©Copyright 2026

David Prendez

Vigilance Decrement Measurement, Prediction, and Intervention Design in Safety-Critical Environments

David Prendez

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2026

Reading Committee:

Ji-Eun Kim, Co-Chair

Linda Ng Boyle, Co-Chair

Prashanth Rajivan

Program Authorized to Offer Degree:

Industrial & Systems Engineering

University of Washington

Abstract

Vigilance Decrement Measurement, Prediction, and Intervention Design

in Safety-Critical Environments

David Prendez

Chairs of the Supervisory Committee:

Ji-Eun Kim

Department of Industrial & Systems Engineering

Linda Ng Boyle

Departments of Industrial & Systems Engineering and Civil & Environmental Engineering

Vigilance is defined as the ability to remain attentive and alert over a sustained period of time, particularly toward a specific task. The deterioration of vigilance, also known as vigilance decrement, is often experienced by those completing tasks over an extended period of time. In high-risk domains, such as healthcare and driving, measuring vigilance decrement is particularly important given the resulting safety impacts. However, most research studies primarily use behavioral responses during vigilance tasks to assess vigilance levels (e.g., response time, hit rate). While these measures have proven to be effective, they are limited to post-hoc analyses only. Given recent advances in physiological sensing tools and predictive modeling strategies, it

is important to also consider contextual factors and physiological indicators, allowing for more robust, continuous, and objective evaluation of vigilance decrement. For effective prevention and resolution to cases of vigilance decrement in safety-critical settings, interventions must be designed to be proactive and sustainable. This dissertation uses a combination of scoping reviews, healthcare-based studies, and a driving simulator study to answer the following research questions: 1) What contextual or contributing factors, together, affect vigilance levels in safety-critical settings?, 2) What combination of physiological measures can be used to monitor vigilance?, 3) What modeling features and strategies can be used to predict vigilance decrement?, and 4) What types of interventions can be used to maintain vigilance in a sustainable manner?

Findings across the healthcare and driving studies revealed that vigilance is best predicted through a hybrid approach combining contextual factors and multi-sensor physiological measures. In the healthcare domain, sleep-related and environmental factors were significant predictors of vigilance, while heart rate, electrodermal activity, skin temperature, and eye movement-related metrics emerged as key physiological indicators. In the driving domain, driver experience, distraction susceptibility, and workload were the most influential contextual factors, while eye fixations toward a virtual meeting-based secondary task serving as the primary physiological predictor. Continuous measures of driving behavior were used as effective substitutes for peripheral physiological sensing. In the healthcare domain, probabilistic models with temporal elements, such as Dynamic Bayesian Networks, performed well for longer-duration vigilance tasks. In the driving domain, simpler linear models were most effective for single, critical event-based vigilance assessments. A scoping review of 52 studies identified 18 vigilance intervention strategies where sleep/breaks, ingestion, meditation/mindfulness, and task context changes represented the most sustainable options for real-world deployment. In the driving study, an emergency-style warning intervention produced significantly faster response times during the takeover event compared to the less urgent planned-style warning, demonstrating that salient, system-integrated alerts can serve as effective and sustainable interventions for enhancing vigilance prior to safety-critical events in semi-autonomous vehicles.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADS	Automated Driving System
AIC	Akaike Information Criterion
ANOVA	Analysis of Variance
BN	Bayesian Network
BORIS	Behavioral Observation Research Interactive Software
BVP	Blood Volume Pulse
CI	Confidence Interval
CPT	Continuous Performance Test
CSV	Comma-Separated Values
DBN	Dynamic Bayesian Network
DSRI	Driving Safety Research Institute
ECG	Electrocardiogram
EDA	Electrodermal Activity
EEG	Electroencephalogram
fMRI	Functional Magnetic Resonance Imaging
fNIRS	Functional Near-Infrared Spectroscopy
HR	Heart Rate
HRV	Heart Rate Variability
IEEE	Institute of Electrical and Electronics Engineers
IRB	Institutional Review Board
ISAT	Interactive Scenario Authoring Tool
LASSO	Least Absolute Shrinkage and Selection Operator
LED	Light-Emitting Diode
LOOCV	Leave-One-Out Cross Validation
LR	Linear Regression
MI	Myocardial Infarction
ML	Machine Learning
NASA	National Aeronautics and Space Administration

OLS	Ordinary Least Squares
PDF	Portable Document Format
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSQI	Pittsburgh Sleep Quality Index
PVT	Psychomotor Vigilance Task
RF	Random Forests
RLM	Robust Linear Model
RMSE	Root Mean Squared Error
RT	Reaction/Response Time
RQ	Research Question
SAE	Society of Automotive Engineers
SAGAT	Situation Awareness Global Assessment Technique
SART	Situation Awareness Rating Technique
SD	Standard Deviation
SDDQ	Susceptibility to Driver Distraction Questionnaire
SDT	Signal Detection Theory
STAI	State-Trait Anxiety Inventory
SVM	Support Vector Machines
tACS	Transcranial Alternating Current Stimulation
tDCS	Transcranial Direct Current Stimulation
TEMP	Skin Temperature
TLX	Task Load Index
TMT	Tile Mosaic Tool
tRNS	Transcranial Random Noise Stimulation
UW	University of Washington

Acknowledgements

This dissertation, and the completion of my Ph.D. degree, would not have been possible without the support of my countless mentors, colleagues, family, friends, and loved ones. First, and foremost, I would like to extend the utmost gratitude to my Ph.D. advisors, Prof. Linda Ng Boyle and Prof. Ji-Eun Kim. To Linda, thank you for taking a chance on me as an undergrad in your lab and then welcoming me again as your Ph.D. student. Your mentorship throughout my career has been invaluable and I've thoroughly enjoyed brainstorming our research together. To Ji-Eun, thank you for taking me on as your student halfway through my Ph.D. I appreciate all the hands-on time you've spent helping me formulate my dissertation topic, break through into new research topics, and grow as a lead researcher. I'm also grateful to my committee members, Prof. Prashanth Rajivan and Prof. Andrea Stocco, for the time they invested into my doctoral work. Your feedback and genuine interest brought excellent perspectives to this dissertation and made it a much better final product.

At the beginning of my Ph.D., I was fortunate to join a lab full of amazing and supportive doctoral researchers. Thank you to my fellow HFSM lab members, Mayuree Binjolkar, Yilun Xing, Sami Park, Grace Douglas, and Jiaxin Li, for your friendship during this time and for showing me how to succeed in my degree. I'll always look back fondly on our weekly coffee get-togethers and holiday gift exchanges. Toward the end of my Ph.D., I was lucky enough to join another lab with a mix of excellent graduate and undergraduates. Thank you to my fellow HAS lab members, Jiaxin Li, Thea Higgins, Veronika Kettel, Ching Wang, Michael Ling, Daisy Xiao, Junghoon Chung, and Sebastian Halicki for the fun times in the lab, going out to collect data, hanging around campus, going to conferences, and writing papers together. I'd also like to

thank my undergraduate researchers who helped me collect and process the data for my driving simulator study: Faliha Amjad, Juan Valle, Shelly Slater, and Ella Grooters. I really enjoyed mentoring all of you and appreciate all the work you contributed to make this portion of my dissertation possible.

Next, I would like to thank the colleagues and friends I made while working in industry prior to, and during, my Ph.D. During my time as a Human Factors Engineer at Battelle, my coworkers constantly pushed me to grow as a researcher, gave me the tools I needed to succeed, and taught me whatever they could to advance my career growth. Thank you to Christian Richard, John Campbell, Vindhya Venkatraman, Joonbum Lee, Alex Noble, Liberty Hoekstra-Atwood, Jim Brown, Kelly Magee, Krista Taylor, Jessica McDonough, Maddy Bennett, Alex Schonbaum, Claire Textor, and Dan Berler, for making my time at Battelle so rewarding and a very happy time in my life. And thank you to the many of you who continued to push me to go back to school for my Ph.D. even when I was reluctant—it seems you were right after all! Before the final year of my degree, I had the pleasure of interning at Amazon for the summer. I want to thank Eric Weston for connecting me with the opportunity, Mauricio Villa for his mentorship, and the rest of the Noise, Vibration, and Harshness team—Keaten Tuttle, Bryan Ross, Mandip Goswami, Nico Deforeit, Alexis Haire, and Suresh Krishnan—for making my internship an exciting and challenging experience and a refreshing break from my typical research work. I think of you all every time I order something off of Amazon (which is honestly far too often).

I cannot thank my loving family enough for their unwavering support while I was completing my Ph.D. To my parents, Rosalind, Alvin, and Jake, thank you for always believing in me and instilling in me the values of kindness, empathy, and community. To my grandparents, Dave, Laurie, Conrad, and Sally, thank you for spoiling me, for teaching me to always do the

right thing and work hard, and reminding me to take time to care for myself. To my siblings, Alena, Nico, and Denila, thank you for cracking me up, for letting me tease you, for spending quality time with me, and giving me distractions from my work when I needed it. I hope you see this dissertation as proof that if I can accomplish a feat like this, you’re each capable of even greater. To my in-laws, Sophia, Richard, Hung, Lan, Thai, Tammy, Rishabh, Julie, Muneeb, Sophie, and Xuyen, thank you for encouraging me and supporting me during this time. I’m incredibly lucky to be part of two sets of families who care for one another so deeply.

Finally, I extend my most profound thanks to my amazing wife, Tiffany, and our fur babies, Charlie and Ruby. You three are the reason I wake up every day with a big smile on my face. Tiffany, I would not have been able to complete this dissertation without you by my side. Thank you for being my biggest cheerleader, for giving me the confidence to start my Ph.D., and the motivation I needed to finish it.

Chapter 1: Introduction

1.1 What is Vigilance?

Vigilance is defined as the ability to remain attentive and alert over a sustained period of time, particularly toward a specific task (Davies & Parasuraman, 1982; Warm et al., 2018). Being vigilant involves paying close and continuous attention to what is happening around oneself, whether it is in regard to your immediate surroundings, within a system you are overseeing, or even in your own thoughts and actions. People display vigilance in their everyday life; while reading a book, watching movies and TV shows, driving a car, or even at work. For example, a TSA agent constantly scanning luggage at an airport security checkpoint or a quality control inspector carefully examining products for flaws. Vigilance helps us notice important details in order to react accordingly. This ability, also referred to as sustained attention, is crucial for effective interaction with our surroundings and successful completion of tasks.

1.2 Brief History of Vigilance Research

The concept of vigilance began during World War II, motivated by the observation that radar and sonar operators frequently missed infrequent but critical events toward the end of their monitoring shifts (Mackworth, 1948). To explore this phenomenon, British psychologist Norman Mackworth developed the "Mackworth Clock" task, where participants observed an unmarked clock face and were instructed to report occasional double-length movements of the clock's hand (Mackworth, 1948). Mackworth found that participants consistently missed more clock hand movements over time. This decline in sustained attention over time is now referred to as vigilance decrement. Over the following decades, vigilance-related research expanded significantly to better understand the underlying causes and effects of vigilance decrement.

1.3 The Importance of Vigilance Decrement

Vigilance decrement has drawn concerns specifically from the field of human factors and ergonomics for its negative effects on worker safety, well-being, and performance (Warm et al., 2018; Weinger & Englund, 1990). Numerous studies have found vigilance decrement to occur within as little as 5-15 minutes of task engagement (e.g., Nuechterlein et al., 1983; Hancock, 2013; Warm et al., 2018). Vigilance decrement has resulted in a significant number of major and minor accidents in high-risk work environments including healthcare, nuclear power plant monitoring, military surveillance, aviation, and ground transportation (Hancock & Hart, 2002; Molloy & Parasuraman, 1996; Warm et al., 2018). For workers in these roles, a failure to detect critical events due to vigilance decrement can have severe consequences, such as accidents, loss of life, and system failures (Anderson et al., 2023). In healthcare settings, for example, clinicians working long shifts are susceptible to vigilance decrement, potentially affecting their productivity, the safety of their patients, and the overall quality of care provided (Li et al., 2024). At the same time, the well-being of workers themselves is threatened by vigilance decrement (Johannes, 2020), resulting in consequences such as long-term sick leave (Demerouti et al., 2009). Consequently, developing effective tools for detecting vigilance and providing proactive interventions are a critical component for the safety, performance, and well-being of workers.

Chapter 2: Review of Gaps in Vigilance Decrement Research

Abstract

This chapter aims to explore the underlying causes of vigilance decrement, methods for detecting and predicting vigilance decrement, and intervention designs that can mitigate the effects of vigilance decrement. The gaps in these areas of vigilance research are then used to propose the four research questions to be addressed in this dissertation.

2.1 Theoretical Mechanisms of Vigilance Decrement

Two dominant theories have been used to explain the mechanism of vigilance decrement: the overload and underload theories. The overload theory assumes that being vigilant consumes the limited resources available for information processing, gradually depleting cognitive resources over time (Davies & Parasuraman, 1982; Neigel et al., 2020). The increased demand for working memory is interrelated with sustained attention, particularly during tasks requiring detection of infrequent signals, and thus results in vigilance decrement (Helton & Russell, 2015; Murray & Amaya, 2024). Brain imaging studies support such explanations by showing increased cerebral blood flow in the prefrontal and anterior cingulate cortex during both working memory and vigilance tasks (D'Esposito et al., 1995; Helton & Russell, 2015; Lawrence et al., 2003).

The underload theory explains that reduced motivation, interests, and task-unrelated thoughts detract from attentional resources and cause vigilance decrement (Greenlee et al., 2020; Neigel et al., 2020; Thomson et al., 2015). Furthermore, it suggests that vigilance decrement arises from the under-stimulating nature of sustained attention tasks, as repetitive and monotonous stimuli can reduce an individual's overall arousal (Gartenberg et al., 2018). Thus, when arousal levels decline, the ability to effectively process information and detect critical signals diminishes, resulting in the vigilance decrement (Gartenberg et al., 2018). Both theories claim that sustained attention is a key component of vigilance, wherein performance tends to decline over time rather than in a single instance (Murray & Amaya, 2024).

2.2 Behavioral Indicators of Vigilance Decrement

A majority of studies measure vigilance using behavioral responses during sustained attention tasks (e.g., Hancock, 2013; Meeker et al., 2021). These tasks often require participants to monitor a stream of information over extended periods, detecting target stimuli embedded within a series of non-target stimuli. Some common types of vigilance tasks include the psychomotor vigilance task (PVT) and continuous performance test (CPT). The PVT is a reaction-time task that requires participants to respond as quickly as possible to a visual stimulus (e.g., a shape appearing on a screen) that occurs at unpredictable intervals (e.g., Drummond et al., 2005). CPTs typically involve the rapid presentation of letters or numbers, with participants instructed to respond only to a specific target stimulus (e.g., Bearden et al., 2004). Outside of traditional vigilance tasks, some researchers simulate real-world tasks by employing more complex scenarios that mimic real-world vigilance demands, such as simulated air traffic control or baggage screening tasks (e.g., Meuter & Lacherez, 2015).

Performance during these tasks is typically evaluated using several key metrics that indicate vigilance decrement, including response time, response time variability, and signal/stimuli detection (Mackworth, 1948; Hancock, 2013). Many vigilance studies use Signal Detection Theory (SDT) to provide a more nuanced understanding of performance changes. SDT allows researchers to quantify participants' ability to discriminate between signal and noise

(Stanislaw & Todorov, 1999). Key behavioral metrics for vigilance decrement are summarized in *Table 1* below.

Table 1. Common Behavioral Metrics for Vigilance

Behavioral Metric	Description	Relationship with Vigilance
Response time (RT)	Time elapsed between the stimulus presentation and the individual's reaction to the target.	As vigilance decreases, individuals tend to react more slowly to target stimuli.
RT variability	The spread of RTs throughout the task.	As vigilance decreases, RTs become less consistent and more erratic over time.
Hit rate	The proportion of target signals that were correctly detected.	Vigilance decrement is exhibited by a decreased ability to accurately identify target signals, leading to lower hit rates.
Miss rate	The proportion of target signals that were not detected.	Conversely, individuals have higher miss rates during vigilance decrement.
False alarm rate	The proportion of non-target stimuli identified as targets.	Overall accuracy on the task tends to decline as sustained attention lapses.

While behavioral measures can accurately indicate vigilance decrement, analysis consisting of only behavioral responses is limited to post-hoc detection. Thus, it is important to also consider contextual factors and physiological indicators, allowing for more robust, continuous, and objective evaluation of vigilance decrement.

2.3 Contextual Factors That Influence Vigilance Decrement

Contextual/contributing factors refer to the relatively stable characteristics and situational conditions that can significantly influence an individual's ability to sustain attention over time. Contextual information can be useful for modeling vigilance due to the significant effect some factors can have on an individual's ability to sustain their attention. Demographic information such as age, gender, and socioeconomic background can contribute to individual differences in baseline vigilance levels and the rate at which vigilance decrement occurs. For instance, several studies have found age-related declines in sustained attention abilities (Rose et al., 2010) and differences in vigilance performance between genders (Blatter et al., 2006). Similarly, factors related to professional background, such as years of work experience or job duties, might shape an individual's resilience during vigilance tasks (Zare et al., 2024).

Furthermore, sleep history is a critical determinant of vigilance performance. A substantial body of research has demonstrated that acute or chronic sleep restriction leads to significant impairments in sustained attention and accelerates vigilance decrement during the day (e.g., Blatter et al., 2006; Chua et al., 2012). In safety-critical environments, such as hospitals, sleep deprivation is known to be a major concern. Several studies have examined sleep deprivation and its impact on vigilance but mainly rely exclusively on self-reported questionnaires such as the Pittsburgh Sleep Quality Index (PSQI; Buyesse et al., 1989). A range of other contextual and contributing factors can influence vigilance. High levels of stress and anxiety can impair one's ability to sustain their attention (e.g., Canisius & Penzel, 2007; Meeker et al., 2021). Aspects of the environment, such as ambient noise, temperature, and lighting, can either facilitate or hinder sustained attention (e.g., Szalma & Hancock, 2011). The motivation and engagement an individual have with their task can also significantly impact their ability to maintain focus (e.g., Johnson & Kim, 2020).

Vigilance-related studies rarely account for the complex interplay of contextual factors (e.g., age, gender, sleep history) with other measures of vigilance. Typically, these factors are only considered singularly—as a condition (e.g., comparing vigilance performance after different amounts of sleep deprivation) or inclusion criteria (e.g., Abe et al., 2020)—rather than a potential series of predictors that could explain individual variability in vigilance decrement. Recognizing and incorporating these broader contextual influences in research studies could lead to a more comprehensive understanding of vigilance decrement in safety-critical situations like driving, clinical care, and security surveillance. This gap in the literature leads to *Research Question #1: What contextual factors, together, affect vigilance levels in safety-critical settings?*

2.4 Physiological Indicators of Vigilance Decrement

Physiological indicators refer to measures of individuals' physiological recordings. These measures offer objective and continuous data for assessing vigilance decrement through non-invasive measurements (Andreassi, 2010; Dirican & Göktürk, 2011). The different types of physiological indicators of vigilance decrement can be broadly categorized by the physiological systems they reflect: the central nervous system and the peripheral nervous system.

2.4.1 Neural Responses

Electroencephalography (EEG) is a key neurophysiological technique that measures the brain's electrical activity via scalp electrodes. It captures voltage fluctuations resulting from ionic currents within neurons, revealing distinct frequency bands (delta, theta, alpha, beta) that correlate with varying states of brain activity (Teplan, 2002). With its high temporal resolution, EEGs can track rapid shifts in brain dynamics, providing insights into an individual's alertness and attentional state (e.g., Kamzanova et al., 2014). Specific brainwave patterns, particularly changes in alpha and theta power, often indicate reduced alertness and increased drowsiness associated with vigilance decrement (e.g., Subasi, 2005). Functional magnetic resonance imaging (fMRI) and functional near-infrared spectroscopy (fNIRS) are other techniques that capture hemodynamic changes and have been shown to correlate with changes in vigilance. Studies using fMRI have shown that vigilance decrement is often associated with altered activity in regions like the prefrontal cortex, parietal cortex, and thalamus, as well as changes in the connectivity between these networks (Hinds et al., 2013). fNIRS can reveal alterations in cortical activity, such as decreased prefrontal cortex activation while vigilance decreases over time (Nogueira et al., 2022).

2.4.2 Physiological Responses

Several physiological measures reflect the activity of the peripheral nervous system, particularly the autonomic nervous system, which plays a significant role in regulating arousal (physiological and psychological states of alertness or wakefulness (Robertson & O'Connell, 2010; Thayer, 1989) and physiological responses associated with vigilance. As arousal levels rise, norepinephrine produced by the locus coeruleus in the brain stem inhibits the parasympathetic nervous system, resulting in increased heart rate (HR), higher blood pressure, and increased pupil dilation (Yung et al., 2020). Heart rate (HR), the number of heart beats per minute, is another indicator reflecting overall physiological arousal (Patel et al., 2011). HR variability (HRV), the variation in the time intervals between heartbeats, is highlighted as a possible indicator of vigilance decrement but with mixed findings. Schuurmans and colleagues (2020) found that with increased R-R interval variability being associated with lower vigilance while another study showed that higher HRV is associated with vagus nerve activation and therefore higher vigilance (Wang et al., 2023).

Respiration rate and pattern, measuring breaths per minute and the regularity of breathing, can also indicate fatigue, stress, or changes in arousal levels that impact vigilance (Pattyn et al., 2008). Electrodermal activity (EDA), also known as skin conductance or galvanic skin response, measures the skin's electrical conductivity, is primarily influenced by sweat gland activity driven by sympathetic nervous system arousal (Distefano et al., 2020). Body or skin temperature is another peripheral measure, with significant deviations from baseline potentially indicating physiological states affecting vigilance; specifically, a decrease in skin temperature is noted as vigilance decreases (Romeijn & Van Someren, 2011).

2.4.3 Eye-Related Measures

Eye movement, while involving sensory input and motor output, is also heavily influenced by central cognitive processes related to attention. These techniques utilize specialized sensors or cameras to infer attentional state (Lavine et al., 2002) by continuously recording parameters such as saccades (rapid eye movements between fixations), fixations (periods of sustained gaze), and blink rate (frequency of eyelid closure). Changes in saccadic velocity and amplitude are known to decrease as vigilance declines (Bodala et al., 2016), while fixation duration also tends to increase with reduced alertness (Lavine et al., 2002). Pupil dilation, controlled by autonomic influences, is also a relevant measure, with pupil diameter typically decreasing as vigilance diminishes (McIntire et al., 2014).

The development of portable and unobtrusive wearable physiological sensing techniques has enabled the monitoring of these peripheral indicators, such as HR, skin temperature, EDA, pupil dilation, and eye movement, in real-world settings (Andreassi, 2010). Several of these innovative sensing techniques are able to be used in real-world environments for monitoring due to their portability and relative unobtrusiveness (Matthews et al., 2007). However, a large majority of vigilance measurement studies consider only one physiological indicator at a time. Despite the technological advances in wearable physiological sensors, there can still be issues in accurate predictive modeling due to these sensors' low signal-to-noise ratio. Using multiple physiological indicators can help to combat this issue in vigilance measurement by adding several signals and filling in gaps when one sensor experiences lapses in continuous data collection. Thus, I propose *Research Question #2: What combination of physiological measures can be used to monitor vigilance?*

2.5 Analytical Methods for Monitoring and Detecting Vigilance Decrement

The analytical approaches employed to monitor and detect vigilance decrement span a spectrum of complexity, each with its own strengths and limitations. Fundamental statistical tests, such as t-tests and Analysis of Variance (ANOVA), provide straightforward methods for comparing group means and identifying significant differences in performance metrics under various conditions or across time. These tests are valuable for initial explorations and hypothesis testing regarding the impact of specific factors on vigilance. For example, Hu & Lodewijks (2021) used paired t-tests to explore the effects of task-related fatigue on vigilance performance and eye movement in driving scenarios. Inferential analyses, particularly regression-based techniques like linear and logistic regression, allow for the modeling of relationships between predictor variables (e.g., sleep deprivation) and vigilance measures (e.g., task performance), enabling the quantification of these relationships and the prediction of future performance.

Generalized linear models were used by Jones and colleagues (2024) to assess relationships between vigilance task events and several brain regions' activity via fMRI data.

More sophisticated techniques, such as probabilistic and stochastic modeling, offer frameworks for understanding the underlying probabilistic dependencies between variables and for making inferences under uncertainty. Bayesian networks, for example, use conditional dependencies to graphically present causal relationships between variables. Other machine learning (ML) models, such as random forests and neural networks, represent a class of data-driven algorithms capable of learning complex, nonlinear patterns from high-dimensional datasets (Jordan & Mitchell, 2015). These models excel at prediction and classification tasks, often achieving higher accuracy than traditional statistical methods, especially when dealing with intricate interactions between multiple factors. A study conducted by Li and colleagues (2022) used several ML strategies (e.g., neural networks, support vector machines, decision trees) to assess vigilance in air traffic controllers from eye-tracker data. Like many other vigilance studies that consider ML models, the training data is restricted to just one type of sensor rather than multiple sensors and/or additional contextual factors.

Despite the wide range of analytical methods available, many vigilance studies still tend to focus on simpler statistical comparisons of group means or basic correlational analyses. Relatively few studies, if any, leverage the power of advanced ML algorithms to integrate behavioral, contextual, and physiological factors into comprehensive predictive models of vigilance decrement. Utilizing these more sophisticated analytical approaches holds significant potential for advancing our understanding and ability to predict vigilance decrement. Therefore, I propose *Research Question #3: What modeling features and strategies can be used to predict vigilance decrement?*

2.6 Vigilance Decrement Intervention Design

Several intervention-based approaches for mitigating vigilance decrement have been identified by researchers. Breaks (Bergum & Lehr, 1962; Ross et al., 2014) and changes in types of tasks (e.g., from a visual task to an auditory task) (Greenlee et al., 2022) have been considered effective approaches for handling vigilance decrement. Esterman and colleagues (2016) attempted to counter vigilance decrement through incentive-based motivation. In their study, participants were told their participation compensation would increase or decrease based on their task performance (e.g., lose money for errors). Feedback on operators' performance (Hancock et al., 2016) is likewise recognized as a systematic way to mitigate vigilance decrement in future tasks. Another intervention strategy targeted at the task and/or task environment is changes to lighting. For example, different colors and temperature of lighting, such as blue light, has been used in vigilance studies to boost attention (Ni et al., 2024; Yuda et al., 2017). The illuminance of ambient lighting has also been studied, with Ni and colleagues (2024) finding that higher illuminance improves sustained attention during driving tasks.

More radical, yet interruptive, methods have been proposed, such as the use of transcranial direct current stimulation (tDCS) on targeted brain regions (Al-Shargie et al., 2019; Nelson et al., 2014) to enhance vigilance. In tDCS applications, well-tolerated electrical current is applied to the brain through scalp electrodes. Several other forms of electrical stimulation have been studied, including transcranial alternating current stimulation, transcranial random noise stimulation, and transcranial magnetic stimulation. Transcranial alternating current stimulation (tACS) is similar to tDCS but instead uses sinusoidal alternating electric currents to affect cortical neurons (Löffler et al., 2018; Rostami et al., 2021). Transcranial random noise stimulation (tRNS) applies weak electrical current oscillating at random frequencies to the scalp for brain stimulation (Harty & Cohen Kadosh, 2019). Transcranial magnetic stimulation (TMS) differs from tDCS, tACS, and tRNS by using magnetic fields, rather than electric current, to stimulate nerve cells in the brain (Esterman et al., 2017).

A third group of interventions is more concerned with the overall health of the individual and therefore implements more accessible, repeatable, and longer lasting strategies. Luo and colleagues (2021) studied the effects of tactile attention training on vigilance. This training involved using an adaptive fingertip device for 40 minutes a day across 5 consecutive days and showed promising results. Light to moderate exercise has also been used to prevent vigilance decrement by stimulating the body. For example, Sanchis and colleagues (2020) found that elevating heart rate through indoor exercise biking enhanced performance during vigilance tasks. Likewise, Axelsen and colleagues (2022) implemented daily app-based mindfulness and meditation exercises to improve individuals' vigilance. Ingestion of particular nutrients in carbohydrates and fruits has had positive effects on sustained attention. One study provided participants with tart cherry juice containing flavonoids to promote vigilance (Chai et al., 2019). Similarly, caffeine ingestion has been shown to improve vigilance task performance (Cintineo et al., 2022).

Despite the importance of monitoring vigilance decrement for proper intervention, when and how to use such interventions, and their long-term effects, remains unclear. The lack of knowledge related to the use of practical, sustainable interventions for vigilance decrement stands in stark contrast to the solid theoretical research explaining the mechanism behind vigilance decrement. There are few evaluations in existing research on the sustainability of vigilance interventions. This contrasts with other fields of research, such as public health, where there is growing interest in understanding and evaluating the sustainability of interventions

(Hailemariam et al., 2019; Walugembe et al., 2019). Considering the sustainability of interventions not only addresses long-term effects of vigilance decrement but accounts for constraints such as cost and general usability. This research gap leads to *Research Question #4: What types of interventions can be used to maintain vigilance in a sustainable manner?*

2.7 Research Questions by Chapter and Dissertation Process

The research questions proposed in this dissertation are organized by chapter in *Table 2* below. *Chapter 3* describes a literature review on sensor-based methods for measuring vigilance and is used to answer RQ 2 while providing viable measures for the following chapters. *Chapter 4* describes the methods for developing a vigilance modeling framework using a partial healthcare-related dataset and provides initial answers to RQs 1–3. *Chapter 5* describes the validation of the vigilance modeling framework using the full healthcare dataset and a naturalistic healthcare task to further answer RQs 1–3. *Chapter 6* describes a scoping review on sustainable interventions for vigilance decrement and is used to answer RQ 4. *Chapter 7* applies the vigilance modeling framework to the driving domain and incorporates a sustainable intervention design to promote better vigilance. This chapter answers all four research questions. *Chapter 8* provides conclusions to the dissertation.

Table 2. Research Questions by Chapter

Research Question	**Chapter 3:** Sensor-Based Methods for Measuring Vigilance	**Chapter 4:** Vigilance Modeling Framework	**Chapter 5:** Vigilance Modeling in Healthcare	**Chapter 6:** Sustainable Interventions for Vigilance Decrement	**Chapter 7:** Vigilance Modeling & Interventions in Driving
1. What contextual factors, together, affect vigilance levels in safety-critical settings?		✓	✓		✓
2. What combination of physiological measures can be used to monitor vigilance?	✓	✓	✓		✓
3. What modeling features and strategies can be used to predict vigilance decrement?		✓	✓		✓
4. What types of interventions can be used to maintain vigilance in a sustainable manner?				✓	✓

Figure 1 is a flowchart illustrating the process by which this dissertation addresses each of the four research questions (RQs). The literature review of contextual factors and physiological measures in this chapter, *Chapter 2*, provides baseline answers to RQ 1 and RQ 2. Due to the expansive research conducted on physiological sensors, *Chapter 3* dives deeper into the sensors that can be leveraged outside of lab settings and in safety-critical environments to further answer RQ 2. The findings in *Chapter 2* and *Chapter 3* are then used to inform the design of the vigilance in healthcare study. This study is described in *Chapter 4*, where a partial form of the dataset is used to develop a structured vigilance modeling framework that can be applied in multiple vigilance application domains. The findings from *Chapter 4*, provide initial answers to

RQs 1–3 and the framework for fully addressing the questions in the subsequent chapters. The full dataset from the healthcare study is assessed in *Chapter 5* to answer RQs 1–3 for the healthcare domain. In *Chapter 6*, a literature review is conducted on interventions for vigilance decrement. This review focuses on sustainable interventions that can be implemented in real-world vigilance applications and provides answers to RQ 4. Finally, *Chapter 7* combines the vigilance modeling framework used for the healthcare domain with vigilance intervention design and applies these methods to the driving domain using a driving simulator study. The study focuses on vigilance during semi-autonomous driving, and its findings provide answers to RQs 1–4 for the driving domain.

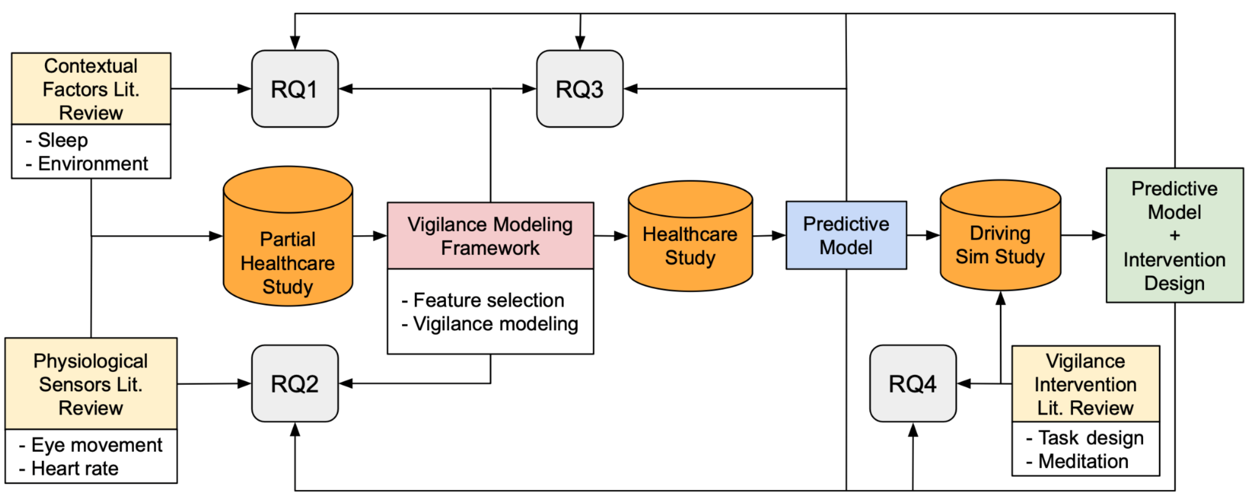

Figure 1. Process Flow of Dissertation Work

Chapter 3: Sensor-Based Methods for Measuring Vigilance

Abstract

This chapter identifies promising sensor-based methods for measuring vigilance. It features a structured review of empirical studies using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method. The types of sensors are organized into electrophysiological, peripheral physiological, and eye-tracking measures. Each sensor was assessed for validity and readiness in naturalistic settings based on its temporal resolution, cost, the number of available metrics associated with the sensor, and the strength of support in the literature. This chapter is used to answer RQ 2.

3.1 Background and Motivation

Attention-requiring tasks that are susceptible to vigilance decrement often require participants to monitor a stream of information over extended periods, detecting target stimuli embedded within a series of nontarget stimuli (see *Chapter 2*). Performance and behavior during vigilance tasks is typically used to evaluate vigilance decrement. While behavioral responses can effectively indicate declines in vigilance-related performance, relying solely on such measures has limitations, particularly outside of laboratory settings. In particular, behavioral assessments typically capture reflexive responses to specific attention-requiring tasks (Hancock, 2013), making their findings less generalizable across different task types and environments (Riva et al., 2003). Moreover, behavioral data often involve low event rates and are analyzed retrospectively (Riva et al., 2003), hindering their ability to support continuous, real-time monitoring and predictive modeling. This is especially problematic given vigilance decrement's concept of performance deterioration over extended periods (Mackworth, 1948). Therefore, developing and using effective sensor-based tools to accurately and continuously measure vigilance decrement across various work settings is critical to ensuring the safety and well-being of individuals.

3.2 Physiological Responses Indicating Vigilance Decrement

Neuroscientific evidence indicates that the integrity of the locus coeruleus (LC)–norepinephrine (NE) system, among other systems, plays a critical role in maintaining optimal vigilance and performance levels (Martin et al., 2022; Mather & Harley, 2016). The LC is a small nucleus located at the rostral pons and is part of a larger brainstem network (Maness et al., 2022). Changes in the firing rates of cholinergic and noradrenergic neurons are key to state transitions, as observed through electroencephalography (EEG) and behavioral changes associated with vigilance (Ross & Van Bockstaele, 2021; Steriade et al., 1989). The LC synthesizes NE, which plays a central role in modulating arousal and regulating physiological functions (Mather & Harley, 2016). Given that NE modulates arousal levels, heart rate, blood pressure, salivary function, electrodermal activity, pupil size, and blink rate, physiological correlates of vigilance levels have been identified in prior empirical studies. In a state of heightened arousal, increased LC activation elevates the NE level, which in turn inhibits parasympathetic activity (Esler et al., 1985; Pajcin et al., 2019), resulting in elevated heart rate, increased blood pressure, pupil dilation, lower body temperature, and decreased eye blink rate (Coimbra et al., 2019; Craig & Klein, 2019; Gabay et al., 2011; Terkelsen et al., 2013; Yung et al., 2020).

