Dynamic Alert Design Based on Driver's Cognitive State for Take-over Request in Automated Vehicles

by

Wachirawit Umpaipant

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2024

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis investigates the effectiveness of dynamic alert systems tailored to drivers' cognitive states in automated driving environments, focusing on enhancing takeover readiness during critical transitions. Utilizing a large-scale immersive driving simulation, the study evaluated drivers' response times and physiological measures when reacting to various alert intensities and the presence of a secondary typing task.

The experiment revealed that dynamic alerts significantly improved response times and takeover performance, especially in high-distraction scenarios. Drivers responded more effectively when alerts were adjusted to their cognitive load, with strong alerts resulting in the fastest reaction times under distracted conditions. On average, dynamic alerts reduced response times by approximately 1.75 seconds compared to static alerts. Additionally, higher lateral accelerations were observed under strong alerts, indicating more decisive maneuvering.

Self-rated attention-capturing scores were notably higher with dynamic alerts, particularly under strong alert conditions and in the presence of secondary tasks. The ANOVA results showed significant improvements in attention capturing and overall alert effectiveness when dynamic alerts were employed, demonstrating the robust design's ability to capture attention and enhance driver responsiveness. The study confirmed that adaptive alert designs, which adjust based on the driver's cognitive state, can markedly enhance overall driving experience and safety. Participants reported higher levels of confidence with dynamic alerts, especially in scenarios involving secondary tasks. Despite the strong alerts, annoyance levels remained low, indicating that dynamic alerts are effective without causing undue stress.

These results underscore the potential of using adaptive systems to improve safety and efficiency in automated driving, advocating for a more nuanced approach to system alerts that considers the variable cognitive states of drivers. Future research should validate these findings with on-road studies, explore a broader range of alert modalities, and refine physiological monitoring techniques to further enhance adaptive alert systems.

Acknowledgments

I would like to thank all the people who made this thesis possible. My sincere gratitude goes to Prof. Samuel and Prof. Burns for their excellent supervision and guidance. I am also thankful to Prof. John McPhee for providing access to the necessary equipment. I appreciate the support from the VI-Grade Company team and Xavier for their help with simulation development. Lastly, I am deeply grateful to my wife and family for their unwavering support and encouragement throughout this journey.

Dedication

This thesis is dedicated to my beloved family. Your unwavering love and support have made this journey possible. Thank you for always being there for me.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACC	Adaptive Cruise Control
ADAS	Advanced Driver Assistance Systems
ADS	Automated Driving System
AOI	Area of Interest
AP	Autopilot
BVP	Blood Volume Pulse
HMI	Human-Machine Interface
HR	Heart Rate
IBI	Inter-Beat Interval
NDRT	Non-Driving Related Task
SA	Situation Awareness
SAE	Society of Automotive Engineers
TOR	Takeover Request

Chapter 1

Introduction

1.1 Automated Driving System (ADS)

Advancements in the automotive industry have made significant progress over the years. As there are stronger regulations for road safety and increased market demands for comfort, autonomous driving opens up multiple potential areas for research and development, as it not only promises both aspects, but many more [1]. Progress in developing a fully autonomous vehicle starts with improving partially automated systems where the involvement of the driver remains intact. Therefore, not only the aspects of driver performance but also the driving experience in advanced automated systems are important to address.

According to the Society of Automotive Engineers (SAE), a vehicle with Automated Driving System (ADS) can be classified into five levels, ranging from simple driver assistance systems, such as automatic emergency braking, to fully automated systems where the vehicle fully drives itself and handles all driving tasks. In terms of driver responsibility, it is defined that at levels 2 (partial driving automation) or less, the driver is required to drive at all times whereas the ADS should only be perceived as an assistant. With a higher level of automation, the need for driver input is significantly less, ranging from level-3 automation where the driver is only required to drive when necessary to where driver input is never needed to safely drive for level-5 automation [2]. However, in level-2 automation with advanced ADS, such as highway assist or Autopilot (AP) feature, most driving tasks are fully handled by automation systems that can include adaptive cruise control Adaptive Cruise Control (ACC), lane tracking, and auto braking. With a much lower workload of the driving task required when cruising on a highway, the driver could be over-trusting in the automation and become drowsy or lose a good vigilance level to maintain the safe driving cognitive state. For instance, the driver may shift to non-driving tasks, such as reading or texting. In case of emergency or unexpected events, the automation system may not safely handle the situation on its own, and it becomes critical to ensure safety by transitioning the driving task from automation to manual driving by humans. This process often involves a "Takeover Request" (TOR) or alert used for prompting the driver to intervene and resume manual driving. Therefore, it is important to consider the alert design for the best possible transition outcome.

1.2 Takeover Request (TOR)

The concept of a TOR is activated as a warning signal when the Advanced Driver Assistance Systems (ADAS) reaches its safety confidence limit, which can occur due to various factors when the situation becomes too intricate for the vehicle to manage independently [3]. While a level 2 vehicle appears proficient in handling driving tasks on a highway or addressing straightforward hazards, it may struggle when faced with more complicated scenarios, such as an unforeseen road closure or sudden road hazard, where the automation system may not fully grasp the dynamic surroundings. TOR plays a critical role in involving the human driver to react safely to such situations. The primary goals of developing TOR are to improve driver performance and ensure a secure and seamless transition during maneuvers. Engaging the driver's attention poses a challenge, particularly when dealing with different cognitive states of drivers. The key considerations in TOR design revolve around determining the optimal timing for alerts and the most effective means through which the system should notify the driver.

1.2.1 Interface Design for Alert in Automated Driving

The primary goal of alerts is to convey unexpected or potential hazards detected by the system to the human operator. The Human-Machine Interface (HMI) enables the system to communicate this message to the operator using various methods. In the context of autonomous driving, common alerting modalities include auditory signals (such as tones and verbal messages)[4], [5], visual cues (such as lights, icons, and animations) [6], [7], and haptic feedback (such as vibrations and object movements) [8], [9]. These modalities are essential for ensuring that drivers can respond promptly and accurately to takeover requests. However, some researchers contend that no universally superior set of TOR modalities can be universally applied, given the dynamic nature of situations and individual preferences. Studies [10] explore how driver states, particularly drowsiness and distraction, affect responses to multimodal alarms. Their findings suggest that adaptive alarm systems, which adjust modalities and intensities based on the driver's state, can improve reaction times and situational awareness during takeover requests in autonomous vehicles. The effectiveness of alerts is heavily influenced by how well the machine communicates with the user and how the user perceives the message. The significance of alert designs in HMI varies depending on the specific situation and system. Insufficient alerting for TOR in autonomous driving could lead to unsafe transitions that jeopardize the well-being of the driver or other road users.

1.2.2 Adaptive Driver Support Design

Each person brings a unique set of experiences, skills, and preferences. Additionally, incorporating intermittent thoughts and different states of mind, different people performing similar tasks will vary in perceptions and established procedures that shape different responses and outcomes [7]. Recent studies [5], [8], [11] suggest the need for adaptive alert systems that account for pre-warning or current driver's cognitive states. In the context of the autonomous driving takeover process, individual cognitive states could significantly affect performance when a driver transitions from a supervisory role to actively engaging in the driving task. The importance of this study revolves around the impact of these factors on the takeover process and how systems can be tailored to align with the individual and recurring cognitive state to improve alert efficiency.

1.3 Objectives and Hypotheses

The objective of this thesis is to explore the design of dynamic alert levels that are adapted to the cognitive state of the driver to optimize the performance of the takeover and the driving experience. There will be a comparison between a single type of alert and dynamic alert in different cognitive states of the driver when using AP on the highway. The effect of the secondary task on takeover reactions and performance will also be investigated using driving behavior, physiological data, and qualitative measures in each alert design.

It is hypothesized that dynamic alert intensity that varies depending on the driver's cognitive state enhances the driving experience while maintaining takeover quality. Specifically:

- **H1**: Dynamic alert systems, which adjust alert intensity based on the driver's cognitive state, will result in enhances the driving experience while having better or similar response times and quality of takeover performance compared to static (single) alert systems.

- **H2**: Drivers will perceive dynamic alerts as more effective and less annoying compared to static alerts, particularly in scenarios involving high cognitive load or distraction.

These hypotheses aim to compare the efficacy of static and dynamic alert systems in enhancing driver performance and safety.

Chapter 2

Literature Review

2.1 TOR in Automated Driving

In Level 2 ADS, which are not fully equipped for all road conditions, driver intervention is often necessary to assist in both urgent and non-urgent driving tasks. Current automated vehicles are equipped with safety features that prompt the driver to refocus on the driving task. Numerous studies have explored the effects of various TOR methods and factors on the handover process. This review will address key aspects to enhance understanding of TOR. Generally, TOR initiation begins when the system recognizes its limitations and seeks to transfer control from automated to manual driving before these limits are reached to prevent hazardous situations. Although drivers are expected to be fully attentive and prepared for takeover, real-life scenarios often differ, as distractions from secondary tasks or drowsiness can lead to delayed responses or inadequate handover, both mentally and ergonomically. Researchers have faced challenges in identifying optimal solutions for such situations. Many studies have been conducted in simulated environments to control external influences and mitigate unknown risks, while some have tested Level 2 ADS on actual roads.

2.1.1 TOR Alert Timing

Situational factors such as TOR lead time, the frequency of takeover requests, and the type of scenario (static vs. dynamic obstacles) significantly influenced takeover criticality. When traveling at speed, timing becomes a critical factor, especially in TOR situations. The timing, determined by the distance to a predefined system limitation, such as a construction site ahead, and the vehicle's current condition like speed and trajectory, can be defined as the time to collision (TTC) in seconds (s). It is important that the TOR is issued before TCC, but how long before is the main factor that researchers have discussed. In early studies, Gold et al. compared different predefined times to find the optimal time for TOR situations [12], but some others argue that there is no single TOR timing that suits all driving situations when considering the varieties in road conditions and individual factors such as type of trajectory and TOR performance [3]. Shorter TOR lead times and dynamic obstacles were associated with higher takeover criticality, suggesting that drivers need more time to respond effectively in more complex or sudden scenarios [13].

2.1.2 Takeover Performance

The safety and quality with which a driver handles the control transition from the automation system is referred to as takeover performance. Several aspects of takeover performance are measured to assess the proper transfer of vehicle control, which typically occurs when a driver reaches and manipulates the steering wheel or uses the brake pedal to avoid a hazard. Reaction time and the time margin to a collision are key metrics compared across various designs of the TOR process measured in seconds; studies have shown that the longest time margin may not always be the best time allotted to the driver.[14] Longitudinal and lateral acceleration measures, in g or m/s, are used in experiments as indicators of a safe takeover, with less acceleration generally being safer.[14] In simulated environments, where driving scenarios can be more severe, the number of collisions is also counted to determine the success of the takeover.[3]

2.2 In-Vehicle Warning System

The modern vehicle is equipped with a HMI that allows the system to communicate with the driver through various modes of communication. These can include auditory signals through vehicle speakers [15], tactile or haptic feedback on the steering wheel [16], and visualizations on the vehicle cluster or display [17], [18]. These modalities can be designed to function individually (unimodal) or in combination (multimodal). When a vehicle requires handover, it communicates with the driver through these channels to facilitate a successful takeover process. The amount of information and how it is delivered to the driver are critical factors considered in the design of these modalities, as explored in several studies

[19]. Research by Seppelt and Lee focused on dynamic feedback mechanisms that are designed to keep the driver engaged, particularly in the context of semi-automated driving systems. Their study underscores the importance of feedback systems that adapt based on the driver's level of engagement and the vehicle's operational status. Such dynamic systems are essential to ensure that drivers are not only aware but also prepared to take control of the vehicle in situations where automation may fail. The study supports the development of feedback that responds to the user's state and external conditions, thereby enhancing safety and effectiveness in automated driving systems [20].

Specifically for the design of HMI in TOR situations, a comprehensive review of the literature examines how advanced in-vehicle HMIs convey additional semantic information from driving automation systems to drivers [21]. The focus is on the transfer of control in automated driving, particularly just before, during, and immediately after a TOR. The paper categorizes this additional semantic information into three main areas: enhancing mode awareness, enhancing situation awareness, and assisting takeover maneuvers. This information helps drivers understand the vehicle's current status, the surrounding traffic conditions, and the actions they may need to undertake. Advanced HMIs are crucial for delivering contextually rich and timely information, aiding drivers in making informed decisions during critical phases of driving automation. This is particularly important during the transition from automated driving to manual control, where there is a risk of drivers being "out of the loop" in terms of engagement and situational awareness. The review identifies gaps in current studies, such as the limited exploration of the cognitive impacts of HMI on driver behavior during TOR, and suggests areas for further research. It emphasizes the need for HMIs that can adaptively deliver information tailored to the driving context and driver state to enhance both safety and the driver experience.

2.2.1 Alert Modalities

Auditory Alerts

Auditory alerts are a common modality in driving contexts, known for their ability to rapidly capture attention. Studies have shown that auditory alerts can be highly effective in prompting driver responses during TORs. For example, [12] found that auditory alerts led to faster reaction times compared to visual alerts in high-autonomy driving scenarios. However, the effectiveness of auditory alerts can be influenced by factors such as volume, pitch, and the presence of other auditory stimuli in the vehicle's environment. Gonzalez et al. investigated how auditory alert on different levels of pulse rate and loudness affects the urgency and annoyance perception of drivers, they found that perceived urgency and annoyance proportionally increase when either pulse rate or loudness levels increase [22].

Not only alert sound is presented in actual driving, Šabić et al. suggested that the design of auditory alert should also consider the level of background noises, e.g., traffic noise, music, engine, and external warning sounds, for an appropriate loudness to ensure to capture driver's attention. They presented the advantage of having adaptive loudness of auditory alert which resulted in [23], despite the mixed results, the approach reflects a significant step towards creating more intelligent and responsive in-vehicle alert systems that enhance user experience and safety by maintaining the intelligibility of critical alerts without increasing annoyance.

Visual Alerts

Visual alerts leverage the driver's visual field to signal the need for a takeover. While effective in providing detailed information, their efficacy can be diminished if the driver's visual attention is not on the display or if the visual stimuli are too subtle. Lees et, al. emphasize the importance of visual alert design, noting that conspicuousness and clarity are crucial to ensure that drivers can quickly interpret and act on the information presented [6], [24].

Visual warnings can potentially reduce the likelihood of accidents involving young drivers, especially in complex driving environments where latent hazards are present, as conducted to investigate the effectiveness of visual warnings on the hazard anticipation and mitigation abilities of young drivers using head-up displays (HUD), it was concluded that the alerts effectively increased the likelihood of drivers glancing toward latent pedestrian and vehicle hazards, allowing the potential benefits of integrating advanced visual warning systems into vehicles to support young drivers [18], [25].

Tactile Alerts

Tactile alerts, or haptic feedback, offer a distinct advantage by engaging the sense of touch, which is less likely to be overloaded compared to auditory and visual senses. Haptic steering wheels, for example, can provide immediate and intuitive cues for drivers to take control. Multiple studies demonstrate that tactile alerts can be effective. Gomes et al. presented an approach to enhance driver situation awareness upon TOR through continuous-time feedback using a device called the Adaptive Tactile Device (ATD) through the steering wheel or driver's seat, with the aim of improving the transitional phase where the driver takes over from the automated system by providing force haptic feedback that adjusts to the driving conditions, it resulted in improved reaction time during TOR events [26]. However, studies by Geitner et al. found that tactile alerts alone were responded by drivers significantly slower than others and also significantly induced more alarm startling when combined with auditory alerts.[27].

Multimodal Alerts

Multimodal alerts refer to the combination of two or more monomodal alerts together to represent the same alert message, combining multiple alert modalities can capitalize on the strengths of each to ensure that drivers receive an unambiguous takeover request. The combination of auditory, visual, or tactile cues can send information to different sensory channels, resulting in less number of missed warnings or false responses from the driver [27], [28]. Geitner et al. suggested that combining all three modalities would enhance attention capturing which reduces false responses and reduces reaction times when prompted timely [27]. Zhang and Tan conducted a study to compare two unimodal (visual and auditory), three bimodal (visual-auditory, visual-tactile, auditory-tactile), and one trimodal (visual-auditory-tactile), the results highlighted the potential of combining alert modalities, the key finding was significant improvement in driver reaction times, trust and satisfaction ratings for multimodal warnings compared to unimodal. Interestingly, warnings that included tactile signals, despite increasing physiological arousal, also heightened user annoyance [29].

When a driver is under the influence of a different cognitive load. Han and Ju proposed an advanced alarm method tailored to the driver's state in autonomous vehicles, focusing on the reaction of drivers in drowsy and distracted states to various alarm modalities. The study emphasizes the importance of multimodal alarms (visual and auditory) and examines the effectiveness of these alarms through a driving simulation with 38 participants. The results indicate that drivers show different response patterns based on their cognitive state, with auditory alarms having a significant impact on alerting drivers [10]. Also, findings in a study conducted by Van der Heiden et al. agreed that multimodal alerts can effectively capture the driver's attention, especially when engaged in cognitively demanding tasks [5].

2.2.2 Alert Urgency and Intensity

There is a fundamental relationship between perceived urgency and the signal's intensity that is determined by the frequency, wavelength, pace, duration, etc. Higher frequency and/or decreased interval between sound pulses can be perceived as increasingly urgent [30], but this also proportions to annoyance [22], [31]–[33]. This also applies to other types of monomodalities [34] and in a combination of modalities such as auditory-tactile and visual-tactile alert systems, perceived urgency should be an appropriate level of urgency as the literature suggested that high urgency warnings should be used for a critical driving task. Therefore, signal urgency is important to consider in the design for TOR [35]–[37]. The modalities of the signal involving visual cues are perceived as more urgent in the signal, while the tactile modality is rated as more annoying in one of the studies [35].

In the study by Roche and Brandenburg (2018), alert urgency and intensity were measured using subjective unidimensional 8-point rating scales. Participants rated the perceived criticality of the driving situation and the urgency of TOR on scales ranging from 0 (not critical/not urgent) to 7 (very critical/very urgent). Additionally, the visual TORs were designed to vary in urgency by using different wordings: "Please take over soon" for low urgency and "Danger: Take over now" for high urgency [36].

2.2.3 Dynamic Interface for Vehicle Warning System

Considering the aspects of safety and user experience in the design of a vehicle warning system, the most effective alert design created in one research may not be the best application in all use cases when taking into account driver preference and concurrent cognitive states. The adaptive alert concept has been researched for its effectiveness as dynamic levels of alert by varying alert intensities Seppelt et al. studied dynamic feedback to keep drivers in the loop, resulting in increasing proactive responses to incidents, and also found that a visual-auditory interface performed the best compared to a unimodal interface.[20] These studies collectively underscore the importance of developing dynamic alert systems that are sensitive to changes in environmental conditions and user states.

2.3 Cognitive Ergonomics Influencing Takeover Performance

In high-level ADS, the vehicle assumes more responsibility for managing traffic and vehicle control, thereby reducing the driver's workload. However, human errors remain a significant concern because the driver must be prepared to retake control of the vehicle at any moment. The behavior of drivers in partially or highly automated vehicles is expected to differ significantly from that in manual driving. Despite Level 2 automation still requiring a high level of driver attention to respond to system requests in both urgent and normal situations, drivers often lose focus due to distractions, drowsiness, or engagement in non-driving tasks. This presents a major safety concern as ADS advances but has not yet reached the level of full autonomy. Consequently, it is crucial to understand how cognitive ergonomics influence takeover performance in such systems. Several researchers are investigating the interplay between human factors and automated driving to enhance system design. Their goal is to improve driver performance and address safety issues. These studies aim to identify strategies to keep drivers engaged and ready to take over when necessary, thereby mitigating the risks associated with reduced vigilance and increased workload during partial automation.

2.3.1 Situation Awareness (SA)

Introduced in aviation psychology in the 1980s in order to explain human errors related to the operator's ability to achieve and maintain a current understanding of a dynamic situation, then later applied to many other areas. Mica Endsley defined SA into three levels "(1) the perception of the elements in the environment within a volume of time and space, (2) the comprehension of their meaning and 'the projection of their status shortly' and their dependence that the higher levels rely on the success of the lower levels of SA [38]. Based on the assumption of Endley's framework, in the driving context, Matthews et al. introduced the SA model for driving to improve understanding of the driving behavior that driving is directed towards three types of goals: strategic, tactical, and operational [39]. These goals are said to define the information driver needed to decide to achieve the goal. Moreover, as the tasks and goals are different between several types of tasks, different levels of SA also differ in each one. They believe that strategic goals, such as route planning, require a stronger projection SA than tactical or operational driving goals to project the future state. On the other hand, perception and comprehension are likely to be more crucial in selecting the most adequate operation in the current situation than projection. SA can be interpreted as a state of knowledge of the operator's mental representation of the current situation where researchers continue to find ways to measure, one of the methods is called SAGAT (Situation Awareness Global Assessment Technique) introduced by Endsley in 1988 [40], in which applied to the driving context, the driver will be asked questions about the driving state and the environment. During automated driving, the driver may also engage in non-driving tasks, such as using a cell phone, and a decrease in SA may lead to declines in driving performance [41]. In the context of automation systems, Baumann et al. identified key challenges posed by automation, such as the reduction in driver vigilance and engagement, which could compromise the driver's ability to swiftly reassume control when necessary. They discussed various levels of driving automation, from no automation to full automation, and the corresponding shifts in the role of the driver from active participant to supervisor. Crucially, the paper stresses the need for systems that provide adequate feedback on automation states and behaviors to mitigate negative effects on human performance and enhance situation comprehension. This is essential to ensure that drivers are not only kept in the loop but are also prepared to takeover control effectively, maintaining both safety and operational efficiency in increasingly automated vehicles [42].

