in automated driving. According to the underload theories, more engaging tasks should result in fewer attentional lapses over time. More interesting tasks should sustain arousal leading to less mind-wandering caused by underload. Furthermore, added stimulation via alternative tasks would improve overall arousal, preventing unwanted attentional withdrawal and disengagement (Pattyn et al., 2008; Smallwood et al., 2004). However, the overload theories suppose that manipulations of task engagement should have no effect or should increase the vigilance decrement due to increased task demands (Thomson et al., 2016). If the individual is already depleting their pool of resources, engagement should not matter and adding additional tasks would further overload them, leading to more stress and workload (Epling et al., 2019; Wickens, 2002). Both of these implications are informative for the design of intervention strategies. If the vigilance decrement is due to an individual being overloaded, actions can be taken to reduce their cognitive load so that vigilance will be maintained for detecting safety-critical signals. In contrast, if vigilance has waned due to low arousal and mind wandering, stimulating and engaging the individual will benefit sustained attention on the task (Hancock & Verwey, 1997; Warm et al., 2008).

Vigilance in Driving Automation Systems

The level of automation for most current semi-autonomous vehicles is SAE Level 2 (partial) automation (SAE, 2021). Level 2 automation requires the human driver to supervise the driving automation system while the system maintains sustained lateral and longitudinal vehicle motion control. The human driver is responsible for object and event detection and recognition during the driving task and must respond appropriately if the driving automation system is unable to avoid a hazard or object in the roadway. The human must always maintain their attention on the primary driving task to supervise the travel, leaving them with the difficulty of staying vigilant for any unforeseen hazards.

As previously discussed, human drivers using a driving automation system are required to monitor the system's actions, which becomes a vigilance task over time. Greenlee and colleagues (2018) conducted a study to determine if the vigilance decrement could be observed for drivers monitoring the roadway for hazards during PAD. They found that the hazard detection rate declined over time and reaction time (RT) to hazards slowed as the drive went on. Participants' workload and subjective stress ratings indicated that the sustained monitoring task was demanding and distressing. As a follow-up, Greenlee et al. (2019) added manipulations of spatial uncertainty of the signals and event rate to investigate effects of task demands in vigilance performance. Their high spatial uncertainty as well as fast event rate served to increase monitoring demands in comparison to low spatial uncertainty and slow event rate. Their results showed that detection performance was worse with higher monitoring demands, indicating that driver overload is likely the reason for the vigilance decrement in partially automated driving.

While the PAD task makes it more of a vigilance task than manual driving, it also potentially leads the human driver to perform secondary tasks more often. Depending on the underlying root of vigilance (i.e., overload or underload), performing secondary tasks may seemingly cost drivers' performance in their main vigilance task of monitoring the driving environment. However, secondary tasks may be beneficial in terms of vigilance maintenance when strategically designed.

Benefits of Secondary Tasks

There has been some evidence that utilizing secondary tasks can help alleviate the vigilance decrement. In air traffic control, vigilance decrements were negated when a secondary task of clicking on each aircraft as it entered the airspace was added to an air traffic control monitoring task (Pop et al., 2012). In manual driving, the effect of added tasks was examined on truck driving performance in a prolonged simulated driving task (Drory, 1985). When a brief voice-communication task was added every 15 min, driving performance significantly improved. More recently, benefits of an interactive cognitive task were found when participants answered multiple choice questions regarding general knowledge during prolonged, monotonous drives (Gershon et al., 2009). Similarly, drivers' manual driving performance (e.g., lane keeping, steering control) was improved when drivers were engaged in a verbal secondary task of free word association during a prolonged simulated drive (Atchley & Chan, 2011). These results are consistent with predictions of the underload theories when adding secondary tasks to the vigilance task. It is worth noting that these secondary tasks are mostly infrequent and intermittent, and do not overlap with the primary vigilance tasks.

However, studies using more traditional vigilance tasks have shown the cost of secondary tasks. For example, Helton and Russell (2011, 2013) added concurrent verbal and spatial secondary tasks to a vigilance task of monitoring an infrequent letter *O* among displayed letters of *D* and minor reversed *D*. Larger vigilance decrements were found with the secondary tasks that increased working memory load. In the driving domain, a meta-analysis focusing on manual driving performance found performance costs in terms of hazard detection and number of collisions due to cell phone use, which can be considered a secondary task (Caird et al., 2018). These results are consistent with predictions of the overload theories, where the larger vigilance decrements imposed by secondary tasks are likely because they deplete mental resources faster. In contrast to the secondary tasks that are shown to be beneficial, these secondary tasks that worsen the vigilance decrement tend to be continuous, tax working memory, and overlap with the primary vigilance tasks.

CURRENT STUDY

Considering the two theoretical explanations underpinning the vigilance decrement – overload and underload – the current study aimed to examine possible countermeasures leveraging secondary tasks to assuage the decline in vigilance. The theoretical motivation for the current study was to disentangle the overload and underload explanations of the vigilance decrement. The overload explanation predicts that an added secondary task will increase the task demands and worsen the vigilance decrement overtime. The underload explanation predicts that an added secondary task will increase engagement of the driver and thus mitigate the vigilance decrement. Our results comparing the secondary-task conditions with the control condition would provide evidence to distinguish these two theoretical explanations. The practical motivation of the current study was to develop secondary-task-based countermeasures to mitigate the vigilance decrement in PAD. Given this practical motivation, we chose secondary tasks that are infrequent and intermittent. The result was expected to provide insight for designers of partially automated vehicles in consideration of the vigilance decrement.

The performance measures used in these studies for manual driving (e.g., lane keeping) are not suitable for measuring driver performance in driving automation systems, where drivers are more likely to perform secondary tasks. In the context of automated driving, Miller et al. (2015) show that, in comparison to an activity of merely supervising the advanced driver assistance system, seemingly distracting activities (e.g., reading, watching videos) indeed reduce the likelihood of driver drowsiness. However, their measure of driver drowsiness was based on visual coding of driver behavior such as yawns and eye closures, and their focus was on predicted and structured transition to driver control where drivers were given 20 s ahead of the transition. The current study used performance-based vigilance measures as well as unpredicted hazardous events.