These shifts in brain activity and neurotransmitter dynamics provide the foundation for using sensors that collect neural, physiological, and eye-related responses in order to continuously and objectively assess and monitor vigilance states (Andreassi, 2010; Dirican & Göktürk, 2011). The noninvasive nature of such sensors gives researchers the opportunity to measure vigilance while participants engage in tasks without pausing and being interrupted during main tasks. At the same time, sensor-based responses cannot be controlled by participants, removing any subjectivity from the outcomes. Technological advances in sensor equipment have made it possible for researchers to continuously collect and observe data in real time, providing insights into temporal changes in vigilance. The three main types of sensor-based indicators of vigilance are 1) central physiological responses related to brain activity controlled by the central nervous system, 2) peripheral physiological responses regulated by the autonomic nervous system, which controls involuntary biological functions, and 3) eye-related responses (Schmälzle & Grall, 2020; Wessa et al., 2005).

3.2.1 Central Physiological Responses

Central physiological responses can be measured using two approaches: 1) electrophysiological techniques and 2) hemodynamic techniques (Mehta & Parasuraman, 2013; Shibasaki, 2008). Electrophysiological techniques capture changes in electrical potential in the brain in response to neuronal activity. They offer high temporal resolution, providing a granular picture of how brain activity changes over time (Salmelin & Baillet, 2009). Hemodynamic techniques capture changes in the brain's blood flow in response to neuronal activity. Hemodynamic techniques prioritize spatial resolution compared to electromagnetic techniques, revealing the specific brain regions being activated during vigilance tasks (Sheth et al., 2005).

A key electrophysiological technique for measuring vigilance are EEGs which capture voltage fluctuations resulting from ionic currents within the brain's neurons. Continuous EEG signals are often analyzed in terms of their frequency and amplitude. The EEG frequency spectrum is subdivided into delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (> 30 Hz) bands that correlate with varying states of brain activity using a Fourier transform (Teplan, 2002). Given its high temporal resolution, EEG can track rapid shifts in brain dynamics, providing insights into an individual's alertness and attentional state (Kamzanova et al., 2014). In particular, changes in alpha- and theta-band power have been found to indicate reduced alertness and vigilance (Subasi, 2005). EEG connectivity patterns consider interrelated activity among multiple brain sites rather than isolated brain regions and have been used to infer vigilance states (Shi et al., 2013).

Event-related potentials (ERPs) are discrete EEG measures that rely on amplitude and the latency of brain responses to specific events or stimuli (e.g., Blackwood & Muir, 1990; Luck et al., 2000). The P300 (P3 wave) component is a positive wave occurring about 300 milliseconds after a significant or unexpected stimulus (Picton, 1992) and is sensitive to changes in vigilance (Koelega et al., 1992; Schmidt et al., 2009). Similarly, the P50 component, a positive wave occurring about 50 milliseconds post-stimulus (Guterman et al., 1992), has been found to change with vigilance levels (White & Yee, 2006). The N100 (N1 wave) and N200 (N2 wave) are negative waves that occur around 100 and 200 milliseconds (respectively) after event stimuli and have been reported to be correlated with vigilance (Kaarre et al., 2018; Patel & Azzam, 2005; Samima et al., 2017).

There are a broad range of techniques that capture hemodynamic changes in the brain and correspond to states of vigilance. Studies using functional magnetic resonance imaging (fMRI), which detects relative blood oxygenation to capture cerebral activity using magnetic resonance imaging (Kim & Lee, 2020; Mehta & Parasuraman, 2013), have shown that vigilance decrement is often associated with altered activity as well as changes in the connectivity of brain networks (Hinds et al., 2013; Kim et al., 2017; Liu & Falahpour, 2020). In general, fMRI measures the blood oxygen level-dependent (BOLD) signal (Arthurs & Boniface, 2002; Steinbrink et al., 2006), which has been found to significantly correlate with vigilance (Olbrich et al., 2009; Wong et al., 2013) and several peripheral physiological responses, including heart rate (Shmueli et al., 2007) and respiration rate (Tong & Frederick, 2014). Functional connectivity using fMRI is the simultaneous activity patterns of distinct brain regions during tasks (Rogers et al., 2007; Van Den Heuvel & Pol, 2010) and has been associated with vigilance (Patanaik et al., 2018; Wang et al., 2016). One exemplary network in fMRI is the default mode network, which comprises the posterior cingulate cortex, medial prefrontal cortex, and the lateral parietal cortex (Gui et al., 2015; Zhang and Raichle, 2010) being active during mind wandering (Fox et al., 2015; Sood & Jones, 2013). Similar to fMRI, functional near-infrared spectroscopy (fNIRS) measures oxygenated and deoxygenated blood, but it uses near-infrared light in cortical regions to infer regional brain activity (Bogler et al., 2014; Mehta & Parasuraman, 2013). fNIRS studies have found that reduced prefrontal cortex activation (Nogueira et al., 2022) and global signals (Chen et al., 2017) are correlated with decreased vigilance over time.

3.2.2 Peripheral Physiological Responses

Electrocardiograms (ECG) can be used to measure the electrical activity of the heart and is reportedly sensitive to changes in vigilance (Charles & Nixon, 2019; Patel et al., 2011). ECG signals capture the repeating P-Q-R-S-T cardiac cycle, which reflects the heart's polarization and depolarization process (Naseer & Nazeer, 2017). Heart rate variability, or the variation between consecutive heartbeats, has been identified as a key physiological indicator of vigilance decrement. Heart rate variability can be analyzed in both the time and frequency domains (Charles & Nixon, 2019; Patel et al., 2011). In the time domain, variability in R-R intervals is known to be sensitive to vigilance levels (Chua et al., 2012; Schuurmans et al., 2020). In frequency-based analysis, the ratio of low-frequency (0.04–0.15 Hz) to high-frequency (0.15–0.4 Hz) heart rate variability has been used as an indicator of vigilance levels (Batchinsky et al., 2007; Patel et al., 2011).

In addition to ECG, other types of peripheral physiological responses—respiration rate, electrodermal activity, and body or skin temperature—have been linked to changes in vigilance. Respiration rate and pattern, which measure breaths per minute and the regularity of breathing, respectively, indicate changes in arousal levels that impact vigilance (Pattyn et al., 2008). Electrodermal activity, often referred to as skin conductance or galvanic skin response, measures sweat gland activity (Distefano et al., 2020) and has been associated with fluctuations in vigilance (Canisius & Penzel, 2007; Dorrian et al., 2008). Body or skin temperature is another peripheral measure, with significant deviations from baseline potentially indicating physiological states affecting vigilance (Romeijn & Van Someren, 2011). Alpha amylase, a consistent biomarker for sympathetic nervous system activity, is present in saliva and can also be measured to determine vigilance levels (Yoon & Weierich, 2016).

3.2.3 Eye-Related Responses

Eye movements are heavily influenced by central cognitive processes related to attention (Sheliga et al., 1994; Theeuwes et al., 2009). To measure eye-related responses, various eye-tracking techniques rely on videooculography. This method typically utilizes noninvasive monitor-mounted or wearable infrared cameras to capture glints on the cornea and pupil shifts

(Chernorizov et al., 2016). Alternatively, electrooculography (EOG) can be used to measure the electrical activity of the eye muscle through the placement of electrodes that capture eye blink-related measures (Chernorizov et al., 2016; Huang et al., 2018).

The parameters computed based on eye movements include eye fixations, periods of sustained gaze on one area (defined as more than 100–400 milliseconds; Rayner, 1998), and saccades (eye movements between fixations). Fixation duration (Lavine et al., 2002), saccadic velocity and saccadic amplitude (Bodala et al., 2016) are known to be related to vigilance declines. Eye blink activity, such as blink rate, is another common metric that has been used to capture changes in cognitive states, such as distraction (Annerer-Walcher et al., 2020) and fatigue (Stern et al., 1994). Involuntary blinks have been found to be elicited by supraorbital nerve stimulation (Bour et al., 2000), as this stimulation from the brain activates the orbicularis oculi muscles in the eye. Similarly, the NC–LE neuromodulatory system projects directly to the superior colliculus, showing that arousal can have an effect on pupil size (Joshi & Gold, 2020). Several studies have found a significant relationship between pupil size and vigilance (e.g., Martin et al., 2022; van den Brink et al., 2016).

3.3 Research Goal

This chapter aims to identify promising sensor-based vigilance measures to facilitate the translation of vigilance research from laboratory settings to real-world field environments. Since Mackworth (1948)'s early empirical observations of vigilance decrement, there have been few significant advancements in the measurement of vigilance that can account for real-world work settings and operator characteristics. Most sensor-based vigilance research remains grounded in biological perspectives and is typically conducted in a controlled laboratory environment.

Central physiological sensors have been used extensively in the literature to measure and evaluate vigilance. However, despite their capabilities in lab settings, these sensors are not suitable for applications in uncontrolled settings due to their physical constraints on portability/mobility and the drop in data quality during naturalistic tasks requiring movement (e.g., Havsteen et al., 2017; Mehta & Parasuraman, 2013). Thus, this review focuses only on the other two types of sensors for vigilance: peripheral physiological and eye tracking sensors, and their respective measures of vigilance. As such, the term "physiological" will be used to refer solely to peripheral physiological sensors.

3.4 Methods

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method was used to support the thorough analysis of relevant peer-reviewed literature. To obtain and draw results from the literature, a four-step process was followed: 1) identification of research related to sensor-based vigilance measurement using search strings and exclusion criteria, 2) screening of the literature based on titles and abstracts, 3) eligibility determination based on full-text review, and 4) description of key findings from all relevant papers. *Figure 2* shows a flowchart of the PRISMA literature review process.

In the identification phase, four databases—Web of Science, ProQuest, ScienceDirect, and PsycInfo—were used, as they contain peer-reviewed work related to vigilance (e.g., psychology and human factors engineering). Each database's literature search used a common search string based on keywords related to vigilance (i.e., "vigilance OR sustained attention") and physiological or neurological aspects (i.e., "physiolog OR neuro OR psychophysiolog"). Although neurological and central physiological sensors were not considered in the review, they were kept in the search terms to maintain studies that used these sensors along with peripheral physiological or eye tracking sensors. Four trained researchers conducted the literature searches in the four databases, with each researcher being assigned one database. In the screening phase, the researchers reviewed the initial results based on the paper titles and/or abstracts. After the preliminary results were screened, the references were saved and uploaded to a shared reference management library, where duplicates were removed and a separate folder was created for each database's results. There were 372 papers after the initial review.

Following the screening of titles and abstracts, a full-text PDF was added to each result in order to facilitate the full review. Each paper was then critically reviewed by the research team using the exclusion criteria. Studies were only included if they were specific to vigilance or sustained attention, involved physiological sensors in experiments, used behavioral measures of vigilance, and contained significant results with respect to vigilance. Selected studies involving human subjects aged 18 to 65 were considered for generalizability (e.g., Gerber et al., 2023; Little et al., 2024). Since the goal of this review was to highlight contemporary sensor-based strategies, only studies published after 2015 were included to focus on recently published studies (Prendez et al., 2026). Studies were excluded if they were not available in English, if the full text was unavailable, or if they were dissertations, meta-analyses, or not published in peer-reviewed journals.

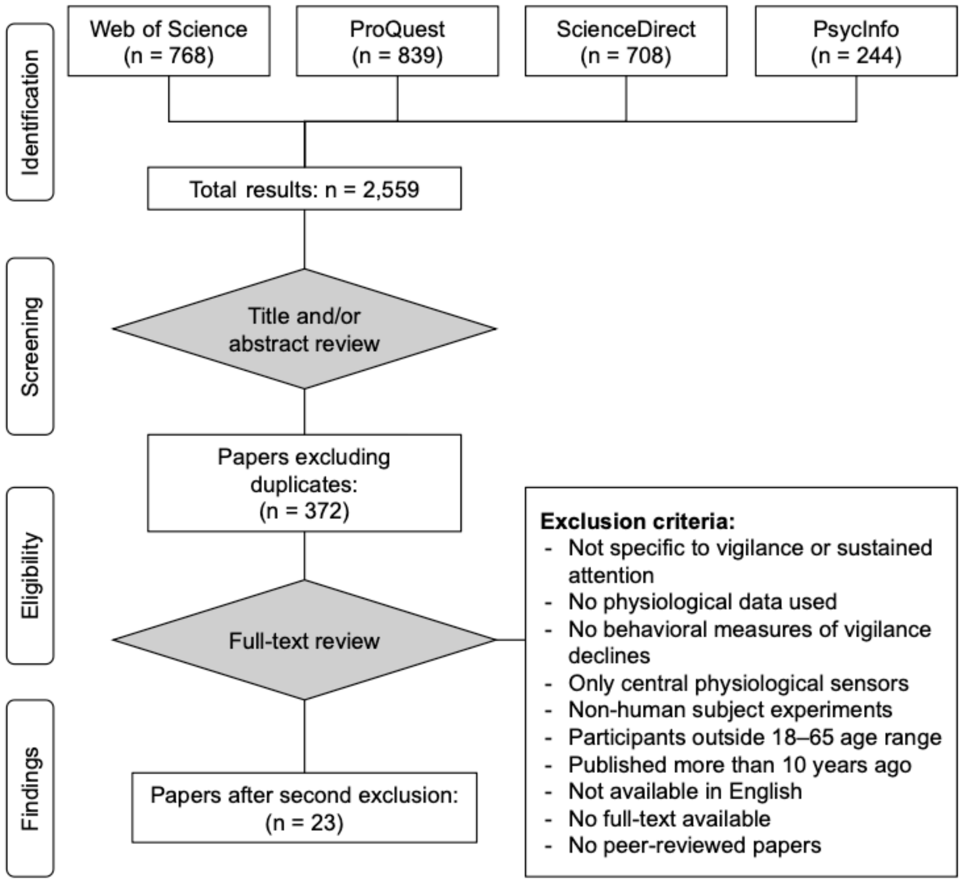

Figure 2. PRISMA Review Process for Vigilance Sensors

3.5 Results

A total of 23 studies were included in the final full-text review. *Table 3* provides a summary of participant demographics, task types, application domains, vigilance measurement modalities, and independent variables. The reviewed studies had sample sizes ranging from 9 to 60 participants. In designing tasks to measure vigilance, most studies employed standardized tests. Of the 11 studies (48%) that utilized standardized vigilance tests, seven used the PVT, three used the CPT (or its variants), and one used the MCT test. The second most frequently used task type was driving-related tasks (5 studies, 22%), which required participants to maintain the lane position, respond to the sudden braking of a lead vehicle, and adjust the speed of the vehicle based on system cues. The remaining seven studies employed a variety of perceptual and cognitive tests that required participants to listen to auditory stimuli, search for visual targets, detect changes, calculate numbers, identify congruent and incongruent stimuli, and conduct multitasking. One study aimed to mimic vigilance in warehouse security operations and implemented a simulation-based intruder identification task with varying levels of difficulty (e.g., adding rain to obscure potential intruders). In two aerial vehicle research studies, typical air traffic controller operations and collision flight path detection were used as vigilance tasks. The application domains for the reviewed studies were limited and lacked diversity. The only attributable application domain was transportation, primarily focusing on in-vehicle driving contexts. Seven of these studies utilized driving simulators in laboratory settings, while two studies were implemented in naturalistic driving conditions. One used an instrumented vehicle, equipped with an eye-tracker, for driving on a highway while the other took place within heavy trucks on a closed-loop test track. Regarding the types of sensors used for assessing vigilance, eye tracking sensors were the most frequently employed (15 studies, 65%). Physiological sensors were used by seven studies (30%), while the one remaining study used both eye tracking and physiological sensors.

Table 3. List of Reviewed Papers for Sensor-based Vigilance Measures

Article (author, year)	**Participants** (*n*)	**Task**	**Domain Application**	**Measures**
Abe et al. (2020)	16	PVT	N/A	Eye tracking
Aeschbach et al. (2024)	9	PVT	N/A	Eye tracking
Bodala et al. (2016)	12	Surveillance task	N/A	Eye tracking
Di Flumeri et al. (2019)	14	Air traffic control	Transportation	Eye tracking
Funke et al. (2017)	36	Unmanned aerial vehicle control task	Transportation	Eye tracking
Guo et al. (2024)	30	Driving task	Transportation	Physiological, eye tracking
He et al. (2023)	32	Driving task	Transportation	Eye tracking
Hu & Lodewijks (2021)	60	Driving task, number calculation	Transportation	Eye tracking
Li et al. (2019)	20	Visual cue task	Transportation	Eye tracking
Li et al. (2022)	20	Reaction time task	Transportation	Eye tracking
Liu et al. (2024)	18	Driving task	Transportation	Eye tracking
Maccora et al. (2019)	18	PVT	N/A	Eye tracking
Maffei & Angrilli (2018)	33	Mackworth Clock Test	N/A	Eye tracking
McGarry et al. (2023)	32	Sensory vigilance task	N/A	Physiological
Molina et al. (2019)	17	PVT	N/A	Physiological
Pajcin et al. (2019)	23	PVT	N/A	Physiological
Pan et al. (2023)	30	Driving task	Transportation	Physiological
Posada Quintero et al. (2018)	25	PVT	N/A	Physiological
Sciaraffa et al. (2021)	12	Visual search	N/A	Eye tracking
van den Brink et al. (2016)	28	CPT	N/A	Eye tracking
Vergara et al. (2017)	19	CPT, Flanker task, Counting task	N/A	Physiological
Vergara et al. (2019)	34	CPT, Flanker task, Counting task	N/A	Physiological
Yamashita et al. (2021)	17	PVT	N/A	Eye tracking

3.5.1 Physiological Sensors

Among the eight studies that utilized physiological sensors, ECG-related measurements allowed for a wide range of features to be analyzed, specifically heart rate variability and frequency-based features, as summarized in *Table 4*. The "+" and "–" symbols under the table's *Relationship with Vigilance* column indicate the direction of the relationship between the measure and vigilance. Three studies utilized derived heart rate variability (HRV) measures as key indicators. In particular, the root mean square of successive differences (RMSSD) between R-peak intervals in the QRS-complex was found to significantly decrease with increasing vigilance levels (Guo et al., 2024; McGarry et al., 2023). Likewise, the average R-peak interval was observed to decrease as reaction times increased, indicating lower vigilance. Two studies presented contradictory results for the standard deviation of normal-to-normal intervals (SDNN) and the percentage of successive normal-to-normal intervals that differ by more than 50 milliseconds (pNN50). Pan and colleagues (2023) reported a positive relationship between both SDNN and pNN50 and vigilance level, whereas Guo and colleagues (2024) found a negative correlation for these same measures.

ECG frequency-based variables were also proven to be significant indicators of vigilance level. High-frequency signals exhibited a positive association with vigilance (Pan et al., 2023; Posada Quintero et al., 2018). Conversely, low-frequency signals showed both positive (Posada Quintero et al., 2018) and negative (Pan et al., 2023) with vigilance. The low-frequency to high-frequency ratio (LF/HF) and sample entropy (i.e., variability of the signal) increased with longer reaction times, reflecting lower vigilance (Pan et al., 2023).

Skin temperature measures also emerged as relevant physiological indicators of vigilance. Distal skin temperature was reported to exhibit both positive and negative associations with vigilance level. Two studies showed that higher distal temperature was associated with reduced vigilance (Molina et al., 2019; Vergara et al., 2019), whereas the other study reported the opposite trend (Vergara et al., 2017). Notably, Vergara et al. (2019) explicitly mentioned that this finding contradicted their previous results reported in Vergara et al. (2017), attributing that discrepancy to small sample size and differences across tasks. In addition, the difference between distal and proximal temperature values was used and demonstrated a negative relationship with vigilance (Molina et al., 2019).

Alpha amylase levels in saliva are controlled by the parasympathetic nervous system and showed a positive relationship with vigilance levels (Pajcin et al., 2019). High frequency signals (0.08 to 0.24 Hz) and high-amplitude skin conductance responses (> 0.05 µS) via EDA sensors also exhibited a positive correlation with vigilance levels (Posada Quintero et al., 2018).

Table 4. Physiological Sensor-based Measures of Vigilance

Measure	Variable	Relationship with Vigilance	References
Temporal ECG	RMSSD of R-peak intervals	–	Guo et al. (2024); McGarry et al. (2023)
	Average R-Peak interval	+	Pan et al. (2023)
	SDNN	+ –	Pan et al. (2023) Guo et al. (2024)
	pNN50	+ –	Pan et al. (2023) Guo et al. (2024)
Frequency ECG	High-frequency signals	+	Pan et al. (2023); Posada Quintero et al. (2018)
	Low-frequency signals	+ –	Posada Quintero et al. (2018) Pan et al. (2023)
	Low/high frequency ratio	–	Pan et al. (2023)
	Sample entropy	–	Pan et al. (2023)

Temperature	Distal skin temperature	+	Vergara et al. (2017)
		–	Molina et al. (2019); Vergara et al. (2019)
	Distal-to-proximal gradient skin temperature	–	Molina et al. (2019)
Salivary	Salivary alpha-amylase level	+	Pajcin et al. (2019)
EDA	High-frequency signals	+	Posada Quintero et al. (2018)
	Rate of high-amplitude skin conductance responses	+	Posada Quintero et al. (2018)

3.5.2 Eye Tracking

Sixteen papers utilized features captured by eye tracking to measure vigilance. Nine of these papers used eye movement-based measures (e.g., fixations, saccades) to determine vigilance levels. As summarized in *Table 5*, fixation duration showed a positive relationship with vigilance (Li et al., 2019), whereas fixation number showed a divergent pattern, being positively related to vigilance under low and moderate mental fatigue conditions and negatively related to vigilance under high mental high mental fatigue condition (Li et al., 2019). Time-to-First-Fixation (TTFF), defined as the time interval between the onset of a new target and the observer's first fixation on that target, has been shown to be negatively correlated with vigilance (Di Flumeri et al., 2019), based on the assumption that higher levels of attentiveness are associated with faster orientation toward newly introduced targets (Buck, 1966). Decreased saccade amplitude (measured by the difference in gaze position from the start to the end of a saccade) was consistently linked to reduced vigilance (Bodala et al., 2016; Li et al., 2022), while saccade duration, the time elapsed between the start and end of a saccade, was found to be negatively correlated with vigilance (Hu & Lodewijks, 2021). Li and colleagues (2019) found that the number of saccadic eye movements increased with higher vigilance under high mental fatigue conditions, whereas the saccadic number decreased with increasing vigilance under moderate mental fatigue conditions. Results for saccade velocity, measured as the average change in gaze position over time, were mixed; Abe and colleagues (2020) and Bodala and colleagues (2016) reported a positive relationship with vigilance, while Hu and Lodewijks (2021) found a negative relationship. Interestingly, Li et al. (2019) reported that such a relationship between saccadic velocity and vigilance varies across levels of perceived mental fatigue, showing a positive association under moderate fatigue but negative association under low and high mental fatigue conditions. Microsaccades, or saccades with rotation angles around the *x*- and *y*-axes of less than 1 degree in magnitude, were also significant, as both the frequency of microsaccades and the microsaccade ratio were observed to decrease with diminishing vigilance (Abe et al., 2020). Furthermore, the proportion of slow eye movements over time, or sinusoidal gaze deflections exceeding 500 milliseconds, was correlated with reduced vigilance (Aeschbach et al., 2024). Changes in gaze patterns, such as decreased gaze entropy (Li et al., 2022), a parameter that indicates scanning efficiency by quantifying the distribution of eye fixations (Jeong et al., 2019; Jónsdóttir et al., 2023; Shiferaw et al., 2019), and increased horizontal gaze dispersion (He et al., 2023), were associated with vigilance.

Seven of the studies focused on using eye blinks or eyelid measures to find significant relationships with vigilance. Increased blink duration and frequency/rate (Bodala et al., 2016; Hu & Lodewijks, 2021; Liu et al., 2024; Maffei & Angrilli (2018); Sciaraffa et al., 2021), increased percentage of time eyes were closed (PERCLOS) (Abe et al., 2020; Funke et al., 2017; Liu et al., 2024), and increased eyelid opening degree were all consistently correlated with lower vigilance. Lower eyelid opening degree variability was associated with lower levels of vigilance (Abe et al., 2020).

Seven papers used pupil-based measures to indicate vigilance levels. Three studies found that pupil diameter grew with decreased vigilance (Guo et al., 2024; Hu & Lodewijks, 2021; van den Brink et al., 2016), yet two other studies observed the opposite relationship (He et al., 2023; Li et al., 2019). Additionally, higher values of the pupillary unrest index (PUI), capturing short-term fluctuations in pupil size, were associated with vigilance levels in opposite directions. Maccora and colleagues (2018) found a positive relationship while Yamashita and colleagues (2021) found a negative relationship.

Table 5. Eye Tracking-based Measures of Vigilance

Measure	Variable	Relationship with vigilance	References
Eye Movement	Fixation duration	+	Li et al. (2019)
	Fixation number	+	Li et al. (2019)
		–	Li et al. (2019)
	Time-To-First-Fixation	–	Di Flumeri et al. (2019)
	Saccade amplitude	+	Li et al. (2022); Bodala et al. (2016)
		–	Li et al. (2019)
	Saccade duration	–	Hu & Lodewijks (2021)
	Saccade number	+	Li et al. (2019)
		–	Li et al. (2019)
	Saccade velocity	+	Abe et al. (2020); Bodala et al. (2016); Li et al. (2019)
		–	Hu & Lodewijks (2021); Li et al. (2019)
	Microsaccade frequency	+	Abe et al. (2020)
	Microsaccade ratio	+	Abe et al. (2020)
	Slow eye movements	–	Aeschbach et al. (2024)
	Gaze entropy	+	Li et al. (2022)

	Horizontal gaze dispersion	–	He et al. (2023)
Blink and Eyelid	Blink duration	–	Hu & Lodewijks (2021); Liu et al. (2024); Sciaraffa et al. (2021)
	Blink rate/frequency	–	Bodala et al. (2016); Hu & Lodewijks (2021); Liu et al. (2024); Maffei & Angrilli (2018); Sciaraffa et al. (2021)
	Average eyelid opening degree	–	Abe et al. (2020)
	Standard deviation of eyelid opening degree	+	Abe et al. (2020)
	PERCLOS	–	Abe et al. (2020); Funke et al. (2017); Liu et al. (2024)
Pupil-Based	Pupil diameter	+ –	Guo et al. (2024); Hu & Lodewijks (2021); van den Brink et al. (2016) He et al. (2023); Li et al. (2019)
	PUI	+ –	Maccora et al. (2018) Yamashita et al. (2021)

3.5.3 Evaluation of Sensor-based Vigilance Research

All sensors cited in the 23 reviewed studies were assessed for their readiness to detect vigilance in real-world applications. Each sensor discussed in the previous sections was assessed in terms of temporal resolution, cost, and wearability. The number of measures used in the analyses and the consistency of findings across studies were additionally included as evaluation criteria, as shown in *Table 6*.

The temporal resolution of a sensor is a critical feature, as vigilance decrement can occur as early as 5–15 minutes after beginning a task. Physiological and eye tracking sensors offer high temporal resolution, capturing data within hundreds of milliseconds (e.g., Kivikangas et al., 2011; Picot et al., 2012). Physiological sensors are relatively low in cost, ranging from hundreds of dollars to a few thousand dollars, which makes them less of a burden to use in various settings

(Mehta & Parasuraman, 2013). Similarly, eye trackers cost a few thousand dollars on average, though prices vary depending on the sampling rates.

High wearability enables comfortable, continuous use during naturalistic tasks, allowing for longer and more accurate measurements of vigilance. Most physiological sensors, ECG, temperature, and EDA sensors were rated highest in wearability, as they can typically be worn on the wrist and are lightweight and compact. Depending on the context, eye trackers can be head-mounted (similar to glasses) or positioned in a fixed position in front of the user (e.g., below a computer screen, above a car's steering column). This range led to more moderate ratings of wearability, as head-mounted eye trackers are typically only comfortable for shorter periods of time compared to physiological sensors on one's wrist. The number of measurements each sensor can provide was also evaluated, as a greater number of features can improve the accuracy and reliability of vigilance monitoring and prediction. Although ECGs provide a variety of analyzable measurements, some inconsistencies were observed. Likewise, eye movement metrics and pupil-based measures contained contrasting findings across more than two studies, whereas blink/eyelid metrics showed more consistency. The observed relationships between each measure and vigilance were used to assess their overall consistency. For the salivary and EDA sensors, only one study was reviewed for each so consistency evaluations for these measures were not carried out.

Table 6. Evaluation of Vigilance Sensors for Real-World Applications

	Measure	Temporal Resolution	Monetary Cost	Wearability	Number of Measures	Consistency in Vigilance Measurement
Physiological Sensors	**ECG**	High	Low	High	High	Moderate
	Temperature	Moderate	Low	High	Low	Moderate
	Salivary	Low	Moderate	Low	Low	—
	EDA	Moderate	Low	High	Moderate	—
Eye Tracking Sensors	**Eye Movement**	High	Moderate	Moderate	High	Moderate
	Blink/Eyelid	High	Moderate	Moderate	Moderate	High
	Pupil-Based	High	Moderate	Moderate	Low	Low

Based on the sensor evaluation, there are several promising candidates for reliable vigilance estimators. The ECG features high temporal resolution at a relatively low cost while offering high wearability and a large set of metrics. However, these measures' relationships with vigilance are still somewhat inconsistent. Additional vigilance studies using ECGs can help solidify these expectations for researchers and practitioners. Within the physiological type of sensors, temperature and EDA may also be effective tools for naturalistic studies due to their low cost and high wearability. Compared to ECGs, they have lower resolution and fewer available measures. Furthermore, temperature sensors lack high consistency in their measures' relationships with vigilance while there aren't enough studies using EDA sensors to judge their consistency. Among eye tracking sensors, all types excel in temporal resolution due to the exceptionally quick nature of eye behavior. All eye tracking sensors also feature moderate cost and wearability. While there are few and inconsistent pupil-based measures, eye movement and blink/eyelid measures offer benefit trade-offs when it comes to the number of available measures and their consistency with vigilance. There are a high number of eye movement-related measures but with only moderate consistency in their relationships with vigilance. There is only a moderate number of blink/eyelid measures available, but they are more consistent in their relationships with vigilance.

3.6 Discussion

This chapter systematically reviewed 23 empirical studies published in the past ten years to examine recent advances in and limitations of physiological and eye-tracking sensor methods for vigilance estimation.

Altogether, there is currently no "best" sensor for measuring vigilance; several types of sensors have the potential to serve as reliable methods for monitoring vigilance. ECGs are promising vigilance sensors due to their high temporal resolution, low costs, high wearability, and high number of available measures. However, the literature findings lack high consistency, so these sensors require additional research to solidify trends with vigilance levels. Two other physiological sensors, temperature and EDA, may also be used to assess vigilance due to moderate temporal resolution, low costs, and high wearability, but likewise require more evidence to support their capability in determining vigilance levels. Salivary sensors are relatively new, lack significant support in the literature, and would be difficult to use in naturalistic settings, making them unsuitable for monitoring of vigilance at this time. Overall, relatively few papers used peripheral physiological sensors (8 of 23), therefore more research is needed to accurately assess these methods' effectiveness in vigilance detection.

Eye tracking sensors were well-represented in this literature review. The most capable types were eye movement sensors and blink/eyelid-based sensors. Both feature high temporal resolution with varying costs and wearability depending on the type of eye tracking device. Eye movement tracking excels in the number of available measures while blink/eyelid tracking feature high consistency in vigilance detection.

Recent vigilance research using sensors reveals the need for sensors' improved applicability to real-world settings. Laboratory environments remain the predominant study setting in recent research on these types of sensors. This underscores the necessity for more naturalistic studies to explore and validate sensor use in more diverse and broader real-world, high-risk work environments, where accurate vigilance measurement and prediction are critical. Only one study used both physiological and eye tracking sensors together to assess vigilance. A multi-sensor approach may be the key to robustly and accurately predicting vigilance decrement, where the low cost and portability of physiological sensors can be combined with the temporal resolution and available metrics from eye tracking. Thus, *Chapters 4, 5* and *7* of this dissertation aim to use several of the vigilance sensors identified in this chapter for applications in the healthcare and driving domains.

Chapter 4: Vigilance Modeling Framework

Abstract

This chapter introduces a vigilance experiment conducted with resident physicians at the University of Washington. The study's methods, contextual and physiological factors, and data processing are discussed. The dataset, analytical methods, and results serve as the basis for my vigilance modeling framework. In this framework, a correlation analysis and regression analysis are used to select features for modeling vigilance in resident physicians. After the feature selection process, several Bayesian Networks with varying architectures were formulated and the prediction performance between the models were compared. The final Bayesian Network is used to answer RQs 1–3 and serves as a baseline model for additional analysis. The framework developed in this chapter is applied in future chapters for modeling in healthcare and driving.

4.1 Background and Motivation

Vigilance decrement can create significant safety concerns, particularly in hospitals. In clinical settings, medical personnel are expected to remain vigilant and provide proper care for their patients. However, resident physicians often endure 60 to 80-hour work weeks during their training and night shifts in which they work overnight, sometimes multiple days in a row, while being directly responsible for several patients a day (Barger et al., 2023). Likewise, a majority of residents' sleepiness scores were similar to those with sleep disorders (Papp et al., 2004). The long working hours and sleep deprivation of resident physicians make it difficult to avoid vigilance decrement, potentially contributing to medical errors and adverse patient outcomes.

Monitoring vigilance decrement in a timely manner is critical in hospital settings given the impact on resident physicians’ job satisfaction, patients’ safety, and healthcare quality.

This study aims to develop a predictive model that considers unique contextual interactions within a hospital setting while employing multiple physiological sensors. The steps used to synthesize the predictive model serve as a framework for modeling vigilance that could be used for applications within the healthcare domain and applied to other safety-critical domains. This chapter outlines the initial effort to model vigilance and validate the vigilance modeling framework using a partial healthcare dataset. *Chapter 5* considers a full form of the healthcare dataset using this chapter’s framework approach.

4.2 Participants

In this initial study, 15 resident physicians (7 female-identified, 6 male-identified, and 2 non-binary) between the ages of 26 and 35 (mean = 28.9, SD = 2.3) were recruited using email and physical flyers. Participants had years of experience ranging from zero to two years as resident physicians. Potential participants were screened using a recruiting survey. Eligible participants were resident physicians with normal or corrected vision and without diagnosed sleep disorders. Each resident visited a closed setting in the hospital twice for data collection, once with over six hours of sleep the previous night (the non-sleep-deprived condition) and once with under five hours of sleep (the sleep-deprived condition). Due to the unpredictable nature of residents’ call schedules, 9 participants completed both their sessions while 6 completed one session, resulting in a total of 24 observations to be used for analysis.

4.3 Apparatus

This section describes the tools and devices used to collect data during the study. The different types include self-reported responses, physiological responses, and behavioral responses.

4.3.1 Self-Reported Responses

At the start of each session, the ambient temperature and humidity were recorded using a room thermometer. A sleep diary was developed to capture participants' self-rated sleep quality, hours in bed and asleep, and number of times they woke up overnight for both the night before the session and the prior night. *Figure 3* shows the sleep diary completed by participants during their sessions. The top grid is used as an example for participants to reference as they fill out their sleep diaries for Night 1 (two nights ago) and Night 2 (last night).