2.3.2 Mental Workload and Distraction

Driving automation reduces the driver's workload as the responsibility of the driving task is transferred to the vehicle. As the vehicle's capability advances, it handles more of this workload. However, higher levels of automation correlate with decreased driver attention, resulting in distractions and off-road glances.[1] Instead of maintaining eye fixation on the road, drivers in highly automated vehicles are more likely to engage in secondary tasks or NDRT, shifting their focus to non-driving activities. This transition from low to high workload, especially in automated settings, can detrimentally affect situational awareness. Drivers may not be fully aware or prepared to take control in critical situations if they are engaged in NDRT that draws their attention away from driving [43]. Studies investigated the impact of NDRT on cognitive workload in automated driving using the Twenty Questions Task, which can significantly distract drivers enough to impair their ability to promptly respond to unexpected driving conditions or emergencies, resulting in poorer takeover performance in terms of reaction time and performance [43], [44]. Therefore, it is important to design appropriate human-machine interaction strategies that support cognitive processes and driver performance in situational awareness to avoid or minimize the potential negatives of automation.

In autonomous driving contexts, distractions impact drivers' mental workload and reduce their engagement and safety perception, which is especially problematic in automated driving settings where the driver's readiness to takeover is crucial. Usually, distractions lead to longer durations of off-road glances, which significantly impair the driver's ability to quickly comprehend the traffic situation [45]. The typical distractions is engaging in NDRT, such as adjusting temperature controls or texting, not only increases mental and physical demands but also extends the time taken for drivers to respond to takeover requests in autonomous vehicles. This delay could potentially lead to safety risks during urgent maneuvering scenarios. It is recommended that the design of in-vehicle technologies consider the cognitive load they impose. For instance, the complexity of tasks that drivers are required to perform should be minimized to reduce the frequency of long off-road glances [46]. Enhanced driver monitoring systems could help in managing the transition of control in autonomous vehicles by assessing the driver's state of attention and readiness to takeover.

2.3.3 Cognitive States

The cognitive state refers to the mental conditions under which a person operates, encompassing aspects such as attention, awareness, and workload management. In the context of driving, particularly with advanced vehicular technologies, the cognitive state significantly influences how drivers interact with both manual and automated systems. Studies indicate that while automated driving systems can alleviate the cognitive load by managing routine tasks, they may also impair a driver's situational awareness and readiness to take control during emergencies, especially if the driver is engaged in NDRT. Furthermore, variations in cognitive load, whether due to overreliance on automation or distractions from non-driving activities, affect drivers' physiological responses, such as blink patterns and heart rate. These responses can be indicative of their engagement and stress levels. Thus, maintaining an optimal cognitive state is crucial for ensuring safety and effective interaction with vehicle automation.[44]

Choe et al. explored how drivers' attentional states are influenced by task demands and individual cognitive capacities. The study employed NDRT (visual and auditory) during driving simulations to manipulate drivers' attentional states and analyzed the effects on driving behaviors and eye movement patterns. Furthermore, individual cognitive capacities were directly measured using various tasks (e.g., simple reaction time, n-back tasks) to examine their interaction with task demands on driving performance. The study found that task demands and cognitive capacities significantly affected driving behaviors and physiological responses, indicating the potential for tailored driver assistance based on individual cognitive profiles [47]. Melnicuk et al. similarly investigated the effects of cognitive load on drivers' state and task performance during transitions from automated to manual driving using the "N-Back" task to manipulate cognitive load while measuring physiological responses and driving performance metrics. The results indicated that higher levels of cognitive load during automated driving impair the driver's ability to regain manual control effectively. Specifically, lateral control of the vehicle and stabilization times were adversely affected by higher cognitive loads.

Over the recent years, emerging research on the automated cognitive state classification systems for driver monitoring systems has been in focus, SAE International has identified various cognitive states of a driver in terms of assessing cognitive distractions, including concentration, mild distraction, and severe distraction. These states served as a basis for the automated detection system. The research aimed to explore the concept of providing varying alerts based on the driver's cognitive state [48]. To effectively detect driver cognitive states, the development of robust, real-time cognitive state monitoring systems is essential for enhancing road safety by providing timely alerts or interventions to distracted drivers, particularly in complex driving environments where the cognitive load is variable and unpredictable. Recent research achieved reliable detection accuracy by leveraging machine learning techniques, employing classifiers like Random Forest, Decision Trees, and Support Vector Machines to analyze data from eye-tracking, physiological signals, and vehicle kinematics [47], [49]. This indicates the need for research in dynamic HMI design to efficiently utilize the most data available to maximize road safety.

Chapter 3

Methodology and Procedures

3.1 Introduction

This study explores how the TOR alert design, customized to match the cognitive state of drivers, examines how TOR alerts affect driver responses depending on their presumed cognitive states in focused attention state or distraction state induced by secondary tasks while driving. The experiment is designed to help investigate the impact of varying alert levels on takeover actions, situational awareness, alert perception, and user satisfaction using different auditory and visual TOR modalities with distinct perceived levels of urgency. Experiments were carried out in a driving simulator environment with a tailored visual driver dashboard and alert configurations based on assigned secondary tasks to analyze the influence of dynamic TOR alert levels in automated driving systems concerning the driver's cognitive condition.

3.2 Methodological Assumptions

In this study, assumptions are made to simplify the classification of the driver's cognitive state into two categories: a concentrated state, where driving without a secondary task allows for heightened mental resources and focused attention on the driving task, and a distracted cognitive state, where NDRT is incorporated during the drive to increase mental and physical demands, the cognitive state and secondary task matching is based on the experiments in recent papers related to distracted cognitive states in drivers. [10], [48], [50]. However, this study does not account for differences in individual workload levels associated with the secondary task. Additionally, it is assumed that alerts will always be issued when a road hazard is encountered, without false positives. Furthermore, it is presumed that all participants, regardless of their driving experience, are equally equipped to complete the driving task in a highway scenario using lane tracing assist (autopilot feature).

3.3 Experimental Research Design

This study, approved by the Research Ethics Board at the University of Waterloo (REB 45254), employs an immersive driving simulator to replicate close-to-reality driving scenarios. Using a between-subject experimental design, the study investigates the differences between dynamic and single alerts across four diverse highway driving scenarios with counterbalanced sequences. Forty participants were randomly assigned to one of two experimental conditions: a dynamic alert group or a single (static) alert group.

3.3.1 Participants

Participants were recruited from University of Waterloo students and alumni, and publicly through the university's recruitment website. Inclusion criteria included having at least an Ontario G2 driver's license or equivalent with at least one year of driving experience and being free from any medical conditions that could affect driving ability. Participants were screened for eligibility prior to the experiment. Forty-one participants with an age range of 18 to 54 years with an average of five years of driving experience completed the experiments.

3.3.2 Apparatus

Driving Simulator

The experiment was conducted using a high-fidelity driving simulator at the Autonomous Vehicle Research and Intelligence Laboratory (AVRIL) of the University of Waterloo. A customized driver dashboard was designed to enhance the driving visualizations for the experiment. Detailed information about the driving simulator setup can be found in Appendix A.1.

- **Driving Instruments:** VI-Grade's STATIC driving simulator utilizes the cockpit of a 2018 Chevrolet Traverse, equipped with an active seat and active 5-point seat belts that offer haptic feedback determined by the vehicle dynamics calculations. Steering wheel, throttle, and brake pedals from the original equipment manufacturer (OEM) that are similar to a real vehicle.

- **Simulation Displays:** A 278-degree screen is used to create an immersive driving environment by incorporating screens that mimic the driver's dashboard, side mirrors, and rear-view mirrors.

- **Sound Interface and System:** A 5.1 surround sound set-up that uses speakers within a vehicle to deliver environmental, vehicle, and alert sounds.

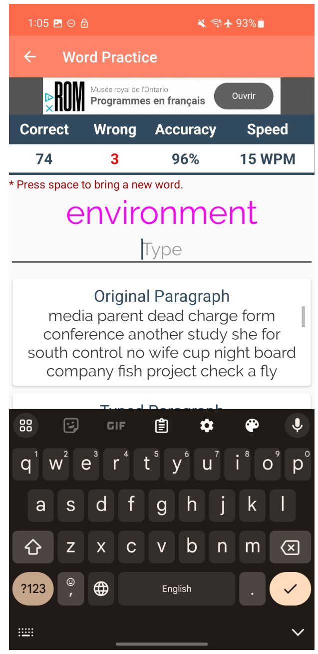
- **Driver Assist Features:** Level-2 autonomous driving system including ACC feature that can be activated and adjusted at any time with a speed greater than 30 km/h with the ability to follow the leading vehicle's speed, the ACC target speed is adjustable using the plus (+) or minus (-) buttons on the gear knob. The AP that offers the lane tracing feature can be activated while the ACC is active; this feature keeps the vehicle in the center of any lane on the highway at any speed; however, it does not mitigate any collision or make a lane change.

- **Driving Controls and Feedback:** Steering wheel with angle mode providing forces to return the steering wheel to a zero angle after rotation. Buttons on the steering wheel to control the simulation and toggle AP mode with physical buttons on the gear knob to change gear and adjust the target speed for cruise control. The indicator lever can be used to signal and disable AP; however, only visual feedback is displayed on the dashboard without the indicator sound. All ADS features are deactivated once the brake pedal is used. Important controls that mismatch from the real-world vehicle include the input of steering wheel angle and feedback, which is disabled in AP mode, and ACC controls, which are located at the gear knob, due to the limitation of the hardware system.

Physiological Measurements

- **Eye Tracker:** The Ergoneers eye tracker was used on all participants with or without eyeglasses, the device recorded first-person view with videos (with the resolution of 1920 x 1080 pixels at 30 frames per second) of eye movement for pupil detection to determine eye movement (with binocular eye cameras tracking resolution of 648 x 488 pixels at 60 Hz). The device weighs 52 g with connection wires routing behind the

(a) Smartphone device used with eye-tracking marker.



(b) Screenshot of the typing task with example paragraph.

Figure 3.1: Typing task setup of with smartphone with eye tracking marker (left), and the screenshot of the typing task with touchscreen keyboard (right).

driver seat's headrest. Area of Interest (AOI) management using automated marker detection within the Ergoneers D-Lab software.

- **Heart Rate (HR) Monitor:** The Empatica E4 Heart Rate Monitor wristband was used to collect heart rate at 1 Hz on the participant's wrist of the non-dominant hand. The E4 uses photoplethysmography (PPG) to detect Blood Volume Pulse (BVP) and derive HR and Inter-Beat Interval (IBI) time series. The device outputs raw BVP data sampled at 64 Hz and utilizes this data to automatically generate HR and IBI measurements. Despite its capabilities, the E4 can experience significant missing IBI data due to movement or baseline shifts in the BVP signal, which may affect short-term heart rate variability metrics [51], [52].

Device for NDRT

An Android phone with a 6.1-inch touchscreen display was used as a NDRT during the autonomous driving simulation. Participants engaged with the Typing Speed Test application in word practice mode, which randomly selected words for the user to type, as shown in

Figure 3.1. The application displayed a single word at a time, which the participants were required to type as quickly and accurately as possible on the phone's virtual keyboard. This task was designed to simulate a realistic distraction, thereby creating a distracted cognitive state similar to real-world driving conditions where drivers might engage with their phones. Engaging in this secondary task aimed to replicate the cognitive load and attentional shift experienced during actual driving distractions. More information related to the device used for the typing task, including an example paragraph of the typing task, can be found in Section A.1.3.

3.3.3 Driving Scenarios

Participants were allowed 10-15 minutes to familiarize themselves with the driving and control maneuvers of the simulator on the same highway condition as the actual scenarios with guidelines on AP controls and ADS limitations. Each participant performed four driving scenarios in the simulation in highway conditions with the target speed between 110-120 km/h, each consisted of one TOR event issued for the participant to regain manual control with constant TOR locations based on 100% alert accuracy upon road hazards. Two sets of TOR alert settings were assigned conditionally according to the subject group, and a secondary task was assigned that varies mild or strong alert settings.

Participants were given control over the start button, the ACC, and the AP and were advised to activate the AP to the greatest extent possible based on their evaluations of the safety and limitations of the system. Approximately halfway through the third and fourth scenarios, participants were asked to perform a secondary task while maintaining the priority of the driving task and be ready for any TOR. After driving for approximately 5-7 minutes, depending on the cruising speed before the TOR event, participants were alerted by TOR to resume manual driving and then asked to stop the simulation shortly after mitigating the hazard.

All driving situations also included background noise from the simulated engine and the surrounding environment with a maximum volume of 70 dB. After each scenario, participants were asked to complete post-scenario questionnaires related to alert effectiveness and situation awareness. Each driving scenario lasts approximately 5-8 minutes with conditions shown in Figure A.2 with details as follows:

Scenario 1 (S1): Road closure due to construction on two left lanes

- **Hazard Condition:** Static Distance with warning signs every 50m starting at 300m before the construction zone on a straight road, driving in good visibility condition.
- **Ego Vehicle Driving Lane:** 2nd lane from the left (left lane for passing)
- **TOR Alert location:** 300m before crash
- **TTC threshold at 110-120 km/h:** 8-9 seconds
- **Condition to Mitigate Hazard:** Stop before crash and/or steer to the 3rd lane.
- **Scenario Order:** 1st or 4th scenario

Scenario 2 (S2): Emergency stop for an animal on the road in low visibility

- **Hazard Condition:** Static distance for to stop due to animal (moose) moving in perpendicular to the straight road, visible at 350m, driving in fog condition with 400m visibility. The leading vehicle starts full braking 200m before the crash (then fully stops at 20m before the crash).
- **Ego Vehicle Driving Lane:** 2nd lane from the left (left lane for passing)
- **TOR Alert location:** 250m before crash into the leading car, 300m before hitting the animal
- **TTC threshold at 110-120 km/h:** 8-9 seconds
- **Condition to Mitigate Hazard:** Full stop before crash
- **Scenario Order:** 2nd or 3rd scenario

Scenario 3 (S3): Slower vehicle suddenly changing lanes from the right, emerging from the blind spot

- **Hazard Condition:** Large truck driving on the 3rd lane at 90 km/h on a straight road, cutting off to the ego vehicle's lane at 40m in front of the ego vehicle with left signaling, driving in good visibility condition.
- **Ego Vehicle Driving Lane:** 2nd lane from the left (left lane for passing)

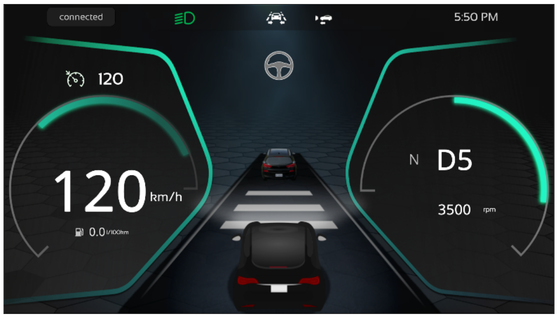
- **TOR Alert location:** 40m behind the merging truck (longitudinal distance)
- **TTC threshold at 110-120 km/h:** 6-8 seconds
- **Condition to Mitigate Hazard:** Slow down to less than 90 km/h or steer to the 1st (left) lane
- **Scenario Order:** 2nd or 3rd scenario

Scenario 4 (S4): Pulled over vehicles merging onto the right lane behind blind curved road

- **Hazard Condition:** Pulled over and cop vehicles merging with a visible light signal at 260m started moving shortly before the alert, moving at 5-10 km/h and accelerating, driving in good visibility condition.
- **Ego Vehicle Driving Lane:** 4th lane (right lane, allowing to pass on the right).
- **TOR Alert location:** 250m before merging location
- **TTC threshold at 110-120 km/h:** 7-8 seconds
- **Condition to Mitigate Hazard:** Stop before crash and/or steer to the 3rd lane
- **Scenario Order:** 1st or 4th scenario

3.3.4 Driver Dashboard Designs

The instrumental cluster was designed using the VI-Grade ADAS dashboard (Figure 3.3a) as a reference which incorporated the icon components that were derived from ADAS icons in a Chevrolet vehicle [53]. The dashboard includes a digital speedometer with an ACC status indicator and a target speed on the left side of the dashboard, a current gear on the right, a vehicle-road display in the middle for visual lane departure warnings, and an angle-tracking steering wheel icon. Utility lights include turn signals, autopilot status, and active automatic emergency brake (AEB) features located at the top. The use of color in normal driving mode utilizes light color (HEX #24FACB) highlights and white icons to indicate inactive automation, for instance, a white steering wheel icon to indicate manual steering mode as shown in Figure 3.2a similar to the existing design. Active automation features are presented with corresponding icons changed to green color (HEX #32D74B) with a



(a) Dashboard in manual mode

(b) Dashboard in Autopilot (AP) mode

Figure 3.2: Example of visual dashboard design

visual message to inform about the activated feature, furthermore, the overall dashboard highlight color changes to green color only when AP feature is activated. The green color was designed based on GM's Super Cruise[54] active autopilot status as shown in Figure 3.3b.

3.3.5 TOR Alert Designs

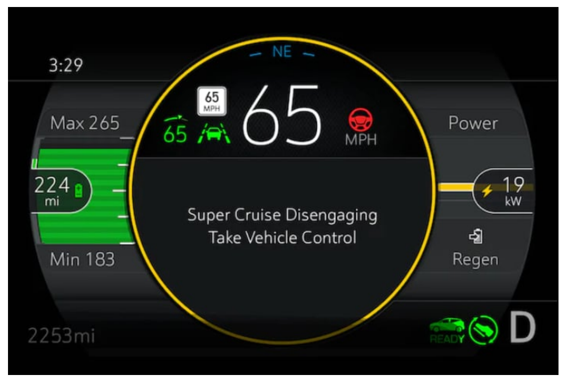
Both TOR alert designs consist of visual messages appearing on the dashboard and a nonverbal sound alert broadcast through surround speakers in the vehicle. The visual message components were reused from the existing VI-Dashboard component that imitates the GM dashboard warning message design, it was modified to display a warning message for 8 seconds using bold and red fonts for the action message along with a white message for short alert reasons. In addition, the red hands icon (HEX #FF5F5F) flashes over the steering wheel icon on the steering wheel during TOR with a frequency of 4 Hz. However, the dashboard color remains green to indicate active autopilot until the driver takes over to resume manual control. The same designs were used for the single and dynamic alert groups of participants and did not include graphic illustrations for specific TOR scenarios on the dashboard. The comparison between mild and strong alert can be found in 3.1

Mild Alert Design

Mild alert visual messages include white text above stating "Potential takeover required" with a red action message stating "STAY VIGILANT!" as an indirect action message aimed at getting the driver's attention to the surroundings. A short low-to-high-pitched auditory alert plays for 0.5 seconds when the visual message starts to appear on the dashboard and

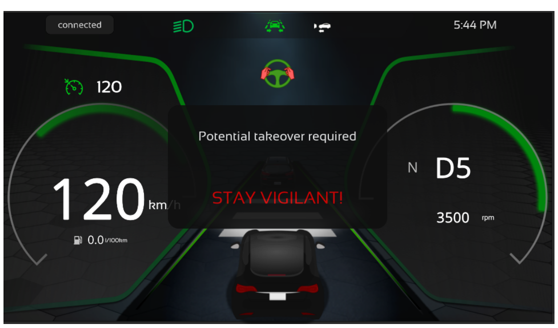
(a) VI-Grade's original ADAS dashboard design (with lane keep warning) [53]

(b) GM's active Super Cruise design [54]

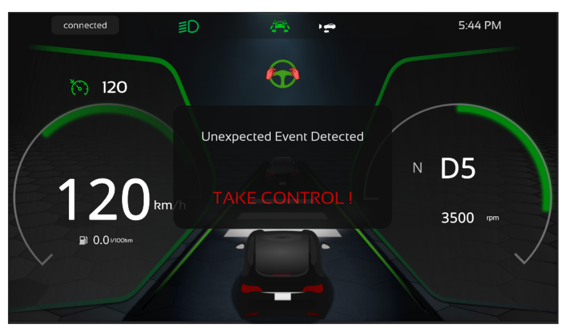


(c) GM's TOR alert design [54]

Figure 3.3: Referenced visual dashboard design.