Furthermore, previous studies on driver vigilance in the driving domain only used driving-irrelevant secondary tasks (Atchley & Chan, 2011; Drory, 1985; Gershon et al., 2009), and no study has examined the effects of secondary tasks on the vigilance decrement in the context of PAD. We tested two types of verbal prompts, non-driving related (NDR) and driving-related (DR), interjected throughout the drive to redirect drivers' attention to the driving task at various stages. Participants were in a simulated SAE (2021) Level 2 partially automated vehicle down the roadway while cars passed by at a consistent rate in the opposite lane. Participants needed to monitor the roadway environment and the passing vehicles and were required to respond when a vehicle in the opposing lane crossed over the centerline (i.e., a hazardous event). In the NDR task, participants were asked non-driving related general knowledge questions; in the DR task, participants were asked driving and roadway relevant questions, which required the participants to scan the driving environment in order to answer the questions. According to the underload theories, both types of secondary tasks would mitigate the vigilance decrement in comparison to the control condition, and the DR would mitigate the decrement further than the NDR due to the former being more engaging. According to the overload theories, both types of secondary tasks would worsen the vigilance decrement due to added task demands in comparison to the control condition, and the DR task would lead to a larger decrement than the NDR task due to the former being more demanding visually and taxing the same resources the vigilance task used (Wickens, 2002). Our goal was to determine if these tasks would help maintain vigilance to improve accuracy and RT to the hazardous events without increasing workload compared to a control condition with no intervention.

Method

Participants. A total of 117 participants (age: M = 20.50, SD = 3.98; 86 female, 31 male) were recruited through SONA, an online research participation system. Participants were required to have a valid driver's license. This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board at Old Dominion University. Informed consent was obtained from each participant. All participants received credit toward course requirements.

Materials. The study was presented through E-prime 3.0 (pstnet.com/products/e-prime) and contained videos of a simulated driving environment created in STISIM driving simulation software (stisimdrive.com). The study was presented on a Dell 27-inch monitor with a 1920×1080 resolution. The Short Stress State Questionnaire (SSSQ; Helton, 2004) was used to measure subjectively reported distress, worry, and engagement before and after the experiment. The NASA-TLX (Hart & Staveland, 1988) was used to measure participants' workload at the end of the experiment.

Procedure. Participants started by filling out a consent form and demographics information, then took the pre-task SSSQ. Once they completed these questions, they were given instructions that they would be assisting with the training of an automated vehicle and would need to press the spacebar when they saw a vehicle coming from the opposite lane cross the centerline. They were encouraged to respond as quickly as possible once they saw the signal and had a five-second window to respond or else their response was counted as a miss. During a two-minute practice session, participants were shown the situation where they were supposed to respond to the vehicle crossing the centerline (i.e., the hazardous event), and there was a total of six hazardous vehicles among a total of 63 vehicles during this practice drive. The first hazardous vehicle was accompanied by an arrow (see Figure 1) and the researcher walked them through the scenario. The participant responded to the subsequent five hazardous vehicles on their own to ensure their understanding of the task, with the researcher watching to verify that they responded appropriately. There was no training on the secondary task other than instructions given to them at the start of the experimental drive. Participants were able to redo the practice session if they did not fully understand the experiment. Participants proceeded to the experimental drive after they successfully completed the training, identifying at least four out of the remaining five hazards.

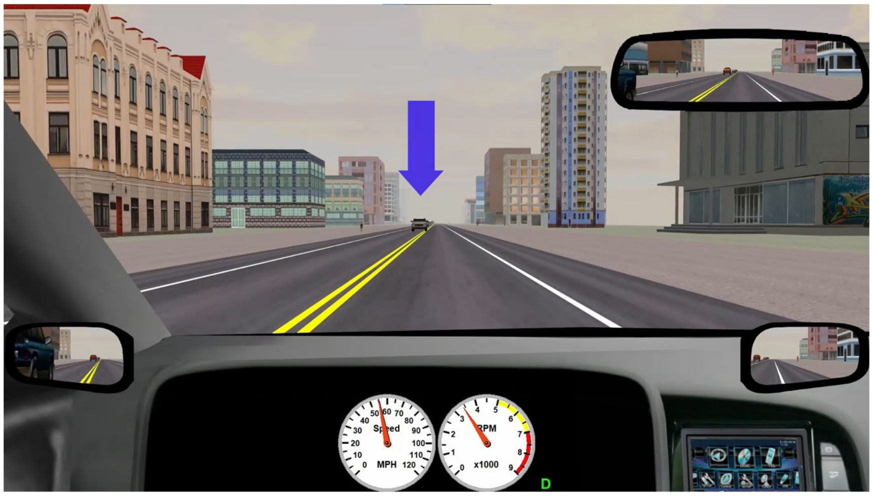

Figure 1. *Roadway with the hazardous vehicle over the centerline (the arrow was not present during the experimental drive)*

The experimental section consisted of a 45-minute drive. A total of 1350 vehicles in the opposite lane passed by the participants' vehicle, resulting in a car passing about every two seconds. Among the passing vehicles, there were 68 hazardous ones, resulting in a signal rate of about 5%. The 68 hazardous events were randomly distributed throughout the drive, with none occurring in the first or last two and a half minutes. Participants' responses were indicated by hitting the spacebar and the system would honk in response to the input.