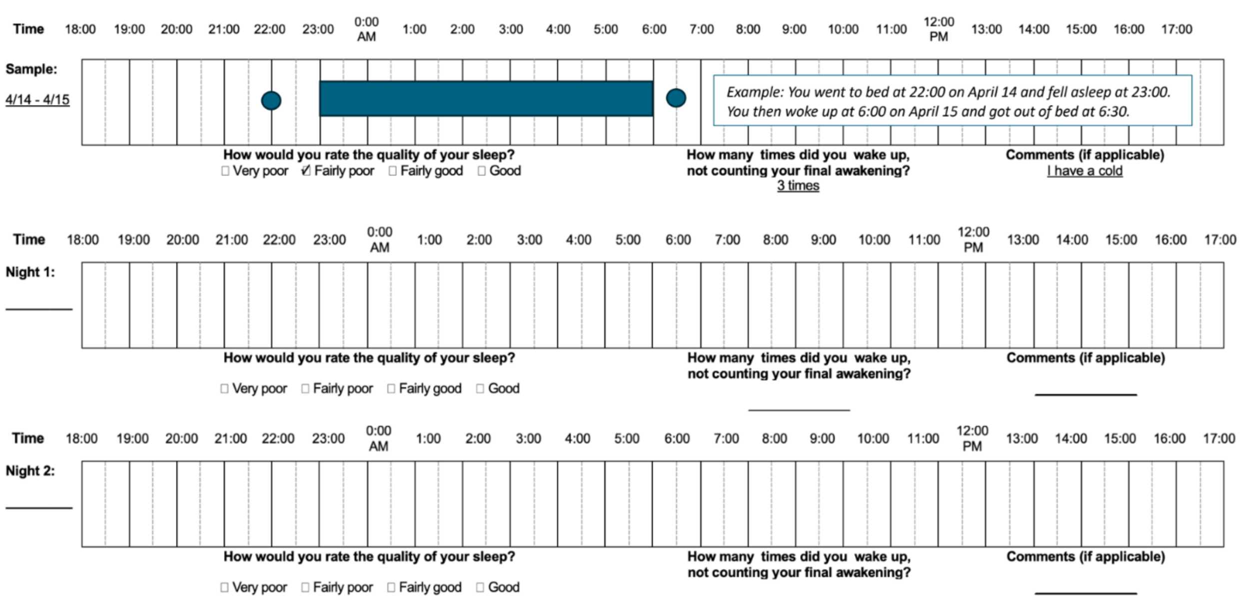

Figure 3. Sleep Diary Worksheet

A PSQI was used to supplement the sleep diary with a validated scale for participants' sleep quality over the past month and the past week leading up to their session. For each time period, the participants' usual bedtime, actual hours of sleep (not including being in bed), and average sleep quality were recorded and used to calculate their PSQI score.

A State-Trait Anxiety Inventory (STAI; Spielberger et al., 1983) was collected in order to assess participants' self-reported presence and severity of anxiety. The "state" portion of the questionnaire measures the current symptoms of anxiety while the "trait" portion aims to reveal their general propensity to be anxious. A short form of the STAI, developed by Zsido and colleagues (2020), which consists of five statements about state anxiety and five statements about trait anxiety. Participants were asked to rate their agreement with each statement using values ranging from 1 (Not at all) to 4 (Very much so).

Hart & Staveland's (1988) NASA Task Load Index (TLX) was administered to evaluate participants' workload during the study's vigilance task. The NASA TLX assesses workload across six different 20-point Likert scales from "Very Low" to "Very High". The six workload areas included Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. These six assessments are then used to calculate an overall workload index.

4.3.2 Physiological Responses

An Empatica E4 wristband (see *Figure 4*) was used to measure and record the participant's blood volume pulse (BVP), electrodermal activity (EDA), skin temperature (TEMP), and heart rate (HR). To minimize noise from arm movements, participants were asked to wear the wristband on their nondominant hand throughout the experiment.

Figure 4. The Empatica E4 Wristband

A 15-inch Dell XPS laptop with a resolution of 1920 x 1080 was used for the experiment task and eye tracking. The laptop was elevated off the desk using a laptop stand, bringing it to eye level for participants and aligning the eye tracker for optimal calibration. A Tobii Pro X3-120 eye tracker (see *Figure 5*) was used to collect participants' eye tracking data at a sampling rate of 120 Hz. The eye tracker was mounted below the screen of the Dell XPS laptop. Eye-related measures were recorded using the Tobii Pro Lab software. To help ensure high-quality data, participants were instructed to maintain a consistent distance (approximately 40 cm) from the screen after calibration.

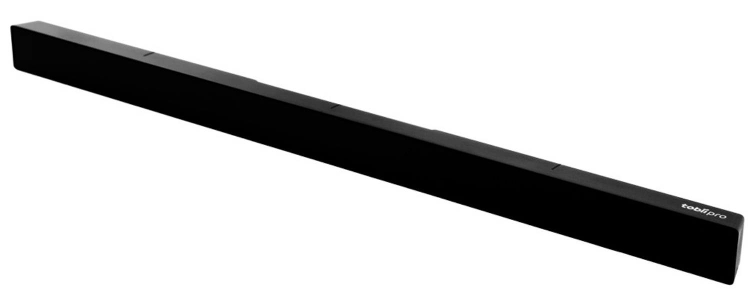

Figure 5. The Tobii Pro X3-120

4.3.3 Behavioral Responses

A PVT was used as the primary vigilance task in this study. The PVT consisted of a simple, monotonous task in which participants were prompted to press the computer's spacebar every time they saw a red dot appear on the screen (see *Figure 6*). The red dot appeared at random time intervals between 1 and 5 seconds.

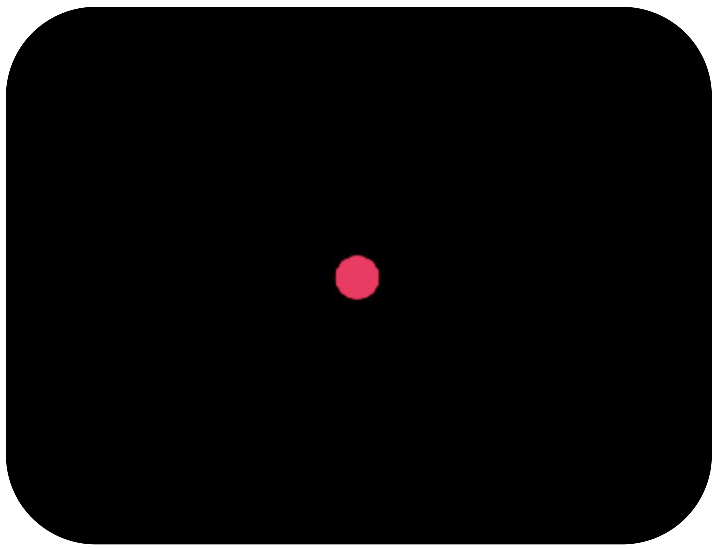

Figure 6. View of the PVT as the Red Dot Appears

4.4 Procedure

Participants began by reviewing and signing the study's consent form, approved by the University of Washington Institutional Review Board. Next, participants completed the sleep diary for the previous two nights, the STAI survey, and PSQI form. While the participant completed the questionnaires, a researcher recorded the current ambient temperature and humidity. Following a one-minute practice session, the participants completed a 15-minute Psychomotor Vigilance Test (PVT) (Basner & Dinges, 2011). After the PVT, participants were instructed to fill out a NASA TLX questionnaire. *Figure 7* provides an overview of the experiment procedure.

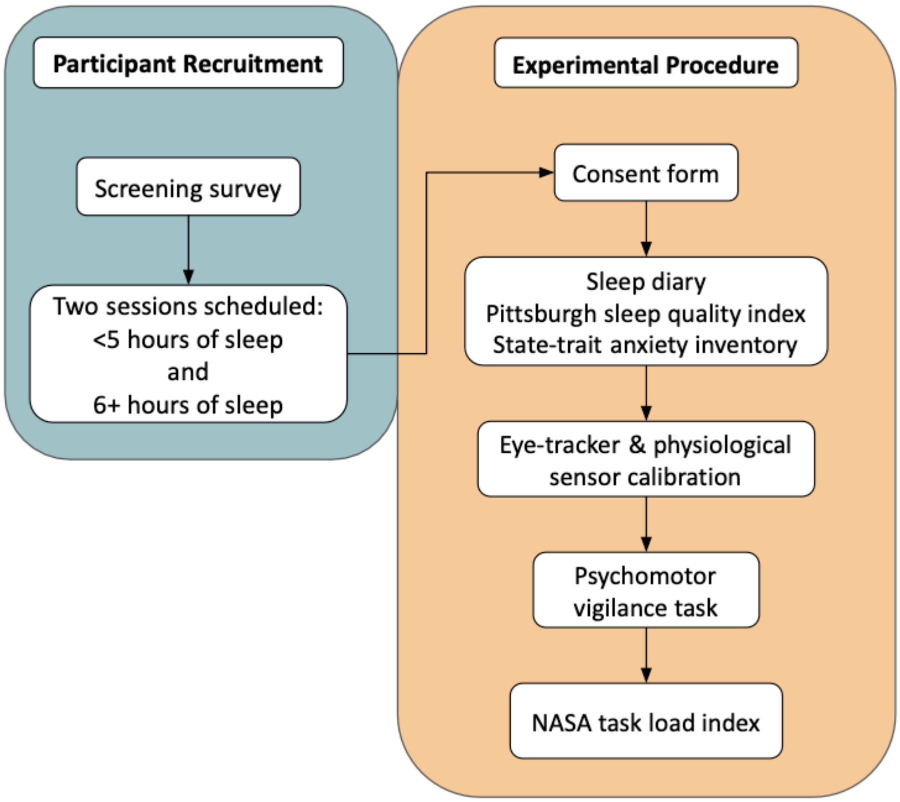

Figure 7. Vigilance Experiment Procedure

4.5 Experimental Design

Given the relatively small and specialized participant pool of resident physicians, a within-subject design was used in this study. Each participant visited the experimental space twice: one visit included over six hours of sleep the previous night (the low sleep-deprived condition) and the other visit included less than five hours of sleep (the high sleep-deprived condition). The order of low and high sleep-deprived was randomized for participants, and dependent on their shift schedule.

4.6 Data Processing

Physiological data, including BVP (64 Hz), EDA (4 Hz), TEMP (4 Hz), and HR (1 Hz), were processed and synchronized by first reading the data from CSV files and unifying each signals' timestamps. Any missing values were imputed using linear interpolation. To reduce noise from the signals, a lowpass, 5th order Butterworth filter was applied. Butterworth filters are maximally flat filters that are commonly used to flatten ripples in signal passbands and stopbands (e.g., Mello et al., 2007). Finally, the filtered physiological data were downsampled to 1 Hz by averaging the data within one-second intervals.

The eye tracking data were cleaned using Tobii Pro Lab then exported as CSV files. Any invalid data points were removed. To synchronize with the physiological data, the eye tracking data were matched by timestamp and downsampled to 1 Hz by calculating, for each second, the mean pupil diameters, saccade durations, fixation durations, saccade counts, and fixation counts. In addition, all physiological and eye tracking data were subjected to a z-score normalization in order to standardize the biometric data across participants.

For this study, participants' vigilance decrement was measured by their average reaction time while completing the PVT (RT_mean). Reaction time was defined as the time interval between the appearance of the red dot on screen and the moment the participant responded by pressing the spacebar. This reaction time was recorded in milliseconds. All data processing, feature selection, and analysis were performed in Python and R programming languages.

4.7 Data Analysis

A total of 38 candidate predictor variables for vigilance were collected during this study, ranging from questionnaire results to eye tracking and physiological sensor-based data. Due to the large number of variables collected, the dataset is summarized in tabular form in the Appendix. The next section, *Feature Selection using Mixed-Effects Modeling*, covers the process for reducing the number of significant variables in the predictive models.

After feature selection, resident physicians' vigilance was modeled using a BN. BNs model causal inference using Bayes' theorem, in which the probability of a cause is inferred by the observed effect (Puga et al., 2015). A BN is a directed acyclic graph made up of interacting nodes that are connected by edges. For example, an edge A→B implies that A has an effect on B. The nodes in a BN are defined by the probability of finding it in a given state (e.g., high or low). If the state of a node is dependent on another node's state, conditional probability is used. These dependencies allow for continuous updates of the entire network when new information is provided, making BNs effective probabilistic models (Puga et al., 2015). The BN in this study followed the architecture proposed by Garg and colleagues (2000): a top layer for contextual factors (e.g., sleep deprivation), a middle layer for performance (e.g., vigilance), and a bottom layer for observable features (e.g., eye movements and physiological measures). The BNs considered in this dissertation are modeled as Conditional Linear Gaussian BNs, which are able to handle datasets with a mix of discrete and continuous variables to make predictions. In these BNs, the continuous nodes are modeled as linear regressions on their continuous parent nodes, with coefficients and error variances depending on the configurations of discrete parent nodes. This BN formulation eliminates the need to discretize any data during pre-processing.

4.8 Results

This section highlights the findings from this chapter's feature selection process and BN model formulation for vigilance prediction.

4.8.1 Feature Selection using Mixed-Effects Modeling

Physiological data were standardized to account for individual differences, such as normal HR range and pupil diameters. The data were transformed using a z-score normalization:

$$Z = \frac{x - \mu}{\sigma}$$

where:

Z is the standardized value

x is the original value

μ is the mean of the sample

and σ is the standard deviation of the sample

Next, the variables for feature selection were assessed for multicollinearity issues using a Spearman's correlation analysis. The Spearman method was selected to account for nonlinear relationships between the outcome variable and predictor variables, particularly because there were several types of variables in the predictor set (e.g., continuous, categorical, integer). The correlation plot in *Figure 8* shows the correlation between variables in the physiological measures while the correlation plot in *Figure 9* shows the correlation between variables in the contextual factors of the study. Note that these figures used the data code names for each variable in the dataset (see *Appendix* for additional details).

All highly collinear variables were identified as those with a Spearman's rank correlation coefficient, ρ, of $|0.8|$ or greater. The variables with multicollinearity issues were removed from the set of variables used in the next feature selection steps. For example, some of these variables included the average ($\rho = 0.91$) and standard deviation ($\rho = 0.84$) of the left pupil's diameter, as they were highly correlated with the right pupil's diameter measures. The standard deviation of eye fixation durations ($\rho = 0.98$), the standard deviation of eye fixation counts ($\rho = -0.83$), and standard deviation of saccade counts ($\rho = -0.83$) were also removed due to their correlation with average fixation duration, standard deviation of fixation duration, and standard deviation of saccade duration, respectively. The participant's visit condition ($\rho = -0.86$) was also removed due to being correlated with the previous night's hours of sleep variable.

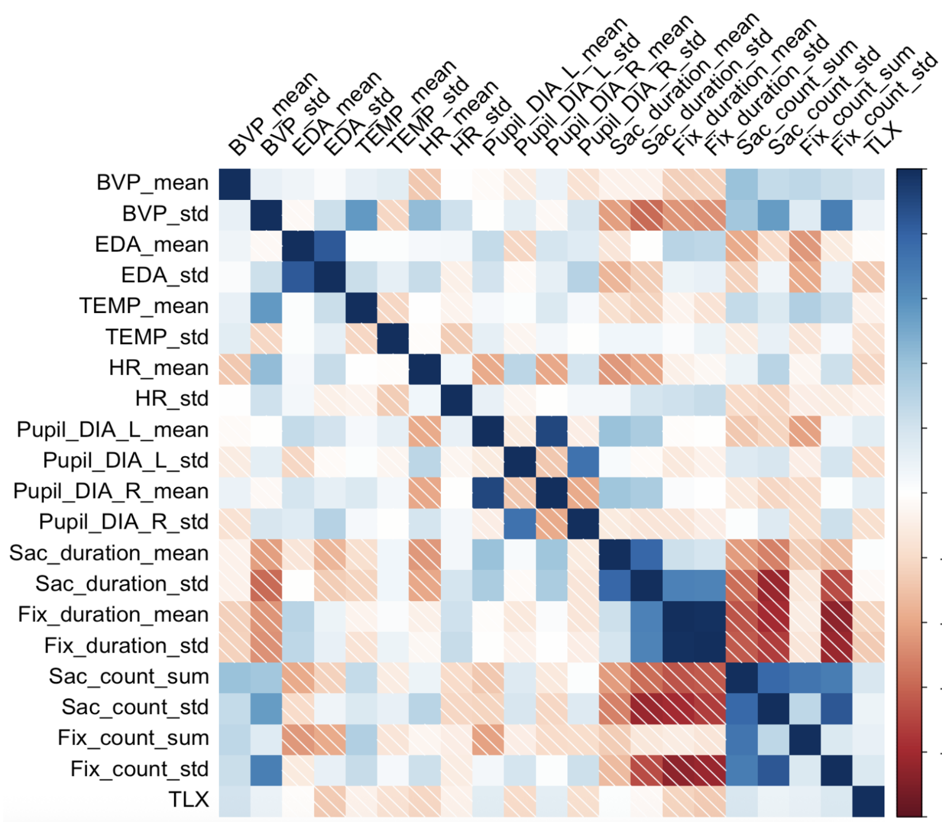

Figure 8. Correlation Plot for Physiological Measures

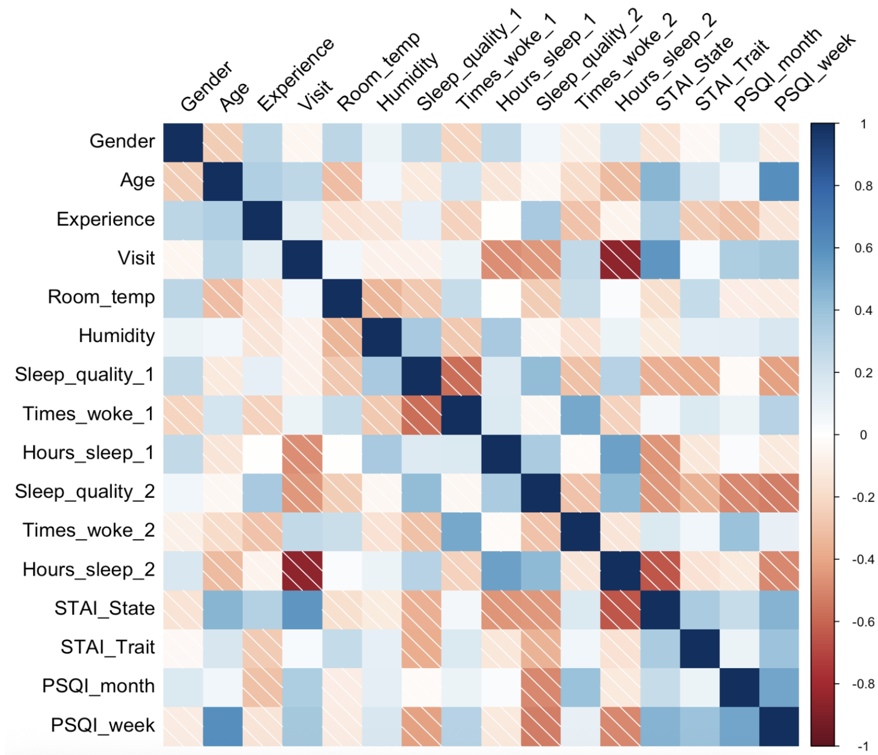

Figure 9. Correlation Plot for Contextual Variables

A mixed-effects regression was used to select the key variables for predicting vigilance while accounting for random effects from individual participants. The model was formulated as:

$$y_{ij} = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ij} + \gamma_j + \varepsilon_{ij}$$

where:

y_{ij} is the outcome of the ith observation of the jth participant

β_0 is the model's intercept

$\beta_k x_{ij}$ is the effect of the kth predictor variable in the ith observation of the jth participant

γ_j is the random intercept for the jth participant

and ε_{ij} is the residual error term of the ith observation of the jth participant

The final mixed-effects model included the following significant predictor variables ($p < 0.05$): average BVP, average EDA, average TEMP, HR variability from the wearable physiological sensor; total saccade count and saccade count variability from the eye tracker; and room temperature, the prior night's sleep quality (via sleep diary), state anxiety (via STAI), and sleep quality over the past month (via PSQI) for the contextual factors. The outcome variable in this model was average reaction time (in ms) during the PVT. A summary of the model's results is shown below in *Table 7*. The relationships between the features and vigilance are further explored in the *Discussion* section.

Table 7. Mixed-Effects Regression Model Results for Partial Healthcare Dataset

Variable Type	Variable	Estimate	Std. Error	p-value
	(Intercept)	333.8	21.9	< 0.01
Contextual Factors	Room Temperature	–7.65	0.29	< 0.01
	Prior Night's Sleep Quality	–22.06	4.90	< 0.01
	Past Month's Sleep Quality	–4.15	4.15	0.02
	State-level Anxiety	11.74	1.85	< 0.01
Physiology	Average BVP	5.05	0.83	< 0.01
	Average EDA	–32.33	0.50	< 0.01
	Average TEMP	–7.53	1.79	< 0.01
	HR Variability	–16.16	0.17	< 0.01
Eye Movement	Total Saccades	–0.02	0.01	< 0.01
	Saccade Variability	29.92	0.01	< 0.01

4.8.2 Vigilance Modeling Using Bayesian Networks

To evaluate the effectiveness of the three-layered BN, five variant BN models were assembled, and their model performance was compared using root mean squared error (RMSE). A leave-one-out cross validation (LOOCV) was used to ensure robustness and generalizability of the results while preventing any overfitting of the dataset. During this process, RMSE values were collected per iteration and then averaged for each BN.

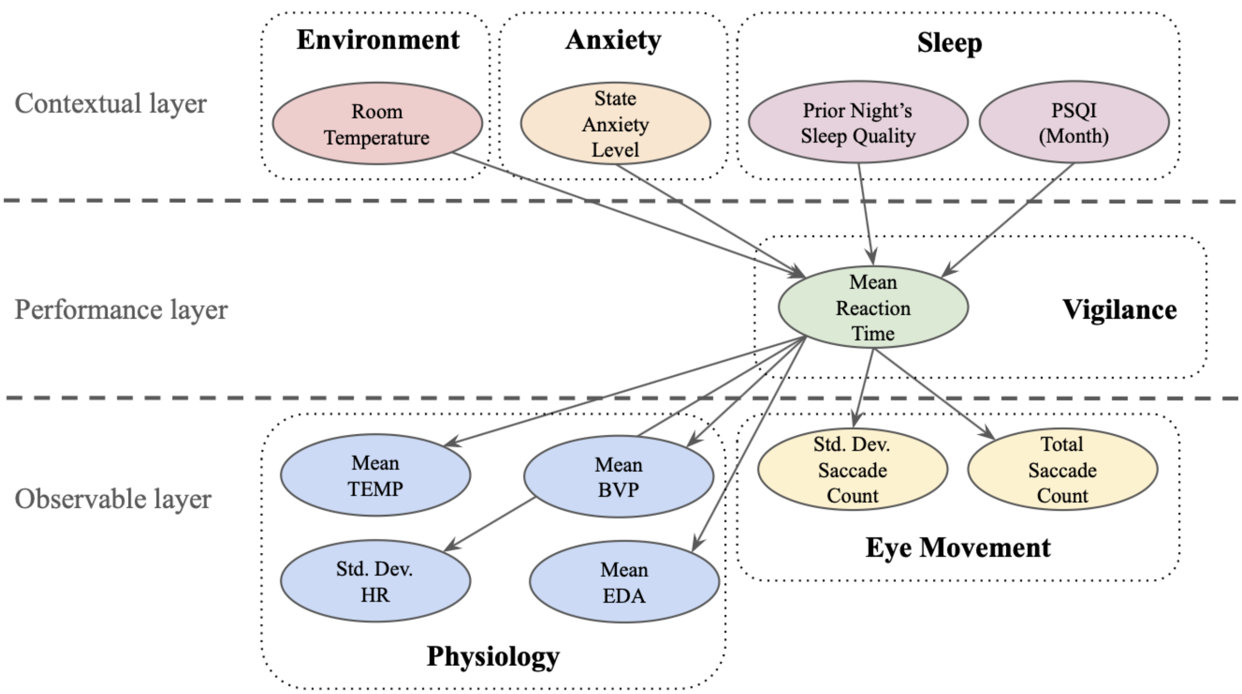

Figure 10. Full BN for Vigilance (BN1)

Figure 10 shows the three-layer network containing the contextual, performance, and observable layers (BN 1). To test against the three-layer structure, two-layer networks were also developed. One included only the contextual and performance layers (*Figure 11a*), while the other was composed of only the performance and observable layers (*Figure 11b*). Two additional three-layer networks were constructed to evaluate the effects of different sensors. BN 4 (*Figure 11c*) and BN 5 (*Figure 11d*) only used measures from the eye tracker or the physiological sensors, respectively, in the observable layer.

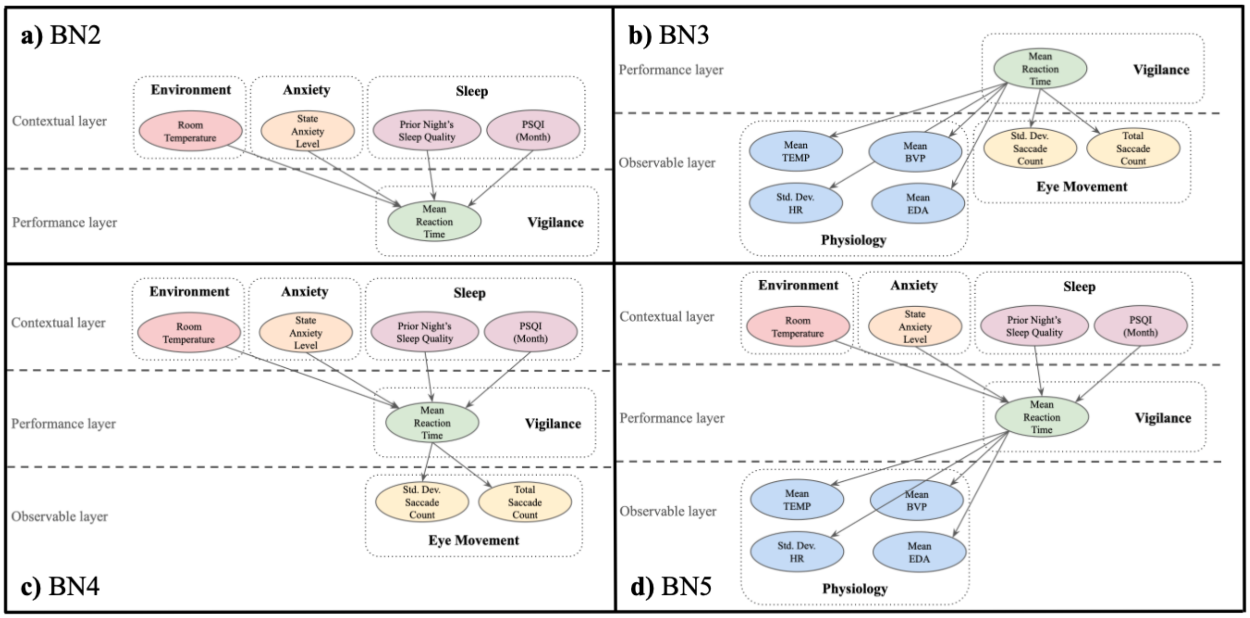

Figure 11. BN 2 (top left), BN 3 (top right), BN 4 (bottom left), and BN 5 (bottom right)

Table 8 shows the RMSE and 95% confidence intervals of the RMSE across all five BNs as a result of the LOOCV. The results indicated that BN 1 had the best prediction performance, followed by BN 5, as indicated by having the lowest RMSE among the five BNs (20.4 ms).

Table 8. Bayesian Network Models Performance (in ms)

Model	**BN 1:** Three Layers	**BN 2:** Contextual + Performance Layers	**BN 3:** Performance + Observable Layers	**BN 4:** Three Layers (Eye Movement Only)	**BN 5:** Three Layers (Physiology Only)
RMSE	20.4	22.5	24.8	22.5	20.9
95% CI	[13.0, 26.9]	[12.0, 30.9]	[15.7, 31.1]	[13.7, 29.7]	[11.3, 28.5]

4.9 Discussion

This chapter used a partial healthcare dataset to develop a vigilance modeling framework for safety-critical environments and uncover significant contextual and physiological measures for predicting vigilance. The results from this chapter are used to provide initial insights and answers to this dissertation's RQs 1–3.

4.9.1 Contextual Factors Affecting Vigilance

Among contextual factors, the environment (room temperature), sleep quality (nightly and monthly), and acute anxiety level all played a role in predicting vigilance for this chapter's models. Starting with the contextual factors, room temperature had a negative relationship with reaction time; higher temperatures were associated with lower reaction times and thus better vigilance. Given that the room temperatures only ranged between 19.5°C and 24°C, this finding may point to the fact that slightly warmer temperatures in the hospital can help practitioners maintain their focus. Sleep quality also had a significant effect on participants' vigilance, as there was a negative relationship between reaction time and sleep quality during the previous night (via sleep diary) and across the past month (via PSQI). This highlights the importance of sleep quality on residents' ability to sustain attention, in which poor sleep quality results in longer reaction times (lower vigilance). It should also be noted that the effects of poor sleep quality can be long lasting, indicated by the significance of the monthly sleep quality variable. State-level anxiety, participants' current feelings of stress and anxiousness via STAI, also had an impact on their vigilance. The positive relationship shows that higher anxiety can be associated with longer reaction times, and thus worse vigilance. It is noteworthy that trait-level anxiety was not a significant factor in the model, which could point to the fact that anxiousness associated with an individual's intrinsic characteristics or personality does not necessarily affect how they experience vigilance decrement compared to others.

4.9.2 Physiological Measures for Monitoring Vigilance

Several physiological measures from the Empatica wristband and Tobii eye tracker were significant predictors of vigilance in this chapter. From the Empatica wristband, average BVP, average EDA, average TEMP, and HR variability were included in the vigilance models. Average BVP had a positive relationship with reaction time, indicating that higher BVP is related to lower levels of vigilance. Conversely, average EDA had a negative relationship with reaction time—as EDA increases, so does vigilance. Average TEMP also had a negative relationship with reaction time, showing that warmer skin temperature can indicate higher vigilance levels. As HR variability increased, reaction times decreased, showing that unstable HR values may indicate higher vigilance and increased mental effort.

The significant variables obtained from the Tobii eye tracker were both related to saccadic eye movement. Higher variability in saccade count was associated with longer reaction times (lower vigilance). Total saccade count had a negative relationship with reaction time, showing that a greater number of saccades is indicative of more active visual processing and therefore increased vigilance.

4.9.3 Predictive Modeling Features and Strategies for Vigilance

The results from the mixed-effects regression and BN models highlight the importance of considering both continuous, physiological sensors and contextual factors when assessing the vigilance of workers in safety-critical environments, such as resident physicians in hospitals. This is supported by the various types of significant features in the mixed-effects model and the higher errors observed in the two-layer BNs (BN 2 and BN 3) compared to the three-layered BN 1. In addition, the use of multiple types of real-time sensors enhanced the prediction accuracy of the vigilance models, as shown by the lower errors of BN 1 (both types of sensors) compared to BN 4 and BN 5 (one sensor type each). Overall, the RMSEs of all BNs ranged between 20 to 25 ms for participants' reaction time during the PVT, highlighting the promising opportunity for modeling vigilance levels in real-time using probabilistic models.

4.9.4 Summary of Vigilance Modeling Framework

The process for analyzing vigilance data and developing the predictive models in this chapter forms the basis for the vigilance modeling framework used in *Chapters 5* and *7*. The initial framework follows the following steps:

1. Conduct a correlation analysis to remove highly correlated independent variables from the dataset to reduce multicollinearity risks.
2. Build a regression model to find and select features for predictive modeling.

3. Develop multiple forms of predictive models using the significant features in *Step 2* as predictors and a measure of vigilance as the outcome variable.
4. Implement a cross-validation method to assess prediction performance across the predictive models.
5. Use a key indicator of predictive performance to compare predictive models.

In the context of this chapter, the dataset's characteristics allowed for a Spearman's correlation analysis in *Step 1* of the vigilance modeling framework, a mixed-effects linear regression in *Step 2*, five unique BN structures in *Step 3*, a LOOCV in *Step 4*, and RMSE as the key indicator of performance in *Step 5*. The general steps of this framework are followed and expanded upon in *Chapters 5* and *7*.

Chapter 5: Vigilance Modeling in Healthcare

Abstract

This chapter applies the vigilance modeling framework to the full healthcare dataset obtained from the study conducted in *Chapter 4*. It also introduces a second vigilance task, the electrocardiogram reading task, to serve as a naturalistic task for healthcare professionals. The two tasks' datasets are used to develop vigilance models for healthcare settings. This chapter is used to answer RQs 1–3 for the healthcare domain.

5.1 Background and Motivation

To expand on *Chapter 4*'s development and validation of a vigilance modeling framework, this chapter uses a more complete version of the dataset to assess vigilance modeling the healthcare domain. The dataset is obtained from the same study but incorporates additional Internal Medicine residents and healthcare professionals in different roles and departments to increase the participant sample size (see *Participants* section). In addition, this complete dataset includes a second, naturalistic vigilance task for healthcare settings: an electrocardiogram (ECG) reading task (see *Electrocardiogram Reading Task* section). The vigilance modeling framework is applied to this dataset by developing vigilance models for the ECG reading task and the PVT.

5.2 Healthcare Dataset

This section describes the new features and updates to the dataset from *Chapter 4*. Unless explicitly stated, all other aspects of the dataset and study procedures remained the same.

5.2.1 Participants

Eight new participants completed the healthcare study and participated in 14 sessions total, bringing the complete dataset sample size to 23 participants with 38 sessions completed. In total, the participant pool consisted of 8 male identified, 13 female identified, and 2 non-binary participants. The participants had an average age of 30.1 (SD = 5.7) years old. While most participants were still resident physicians from the Internal Medicine department at UW, several new participants were registered nurses and respiratory therapists.

5.2.2 Electrocardiogram Reading Task

In this experiment, a naturalistic hospital task was needed to provide representative data of healthcare workers' vigilance. In collaboration with leading doctors and researchers at UW Medicine, an ECG reading task was proposed and developed, as various healthcare professionals regularly review ECG data for their patients. Additionally, minimal training was needed for this task. Although all study participants were expected to be familiar with ECGs as part of their basic medical training, the research team used the screening survey to ensure participants' familiarity with ECG reading prior to the study.

An ECG records the heart's electrical activity through repeated cardiac cycles and can be used to identify signs of a myocardial infarction (MI) rhythm—in layman's terms, a heart attack. In the ECG reading task, participants were presented with a series of 12-lead ECG readings and used the laptop keyboard to answer "yes" if they saw signs of a heart attack or "no" if the signs were absent. The task's software recorded reaction time and the accuracy of the answer. *Figure 12* below shows an example of an ECG reading task sample.

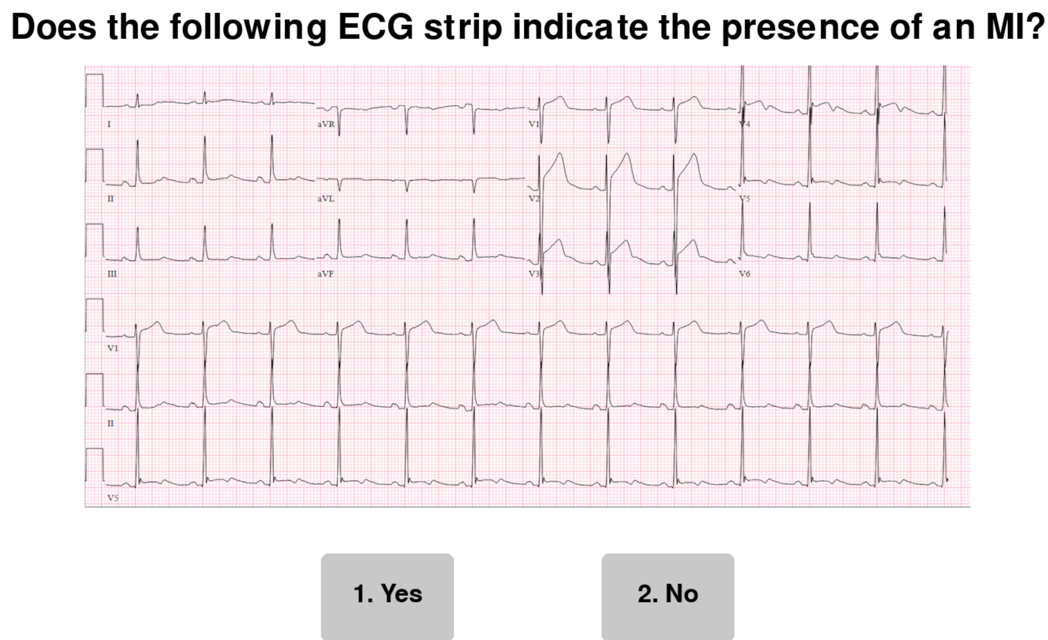

Figure 12. ECG Reading Task Sample

Prior to the full ECG reading task, participants completed a practice test consisting of 15 ECG readings. Participants were informed of their practice test results before the main task. The full task used 60 static ECG samples, 30 from basic life support and 30 from advanced cardiovascular life support materials. Half of the samples (30) showed normal ECG signals while the remaining samples showed signs of a heart attack. In a 15-minute continuous sequence, participants were presented with ECG samples in random order. There was a 15-second time limit on each ECG sample, making the full session a maximum of 15 minutes. Participants' reaction time when making a decision for each sample in the ECG reading task was used to measure vigilance.