(a) Mild alert message

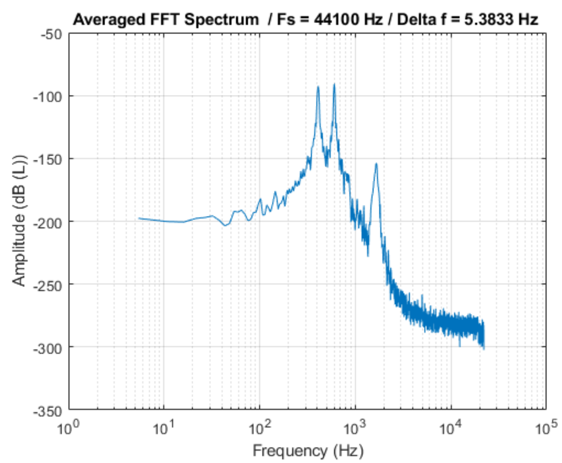


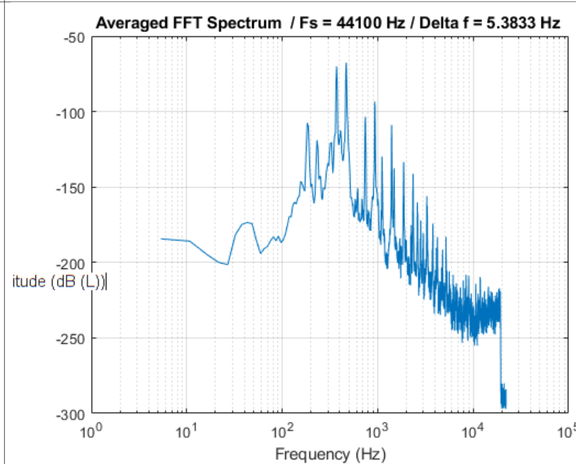
(b) Strong alert message

Figure 3.4: Example pop-up alert designs on dashboard.

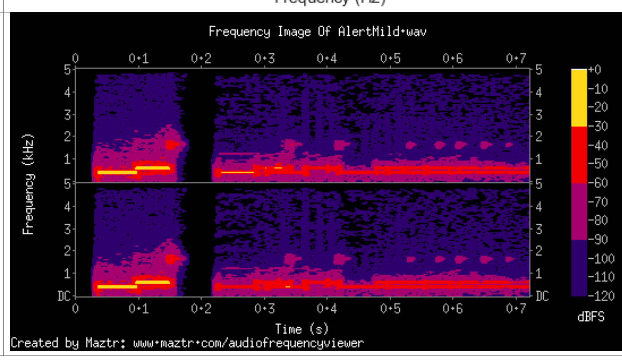
Attribute	Mild Alert	Strong Alert
Visual Message	"Potential takeover required" (white) "STAY VIGILANT!" (red)	"Unexpected Event Detected" (white) "TAKE CONTROL!" (red)
Sound Profile	Short low-to-high-pitched auditory alert	Three cycles of ascending pitch tones
Sound Duration	0.5 seconds	1.2 seconds
Peak Frequencies (Hz)	400 Hz, 600 Hz	380 Hz, 480 Hz, 750 Hz, 950 Hz
Spectrogram Intensity (dB)	Maximum alert volume at 85 dB	Volume capped at 90 dB
Sound Pulses Duration (s)	0.2 seconds	0.4 seconds each
Sound Frequency Analysis	High intensity in brief pulses at 400 Hz and 600 Hz within 0.2 seconds	Increased intensity and extended duration across various frequencies
Visual Icon Animation	Red hands icon (HEX #FF5F5F) flashes over the steering wheel icon on the dashboard at 4 Hz	Red hands icon (HEX #FF5F5F) flashes over the steering wheel icon on the dashboard at 4 Hz
Dashboard Background	Stay green to indicate on-going autopilot until takeover	Stay green to indicate on-going autopilot until takeover
Visual Message Intention	To gain the driver's attention to the surroundings without immediate takeover required	Advises the driver to take control immediately due to unexpected event

Sound Analysis Illustration





Spectrogram Illustration



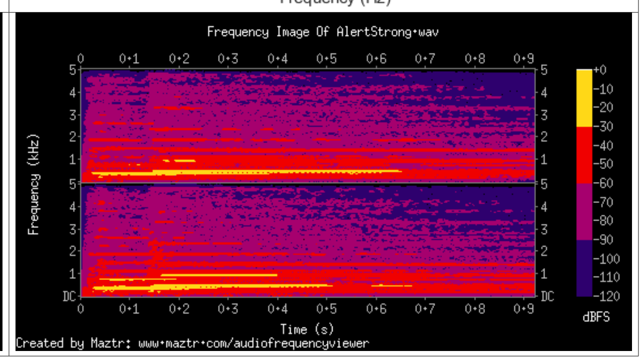


Table 3.1: Comparison of Mild vs Strong Alerts Based on Frequency Analysis and Design Attributes

shows for 8 seconds or until the driver takes over. The Fast Fourier Transform (FFT) spectrum and spectrogram are illustrated in Figure **??** and **??** respectively. The spectrogram delineates that the mild alert sound has high intensity in brief pulses at frequencies of 400 Hz and 600 Hz, occurring within a duration of 0.2 seconds with the maximum alert volume at 85 dB.

Strong Alert Design

Strong Alert visual messages include white text stating "Unexpected Event Detected" with a red action message stating "TAKE CONTROL!" aiming to advise the driver to takeover immediately. Three cycles of ascending pitch tones, each enduring 0.4 seconds, featuring prominent frequencies at 380, 480, 750, and 950 Hz, continue for a total of 1.2 seconds at a volume capped at 90 dB. According to the spectrogram in Figure **??**, the sound spectrum exhibits increased intensity and extended duration across various frequencies, suggesting a more robust sound profile compared to the mild alert.

Alert Perception Test

An alert perception test was created to validate the urgency perception of the alert in a controlled environment without traffic using an identical simulation environment as the scenarios. Following the fourth scenario, thirty participants were asked to take the alert perception test. Fifteen participants in the dynamic alert group were presented with both types of alert in a random sequence, while participants in the single alert group were first shown the exposed type and followed by the other alert type. Short questions related to perceived urgency and annoyance were asked immediately after each alert was presented.

3.3.6 Independent Variable(s)

- **TOR Alert Intensity Design:** This study utilizes two alert designs, categorized as mild and strong. The 'single alert' group receives a consistent alert type, either mild or strong, regardless of their cognitive state. The 'dynamic alert' group receives alerts tailored to their cognitive state: mild alerts are issued when the driver is presumed to be fully attentive (without a secondary task), and strong alerts are issued when the driver is presumed to be distracted (engaged in a secondary task). This design tests the hypothesis that dynamic alerts, which adjust based on the driver's cognitive state, enhance driving experience and takeover performance.

- **Presumed Cognitive State:**Drivers are classified into two cognitive states: concentrated and distracted. Concentrated states are assumed in scenarios without a secondary task, allowing the driver to focus fully on driving. Distracted states are assumed when the driver engages in a secondary task, potentially impairing their focus on driving. This categorization tests the hypothesis that drivers' cognitive states affect their response to different alert intensities.

3.3.7 Dependent Variable(s)

- **Takeover Performance:** This is assessed by measuring the driver's reaction times and the quality of maneuvers during takeover incidents, with metrics such as acceleration and distance from potential hazards recorded. These measures provide empirical evidence to support or refute the hypothesis concerning the effectiveness of dynamic versus single alerts in varying cognitive states.
- **Alert Perception and Satisfaction:** Drivers' satisfaction with and perception of the alert system are evaluated through post-experiment questionnaires. This assessment helped determine whether the alert intensity was perceived as too strong, appropriate, or insufficient in different cognitive states to directly test the hypothesis on alert perception in different cognitive states.

3.3.8 Controlled Variable(s)

- **Driving Simulator Configuration:** Ensure consistent simulator settings across all experimental sessions, including visual design and resolution, simulator hardware including simulation controls, and software version of ACC and AP.
- **Driving Scenarios and TOR Timing:** Four driving scenarios were exposed to all participants on the same four-lane highway condition, the order of the scenarios being counterbalanced between scenarios 1 and 4, and scenarios 2 and 3. Variations in road conditions, traffic density, and environmental factors are controlled to be similar for each scenario. TOR alerts were set to be issued based on the location that was timed to be similar across all scenarios, whereas the alert trigger locations were placed at the relevant location to the incident.
- **TOR Modalities:** The two multimodal TOR alert designs were presented using auditory and visual modalities as recommended in related studies to reduce reaction time with lower alert startle [27], [29]. Auditory alerts were delivered using the

vehicle's built-in sound system, and alert messages were visually displayed on the dashboard in front of the driver. Alerts involving both auditory and visual components were exclusively used for TOR alerts. In addition, the dashboard displayed other visual messages, such as activating or deactivating ACC or AP features. To reduce the cognitive load associated with signal detection, each TOR alert was designed to always warn the driver of the predicted incident in each scenario.

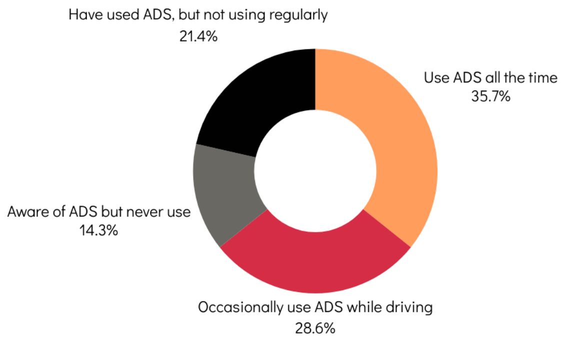
- **Driving Lane and ACC Target Speed:** A consistent lane use and speed control minimize variations in driving behavior not related to cognitive state or alert design. Participants were directed to use a designated lane with an ACC target speed ranging from 110 to 120 km/h to control similar driving duration in each scenario and the time to collision threshold.

- **Non-driving-related Secondary Task and Workload Level:** The introduction of a typing task as a controlled secondary task simulates increased cognitive load, facilitating the examination of its effect on the efficiency of different alert intensities. Participants were asked to perform a typing task on the touch screen phone as a secondary task approximately halfway through each scenario for the third and fourth scenarios. The complexity of the tasks will be similar for all experiments, which is to type as the appearing word and continue to the next word using the spacebar. By performing a secondary task while driving, the driver is assumed to have a higher workload with a lower vigilance level on the road than merely performing a driving task due to increasing off-road glance and occupied physical workload [1], [45], [55].

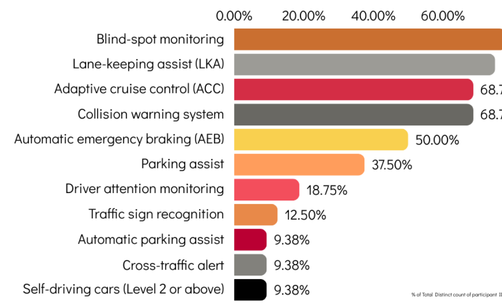
3.3.9 Data Collection

- **Demographics and screening:** Forty-one participants aged between 18-54 years old with at least an Ontario G2 driver's license or equivalent and one year of driving experience. The age groups of the participants consisted of 72.5% of 18-24 years (avg. 4.1 years of driving experience), 25% of 24-34 years (avg. 7.9 years of driving experience, and 2.5% of 45-54 years (avg. 38 years of driving experience).

- **Pre-study questionnaires:** Participants were asked to complete questions on their experience with ADS, a self-assessment questionnaire on driving behavior, and the Short Motion Sickness Susceptibility Questionnaire (Short-MSSQ) before the experiment. Short-MSSQ was used to inform registered participants who are at risk of experiencing simulator sickness. Most participants were familiar with ADS features



(a) Participants' experience with ADS features.



(b) Level of ADS usage in daily lives.

Figure 3.5: Driver assistance features familiarity of the participants.

where 35.7% of them use ADS all the time, with over 68% having used ACC and lane-keep assist (LKA) features, as shown in Figure 3.5.

- **Quantitative measures:** Data obtained from the driving simulation on speed, accelerations, driving maneuvers (throttle, brake, steering, indicators, and AP control), and calculated response times in relation to the alert trigger were gathered in CSV format for each specific scenario at 1000 Hz sampling rate.

- **Eye Tracker:** Participants' eye movements with video recordings in first-person view were collected for eye tracking to calculate gaze movements during driving in each scenario. The eye tracker was calibrated at the beginning of the experiment for each participant.

- **Heart Rate and Electrodermal Activities:** Participants were asked to wear the heart rate monitor device on their non-dominant wrist to minimize interference with the secondary task throughout the study. The HR data was collected at a 64 Hz sampling rate.

- **Post-Scenario Questionnaires:** After each scenario, participants completed brief questionnaires to assess different aspects of the TOR alerts. These evaluations included attention-capturing, timing relevance of TOR alerts, assistance in understanding the takeover situation, alert comprehension, confidence in takeover abilities, level of alert-induced annoyance, and overall satisfaction with the alert design, all rated on a 5-point Likert scale. The questionnaire mentioned is located in Appendix Section A.2.

Chapter 4

Results and Discussion

This chapter delves into the analysis and discussions regarding the outcomes of the experiments using simulation data to assess takeover performance, eye tracking data to gauge the duration of glances at various stages of the takeover process, and qualitative analysis of alert effectiveness and perceptions from post-scenario questionnaires. The results of certain participants in the experiment were omitted due to no alerts because of the system or incomplete data during the scenario. Additionally, all heart rate data were deemed unreliable because they were inconsistent and undetectable when there were movements on the wrist to which the device was attached while driving. It is important to note that data omission due to pre-takeover actions may cause missing data, particularly in scenario 1 where the TOR was triggered by the presence of a construction zone. This preemptive takeover by participants resulted in a lack of relevant data for analysis in some instances.

4.1 Experiment Results and Data Analysis

4.1.1 Analysis of Quantitative Results

To analyze the simulation data, driver reactions were extracted using a Simulink model that derives driver response following the TOR as braking, steering, and using indicators in each scenario. Simulated driving measures related to speed, acceleration, and AP toggle control were included in the analysis to determine takeover performance. Data from each scenario were correlated with the assigned group and secondary task for comparison. As shown in Figure 4.1, the mean cruising velocities while under AP control prior to TOR were 116.15 km/h (SD = 4.32) in Scenario 1, 113.87 km/h (SD = 7.45) in Scenario 2, 116.92 km/h (SD = 3.27) in Scenario 3, and 115.45 km/h (SD = 4.60).

Participants were allowed to control the AP based on their judgment of safety, resulting in deactivation of the AP preceding TOR incidents in 22.22% of 144 instances, as shown in Figure 4.2. The timing of the takeover depended on the nature of the driving scenario. The rate of takeover before TOR was higher in scenarios with static incident distance, leading to 60.61% (20 of 33 sessions) of drivers deactivating the AP before the alert in Scenario 1, where the construction zone was visible more than 500 m ahead of the alert at 300 m. The rate of takeover before the onset of TOR was lower in situations where incidents were more difficult to detect: 15.79% (6 of 38 sessions), 13.89% (5 of 36 sessions), and 2.70% (1 of 37 sessions) in Scenarios 2, 3, and 4, respectively. Some drivers were observed to take control before receiving an alert due to the actions of surrounding AI vehicles or driving conditions unrelated to the designed incidents. To conduct a comprehensive takeover analysis within the scope of this study, it was imperative to exclude certain data from the driving sessions. Specifically, this entailed removing instances wherein drivers transitioned to manual control before the initiation of the TOR, as well as any incomplete data resulting from data collection anomalies, such as instances where a TOR was not issued due to system errors. It is important to note that data omission due to pre-takeover actions may cause missing data, particularly in Scenario 1 where the TOR was triggered by the presence of a construction zone.

Takeover Reactions Times

The measurement of driver reaction times was calculated based on the time stamp of the issued TOR across three distinct types of reactions needed for takeover: braking, turn signaling, and steering, whereas the first reaction indicates the takeover time (fastest response to takeover). Table 4.1 displays the average takeover times for each scenario, covering 41 participants and omitting instances where the driver regained manual control prior to the TOR. The overall average takeover time was 2.85 seconds (SD = 1.88) in the Construction Zone Scenario (S1), 4.32 seconds (SD = 1.28) in the Animal on the Road Scenario (S2), 2.04 seconds (SD = 1.10) in the Blind Spot Lane Change Scenario (S3), and 3.59 seconds (SD = 1.70) in the Pulled Over Vehicle Merging Scenario (S4). When contrasting the mean response durations across scenarios with and without NDRT, it was observed that, in certain instances, prolonged average takeover times were evident in the absence of secondary tasks. Notably, in the Construction Zone Scenario (S1), average reaction times without NDRT were 5.29 seconds (compared to 2.35 seconds when engaging in typing activities) during mild alerts, while in the Pulled Over Vehicle Merging Scenario

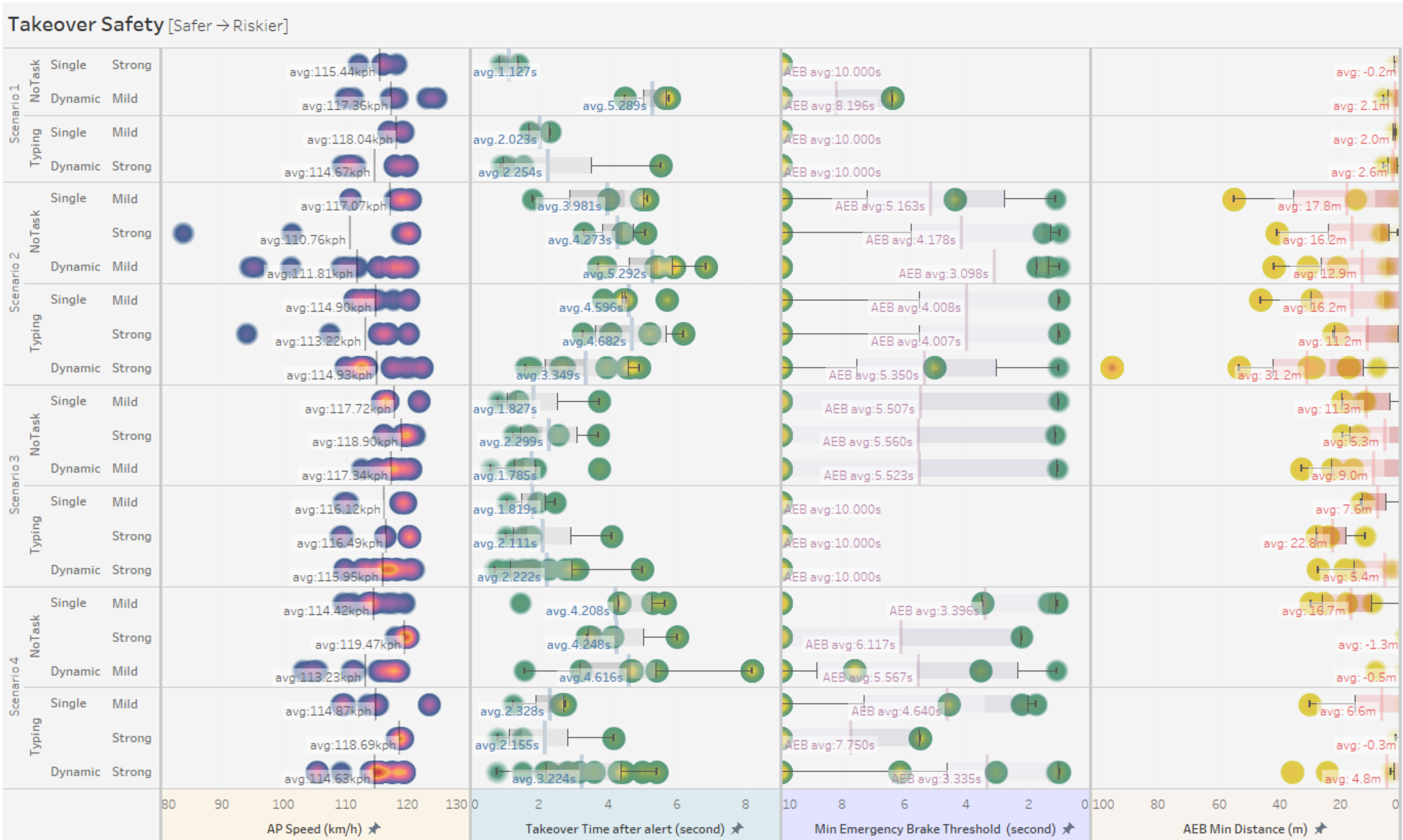


Figure 4.1: Takeover Safety Analysis for each scenario, including the following metrics: average AP speed (km/h), takeover time after alert (seconds), minimum emergency brake threshold time (seconds), and minimum distance to activation of the automatic emergency braking (AEB) system (meters). The data is segmented by alert type (mild and strong) and the presence of secondary tasks (no task and typing). This comprehensive view allows for a detailed comparison of takeover performance and safety across different scenarios and conditions.