The independent variable manipulated in the experimental drive was the vigilance-intervention strategy (between-subjects). Each participant was randomly assigned to one of the three conditions, the control, NDR, or DR condition. The control condition only required participants to respond to the hazardous vehicles by pressing the spacebar on the keyboard. The NDR and DR conditions added secondary voice tasks by asking either driving relevant or non-driving relevant questions at eight unique locations randomly spaced throughout the drive. All visual stimuli for the secondary tasks in the DR condition were located on the right side of the road to ensure they were clearly seen by participants, but were randomly distributed in time to ensure that their appearance could not be predicted. All questions were pre-recorded voices and required participants to answer “yes” or “no” verbally. For the NDR condition, the questions were about simple knowledge (e.g., “Is January the first month of the year”). The DR condition contained questions about driving-related objects in the driving environment (e.g., “Is the current speed limit 55 miles per hour?”). The questions in both conditions were matched to be similar for word count and presentation time. There was a total of eight questions in each of the two experimental conditions, and they were randomly distributed throughout the drive to avoid participants predicting the timing of the questions. The questions were not asked at the same time as any hazardous vehicles. Participants answered the questions verbally, and their response was logged on a response sheet by the researcher. At the end of the experiment, participants completed the post-SSSQ and NASA-TLX.

The dependent variables included the RT to the hazardous events, hit rate, and false alarm rate of responses. RT was recorded from the time the vehicle started divulging from its lane to cross the centerline until the spacebar was pressed. A response was a hit if the participant correctly identified the hazardous vehicle and pressed the spacebar. A false alarm was defined by the participant indicating a response when there was no hazardous vehicle. We also measured response criterion and sensitivity using SDT calculated with the hit rate and false alarm rate (Stanislaw & Todorov, 1999). Additionally, the ratings for the pre-SSSQ and post-SSSQ were divided into three sections – distress, worry, and engagement – and the mean score changes from pre- to post-SSSQ were calculated. Finally, the NASA-TLX was computed for each aspect (mental workload, effort, physical workload, frustration, temporal workload, performance) and for the overall mean.

Results

The RT, hit, and false alarm data from the individual 68 hazardous events were evenly divided into four periods of watch (POW) with 17 responses in each. Separate 3 (control, NDR, DR) × 4 (period of watch, POW 1-4) repeated-measures Analyses of Variance (ANOVAs) were conducted on each of the DVs and reported below. Arcsine transformation was conducted on the hit and false alarm rates for the ANOVAs. In addition, trend analyses were conducted across the POWs to determine if RTs, hits, and false alarms increased or decreased linearly over time.

Response Time. The main effect of POW was significant (see Figure 2; *M*s = 1027.87 ms, 1051.08 ms, 1115.10 ms, 1114.25 ms, for each period 1-4, and *SDs* = 201.27 ms, 212.01 ms, 212.82 ms, 230.72 ms, respectively), $F(3,342) = 20.27$, $p < .001$, $\eta_p^2 = .15$. There was a significant linear trend, $F(1,114) = 33.12$, $p < .001$, $\eta_p^2 = .23$, indicating that RT increased (responses were slower) over time (as indicated by POW). Neither the main effect of vigilance-intervention strategy, $F < 1$, nor the interaction between the two factors was significant, $F(6,342) = 1.26$, $p = .275$ $\eta_p^2 = .02$.

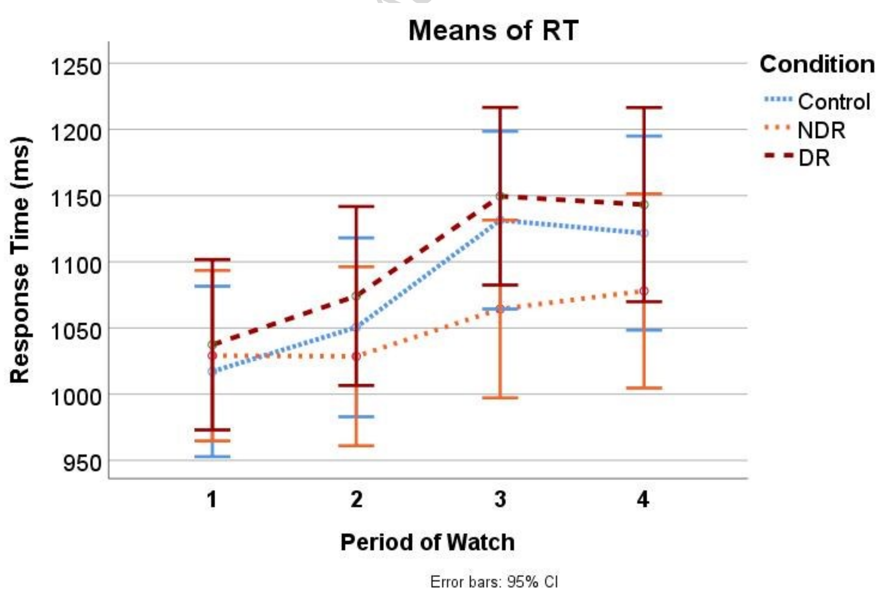

Figure 2. *Mean response time as a function of period of watch and vigilance-intervention strategy (NDR = non driving related; DR = driving related). Error bars are 95% CIs.*

Hit Rate. The ANOVA on the hit rate (see Figure 3) indicated a significant main effect of POW (*M*s = 98.69%, 96.63%, 93.41%, .93.82%, and *SDs* = 3.71%, 6.19%, 11.46%, 9.98%, for each period 1-4, respectively), $F(3,342) = 17.73$, $p < .001$ $\eta_p^2 = .13$. There was a significant linear trend of hit rate, $F(1,114) = 40.61$, $p < .001$, $\eta_p^2 = .26$, showing that the hit rate decreased (more hazardous events were missed) over time. The main effect of vigilance-intervention strategy was significant (*M*s = 94.53%, 97.97%, 94.54%, for control, NDR, and DR, respectively), $F(2,114) = 5.42$, $p = .006$, $\eta_p^2 = .09$. Post hoc comparisons showed the hit rate for the NDR was significantly different from both the control and DR conditions, *p*s = .003 and .010, respectively. The interaction between the two factors was also significant, $F(6,342) = 2.34$, $p = .031$, $\eta_p^2 = .04$. Simple main effects analyses showed that POW had a significant effect on hit rate for the control condition, $F(3,112) = 10.83$, $p < .001$, $\eta_p^2 = .23$, the NDR condition, $F(3,112) = 2.97$, $p = .035$, $\eta_p^2 = .07$, and the DR condition, $F(3,112) = 6.50$, $p < .001$, $\eta_p^2 = .15$. Pairwise comparisons showed that for the control condition, hit rate between the POWs significantly decreased for periods 1 to 2, $p = .003$, and 2 to 3, $p = .043$, but not from 3 to 4, $p = .462$. Similarly for the DR condition, hit rate between the POWs significantly decreased for periods 1 to 2, $p < .001$, and 2 to 3, $p < .001$, but not from 3 to 4, $p = .243$. However, for the NDR condition hit rate only decreased from period one to two, $p = .013$ but not for the subsequent POWs.