5.3 Procedure

The experimental procedure and apparatus in this chapter were the same as in *Chapter 4*, except for the ECG reading task (see *Figure 13*). Each participant completed the PVT and ECG reading task in each sleep-deprived and non-sleep-deprived session. The order of the two tasks was randomized among the participants. With two sessions per participant, and both tasks completed per session, each participant provided four observations to the dataset: 1) PVT during the sleep deprivation session, 2) ECG reading during the sleep deprivation session, 3) PVT during the non-sleep-deprived session, and 4) ECG reading during the non-sleep-deprived session.

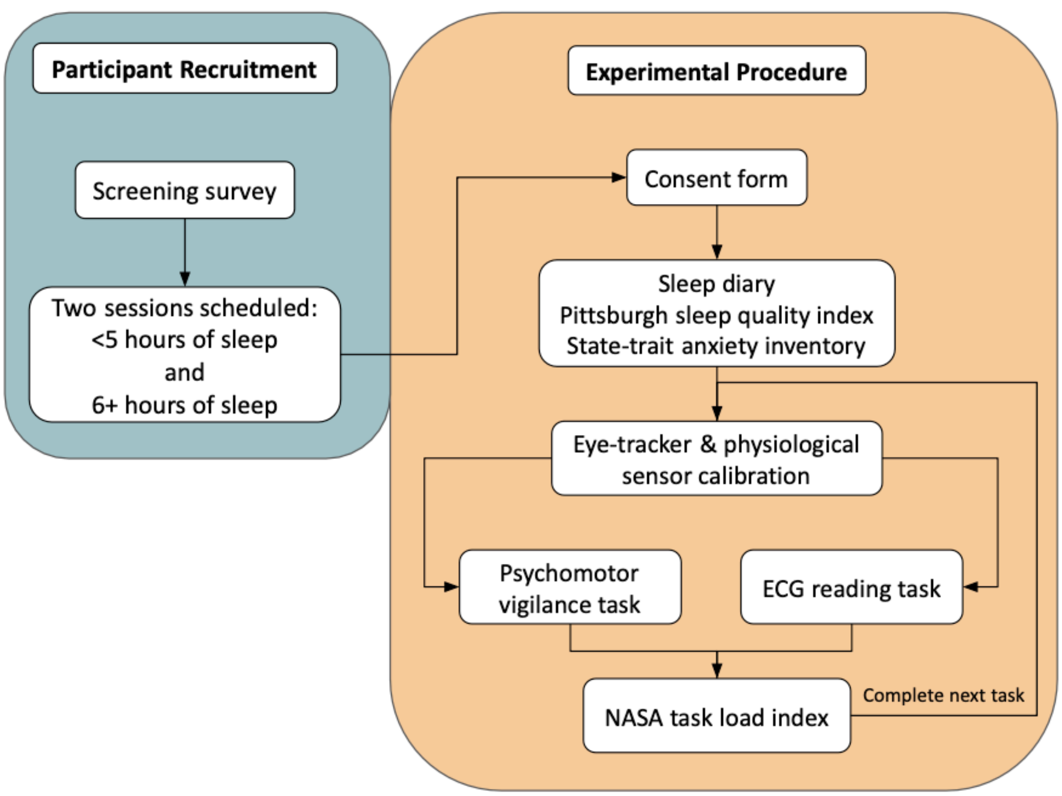

Figure 13. Full Healthcare Experiment Procedure

5.4 Data Processing

The data processing for this dataset followed the same procedure used in *Chapter 4* due to identical variables across the initial and complete datasets. However, there were two changes. Due to required software updates to the Empatica E4 wristband, the new participants' heart rate data required a different processing technique. With this update, systolic peak data were recorded instead of heart rate data. To maintain the heart rate variable in our dataset, the time gap between consecutive systolic peaks was calculated and converted into heart rate values. Since this process did not yield a fixed sampling rate, the heart rate values were resampled by aggregating the data to 1 Hz. To further stabilize the signal and reduce the influence of noise, a moving average filter was applied. Missing values in the data were imputed using linear interpolation.

Next, the data from each session was aggregated at two levels: 1) per session (mirroring *Chapter 4*'s dataset) and 2) per minute to create a dataset suitable for modeling techniques that incorporate a dynamic time element. For the per-minute aggregation, each row represented one minute of data within each session. At both levels, the mean and standard deviation of each variable were calculated. To mitigate the influence of extreme values during aggregation, a winsorization procedure was applied, thus data points falling below the lower bound or above the upper bound of 1.5 times the interquartile range (IQR) were replaced by the respective boundary values. Finally, the aggregated datasets from all participants were merged into a single dataset to serve as the input for the subsequent vigilance model training and evaluation.

5.5 Data Analysis

The analysis of the full healthcare dataset follows the vigilance modeling framework developed in *Chapter 4* and extends the framework by applying multiple modeling techniques to analyze predictive performance and select an optimal vigilance model. Feature selection for both the PVT and ECG datasets was conducted using the PVT data only since it is based on a validated vigilance task and would therefore yield the most significant variables directly related to vigilance. This common feature set also created consistency and comparability between the two datasets' resulting models. Although the final feature set was based on the PVT, the training and testing of each model was specific to each task (i.e., only ECG data were used to train and test the ECG-based models of vigilance). The assessment of the models has been extended to include multiple statistical tests and data visualizations as described in the *Vigilance Models and Analysis* section.

5.5.1 Feature Selection

The full set of predictors was first organized into two categories matching the two-layer BN developed in *Chapter 4*: contextual features, comprising environmental measurements and self-reported survey responses, and observable features, consisting of all physiological and eye-movement measures. Multicollinearity was mitigated within each group using a Spearman correlation analysis. Spearman correlation coefficients were calculated and features with paired correlation coefficients ρ above $|0.8|$ were examined for redundancy and removed primarily based on their correlation with the vigilance measure, reaction time. In particular, the standard deviations of electrodermal activity, pupil diameter of each eye, saccade duration, and fixation duration were among the variables excluded due to high multicollinearity. In cases where both variables were highly correlated with reaction time, the variable with clearer theoretical relevance or greater potential utility for modeling was prioritized.

After multicollinearity in the PVT dataset was addressed, a mixed-effects linear regression model was used to identify features that were significant. The significance of each feature was determined with session-level reaction time serving as the primary dependent variable and participant ID as the random effect. A Backward Stepwise Selection algorithm was used to establish the final feature set. This method reduces the number of variables in a model by reducing the Akaike Information Criterion (AIC), a common metric that balances model fit against complexity to prevent overfitting. *Table 9* summarizes the mixed-effects linear regression model using the final feature set as predictor variables for PVT reaction time. The relationships between these features and vigilance are explored in the *Discussion* section.

Table 9. Mixed-Effects Regression Model Results for Full Healthcare PVT Dataset

Variable Type	Variable	Estimate	Std. Error	p-value
	(Intercept)	668.83	25.47	< 0.01
Contextual Factors	Prior Night's Sleep Quality	–8.97	1.14	<0.01
	Room Temperature	–12.75	1.03	< 0.01
Physiology	HR Variability	–10.26	0.90	< 0.01
	Average BVP	5.06	0.85	< 0.01
	BVP Variability	3.90	1.14	0.01
	Average EDA	–10.04	1.77	< 0.01
	Average TEMP	–8.78	1.56	< 0.01
Eye Movement	Average Pupil Diameter	–10.53	4.16	0.02
	Average Saccade Duration	0.18	0.04	< 0.01
	Average Fixation Duration	–0.02	< 0.01	< 0.01
	Total Fixation Count	–0.02	< 0.01	0.04

5.5.2 Vigilance Models and Analysis

Using the final feature set as predictors for vigilance, five unique modeling techniques were developed. These five types of models included Linear Regression (LR), Support Vector Machines (SVM), Random Forests (RF), Bayesian Networks (BN), and Dynamic Bayesian Networks (DBN). All vigilance models were applied to both the PVT and ECG reading task datasets, resulting in 10 total models.

The LR served as the baseline model, representing the same scheme used for the final feature set selection. The LR model derives the relationship between fixed-effects predictors and reaction time while designating participants as a random-effect. The formulation for this model is described in *Chapter 4*'s *Feature Selection using Mixed-Effects Modeling* section. The SVM model was employed using a radial basis function kernel to capture potential nonlinear relationships in the datasets. These SVM are able to map predictors into a higher-dimensional space to find an optimal nonlinear decision boundary for regression. The SVM model was formulated as:

$$f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i{}^*) K(x_i, x_j) + b$$

$$K(x_i, x_j) = exp(-\gamma \|x_i - x_j\|^2 + C)$$

where:

$f(x)$ is the model's predicted outcome

$\alpha_i - \alpha_i{}^*$ are the model's Lagrange multipliers for points around the decision boundary

$K(x_i, x_j)$ is the model's radial basis function kernel

γ is a hyperparameter defining the "smoothness" of the decision boundary

$\|x_i - x_j\|^2$ is the squared Euclidean distance between two points in the input space

C is a regularization hyperparameter that balances boundary margins and errors

and b is the model's geometric intercept

The RF model was selected as an ensemble machine learning method for modeling vigilance. The RF aggregates the outcomes of a forest of decision trees, thereby reducing variance and improving predictive reliability compared to individual trees. The RF model was formulated as:

$$f(x) = \frac{1}{M}\sum_{m=1}^{M} T_m(x, \theta_m)$$

where:

$f(x)$ is the model's predicted outcome

M is the total number of decision trees in the forest

and $T_m(x, \theta_m)$ is the prediction of the *m*-th individual decision tree

Similar to *Chapter 4*'s final BN, a three-layer BN made up of contextual, performance, and observable layers was used as a probabilistic model of vigilance. *Figure 14* shows the graphical form of the BN using the full healthcare dataset for the PVT and ECG reading tasks.

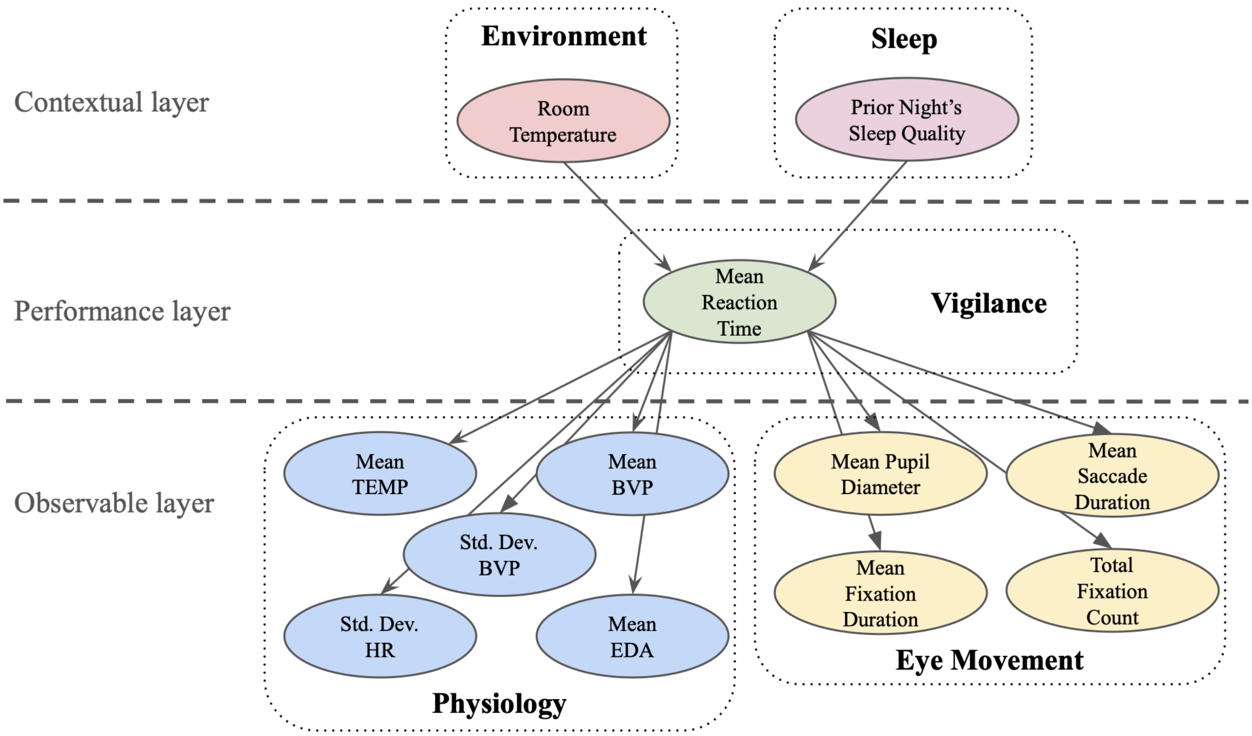

Figure 14. BN for the Full Healthcare Dataset

To account for time-related effects in the data, a DBN was developed. In general, DBNs differ from static BNs by incorporating temporal dependencies in their structure (Ghahramani, 1997; Murphy, 2002). The DBN in this study used a two-time-step structure (*t* and *t + 1*), allowing the model to learn how changes in various measures across time influence subsequent changes in vigilance via reaction time. While the previous four models used the session-level dataset, the DBN used the minute-aggregated dataset to incorporate the time element into the model. *Figure 15* illustrates how the DBN uses the same structure as the BN but extends the network by adding the time step.

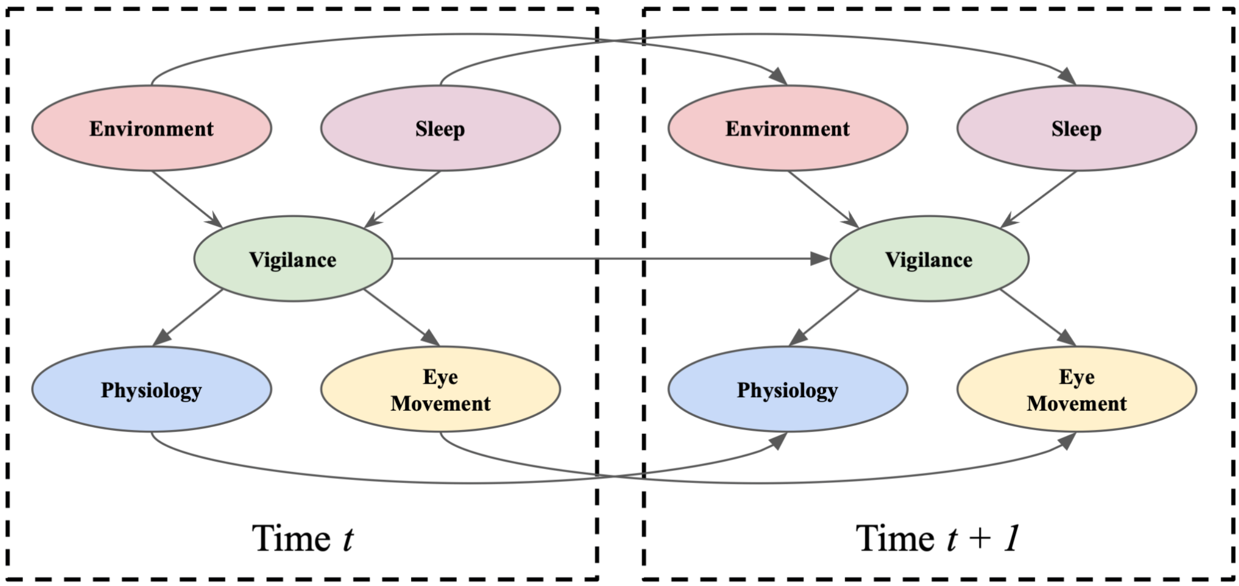

Figure 15. Simplified DBN for the Full Healthcare Dataset

The predictive performance of the vigilance models was obtained by training and testing using a LOOCV. Each model's RMSE, the RMSE 95% confidence interval, boxplot visualizations of absolute errors, Friedman's tests (a non-parametric repeated measures analysis of variance), and Levene's tests (to assess equality of variance) were used to compare the vigilance models within each task.

5.6 Results

This results section covers comparisons between the five vigilance models within the PVT dataset and the ECG reading task dataset and draws conclusions based on the findings.

5.6.1 Model Performance for the PVT

The RMSE results from the vigilance models using the PVT dataset are summarized in *Table 10*, including the RMSE value and its 95% confidence interval for each model.

Table 10. Healthcare Vigilance Models Performance for the PVT (in ms)

Model	Linear Regression	Support Vector Machine	Random Forest	Bayesian Network	Dynamic Bayesian Network
RMSE	28.66	30.05	29.72	30.50	16.99
95% CI	[21.18, 34.56]	[21.48, 36.68]	[23.11, 35.11]	[21.36, 37.47]	[15.55, 18.32]

Following the LOOCV, the DBN had the lowest RMSE of 16.99 ms whereas the other four models had higher RMSE values between 28.66–30.50 ms. However, a Friedman's test across all models was insignificant ($\chi^2 = 4.21$, df = 4, $p = 0.38$), indicating that there was no significant difference in errors across the models. This is further supported in the boxplot in *Figure 16*, in which the models' median absolute errors are much more similar than the RMSE. Although there isn't one model with the lowest prediction error, *Figure 16* also shows that the spread of errors for the DBN is much smaller than the other models. A Levene's test proved that this finding was statistically significant ($F = 2.80$, df = 4, $p = 0.03$), meaning that the DBN's predictions were more consistent and reliable compared to the other models.

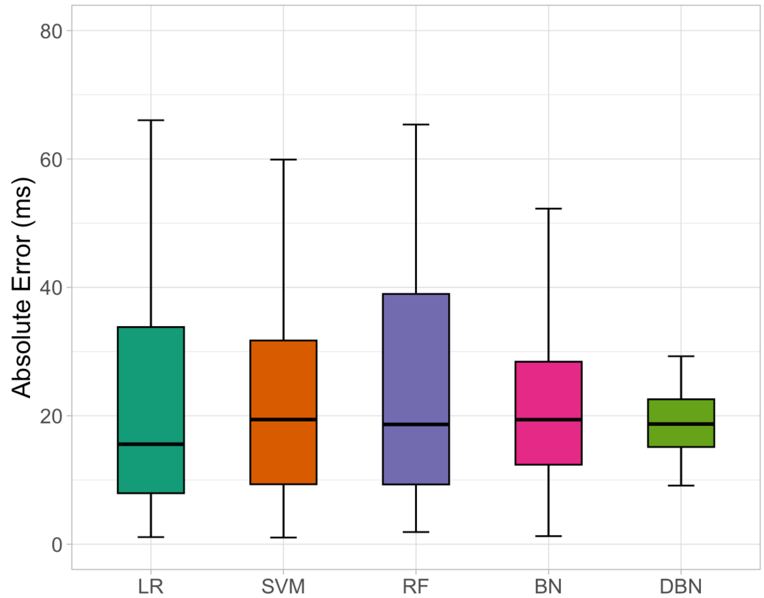

Figure 16. Box Plot of the PVT Models' Absolute Errors

5.6.2 Model Performance for the ECG Reading Task

Just as in the previous results section for the PVT, RMSE results from the vigilance models using the ECG reading task dataset are summarized in *Table 11*.

Table 11. Healthcare Vigilance Models Performance for the ECG (in s)

Model	**Linear Regression**	**Support Vector Machine**	**Random Forest**	**Bayesian Network**	**Dynamic Bayesian Network**
RMSE	1.35	0.95	1.10	1.20	1.37
95% CI	[1.03, 1.62]	[0.73, 1.13]	[0.93, 1.44]	[0.94, 1.41]	[1.12, 1.48]

The results from this analysis revealed somewhat similar findings compared to the PVT results. While the SVM had the lowest RMSE of 0.95 s, a Friedman's test across all models was insignificant (χ^2 = 9.16, df = 4, p = 0.06) again, but approached significance. *Figure 17* shows that the models' median absolute errors are similar in this dataset as well. Despite the visually similar spread in errors, a Levene's test showed that there was a statistically significant difference between the models (F = 2.71, df = 4, p = 0.03), where the DBN's predictions were still the most consistent and reliable among the vigilance models.

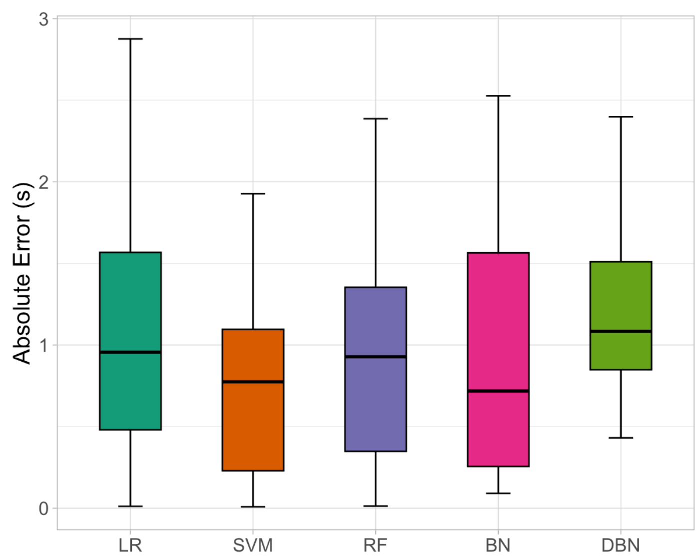

Figure 17. Box Plot of the ECG Models' Absolute Errors

5.7 Discussion

In this chapter, the initial dataset in *Chapter 4* was expanded to include additional participants and add a second, more naturalistic vigilance task: the ECG reading task. The two tasks, the PVT and ECG reading tasks, were contained in separate datasets for analysis. The vigilance modeling framework in *Chapter 4* was applied to these datasets and expanded by adding another step in the process: conducting statistical tests and creating data visualizations to compare predictive modeling strategies based on error magnitude and variance. The outcomes from this chapter are used to provide answers to this dissertation's RQs 1–3, particularly for the healthcare domain.

5.7.1 Contextual Factors Affecting Vigilance

In this chapter, the significant contextual factors for vigilance were sleep quality and room temperature. Based on results from the mixed-effects model, participants' sleep quality during the previous night and room temperature were both negatively associated with reaction time (i.e., positively associated with vigilance). These relationships indicate that having better sleep quality prior to hospital shifts and warm room temperatures can improve vigilance in healthcare settings.

5.7.2 Physiological Measures for Monitoring Vigilance

Among the metrics from the Empatica wristband, average BVP, BVP variability, average EDA, average skin temperature, and HR variability were significant variables in the predictive models. Higher heart rate variability and average electrodermal activity were associated with shorter reaction times and higher vigilance. Conversely, increases in average blood volume pulse and blood volume pulse variability predicted longer reaction times, signaling a decrease in vigilance.

For the eye tracking measures, average saccade duration, average pupil diameter, fixation count, and average fixation duration were significant variables for predicting vigilance. Longer average saccade duration was linked to longer reaction times and lower vigilance. In contrast, higher average pupil diameter, total fixation counts, and average fixation durations were all associated with better vigilance via short reaction times.

5.7.3 Predictive Modeling Features and Strategies for Vigilance

When comparing the two types of vigilance tasks in this chapter, the predictive models behaved similarly across both the PVT and ECG reading task datasets. While all models produced relatively similar predictive errors by magnitude, the DBN, which incorporates a time element to the analysis, had the lowest variance in its errors. This suggests that the DBN's ability to handle uncertainty and consider temporal relationships in the data makes it a more reliable and stable model for vigilance. The similar outcomes between the two vigilance tasks also points to the fact that the ECG reading task can be treated as a reliable, naturalistic task for vigilance in healthcare settings. Just as in *Chapter 4*, both contextual factors and physiological measures were needed to model vigilance.

Chapter 6: Sustainable Intervention Strategies for Vigilance Decrement

Abstract

This chapter identifies intervention tools for enhancing vigilance and considers the sustainability of available designs. It highlights the importance of considering sustainability when designing intervention tools for enhancing vigilance. Using a PRISMA literature review, 18 vigilance intervention strategies were identified across 52 studies. Each intervention was evaluated based on its sustainability, defined as the ability for vigilance to be maintained after external support is withdrawn. This chapter is used to answer RQ 4.

6.1 Background and Motivation

Despite the importance of promoting proper interventions during cases of vigilance decrement, when and how to use them remains unclear. Discrete and retrospective behavioral datasets that are only available after the point of vigilance decrement may contribute to the limited understanding and availability of interventions. Still, several approaches to mitigating vigilance decrement have been identified. Breaks (Helton and Russell, 2015; Ross et al., 2014) and changes in types of tasks (Greenlee et al., 2022) have been considered ideal approaches for handling vigilance decrement. Feedback on operators' performance (Hancock et al., 2016) is likewise recognized as a systematic way to mitigate vigilance decrement. Researchers have also proposed a radical and interruptive method of transcranial direct current stimulation to be used in targeted brain regions (Al-Shargie et al., 2019; Nelson et al., 2014).

There are few evaluations in the existing research on the sustainability of vigilance interventions. Such limited focus stands in contrast to the more robust body of research in public health, a field in which there is growing interest in understanding and evaluating the sustainability of interventions (Brownson et al., 2018; Hailemariam et al., 2019; Scheirer, 2013; Shelton et al., 2018; Walugembe et al., 2019). This strong interest has been driven by challenges that emerged when implementing and sustaining evidence-based healthcare interventions in practice, outside controlled environments (Shelton et al., 2018).

Although there is growing interest in intervention sustainability, there is no consensus in public health research on how to define the sustainability of interventions; rather, numerous methods of conceptualization, definition, and measurement exist (Shelton et al., 2018). One way to define intervention sustainability is as the capacity to maintain an intervention program at an adequate level after the termination of external support for the intervention, whether financial, technical, or managerial (LaPelle et al., 2006; US Agency for International Development, 1988). It has been found that more than 80% of existing articles include program activities that are maintained even after the cessation of external support in their definitions of sustainability (Lennox et al., 2018, Moore et al., 2017, Walugembe et al., 2019).

6.2 Research Goal

This literature review explores available interventions for mitigating vigilance decrement and then evaluates the sustainability of individual vigilance interventions. There is a need for a more comprehensive understanding of the sustainability of vigilance interventions to better inform when and how to implement the various types of interventions, as there are unique factors and circumstances influencing when vigilance decrement is observed. Understanding multiple aspects of intervention strategies for enhancing vigilance would help researchers design, implement, and adapt successful interventions that can be used continuously, even after the termination of external support.

6.3 Methods

Similar to *Chapter 3*, the PRISMA method was used to obtain and draw results from the literature. The review followed the four main steps in the PRISMA method: 1) identification of literature related to vigilance interventions based on search strings and exclusion criteria, 2) screening of literature using titles and abstracts, 3) eligibility determination based on a full-text review, and 4) analysis of results from the eligible articles. *Figure 18* shows an overview of the review process and results.

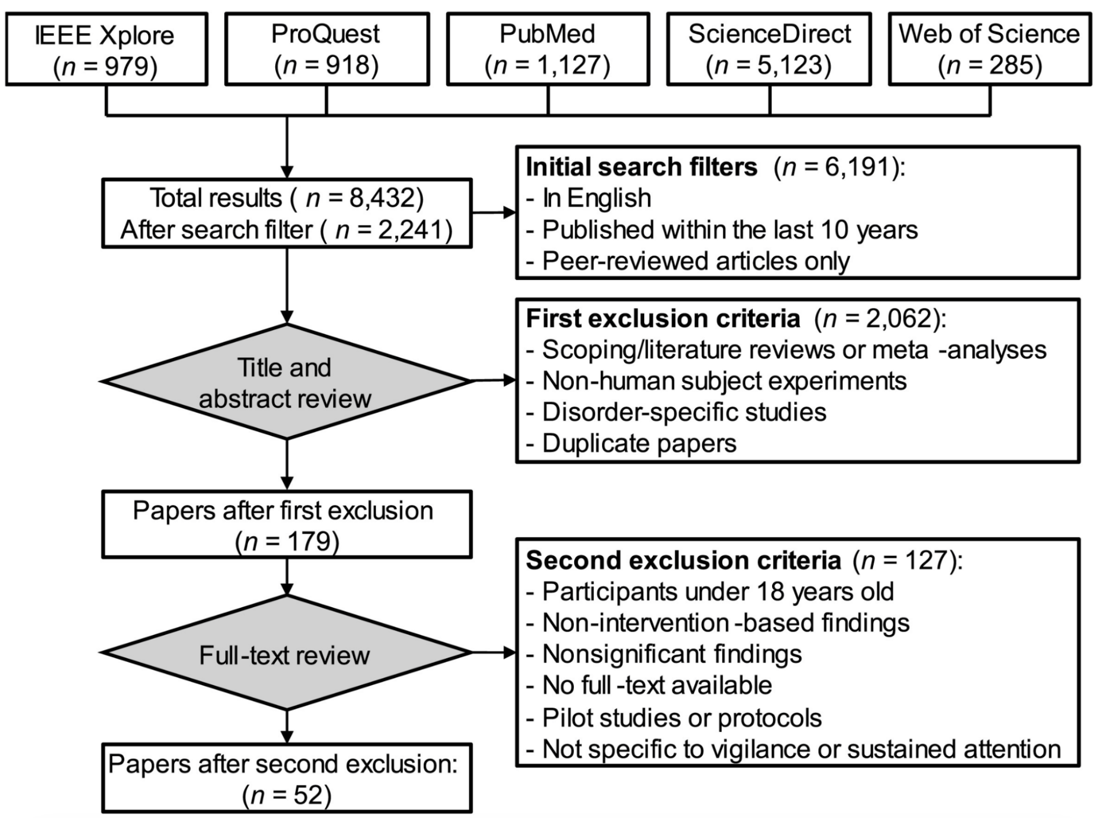

Figure 18. PRISMA Review Process for Vigilance Interventions

Five major databases—IEEE Xplore, ProQuest, PubMed, ScienceDirect, and Web of Science—were selected based on prior studies that reviewed interventions in the field of applied ergonomics and human factors (e.g., Choi et al., 2017; Islam & Lim, 2022). Within each of these databases, a common search string and exclusion criteria were used to find and filter articles relevant to vigilance interventions. The search string used was (vigilan∗ OR “sustained attention”) AND (interven∗ OR enhanc∗ OR improv∗) NOT (fatigue∗ OR drows∗ OR distract∗ OR sleep∗). The search string was developed with two considerations in mind. First, it includes a string for intervention (“interven∗“) as well as strings for enhancements or improvements (“enhanc∗” and “improv∗“) to account for alternative wording. Second, it isolates vigilance as the focus of interventions not only by including both “vigilan∗” and “sustained attention” but also by excluding related topics such as fatigue and drowsiness. The keywords “fatigue” and “drowsiness” were excluded for empirical purposes, as during the keyword formulation phase of our literature review, including the terms “fatigue” and “drowsiness” yielded a substantial number of papers that were not specifically focused on vigilance interventions (e.g., drowsy driving interventions). Also, from a theoretical perspective, the concepts of fatigue and drowsiness can be aligned with overload and underload theories, respectively (Pattyn et al., 2008). Each of these states only partially explains aspects of vigilance, further justifying the exclusion of the terms. This search yielded a total of 8,432 sources.

Excluded from the review were papers published more than 10 years ago, papers unavailable in English, and non-peer-reviewed works (i.e., the included papers were required to be journal articles or conference proceedings). Only relatively recent studies were included in this review to highlight the state-of-the-art in vigilance interventions. As shown in *Figure 18*, the exclusion criteria were applied over various phases of the search process to produce the most valuable results among the literature, based on recency, quality, applicability, and generalizability.

Four trained researchers divided up the papers from the databases and reviewed them based on their titles and/or abstracts. After screening the results, the research team uploaded the studies that passed the preliminary review to a shared reference management system. Once the articles had been loaded into this shared library, duplicates were removed, and a separate folder was created for each database's results. The researchers then swapped which databases they were working on in order to conduct an extended review of the articles' abstracts. This way, two researchers screened each database's results prior to the full review. Excluded from this phase were other literature reviews, meta-analyses, animal experiments, and disorder-specific experiments. At the end of this step, 179 articles remained for the full review.

Following the extended review of abstracts, a full-text PDF was added to each result in order to facilitate the full review. Papers with the following features were excluded during this phase: studies with participants under the age of 18 years old, studies with non-intervention findings or insignificant findings related to vigilance, studies whose full text was not available, pilot studies, protocol-based papers, and papers that were not directly related to vigilance or sustained attention (e.g., divided attention). Each paper was then critically reviewed by the research team using the review table in *Table 12*. After the full-text review was completed, 52 papers remained.

Table 12. List of Reviewed Papers for Vigilance Interventions

Article (author, year)	**Participants** (*n*)	**Task**	**Domain Application**	**Intervention strategy**
Abbasi et al. (2017)	6	Driving task	Transportation	Visual cue, visual-haptic cue
Al-Shargie et al. (2021)	20	Stroop color-word task	N/A	Audio stimulation of pure tone (250 Hz)
Axelsen et al. (2022)	459	Sustained Attention to Response Task	N/A	30-day app-based mindfulness and music intervention
Bajestani et al. (2024)	60	Automated Operation Span Task	Classroom	Mindfulness training
Barba et al. (2018)	24	PVT	N/A	Exercise-based arousal
Berding et al. (2021)	18	Rapid Visual Information Processing	Workplace	Dietary fiber supplement
Besten et al. (2023)	80	Sustained Attention to Response Task	N/A	Negative and positive mood induction
Bodala et al. (2016)	12	Intruder surveillance monitoring task	N/A	Challenge integration
Braun et al. (2015)	34	Sustained Attention to Response Task	N/A	Self-alert training using EDA feedback
Bruce et al. (2014)	60	Go/No-Go Task, CPT, Digit symbol substitution test	N/A	Citicoline-caffeine beverage
Caravati et al. (2024)	10	CPT	N/A	Transcranial direct current stimulation
Chai et al. (2019)	37	Rapid Visual Information Processing	N/A	Tart cherry juice consumption
Chaudhari et al. (2023)	22	PVT	N/A	LED-based transcranial photobiomodulation
Chen et al. (2018)	51	PVT	N/A	Sleep deprivation or not
Chen et al. (2023)	50	PVT	N/A	Acute Tai Chi Chuan exercise
Chin et al. (2021)	137	Dichotic listening sustained attention task	N/A	Meditation
Cintineo et al. (2022)	49	Go/No-Go Task	Military	Caffeine, methylliberine, and theacrine
Claypoole et al. (2019)	284	Two-digit numbers monitoring	N/A	Social presence: evaluative observer, merely present observer, electronic observer or no observer present
Esterman et al. (2016)	36	CPT	N/A	Task reward motivation strategies (monetary losses for errors)

Esterman et al. (2017)	15	CPT	N/A	Transcranial magnetic stimulation
Finkbeiner et al. (2016)	101	Abbreviated vigilance task	Workplace	Different types of task breaks
Flood et al. (2015)	95	Target detection task	N/A	Task-irrelevant emotive picture stimuli
Ford & Nagamatsu (2024)	38	Sustained Attention to Response Task	N/A	Meditation training
Harty & Cohen Kadosh (2019)	72	Continuous monitoring task	N/A	High-frequency transcranial random noise stimulation (1mA and 2mA)
Hwang et al. (2016)	60	PVT, Exercise Stress Test, Kaufman brief intelligence test	N/A	Laser light therapy, exercise
Jones et al. (2022)	40	NeuroRacer game, CPT	N/A	Transcranial electrical stimulation
Karampela et al. (2020)	40	CPT	N/A	Sensorimotor synchronization training with feedback
Kazinka et al. (2024)	78	Flanker, N-back, dot pattern expectancy task	N/A	Transcranial direct current stimulation
Kumar et al. (2016)	10	CPT	N/A	Isocaloric carbohydrate (mixed in a drink) ingestion
Lara et al. (2014)	27	PVT, Sustained Attention to Response Task	N/A	Sleep time (morning vs evening chronotype)
Liu et al. (2017)	5	Shooting, Dauf test	Professional shooting	Neurofeedback training
Löffler et al. (2018)	24	PVT	N/A	40 Hz gamma-transcranial alternating current stimulation
Luo et al. (2021)	14	Sustained Attention to Response Task	N/A	Tactile attention training
MacQueen et al. (2018)	69	CPT, Iowa Gambling Task, Wisconsin Card Sorting Task	N/A	10 mg or 20 mg of d-amphetamine
Ni et al. (2024)	19	D2 Test of Attention	Transportation	In-vehicle lighting (illuminance, correlated color temperature)
Pasanen et al. (2018)	272	Sustained Attention to Response Task	N/A	Nature walk with restoration-enhancement tasks/alternative tasks
Qi et al. (2021)	48	CPT	N/A	Qigong exercise
Ralph et al. (2017)	160	Mackworth Clock Test	N/A	Rest break, intervening task (a demanding visuospatial task)
Ritland et al. (2019)	50	PVT	Military	Sleep extension
Rostami et al. (2021)	21	Rapid Visual Information Processing	Security and surveillance	Transcranial alternating current stimulation

Ru et al. (2022)	18	PVT, N-back task	Workplace	Afternoon nap
Sanchis et al. (2020)	24	Attentional networks test	N/A	Caffeine ingestion, cyclergometer exercise
Schmalzl et al. (2018)	40	Response inhibition task	N/A	Meditation/ yoga
Sowinski et al. (2021)	26	Berg-Wisconsin Card Sorting Task, Go/No-Go Task, Sternberg Task Test, PVT, Cambridge Brain Sciences Reasoning and Concentration Tests, Light tracking test	E-gaming	Inositol-stabilized arginine silicate
Walters et al. (2019)	28	Sustained Attention to Response Task	Workplace	Cognitive training
Wang et al. (2015)	37	Video lecture and post-test for learning performance	School/ e-learning	Video lecture formats
Weisgerber et al. (2017)	21	Driving task	Transportation	Blue light illuminance
Xu et al. (2024)	33	2-back test, PVT	N/A	Neurofeedback training
Yuda et al. (2017)	20	PVT	N/A	Exposure to blue wavelength light
Yuda et al. (2018)	13	PVT	N/A	Chewing gum
Zhang et al. (2016)	20	Test of Variables of Attention	N/A	Rhythmic haptic stimuli
Ziegler et al. (2019)	42	Test of Variables of Attention	N/A	Meditation-based training via iPad

6.4 Results

Table 12 synthesizes the contexts, measurements of vigilance, and interventions for mitigating vigilance decrement in the reviewed papers. It includes the following five columns:

1. Article (year): Author information and publication year for the selected source.
2. Participants: Sample size of the study.
3. Task: Vigilance-related tasks performed by the participants during the experiment (if any).