Alert Type	Group	Measure	All Scenarios Average	Scenario 1 (n: 12) Average	No Task	Typing	Scenario 2 (n: 31) Average	No Task	Typing	Scenario 3 (n: 31) Average	No Task	Typing	Scenario 4 (n: 34) Average	No Task	Typing
Mild Alert	Dynamic	Average	4.15	5.29	5.29		5.29	5.29		1.79	1.79		4.62	4.62	
		5th %iles	1.31	4.47	4.47		3.70	3.70		0.60	0.60		1.58	1.58	
		95th %iles	6.85	5.77	5.77		6.85	6.85		3.74	3.74		8.16	8.16	
		SD	1.98	0.58	0.58		1.04	1.04		0.96	0.96		2.02	2.02	
	Single	Average	3.28	2.35		2.35	4.32	3.98	4.60	1.82	1.83	1.82	3.63	4.50	2.33
		5th %iles	1.06	1.72		1.72	1.81	1.81	3.87	0.81	0.81	1.06	1.24	1.44	1.24
		95th %iles	5.70	2.99		2.99	5.70	5.13	5.70	3.74	3.74	2.45	5.99	5.99	2.75
		SD	1.57	0.52		0.52	1.04	1.32	0.60	0.93	1.13	0.58	1.63	1.50	0.63
	All Groups	Average	3.65	3.82	5.29	2.35	4.75	4.82	4.60	1.81	1.80	1.82	4.00	4.56	2.33
		5th %iles	1.06	1.72	4.47	1.72	1.81	1.81	3.87	0.60	0.60	1.06	1.24	1.44	1.24
		95th %iles	5.99	5.77	5.77	2.99	6.85	6.85	5.70	3.74	3.74	2.45	8.16	8.16	2.75
		SD	1.81	1.57	0.58	0.52	1.14	1.31	0.60	0.95	1.03	0.58	1.85	1.78	0.63
Strong Alert	Dynamic	Average	2.83	2.25		2.25	3.35		3.35	2.22		2.22	3.22		3.22
		5th %iles	0.78	0.89		0.89	1.60		1.60	0.70		0.70	0.78		0.78
		95th %iles	5.41	5.52		5.52	4.91		4.91	4.99		4.99	5.41		5.41
		SD	1.48	1.90		1.90	1.25		1.25	1.23		1.23	1.41		1.41
	Single	Average	3.05	1.13	1.13		4.44	4.27	4.58	2.21	2.30	2.11	3.21	4.08	2.16
		5th %iles	0.86	0.86	0.86		3.27	3.31	3.27	1.03	1.26	1.03	0.79	3.39	0.79
		95th %iles	5.99	1.40	1.40		6.19	5.09	6.19	4.10	3.70	4.10	5.99	5.99	4.15
		SD	1.57	0.27	0.27		0.95	0.74	1.07	1.07	0.93	1.18	1.53	0.93	1.44
	All Groups	Average	2.92	1.88	1.13	2.25	3.85	4.27	3.74	2.22	2.30	2.19	3.22	4.08	3.00
		5th %iles	0.79	0.86	0.86	0.89	1.60	3.31	1.60	0.70	1.26	0.70	0.78	3.39	0.78
		95th %iles	5.41	5.52	1.40	5.52	6.19	5.09	6.19	4.99	3.70	4.99	5.41	5.99	5.41
		SD	1.52	1.65	0.27	1.90	1.25	0.74	1.33	1.16	0.93	1.22	1.45	0.93	1.48
Overall	All Groups	Average	3.27	2.85	3.63	2.29	4.32	4.70	4.00	2.04	1.94	2.13	3.59	4.45	2.85
		5th %iles	0.86	0.86	0.86	0.89	1.77	1.81	1.60	0.70	0.60	0.70	0.79	1.44	0.78
		95th %iles	5.90	5.77	5.77	5.52	6.19	6.85	5.70	4.10	3.74	4.99	5.99	8.16	5.41
		SD	1.71	1.88	2.10	1.48	1.28	1.23	1.22	1.10	1.03	1.14	1.70	1.63	1.37

Table 4.1: Average takeover time in seconds after TOR in each scenario separated by alert type (Mild and Strong) with 5th and 95th percentiles. Color coding indicates longer duration in red and shorter duration in green, based on the median values of average reaction time. Note: Missing data in single groups was caused by omitted data from takeover before TOR.

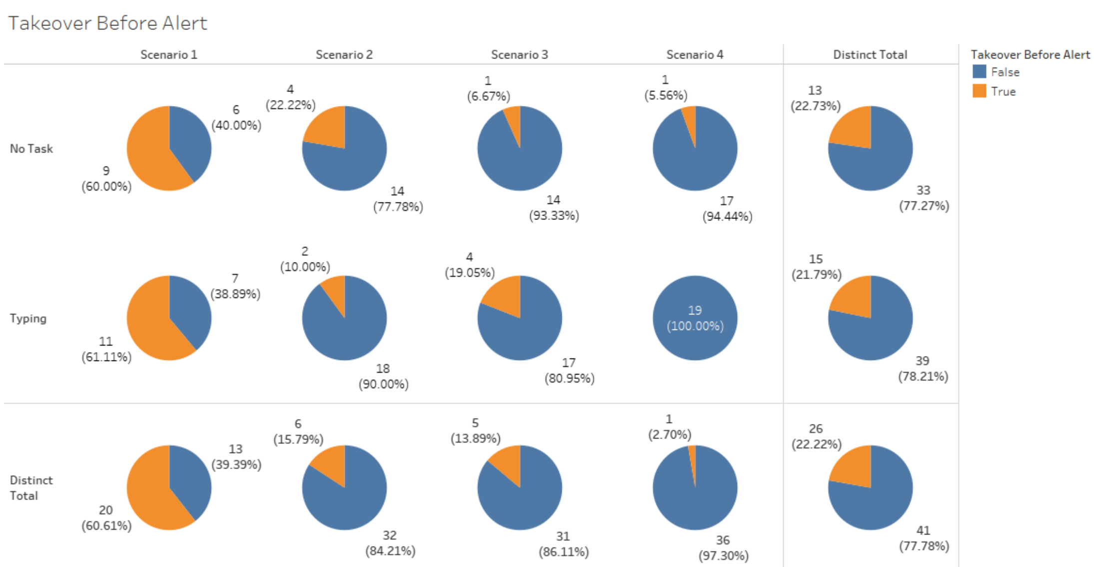


Figure 4.2: This pie chart illustrates the percentage of total cases with a distinct count of participants who decided to takeover before the takeover request (TOR) (orange) or takeover after TOR (blue). The data is segmented by scenario and task condition, providing a comprehensive view of the participants' takeover decisions across different contexts.

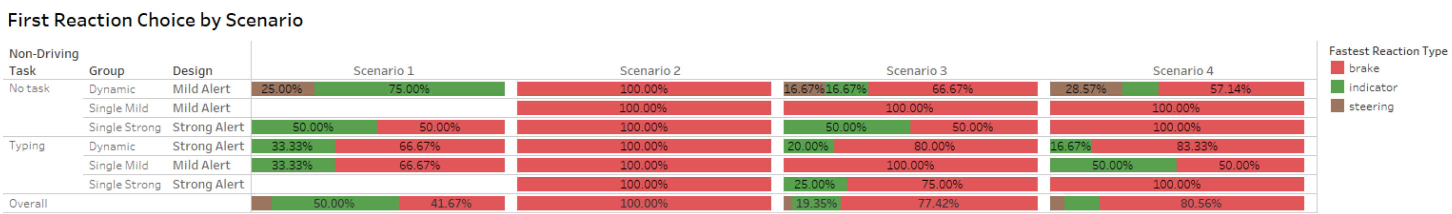


Figure 4.3: Percentage distribution of first takeover maneuvers by scenario and secondary task, categorized by reaction types: braking (red), turn signaling (green), and steering (brown). The figure compares dynamic and single alert groups across different scenarios (S1: Construction Zone, S2: Animal on the Road, S3: Blind Spot Lane Change, S4: Pulled Over Vehicle Merging) and highlights the prevalence of each reaction type under various alert conditions and secondary task engagements (no task vs. typing). This detailed breakdown provides insights into how different alert designs and secondary tasks influence the initial driver response during takeover requests.

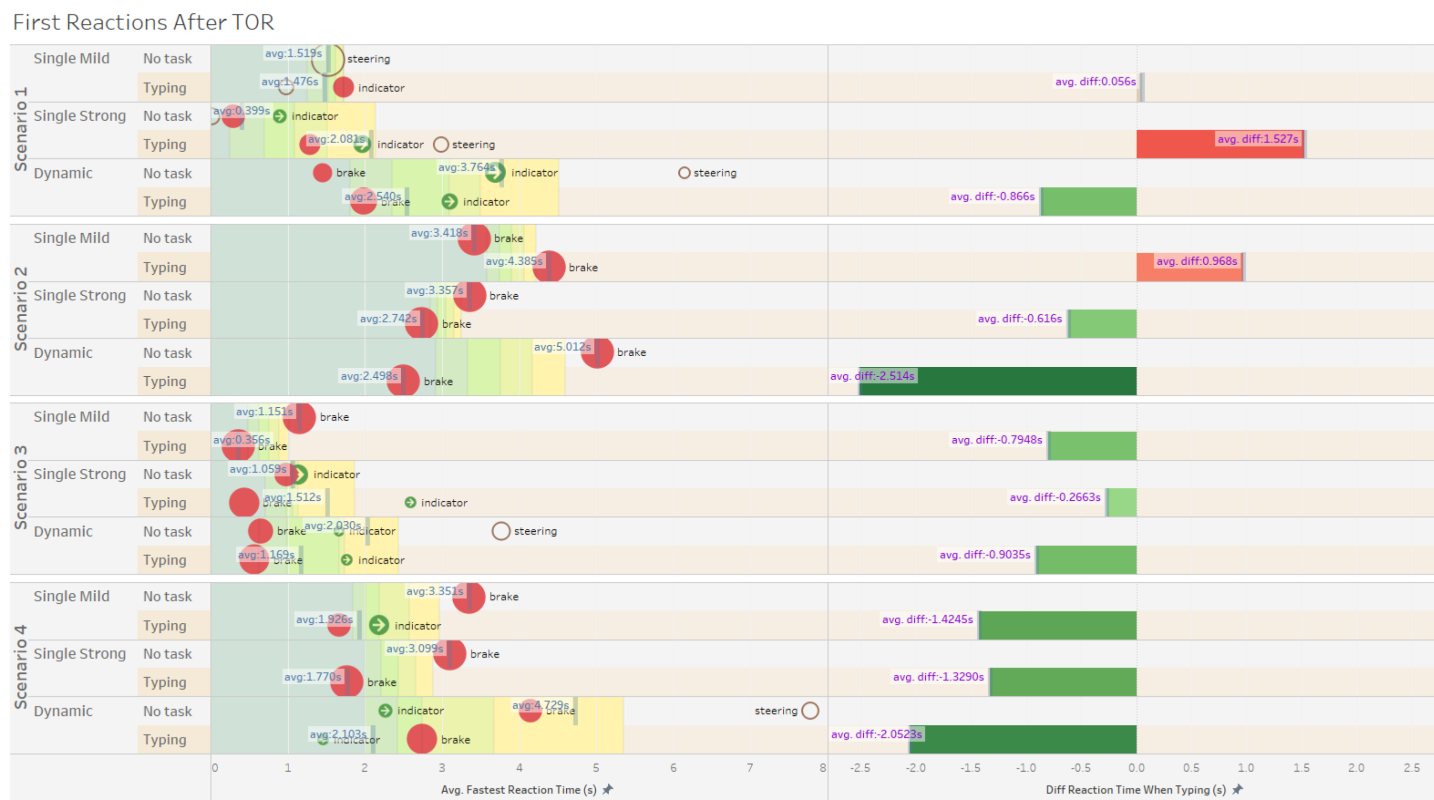


Figure 4.4: Comparison of average response times for different first reactions (braking in red, steering in brown, and using indicators in green) following the takeover request (TOR), measured in seconds (left). The figure also illustrates the difference in reaction times during typing tasks compared to scenarios without secondary tasks, with differences quantified in seconds (right). Each scenario (1 to 4) is displayed with distinctions between single mild, single strong, and dynamic alerts, highlighting how various alert types and secondary task engagements influence driver response times. Note the average reaction time differences are indicated with green for faster response times during typing and red for slower response times during typing.

(S4), the overall average reaction time was 4.45 seconds (compared to 2.85 seconds when typing). However, in the Construction Zone Scenario (S1) under heightened alertness, it was observed that the average response time shortened when participants were not engaged in a secondary task, with a recorded duration of 1.13 seconds (SD = 0.27) compared to 2.25 seconds (SD = 1.90) during typing. On average, strong alerts were found to induce faster reactions in the Construction Zone Scenario (S1; M = 1.88 seconds, SD = 1.65), the Animal on the Road Scenario (S2; M = 3.85 seconds, SD = 1.25), and the Pulled Over Vehicle Merging Scenario (S4; M = 3.22 seconds, SD = 1.45) when compared to mild alerts in the Construction Zone Scenario (S1; M = 3.82 seconds, SD = 1.57), the Animal on the Road Scenario (S2; M = 4.75 seconds, SD = 1.14), and the Pulled Over Vehicle Merging Scenario (S4; M = 4.00 seconds, SD = 1.85).

Comparing the average takeover times of each participant with and without NDRT, significantly faster takeover responses were observed in all scenarios with NDRT (0.316 seconds in the Construction Zone Scenario (S1), 1.177 seconds in the Animal on the Road Scenario (S2), 0.672 seconds in the Blind Spot Lane Change Scenario (S3), and 1.559 seconds in the Pulled Over Vehicle Merging Scenario (S4)), t = -2.94, p = .0058. Furthermore, when analyzing among the participant groups, the dynamic group had the strongest significance in response time reduction with an average of 1.749 seconds faster, t = -2.82, p = .0117, while there was no significant difference among participants in the single group: an average of 0.380 seconds faster, t = -1.33, p = .2195 in the mild alert group, and an average of 0.276 seconds faster, t = -0.29, p = .7810 in the strong alert group.

The analysis of reaction times during takeover is illustrated in Figure 4.4 and summarized in Table 4.1. Notably, drivers responded an average of 1.179 seconds faster when presented with strong alerts across all groups. It was observed that engaging in a typing task generally led to faster reaction times in all groups, except in the single alert group during the Construction Zone Scenario (S1) and the mild alert group in the Animal on the Road Scenario (S2), as detailed in Figure 4.4. Furthermore, in the dynamic group, drivers reacted significantly faster by an average of 1.749 seconds when responding to strong alerts during typing tasks, compared to a modest improvement of 0.309 seconds among drivers in the single alert groups, with the most pronounced difference noted in the Animal on the Road Scenario (S2).

Analysis of the type and timing of the first response revealed that braking was the predominant initial reaction across scenarios, regardless of secondary task involvement, as depicted in 4.3. The use of turn signals was the next most frequent response. However, there was no significant difference in the takeover times among reaction types. Exceptionally quick reactions (within 1 second) occurred under strong alert conditions in the Construction Zone Scenario (S1) without a secondary task and in the Blind Spot Lane Change Scenario (S3)

Group	Reaction Type	T-Statistic	p-Value
Dynamic	Indicator	1.04	0.32
Dynamic	Steering	-2.89	0.013*
Dynamic	Braking	-2.16	0.05*
Single	Indicator	-2.01	0.079.
Single	Steering	-0.6	0.558
Single	Braking	-2.04	0.059.

Table 4.2: Paired T-test (ttest_rel in scipystats Python library) on the differences in reaction times of takeover maneuvers (using indicator, braking, and steering) in all scenarios with and without NDRT across the dynamic and single groups.

with typing. Conversely, very delayed reactions (exceeding 6 seconds) were recorded in the Construction Zone Scenario (S1) with mild alerts. There were no significant performance differences between the dynamic and single alert groups when averaged across all scenarios. Notably, reaction times were reduced in the Blind Spot Lane Change (S3) and Pulled Over Vehicle Merging (S4) Scenarios when drivers were engaged with their phones, except for the Blind Spot Lane Change Scenario (S3). However, slower reactions were observed with mild alerts when no secondary task was present. Regarding safety margins, mild alerts were more effective at maintaining safer distances from hazards in scenarios without secondary tasks, whereas the strongest safety performance was observed with strong alerts in scenarios involving typing tasks.

When comparing the difference in reaction times between the scenarios with no secondary task and the scenarios with NDRT within the same participant groups, faster response times were observed in general except for the Construction Zone Scenario (S1) where the response time is much slower for the strong alert group. Noticeably faster response times were observed among participants in the dynamic alert group that alerts changed to strong alert design in all scenarios compared to the single alert groups as shown in Figure 4.4. The largest difference was found in the Animal on the Road Scenario (S2) with over 2 seconds faster on average for the strong alerts in the dynamic group in Scenarios 2 and 4. There were no significant differences when comparing within the single groups. However, as mentioned, reaction times were significantly faster when comparing the secondary task with typing, using ANOVA, resulting in $F = 6.25$, $p = .014$ for braking reactions and $F = 9.12$, $p = .003$ for steering reactions. By comparing the differences in reaction times between the no-task and typing scenarios among the dynamic and single groups using paired t-test, as shown in Table 4.2, steering and braking were significantly faster with a p-value of .013 and .05 respectively, with marginally close to significant in braking ($p = .066$) among participants in the dynamic group.

Scenario	Metric	Factor	Sum of Squares	F-value	p-value
1 (Construction Zone)	Min Front Distance	Design	4352.00	2.37	.129
1 (Construction Zone)	Min Front Distance	Secondary Task	3943.00	2.15	.149
1 (Construction Zone)	Min Front Distance	Group	9486.00	5.17	.027*
1 (Construction Zone)	Max Brake Acceleration	Design	0.0042	0.20	.655
1 (Construction Zone)	Max Brake Acceleration	Secondary Task	0.0090	0.44	.511
1 (Construction Zone)	Max Brake Acceleration	Group	0.0404	1.98	.170
1 (Construction Zone)	Max Lateral Acc. Neg	Design	0.0092	1.85	.180
1 (Construction Zone)	Max Lateral Acc. Neg	Secondary Task	0.0111	2.24	.141
1 (Construction Zone)	Max Lateral Acc. Neg	Group	0.0105	2.12	.152
1 (Construction Zone)	Max Lateral Acc. Pos	Design	0.0007	0.16	.691
1 (Construction Zone)	Max Lateral Acc. Pos	Secondary Task	0.0001	0.02	.879
1 (Construction Zone)	Max Lateral Acc. Pos	Group	0.0010	0.22	.645
2 (Animal on Road)	Min Front Distance	Design	2930.30	2.86	.094.
2 (Animal on Road)	Min Front Distance	Secondary Task	1136.67	1.11	.295
2 (Animal on Road)	Min Front Distance	Group	2349.84	2.30	.134
2 (Animal on Road)	Max Brake Acceleration	Design	0.0765	1.85	.177
2 (Animal on Road)	Max Brake Acceleration	Secondary Task	0.0027	0.07	.800
2 (Animal on Road)	Max Brake Acceleration	Group	0.0018	0.04	.834
2 (Animal on Road)	Max Lateral Acc. Neg	Design	0.0074	2.19	.143
2 (Animal on Road)	Max Lateral Acc. Neg	Secondary Task	0.0148	4.38	.040*
2 (Animal on Road)	Max Lateral Acc. Neg	Group	0.0057	1.69	.198
2 (Animal on Road)	Max Lateral Acc. Pos	Design	0.0200	16.21	.0001***
2 (Animal on Road)	Max Lateral Acc. Pos	Secondary Task	0.0246	20.35	.00002***
2 (Animal on Road)	Max Lateral Acc. Pos	Group	0.0000	0.00	.994
3 (Blind Spot Lane Change)	Min Front Distance	Design	6885.51	2.02	.160
3 (Blind Spot Lane Change)	Min Front Distance	Secondary Task	305.08	0.09	.766
3 (Blind Spot Lane Change)	Min Front Distance	Group	18.59	0.01	.941
3 (Blind Spot Lane Change)	Max Brake Acceleration	Design	0.0961	2.81	.099
3 (Blind Spot Lane Change)	Max Brake Acceleration	Secondary Task	0.0315	0.92	.342
3 (Blind Spot Lane Change)	Max Brake Acceleration	Group	0.0094	0.27	.603
3 (Blind Spot Lane Change)	Max Lateral Acc. Neg	Design	0.0068	2.04	.158
3 (Blind Spot Lane Change)	Max Lateral Acc. Neg	Secondary Task	0.0003	0.09	.763
3 (Blind Spot Lane Change)	Max Lateral Acc. Neg	Group	0.0491	14.65	.0003***
3 (Blind Spot Lane Change)	Max Lateral Acc. Pos	Design	0.0005	0.09	.766
3 (Blind Spot Lane Change)	Max Lateral Acc. Pos	Secondary Task	0.0007	0.13	.716
3 (Blind Spot Lane Change)	Max Lateral Acc. Pos	Group	0.0471	9.05	.0039**
4 (Pulled Over Vehicle)	Min Front Distance	Design	753.22	0.26	.609
4 (Pulled Over Vehicle)	Min Front Distance	Secondary Task	3686.84	1.29	.259
4 (Pulled Over Vehicle)	Min Front Distance	Group	2038.55	0.71	.401
4 (Pulled Over Vehicle)	Max Brake Acceleration	Design	0.0448	1.22	.274
4 (Pulled Over Vehicle)	Max Brake Acceleration	Secondary Task	0.0024	0.07	.800
4 (Pulled Over Vehicle)	Max Brake Acceleration	Group	0.1242	3.37	.071.
4 (Pulled Over Vehicle)	Max Lateral Acc. Neg	Design	0.1516	11.81	.001***
4 (Pulled Over Vehicle)	Max Lateral Acc. Neg	Secondary Task	0.0183	1.43	.236
4 (Pulled Over Vehicle)	Max Lateral Acc. Neg	Group	0.0127	0.99	.324
4 (Pulled Over Vehicle)	Max Lateral Acc. Pos	Design	0.0280	5.79	.019*
4 (Pulled Over Vehicle)	Max Lateral Acc. Pos	Secondary Task	0.0052	1.08	.303
4 (Pulled Over Vehicle)	Max Lateral Acc. Pos	Group	0.0005	0.11	.738

Table 4.3: Type II ANOVA analysis (using ols model in statsmodels Python library) of takeover safety metric interactions on group (dynamic vs. single), alert design (mild vs. strong), and secondary task (no NDRT vs. with NDRT) for each scenario (degrees of freedom = 1).