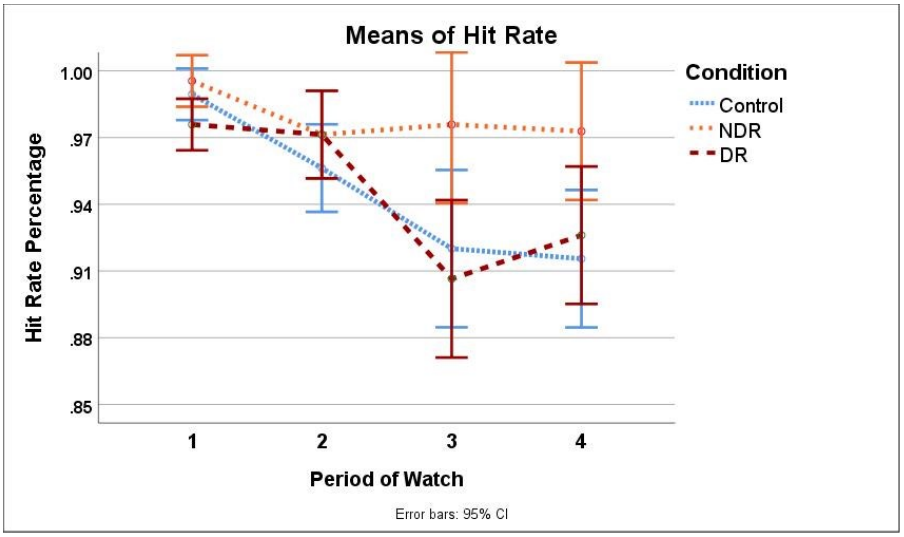

Figure 3. *Mean hit rate as a function of period of watch and vigilance-intervention strategy (NDR = non driving related; DR = driving related). Error bars are 95% CIs.*

False Alarms. The analysis on false alarm rate showed no significant main effect of POW (*M*s = 2.67%, 2.67%, 3.49%, 2.05% and *SDs* = 4.56%, 4.13%, 4.87%, 3.71%, for each period 1-4, respectively), $F(3,342) = 2.22$, $p = .085$, $\eta_p^2 = .02$, nor was the vigilance-intervention strategy or interaction significant, *F*s < 1.

Signal Detection Theory Measures. SDT analyses were conducted using the non-parametric analysis with *A'* (response sensitivity) and B''_D (response criterion; See et al., 1997). Response sensitivity (*A*'; see Figure 4) showed a significant main effect of POW (*M*s = .991, .981, .973, .978, and *SD* = .01, .02, .04,.03, for each period 1-4, respectively), $F(3,342) = 13.24$, $p < .001$, $\eta_p^2 = .10$. There was a significant linear trend, $F(1,114) = 22.11$, $p < .001$, $\eta_p^2 = .16$, indicating that sensitivity generally decreased over time. The main effect of vigilance-intervention strategy was significant, $F(1,114) = 3.15$, $p = .047$, $\eta_p^2 = .05$. Post-hoc comparisons showed that sensitivity for the NDR (*M* = .987) was significantly higher than both the control (*M* = .978) and DR (*M* = .978) conditions, *p*s = .030 and .034, respectively. The interaction between

POW and vigilance-intervention strategy was not significant, $F(6,342) = 1.73, p = .114, \eta_p^2 = .03$.

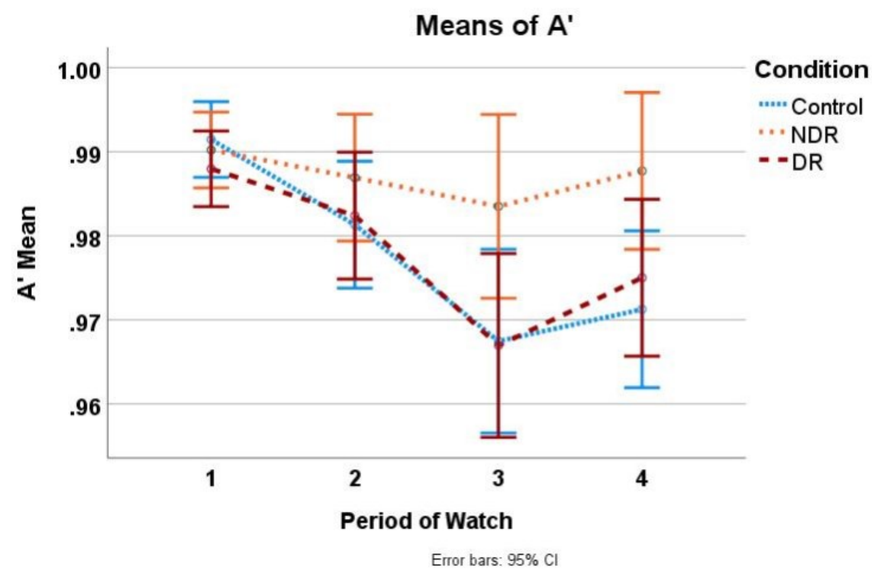

Figure 4. *Mean response sensitivity as a function of period of watch and vigilance-intervention strategy (NDR = non driving related; DR = driving related). Error bars are 95% CIs.*