4. Domain application: Targeted domain (if any).
5. Intervention strategy: Intervention strategy used to enhance vigilance during the study.

The average number of participants across the review studies was 57 (SD = 78). Of these participants, 42% were men and 58% were women. Three studies recruited only women, and another three studies recruited only men. Participants were on average 28 years old (range: 18–81 years old, SD = 16), with the majority of the studies recruiting younger adults. However, several studies specifically targeted older adults (aged 50 or above). These studies employed intervention strategies including Qigong exercises (Qi et al., 2021), meditation training (Ford and Nagamatsu, 2024), and tart cherry juice supplementation (Chai et al., 2019). Of the 52 studies, 71% were conducted in controlled laboratory settings to minimize the influence of extraneous factors. Nearly all studies (47) utilized standardized vigilance tasks to assess the vigilance level of each participant. Such tasks included the Sustained Attention to Response Task (SART), PVT, and CPT. Some studies used naturalistic tasks to simulate routine activities relevant to the target participant group or application domain. For example, Wang et al. (2015) employed a learning performance test sheet to evaluate the vigilance levels of college students following interventions related to lecture-based format. Using a simulated rural driving task in which participants were asked to brake for visual stimuli, Abbasi et al. (2017) assessed the impact of sensory cues on vigilance levels.

6.4.1 Vigilance Intervention Strategies

The 18 vigilance intervention strategies identified during the review are described with examples below.

1. Monetary rewards (*n* = *1*): monetary gain or loss due to performance; e.g., participants lost a small amount of money for each error they made (Esterman et al., 2016)
2. Peer pressure (*n* = *1*): the presence of others during tasks; e.g., a research assistant sitting behind participants while they were completing tasks (Claypoole et al., 2019)
3. Task context changes (*n* = *5*): alterations to the structure or components of tasks, such as providing emotive picture stimuli (Flood et al., 2015), using different video formats during lectures (Wang et al., 2015), mood induction during tasks (Besten et al., 2023), changing task difficulty (Bodala et al., 2016), and intervening demanding tasks (Ralph et al., 2017)
4. Sleep (*n* = *4*): sleep characteristics; e.g., no sleep deprivation (Chen et al., 2018), sleep extensions (Ritland et al., 2019), afternoon naps (Ru et al., 2022), or different sleep schedules (Lara et al., 2014)
5. Breaks (*n* = *2*): breaks during work activities; e.g., providing 45-second video breaks (Finkbeiner et al., 2016) or rest breaks (Ralph et al., 2017)
6. Changes in light color (*n* = *3*): altering the setting's color; e.g., via blue light (Weisgerber et al., 2017; Yuda et al., 2017) or different color temperatures (Ni et al., 2024)
7. Changes in illuminance (*n* = *1*): changes to the intensity of light provided during vigilance tasks; e.g., providing different levels of illuminance (Ni et al., 2024)
8. Electrical stimulation (*n* = *7*): transcranial stimulation; e.g., performing transcranial direct current stimulation (Caravati et al., 2024; Kazinka et al., 2024), transcranial alternating current stimulation (Löffler et al., 2018; Rostami et al., 2021; Jones et al., 2022), transcranial magnetic stimulation (Esterman et al., 2017), or transcranial random noise stimulation (Harty and Cohen Kadosh, 2019)

9. Sensory stimulation (*n = 3*): visual, auditory, or haptic stimulation; e.g., visual cues during a driving study (Abbasi et al., 2017), auditory stimulation at 250 Hz (Al-Shargie et al., 2021), or rhythmic haptic stimuli (Zhang et al., 2016)
10. Transcranial light (*n = 2*): LED/laser light delivery to the head of a participant; e.g., transcranial photobiomodulation (Chaudhari et al., 2023) and transcranial infrared laser stimulation (Hwang et al., 2016)
11. Neuro/biofeedback (*n = 3*): feedback signals to a participant's brain based on brain activity (Liu et al., 2017, Xu et al., 2024) or electrodermal feedback (Braun et al., 2015)
12. Sugar-based drinks (*n = 3*): ingestion of particular nutrients in carbohydrates and fruits, by providing participants with tart cherry juice that contains flavonoids, such as anthocyanins and proanthocyanins (Chai et al., 2019), isocaloric carbohydrate drinks (Kumar et al., 2016), or dietary fiber supplements (Berding et al., 2021)
13. Caffeine (*n = 3*): ingestion of caffeine (Bruce et al., 2014, Sanchis et al., 2020), such as caffeine capsules (300 mg) prior to performing vigilance tasks (Cintineo et al., 2022)
14. Drugs (*n = 2*): prescription or over-the-counter drugs; such as dextroamphetamine, a stimulant used to treat attention-deficit/hyperactivity disorder (MacQueen et al., 2018) or inositol stabilized arginine silicate, a supplement for athletic and cognitive performance (Sowinski et al., 2021)
15. Chewing gum (*n = 1*): chewing gum during tasks; e.g., giving participants spearmint gum and asking them to chew continuously for 25 minutes (Yuda et al., 2018)
16. Training (*n = 3*): learning a technique or skill aimed at improving vigilance through feedback; e.g., providing participants tactile attention training involving the use of an adaptive fingertip device for 40 minutes per day across 5 consecutive days (Luo et al.,

2021), sensorimotor synchronization training (Karampela et al., 2020), and cognitive training (Walters et al., 2019)

17. Exercise (*n* = *7*): light to moderate exercise, as determined by participants' physiological responses to these activities, such as elevated heart rate (Barba et al., 2018, Chen et al., 2023, Hwang et al., 2016, Pasanen et al., 2018, Qi et al., 2021, Schmalzl et al., 2018); e.g., having participants exercise using a cycle ergometer, similar to an indoor exercise bike, during vigilance tasks (Sanchis et al., 2020)
18. Meditation (*n* = *6*): meditation and/or mindfulness (Bajestani et al., 2024, Chin et al., 2021, Ford and Nagamatsu, 2024, Schmalzl et al., 2018, Ziegler et al., 2019); e.g., creating a mindfulness app for participants to use daily for 30 days (Axelsen et al., 2022)

Of the 18 intervention strategies, electrical stimulation and exercise were used the most, at 12% of all studies, each. Meditation was next at 11% of all interventions. Task context changes were the fourth most popular at 9%, followed by sleep at 7%. Training, changes in light color, caffeine, sugar-based drinks, sensory stimulation, and neuro/biofeedback interventions each accounted for 5% of all interventions observed in this review. Breaks, drugs, and transcranial light accounted for 4% each. The least used interventions were changes in illuminance, chewing gum, peer pressure, and monetary rewards at just 2% each.

6.4.2 Characteristics and Sustainability of Interventions

As sustainability is mainly defined by the ability to maintain interventions after the termination of external supports, it is critical to understand the characteristics of an intervention to evaluate its sustainability. *Table 13* summarizes each intervention's characteristics as they pertain to sustainability. Since interventions that are designed to mitigate vigilance decrement have characteristics that make them distinct from those used in healthcare and public health, criteria were developed to evaluate each intervention's characteristics in the context of vigilance research. Each intervention was evaluated according to the following characteristics: duration, frequency, assistance required, and equipment required for intervention delivery.

Table 13. Characteristics and Sustainability of Intervention Types

Intervention		Characteristics				Sustainability
		Duration	Frequency	Assistance required?	Equipment required?	
Incentives	**Monetary rewards**	Mid	Low	Y	N	Mid
	Peer pressure	Mid	Low	Y	N	Mid
Task changes	**Task context changes**	Mid	Low	Y/N	N	Mid
Sleep and breaks	**Sleep**	Long	Low/Mid	N	N	Mid
	Breaks	Short/Mid	Mid	N	N	High
Lighting	**Light color change**	Mid	Low	N	Y	Mid
	Illuminance change	Mid	Low	N	Y	Mid
Stimulation	**Electrical stimulation**	Mid	Low	Y	Y	Low
	Sensory stimulation	Mid	Mid	Y/N	Y	Mid
	Transcranial light	Mid	Low/Mid	Y	Y	Low
	Neuro/biofeedback	Mid	Mid	Y	Y	Low
Ingestion	**Sugar-based drinks**	Mid	Mid	N	N	High
	Caffeine	Short/Mid	Low	N	N	High
	Drugs	Short	Low	Y	N	Mid
	Chewing gum	Mid	Low	N	N	High
Training	**Training**	Mid	Mid	Y	Y/N	Mid
Exercise and meditation	**Exercise**	Mid	High	Y/N	Y/N	Mid
	Meditation	Mid	High	Y/N	N	High

The first criterion used for evaluating intervention structure was how long an intervention activity took ("Duration" in *Table 13*), where "short", "mid", and "long" refer to less than 5 minutes, between 5 minutes and 2 hours, and longer than 2 hours, respectively. This temporal scale takes into account that vigilance decrement is usually observed between 5 and 30 min after starting a task (Hancock, 2013; Mackworth, 1948; Nuechterlein et al., 1983; Warm et al., 2018). The average duration of an intervention activity across the 52 articles was 37 minutes. Sleep interventions were the longest in duration, while taking a break and ingesting caffeine were among the shortest.

The frequency with which an intervention was repeated (“Frequency” in *Table 13*) was the second criterion we used for evaluating intervention structure. We categorized the intervention frequency as “low”, “mid”, or “high” depending on whether the intervention occurred once, between one and seven times, or more than seven times, respectively. This scale was set up based on our analyses of experimental and intervention procedures in the reviewed papers. In these studies, intervention strategies were repeated seven times on average. Physical exercise required more than seven instances to produce effects, while incentives, changes in lighting, and most ingestion-related interventions required only one instance.

Evaluating whether assistance from others was necessary for the person receiving the intervention was the third evaluation criterion (“Assistance required” in *Table 13*). Incentives, both monetary and peer pressure-based, required others to evaluate the intervention recipients’ performance and to compete with the recipients. Brain stimulation required skilled staff to set up devices and follow technical procedures. Training typically required an instructor, whether in-person or online, and drug ingestion required clinicians who could prescribe the drugs. In contrast, task changes, sensory stimulation, exercise, and meditation required little assistance, and the intervention recipients were able to deliver and receive the intervention by themselves after minimal training.

Whether special equipment or devices were required to conduct an intervention was the final criterion we used to analyze intervention structure ("Equipment required" in *Table 13*). Changing the light color or luminance of an environment required multilight and dimmable light bulbs and fixtures, and simulation interventions could not be conducted without neuroimaging and sensory/brain stimulation devices. Training and exercise sometimes required computers or fitness equipment, depending on the training context and exercise type.

Finally, the sustainability of each intervention was evaluated in reference to the most frequently used and accepted definition of sustainability (Lennox et al., 2018; Moore et al., 2017): the intervention's ability to be maintained after the withdrawal of external support ("Sustainability" in *Table 13*). Based on this definition, the aforementioned intervention characteristics—duration, frequency, assistance required, and equipment required—were taken together to assess sustainability. The four characteristics selected were inferred from healthcare intervention research findings showing that capacity, funding, staffing/turnover, training, and technical support are significant factors in determining an intervention's sustainability (Shelton et al., 2018; Walugembe et al., 2019). "High" sustainability describes an intervention strategy that requires less than 2 h to implement (short to mid-duration), fewer than seven repetitions (low-to mid-frequency), and does not necessitate assistance or equipment that could incur additional costs for training and technical support. "Low" sustainability describes an intervention that takes longer than 2 h (mid-to long duration) or more than seven repetitions (mid-to high frequency) and requires both assistance and equipment. "Intermediate" sustainability describes all interventions that cannot be categorized as either "high" or "low."

Overall, task breaks, sugar-based drinks, caffeine, chewing gum, and meditation were evaluated as highly sustainable interventions for mitigating vigilance decrement based on participants' ability to maintain them after the termination of external support. In contrast, electrical stimulation, transcranial light, and neuro/biofeedback were considered to have low sustainability due to participants' inability to perform such lengthy and repeated activities without the support of trained assistant(s) and stimulation devices.

6.5 Discussion

This scoping review was conducted to understand and evaluate the sustainability of interventions observed in vigilance research. Following the literature search, 52 studies were reviewed and 18 types of intervention strategies were identified. These intervention strategies were then assessed based on their sustainability—their ability to remain effective over longer periods of time and without continuous support. None of the 52 articles evaluated or suggested the importance of sustainability for vigilance interventions, leading to a unique, structured evaluation of sustainability.

Based on the vigilance intervention evaluations, the less sustainable strategies—electrical stimulation, transcranial light, and neuro/biofeedback—were all technology-based, suggesting that the need for equipment or devices to perform complex procedures hinders intervention recipients from consistently receiving the intervention's benefits. The highly sustainable interventions were task breaks, sugar-based drinks, caffeine, chewing gum, and meditation. From a human factors perspective, these interventions are often self-imposed and difficult for researchers to reliably replicate in naturalistic settings. Instead, another intervention to consider is the borderline-highly sustainable "task context changes" strategy, where if the intervention is designed to avoid assistance requirements, it can be regarded as highly sustainable. This intervention strategy is thus implemented in *Chapter 7*.

Chapter 7: Vigilance Modeling and Intervention Design in the Driving Environment

Abstract

This chapter applied the vigilance modeling framework to the driving domain and explored the effectiveness of a sustainable intervention. In this chapter, a driving simulator was used to simulate a scenario in semi-autonomous driving: engaging in a virtual meeting task while the system is in control. A takeover event, in which the driver retakes manual control of the vehicle due to a system failure, served as a means for measuring drivers' vigilance toward the end of the virtual meeting. Two takeover warning designs were implemented as interventions during the virtual meeting and compared based on their respective effects on drivers' vigilance. This chapter is used to answer RQs 1–4 for the driving domain.

7.1 Background and Motivation

The number of people who work remotely has increased substantially. According to the 2020 U.S. Census data, only 3% of people worked remotely. In 2020, the number increased to 7% and in 2021, the number was 18% (Silver, 2023). The sophistication of existing video conferencing software (e.g., Zoom, Webex, and Microsoft Teams) has made it easier for individuals to have a more immersive online meeting experience. This has also promoted more virtual meetings (Tolliver & Sass, 2024; Zaki, 2023) and allowed workers to be more efficient outside the workplace, thereby reducing commute times (Kugler, 2022). Many software packages for virtual meetings also include recording, AI note-taking, whiteboards, and screen-sharing to make meetings more effective (Kugler, 2022). However, virtual meetings from home may create distractions due to roommates, family, or others who are nearby (Karl et al., 2022).

Semi-autonomous vehicles can provide a comfortable space for drivers to participate in meetings during their commute or between errands.

Recent advancements in vehicle automation have provided an opportunity for more seamless virtual meeting experiences in the car. In 2023, the Mercedes-Benz Drive Pilot became the first production-approved SAE Level 3 Automated Driving System (ADS) which allows the driver to relinquish full control of the vehicle for short periods of time (Edward, 2023; SAE International, 2021). Hence, more drivers will be able to hand over control to the ADS without the need to continually attend to the roadway. Although the ADS can handle most aspects of the driving task, drivers must be prepared to resume manual driving when needed. These transfers of control from the ADS to the driver are called takeovers. Takeovers during secondary task engagement are significantly more complex for drivers to respond to since they are not constantly monitoring the ADS (Morales-Alvarez et al., 2020). Thus, drivers' behavior during takeovers may be used to assess their vigilance levels. In particular, McWilliams and Ward (2021) identified response times during safety-critical events as the most commonly used measure for vigilance decrement during partially-automated driving.

The human-machine interaction of takeovers has also been explored by safety researchers to design takeover warnings and present them to the driver in an appropriate manner. These takeover warnings have been designed using a variety of message modalities. For example, visual-based messages have used icons (Radlmayr et al., 2014), text (Roche et al., 2019), or combinations of the two (Naujoks et al., 2019) in order to direct drivers' attention to the takeover. Although there is little consensus on message timing specific to Level 3 ADS, several literature reviews and meta-analyses have identified 2–6 seconds as the average lead time for takeover warnings across multiple automation levels (Eriksson & Stanton, 2017; Zhang et al., 2019). In the context of vigilance, these alerts could serve as a sustainable intervention for drivers experiencing vigilance decrement during semi-autonomous driving (see *Chapter 6)*.

Studies exploring secondary task engagement in ADS generally focus on the secondary task type. Different secondary task modalities can range from visual tasks, such as watching movies or TV shows (Du et al., 2020), to auditory tasks, such as listening to audio books (Naujoks et al., 2019), or visual-manual tasks, such as playing video games (Ou et al., 2021). Other common types of secondary tasks that mimic real-life interactions include calling/texting or reading physical media (Naujoks et al., 2019). Despite the wide array of secondary tasks used in takeover studies, virtual meetings have been left unexplored. Vigilance research in the driving domain has identified driving automation as a key source of vigilance decrement due to the monotony and cognitive underload associated with typical automated driving (McWilliams & Ward, 2021). To counter this situation, Mishler and Chen (2023) used secondary tasks to effectively mitigate vigilance decrement in Level 2 automated vehicles. However, these verbal tasks only consisted of general knowledge and roadway-specific questions during the drive. Engaging in a more naturalistic task, such as virtual meetings, while driving in a vehicle with a higher level of automation may yield different outcomes.

Vehicles equipped with Level 3 ADS present a unique environment for virtual meetings to be held. While the ADS handles the driving task, the driver can actively participate in a short, impromptu virtual meeting without having to constantly supervise the ADS. Thus, this study uses a virtual meeting as a naturalistic task for drivers during semi-autonomous driving. At the same time, a takeover situation is used to measure drivers' vigilance levels while takeover warning prototypes are designed and assessed as potential interventions. This chapter aims to answer RQs 1–4 by applying the vigilance modeling framework from the healthcare domain (see

Chapters 4 and 5) to identify the contextual factors and physiological measures associated with vigilance in driving, build a predictive model specific to the driving domain, and compare takeover warning designs and their effect on vigilance.

7.2 Participants

Participants for this study were recruited from regions within driving distance of Seattle. To be eligible, participants were required to have a valid U.S. driver's license, be at least 18 years old, and have prior experience participating in virtual meetings and designing presentation slides. Participant age groups were divided between 18–29 year-olds and 30–49 year-olds based on U.S. census data from a Pew Research report on technology use (McClain et al., 2021). The report showed that these two census groups have nearly identical population densities across the U.S. and were the only age groups with over half of the population having experience with virtual meetings. The research team aimed to sample the participant population as equally as possible across age and gender categories. However, due to some data quality issues and various technical issues while using the driving simulator, 11 participants' data were excluded from the final dataset, resulting in a sample size of 36 participants. Thus, there were 19 participants in the 18–29 year-old group and 17 participants in the 30–49 year-old group. The average age of the participant pool was 29.5 (SD = 8.7) years old. For gender, 22 participants were female, 14 participants were non-female.

This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board at the University of Washington. Participant outreach included social media and online forums (e.g., Reddit, Facebook, Craigslist) ad posting, posting physical flyers around the university, sending recruitment emails to the lab's past participants and various departments across the university, and through word-of-mouth referrals. Interested individuals filled out an online eligibility survey and were contacted by the research team to sign up if eligible. Participants received IRB-approved consent forms via email prior to their driving session and had a second opportunity to review the form upon arrival to the research lab. Each participant's informed consent was obtained and recorded before any experiment procedures could begin. All participants were paid $25 per hour through Tango gift cards.

7.3 Apparatus and Materials

This section describes the different apparatus and materials used to collect data in this driving experiment.

7.3.1 Driving Simulator

The experiment used a fixed-base miniSim™ driving simulator to mimic real-world takeover situations in semi-autonomous vehicles. The software and hardware were developed by the University of Iowa's Driving Safety Research Institute (DSRI). The driving simulator features a Quarter Cab configuration (see *Figure 19*), offering three 48" 1920x1080 displays, a 138° horizontal and 27° vertical field-of-view, a real vehicle steering wheel and seat, a dashboard and instrument panel, and various other car features (e.g., turn signals, gear shifter). In addition, the driving simulator was equipped with the miniSim AutoDriver™ Vehicle Automation package to create scenarios with autonomous driving capabilities. Driving scenarios were built and designed using miniSim's Tile Mosaic Tool (TMT) and Interactive Scenario Authoring Tool (ISAT).

Figure 19. MiniSim Quarter-Cab Driving Simulator

7.3.2 Takeover Warning-Based Intervention Design

The research team designed a series of continuous visual and auditory messages to indicate the status of the ADS to the driver. Using miniSim's ISAT, the message designs were implemented into the driving scenarios so that visuals and sounds activated based on changes in the automated driving system's state (on/off) and within a specific time-to-arrival of a work zone. Work zones have been identified as potential sources of failure in semi-autonomous vehicles due to sensor identification issues (Ansarinejad et al., 2025; Shi & Rajkumar, 2021). In this study, the work zone is depicted by maintenance crews and vehicles blocking a lane on the highway and serves as the takeover event. The four icons used in the ADS status display are shown in *Figure 20*. The Off Status (Icon A) depicts a gray steering wheel and indicates that the ADS is off and inactive. The On Status (Icon B) was a blue steering wheel icon with ripples on the sides to show that the ADS is on and active. To enhance drivers' vigilance ahead of the work zone, two takeover warning messages were designed with unique timing, visuals, and auditory features: a less urgent, *Planned* warning and an *Emergency* warning. Half of participants in this experiment received the *Planned* warning while the other half received the *Emergency* warning.

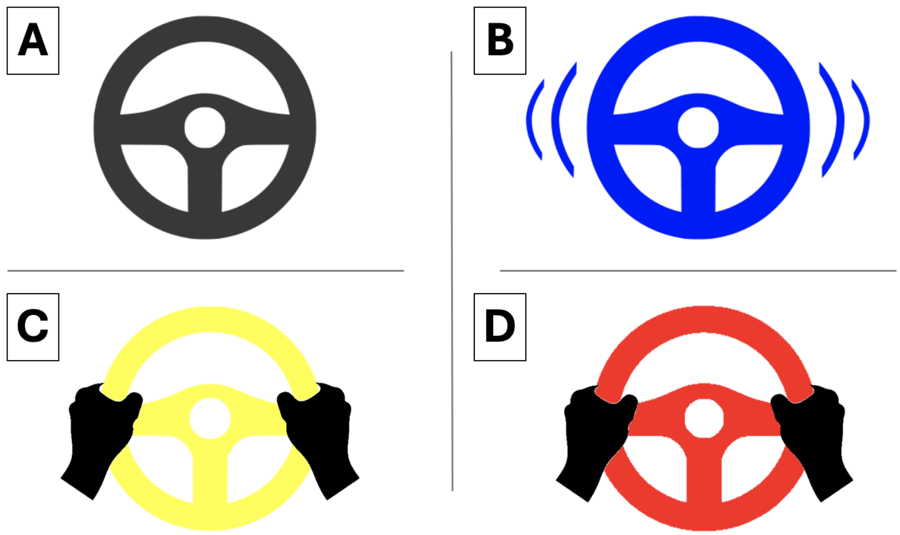

Figure 20. ADS Status Icons

The *Planned* takeover alert (Icon C) was a yellow steering wheel with hands on the wheel. At 15 seconds from the work zone, two successive auditory beeps were played with Icon C—similar to a notification-style alert. The *Emergency* takeover alert (Icon D) was a red steering wheel with hands on the wheel that appeared 6 seconds from the work zone. *Figure 21* shows the timeline of both takeover warning designs leading up to the work zone. The audio that accompanied Icon D was a series of consecutive beeps that played until the participant began the lane change. To summarize, the *Planned* warning prioritized giving drivers additional time with less urgency to perform the takeover while the *Emergency* warning used a more direct, salient approach to promote vigilance.

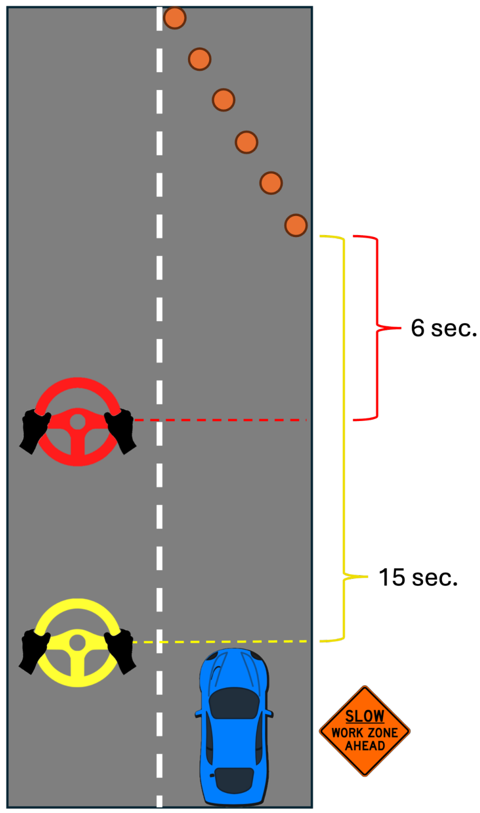

Figure 21. Timeline of Takeover Warnings

The takeover warning-based interventions in this study resemble a highly sustainable version of the "task context changes" intervention type from *Chapter 6*. Since no external assistance is needed to activate the takeover warning intervention, the system-integrated nature makes the designs highly sustainable. Although the scoping review revealed several other highly sustainable interventions, including sleep, ingestion, and meditation, a task context-related intervention was selected for this semi-autonomous driving scenario since human-machine interaction is one of the primary design aspects used by human factors professionals to promote safety. The main goal of task context-related interventions is to enhance vigilance by affecting the vigilance task itself. Thus, the two designs are compared in order to highlight the key aspects of the alerts that most effectively contribute to higher levels of vigilance.

7.3.3 Virtual Meeting Task

The virtual meeting in this experiment simulated a real-world virtual work meeting. The participants' main task was to work with other meeting attendees to create a slide deck for presentation to a management team. The meeting consisted of four members, the meeting host, the participant, and two other virtual attendees. The responsibility of the virtual meeting host was to guide the attendees (including the participant) through slide designs and present the criteria for the best design. The two virtual attendees shared their pre-designed slides along with the virtual host. The participant was responsible for choosing the best design of three. For example, one slide selection began with the host giving the following two criteria: "the title slide should have a large font and only one picture". The participant would then observe the host and other two attendees' present their slide designs one at a time. After the slide presentations, the meeting host asked the participant "which slide design do you think we should use, slide 1, slide 2, or slide 3?". The participant would then respond with the slide they felt best matched the two criteria in the host prompt. *Figure 22* shows the participant's view of the virtual meeting.

Figure 22. View of the Driver During the Main Drive

The theme of the virtual meeting was focused on how to make cheeseburgers. The meeting was separated into eight slide selection subtasks: 1) title slide, 2) outline slide, 3) definition slide, 4) meat patty slide, 5) cheese slide, 6) toppings slide, 7) condiments slide, and 8) cooking instructions slide. For each subtask, only one slide matched the two criteria presented by the meeting host. The other two had either one match or no match. The correct slide was randomly ordered. Depending on the participant’s answer, the response was recorded as correct, semi correct, or incorrect. These responses were recorded in real time using a Google Form. No feedback was provided to the participant on their answers.

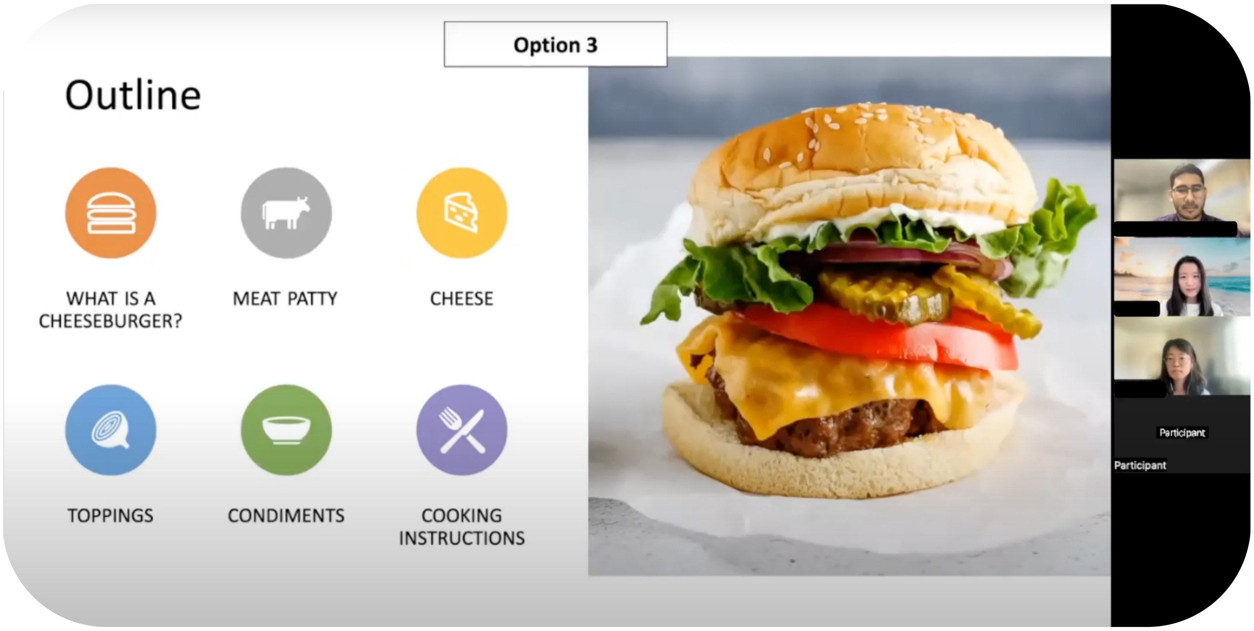

Figure 23. Sample Frame from the Virtual Meeting

For consistency across all participant interactions, the virtual meeting functioned as a pseudo-live meeting—the meeting was scripted and prerecorded over Zoom with the other 3 attendees (all members of the research team) then played for the participant during the drives (see *Figure 23*). Timed pauses were used in the script to allow for participant responses. The virtual meeting video was played remotely on an Apple iPad Pro using a Magic Mouse. The iPad was mounted on a stand next to the participant in the general location that an in-vehicle display would be. Participants were allowed to adjust the iPad position so that they could see the screen properly during the drives.

7.3.4 Questionnaires

During this study, six questionnaires were used to collect contextual factors related to the participants. Prior to enrolling in the study, participants filled out an eligibility survey. The eligibility survey covered demographic information (e.g., age, gender, education), driving activity, and prior experience with virtual meetings.

The other five questionnaires were adopted from validated scales and used to assess trust in ADS, situation awareness, workload, distraction susceptibility, and acceptance of ADS technology. The trust questionnaire was adapted from Lee and See's (2004) trust in automation questionnaire and consisted of twelve 10-point Likert scales.

Situation awareness was assessed using the Situation Awareness Rating Technique (SART). The SART was developed by Taylor (1990) as a measure of situation awareness that can be administered after study tasks—compared to the popular Situation Awareness Global Assessment Technique (SAGAT) which requires pauses during tasks. The questionnaire featured 7-point Likert scales within three main domains related to situation awareness 1) demand (on attentional resources), 2) supply (of attentional resources), and 3) understanding (of the current situation).

Participants' mental workload during the experiment was assessed using an abbreviated version of the NASA TLX (Hart & Staveland, 1988). Just as in the healthcare vigilance experiment described in *Chapter 4*, this adapted NASA TLX measured six key aspects of workload: 1) mental demand, 2) physical demand, 3) temporal demand, 4) performance (during the task), 5) effort (required to complete the task), and 6) frustration (during the task).

Participants' susceptibility to distraction was measured using the Susceptibility to Driver Distraction Questionnaire (SDDQ; Feng et al., 2014). The SDDQ consisted of 5-point Likert scales grouped into three main sections: 1) engagement in distraction while driving, 2) attitudes and beliefs about voluntary distraction, and 3) susceptibility to involuntary distraction.

Lastly, acceptance of ADS technology (i.e., attitudes and willingness to use ADS) was assessed using Van der Laan and colleagues' (1997) technology acceptance questionnaire. The questionnaire included nine 5-point Likert scales with questions regarding the participants' opinions of ADS.

7.3.5 Eye Fixations

Participants' eye fixations during the experiment were collected using manual coding of video data recorded at 60 FPS. The camera was positioned directly above the driving simulator's center display to capture the front-view of participants during their drives. Eye fixations between the road and the iPad (i.e., the virtual meeting) were coded using Behavioral Observation Research Interactive Software (BORIS; Friard & Gamba, 2016). BORIS is a user-friendly, open-source software that uses audio and video inputs for event logging. Within BORIS, significant events during the drive were logged, including each virtual meeting task, the entire virtual meeting, and the takeover event. This method expedited the eye fixation data processing procedure (see *Data and Processing* section). Three researchers used BORIS to process the data and completed synchronized training together to ensure consistency prior to data processing.

7.4 Procedure

This section describes the study timeline from participant arrival to the lab and post-experiment participant debriefing. *Figure 24* below provides an overview of the experiment procedure.