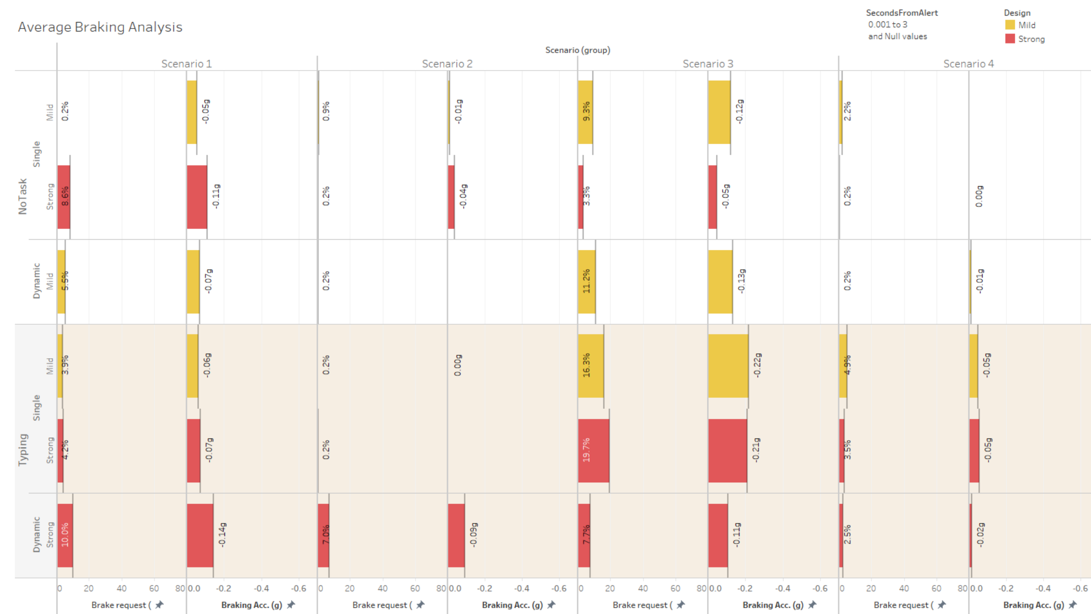


Figure 4.5: Average braking measures during the 0-3 seconds after takeover in each scenario. The graph presents two key metrics: the percentage of brake requests and the average braking acceleration (g). Each scenario is analyzed under different alert conditions (Mild and Strong) and secondary task conditions (No task and Typing), allowing for a comprehensive comparison of braking behavior across varying conditions.

Takeover Safety Analysis

Safe takeover is indicated by how hard the brake is applied, the lateral accelerations, and the use of an indicator on a change of direction [3]. The average of the measures focuses on the time range of 1-8 seconds after TOR which was determined based on the rounded average takeover time with a 90% confidence interval. As observed, six collisions occurred during the takeover in the Animal on the Road Scenario (S2) and the Pulled Over Vehicle Merging Scenario (S4) combined. The dynamic group experienced two frontal collisions in Scenario 2, one frontal collision and a rear-end collision after lane change in Scenario 4, characterized by mild alerts with average collision times after takeover of 7.24 seconds and 4.53 seconds respectively. Additionally, within Scenario 4, the dynamic group experienced one collision marked by a strong alert at 1.55 seconds post-takeover while the single group encountered another collision also following the strong alert after 2.68 seconds resuming control.

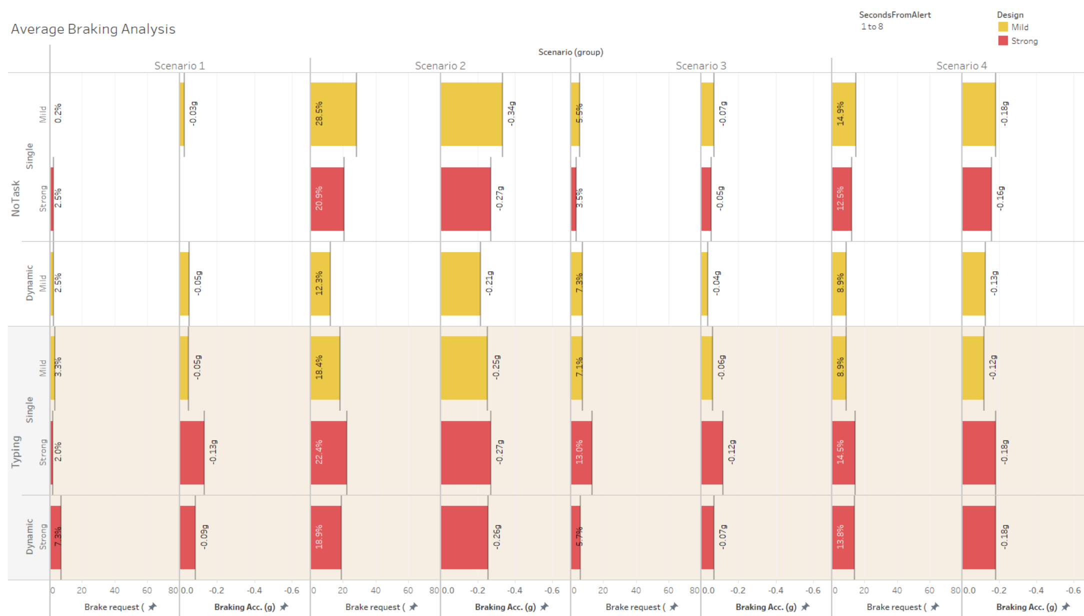


Figure 4.6: Average braking measures during the 1-8 seconds after takeover in each scenario. The graph presents two key metrics: the percentage of brake requests and the average braking acceleration (g). Each scenario is analyzed under different alert conditions (Mild and Strong) and secondary task conditions (No task and Typing), allowing for a comprehensive comparison of braking behavior across varying conditions.

- **Braking**

  Harsh braking indicates a more dangerous takeover because this may cause rear-ended accidents on the road. This study investigates how hard the driver brakes in percent and the acceleration of the braking to determine the safety of the takeover. In this study, the brake is the main control for disabling ACC and AP to gain manual control. Additionally, all takeover scenarios include hazards in the front. The analysis of the use of brakes in the first three seconds after the alert shows that harsher brakes were used in the Blind Spot Lane Change Scenario (S3) with a secondary task in single alert groups (-0.22 g) and in the Construction Zone Scenario (S1) in the dynamic group (-0.14 g) as shown in Figure 4.5. Overall, the highest brake accelerations were observed in the Animal on the Road Scenario (S2) with a mild alert as 0.61 (SD = 0.26) g without NDRT and 0.60 (SD = 0.005) g while typing. In all situations, the strong alerts led to faster and slightly more intense braking in the initial 3 seconds following the alert compared to the mild alerts. However, a more extended braking period was observed in the Animal on the Road Scenario (S2) and the Pulled Over Vehicle Merging Scenario (S4) after the takeover request, as illustrated in Figure 4.6.

  The minimum front distances were measured to analyze risky takeovers. However, across all scenarios, none of the factors, including design, secondary task, or group, showed a significant effect on minimum front distance or maximum brake acceleration. This consistency across scenarios suggests that these specific categorical distinctions in the experiment do not influence the intensity of braking behavior under the conditions tested.

- **Steering**

  Steering angle velocities can be used to analyze fast steering turns during the takeover, while lateral accelerations can be used to measure how hard the vehicle turns. Fast steering velocities and high lateral accelerations during takeover can cause unstable steering, which could lead to unsafe situations. Figure 4.7 shows how drivers mitigated road hazards using steering in each scenario. In the Construction Zone Scenario (S1), there were fewer instances of quick steering maneuvers, whereas in the other scenarios, particularly in the Pulled Over Vehicle Merging Scenario (S4) where a car pulled over was merging from the right shoulder, most drivers turned left. In the Blind Spot Lane Change Scenario (S3), there were strong alerts on both the left and right sides as evidenced by rapid steering turns (0.341 and 0.4613 rad/s respectively).

  According to Table 4.3, higher lateral accelerations were significantly influenced by alert design, particularly noted in scenarios involving mild alerts. Specifically, during

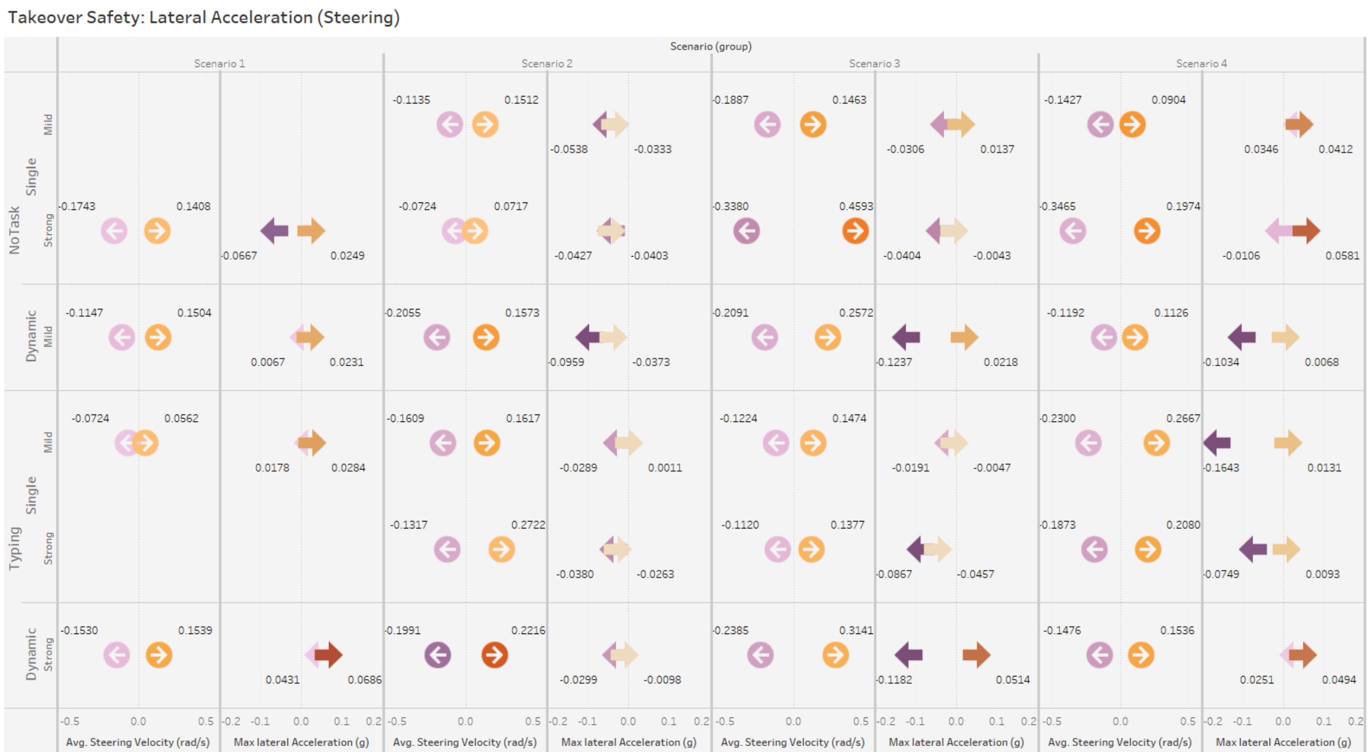


Figure 4.7: Steering measures during 1-8 seconds after the takeover in each scenario and assigned secondary task. Measurements include steering angle velocity (rad/s) and the average of maximum lateral acceleration (g). Left (purple) and right (orange) icons indicate steering to left or right respectively, higher color contrast indicates higher maximum values among drivers.

the Pulled Over Vehicle Merging Scenario (Scenario 4), the most pronounced lateral accelerations were observed. In this scenario, the mean lateral acceleration for the single group was notably higher at -0.164 g compared to -0.103 g for the dynamic group, indicating a considerable effect of group categorization on the response to mild alerts. Furthermore, significant differences were also identified in the Blind Spot Lane Change Scenario (Scenario 3). In this scenario, the dynamic group exhibited greater lateral acceleration when turning left, reaching a mean of 0.121 g, with this finding being statistically significant ($p < 0.001$, degrees of freedom = 1). Additionally, the mild alert design was found to influence stronger right lateral acceleration in scenarios involving animals on the road, highlighting how alert intensity can differentially impact driver responses based on the nature of the alert and the immediate driving context.

4.1.2 Physiological Response during Takeover

Eye movement data was collected simultaneously with video footage of the front part of the vehicle. Unfortunately, out of the total 160 eye-tracking sessions, 107 were deemed usable for analysis, with 15 sessions being excluded due to incomplete data, and sessions in which participants disabled AP before alerts were also excluded from the takeover analysis, which is boiled down to a total of 107 sessions as 13, 31, 29, and 34 sessions for Scenario 1, 2, 3, and 4 respectively.

The analysis of the HR data reveals significant inconsistencies during the periods of TOR events for all participants, as shown in Figure 4.8. Although the HR data was recorded at a frequency of 1 Hz using the Empatica E4 wristband, the HR data during the short takeover periods (10-20 seconds) exhibits a smoothing effect. This effect is markedly different from the more variable HR patterns observed during regular driving intervals before TOR events. The smoothing is often followed by a sudden drop in heart rate, suggesting potential inaccuracies or artifacts in the data during these critical periods. Such discrepancies indicate that the HR data collected during takeover events may not reliably reflect the participants' physiological responses. After carefully analyzing each participant's data, it was determined that the heart rate data recorded during these takeover periods should be excluded from the final analysis to ensure the reliability of the study's conclusions when comparing between alert types or scenarios. However, general trends suggested an increased HR shortly after TOR compared to HR during cruising in all participants.

In eye movement analysis, the takeover process was segmented into three parts, 20 s before the TOR, between the TOR and the first response (takeover period), and post-takeover until the hazard situation was mitigated. Area of Interest (AOI) were segmented using automated marker detection with manual area tracking when eye movements were not found in the front-view footage, usually when the phone was being used. The AOI included a road, dashboard, phone, and mirrors. Glance and AOI attention ratio calculations were performed to analyze driver behavior during the takeover.

Pre-Alert AOI Attention Ratio

During the experiment, drivers were asked to type on the phone as much as possible when they felt it was safe enough to do so to replicate the realistic typing task in the real world. According to the attention ratio in Figures 4.9 and 4.10, the most attention was highest on the road when there was no secondary task; however, when typing (a secondary task), the attention of drivers to the road decreased substantially, observations indicated similar phone usage with an average 34% attention ratio that reduced road attention to less than half before alerts, which is expected as they were engaged with their phones. Interestingly, the road attention in the typing condition with strong alerts is still quite high, indicating that these alerts may have been effective in capturing the driver's attention even during a distracting task.

During Takeover AOI Attention Ratio (from alert to the first response)

During the takeover period among drivers in scenarios without NDRT, increased attention on the road following the strong alert in all conditions, indicating that drivers were refocusing on driving in response to the alerts, while dashboard attention had an increase in attention ratio, suggesting that drivers were looking at the dashboard possibly to get more information about the alert or the situation on the road. The single alert groups, whether under mild or strong alert conditions, show an increase in road attention, but the strong alert condition seems to elicit slightly more focused attention on the road compared to the mild condition. This is especially noticeable in scenarios with a secondary task, where strong alerts appear to be more effective in redirecting attention.

Post-Takeover AOI Attention Ratio (from the first response until the hazard is mitigated)

After taking control, the dynamic group maintains high road attention in all scenarios, indicating that the previous variability in alerts may have established a heightened sense

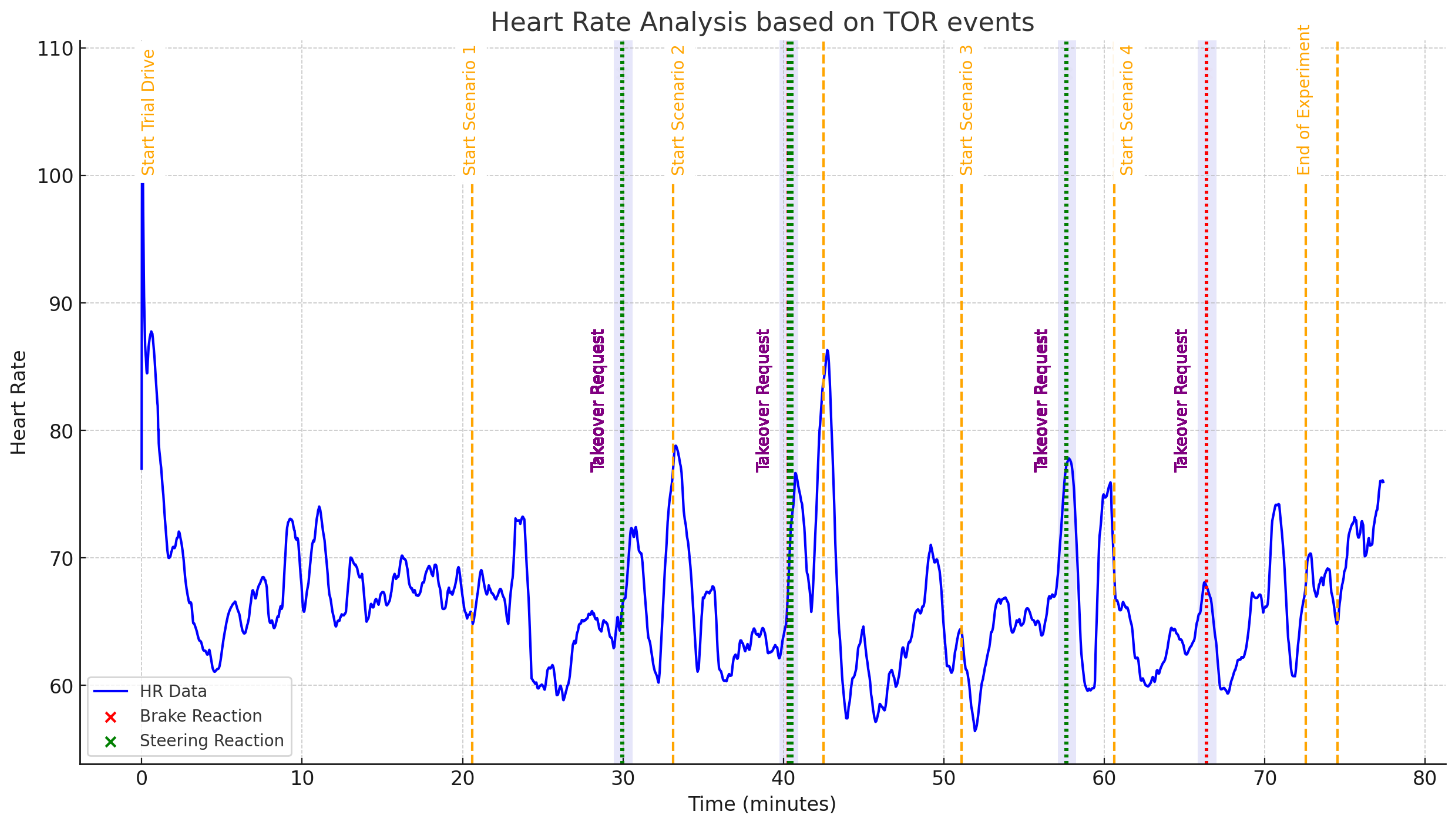


Figure 4.8: Heart Rate Analysis based on TOR Events. The plot illustrates an example of the HR data over time in minutes, represented by the blue line. Red and green dots and corresponding dotted vertical lines mark the instances of reactions detected on brake and steering respectively. Purple-shaded areas highlight the periods of TOR events. Orange dashed vertical lines indicate the timestamps from the event tags in orange, label specific scenarios and milestones: "Start Trial Drive", "Start Scenario 1", "Start Scenario 2", "Start Scenario 3", "Start Scenario 4", and "End of Experiment".