Analysis on response criterion (B''_D; see Figure 5) showed a significant main effect of POW, (*M*s = -.046, .018, .073, .114, and *SDs* = .20, .22, .33, .28 for each period 1-4, respectively), $F(3,342) = 10.06, p < .001, \eta_p^2 = .08$. There was a significant linear trend, $F(1,114) = 24.93, p < .001, \eta_p^2 = .18$, indicating that criterion shifted to become more conservative over time. The main effect of vigilance-intervention strategy was also significant (*M*s = .075, -.025, .069, for control, NDR, and DR, respectively), $F(1,114) = 5.06, p = .008, \eta_p^2 = .08$. Post-hoc comparisons showed that NDR was significantly less conservative than both control ($p = .006$) and DR ($p = .009$). The interaction between POW and vigilance-intervention strategy was also significant, $F(6,342) = 2.17, p = .045, \eta_p^2 = .04$. The simple main effects analyses showed similar result patterns to the hit rate. POW had a significant effect on response criterion for the control condition, $F(3,112) = 5.91$, $p = .001$, $\eta_p^2 = .14$, the NDR condition, $F(3,112) = 3.20$, $p = .026$, $\eta_p^2 = .08$, and the DR condition, $F(3,112) = 4.36$, $p = .006$, $\eta_p^2 = .11$. Pairwise comparisons showed that for the control condition, response criterion between the POWs significantly shifted more conservative for periods 1 to 2, $p = .031$, 1 to 3, $p = .014$, 1 to 4, $p < .001$, and 2 to 4, $p = .034$. Similarly for the DR condition, response criterion between the POWs significantly shifted more conservative for periods 1 to 3, $p = .007$, 1 to 4, $p = .009$, 2 to 3, $p = .002$, and 2 to 4, $p = .005$. However, for the NDR condition, the response criterion between the POWs only significantly shifted more conservative for periods 1 to 2, $p = .018$ and 1 to 4, $p = .027$. No other comparisons were significant, ps > .05. The response criterion in the control and DR conditions shifted to be more conservative over time but shifted less for the NDR condition.

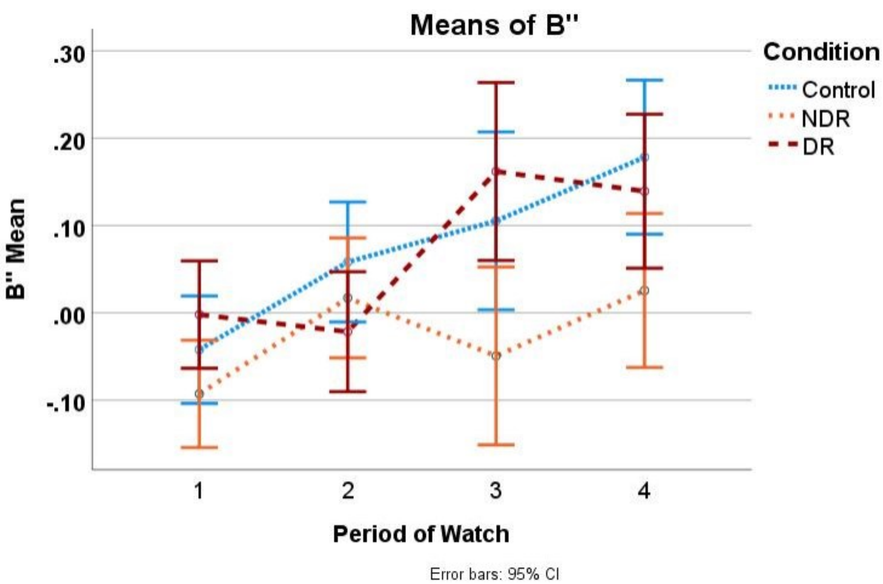

Figure 5. *Mean response criterion as a function of period of watch and vigilance-intervention strategy (NDR = non driving related; DR = driving related). Error bars are 95% CIs.*

SSSQ. The standardized SSSQ change scores were calculated for each of the three scales using the formula, (Post-score – Pre-score) / σ of the Pre-scores (Helton, 2004). A 3 (vigilance-intervention strategy; control, NDR, and DR) × 3 (scales; distress, engagement, and worry) repeated measures ANOVA was performed on these change scores (see Figure 6). A significant main effect of scales was found (*M*s = .437, -.725, -.109, for distress, engagement, and worry, respectively), $F(2,228) = 40.10$, $p < .001$, $\eta_p^2 = .26$. No significant effects were found for vigilance-intervention strategy, $F(1,114) = 2.49$, $p = .089$, $\eta_p^2 = .04$, or for the interaction, $F(4,228) = 1.90$, $p = .112$, $\eta_p^2 = .03$. For distress, the changes in score were significantly above zero, $t(44) = 4.36$, $p < .001$, $t(44) = 2.09$, $p = .021$, $t(44) = 2.78$, $p = .004$, for the control, NDR, and DR conditions, respectively. For engagement, the changes in score were below zero, $t(44) = 5.43$, $p < .001$, $t(44) = 2.92$, $p = .003$, $t(44) = 4.55$, $p<.001$, for the control, NDR, and DR conditions, respectively. For worry, none of the conditions had a change score different from zero *t*s < 1.43.