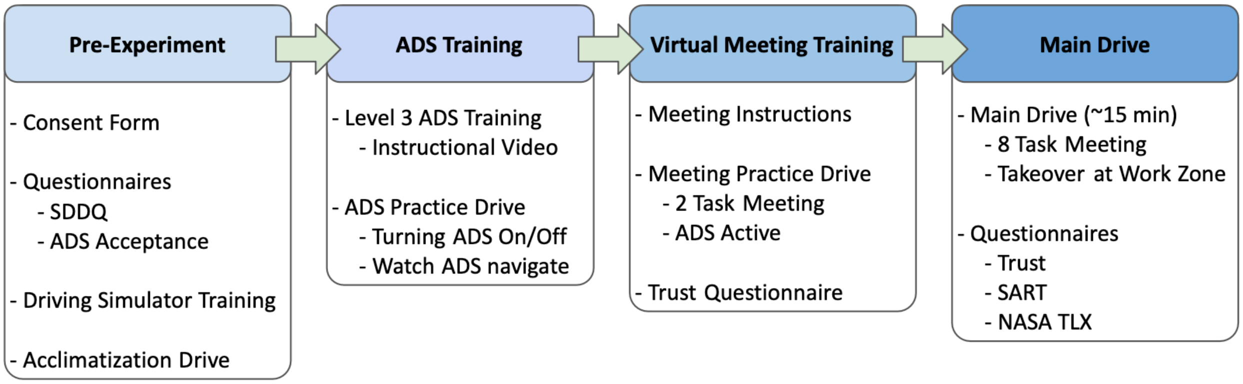

Figure 24. Overview of the Driving Study Procedure

7.4.1 Pre-Experiment

Participants received an online consent form via email after their appointment was confirmed. Upon arrival at the research lab, participants met with a member of the research team and provided their informed consent prior to the start of the study. After the initial intake process, the participants completed a pre-experimental questionnaire which included the SDDQ and automated vehicle acceptance scales. Next, participants were directed to the driving simulator to familiarize themselves with its features and how to drive in it. Participants completed a short acclimatization drive in which they practiced turning on the car, accelerating, maintaining speed, slowing down, and coming to a full stop.

7.4.2 ADS Training and Practice

Following the acclimatization drive, participants were trained on how to use the driving simulator's Level 3 ADS using a YouTube video with an overview of driving automation. After watching the video, participants also viewed an information sheet that explained the driving simulator's automation status display. Once the participant had sufficient time to review the information sheet, they were escorted back to the driving seat of the simulator to learn about how to use the ADS. Once the participants were comfortable with the capabilities, displays, and mechanisms of the driving simulator's ADS, they drove through a practice drive. The research team designed the ADS practice drive so that the driver could gain familiarity with the feel of the ADS (e.g., observing the vehicle follow curves) and the different ways they can turn the ADS off. This practice was important in preparation for the takeover event in the main drive.

7.4.3 Virtual Meeting Training and Practice

Participants practiced engaging in the virtual meeting task. Participants were escorted back to the driving simulator, and the iPad was turned on and swiveled on its stand for the participant's preferred view. Verbal instructions were given by the research assistant regarding the virtual meeting format. Participants then drove through the meeting practice drive. The drive featured the same loop as the practice ADS drive, but without any instructions. Participants were instructed to turn on the ADS at the beginning of the drive and to allow it to drive for them while the virtual meeting was held. The research assistant started the virtual meeting video on the iPad remotely after the ADS was activated and recorded the participant's responses. The practice meeting matched the format of the virtual meeting in the main drive, except the practice was limited to two questions and the theme was about cats and dogs. A trust questionnaire was administered to capture drivers' pre-main drive levels of trust in the ADS.

7.4.4 Main Drive

Participants were instructed to begin the drive the same as the virtual meeting practice drive: with ADS on and the car self-driving. The drive starts with the vehicle navigating to the highway. Once on the highway, the research assistant remotely started the virtual meeting.

Toward the end of the meeting, each participant encountered a work zone that blocked their lane on the highway, necessitating a manual takeover of the vehicle. Participants experienced either the planned or emergency takeover warning. The planned version was a less urgent, informative alert presented 15 seconds before the work zone. The emergency version was an urgent alert presented 6 seconds from the work zone.

The takeover event was designed to occur in the middle of the 7th subtask so that the participant had to manage the two tasks simultaneously (see *Figure 25*). Participants were not warned of the takeover prior to the drive in order to elicit as natural of a response as possible. Instead, they were provided sufficient knowledge and information about the ADS through the practice drives. After the virtual meeting ended, participants were instructed to turn off the ADS and park the car. The main drive typically lasted 15 minutes.

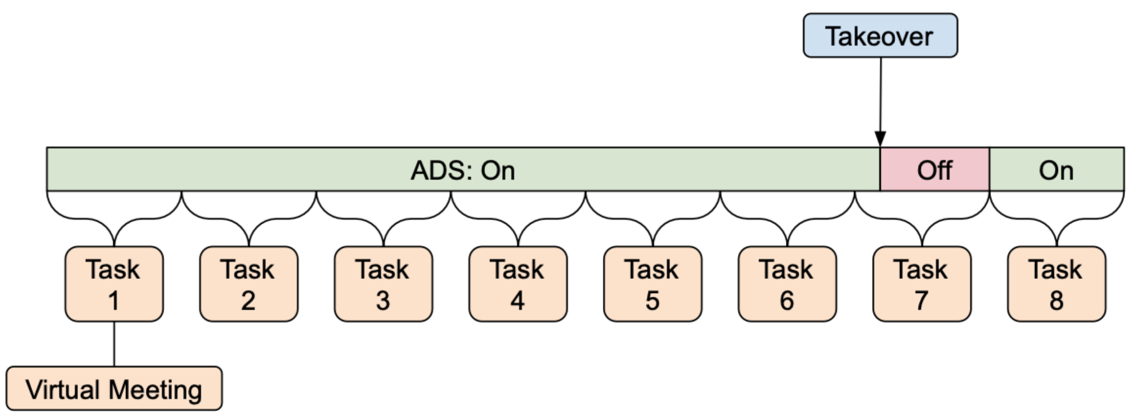

Figure 25. Overview of Events During the Main Drive

After completing the main drive, participants filled out the trust, situation awareness, and workload questionnaires were administered to collect pre- and post-takeover measures of these constructs. The research assistant explained the takeover situation, plans for the collected data, and answered any questions the participant had, then paid them for their time.

7.5 Experimental Design

The experiment was set up as a 2 (takeover warning; between) x 2 (virtual meeting stage; within) mixed-factorial design. Participants either received the planned or emergency takeover warning and all took part in the virtual meeting. The virtual meeting had two levels: prior to the takeover and during the takeover. The 'prior to takeover' level was based on the average of 6 meeting subtasks. These subtasks were grouped together because they all occurred while the ADS was active and prior to the takeover situation. The 'during takeover' level contained the 7th meeting subtask and contained the takeover event entirely.

7.6 Data and Processing

This section describes the study's methods for cleaning and aggregating the raw data from different sources and characteristics of the resulting dataset.

7.6.1 Driving Simulator Data

Data from the driving simulator were saved after each drive as Data Acquisition (DAQ) files. The DAQ files include time-series data for many vehicle variables, including speed, brake/gas pedal inputs, and steering inputs. The simulator data were collected by the system at 60 Hz. Each participants' main drive DAQ files were processed in MATLAB using miniSim's nDaqTools package. First, since the ADS was in full control of the driving task prior to the takeover event, the data were segmented to isolate the takeover. During this portion of the drive, participants were transitioning toward controlling, and then in control of, the driving task. Thus, the data from the simulator reflected the participants' inputs during this time rather than the ADS's. The raw data were then aggregated to obtain mean and standard deviation values of these measures during the takeover event. Instead of another physiological measure, driving behavior data were used as a substitute due to the fact that this type of data is often readily available in vehicles and can provide key insight into vigilance.

7.6.2 Eye Fixation Data

After the main drive's events and participants' fixations were logged using BORIS, the data were exported as an individual CSV file per participant. The data were then processed into four groups of variables for fixations to the road and fixations to the iPad: 1) number of fixations, 2) total duration of fixations, and 3) average duration per fixation. These variables were aggregated over the pre-takeover event period (i.e., virtual meeting subtasks 1–6) and during the takeover event. This resulted in 12 total eye fixation variables.

7.6.3 Virtual Meeting Data

Participants' responses during the virtual meeting were processed and scored according to the meeting script. All subtasks were graded as either 0, 0.5, or 1 if the participant answered incorrectly, semi-correctly, or correctly (respectively). Scores for subtask 1–6 were averaged to show meeting performance prior to the takeover. The score for subtask 7 was used for performance during the takeover event.

7.6.4 Questionnaire Data

All questionnaire data were scored according to the original authors' protocol. For the trust in ADS questionnaire, the Likert scales were averaged and produced a single value for participants' trust level where higher scores indicate higher trust. Situation awareness via the SART was evaluated using the authors' formula SA = Understanding – (Demand – Supply). All Likert scales within the understanding, demand, and supply groups were averaged and then applied to the formula to produce a single value for participants' situation awareness. Participants' workload via the NASA TLX was calculated by averaging the Likert scale responses into a single value for workload. Participants' susceptibility to distraction was calculated by averaging the Likert scale responses in the SDDQ. Similarly, participants' acceptance of ADS was calculated by averaging the Likert scales in the adapted technology acceptance questionnaire.

7.6.5 Vigilance Measure

Participants' response time after receiving the takeover warning was used as the study's measure of vigilance. Response time was defined and calculated as the time between the takeover warning onset to the driver's first response (e.g., hand movement, brake press).

7.7 Data Analysis

After collecting and processing the study's data, the final dataset consisted of 132 variables. See *Appendix* for a detailed, tabular summary of the resulting dataset. This section contains the methods used to select features from the dataset, the different modeling schemes used to predict vigilance, and statistical tests for the takeover warning intervention design. The analysis follows the vigilance modeling framework developed and validated using the healthcare datasets in *Chapter 4* and *Chapter 5*.

7.7.1 Correlation Analysis

To address potential multicollinearity, a systematic feature reduction process was conducted on the initial dataset of 132 variables. Given the mixed nature of the data (containing both categorical and continuous scales), Spearman's rank correlation coefficients were calculated to identify highly redundant features. Variables with a correlation magnitude of $\rho \geq |0.8|$ were flagged for removal. This filtering, supplemented by a manual review for relevance to vigilance, reduced the feature set to 35 potential predictor variables. These 35 variables and their correlations are illustrated in *Figure 26*. Note that several fixation measures with high inter-correlations were intentionally retained for the regression-based feature selection phase due to their relatively high correlation with takeover response time. This approach ensures that while global multicollinearity is minimized, potentially sensitive fixation measures are preserved for vigilance modeling.

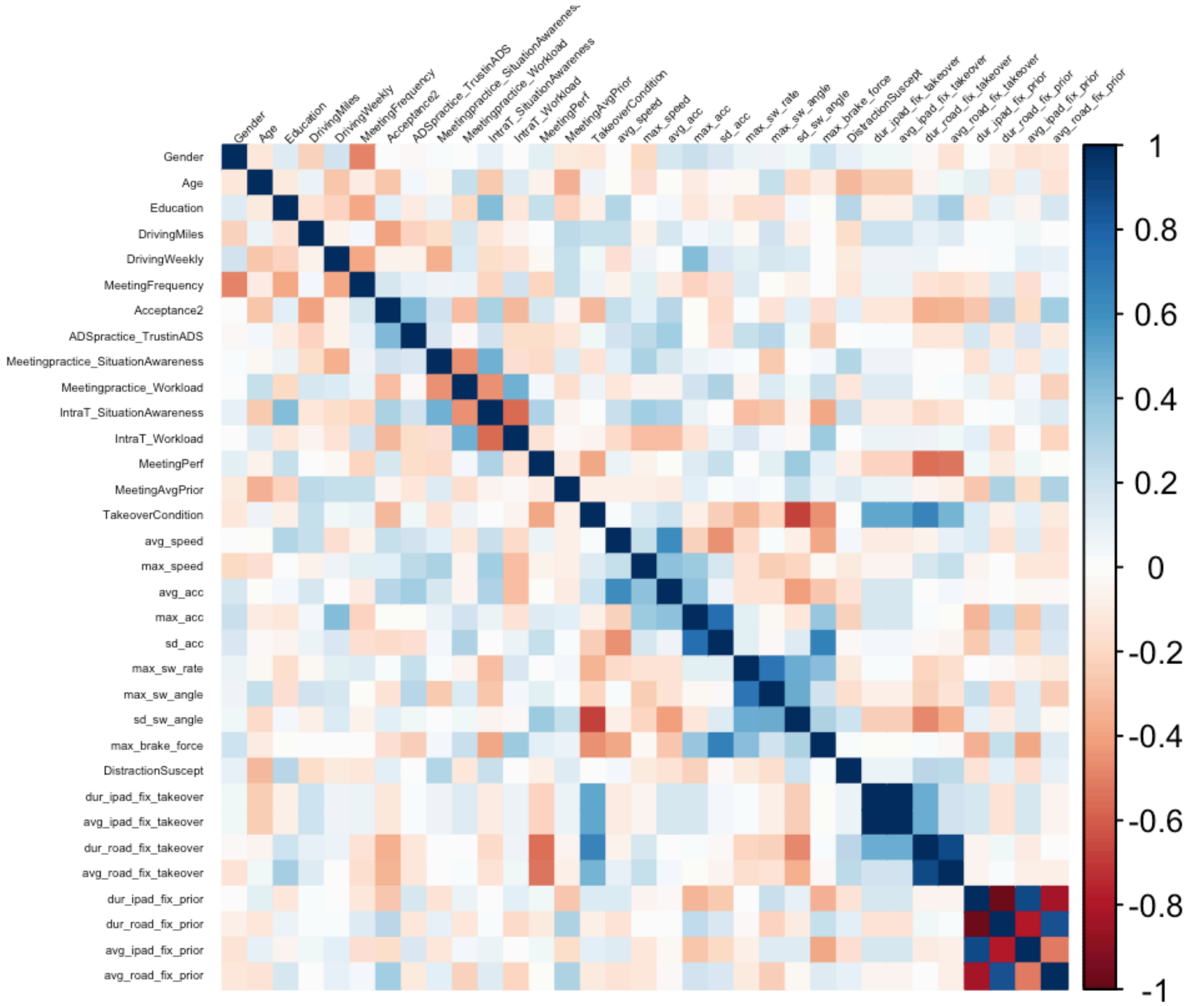

Figure 26. Spearman Correlation Plot for the Driving Dataset

7.7.2 Feature Selection

Prior to conducting the feature selection via linear regression, the distribution of response time was assessed. Visualization of a Q-Q plot revealed significant positive skewness and heavy right-tailed behavior in the distribution (see *Figure 27*), suggesting a violation of the normality assumption required for linear modeling.

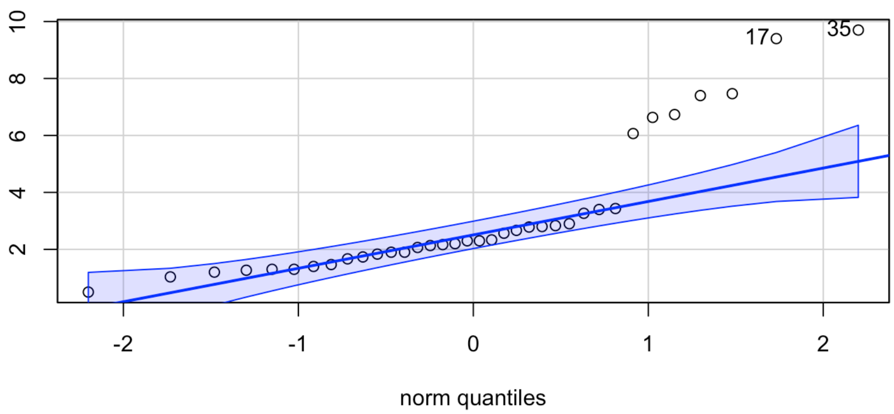

Figure 27. Q-Q Plot of Response Time Prior to Log-Transformation

To address this issue, a log-transformation was applied to the response time metric. Post-transformation visualization of the Q-Q plot indicated a substantially improved alignment with the normal distribution, with data points remaining largely within the 95% confidence bands (see *Figure 28*). This improvement was confirmed by a non-significant Shapiro-Wilk test ($W = 0.95$, $p = 0.10$), meaning the log-transformed variable can be assumed to fit a normal distribution. The log-transformed response time was used in all models for this study's analysis.

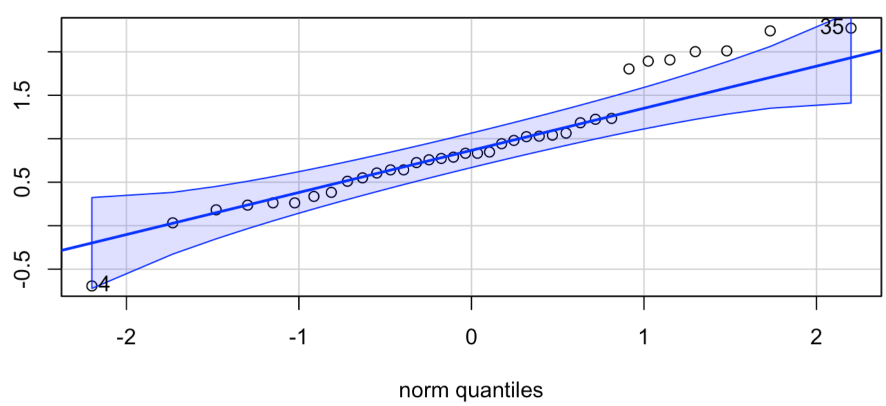

Figure 28. Q-Q Plot of Log-Transformed Response Time

Following the log-transformation of the vigilance measure, two distinct feature selection methodologies were employed within a linear regression framework to determine the best model composition: Backward Stepwise Selection and Least Absolute Shrinkage and Selection Operator (LASSO) Regression. The stepwise selection mirrored the same method used in *Chapter 5* to select the healthcare models' final feature set. The LASSO is an L1 regularization method that shrinks the model coefficients of variables with the least predictive power to zero and was implemented against the stepwise method as a possible improvement to the framework. However, by comparing the results of the two methods, the stepwise algorithm yielded a model with an AIC of 33.19, while the LASSO model's AIC was 41.15. Since a model with a lower AIC corresponds to a better fitting model, the feature set obtained from the stepwise algorithm was ultimately used to model vigilance. The significant variables in the resulting feature set are summarized in the regression table below (see *Table 14*). The relationships between these variables and vigilance are explored in the *Discussion* section.

Table 14. Linear Regression Model Results for Driving Dataset

Variable Type	Variable	Estimate	Std. Error	p-value
	(Intercept)	–1.37	0.84	0.12
Contextual Factors	Annual Miles Driven	0.20	0.08	0.02
	Weekly Driving Frequency	–0.25	0.08	< 0.01
	Virtual Meeting Experience	–0.34	0.11	< 0.01
	Distraction Susceptibility	0.81	0.24	< 0.01
	Workload Prior to the Takeover	0.18	0.05	< 0.01
	Takeover Warning Type	0.69	0.24	< 0.01
Driving Behavior	Maximum Steering Wheel Angle	0.08	0.03	< 0.01
	Maximum Brake Force	–0.01	<0.01	0.02
Eye Movement	Duration of iPad Fixations Prior to the Takeover	0.01	<0.01	0.03

7.7.3 Vigilance Modeling

To evaluate the predictive power of the significant variables identified during the stepwise regression, five unique vigilance models were developed. Similar to *Chapter 5*, these models included 1) LR, 2) Robust Linear Models (RLM), 3) SVM, 4), RF, and 5) BN. Due to the fact that vigilance was only assessed once during the takeover event, a DBN was not feasible with this dataset. Still, these models provided a comprehensive comparison between traditional linear estimation, robust techniques, nonlinear, and probabilistic methods. The LR model served as the baseline, assuming a linear relationship where the target variable is a weighted sum of the predictors. To account for potential outliers that might bias the ordinary least squares-based regression model, an RLM was employed using M-estimation in which observations with large residuals are down-weighted (Wiens, 1994). In previous analyses of this dataset for non-vigilance contexts, there were significant differences between the OLS-based LR model and the M-estimation-based RLM model. SVM was implemented using a linear function kernel, which creates a linear hyperplane boundary to fit data within a predefined error tolerance while maximizing the flatness of the regression line. A radial basis function kernel was initially tested but was outperformed by the linear SVM. To capture nonlinear relationships in the data RF and BN models were used. The RF used an ensemble method that constructs multiple decision trees and utilizes bootstrap aggregating (bagging) to produce an output with the average prediction of the forest. The BN followed the same architecture as the three-layer BNs developed in *Chapter 4* and *Chapter 5* in which the top layer contained the contextual variables, the middle layer contained the vigilance measure, and the bottom layer was dedicated to the observable variables (see *Figure 29*).

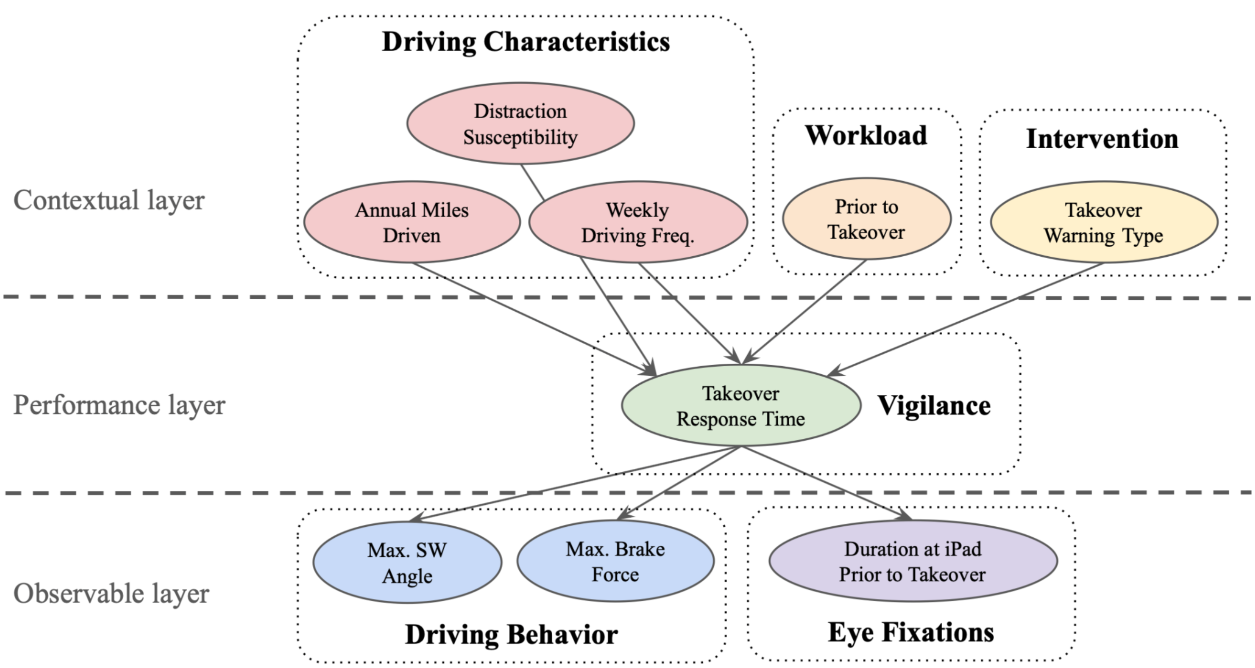

Figure 29. Bayesian Network for Vigilance in Driving

As in *Chapter 5*, model performance and generalizability were assessed using LOOCV. The predictive accuracy of each model was quantified using RMSE, providing a standardized basis for identifying the most reliable model for predicting drivers' vigilance. A Friedman test and Levene's test were used to compare the magnitude and variance of errors between all five models.

7.7.4 Vigilance Intervention Assessment

Lastly, a comparative analysis was performed to assess the impact of the takeover warning-based intervention designs. Due to the non-normal distribution of the response time data, a Wilcoxon rank sum test was used to compare the response times of participants who received the emergency takeover warning compared to those who received the planned takeover warning. Unlike a t-test, the Wilcoxon rank sum test is a non-parametric test that does not assume a normal distribution. Instead, it ranks all observations from both groups together and calculates a test statistic (W) based on the sum of the ranks for each group.

7.8 Results

The predictive performance of the five vigilance models was evaluated by comparing their cross-validated error metrics. The results are summarized in *Table 15*, including the RMSE and 95% confidence interval of the RMSE for each model.

Table 15. Driving Vigilance Models Performance (in s)

Model	Linear Regression	Robust Linear Model	Support Vector Machine	Random Forest	Bayesian Network
RMSE	1.03	1.03	1.08	1.15	1.12
95% CI	[0.67, 2.18]	[0.70, 2.21]	[0.83, 2.34]	[1.20, 2.54]	[1.04, 2.49]

The LR and RLM demonstrated the lowest errors, both achieving an identical RMSE of 1.03 s. These models were followed by the SVM (RMSE = 1.08 s), Bayesian Network (RMSE = 1.12 s), and Random Forest (RMSE = 1.15 s) models. While the LR and RLM models provided the narrowest 95% confidence intervals, all models exhibited overlapping CIs, suggesting a relatively similar error magnitude across all modeling paradigms.

To determine if the differences in errors were statistically significant, a Friedman test was conducted on the absolute errors across all five models. The test yielded an insignificant result ($\chi^2 = 1.04$, df = 4, $p = 0.903$), indicating that no single model significantly outperformed the others in terms of error magnitude. This finding is further illustrated using the boxplot in *Figure 30*, where the models' median absolute errors and their distributions overlapped significantly. Although the RF and BN models seemed to have wider interquartile ranges compared to the LR, RLM, and SVM models, a Levene's test was insignificant ($F = 0.38$, $p = 0.82$).

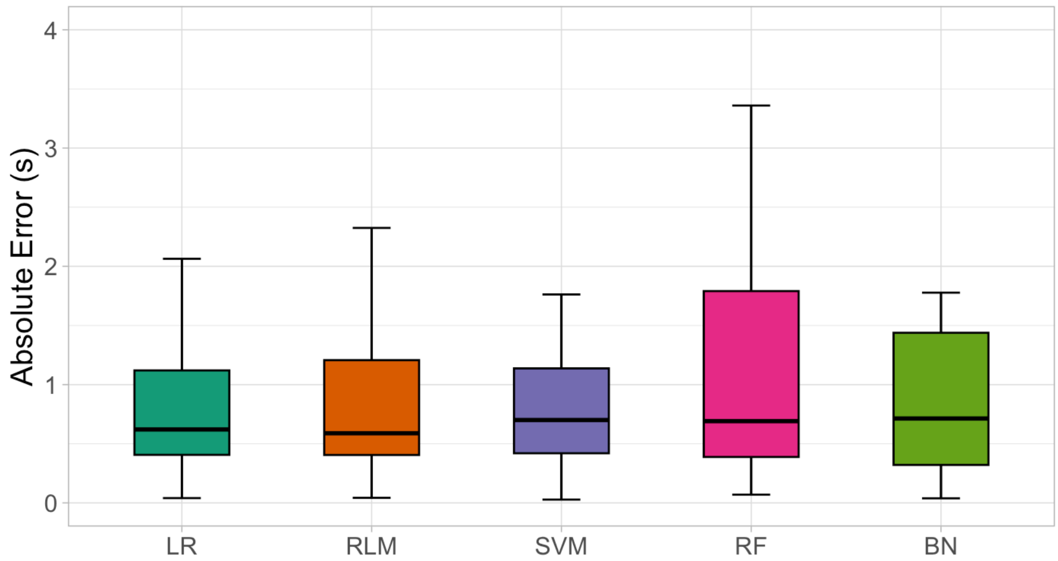

Figure 30. Box Plot of the Driving Models' Absolute Errors

When comparing the takeover warning interventions, a Wilcoxon rank sum test revealed a highly significant difference in performance between the two designs ($W = 59$, $p = 0.001$). Participants who received the Emergency warning had a significantly faster mean response time of 1.92 s, whereas those who received the Planned warning took over twice as long to respond on average, with a mean response time of 4.42 s. As shown in *Figure 31*, the Emergency warning not only resulted in faster response times but also showed much higher consistency, with a tighter interquartile range and lower overall variance.

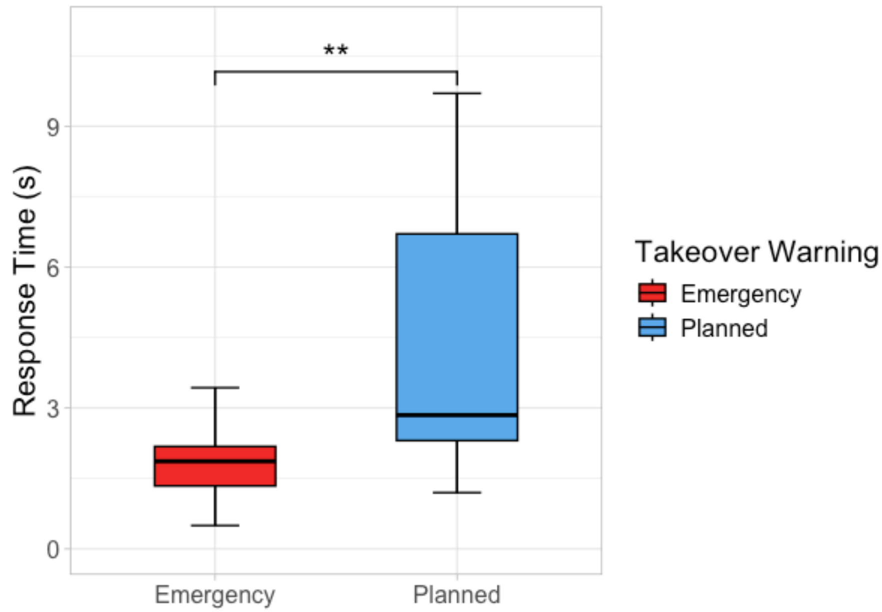

Figure 31. Box Plot of Response Time by Takeover Warning Type

7.9 Discussion

In this chapter, a driving simulator-based study was used to validate the use of my vigilance modeling framework in a different safety-critical domain. The analysis and results from this chapter are used to answer all four research questions in this dissertation, specifically for the driving domain.

7.9.1 Contextual Factors Affecting Vigilance

Among the contextual factors, annual miles driven, weekly driving frequency, virtual meeting experience, distraction susceptibility, workload prior to the takeover, and takeover warning type were statistically significant in the regression model. Higher annual miles driven was associated with slower response times (i.e., lower vigilance) during the takeover event.

Conversely, higher weekly driving frequency was associated with faster response times (i.e., higher vigilance) during the takeover event.

Those with more experience engaging in virtual meetings typically had faster response times during the takeover event. Participants with higher scores in the SDDQ (i.e., those who were more susceptible to distraction), tended to have slower response times during the takeover event. Participants with higher self-reported workload during the takeover also had slower reaction times during the takeover event. This key finding appears to support the overload theory of vigilance. The takeover warning intervention designs also had an effect on participants' vigilance. The planned takeover warning was associated with slower response times compared to the emergency takeover warning.

7.9.2 Physiological Measures for Monitoring Vigilance

For the eye fixation variables, only the duration of iPad fixations was significant for predictive modeling. Longer fixations toward the iPad prior to the takeover event were associated with slower response times during the takeover event. This finding suggests that allocating significant visual attention toward secondary tasks may hinder drivers' vigilance during safety-critical events.

Although they are not physiological measures per-se, driving behavior was used as a unique substitute particular to the driving domain. Of the driving behavior variables, maximum steering wheel angle and maximum brake force were significant in the stepwise regression model. Higher maximum steering wheel angle was associated with slower response times during the takeover event. On the other hand, higher maximum brake force was associated with faster response times during the takeover event. These findings indicate that harsher steering wheel inputs are related to lower levels of vigilance while more forceful brake inputs may be indicative of higher vigilance.

7.9.3 Predictive Modeling Features and Strategies for Vigilance

The results of this chapter showed that there were no significant differences in error magnitude nor variance across all five types of predictive models. Despite the lack of statistical significance, the linear regression and robust linear models had the lowest RMSE and also showed smaller interquartile ranges for absolute errors. This may point to the fact that there are strong linear relationships between the predictors and vigilance in this dataset, causing nonlinear and probabilistic models such as the RF and BN models to struggle in comparison to the simpler linear models. Similar to the healthcare chapter's predictive models, the driving chapter's models required the use of both contextual and physiological (and driving behavioral) measures to accurately predict vigilance. Interestingly, virtual meeting task performance was not one of the significant features affecting vigilance even though it required drivers' attention during the drive. This suggests that the combination of contextual and observable features is able to substitute for the effects of the task, negating the need to capture more behavioral measures to model vigilance.

7.9.4 Sustainable Vigilance Interventions

In this chapter, two types of takeover warning designs were compared as interventions to enhance vigilance ahead of a work zone. The findings showed that the intervention design was a critical determinant of drivers' vigilance during takeover events. The emergency warning featured more urgent visuals and auditory alerts and resulted in quicker reaction times during the takeover. In contrast, the planned warning featured “notification” style visuals and auditory alerts, resulting in longer reaction times. The differences in average reaction times between the participants who received the two types of warnings was 2.5 seconds, showcasing a significant disparity between the designs. Based on these results, in-vehicle warnings need salient and urgent features in order to be effective interventions for enhancing vigilance ahead of safety-critical events. At the same time, these alerts can be considered sustainable interventions since they are integrated into automotive user interfaces by the manufacturers, researchers, and designers for automated deployment when needed. Since the takeover warnings were only presented prior to the takeover, these findings represent a means for enhancing vigilance directly prior to safety-critical events and were therefore analyzed for the takeover only. Future work in this area could explore the long-term effects of these types of warnings by exploring post-intervention effects on vigilance over time.

It is also worth noting that the virtual meeting itself may have functioned as its own vigilance intervention by providing cognitive engagement during the portion of the drive with ADS active. This may have influenced drivers’ vigilance levels prior to the takeover, independent of the takeover warnings. However, the virtual meeting was designed to serve as a naturalistic secondary task in driving rather than a formal intervention, thus its potential effect on vigilance was not directly assessed in this study but accounted for, in part, using meeting performance and eye fixation activity to the iPad as predictors in the vigilance models.

Chapter 8. Conclusions

8.1 Vigilance Modeling Framework

The goal of this dissertation is to provide a framework for predicting vigilance decrement that can be applied to various safety-critical domains. Vigilance decrement can have a significant impact on an individual's ability to make appropriate and timely decisions within these high workload, high stress environments. Performance can be negatively impacted such that individuals may miss important signals or information leading to slow and/or inaccurate responses. The proposed vigilance modeling framework (see *Figure 32*) can provide insights into the characteristics associated with decrements in vigilance and interventions that can help mitigate the impacts.

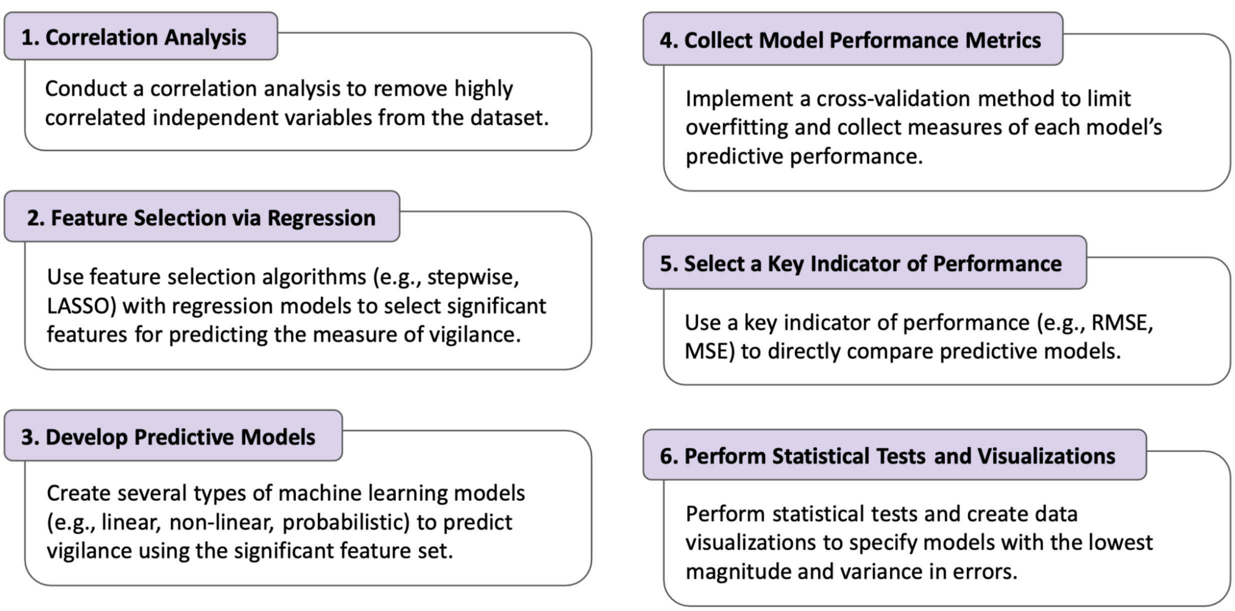

Figure 32. Proposed Vigilance Modeling Framework

The framework's first two steps focus on using correlation analysis and regression to select features. These steps are crucial for vigilance modeling because of the unique mix of contextual, physiological, and behavioral measures needed to predict vigilance as well as the inherent collinearity and high dimensionality in these datasets. Step 3 emphasizes the use of various types of modeling schemes since no single model type is universally optimal. However, these models should still be rooted by the fact that contextual and physiological factors and multiple types of sensors were required by the healthcare and driving studies to effectively measure vigilance. The final three steps focus on analyzing and comparing the vigilance models' prediction error distributions. Most analytical methods prioritize error magnitude while this vigilance modeling framework considers multiple aspects: cross-validated RMSE, error magnitude, and error variance. In real-world safety applications, the consistency in model predictions can be just as crucial as its accuracy. Considering these different model metrics provides a holistic and statistically rigorous assessment of prediction performance in vigilance models. *Chapters 4, 5,* and *7* validated and demonstrated the ability of the framework to be applied in both healthcare and driving environments. The outcomes of this framework are used to answer the four main research questions of this dissertation, as summarized below.