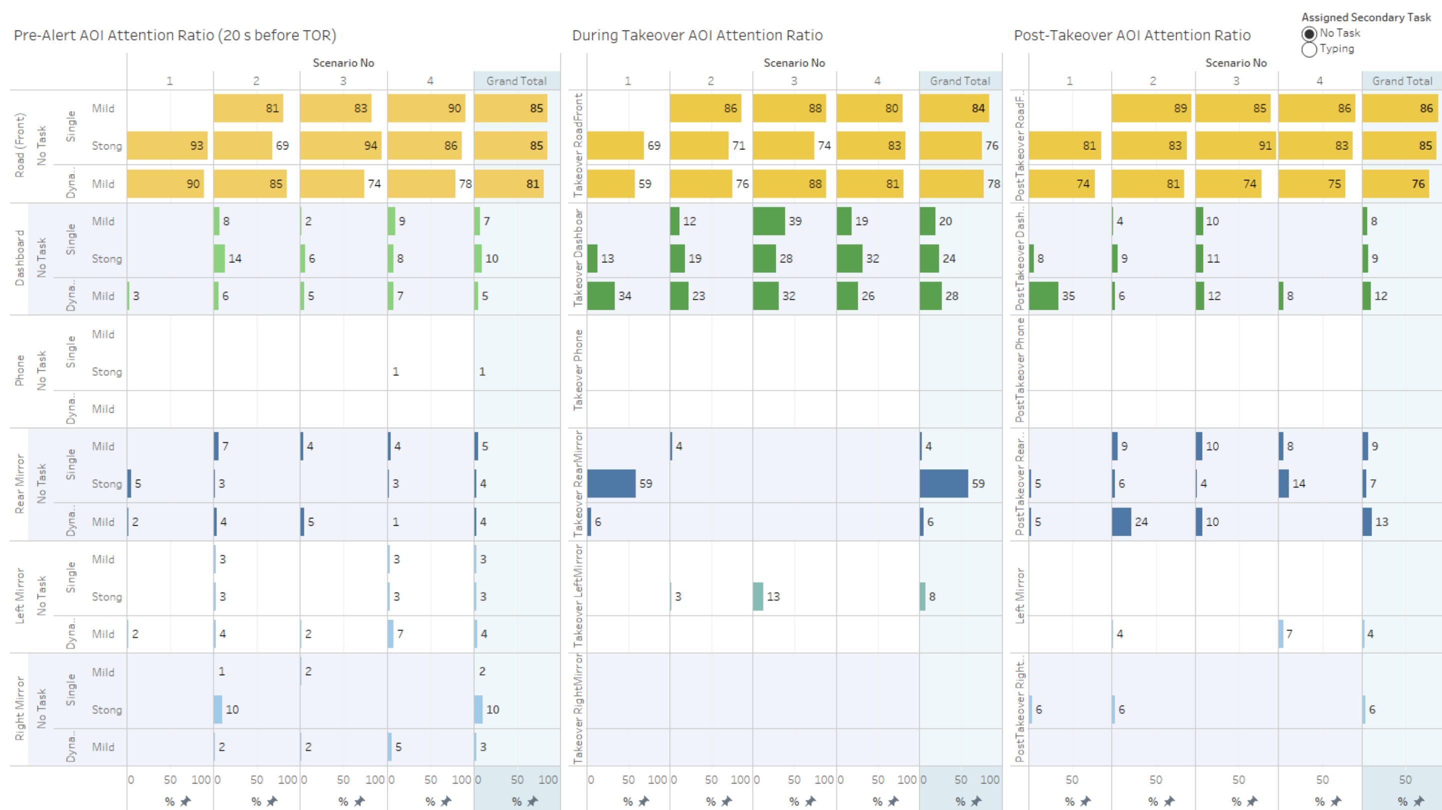


Figure 4.9: Percentage of attention in different Areas of Interest (AOIs) in each scenario with **no secondary task** comparing groups and alert designs. The chart displays data for three time periods: 20 seconds before the alert (Pre-Alert), during the takeover request (Takeover), and post-takeover periods (Post-Takeover). Attention ratios are shown for each AOI including the road (front), dashboard, phone, and mirrors (rear, left, and right). The analysis highlights how participants' focus shifts across different scenarios and conditions, providing insights into their attention distribution before, during, and after the takeover event.

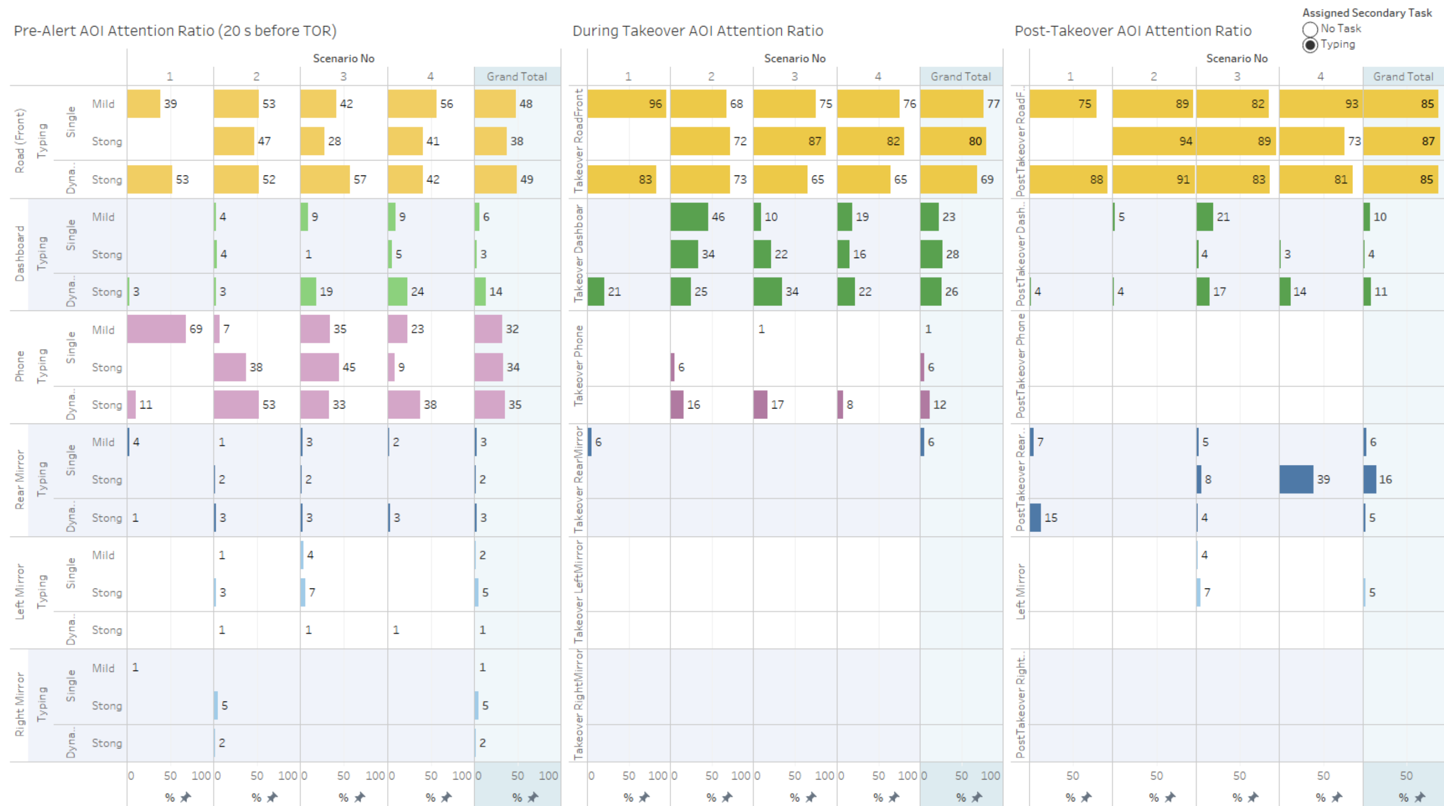


Figure 4.10: Percentage of attention in different Areas of Interest (AOIs) in each scenario with **NDRT (typing task)**, comparing groups and alert designs. The chart displays data for three time periods: 20 seconds before the alert (Pre-Alert), during the takeover request (Takeover), and post-takeover periods (Post-Takeover). Attention ratios are shown for each AOI including the road (front), dashboard, phone, and mirrors (rear, left, and right). The analysis highlights how participants' focus shifts across different scenarios and conditions, providing insights into their attention distribution before, during, and after the takeover event.

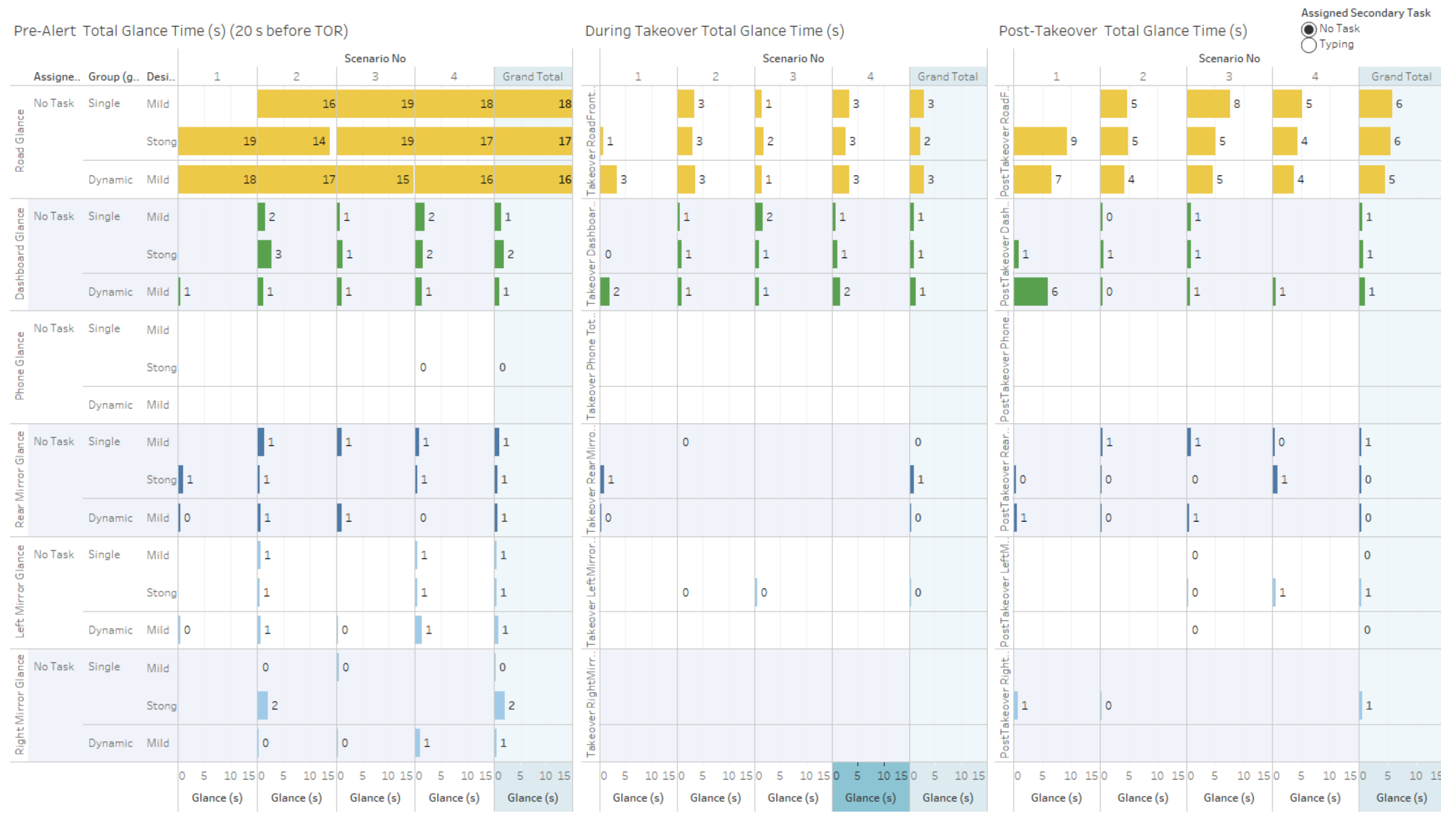


Figure 4.11: Total glance duration (seconds) in different Areas of Interest (AOIs) for each scenario with **no secondary task**. This figure compares the groups and alert designs during three periods: 20 seconds before the alert (Pre-Alert), during the takeover request (Takeover), and post-takeover periods (Post-Takeover). The AOIs include glances at the road (front), dashboard, phone, and mirrors (rear, left, and right). The analysis provides insights into how participants' visual attention is distributed across different areas in various scenarios and conditions, highlighting the impact of alert designs on glance behavior.

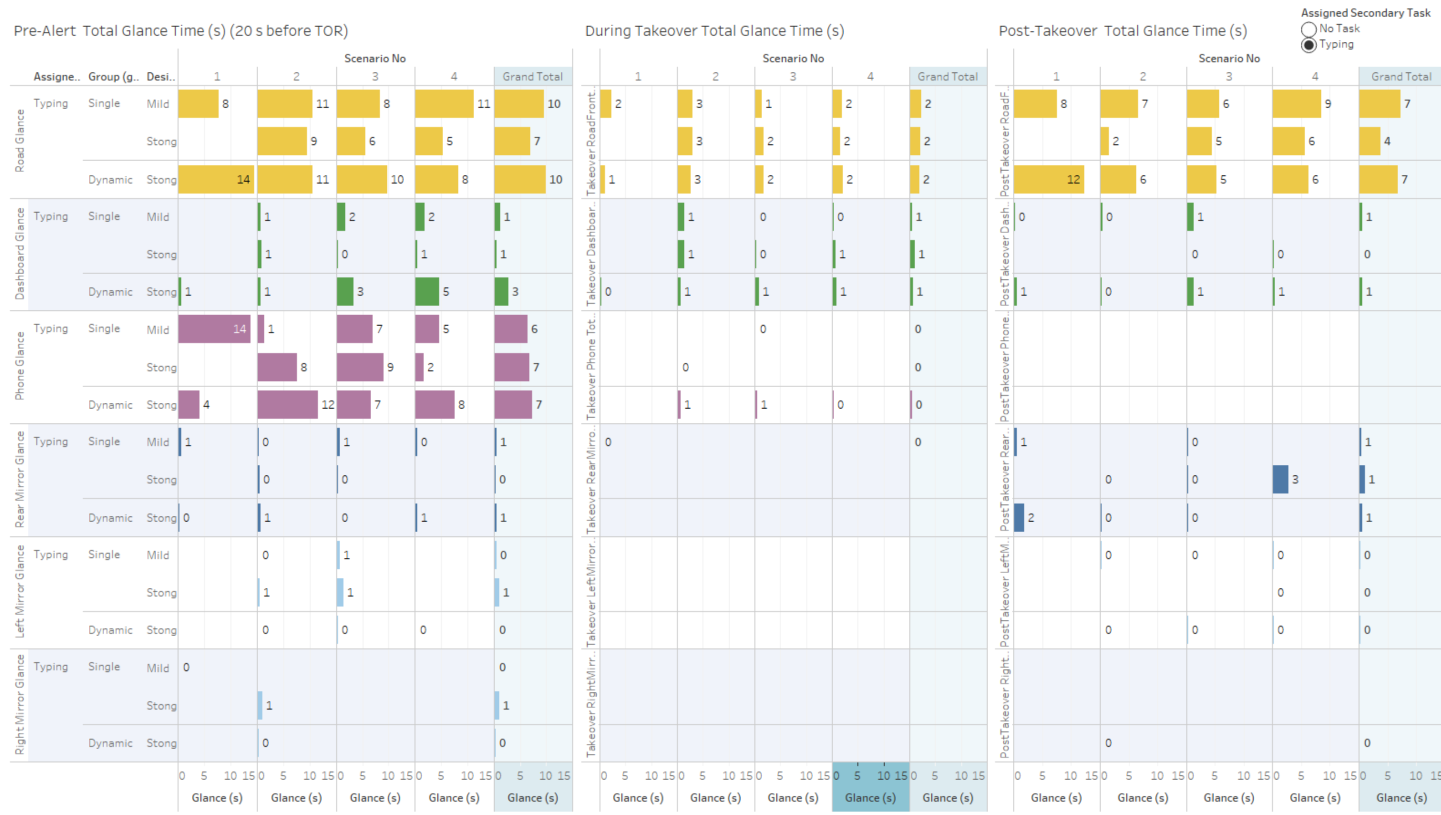


Figure 4.12: Total glance duration (seconds) in different Areas of Interest (AOIs) for each scenario with **NDRT (typing task)**. This figure compares the groups and alert designs during three periods: 20 seconds before the alert (Pre-Alert), during the takeover request (Takeover), and post-takeover periods (Post-Takeover). The AOIs include glances at the road (front), dashboard, phone, and mirrors (rear, left, and right). The analysis provides insights into how participants' visual attention is distributed across different areas in various scenarios and conditions, highlighting the impact of alert designs on glance behavior while engaged in a secondary typing task.

of caution or vigilance. Similarly, single-alert groups maintain high attention on the road after the takeover. However, there is a notable increase in mirror checking in the strong alert conditions, suggesting that the intensity of the alert may prompt more thorough post-takeover checks.

4.1.3 Qualitative Insights

In this comprehensive analysis, we delve into the dynamics among various experimental groups ('Exp Group'), alert designs ('Design'), and the transitions between scenarios with and without secondary tasks, specifically focusing on attention capturing, understanding, SA comprehension, TOR timing, confidence, annoyance, and alert design satisfaction. The summary of the results and statistical analysis, including Type II ANOVA and the group, alert design, and cognitive state interactions on each metric, are shown in Table 4.4 and 4.5 respectively.

Attention Capturing

Attention-capturing scores gauge how much participants felt that alerts grabbed their attention during the TOR. Dynamic groups with mild alerts and no secondary task (concentrated state) reported an initial mean score of 3.59 (SD = 0.98). Introducing a stronger alert in this group significantly enhanced the scores to 4.28 (SD = 0.88), indicating better alert effectiveness. In single groups, transitioning from a concentrated state (no task, M = 4.07, SD = 0.99) to a distracted state (typing, M = 4.67, SD = 0.62) within a mild alert group demonstrated a notable increase in attention capturing. This trend was consistent under strong alerts, where the mean score remained high at about 4.38, irrespective of the presence of a secondary task, highlighting the robust design's ability to capture attention. The ANOVA showed a significant effect, $F = 8.32$, $p = .004$, when accompanied by the NDRT.

TOR Timing

Although all the TORs were issued as designed at the same timing across all groups and designs in each scenario, the perception of well-timed alerts was generally better in individual groups than in dynamic groups. For instance, single groups without NDRT (concentrated state) recorded a mean TOR timing of 4.00 (SD = 1.11), which increased to 4.33 (SD = 0.90) in the third and fourth scenarios. This suggests that cognitive engagement with a secondary task (distracted state) might enhance responsiveness.

Group	Design	NDRT	Alert Effectiveness (Mean, SD)	TOR Timing (Mean, SD)	Understandability (Mean, SD)	SA Comprehension (Mean, SD)	Confidence (Mean, SD)	Annoyance (Mean, SD)	Overall Satisfaction (Mean, SD)	Takeover Before Alert (Mean %, SD)
Dynamic	Mild Alert	No Task	3.59, 0.98	3.45, 1.21	3.59, 1.12	2.52, 1.09	3.03, 1.20	2.14, 1.05	3.21, 0.94	34.5%, 48.4%
Dynamic	Strong Alert	Typing	4.28, 0.88	3.69, 1.17	4.17, 0.85	2.55, 1.30	3.93, 0.85	1.34, 0.49	4.00, 0.89	17.2%, 38.4%
Single	Mild Alert	No Task	4.07, 1.00	4.00, 1.11	3.71, 1.20	2.50, 1.40	3.64, 1.27	1.36, 0.50	3.71, 1.14	7.1%, 26.7%
Single	Mild Alert	Typing	4.67, 0.62	4.33, 0.90	4.13, 0.99	1.93, 1.10	4.27, 0.79	1.53, 0.92	4.07, 0.80	20.0%, 41.4%
Single	Strong Alert	No Task	4.38, 0.89	3.94, 1.24	4.44, 0.81	2.25, 1.18	3.81, 1.17	1.38, 0.50	4.00, 1.10	18.8%, 40.3%
Single	Strong Alert	Typing	4.38, 0.65	3.77, 1.30	4.00, 1.00	1.92, 1.04	4.23, 0.83	1.15, 0.38	3.92, 1.26	30.8%, 48.0%

Table 4.4: The average ratings and standard deviations (SD) for various aspects of the alert experience, including alert effectiveness, TOR timing, understandability, SA comprehension, confidence, annoyance, overall satisfaction, and the percentage of takeovers before the alert. The data is grouped by the experiment group (Dynamic vs. Single), alert design (Mild vs. Strong), and secondary task (No Task vs. Typing).