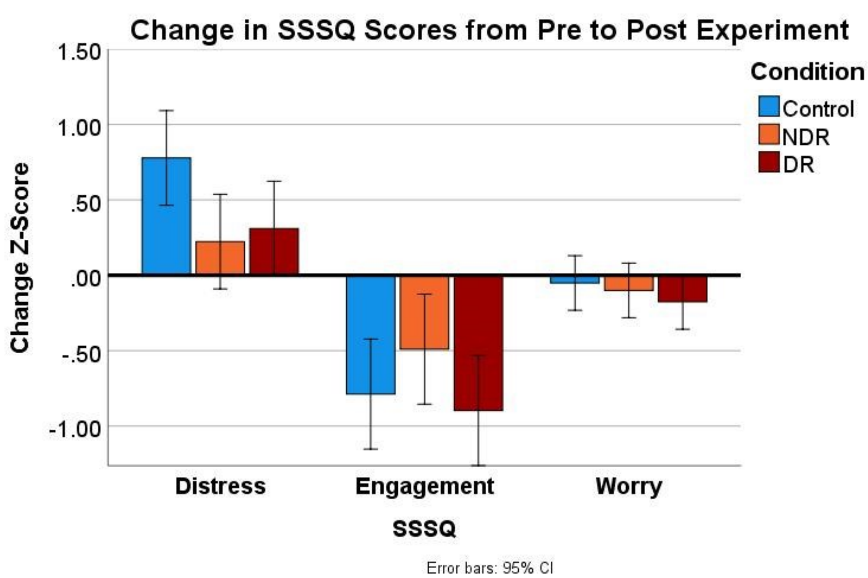

Figure 6. *Mean SSSQ change z-scores for each of the three vigilance-intervention-strategy conditions (NDR = non driving related; DR = driving related). Error bars are 95% CIs.*

NASA-TLX. For the NASA-TLX data, we conducted a one-way ANOVA on the global workload with vigilance-intervention strategy as a between-subjects factor. The analysis showed no significant effect of vigilance-intervention strategy, $F(2,114) = 1.65$, $p = .196$, $\eta_p^2 = .03$, although the mean values (8.40, 7.41, and 8.35 for the control, NDR, and DR conditions, respectively) showed a lower mean for the NDR condition. For the NASA-TLX subscales (Mental, Physical, Temporal, Performance, Effort, and Frustration), there were no significant differences: Physical, $F(2,114) = 1.14$, $p = .241$ $\eta_p^2 = .03$; Effort, $F(2,114) = 1.12$, $p = .330$ $\eta_p^2 = .02$; Frustration, $F(2,114) = 2.02$, $p = .137$ $\eta_p^2 = .03$, and all the others, Fs < 1.

Discussion

This study showed results of the vigilance decrements in PAD systems and potential countermeasures. We found slower RTs and lower hit rates responding to hazardous vehicles on the road over time during the drive, in line with previous literature that indicates Level 2 automation is a monotonous vigilance task (Greenlee et al., 2018; Körber et al., 2015). The SSSQ results also showed the expected result pattern for subjective reports before and after a vigilance task, decreases in engagement, and increases in distress (Greenlee et al., 2018; Helton, 2004). These results are consistent with those found in other domains such as air traffic control, military, and industrial supervisory control monitoring (Parasuraman et al., 1987).

The effect of POW on response time was not moderated by the vigilance-intervention strategy, indicating that the NDR was still subject to some vigilance decrement in RT, similarly to the other conditions. For hit rate, even though the POW had a significant effect in all conditions, it only affected the NDR condition in the first POW, and its hit rate was significantly higher than those for the DR and control conditions in the third and fourth POWs. This result demonstrates a clear benefit of introducing the NDR task over time as the vigilance task continued. The difference between the NDR condition and the other two indicates that the NDR task helped participants maintain sustained attention on hazardous vehicles to perform better than the control and DR conditions. This benefit of the NDR is especially interesting given that it has an added secondary task on top of the control condition. The finding that adding this secondary task did not increase the vigilance decrement but reduced it provides evidence against the overload theories of vigilance, which would predict the secondary task to increase the vigilance decrement.

Both the interventions, NDR and DR, would be expected to perform better than the control condition if the problem of vigilance was one of simply mindlessness or underload (Manly et al., 1999). The difference between the two tasks was that the NDR task simply reengages their attention in general without the need to be redirected visually, whereas the DR task requires them to scan the driving environment visually. Interpreting these results with regard to multiple resource theory, the DR task is more demanding than the NDR task because the primary vigilance task was visual and the DR task competed for the visual resources with the primary vigilance task (Wickens, 2002). Whereas the NDR task only required auditory resources and could be task shared with the primary vigilance task. An underloaded driver could benefit from a secondary task to increase their arousal if that task did not compete for mental resources. However, an overloaded driver would be further taxed by a secondary task, gaining no benefit from increased arousal.

It is worth noting that neither secondary-task intervention made the vigilance decrement worse than the control. A possible reason why the DR condition showed a similar vigilance decrement to the control condition is that the benefit of being more engaged and the cost of being more demanding, specifically in the similar visual resource channel, cancel out for the DR condition. In other words, the benefits of increased arousal due to the added DR task can be negated by the visual resource demand imposed by it, which explains why the NDR task helped alleviate the vigilance decrement while the DR task did not. This explanation requires both disengagement and resource depletion to be sources of the vigilance decrement, which is supported by our SDT analyses.

The results of SDT showed a decrease of response sensitivity along with the response criterion shifting more conservative overtime. These results indicate that the vigilance decrement was partly due to the drivers being less able to discriminate the hazardous vehicles and shifting their response criterion. In addition, both measures showed the advantageous effects of the NDR task in comparison to the DR task and the control condition. Previous research has shown similar results on response sensitivity and has been used to support an explanation of attentional resource depletion (Greenlee et al., 2018, 2019; See et al., 1995). Reduced response sensitivity has been used as evidence for resource depletion that causes the vigilance decrement; and overloaded operator is not able to effectively distinguish the critical signals from the noise, leading to decreased response sensitivity (Caggiano & Parasuraman, 2004; DeLucia & Greenlee, 2022). The subscale of distress in SSSQ also demonstrated increased stress, often caused by high task demands, which is consistent with the resource depletion account.