8.2 Review of Findings

This dissertation used a combination of literature reviews and empirical studies to effectively answer RQs 1–4:

1. What contextual factors, together, affect vigilance levels in safety-critical settings?
2. What combination of physiological measures can be used to monitor vigilance?
3. What modeling features and strategies can be used to predict vigilance decrement?

4. What types of interventions can be used to maintain vigilance in a sustainable manner? The following sections describe the specific contributions of each chapter and discuss what the collective answers mean for future research in the field of vigilance.

8.2.1 Contextual Factors of Vigilance

In the healthcare studies (*Chapters 4* and *5*), vigilance was primarily affected by sleep quality, environmental factors (e.g., room temperature), and feelings of anxiety. In the driving study (*Chapter 7*) the significant contextual factors were related to drivers' task experience, their driving habits, and outcomes during the drive: frequency of driving and miles driven, virtual meeting experience, susceptibility to distraction, cognitive workload leading up to the takeover event, and the type of takeover warning they received. These findings suggest that effective contextual factors that affect vigilance are primarily domain specific. For example, the demanding nature of work in healthcare settings requires proper sleep to maintain high levels of vigilance during shifts. In the driving domain, prior habits and experience with tasks within the vehicle can influence how well the driver manages their attention. Collectively, these chapters demonstrate that the most capable predictors of vigilance are not universal but are highly dependent on the specific demands of the safety-critical environment itself.

8.2.2 Physiological Measures of Vigilance

The findings from the review of physiological sensors in *Chapter 3* were tested in the healthcare (*Chapter 4* and *5*) and driving (*Chapter 7*) domains. The ECG was supported by the findings in *Chapter 4* and *Chapter 5*, where HR-related measures were found to significantly affect vigilance. Likewise, EDA and skin temperature were significant predictors of vigilance in the healthcare study. A significant measure that was not covered in the literature review was BVP, pointing to an additional, relatively unexplored metric that can be used in other applications to assess vigilance. The use of eye tracking measures was also supported by findings in both the healthcare and driving studies. Eye movement metrics, such as fixations, were significant predictors of vigilance across the two domains, while pupil-related measures were found to be effective in the healthcare domain. *Chapter 7* used driving behavior to substitute for the physiological measures used in *Chapters 4* and *5* and maintain a multi-sensor approach to monitoring vigilance. Steering wheel and brake pedal inputs were used with eye movement metrics to effectively assess vigilance in the driving study. While traditional, direct measures of vigilance rely on behavioral responses or central physiological indicators (e.g., via EEGs), the findings from this dissertation suggest that peripheral physiological and eye tracking can serve as similarly direct indicators of vigilance in real-world settings, particularly when leveraging contextual factors in the modeling framework.

8.2.3 Predictive Modeling Features and Strategies for Vigilance

The modeling strategies employed in *Chapters 4*, *5*, and *7* reveal that the "best" predictive model for vigilance may be dependent on the environment. In the healthcare domain study (*Chapters 4* and *5*), there were more longer-duration vigilance tasks, which led to probabilistic models like BNs and DBNs being superior performers since they could account for the nuanced, temporal decays in vigilance. Conversely, the driving study (*Chapter 7*), which featured a more sudden reaction event, found that simpler linear models were highly effective. This suggests that in safety-critical settings, longer-form vigilance tasks (e.g., monitoring vitals during a hospital shift) require models that can account for time and uncertainty while for sudden-event vigilance tasks (e.g., a semi-autonomous vehicle takeover) the relationship between predictors and performance is more direct and linear. While there was no common model that performed best across both domains, the chapters confirmed that a "hybrid" approach—combining contextual and physiological features in the modeling scheme—is essential for minimizing and stabilizing prediction errors. In particular, eye-related measures, such as fixation duration, were significant predictors of vigilance across both healthcare and driving domains. In addition, the three-layer structure of the BNs and DBNs were effective for modeling vigilance across both domains, suggesting that using contextual, performance, and observable layers may serve as a generalizable vigilance modeling structure in other safety-critical environments as well.

8.2.4 Intervention Designs for Vigilance Decrement

The literature review in *Chapter 6* identified a gap in vigilance intervention research: while stimulants or neurofeedback can effectively boost vigilance, they are difficult to sustain in a naturalistic setting. *Chapter 7* addressed this gap by testing an integrated task context change-related intervention: a semi-autonomous vehicle takeover warning. The main findings from this driving study showed that the design aspects of the warning intervention are critical for enhancing vigilance. Effective interventions should be salient and urgent to bypass the effects of vigilance decrement. At the same time, by integrating these alerts directly into the user interface (e.g., the vehicle dashboard, user smartphones), the intervention becomes "sustainable" as it requires no extra effort from the user or external assistance to receive its benefits.

8.3 Impact and Implications

The findings of this dissertation provide a comprehensive foundation for the assessment and management of vigilance in safety-critical environments. This work advocates for a shift toward domain-specific validity in vigilance research. While standard tasks like the PVT and CPT are useful for baseline measurements of vigilance, this research illustrates that naturalistic tasks (such as ECG reading or virtual meetings in semi-autonomous vehicles) yield more representative results that reflect the unique demands of the respective domain. By bridging the gap between theoretical findings and naturalistic applications, this work offers several critical contributions for human factors researchers and practitioners.

A significant implication of this research is the necessity of a hybrid approach to vigilance modeling. The results consistently demonstrate that neither physiological sensors nor contextual factors are sufficient alone. Instead, they should be used in tandem to produce reliable predictions. For practitioners, this means that vigilance monitoring systems should not only track real-time biomarkers like heart rate or eye gaze activity but must also account for effects such as sleep history, task experience, and environmental factors.

A theoretical contribution from *Chapter 7* was the significant relationship between mental workload and vigilance. In the driving context specifically, the overload theory of vigilance was supported, as higher workload was associated with vigilance decrement. However, mental workload was not a significant predictor of vigilance in the healthcare study. This suggests that measuring overload and underload behind vigilance decrement may require different measurement strategies. Overload-related vigilance decrement may be better captured through discrete performance measures at high-demand moments, such as the takeover event in *Chapter 7*, while underload-related vigilance decrement may be better reflected in continuous vigilance tasks over time. Furthermore, workload may function as a mediating factor between contextual conditions and vigilance outcomes, where the degree of cognitive load imposed by the environment (e.g., the virtual meeting) determines which theoretical mechanism is most active.

This work builds on traditional reliance on mean error metrics (e.g., RMSE) as the sole validator of vigilance model performance. A key takeaway from the longitudinal analysis is that error variance is a critical measure in combination with error magnitude. As seen with the DBN in *Chapter 5*, the incorporation of a temporal element may not always significantly reduce absolute error, but it can significantly stabilize the model's predictions. In safety-critical settings, a reliable vigilance model that provides consistent, low-variance results can be more valuable than one with slightly lower errors but higher volatility.

Finally, the findings emphasize that for an intervention to be effective, it must be sustainable and salient. Particularly for safety-critical environments, moving away from expensive brain stimulants toward intuitive, system-integrated alerts (e.g., urgent takeover warnings) provides a scalable path for maintaining vigilance. Ultimately, this dissertation suggests that the future of vigilance modeling and safety engineering lies in multimodal and flexible systems that proactively assess the user's cognitive state through context-aware design.

8.4 Limitations and Future Research

While the promising results of this dissertation can be used to significantly improve the state of vigilance research, there are several limitations to consider. *Chapter 4* and *Chapter 5* leverage a wide range of sensor-based methods and contextual factors but are skewed toward a specific demographic in hospitals: resident physicians. Although additional participants from nursing and therapy were included, a larger sample covering even distributions of hospital workers would help make the results more generalizable to the entire healthcare domain. *Chapter 7* used an innovative approach to future problems in semi-autonomous driving but was not designed specifically for vigilance. This dissertation used the pre-existing dataset and molded it to suitably fit the current context and answer research questions specific to vigilance. Re-designing the experiment to include more features similar to the ones in *Chapters 4–5* (e.g., assessments of sleep quality, environmental conditions, wearable sensors) would have made the *Chapter 7*'s results more comparable. At the same time, temporal features would have been considered and allowed for predictive modeling using DBNs or other similar analytical strategies. In general, these chapters used suitable but low sample sizes for the machine learning-based models of vigilance. Increasing the study sample sizes from n = 30–40 to n ≥ 100 would undoubtedly strengthen the robustness of the modeling efforts and lead to more remarkable differences in the results. Finally, the literature reviews covered a substantial spread of studies in vigilance but were limited to the past 10 years to highlight the state-of-the-art in sensing technologies and intervention design. Expanding these reviews to include older papers may uncover key research findings that were not covered in this dissertation's analysis.

While this dissertation establishes a robust framework for modeling vigilance, several avenues remain for extending this work to enhance safety. The success of the DBN in reducing error variance in *Chapter 5* suggests that temporal relationships are key to reliable monitoring. Future research should consider exploring more complex temporal architectures, such as Neural Networks or Long Short-Term Memory networks, which may be better suited to capturing the time-dependent nature of vigilance decrement. While *Chapter 7* demonstrated that urgent, multimodal alerts are highly effective for immediate vigilance recovery, it remains unclear how frequent exposure to this style of warning affects driver trust and cognitive load over extended periods of time. Future studies should investigate whether the efficacy of these salient alerts diminishes through alarm fatigue and explore adaptive intervention strategies that vary in intensity based on the severity of the predicted vigilance decrement. At the same time, future work could explore the longitudinal effects of takeover warning designs on vigilance. Additionally, since all participants in the driving study engaged in the virtual meeting, there is no baseline condition to assess the effects of the virtual meeting itself on vigilance during the takeover. Future studies that incorporate naturalistic secondary tasks should use conditions without the task to better evaluate its contributions to vigilance during semi-autonomous driving. Finally, while this research successfully compared the healthcare and driving domains, the framework should continue to be tested in other safety-critical environments, such as air traffic control or power plant operation. Applying the vigilance modeling framework in more diverse scenarios can continue to improve the framework's ability to accurately assess and predict vigilance in these types of domains.

References

Abbasi, N. I., Bodala, I. P., Bezerianos, A., Sun, Y., Al-Nashash, H., & Thakor, N. V. (2017). Role of multisensory stimuli in vigilance enhancement- a single trial event related potential study. *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2446–2449.

Abe, T., Mishima, K., Kitamura, S., Hida, A., Inoue, Y., Mizuno, K., Kaida, K., Nakazaki, K., Motomura, Y., Maruo, K., Ohta, T., Furukawa, S., Dinges, D. F., & Ogata, K. (2020). Tracking intermediate performance of vigilant attention using multiple eye metrics. *Sleep: Journal of Sleep and Sleep Disorders Research*, *43*(3), 1–17.

Aeschbach, D., Cohen, D. A., Lockyer, B. J., Chellappa, S. L., & Klerman, E. B. (2024). Spontaneous attentional failures reflect multiplicative interactions of chronic sleep loss with acute sleep loss and circadian misalignment. *Festschrift of Dr. Charles A. Czeisler*, *10*(1, Supplement), S89–S95.

Al-Shargie, F., Tariq, U., Babiloni, F., & Al-Nashash, H. (2021). Cognitive Vigilance Enhancement Using Audio Stimulation of Pure Tone at 250 Hz. *IEEE Access*, *9*, 22955–22970.

Al-Shargie, F., Tariq, U., Mir, H., Alawar, H., Babiloni, F., & Al-Nashash, H. (2019). Vigilance decrement and enhancement techniques: a review. *Brain sciences, 9*(8), 178.

Anderson, B. A., Kim, N., Gregoire, L., Yan, N., & Ahn, C. R. (2023). Attention Failures Cause Workplace Accidents: Why Workers Ignore Hazards and What To Do About It. *Policy Insights from the Behavioral and Brain Sciences, 11*(1), 27-35.

Andreassi, J. L. (2010). *Psychophysiology: Human behavior and physiological response*. Psychology Press.

Annerer-Walcher, S., Körner, C., Beaty, R.E. *et al.* (2020). Eye behavior predicts susceptibility to visual distraction during internally directed cognition. *Atten Percept Psychophys* 82, 3432–3444.

Ansarinejad, M., Ansarinejad, K., Lu, P., & Huang, Y. (2025). User experience of navigating work zones with automated vehicles: Insights from YouTube on challenges and strengths. *Smart Cities, 8*(4).

Arthurs, O. J., & Boniface, S. (2002). How well do we understand the neural origins of the fMRI BOLD signal?. *TRENDS in Neurosciences*, *25*(1), 27-31.

Axelsen, J. L., Meline, J. S. J., Staiano, W., & Kirk, U. (2022). Mindfulness and music interventions in the workplace: Assessment of sustained attention and working memory using a crowdsourcing approach. *BMC Psychology*, *10*(1), 108.

Bajestani, G. S., Ghanizadeh, A., Makhloughi, F., Hosseinpour Kharrazi, F., Hosseini, A., & Toosi, M. B. (2024). The Impact of Blended Mindfulness Intervention (BMI) on University Students' Sustained Attention, Working Memory, Academic Achievement, and Electroencephalogram (EEG) Asymmetry. *Mindfulness*, *15*(3), 675–688.

Barba, A., Padilla, F., Luque-Casado, A., Sanabria, D., & Correa, Á. (2018). The Role of Exercise-Induced Arousal and Exposure to Blue-Enriched Lighting on Vigilance. *Frontiers in Human Neuroscience*. Natural Science Collection; Publicly Available Content Database.

Batchinsky, A. I., Cooke, W. H., Kuusela, T. A., Jordan, B. S., Wang, J. J., & Cancio, L. C. (2007). Sympathetic nerve activity and heart rate variability during severe hemorrhagic shock in sheep. *Autonomic Neuroscience, 136*(1-2), 43-51.

Bearden, T.S., Cassisi, J.E., & White, J.N. (2004). Electrophysiological Correlates of Vigilance During a Continuous Performance Test in Healthy Adults. *Appl Psychophysiol Biofeedback, 29*, 175–188.

Berding, K., Long-Smith, C. M., Carbia, C., Bastiaanssen, T. F. S., Van De Wouw, M., Wiley, N., Strain, C. R., Fouhy, F., Stanton, C., Cryan, J. F., & Dinan, T. G. (2021). A specific dietary fibre supplementation improves cognitive performance—An exploratory randomised, placebo-controlled, crossover study. *Psychopharmacology*, *238*(1), 149–163.

Bergum, B. O., & Lehr, D. J. (1962). Vigilance performance as a function of interpolated rest. *Journal of Applied Psychology, 46*, 425–427.

Besten, M. E., van Tol, M.-J., van Rij, J., & van Vugt, M. K. (2023). The impact of mood-induction on maladaptive thinking in the vulnerability for depression. *Journal of Behavior Therapy and Experimental Psychiatry*, *81*, 101888.

Blackwood, D. H. R., & Muir, W. J. (1990). Cognitive brain potentials and their application. *The British Journal of Psychiatry*, *157*(S9), 96-101.

Blatter, K., Graw, P., Münch, M., Knoblauch, V., Wirz-Justice, A., & Cajochen, C. (2006). Gender and age differences in psychomotor vigilance performance under differential sleep pressure conditions. *Behavioural brain research*, *168*(2), 312-317.

Bodala, I. P., Li, J., Thakor, N. V., & Al-Nashash, H. (2016). EEG and eye tracking demonstrate vigilance enhancement with challenge integration. *Frontiers in human neuroscience*, *10*, 273.

Bogler, C., Mehnert, J., Steinbrink, J., & Haynes, J. D. (2014) Decoding Vigilance with NIRS. *PLOS ONE 9*(7): e101729.

Bour, L. J., Aramideh, M., & Ongerboer De Visser, B. W. (2000). Neurophysiological Aspects of Eye and Eyelid Movements During Blinking in Humans. *Journal of Neurophysiology*, *83*(1), 166–176.

Brownson, R. C., Colditz, G. A., & Proctor, E. K. (Eds.). (2018). Dissemination and implementation research in health: translating science to practice. *Oxford University Press.*

Braun, N., Debener, S., Sölle, A., Kranczioch, C., & Hildebrandt, H. (2015). Biofeedback-based self-alert training reduces alpha activity and stabilizes accuracy in the Sustained Attention to Response Task. *Journal of Clinical and Experimental Neuropsychology*, *37*(1), 16–26.

Bruce, S. E., Werner, K. B., Preston, B. F., & Baker, L. M. (2014). Improvements in concentration, working memory and sustained attention following consumption of a natural citicoline-caffeine beverage. *International Journal of Food Sciences and Nutrition*, *65*(8), 1003–1007.

Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193-213.

Canisius, S., & Penzel, T. (2007) Vigilance monitoring--review and practical aspects. *Biomed Tech (Berl)*, *52*(1), 77-82.

Caravati, E., Barbeni, F., Chiarion, G., Raggi, M., & Mesin, L. (2024). Closed-Loop Transcranial Electrical Neurostimulation for Sustained Attention Enhancement: A Pilot Study towards Personalized Intervention Strategies. *Bioengineering (Basel, Switzerland)*, *11*(5), 467.

Chai, S. C., Jerusik, J., Davis, K., Wright, R. S., & Zhang, Z. (2019). Effect of Montmorency tart cherry juice on cognitive performance in older adults: A randomized controlled trial. *Food & Function*, *10*(7), 4423–4431

Charles, R. L., & Nixon, J. (2019). Measuring mental workload using physiological measures: A systematic review. *Applied ergonomics, 74*, 221-232.

Chaudhari, A., Wang, X., Wu, A., & Liu, H. (2023). Repeated Transcranial Photobiomodulation with Light-Emitting Diodes Improves Psychomotor Vigilance and EEG Networks of the Human Brain. *Bioengineering (Basel, Switzerland)*, *10*(9), 1043.

Chen, R., Wang, S., Fan, Y., Liu, X., Wang, J., Lv, Y., Wang, D., Wu, D., Cao, W., & Zou, Q. (2023). Acute Tai Chi Chuan exercise enhances sustained attention and elicits increased cuneus/precuneus activation in young adults. *Cerebral Cortex (New York, N.Y.: 1991)*, *33*(6), 2969–2981.

Chen, W.-H., Chen, J., Lin, X., Li, P., Shi, L., Liu, J.-J., Sun, H.-Q., Lu, L., & Shi, J. (2018). Dissociable effects of sleep deprivation on functional connectivity in the dorsal and ventral default mode networks. *Sleep Medicine*, *50*, 137–144.

Chen, Y., Farrand, J., Tang, J., Chen, Y., O'Keeffe, J., & Shou, G. (2017). Relationship between amplitude of resting-state fNIRS global signal and EEG vigilance measures. In the *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 537-540.

Chernorizov, A. M., Isaychev, S. A., Zinchenko, Y. P., Galatenko, V. V., Znamenskaya, I. A., Zakharov, P. N., ... & Gradoboeva, O. N. (2016). Psychophysiological methods for the diagnostics of human functional states: New approaches and perspectives. *Psychology in Russia, 9*(4), 56.

Chin, B., Lindsay, E. K., Greco, C. M., Brown, K. W., Smyth, J. M., Wright, A. G. C., & Creswell, J. D. (2021). Mindfulness interventions improve momentary and trait measures of attentional control: Evidence from a randomized controlled trial. *Journal of Experimental Psychology: General*, *150*(4), 686–699.

Choi, S.D., Guo, L., Kang, D., & Xiong, S. (2017). Exergame technology and interactive interventions for elderly fall prevention: a systematic literature review. *Applied Ergonomics 65*, 570–581.

Chua, E. C. P., Tan, W. Q., Yeo, S. C., Lau, P., Lee, I., Mien, I. H., ... & Gooley, J. J. (2012). Heart rate variability can be used to estimate sleepiness-related decrements in psychomotor vigilance during total sleep deprivation. *Sleep*, *35*(3), 325-334.

Cintineo, H. P., Bello, M. L., Chandler, A. J., Cardaci, T. D., McFadden, B. A., & Arent, S. M. (2022). Effects of caffeine, methylliberine, and theacrine on vigilance, marksmanship, and hemodynamic responses in tactical personnel: A double-blind, randomized, placebo-controlled trial. *Journal of the International Society of Sports Nutrition*, *19*(1), 543–564.

Claypoole, V. L., Neigel, A. R., Waldfogle, G. E., & Szalma, J. L. (2019). Evaluative social presence can improve vigilance performance, but vigilance is still hard work and is stressful. *Journal of Experimental Psychology. Human Perception and Performance*, *45*(5), 616–627.

Coimbra, K. T. F., de Freitas, F. G. R., Bafi, A. T., Pinheiro, T. T., Nunes, N. F., de Azevedo, L. C. P., & Machado, F. R. (2019). Effect of increasing blood pressure with noradrenaline on the microcirculation of patients with septic shock and previous arterial hypertension. *Critical Care Medicine*, *47*(8), 1033-1040.

Craig, C. M., & Klein, M. I. (2019). The abbreviated vigilance task and its attentional contributors. *Human factors*, *61*(3), 426-439.

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. Academic Press.

Demerouti, E., Le Blanc, P. M., Bakker, A. B., Schaufeli, W. B., & Hox, J. (2009). *Present but sick: a three-wave study on job demands, presenteeism and burnout*. Career Development International.

D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S, & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature, 378*, 279–281.

Di Flumeri, G., De Crescenzio, F., Berberian, B., Ohneiser, O., Kramer, J., Aricò, P., Borghini, G., Babiloni, F., Bagassi, S., & Piastra, S. (2019). Brain–Computer Interface-Based Adaptive Automation to Prevent Out-Of-The-Loop Phenomenon in Air Traffic Controllers Dealing With Highly Automated Systems. *Frontiers in Human Neuroscience*.

Dirican, A. C., & Göktürk, M. (2011). Psychophysiological measures of human cognitive states applied in human computer interaction. *Procedia Computer Science*, *3*, 1361–1367.

Distefano, N., Leonardi, S., Pulvirenti, G., Romano, R., Merat, N., Boer, E., & Woolridge, E. (2020). Physiological and driving behaviour changes associated to different road intersections. *European Transport*, *77*.

Dorrian, J., Lamond, N., Kozuchowski, K., & Dawson, D. (2008). The driver vigilance telemetric control system (DVTCS): Investigating sensitivity to experimentally induced sleep loss and fatigue. *Behavior research methods, 40*(4), 1016-1025.

Drummond, S., Bischoff-Grethe, A., Dinges, D. F., Ayalon, L., Mednick, S. C., & Meloy, M. J. (2005). The Neural Basis of the Psychomotor Vigilance Task. *Sleep, 28*(9), 1059–1068.

Du, N., Zhou, F., Pulver, E. M., Tilbury, D. M., Robert, L. P., Pradhan, A. K., & Yang, X. J. (2020). Examining the effects of emotional valence and arousal on takeover performance in conditionally automated driving. *Transportation Research Part C: Emerging Technologies, 112*, 78–87.

Edward, K. (2023). Mercedes-Benz first to gain U.S. approval for level 3 automated driving system. *Forbes, Lifestyle*(Cars & Bikes). https://www.forbes.com/sites/kyleedward/2023/09/28/mercedes-benz-first-to-gain-us-approval-for-level-3-automated-driving-system/.

Eriksson, A., & Stanton, N. A. (2017). Takeover time in highly automated vehicles: Noncritical transitions to and from manual control. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 59*(4), 689–705.

Esler, M. D., Hasking, G. J., Willett, I. R., Leonard, P. W., & Jennings, G. L. (1985). Noradrenaline release and sympathetic nervous system activity. *Journal of hypertension*, *3*(2), 117-129.

Esterman, M., Grosso, M., Liu, G., Mitko, A., Morris, R., & DeGutis, J. (2016). Anticipation of Monetary Reward Can Attenuate the Vigilance Decrement. *PLOS ONE*, *11*(7), e0159741.

Esterman, M., Thai, M., Okabe, H., DeGutis, J., Saad, E., Laganiere, S. E., & Halko, M. A. (2017). Network-targeted cerebellar transcranial magnetic stimulation improves attentional control. *NeuroImage*, *156*, 190–198.

Finkbeiner, K. M., Russell, P. N., & Helton, W. S. (2016). Rest improves performance, nature improves happiness: Assessment of break periods on the abbreviated vigilance task. *Consciousness and Cognition*, *42*, 277–285.

Flood, G., Näswall, K., & Helton, W. S. (2015). The effects of emotional stimuli on visuo-spatial vigilance. *Psychological Research*, *79*(5), 861–871.

Ford, S. D., & Nagamatsu, L. S. (2024). Four weeks of meditation training improves sustained attention in community-dwelling older adults: A proof-of-concept randomized controlled trial. *Frontiers in Aging*, *5*, 1322705.

Fox, K. C., Spreng, R. N., Ellamil, M., Andrews-Hanna, J. R., & Christoff, K. (2015). The wandering brain: Meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes. *Neuroimage,* 111, 611-621.

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*, 1325–1330.

Funke, M. E., Warm, J. S., Matthews, G., Funke, G. J., Chiu, P. Y., Shaw, T. H., & Greenlee, E. T. (2017). The neuroergonomics of vigilance: Effects of spatial uncertainty on cerebral blood flow velocity and oculomotor fatigue. *Human factors*, *59*(1), 62-75.

Gabay, S., Pertzov, Y., & Henik, A. (2011). Orienting of attention, pupil size, and the norepinephrine system. *Attention, Perception, & Psychophysics*, *73*(1), 123-129.

Garg, A., Pavlovic, V., & Rehg, J. M. (2000). Audio-visual speaker detection using dynamic Bayesian networks. In *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 384–390).

Gartenberg, D., Gunzelmann, G., Hassanzadeh-Behbaha, S., & Trafton, J. G. (2018). Examining the Role of Task Requirements in the Magnitude of the Vigilance Decrement. *Frontiers in Psychology*, *9*.

Gerber, M., Cody, R., Beck, J., Brand, S., Donath, L., Eckert, A., ... & Ludyga, S. (2023). Differences in selective attention and inhibitory control in patients with major depressive disorder and healthy controls who do not engage in sufficient physical activity. *Journal of Clinical Medicine, 12*(10), 3370.

Greenlee, E. T., DeLucia, P. R., & Lui, T. G. (2022). Modality Changes in Vigilance Displays: Further Evidence of Supramodal Resource Depletion in Vigilance. *Human Factors, 66*(2), 424-436.

Greenlee, E. T., Lui, T. G., & Maw, E. L. (2020). Is Physiobehavioral Monitoring Nonintrusive? An Examination of Transcranial Doppler Sonography in a Vigilance Task. *Human Factors, 63*(7), 1256-1270.

Gui, D., Xu, S., Zhu, S., Fang, Z., Spaeth, A. M., Xin, Y., Feng, T., & Rao, H. (2015). Resting spontaneous activity in the default mode network predicts performance decline during prolonged attention workload. *NeuroImage*, *120*, 323–330.

Guo, M., Wei, Y., Zhang, J., Huang, Q., Jin, X., & Ma, J. (2024). Optimizing Driver Vigilance Recognition: Examining the Characterization and Cumulative Effect of Physiological Signals Across Manual and Automated Driving and Durations. *Applied Sciences*, *14*(22), 10482.

Guterman, Y., Josiassen, R. C., & Bashore Jr, T. R. (1992). Attentional influence on the P50 component of the auditory event-related brain potential. *International Journal of Psychophysiology*, *12*(2), 197-209.

Hailemariam, M., Bustos, T., Montgomery, B., Barajas, R., Evans, L. B., & Drahota, A. (2019). Evidence-based intervention sustainability strategies: a systematic review. *Implementation Science, 14*, 1-12.

Hancock, P. A. (2013). In search of vigilance: the problem of iatrogenically created psychological phenomena. *American Psychologist*, *68*(2), 97.

Hancock, P. A., & Hart, S. G. (2002). Defeating Terrorism: What Can Human Factors/Ergonomics Offer? *Ergonomics in Design, 10*(1), 6-16.

Hancock, P. A., Volante, W. G., & Szalma, J. L. (2016). Defeating the vigilance decrement. *IISE Transactions on Occupational Ergonomics and Human Factors, 4*(2-3), 151-163.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). North-Holland.

Harty, S., & Cohen Kadosh, R. (2019). Suboptimal Engagement of High-Level Cortical Regions Predicts Random-Noise-Related Gains in Sustained Attention. *Psychological Science*, *30*(9), 1318–1332.

Havsteen, I., Ohlhues, A., Madsen, K. H., Nybing, J. D., Christensen, H., & Christensen, A. (2017). Are movement artifacts in magnetic resonance imaging a real problem?—a narrative review. *Frontiers in neurology*, *8*, 232.

He, J., Li, Z., Ma, Y., Long, S., & Ma, K.-H. (2023). Physiological and Behavioral Changes of Passive Fatigue on Drivers during On-Road Driving. *Applied Sciences*, *13*(2), 1200.

Helton, W. S., & Russell, P. N. (2015). Rest is best: The role of rest and task interruptions on vigilance. *Cognition, 134*, 165-173.

Hinds, O., Thompson, T. W., Ghosh, S., Yoo, J. J., Whitfield-Gabrieli, S., Triantafyllou, C. & Gabrieli, J. D. E (2013). Roles of default-mode network and supplementary motor area in human vigilance performance: evidence from real-time fMRI. *Journal of Neurophysiology, 109*(5), 1250-1258.

Hu, X., & Lodewijks, G. (2021). Exploration of the effects of task-related fatigue on eye-motion features and its value in improving driver fatigue-related technology. *Transportation Research Part F: Traffic Psychology and Behaviour, 80*, 150–171.

Huang, Q., He, S., Wang, Q., Gu, Z., Peng, N., Li, K., Zhang, Y., Shao, M. & Li., Y. (2018). An EOG-Based Human–Machine Interface for Wheelchair Control. *IEEE Transactions on Biomedical Engineering*, *65*(9), 2023–2032.

Hwang, J., Castelli, D. M., & Gonzalez-Lima, F. (2016). Cognitive enhancement by transcranial laser stimulation and acute aerobic exercise. *Lasers in Medical Science*, *31*(6), 1151–1160.

Islam, M.S., & Lim, S. (2022). Vibrotactile feedback in virtual motor learning: a systematic review. Applied Ergonomics, *101*, 103694.

Jeong, H., Kang, Z., & Liu, Y. (2019). Driver glance behaviors and scanning patterns: Applying static and dynamic glance measures to the analysis of curve driving with secondary tasks. *Human Factors and Ergonomics in Manufacturing & Service Industries, 29*(6), 437-446.

Johannes, N. (2020). Effects of smartphone cues and online vigilance on well-being and performance (Doctoral dissertation, Sl: sn).

Johnson, M., & Kim, J. E. (2020). The Effect of Task Complexity on Eye Movement and Multitasking Performance in Students With and Without ADHD. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 64, No. 1, pp. 786-790).

Jones, H. M., Yoo, K., Chun, M. M., & Rosenberg, M. D. (2024). Edge-based general linear models capture moment-to-moment fluctuations in attention. *The Journal of Neuroscience*, *44*(14).

Jones, K. T., Smith, C. C., Gazzaley, A., & Zanto, T. P. (2022). Research outside the laboratory: Longitudinal at-home neurostimulation. *Behavioural Brain Research*, *428*, 113894.

Jónsdóttir, A. A., Kang, Z., Sun, T., Mandal, S., & Kim, J. E. (2023). The effects of language barriers and time constraints on online learning performance: An eye-tracking study. *Human Factors, 65*(5), 779-791.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*, 255-260.

Joshi, S., & Gold, J. I. (2020). Pupil Size as a Window on Neural Substrates of Cognition. *Trends in Cognitive Sciences*, *24*(6), 466–480.

Kaarre, O., Äikiä, M., Kallioniemi, E., Könönen, M., Kekkonen, V., Heikkinen, N., Kivimäki, P., Tolmunen, T., Määttä, S., & Laukkanen, E. (2018). Association of the N100 TMS-evoked potential with attentional processes: A motor cortex TMS–EEG study. *Brain and Cognition*, *122*, 9–16.

Kamzanova, A. T., Kustubayeva, A. M., & Matthews, G. (2014). Use of EEG workload indices for diagnostic monitoring of vigilance decrement. *Human Factors*, *56*(6), 1136-1149.

Karampela, O., Madison, G., & Holm, L. (2020). Motor timing training improves sustained attention performance but not fluid intelligence: Near but not far transfer. *Experimental Brain Research*, *238*(4), 1051–1060.

Karl, K. A., Peluchette, J. V., & Aghakhani, N. (2022). Virtual work meetings during the Covid-19 pandemic: The good, bad, and ugly. *Small Group Research, 53*(3), 343–365.

Kazinka, R., Roediger, D., Xuan, L., Yu, L., Mueller, B. A., Camchong, J., Opitz, A., MacDonald, A., & Lim, K. O. (2024). tDCS-enhanced cognitive training improves attention and alters connectivity in control and somatomotor networks: A triple blind study. *NeuroImage*, *298*, 120792.

Kim, J. H., Kim, D. W., & Im, C. H. (2017). Brain Areas Responsible for Vigilance: An EEG Source Imaging Study. *Brain Topography, 30*, 343–351.

Kim, J. E., & Lee, T. H. (2020). Neuroergonomics and Its Relation to Psychophysiology. In *Neuroergonomics* (pp. 183-192). Springer, Cham.

Kivikangas, J. M., Chanel, G., Cowley, B., Ekman, I., Salminen, M., Järvelä, S., & Ravaja, N. (2011). A review of the use of psychophysiological methods in game research. *Journal of gaming & virtual worlds, 3*(3), 181-199.

Koelega, H.S., Verbaten, M.N., van Leeuwen, T.H., Kenemans, J.L., Kemner, C., & Sjouw, W. (1992). Time effects on event-related brain potentials and vigilance performance. *Biological Psychology 34*, 59–86.

Kugler, L. (2022). The impact of virtual meetings. *Communications of the ACM, 65*(11), 19–21.

Kumar, N., Wheaton, L. A., Snow, T. K., & Millard-Stafford, M. (2016). Carbohydrate ingestion but not mouth rinse maintains sustained attention when fasted. *Physiology & Behavior*, *153*, 33–39.

LaPelle, N. R., Zapka, J., & Ockene, J. K. (2006). Sustainability of public health programs: the example of tobacco treatment services in Massachusetts. *American journal of public health*, 96(8), 1363-1369.

Lara, T., Madrid, J. A., & Correa, Á. (2014). The Vigilance Decrement in Executive Function Is Attenuated When Individual Chronotypes Perform at Their Optimal Time of Day. *PLoS One*, *9*(2). Natural Science Collection; Publicly Available Content Database.

Lavine, R. A., Sibert, J. L., Gokturk, M., & Dickens, B. (2002). Eye-tracking measures and human performance in a vigilance task. *Aviat Space Environ Med, 73*(4), 367-72.

Lawrence, N. S., Ross, T. J., Hoffman, R., Garavan, H., & Stein, E. A. (2003). Multiple neuronal networks mediate sustained attention. *Journal of Cognitive Neuroscience, 15*, 1028–1038.

Lennox, L., Maher, L., & Reed, J. (2018). Navigating the sustainability landscape: a systematic review of sustainability approaches in healthcare. *Implementation Science*, 13, 1-17.