Metric	Factor(s)	Sum of Squares	Df	F-value	P-value
Alert Effectiveness	Group	1.68	1	2.2	0.141
	Design	2.34	1	3.07	0.083.
	Secondary Task	6.65	1	8.71	0.004**
	Group:Design	1.09	1	1.43	0.234
	Group:Secondary Task	0.21	1	0.28	0.598
	Design:Secondary Task	1.24	1	1.62	0.206
TOR Timing	Group	1.91	1	1.4	0.24
	Design	0.03	1	0.02	0.891
	Secondary Task	1.64	1	1.2	0.276
	Group:Design	1.8	1	1.32	0.253
	Group:Secondary Task	0.3	1	0.22	0.639
	Design:Secondary Task	0.91	1	0.67	0.417
Understandability	Group	0.01	1	0.02	0.897
	Design	5.5	1	6.81	0.01**
	Secondary Task	1.37	1	1.7	0.195
	Group:Design	0.09	1	0.11	0.74
	Group:Secondary Task	0	1	0	0.959
	Design:Secondary Task	0.42	1	0.52	0.473
Confidence	Group	5.53	1	12.15	0.001***
	Design	1.3	1	2.67	0.105
	Secondary Task	0	1	0	0.966
	Group:Design	0.05	1	0.11	0.742
	Group:Secondary Task	0	1	0	0.965
	Design:Secondary Task	0.42	1	0.47	0.494
Annoyance	Group	0.23	1	0.72	0.398
	Design	0.45	1	1.41	0.238
	Secondary Task	1.19	1	3.69	0.057.
	Group:Design	0.37	1	1.15	0.285
	Group:Secondary Task	0.15	1	0.48	0.491
	Design:Secondary Task	0.27	1	0.84	0.362
Satisfaction Overall	Group	0.91	1	0.92	0.34
	Design	4.74	1	4.76	0.031*
	Secondary Task	5.76	1	5.79	0.018*
	Group:Design	1.53	1	1.54	0.218
	Group:Secondary Task	1.22	1	1.22	0.272
	Design:Secondary Task	0.66	1	0.67	0.416

Table 4.5: Type II ANOVA results for various surveyed aspects (using ols model in statsmodels Python library), examining the influence of the experimental group (dynamic vs. single), alert design (mild vs. strong), secondary task (no NDRT vs. with NDRT), and their interactions on metrics such as alert effectiveness, TOR timing, understandability, confidence, annoyance, and overall satisfaction. Significance levels are indicated as follows: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Understandability and SA Comprehension

Scores for understandability remained high across all groups and conditions, with some significance in better understanding upon strong alerts ($p = .01$), peaking in single groups under strong alerts without secondary tasks ($M = 4.44$, $SD = 0.81$). However, the abilities of drivers to comprehend the takeover situation measured as comprehension scores were more variable, typically decreasing in scenarios that involved secondary tasks (distracted state), particularly under strong alerts where the mean shifted from higher without tasks to 1.92 ($SD = 1.04$) with tasks. However, improvements in comprehension scores were observed only in the dynamic group.

Confidence

Confidence levels surged notably in scenarios involving secondary tasks (distracted state) in the third and fourth scenarios, particularly within the dynamic group, where confidence increased from a mean of 3.03 to 3.93 ($SD = 0.85$) with a highly significant difference, $p = .001$. This suggests that experiencing the alerts multiple times may boost participants' confidence in their understanding and reactions, but dynamic alerts could enhance this effect. The ANOVA showed this change to be statistically significant, $F = 12.97$, $p = .001$.

Annoyance

Annoyance levels remained consistently low across all settings, though there was a slight increase in the single group under mild alerts when a secondary task (distracted state) was added, from a mean of 1.36 to 1.53. Other groups experienced a decrease in annoyance levels, particularly in the dynamic group, where it decreased from 2.14 to 1.34. Generally, lower annoyance levels were reported with the strong alert design, although it was not statistically significant according to the ANOVA results. This could be explained by findings from [33], [56] that indicate a strong link between increased urgency levels in alerts and perceived appropriateness in situations requiring high urgency to reflect critical circumstances.

Alert Design Satisfaction

Overall satisfaction with the alert design was generally higher in scenarios involving strong alerts, whether or not tasks were present. This trend was evident in both dynamic (stable mean: 4.00, SD = 0.89) and single groups (mean improvement from 4.07, SD = 0.80), reinforcing the effectiveness of strong alert designs in maintaining user satisfaction even with increased task complexity. The ANOVA indicated significant improvements, $F = 5.67$, $p = .018$, when the participants were assigned a secondary task (distracted state).

Participants generally rated the strong alert as more effective, understandable, and comprehensible than the mild alert, especially in the dynamic group. While drivers felt more confident in single groups, it was observed that dynamic alerts were found less annoying in the dynamic group, especially with the strong alert.

The ANOVA findings from the survey data reveal several significant insights into how participants perceive various aspects of the alert system based on the experimental group (Dynamic vs. Single) and the alert design (Strong vs. Mild). Notably, "Alert Effectiveness" and "Confidence" show significant differences between the experimental groups, indicating that participants' group assignments have a notable impact on how effective and confidence-inspiring they find the alerts. Furthermore, "Understandability" and "Alert Design Satisfaction" are significantly influenced by the alert design (p-value < 0.05), suggesting that the strength of the alert plays a crucial role in how understandable and satisfying participants find the alert system. Additionally, significant interaction effects between the group and the design were observed in "Alert Effectiveness" and "Satisfaction Overall," indicating that the impact of the alert design varies depending on the experimental group. However, aspects like "Timing," "SA Comprehension," and "Annoyance" did not exhibit strong significant differences based on these factors, suggesting a more uniform perception across different conditions for these aspects. In general, the ANOVA results highlight the nuanced ways in which different elements of the alert system influence user perceptions and underscore the importance of considering both individual and contextual factors in the design and evaluation of the alert system.

Post-Experiment: Alert Perception Test

The alert perception test was designed to help validate how the alerts were perceived by 30 participants in random order. As illustrated in Figure 4.13, the average ratings suggest that strong alerts are perceived as slightly more urgent than mild alerts, but the difference is not statistically significant, $p = .46$, while the annoyance ratings are identical on average between mild and strong alerts, supported by a p-value of 1.0, indicating no significant difference. When analyzing the perceived urgency and annoyance based on participant groups, there is also no statistically significant difference between the groups, as indicated by p-values of .15 and .18, respectively. More than half of the participants preferred the strong alert over the mild alert design.

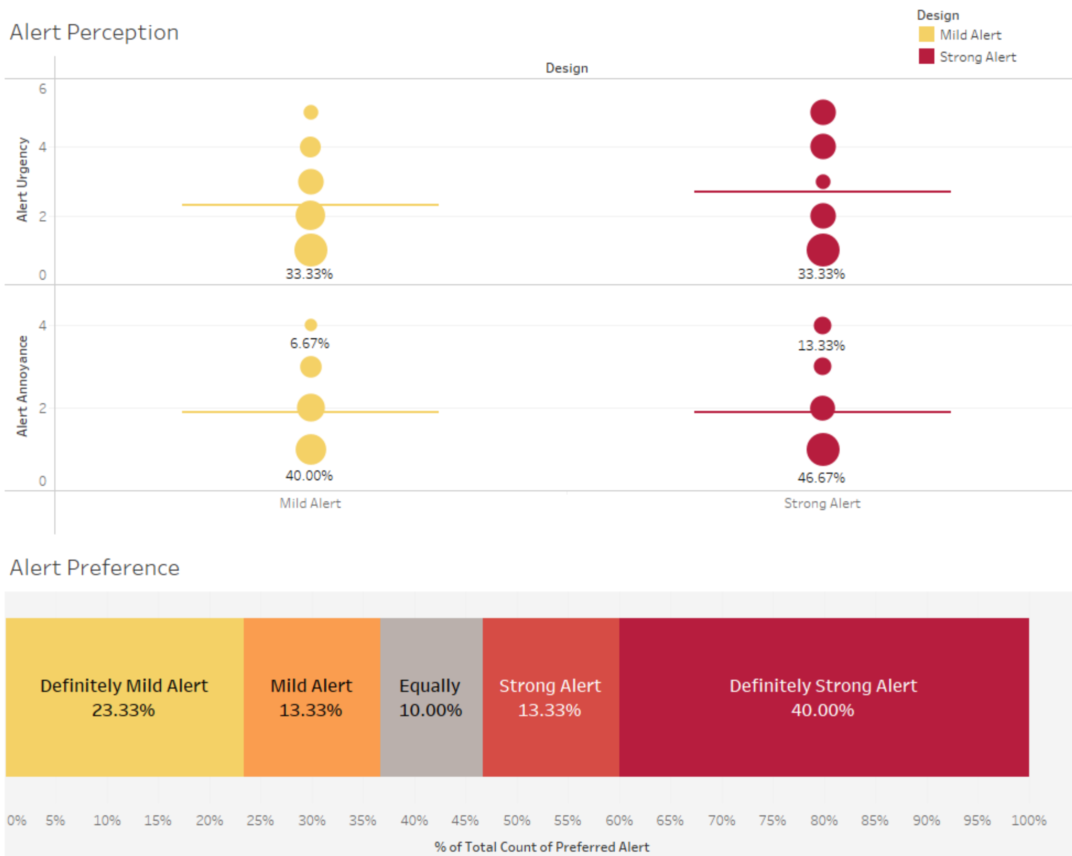


Figure 4.13: Perceived urgency and annoyance of mild and strong alerts (top) and overall alert preference (bottom). The top section illustrates participants' perception of alert urgency and annoyance for both mild and strong alerts, depicted as a distribution of ratings with bubble size representing the percentage of participants who selected each rating. The bottom section shows the overall preference distribution among participants, indicating a higher preference for strong alerts.

4.2 Discussions

4.2.1 Takeover Reactions and Performance

When we look at how drivers react to different types of warnings in automated driving, stronger warnings help drivers take control faster in most cases, this is especially true when drivers are distracted which was in contrast to findings in Melnicuk et al. [43]. Interestingly, how drivers respond to these warnings also depends on how predictable the situation is, they might not react as quickly.

The differences in reaction times between the dynamic and single groups highlight key insights into driver responsiveness under varied conditions. The dynamic group consistently demonstrated quicker responses, especially under strong alerts and while multitasking with typing. This group's reaction time improved by an average of 1.749 seconds, suggesting a high adaptability to demanding driving scenarios. In contrast, the single group showed minimal changes in reaction time, with reductions being modest and not statistically significant, indicating less responsiveness to alert intensity and secondary tasks. These findings aligned with the advantages of dynamic alerts as discussed in [10], [15], indicating that a faster response time was associated with the dynamic alerts.

When cruising on a highway in autonomous driving mode during the study, unlike real car control, the driver only had the option to use the AP toggle button or turning signals to access steering control, or to use the brake pedal to disable all automated driving features. As shown in Figure 4.4, braking was the most frequently used method for taking over in all situations. Similar findings were also documented in [3], with braking being employed in 67% of takeover maneuvers. When looking closer at how drivers respond, we find that strong warnings make them act more quickly and decisively. It seems that these urgent warnings refocus the driver's attention on driving, leading to faster braking and steering. An interesting observation in any of the first scenarios of some participants indicated that when drivers experienced the TOR for the first time, there were signs of confusion or difficulty in trying to take over manual control, although they felt comfortable driving in the trial scenario, similar findings were reported by Kim and Yang that takeover reactions could also be influenced by different takeover scenarios [3].

By tracking where drivers look, we learn about what grabs their attention. Eye movements in distracted drivers showed a growing tendency for off-road glances, yet reaction times, in general, remained unimpeded, aligning with findings in other studies [45], [57] that found no general correlation between lower off-road glances and quicker responses to the sudden takeover request. From the result, it was found that when drivers are distracted from the road or busy with something else, they quickly look back at the road upon receiving a strong warning which could cut through distractions and help drivers focus on the road again.

4.2.2 Alert Perception

Attention Capturing

The results indicated that alert intensity was perceived as more effective when participants were engaged in a typing task compared to when no task was present. Specifically, attention-capturing scores were significantly higher in typing scenarios across both mild and strong alerts. This suggests that alerts need to be more pronounced or intense to capture attention effectively in multitasking environments. The increase in attention-capturing scores from no-task scenarios to typing scenarios, particularly under mild alerts, supports the hypothesis that cognitive load influences alert perception, making participants potentially more responsive to alerts when engaged in a secondary task.

Confidence Levels

Confidence levels increased significantly when participants transitioned from no task to typing tasks, particularly with the dynamic alerts. This increase might indicate that the higher cognitive load during typing tasks requires and consequently elicits a stronger alert to maintain or enhance confidence in task management. This can be interpreted as participants perceiving the alert intensity as appropriate or sufficient under increased cognitive demands.

Annoyance Levels

Despite the variations in alert intensity and cognitive load, annoyance levels remained consistently low, suggesting that the intensity of the alerts was not perceived as excessive. This is critical as it indicates that even with stronger alerts necessitated by higher cognitive states, the alert design did not cross the threshold of being perceived as too intrusive or bothersome. However, this contradicts the second hypothesis that the level of annoyance is higher with a strong alert.

4.2.3 Scenario and Alert Designs

Managing all variables in a driving experiment is complex, particularly when a secondary task is involved, as it is challenging to regulate the level of mental and physical distractions experienced by each individual while driving. In an open-goal typing task where participants were encouraged to prioritize driving, they tended to be more attentive to the road. Consequently, more than half of the participants disengaged the AP and resumed manual control before the Takeover Request (TOR) in Scenario 1, suggesting room for improvement in this scenario. Despite expectations of an early takeover rate in Scenario 1 due to traffic signs, eye-tracking data focusing on attention AOI revealed that participants were still using their phones up to 20 seconds before the TOR, underscoring the authenticity of distracted driving scenarios. The duration of the driving scenarios, lasting between 6-8 minutes, aimed to mirror the cognitive state of highway cruising and minimize the impact of alerts from previous scenarios; however, this duration may inadvertently shift the cognitive focus from concentration to drowsiness in some drivers, potentially influencing reaction times in scenarios without a secondary task.

To summarize, a dynamic alert was perceived differently when investigating the reaction time, as the stronger the alert gets, the more effective in grabbing attention and reducing reaction time when drivers are distracted, which means, when it comes to safety, a stronger alert would be preferred in most cases; however, it still depends on individual preference of how they perceive alert, as well as how urgent or harmful the situation is.

Chapter 5

Conclusion

5.1 Summary

With the emerging technology of autonomous driving, it is very important to understand how humans and machines can work together in the most efficient ways, especially during crucial situations in which lives depend on it. This study aims to investigate the potential benefits of a user-centered design approach using dynamic alert designs which can vary the intensity based on the cognitive state to optimize how the driver perceives the alert and prioritizing thoughts for the best possible outcome.

The key results of this study are summarized as follows:

- **Reaction Times**: Dynamic alerts resulted in significantly faster reaction times compared to single alerts. On average, strong alerts produced faster responses (M = 2.22 s, SD = 1.03) compared to mild alerts (M = 3.65 s, SD = 1.81), with $p < 0.01$.

- **Accelerations**: Higher lateral accelerations were observed under strong alerts, especially in dynamic groups. The greatest values were recorded during strong alerts in the Construction Zone Scenario (S1), indicating more aggressive maneuvering.

- **Physiological Stats**: Heart rate data collected during takeover events were found to be unreliable due to motion artifacts. However, general trends indicated an increased heart rate shortly after TORs compared to cruising periods. Eye-tracking data showed variations in attention to different areas of interest (AOI) before, during, and after the takeover. Participants spent more time looking at the road and dashboard during takeovers in the absence of secondary tasks, while attention was more divided with secondary tasks.

- **Attention Capturing**: Dynamic alerts significantly improved attention-capturing scores, especially under stronger alert conditions and involving NDRT (e.g., typing). This improvement was supported by ANOVA results showing a significant effect, $F = 8.32$, $p = 0.004$.

- **TOR Timing**: Single groups showed a slight better perception of well-timed alerts compared to dynamic groups in self-rated timing on TOR, suggesting that cognitive engagement with secondary tasks might enhance alert responsiveness.

- **Alert Understandability and SA Comprehension**: Strong alerts significantly performed better to be understood by users scores ($p = 0.01$), particularly in single groups without secondary tasks. However, SA comprehension scores varied, typically decreasing in scenarios involving secondary tasks under strong alerts, while improvements were observed only in the dynamic group.

- **Confidence**: Confidence levels surged notably in dynamic groups with secondary tasks, indicating that experiencing dynamic alerts multiple times boosts driver confidence. This change was statistically significant, $F = 12.97$, $p = 0.001$.

- **Annoyance**: Annoyance levels remained consistently low, with slight increases in single groups under mild alerts with secondary tasks. Dynamic alerts were generally found to be less annoying, particularly under strong alert conditions.

- **Overall Satisfaction**: Satisfaction with alert design was higher in scenarios involving strong alerts, reinforcing the effectiveness of strong alert designs in maintaining user satisfaction despite increased task complexity ($F = 5.67$, $p = 0.018$).

What we observed in this study was that changing the alert to be more intense can help the driver in the dynamic group to react faster when distracted compared to only using the strong or mild alert alone in all situations; this suggests that the perception of the alert varied when the alert changed, even if the alert was intended to convey the same information for a takeover. Dynamic alerts, which adjust their intensity based on the situation or the driver's level of engagement, could be the key to maintaining optimal driver awareness, regardless of what is happening inside or outside the vehicle. When investigating the physiological response, all designs in this study gave a similar result in the driver's area of attention during the takeover.

Self-rating scores reinforce the analysis of driver reactions. There was a clear preference for strong alerts, which participants perceived as more effective. Importantly, this preference does not translate into increased annoyance in stronger alerts as hypothesized. Nevertheless, if we take into account the occurrence of repeated alarms, which was not covered in this study, solely with the strong alert with auditory, it could result in increased levels of annoyance [58], thus posing a potential research inquiry in the autonomous driving setting. Concerning the statistical analysis of the self-rating scores, the results indicated that engaging in a secondary task might heighten drivers' sensitivity to external stimuli, making them more receptive to alerts. Moreover, high scores for understandability emphasize the importance of clear and direct alerts in effectively conveying information, which is essential for safe takeover maneuvers. Furthermore, the significant rise in confidence within the dynamic group, where alerts were tailored based on the cognitive state of the driver, suggests that dynamic alerts are particularly effective in boosting driver confidence. This could be due to the adaptive nature of the alerts, which might better align with the drivers' perceived needs and expectations during different driving conditions.

Overall, the results support the hypothesis that the intensity of alerts needs to be stronger in multitasking environments to be perceived as effective. This supports the theory that cognitive load can conceal or weaken the perceived intensity of sensory inputs, necessitating stronger signals to reach the same level of awareness as in less challenging circumstances.

5.2 Limitations, Recommendations, and Future Research Directions

Sample Size and Diversity

The study involved 41 participants, but this sample size may not be sufficient to generalize the findings to a broader population. Although the participants' ages varied, the sample was skewed towards younger individuals. This skewness could impact the generalizability of the results across all age groups. Future studies should strive for a larger and more balanced sample to ensure the findings are representative of different demographics.

Experience Level

While all participants held at least an Ontario G2 driver's license or equivalent, their levels of driving experience and familiarity with advanced driver assistance systems (ADAS) varied. Such variability could influence the results, as more experienced drivers might respond differently to take-over requests compared to less experienced drivers. Future research should consider stratifying participants based on their driving experience and familiarity with ADAS to better understand how these factors affect driver responses.

Alert and Scenario Design

The study did not incorporate false positive alerts, which could have provided a more comprehensive understanding of takeover reactions. Including scenarios where alerts are triggered without an imminent need for a takeover could offer a more nuanced view of driver alertness and response strategies. Furthermore, more noticeable differences between strong and mild alerts may result in a clearer pattern in driver behavior. Overall driving scenarios could be shorter to avoid inducing drowsiness in drivers in the scenarios without secondary tasks, which could help maintain a concentrated cognitive state.

Future studies could benefit from experimenting with more distinct variations in the intensity and modality of alerts, including verbal auditory alerts or tactile alerts, between the strong and mild categories. This approach would help better understand their effectiveness across different cognitive loads. Additionally, research should explore the impact of alert intensity in multitasking environments to confirm the hypothesis that stronger alerts are necessary in such settings to maintain driver awareness. Furthermore, integrating false positive alerts into the experimental design could offer deeper insights into how drivers respond when no immediate takeover is necessary. This would enhance understanding of driver alertness and help refine alert systems to minimize potential distractions or overreactions.

Cognitive State and Secondary Task

This study classifies the driver's cognitive state into two broad categories: a concentrated state and a distracted state, based on the presence of a NDRT. While this simplification facilitates the experimental design and aligns with recent research [10], [48], [50], it does not capture the nuances of individual differences in cognitive load and task handling capabilities, which can significantly affect driver behavior and alert responsiveness. Additionally, the study presumes uniform response capabilities among participants, disregarding variations in driving experience and personal aptitude for multitasking in an autopilot-supported highway scenario.

The choice of selecting texting as a secondary task was made due to the limitation where an in-vehicle touchscreen was not available. The use of the phone as a secondary task lacked stringent controls, leading to variations in the way participants engaged with the device. The timing of phone usage in relation to the TOR alert was not standardized, potentially affecting the consistency of our findings related to driver distraction. More standardized secondary tasks may help to maintain a similar level of distraction among drivers.

To mitigate the variations in how participants engage with secondary tasks, future research should standardize these activities with a specific goal for the secondary task based on the individual capability to maintain a similar workload level. This would help maintain a consistent level of distraction across participants and provide a clearer picture of how secondary tasks impact driver response to alerts.

Investigating the influence of various cognitive states on the quality and speed of driver takeover actions will shed light on the intricate dynamics between human cognitive processes and autonomous vehicle control mechanisms. Future studies should incorporate a more nuanced classification of cognitive states, perhaps by utilizing real-time monitoring tools that can measure physiological and neurological indicators of stress, attention, and workload. This could enable a dynamic adjustment of alert types and intensities based on real-time assessment rather than preset conditions. To effectively tailor alerts and interventions, long-term monitoring of driver behavior and cognitive state through in-vehicle systems could be instrumental. These systems would benefit from machine learning algorithms capable of learning and adapting to individual driver patterns over time.