The shift in criterion could be due to the frequency of the hazardous vehicles being greater during the practice session (3 per min) than during the experimental session (1.7 per min). However, the significantly more conservative criteria for periods 3 and 4 in the control and DR conditions than in the NDR condition indicates that the shift of criteria was not merely due to the change in signal rate. In contrast to response sensitivity, response criterion has been used as evidence for the underload explanation of the vigilance decrement. The response criterion shifting towards more conservative is an indication of the operator becoming more disengaged and less likely to respond to both the critical signals and noise (Thomson et al., 2015). The subscale of engagement in SSSQ also provides convergent evidence that participants become less engaged during the task. As a result, our findings provide evidence for both sources of resource depletion and disengagement in the vigilance decrement.

Based on precedent from prior literature, along with our current results, we propose that both resource depletion and disengagement are sources of the vigilance decrement. For example, Thomson et al. (2015, 2016) draw upon and address the explanatory limitations of both the mind-wander and resource-depletion theories and propose a resource-control theory of mind wandering to explain vigilance performance. One central tenet of this theory is that mind wandering takes attentional resources, which results in poor vigilance performance. When the limited attentional resources are absorbed by mind wandering, performance for the primary task may be sacrificed if it requires the full complement of attentional resources. For Level 2 automation where human drivers seek to monitor the automation, if the task becomes too monotonous, boring, or excessively demanding over time, drivers begin to withdraw and reallocate their resources to a more interesting task. Withdrawing from the primary task leads to task unrelated thoughts, which further disrupt performance on the main task (Forster & Lavie, 2009). By introducing a secondary task as an intervention in the current study, it may cut down on user-generated task unrelated thoughts and allow them a brief, semi-structured break to then reorient their thoughts on the task, consistent with the benefits in performance caused by intermittent breaks shown in prior studies (Atchley & Chan, 2011; Drory, 1985; Pop et al., 2012). However, our measure of disengagement with the SSSQ involved more than simply task unrelated thoughts and might have clouded our ability to fully explain this. The participants might not have been able to self-monitor their engagement or disengagement as it relates to performance and sufficiently report the potential changes via subjective measures. The performance measures showed clear differences between the groups, but the distress measure might not have been sensitive enough to detect the difference if all the participants were not cognizant of their mental processes.

One limitation of this study is that it was not conducted on a driving simulator where participants could fully interact with the vehicle. However, this was done intentionally to closely replicate a classic vigilance task while controlling as many variables as possible. Future research can validate the results in a driving simulator, which allows human drivers to take over control at critical instances and more directly interact with the automation system. Another limitation is that this study only used one type of vigilance signal (the hazardous vehicles). Although it was presented at random times, it does not fit real-world situations where there can be various types of events that occur at different locations. Future research can incorporate different types of hazards to validate the current findings on the different vigilance-decrement interventions. Finally, the number of false alarms was quite low with large variances in the study which may influence the use of signal detection theory. However, some recent research with vigilance has failed to find a large number of false alarms in a similar way to the current results (Epling et al., 2019; Körber et al., 2015). This result calls for caution when using the SDT measures in vigilance studies.

Conclusions

This study shows that human drivers in charge of monitoring driving automation systems during PAD are subject to the vigilance decrement. However, the study demonstrated potential for decreasing the negative performance associated with the vigilance decrement through the implementation of a non-driving related intervention task. Additionally, this work contributed to the theoretical underpinnings of vigilance, providing evidence for both resource depletion and disengagement being causes of the vigilance decrement. The interventions implemented in this study provide insight for how designers of automation systems could alleviate the problems of vigilance and keep the human driver in the loop.

Key Points

- The added driving-related secondary task showed similar vigilance decrement to the control condition.
- The non-driving related secondary task mitigated the vigilance decrement in partially automated driving.
- Signal detection theory measures and subjective reports on task-induced stress provided convergent evidence for resource depletion and disengagement causing the vigilance decrement.
- Infrequent and intermittent breaks using secondary tasks may be utilized in partially automated driving system design to help maintain driving vigilance.

References

Atchley, P., & Chan, M. (2011). Potential benefits and costs of concurrent task engagement to maintain vigilance: A driving simulator investigation. *Human Factors*, *53*(1), 3–12.

Caggiano, D. M., & Parasuraman, R. (2004). The role of memory representation in the vigilance decrement. *Psychonomic Bulletin & Review*, *11*(5), 932–937.

Caird, J. K., Simmons, S. M., Wiley, K., Johnston, K. A., & Horrey, W. J. (2018). Does talking on a cell phone, with a passenger, or dialing affect driving performance? An updated systematic review and meta-analysis of experimental studies. *Human Factors*, *60*(1), 101–133.

Deaton, J. E., & Parasuraman, R. (1993). Sensory and cognitive vigilance: Effects of age on performance and subjective workload. *Human Performance*, *6*(1), 71–97.

DeLucia, P. R., & Greenlee, E. T. (2022). Tactile Vigilance Is Stressful and Demanding. *Human Factors*, *64*(4), 732–745.

Drory, A. (1985). Effects of rest and secondary task on simulated truck-driving task performance. *Human Factors*, *27*(2), 201–207.

Epling, S. L., Edgar, G. K., Russell, P. N., & Helton, W. S. (2019). Is Semantic Vigilance Impaired by Narrative Memory Demands? Theory and Applications. *Human Factors*, *61*(3), 451–461. https://doi.org/10.1177/0018720818805602

Forster, S., & Lavie, N. (2009). Harnessing the wandering mind: The role of perceptual load. *Cognition*, *111*(3), 345–355. https://doi.org/10.1016/j.cognition.2009.02.006

Gershon, P., Ronen, A., Oron-Gilad, T., & Shinar, D. (2009). The effects of an interactive cognitive task (ICT) in suppressing fatigue symptoms in driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(1), 21–28.