Li, F., Chen, C.-H., Lee, C.-H., & Feng, S. (2022). Artificial intelligence-enabled non-intrusive vigilance assessment approach to reducing traffic controller's human errors. *Knowledge-Based Systems*, *239*, 108047.

Li, F., Lee, C., Chen, C., & Khoo, L. (2019). Hybrid data-driven vigilance model in traffic control center using eye-tracking data and context data. *Advanced Engineering Informatics*, *42*.

Li, J., Kettel, V., Higgins, E. A., Xiao, C., Kwon, Y., Meno, M., Roach, V. A., & Kim, J.-E. (2024). Modeling Sensor-based Vigilance Decrement in the Healthcare Environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *68*(1), 1655-1659.

Liu, C., Chen, B., Guo, Y., Chen, X., Cruz Nacpil, E. J., Hou, W., & Zheng, R. (2024). Evaluation of a user-accessible countermeasure: Effect of manual dexterity gymnastics on passive driver fatigue. *Transportation Research Part F: Traffic Psychology and Behaviour*, *101*, 387–402.

Liu, T. T., & Falahpour, M. (2020). Vigilance Effects in Resting-State fMRI. *Frontiers in Neuroscience*, *14*.

Liu, Y., Subramaniam, S. C. H., Sourina, O., Shah, E., Chua, J., & Ivanov, K. (2017). Neurofeedback Training for Rifle Shooters to Improve Cognitive Ability. *2017 International Conference on Cyberworlds (CW)*, 186–189.

Little, B., Anwyll, M., Norsworthy, L., Corbett, L., Schultz-Froggatt, M., & Gallagher, P. (2024). Processing speed and sustained attention in bipolar disorder and major depressive disorder: A systematic review and meta-analysis. *Bipolar disorders, 26*(2), 109-128.

Löffler, B. S., Stecher, H. I., Fudickar, S., Sordi, D. de, Otto-Sobotka, F., Hein, A., & Herrmann, C. S. (2018). Counteracting the Slowdown of Reaction Times in a Vigilance Experiment With 40-Hz Transcranial Alternating Current Stimulation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(10), 2053–2061.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in cognitive sciences*, *4*(11), 432-440.

Luo, Y., Chen, H., & Zhang, J. (2021). Enhanced Connectivity and Reduced Mind Wandering after Tactile Training in Young Adults. *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 285–289.

Maccora, J., Manousakis, J. E., & Anderson, C. (2019). Pupillary instability as an accurate, objective marker of alertness failure and performance impairment. *Journal of Sleep Research*, *28*(2), 1–11.

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly journal of experimental psychology, 1*(1), 6-21.

MacQueen, D. A., Minassian, A., Henry, B. L., Geyer, M. A., Young, J. W., & Perry, W. (2018). Amphetamine Modestly Improves Conners' Continuous Performance Test Performance in Healthy Adults. *Journal of the International Neuropsychological Society*, *24*(3), 283–293.

Maffei, A., & Angrilli, A. (2018). Spontaneous eye blink rate: An index of dopaminergic component of sustained attention and fatigue. *International Journal of Psychophysiology*, *123*, 58–63.

Maness, E. B., Burk, J. A., McKenna, J. T., Schiffino, F. L., Strecker, R. E., & McCoy, J. G. (2022). Role of the locus coeruleus and basal forebrain in arousal and attention. *Brain research bulletin*.

Martin, J. T., Whittaker, A. H., & Johnston, S. J. (2022). Pupillometry and the vigilance decrement: Task-evoked but not baseline pupil measures reflect declining performance in visual vigilance tasks. *European Journal of Neuroscience*, *55*(3), 778–799.

Mather, M., & Harley, C. W. (2016). The locus coeruleus: essential for maintaining cognitive function and the aging brain. *Trends in cognitive sciences*, *20*(3), 214-226.

McClain, C., Vogels, E. A., Andrew Perrin, S. S., & Rainie, L. (2021). The internet and the pandemic (tech. rep.). *Pew Research Center*.

McGarry, S. P. D., Neilson, B. N., Brown, N. L., Strong, K. D., Greenlee, E. T., Klein, M. I., & Coyne, J. T. (2023). An investigation of cardiac vagal tone over time and its relation to vigilance performance: A growth curve modeling approach. *Frontiers in Neuroergonomics*, *4*, 1244658.

McIntire, L. K., McIntire, J. P., McKinley, R. A., & Goodyear, C. (2014). Detection of vigilance performance with pupillometry. In *Proceedings of the symposium on eye tracking research and applications* (pp. 167-174).

Meeker, T. J., Emerson, N. M., Chien, J. H., Saffer, M. I., Bienvenu, O. J., Korzeniewska, A., ... & Lenz, F. A. (2021). During vigilance to painful stimuli: slower response rate is related to high trait anxiety, whereas faster response rate is related to high state anxiety. *Journal of neurophysiology, 125*(1), 305-319.

Mehta, R. K., & Parasuraman, R. (2013). Neuroergonomics: A review of applications to physical and cognitive work. *Frontiers in Human Neuroscience, 7,* Article 889.

Meuter, R. F. I., & Lacherez, P. F. (2015). When and Why Threats Go Undetected: Impacts of Event Rate and Shift Length on Threat Detection Accuracy During Airport Baggage Screening. *Human Factors, 58*(2), 218-228.

Molina, E., Sanabria, D., Jung, T. P., & Correa, A. (2019). Electroencephalographic and peripheral temperature dynamics during a prolonged psychomotor vigilance task. *Accident Analysis & Prevention*, 126, 198-208.

Molloy, R., & Parasuraman, R. (1996). Monitoring an Automated System for a Single Failure: Vigilance and Task Complexity Effects. *Human Factors, 38*(2), 311-322.

Moore, J.E., Mascarenhas, A., Bain, J., & Straus, S.E. (2017). Developing a comprehensive definition of sustainability. *Implementation Science 12*(110).

Morales-Alvarez, W., Sipele, O., Léberon, R., Tadjine, H. H., & Olaverri-Monreal, C. (2020). Automated driving: A literature review of the take over request in conditional automation. *Electronics, 9*(2087).

Murray, S., & Amaya, S. (2024). The strategic allocation theory of vigilance. *Wiley Interdisciplinary Reviews: Cognitive Science*, e1693.

Naseer, N., & Nazeer, H. (2017, May). Classification of normal and abnormal ECG signals based on their PQRST intervals. In *2017 International conference on mechanical, system and control engineering (ICMSC)* (pp. 388-391). IEEE.

Naujoks, F., Purucker, C., Wiedemann, K., & Marberger, C. (2019). Noncritical state transitions during conditionally automated driving on german freeways: Effects of non-driving related tasks on takeover time and takeover quality. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 61*(4), 596–613.

Neigel, A. R., Claypoole, V. L., Smith, S. L., Waldfogle, G. E., Fraulini, N. W., Hancock, G. M., & Szalma, J. L. (2020). Engaging the human operator: a review of the theoretical support for the vigilance decrement and a discussion of practical applications. *Theoretical issues in ergonomics science, 21*(2), 239-258.

Nelson, J. T., McKinley, R. A., Golob, E. J., Warm, J. S., & Parasuraman, R. (2014). Enhancing vigilance in operators with prefrontal cortex transcranial direct current stimulation (tDCS). *Neuroimage, 85*, 909-917.

Ni, Y., Weirich, C., & Lin, Y. (2024). Enhanced Visual Performance for In–Vehicle Reading Task Evaluated by Preferences, Emotions and Sustained Attention. *Applied Sciences*, *14*(8), 3513.

Nogueira, M.G., Silvestrin, M., Barreto, C.S.F., Sato, J. R., Mesquita, R. C., Biazoli, C., & Baptista, A. F. (2022). Differences in brain activity between fast and slow responses on psychomotor vigilance task: an fNIRS study. *Brain Imaging and Behavior, 16*, 1563–1574.

Nuechterlein, K. H., Parasuraman, R., & Jiang, Q. (1983). Visual sustained attention: Image degradation produces rapid sensitivity decrement over time. *Science, 220*(4594), 327-329.

Olbrich, S., Mulert, C., Karch, S., Trenner, M., Leicht, G., Pogarell, O., & Hegerl, U. (2009). EEG-vigilance and BOLD effect during simultaneous EEG/fMRI measurement. *neuroimage*, *45*(2), 319-332.

Ou, Y.-K., Huang, W.-X., & Fang, C.-W. (2021). Effects of different takeover request interfaces on takeover behavior and performance during conditionally automated driving. *Accident Analysis & Prevention, 162*, 106425.

Pajcin, M., White, J. M., Banks, S., Dorrian, J., Paech, G. M., Grant, C. L., ... & Della Vedova, C. B. (2019). Effects of strategic early-morning caffeine gum administration on association between salivary alpha-amylase and neurobehavioural performance during 50 h of sleep deprivation. *Accident Analysis & Prevention*, *126*, 160-172.

Pan, Y., Li, Z., Zhang, E., & Guo, Z. (2023). A vigilance estimation method for high-speed rail drivers using physiological signals with a two-level fusion framework. *Biomedical Signal Processing and Control*, *84*.

Pasanen, T., Johnson, K., Lee, K., & Korpela, K. (2018). Can Nature Walks With Psychological Tasks Improve Mood, Self-Reported Restoration, and Sustained Attention? Results From Two Experimental Field Studies. *Frontiers in Psychology*, *9*, 2057.

Patanaik, A., Tandi, J., Ong, J. L., Wang, C., Zhou, J., & Chee, M. W. (2018). Dynamic functional connectivity and its behavioral correlates beyond vigilance. *Neuroimage*, *177*, 1-10.

Patel, M., Lal, S. K., Kavanagh, D., & Rossiter, P. (2011). Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert systems with Applications*, *38*(6), 7235-7242.

Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: selected studies of the Event-Related Potential. *International journal of medical sciences*, *2*(4), 147–154.

Pattyn, N., Neyt, X., Henderick, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue?. *Physiology & Behavior, 93*, 369-378.

Picot, A., Charbonnier, S., Caplier, A., & Vu, N. S. (2012). Using retina modelling to characterize blinking: comparison between EOG and video analysis. *Machine Vision and Applications, 23*(6), 1195-1208.

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of clinical neurophysiology*, *9*, 456-456.

Posada-Quintero, H. F., Bolkhovsky, J. B., Qin, M., & Chon, K. H. (2018). Human performance deterioration due to prolonged wakefulness can be accurately detected using time-varying spectral analysis of electrodermal activity. *Human Factors*, *60*(7), 1035–1047.

Prendez, D., Li, J., Higgins, E. A., & Kim, J. E. (2026). Toward sustainable interventions for enhancing vigilance: A scoping review. *Applied Ergonomics*, *131*, 104628.

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58* (1), 2063–2067.

Roche, F., Somieski, A., & Brandenburg, S. (2019). Behavioral changes to repeated takeovers in highly automated driving: Effects of the takeover-request design and the nondriving-related task modality. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 61*(5), 839–849.

Qi, D., Wong, N. M. L., Shao, R., Man, I. S. C., Wong, C. H. Y., Yuen, L. P., Chan, C. C. H., & Lee, T. M. C. (2021). Qigong exercise enhances cognitive functions in the elderly via an interleukin-6-hippocampus pathway: A randomized active-controlled trial. *Brain, Behavior, and Immunity*, *95*, 381–390.

Ralph, B. C. W., Onderwater, K., Thomson, D. R., & Smilek, D. (2017). Disrupting monotony while increasing demand: Benefits of rest and intervening tasks on vigilance. *Psychological Research*, *81*(2), 432–444.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, *124*(3), 372.

Ritland, B. M., Simonelli, G., Gentili, R. J., Smith, J. C., He, X., Mantua, J., Balkin, T. J., & Hatfield, B. D. (2019). Effects of sleep extension on cognitive/motor performance and motivation in military tactical athletes. *Sleep Medicine*, *58*, 48–55.

Riva, G., Davide, F., & IJsselsteijn, W. A. (2003). Measuring presence: Subjective, behavioral and physiological methods. *Being there: Concepts, effects and measurement of user presence in synthetic environments, 5*, 110-118.

Robertson, I. H., & O'Connell, R. (2010). Vigilant attention. *Attention and time,* 79-88.

Rogers, B. P., Morgan, V. L., Newton, A. T., & Gore, J. C. (2007). Assessing functional connectivity in the human brain by fMRI. *Magnetic resonance imaging*, *25*(10), 1347-1357.

Romeijn, N., & Van Someren, E. J. (2011). Correlated fluctuations of daytime skin temperature and vigilance. *Journal of biological rhythms*, *26*(1), 68-77.

Rose, N. S., Rendell, P. G., McDaniel, M. A., Aberle, I., & Kliegel, M. (2010). Age and individual differences in prospective memory during a" Virtual Week": the roles of working memory, vigilance, task regularity, and cue focality. *Psychology and aging*, *25*(3), 595.

Ross, H. A., Russell, P. N., & Helton, W. S. (2014). Effects of breaks and goal switches on the vigilance decrement. *Experimental Brain Research, 232*, 1729–1737.

Ross, J. A., & Van Bockstaele, E. J. (2021). The locus coeruleus-norepinephrine system in stress and arousal: unraveling historical, current, and future perspectives. Frontiers in psychiatry, 11, 601519.

Rostami, R., Kazemi, R., Mozaffarinejad, F., Nasiri, Z., Rostami, M., L Hadipour, A., & Sadeghihassanabadi, F. (2021). 6 Hz transcranial alternating current stimulation of mPFC improves sustained attention and modulates alpha phase synchronization and power in dorsal attention network. *Cognitive Neuroscience*, *12*(1), 1–13.

Ru, T., Qian, L., Chen, Q., Sun, H., & Zhou, G. (2022). Effects of an afternoon nap on sustained attention and working memory: The role of physiological arousal and sleep variables. *International Journal of Psychophysiology*, *179*, 21–29.

SAE International (2021, April). *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*. https://www.sae.org/standards/content/j3016_202104/.

Salmelin, R., & Baillet, S. (2009). Electromagnetic brain imaging. *Human brain mapping*, *30*(6), 1753.

Samima, S., Sarma, M., & Samanta, D. (2017). Detecting Vigilance in People Performing Continual Monitoring Task. In P. Horain, C. Achard, & M. Mallem (Eds.), *Intelligent Human Computer Interaction* (pp. 202–214). Springer International Publishing.

Sanchis, C., Blasco, E., Luna, F. G., & Lupiáñez, J. (2020). Effects of caffeine intake and exercise intensity on executive and arousal vigilance. *Scientific Reports*, *10*(1), 8393.

Scheirer, M. A. (2013). Linking sustainability research to intervention types. *American journal of public health*, 103(4), e73-e80.

Schmälzle, R., & Grall, C. (2020). Psychophysiological methods: Options, uses, and validity. The international encyclopedia of media psychology, 1-8.

Schmalzl, L., Powers, C., Zanesco, A. P., Yetz, N., Groessl, E. J., & Saron, C. D. (2018). The effect of movement-focused and breath-focused yoga practice on stress parameters and sustained attention: A randomized controlled pilot study. *Consciousness and Cognition*, *65*, 109–125.

Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accident Analysis & Prevention, 41*(5), 1087-1093.

Schuurmans, A. A., de Looff, P., Nijhof, K. S., Rosada, C., Scholte, R. H., Popma, A., & Otten, R. (2020). Validity of the Empatica E4 Wristband to Measure Heart Rate Variability (HRV) Parameters: a Comparison to Electrocardiography (ECG). *Journal of medical systems*, *44*(11), 1-11.

Sciaraffa, N., Borghini, G., Flumeri, G. D., Cincotti, F., Babiloni, F., & Aricò, P. (2021). Joint Analysis of Eye Blinks and Brain Activity to Investigate Attentional Demand during a Visual Search Task. *Brain Sciences*, *11*(5), 562.

Sheliga, B. M., Riggio, L., & Rizzolatti, G. (1994). Orienting of attention and eye movements. *Experimental Brain Research*, *98*(3), 507–522.

Shelton, R.C., Cooper, B.R., & Stirman, S.W. (2018). The sustainability of evidence-based interventions and practices in public health and health care. *Annual Review of Public Health 39*(1), 55–76.

Sheth, S. A., Nemoto, M., Guiou, M. W., Walker, M. A., & Toga, A. W. (2005). Spatiotemporal evolution of functional hemodynamic changes and their relationship to neuronal activity. *Journal of Cerebral Blood Flow & Metabolism*, *25*(7), 830-841.

Shi, L. C., Jiao, Y. Y., & Lu, B. L. (2013). Differential entropy feature for EEG-based vigilance estimation. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 6627-6630). IEEE.

Shi, W., & Rajkumar, R. R. (2021). Work zone detection for autonomous vehicles. *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 1585–1591.

Shibasaki, H. (2008). Human brain mapping: hemodynamic response and electrophysiology. *Clinical Neurophysiology*, *119*(4), 731-743.

Shiferaw, B., Downey, L., & Crewther, D. (2019). A review of gaze entropy as a measure of visual scanning efficiency. *Neuroscience & Biobehavioral Reviews, 96*, 353-366.

Shmueli, K., van Gelderen, P., de Zwart, J. A., Horovitz, S. G., Fukunaga, M., Jansma, J. M., & Duyn, J. H. (2007). Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI BOLD signal. *Neuroimage*, *38*(2), 306-320.

Silver, H. (2023). Working from home: Before and after the pandemic. *Contexts (Berkeley, Calif.), 22*, 66–70.

Sood, A., & Jones, D. T. (2013). On mind wandering, attention, brain networks, and meditation. *Explore, 9*(3), 136-141.

Sowinski, R., Gonzalez, D., Xing, D., Yoo, C., Jenkins, V., Nottingham, K., Dickerson, B., Humphries, M., Leonard, M., Ko, J., Faries, M., Kephart, W., Rasmussen, C.J., & Kreider, R.B. (2021). Effects of inositol-enhanced bonded arginine silicate ingestion on cognitive and executive function in gamers. *Nutrients 13*(11), 3758.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*, 137–149.

Steinbrink, J., Villringer, A., Kempf, F., Haux, D., Boden, S., & Obrig, H. (2006). Illuminating the BOLD signal: combined fMRI–fNIRS studies. *Magnetic resonance imaging*, *24*(4), 495-505.

Steriade, M., Paré, D., Bouhassira, D., Deschênes, M., & Oakson, G. (1989). Phasic activation of lateral geniculate and perigeniculate thalamic neurons during sleep with ponto-geniculo-occipital waves. *Journal of Neuroscience*, *9*(7), 2215-2229.

Stern, J. A., Boyer, D., & Schroeder, D. (1994). Blink Rate: A Possible Measure of Fatigue. *Human Factors*, *36*(2), 285-297.

Subasi, A. (2005). Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. *Expert Systems with Applications, 28*, 701-711.

Szalma, J. L., & Hancock, P. A. (2011). Noise effects on human performance: a meta-analytic synthesis. *Psychological Bulletin*, *137*(4), 682.

Teplan, M. (2002). Fundamentals of EEG Measurement. *Measurement Science Review, 2*(2), 1-11.

Terkelsen, A. J., Gierthmühlen, J., Petersen, L. J., Knudsen, L., Christensen, N. J., Kehr, J., ... & Jensen, T. S. (2013). Cutaneous noradrenaline measured by microdialysis in complex regional pain syndrome during whole-body cooling and heating. *Experimental neurology*, *247*, 456-465.

Thayer, R. (1989). *The psychobiology of mood and arousal.* Oxford: Oxford University Press.

Theeuwes, J., Belopolsky, A., & Olivers, C. N. L. (2009). Interactions between working memory, attention and eye movements. *Spatial Working Memory and Imagery: From Eye Movements to Grounded Cognition*, *132*(2), 106–114.

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.

Tolliver, M., & Sass, J. (2024). Hybrid work has changed meetings forever. *Harvard Business Review, Meeting Management*. https://hbr.org/2024/06/hybrid-work-has-changed-meetings-forever.

Tong, Y., & Frederick, B. D. (2014). Studying the spatial distribution of physiological effects on BOLD signals using ultrafast fMRI. *Frontiers in human neuroscience*, *8*, 196.

US Agency for International Development. (1988). Sustainability of development programs: a compendium of donor experience.

Van Den Brink, R. L., Murphy, P. R., & Nieuwenhuis, S. (2016). Pupil Diameter Tracks Lapses of Attention. *PLOS ONE*, *11*(10), e0165274.

Van Den Heuvel, M. P., & Pol, H. E. H. (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *European neuropsychopharmacology*, *20*(8), 519-534.

Vergara, R. C., Moënne-Loccoz, C., Ávalos, C., Egaña, J., & Maldonado, P. E. (2019). Finger Temperature: A Psychophysiological Assessment of the Attentional State. *Frontiers in Human Neuroscience*.

Vergara, R. C., Moënne-Loccoz, C., & Maldonado, P. E. (2017). Cold-Blooded Attention: Finger Temperature Predicts Attentional Performance. *Frontiers in Human Neuroscience*.

Walters, J., Sheft, S. E., Stellmack, M. A., & Abele, A. (2019). Improving Attentiveness: Effect of Cognitive Training on Sustained Attention Measures. *Professional Safety*, *64*, 31–35.

Walugembe, D. R., Sibbald, S., Le Ber, M. J., & Kothari, A. (2019). Sustainability of public health interventions: where are the gaps?. *Health research policy and systems, 17*(1), 8.

Wang, C., Ong, J. L., Patanaik, A., Zhou, J., & Chee, M. W. (2016). Spontaneous eyelid closures link vigilance fluctuation with fMRI dynamic connectivity states. *Proceedings of the National Academy of Sciences*, *113*(34), 9653-9658.

Wang, W.-F., Chen, C.-M., & Wu, C.-H. (2015). Effects of Different Video Lecture Types on Sustained Attention, Emotion, Cognitive Load, and Learning Performance. *2015 IIAI 4th International Congress on Advanced Applied Informatics*, 385–390.

Warm, J. S., Matthews, G., & Finomore Jr, V. S. (2018). Vigilance, workload, and stress. In *Performance under stress* (pp. 131-158). CRC Press.

Weinger, M. B., & Englund, C. E. (1990). Ergonomic and human factors affecting anesthetic vigilance and monitoring performance in the operating room environment. *Anesthesiology, 73*(5), 995-1021.

Weisgerber, D. M., Nikol, M., & Mistlberger, R. E. (2017). Driving home from the night shift: A bright light intervention study. *Sleep Medicine*, *30*, 171–179.

Wessa, M., Karl, A., & Flor, H. (2005). Central and peripheral psychophysiological responses to trauma-related cues in subclinical posttraumatic stress disorder: a pilot study. *Experimental brain research, 167*(1), 56-65.

White, P. M., & Yee, C. M. (2006). P50 sensitivity to physical and psychological state influences. *Psychophysiology*, *43*(3), 320-328.

Wong, C. W., Olafsson, V., Tal, O., & Liu, T. T. (2013). The amplitude of the resting-state fMRI global signal is related to EEG vigilance measures. *Neuroimage*, *83*, 983-990.

Xu, J., Zhang, W., Yu, J., Li, G., Cui, J., Qi, H., Zhang, M., Li, M., Hu, Y., Wang, H., Min, H., Xu, F., Xu, X., Zhu, C., Xiao, Y., & Zhang, Y. (2024). Functional near-infrared spectroscopy-based neurofeedback training regulates time-on-task effects and enhances sustained cognitive performance. *Cerebral Cortex*, *34*(6), bhae259.

Yamashita, J., Terashima, H., Yoneya, M., Maruya, K., Koya, H., Oishi, H., Nakamura, H., & Kumada, T. (2021). Pupillary fluctuation amplitude before target presentation reflects short-term vigilance level in Psychomotor Vigilance Tasks. *PLoS One*, *16*(9), e0256953.

Yoon, S. A., & Weierich, M. R. (2016). Salivary biomarkers of neural hypervigilance in trauma-exposed women. *Psychoneuroendocrinology*, *63*, 17-25.

Yuda, E., Yoshida, Y., & Hayano, J. (2017). Blue glasses increase your alertness: Effects of colored glasses on psychomotor performance. *2017 6th International Conference on Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT)*, 1–4.

Yuda, E., Yoshida, Y., & Hayano, J. (2018). Effects of Chewing Gum on Differential Components of Psychomotor Vigilance: Improved Sustained Attention with Prolonged Reaction Time. *2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*, 55–58.

Yung, T. W., Lai, C. Y., Chan, J. Y., Ng, S. S., & Chan, C. C. (2020). Neuro-physiological correlates of sluggish cognitive tempo (SCT) symptoms in school-aged children. *European Child & Adolescent Psychiatry, 29*, 315-326.

Zaki, A. (2023). Remote meetings up 60% since 2020: Weekly stat. *CFO, Corporate Finance*. https://www.cfo.com/news/remote-meetings-up-60-since-2020-weekly-stat/654749/.

Zare, S., Esmaeili, R., & Zandsalimi, F. (2024). Modeling and Weighting of Factors Affecting Sustained Attention and Working Memory of Workers Exposed to Different Sound Pressure Levels using Deep Learning and Random Forest Algorithms: A Case Study of a Steel Industry. *Journal of Health & Safety at Work*, *14*(3).

Zhang, B., De Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation Research Part F: Traffic Psychology and Behaviour, 64*, 285–307.

Zhang, D., & Raichle, M. E. (2010). Disease and the brain's dark energy. *Nature Reviews Neurology*, 6(1), 15-28.

Zhang, S., Wang, D., Afzal, N., Zhang, Y., & Wu, R. (2016). Rhythmic Haptic Stimuli Improve Short-Term Attention. *IEEE Transactions on Haptics*, *9*(3), 437–442.

Ziegler, D. A., Simon, A. J., Gallen, C. L., Skinner, S., Janowich, J. R., Volponi, J. J., Rolle, C. E., Mishra, J., Kornfield, J., Anguera, J. A., & Gazzaley, A. (2019). Closed-loop digital meditation improves sustained attention in young adults. *Nature Human Behaviour*, *3*(7), 746–757.

Zsido, A. N., Teleki, S. A., Csokasi, K., Rozsa, S., & Bandi, S. A. (2020). Development of the short version of the spielberger state-trait anxiety inventory. *Psychiatry research*, 291, 113223.

Appendix

The following table, *Table 16*, summarizes the predictor variables in the healthcare study's dataset as well as the outcome variable for vigilance (reaction time during the PVT). The *Data Source* column refers to the task, questionnaire, or sensor that the variable's data is obtained from. The *Variable* column is the variable's name, while the *Data Code* column shows the variable's coded name during analysis in Python and R programming languages. These coded names are used in some of the data visualizations in *Chapter 4*. The *Levels* column indicates the different levels for categorical variables and the frequency of that group by percentage. The *Mean(SD)* column reflects the average and standard deviation values of the continuous or numerical variables in the dataset. The *Units* column refers to the unit used to measure the variable.

Table 16. Healthcare Study Dataset Summary

Data Source	Variable	Data Code	Levels (%)	Mean (SD)	Units
PVT	Average reaction time	RT_mean	-	337.8 (34.8)	ms
Empatica E4	Average blood volume pulse	BVP_mean	-	0.1 (0.4)	N/A
Empatica E4	Std. dev. blood volume pulse	BVP_std	-	6.8 (5.2)	N/A
Empatica E4	Average electrodermal activity	EDA_mean	-	1.2 (1.4)	µS
Empatica E4	Std. dev. electrodermal activity	EDA_std	-	0.3 (0.3)	µS

Data Source	Variable	Data Code	Levels (%)	Mean (SD)	Units
Empatica E4	Average skin temperature	TEMP_mean	-	32.1 (1.5)	°C
Empatica E4	Std. dev. skin temperature	TEMP_std	-	0.3 (0.5)	°C
Empatica E4	Average heart rate	HR_mean	-	75.8 (8.3)	bpm
Empatica E4	Heart rate variability	HR_std	-	2.9 (3.1)	bpm
Tobii Eye Tracker	Average left pupil diameter	Pupil_DIA_L_mea n	-	4.9 (0.7)	mm
Tobii Eye Tracker	Std. dev. left pupil diameter	Pupil_DIA_L_std	-	0.3 (0.2)	mm
Tobii Eye Tracker	Average right pupil diameter	Pupil_DIA_R_mea n	-	4.9 (0.7)	mm
Tobii Eye Tracker	Std. dev. right pupil diameter	Pupil_DIA_R_std	-	0.4 (0.2)	mm
Tobii Eye Tracker	Average saccade duration	Sac_duration_mea n	-	29.6 (24.2)	ms
Tobii Eye Tracker	Std. dev. saccade duration	Sac_duration_std	-	11.9 (16.6)	ms
Tobii Eye Tracker	Average eye fixation duration	Fix_duration_mean	-	727.4 (647.5)	ms
Tobii Eye Tracker	Std. dev. eye fixation duration	Fix_duration_std	-	427.03 (412.0)	ms
Tobii Eye Tracker	Total saccade count	Sac_count_sum	-	1902.0 (1001.9)	-
Tobii Eye Tracker	Std. dev. saccade count	Sac_count_std	-	2.5 (1.2)	-
Tobii Eye Tracker	Total fixation count	Fix_count_sum	-	1041.4 (435.9)	-

Data Source	Variable	Data Code	Levels (%)	Mean (SD)	Units
Tobii Eye Tracker	Std. dev. fixation count	Fix_count_std	-	0.9 (0.2)	-
N/A	Participant ID	PID	-	-	-
Screening Survey	Gender	Gender	Male (50%), Female (42%), Non-Binary (8%)	–	–
Screening Survey	Age	Age	–	29.3 (2.3)	years
Screening Survey	Years of experience	Experience	–	0.6 (0.8)	years
N/A	Visit Condition	Visit	Sleep deprived (42%), Non-sleep deprived (58%)	–	–
Room Thermo-meter	Temperature of the room	Room_temp	–	21.9 (0.8)	°C
Room Thermo-meter	Humidity of the room	Humidity	–	42 (0.1)	%
Sleep Diary	Quality of sleep two nights prior	Sleep_quality_1	Very poor (4%), Fairly poor (29%), Fairly good (63), Very good (4%)	–	–
Sleep Diary	Times woken up two nights prior	Times_woke_1	–	1.4 (1.2)	–
Sleep Diary	Hours slept two nights prior	Hours_sleep_1	–	7.2 (1.1)	hours
Sleep Diary	Quality of sleep prior night	Sleep_quality_2	Very poor (12%), Fairly poor (33%), Fairly good (38%), Very good (17%)	–	–

Data Source	Variable	Data Code	Levels (%)	Mean (SD)	Units
Sleep Diary	Times woken up prior night	Times_woke_2	–	1.3 (2.1)	–
Sleep Diary	Hours slept prior night	Hours_sleep_2	–	6.5 (2.1)	hours
STAI	State-anxiety level	STAI_State	–	6.7 (1.7)	–
STAI	Trait-anxiety level	STAI_Trait	–	9.0 (2.7)	–
PSQI	Sleep quality over past month	PSQI_month	–	2.2 (0.9)	–
PSQI	Sleep quality over past week	PSQI_week	–	2.5 (1.2)	–
NASA TLX	Workload	TLX	–	7.7 (3.3)	–

The next table, *Table 17*, summarizes the predictor variables in the driving study's dataset as well as the outcome variable for vigilance (response time during the takeover). The format of the table matches the previous format for the healthcare study dataset.

Table 17. Driving Study Dataset Summary

Data Source	Variable	Data Code	Levels (%)	Mean (SD)	Units
Screening Survey	Gender	Gender	Female (61%), Non-Female (39%)	–	–
Screening Survey	Age	Age	–	29.5 (8.7)	years

Data Source	Variable	Data Code	Levels (%)	Mean (SD)	Units
Screening Survey	Education Level	Education	High School (28%), Undergrad. (25%), Graduate (47%)	–	–
Screening Survey	Average Miles Driven Per Year	DrivingMiles	Under 3k (33%), 3k-5k (25%), 5k-10k (22%), Over 10k (19%)	–	–
Screening Survey	Weekly Driving frequency	DrivingWeekly	Less than Weekly (22%), Some Days (39%), Most Days (17%), Every Day (22%)	–	–
Screening Survey	Virtual Meeting Experience	MeetingFrequency	Less than Weekly (17%), More than Weekly (61%), Every Day (22%)	–	–
Technology Acceptance Questionnaire	AV Acceptance	Acceptance2	–	3.4 (1.0)	–
SDDQ	Distraction Susceptibility	DistractionSuscept	–	3.2 (0.3)	–
Trust Questionnaires	Pre-Takeover Trust in AVs	PreT_TrustinADS	–	7.4 (2.3)	–
NASA-TLX	Pre-Takeover Workload	PreT_Workload	–	3.4 (1.4)	–
NASA-TLX	Intra-Takeover Workload	IntraT_Workload	–	4.5 (1.3)	–

Data Source	Variable	Data Code	Levels (%)	Mean (SD)	Units
SART	Pre-Takeover Situation Awareness	PreT_SituationAwareness	–	5.6 (2.4)	–
SART	Intra-Takeover Situation Awareness	IntraT_SituationAwareness	–	3.1 (2.5)	–
N/A	Takeover Warning	TakeoverCondition	Planned (50%), Emergency (50%)	–	–
Driving Simulator	Average Speed	avg_speed	–	53.9 (5.9)	mph
Driving Simulator	Maximum Speed	max_speed	–	60.8 (2.9)	mph
Driving Simulator	Standard Deviation of Speed	sd_speed	–	4.2 (3.7)	mph
Driving Simulator	Average Acceleration	avg_acc	–	-1.1 (1.4)	ft/s^2
Driving Simulator	Maximum Acceleration	max_acc	–	2.6 (3.6)	ft/s^2
Driving Simulator	Maximum Steering Wheel Angle Rate	max_sw_rate	–	40.6 (30.9)	deg/s
Driving Simulator	Standard Deviation of Steering Wheel Angle Rate	sd_sw_rate	–	10.9 (6.9)	deg/s
Driving Simulator	Maximum Steering Wheel Angle	max_sw_angle	–	7.9 (4.5)	deg

Data Source	Variable	Data Code	Levels (%)	Mean (SD)	Units
Driving Simulator	Standard Deviation of Steering Wheel Angle	sd_sw_angle	–	5.5 (3.4)	deg
Driving Simulator	Maximum Brake Force	max_brake_force	–	21.8 (36.1)	lbs
Video Data	Response Time	response_time	–	3.2 (2.4)	s
Virtual Meeting	Pre-Takeover Virtual Meeting Performance	MeetingAvgPrior	–	0.9 (0.1)	–
Virtual Meeting	Intra-Takeover Virtual Meeting Performance	MeetingPerf	–	0.6 (0.4)	–
Eye Tracking	Pre-Takeover Average Fixation Duration (iPad)	avg_ipad_fix_prior	–	3.6 (4.7)	s
Eye Tracking	Pre-Takeover Average Fixation Duration (Road)	avg_road_fix_prior	–	4.5 (3.3)	s
Eye Tracking	Pre-Takeover Total Fixation Duration (iPad)	dur_ipad_fix_prior	–	29.2 (18.6)	s
Eye Tracking	Pre-Takeover Total Fixation Duration (Road)	dur_road_fix_prior	–	45.2 (19.0)	s

Data Source	Variable	Data Code	Levels (%)	Mean (SD)	Units
Eye Tracking	Pre-Takeover Number of Fixations (iPad)	num_ipad_fix_prior	–	12.0 (4.9)	–
Eye Tracking	Pre-Takeover Number of Fixations (Road)	num_road_fix_prior	–	12.2 (4.9)	–
Eye Tracking	Intra-Takeover Average Fixation Duration (iPad)	avg_ipad_fix_takeover	–	0.1 (0.2)	s
Eye Tracking	Intra-Takeover Average Fixation Duration (Road)	avg_road_fix_takeover	–	6.0 (3.6)	s
Eye Tracking	Intra-Takeover Total Fixation Duration (iPad)	dur_ipad_fix_takeover	–	0.2 (0.4)	s
Eye Tracking	Intra-Takeover Total Fixation Duration (Road)	dur_road_fix_takeover	–	7.7 (4.6)	s
Eye Tracking	Intra-Takeover Number of Fixations (iPad)	num_ipad_fix_takeover	–	0.4 (0.5)	–
Eye Tracking	Intra-Takeover Number of Fixations (Road)	num_road_fix_takeover	–	1.3 (0.5)	–