Green, D. G., & Swets, J. A. (1966). Signal detection theory and psychophysics. New York, NY: John Wiley and Sons Inc. In *Wiley & Sons, Inc.*

Greenlee, E. T., DeLucia, P. R., & Newton, D. C. (2018). Driver Vigilance in Automated Vehicles: Hazard Detection Failures Are a Matter of Time. *Human Factors*. https://doi.org/10.1177/0018720818761711

Greenlee, E. T., DeLucia, P. R., & Newton, D. C. (2019). Driver Vigilance in Automated Vehicles: Effects of Demands on Hazard Detection Performance. *Human Factors*, *61*(3), 474–487. https://doi.org/10.1177/0018720818802095

Grier, R. A., Warm, J. S., Dember, W. N., Matthews, G., Galinsky, T. L., Szalma, J. L., & Parasuraman, R. (2003). The Vigilance Decrement Reflects Limitations in Effortful Attention, Not Mindlessness. *Human Factors*, *45*(3), 349–359. https://doi.org/10.1518/hfes.45.3.349.27253

Hancock, P. A., & Verwey, W. B. (1997). Fatigue, workload and adaptive driver systems. *Accident Analysis and Prevention*, *29*(4 SPEC. ISS.), 495–506. https://doi.org/10.1016/s0001-4575(97)00029-8

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, *52*(C), 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Helton, W. S. (2004). Validation of a Short Stress State Questionnaire. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *48*(11), 1238–1242. https://doi.org/10.1177/154193120404801107

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental Brain Research*, *212*(3), 429–437. https://doi.org/10.1007/s00221-011-2749-1

Helton, W. S., & Russell, P. N. (2013). Visuospatial and verbal working memory load: effects on visuospatial vigilance. *Experimental Brain Research*, *224*(3), 429–436.

Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica*, *129*(1), 18–25. https://doi.org/10.1016/j.actpsy.2008.04.002

Johnson, A., & Proctor, R. W. (2004). *Attention: Theory and practice*. Sage.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Citeseer.

Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving. *Procedia Manufacturing*, *3*(Ahfe), 2403–2409. https://doi.org/10.1016/j.promfg.2015.07.499

Larue, G. S., Rakotonirainy, A., & Pettitt, A. N. (2010). Predicting driver's hypovigilance on monotonous roads: literature review. *1st International Conference on Driver Distraction and Inattention*.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia*, *37*(6), 661–670. https://doi.org/10.1016/S0028-3932(98)00127-4

Miller, D., Sun, A., Johns, M., Ive, H., Sirkin, D., Aich, S., & Ju, W. (2015). Distraction becomes engagement in automated driving. *Proceedings of the Human Factors and Ergonomics Society*, *2015-January*, 1676–1680. https://doi.org/10.1177/1541931215591362

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science*, *205*(4409), 924–927.

Parasuraman, R. (1986). Vigilance, monitoring, and search. In *Handbook of perception and human performance, Vol. 2: Cognitive processes and performance.* (pp. 1–39). John Wiley & Sons.

Parasuraman, R., & Davies, D. R. (1977). *A Taxonomic Analysis of Vigilance Performance BT - Vigilance: Theory, Operational Performance, and Physiological Correlates* (R. R. Mackie, Ed.; pp. 559–574). Springer US. https://doi.org/10.1007/978-1-4684-2529-1_26

Parasuraman, R., Warm, J. S., & Dember, W. N. (1987). *Vigilance: Taxonomy And Utility BT - Ergonomics and Human Factors: Recent Research* (L. S. Mark, J. S. Warm, & R. L. Huston, Eds.; pp. 11–32). Springer New York. https://doi.org/10.1007/978-1-4612-4756-2_2

Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiology and Behavior*, *93*(1–2), 369–378. https://doi.org/10.1016/j.physbeh.2007.09.016

Pop, V. L., Stearman, E. J., Kazi, S., & Durso, F. T. (2012). Using engagement to negate vigilance decrements in the NextGen environment. *International Journal of Human-Computer Interaction*, *28*(2), 99–106.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*(6), 747–758.

SAE. (2021). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. In *SAE International* (Vol. J3016, pp. 1–41). https://doi.org/https://doi.org/10.4271/J3016_202104

See, J. E., Howe, S. R., Warm, J. S., & Dember, W. N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin*, *117*(2), 230–249. https://doi.org/10.1201/9780203872512.ch23

See, J. E., Warm, J. S., Dember, W. N., & Howe, S. R. (1997). Vigilance and signal detection theory: An empirical evaluation of five measures of response bias. *Human Factors*, *39*(1), 14–29. https://doi.org/10.1518/001872097778940704

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition*, *13*(4), 657–690.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946–958. https://doi.org/10.1037/0033-2909.132.6.946

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Szalma, J. L., Warm, J. S., Matthews, G., Dember, W. N., Weiler, E. M., Meier, A., & Eggemeier, F. T. (2004). Effects of sensory modality and task duration on performance, workload, and stress in sustained attention. *Human Factors*, *46*(2), 219–233. https://doi.org/10.1518/hfes.46.2.219.37334

Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis & Prevention*, *35*(3), 381–391.

Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. *Perspectives on Psychological Science*, *10*(1), 82–96.

Thomson, D. R., Besner, D., & Smilek, D. (2016). A critical examination of the evidence for sensitivity loss in modern vigilance tasks. *Psychological Review*, *123*(1), 70–83. https://doi.org/10.1037/rev0000021

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). *Vigilance and workload in automated systems.*

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, *50*(3), 433–441. https://doi.org/10.1518/001872008X312152

Welford, A. T. (1968). *Fundamentals of skill.*

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

Biographies

Scott Mishler is a PhD candidate in Human Factors Psychology at Old Dominion University. He received his master's degree in Psychology at Old Dominion University in 2019. His areas of research include trust in automation, autonomous driving, and anti-phishing automation.

Jing Chen is an Assistant Professor of Human Factors/Human-Computer Interaction in the Department of Psychological Sciences at Rice University. She received her Ph.D. in Cognitive Psychology and M.S. in Industrial Engineering at Purdue University in 2015. Her areas of research include fundamental principles of human performance and decision-making, and applications of these principles to cybersecurity problems and human-system design. ORCID ID: https://orcid.org/0000-0003-0394-0375