Physiological Measures

The eye-tracking technology, while insightful, had its limitations. Limited viewing angles and sub-optimal lighting conditions led to periods where tracking data was not available, potentially interfering with the data analysis of drivers' visual focus during critical moments. Most participants felt uncomfortable after wearing the eye tracker for a long period due to a non-adjustable headband, which could obstruct the driving in-vehicle eye tracking system and would be ideal for monitoring eye movements in the car cockpit. Moreover, the data obtained from the heart rate monitors during the takeover events was found to be unreliable for analysis, thus limiting our ability to incorporate physiological responses to alerts and takeover events.

While the average heart rate data collected during cruising periods were generally reliable, significant challenges were encountered during Takeover Request (TOR) scenarios. These challenges primarily arose due to the increased physical movements required for vehicle control, such as steering and manipulating other controls, which introduced substantial noise in the heart rate measurements captured by photoplethysmographic (PPG) sensors.

This interference was particularly problematic during TORs, where precise heart rate measurements are critical for assessing physiological responses to different alert types. The analysis of the HR data revealed significant inconsistencies during TOR events, leading to the exclusion of HR data from all participants and limiting the physiological insights into the drivers' cognitive states and stress levels during the takeover process. Several studies [51], [52] have reported that the E4 device experiences significant missing IBI data due to movements, likely contributing to the observed data smoothing during TOR events. This highlights the need to account for wearable device limitations in dynamic environments. Additionally, analyzing electrodermal activity (EDA) could be useful for future research to provide a more comprehensive understanding of physiological responses during TOR events. However, this analysis was not within the scope of the current study.

The TOR periods, lasting only 5 to 15 seconds, are crucial yet brief windows during which accurate physiological data is essential to understanding driver responses. The short duration means that any inaccuracies in heart rate measurement during these moments can disproportionately impact the interpretation of how alert types influence driver stress and arousal levels. This limitation is noted in similar studies, where motion artifacts significantly affect the accuracy of physiological measurements obtained from wearable sensors during physical activities [52]. For example, Ravindran et al. (2022) discuss how physical movements can lead to interpolated or estimated heart rate readings that may not accurately reflect the true physiological state [51].

Given this context, the potential for inaccurate interpretations of physiological responses during these critical moments is a significant concern. This study's findings regarding physiological responses to alerts must, therefore, be considered with caution. Future research should focus on employing more robust physiological monitoring technologies that can withstand the rigors of active driving tasks or explore alternative methods less prone to such errors. Enhancements in sensor technology or data processing techniques that can more effectively isolate true physiological signals from noise induced by movement could greatly improve the accuracy of findings in similar scenarios.

Driving Simulation

- **Hardware:** In this study, the simulation setup had its drawbacks, with an inactive steering wheel during autopilot mode and counter-intuitive icons on the steering wheel due to limited controls available, which could cause confusion and affect the realism of the driving experience. Furthermore, due to limitations of hardware that only allow angle-mode steering configuration with autopilot, resulting in low torque feedback when steering, most participants found it challenging and needed some time to get used to the sensitive steering wheel.

- **Simulator Sickness:** Immersive simulation caused simulator sickness in some participants. The chance of getting motion sickness could be optimized by increasing screen frame rates and using scenarios with fewer turns.

- **Software:** The simulation software encountered issues with skipped TOR alerts due to high computer processing load, leading to data loss and imbalance in results. Furthermore, the driving scenario editor lacked flexibility with non-removable components such as traffic signs and a scarcity of usable road signs that align with real-world driving scenarios.

- **NPC driving behavior:** The behavior of vehicles of non-player characters (NPC) within the simulation occasionally triggered false alarms, contributing to unnatural driving responses, particularly on curved roads where repeated braking was observed.

- **ACC and AP Features:** The ACC and AP systems were not perfect and exhibited rough handling when the target vehicle was detected in curved lanes, potentially influencing unnecessary takeover and trust level in the automated system, which may affect the outcome.

- **Simulation Realism:** The high-fidelity driving simulator used in this study, while advanced, may not perfectly replicate real-world driving conditions. Responses in a simulated environment can differ from those in actual driving, especially regarding stress and urgency levels. Therefore, the simulator findings may not fully capture the complexities and nuances of real driving experiences. Future research should validate these results with on-road studies to ensure applicability in real-world contexts. Additionally, while the scenarios in this study are designed to be realistic, they might not fully capture the complexity and unpredictability of real-world driving. The frequency and nature of TORs should be more varied and spontaneous to better reflect actual driving conditions. Future research should incorporate a wider range of scenarios to enhance the realism and applicability of the findings.

Refining the simulation setup is crucial for creating a more realistic and engaging driving experience. Enhancements could include active steering feedback during autopilot modes, high frame rate scenarios to reduce simulator sickness, and more flexible scenario editing software that allows for better replication of real-world driving conditions. Real-time classification and analysis of driver cognitive states through advanced monitoring systems could enhance the development of adaptive alert systems tailored to the driver's current state, potentially improving the evaluation of the adaptive alert system.

References

[1] H. Jamson, N. Merat, O. Carsten, and F. Lai, "Fully-automated driving: The road to future vehicles," en, in *Proceedings of the 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design: driving assessment 2011*, Olympic Valley-Lake Tahoe, California, USA¿: University of Iowa, 2011, pp. 2–9. DOI: 10.17077/drivingassessment.1370. [Online]. Available: http://pubs.lib.uiowa.edu/drive/article/id/27534/.

[2] S. International, "Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles," 2014. DOI: 10.4271/j3016_202104.

[3] H. J. Kim and J. H. Yang, "Takeover requests in simulated partially autonomous vehicles considering human factors," en, *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 5, pp. 735–740, Oct. 2017, ISSN: 2168-2291, 2168-2305. DOI: 10.1109/THMS.2017.2674998.

[4] Y. Ma, D. Zhao, and Z. Wang, "Improved ux design of self-autopilot driving takeover audio prompt based on ucd method," en, in *Advances in Usability, User Experience, Wearable and Assistive Technology* (Lecture Notes in Networks and Systems), T. Z. Ahram and C. S. Falcão, Eds., Lecture Notes in Networks and Systems. Cham: Springer International Publishing, 2021, vol. 275, pp. 301–309, ISBN: 978-3-030-80090-1. DOI: 10.1007/978-3-030-80091-8_35. [Online]. Available: https://link.springer.com/10.1007/978-3-030-80091-8_35.

[5] R. M. A. Van der Heiden, J. L. Kenemans, S. F. Donker, and C. P. Janssen, "The effect of cognitive load on auditory susceptibility during automated driving," *Human Factors*, vol. 64, no. 7, pp. 1195–1209, Nov. 2022, ISSN: 0018-7208. DOI: 10.1177/0018720821998850.

[6] K. Zeeb, A. Buchner, and M. Schrauf, "Is take-over time all that matters? the impact of visual-cognitive load on driver take-over quality after conditionally automated driving," en, *Accident Analysis Prevention*, vol. 92, pp. 230–239, Jul. 2016, ISSN: 00014575. DOI: 10.1016/j.aap.2016.04.002.

[7] J. H. Yang, S. C. Lee, C. Nadri, J. Kim, J. Shin, and M. Jeon, "Multimodal displays for takeover requests," en, in *User Experience Design in the Era of Automated Driving* (Studies in Computational Intelligence), A. Riener, M. Jeon, and I. Alvarez, Eds., Studies in Computational Intelligence. Cham: Springer International Publishing, 2022, vol. 980, pp. 397–424, ISBN: 978-3-030-77725-8. DOI: 10.1007/978-3-030-77726-5_15. [Online]. Available: https://link.springer.com/10.1007/978-3-030-77726-5_15.

[8] J. Izquierdo-Reyes, R. A. Ramirez-Mendoza, and M. R. Bustamante-Bello, "A study of the effects of advanced driver assistance systems alerts on driver performance," en, *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 12, no. 1, pp. 263–272, Feb. 2018, ISSN: 1955-2505. DOI: 10.1007/s12008-016-0368-6.

[9] A. P. van den Beukel, M. C. van der Voort, and A. O. Eger, "Supporting the changing driver's task: Exploration of interface designs for supervision and intervention in automated driving," en, *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 43, pp. 279–301, Nov. 2016, ISSN: 13698478. DOI: 10.1016/j.trf.2016.09.009.

[10] J.-H. Han and D.-Y. Ju, "Advanced alarm method based on driver's state in autonomous vehicles," en, *Electronics*, vol. 10, no. 2222, p. 2796, Jan. 2021, ISSN: 2079-9292. DOI: 10.3390/electronics10222796.

[11] S. Agrawal and S. Peeta, "Evaluating the impacts of driver's pre-warning cognitive state on takeover performance under conditional automation," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 83, pp. 80–98, Nov. 2021, ISSN: 1369-8478. DOI: 10.1016/j.trf.2021.10.004.

[12] C. Gold, D. Damböck, L. Lorenz, and K. Bengler, ""take over!" how long does it take to get the driver back into the loop?" *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 57, no. 1, pp. 1938–1942, Sep. 2013, ISSN: 2169-5067. DOI: 10.1177/1541931213571433.

[13] B. Yi, H. Cao, X. Song, S. Zhao, W. Guo, and M. Li, "How to identify the takeover criticality in conditionally automated driving? an examination using drivers' physiological parameters and situational factors," en, *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 85, pp. 161–178, Feb. 2022, ISSN: 13698478. DOI: 10.1016/j.trf.2021.12.007.

[14] J. Wan and C. Wu, "The effects of lead time of take-over request and nondriving tasks on taking-over control of automated vehicles," en, *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 6, pp. 582–591, Dec. 2018, ISSN: 2168-2291, 2168-2305. DOI: 10.1109/THMS.2018.2844251.

[15] E. Šabić, J. Chen, and J. A. MacDonald, "Toward a better understanding of in-vehicle auditory warnings and background noise," en, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 63, no. 2, pp. 312–335, Mar. 2021, ISSN: 0018-7208, 1547-8181. DOI: 10.1177/0018720819879311.

[16] F. Meng and C. Spence, "Tactile warning signals for in-vehicle systems," en, *Accident Analysis Prevention*, vol. 75, pp. 333–346, Feb. 2015, ISSN: 00014575. DOI: 10.1016/j.aap.2014.12.013.

[17] M. AblaBmeier, T. Poitschke, F. W. K. Bengler, and G. Rigoll, "Eye gaze studies comparing head-up and head-down displays in vehicles," in *Multimedia and Expo, 2007 IEEE International Conference on*, IEEE, 2007, pp. 2250–2252, ISBN: 9781424410163.

[18] F. Hajiseyedjavadi, T. Zhang, R. Agrawal, M. Knodler, D. Fisher, and S. Samuel, "Effectiveness of visual warnings on young drivers hazard anticipation and hazard mitigation abilities," en, *Accident Analysis Prevention*, vol. 116, pp. 41–52, Jul. 2018, ISSN: 00014575. DOI: 10.1016/j.aap.2017.11.037.

[19] A. P. van den Beukel, M. C. van der Voort, and A. O. Eger, "Supporting the changing driver's task: Exploration of interface designs for supervision and intervention in automated driving," en, *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 43, pp. 279–301, Nov. 2016, ISSN: 13698478. DOI: 10.1016/j.trf.2016.09.009.

[20] B. D. Seppelt and J. D. Lee, "Keeping the driver in the loop: Dynamic feedback to support appropriate use of imperfect vehicle control automation," en, *International journal of human-computer studies*, vol. 125, pp. 66–80, 2019, ISSN: 1071-5819. DOI: 10.1016/j.ijhcs.2018.12.009. [Online]. Available: https://www-sciencedirect-com.proxy.lib.uwaterloo.ca/science/article/pii/S1071581918301277?ref=pdf_download&fr=RR-2&rr=77890ac29e8ea24c.

[21] W. Liu, Q. Li, Z. Wang, W. Wang, C. Zeng, and B. Cheng, "A literature review on additional semantic information conveyed from driving automation systems to drivers through advanced in-vehicle hmi just before, during, and right after takeover request," en, *International Journal of Human–Computer Interaction*, pp. 1–21, May 2022, ISSN: 1044-7318, 1532-7590. DOI: 10.1080/10447318.2022.2074669.

[22] C. Gonzalez, B. A. Lewis, D. M. Roberts, S. M. Pratt, and C. L. Baldwin, "Perceived urgency and annoyance of auditory alerts in a driving context," en, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, no. 1, pp. 1684–1687, Sep. 2012, ISSN: 2169-5067, 1071-1813. DOI: 10.1177/1071181312561337.

[23] E. Šabić, D. Henning, and J. MacDonald, "Adaptive auditory alerts for smart in-vehicle interfaces," en, *Proceedings of the Human Factors and Ergonomics Society ... Annual Meeting. Human Factors and Ergonomics Society. Annual Meeting*, vol. 63, no. 1, pp. 1545–1549, 2019, ISSN: 1071-1813. DOI: 10.1177/1071181319631404. [Online]. Available: http://dx.doi.org/10.1177/1071181319631404.

[24] J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," *Human Factors*, vol. 46, no. 1, pp. 50–80, Mar. 2004, ISSN: 0018-7208. DOI: 10.1518/hfes.46.1.50_30392.

[25] G. Huang and B. Pitts, "Age-related differences in takeover request modality preferences and attention allocation during semi-autonomous driving," en, in *Human Aspects of IT for the Aged Population. Technologies, Design and User Experience* (Lecture Notes in Computer Science), Q. Gao and J. Zhou, Eds., Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020, vol. 12207, pp. 135–146, ISBN: 978-3-030-50251-5. DOI: 10.1007/978-3-030-50252-2_11. [Online]. Available: http://link.springer.com/10.1007/978-3-030-50252-2_11.

[26] G. D. Gomes, R. Flynn, and N. Murray, "Continuous-time feedback device to enhance situation awareness during take-over requests in automated driving conditions," in *Proceedings of the 13th ACM Multimedia Systems Conference*, New York, NY, USA: ACM, 2022.

[27] C. Geitner, F. Biondi, L. Skrypchuk, P. Jennings, and S. Birrell, "The comparison of auditory, tactile, and multimodal warnings for the effective communication of unexpected events during an automated driving scenario," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 65, pp. 23–33, Aug. 2019, ISSN: 1369-8478. DOI: 10.1016/j.trf.2019.06.011.

[28] D. Yu, C. Park, H. Choi, D. Kim, and S.-H. Hwang, "Takeover safety analysis with driver monitoring systems and driver–vehicle interfaces in highly automated vehicles," en, *Applied Sciences*, vol. 11, no. 1515, p. 6685, Jan. 2021, ISSN: 2076-3417. DOI: 10.3390/app11156685.

[29] Y. Zhang and H. Tan, "Effects of multimodal warning types on driver's task performance, physiological data and user experience," en, in *Cross-Cultural Design. Applications in Cultural Heritage, Tourism, Autonomous Vehicles, and Intelligent Agents* (Lecture Notes in Computer Science), P.-L. P. Rau, Ed., Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021, vol. 12773, pp. 304–315, ISBN: 978-3-030-77079-2. DOI: 10.1007/978-3-030-77080-8_25. [Online]. Available: https://link.springer.com/10.1007/978-3-030-77080-8_25.

[30] J. Edworthy, S. Loxley, and I. Dennis, "Improving auditory warning design: Relationship between warning sound parameters and perceived urgency," en, *Human factors*, vol. 33, no. 2, pp. 205–231, 1991, ISSN: 0018-7208. DOI: 10.1177/001872089103300206. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/1860703/.

[31] C. L. Baldwin and J. F. May, "Loudness interacts with semantics in auditory warnings to impact rear-end collisions," en, *Transportation research. Part F, Traffic psychology and behaviour*, vol. 14, no. 1, pp. 36–42, 2011, ISSN: 1369-8478. DOI: 10.1016/j.trf.2010.09.004. [Online]. Available: https://www-sciencedirect-com.proxy.lib.uwaterloo.ca/science/article/pii/S1369847810000690?via%3Dihub.

[32] D. Marshall, J. D. Lee, and P. A. Austria, "Annoyance and urgency of auditory alerts for in-vehicle information systems," en, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 45, no. 23, pp. 1627–1631, Oct. 2001, ISSN: 2169-5067, 1071-1813. DOI: 10.1177/154193120104502314.

[33] D. C. Marshall, J. D. Lee, and P. A. Austria, "Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness," en, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 49, no. 1, pp. 145–157, Feb. 2007, ISSN: 0018-7208, 1547-8181. DOI: 10.1518/001872007779598145.

[34] C. L. Baldwin and B. A. Lewis, "Perceived urgency mapping across modalities within a driving context," en, *Applied Ergonomics*, vol. 45, no. 5, pp. 1270–1277, Sep. 2014, ISSN: 00036870. DOI: 10.1016/j.apergo.2013.05.002.

[35] I. Politis, S. Brewster, and F. Pollick, "Evaluating multimodal driver displays of varying urgency," in *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13*, New York, New York, USA: ACM Press, 2013, ISBN: 9781450324786.

[36] F. Roche and S. Brandenburg, "Should the urgency of auditory-tactile takeover requests match the criticality of takeover situations?" In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, 2018, pp. 1035–1040, ISBN: 9781728103211.

[37] F. Roche and S. Brandenburg, "Should the urgency of visual-tactile takeover requests match the criticality of takeover situations?" *IEEE transactions on intelligent vehicles*, vol. 5, no. 2, pp. 306–313, 2020, ISSN: 2379-8858. DOI: 10.1109/tiv.2019.2955906. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8911449&tag=1.

[38] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," en, *Human factors*, vol. 37, no. 1, pp. 32–64, 1995, ISSN: 0018-7208. DOI: 10.1518/001872095779049543.

[39] M. L. Matthews, D. J. Bryant, R. D. G. Webb, and J. L. Harbluk, "Model for situation awareness and driving: Application to analysis and research for intelligent transportation systems," en, *Transportation research record*, vol. 1779, no. 1, pp. 26–32, 2001, ISSN: 0361-1981. DOI: 10.3141/1779-04. [Online]. Available: http://dx.doi.org/10.3141/1779-04.

[40] M. R. Endsley, "Situation awareness global assessment technique (sagat)," in *Proceedings of the IEEE 1988 National Aerospace and Electronics Conference*, IEEE, 1988, 789–795 vol.3.

[41] R. Ma and D. B. Kaber, "Situation awareness and workload in driving while using adaptive cruise control and a cell phone," en, *International journal of industrial ergonomics*, vol. 35, no. 10, pp. 939–953, 2005, ISSN: 0169-8141. DOI: 10.1016/j.ergon.2005.04.002. [Online]. Available: https://people.engr.ncsu.edu/dbkaber/papers/Ma_Kaber_IJIE_05.pdf.

[42] M. Baumann, J. Krems, and L. K. Heinrich, "Automation, situation awareness and mental workload," in *Studies in Computational Intelligence*. Cham: Springer International Publishing, 2022, pp. 3–27, ISBN: 9783030777258.

[43] V. Melnicuk, S. Thompson, P. Jennings, and S. Birrell, "Effect of cognitive load on drivers' state and task performance during automated driving: Introducing a novel method for determining stabilisation time following take-over of control," *Accident Analysis Prevention*, vol. 151, p. 105 967, Mar. 2021, ISSN: 0001-4575. DOI: 10.1016/j.aap.2020.105967.

[44] N. Merat, A. H. Jamson, F. C. H. Lai, and O. Carsten, "Highly automated driving, secondary task performance, and driver state," *Human Factors*, vol. 54, no. 5, pp. 762–771, Oct. 2012, ISSN: 0018-7208. DOI: 10.1177/0018720812442087.

[45] S. Yang, J. Kuo, and M. G. Lenné, "Effects of distraction in on-road level 2 automated driving: Impacts on glance behavior and takeover performance," *Human Factors*, vol. 63, no. 8, pp. 1485–1497, Dec. 2021, ISSN: 0018-7208. DOI: 10.1177/0018720820936793.

[46] Y. Yamani, P. Bıçaksız, J. Unverricht, and S. Samuel, "Impact of information bandwidth of in-vehicle technologies on drivers' attention maintenance performance: A driving simulator study," en, *Transportation Research Part F: Traffic Psychology*

and Behaviour, vol. 59, pp. 195–202, Nov. 2018, ISSN: 13698478. DOI: 10.1016/j.trf.2018.09.004.

[47] S. Choe, H.-J. Jeon, and O.-S. Kwon, "Estimating attentional state of a driver: Interacting effects of task demands and cognitive capacities," in *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, Oct. 2018, pp. 999–1004. DOI: 10.1109/ICTC.2018.8539676.

[48] J. Li, Y. Liu, X. Ji, and S. Tao, "Detection of driver's cognitive states based on lightgbm with multi-source fused data," in *WCX SAE World Congress Experience*, SAE International, Mar. 2022. DOI: https://doi.org/10.4271/2022-01-0066. [Online]. Available: https://doi.org/10.4271/2022-01-0066.

[49] A. Misra, S. Samuel, S. Cao, and K. Shariatmadari, "Detection of driver cognitive distraction using machine learning methods," *IEEE Access*, vol. 11, pp. 18 000–18 012, 2023, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2023.3245122.

[50] H. Yang, J. Wu, Z. Hu, and C. Lv, "Real-time driver cognitive workload recognition: Attention-enabled learning with multimodal information fusion," *IEEE Transactions on Industrial Electronics*, vol. 71, no. 5, pp. 4999–5009, May 2024, ISSN: 1557-9948. DOI: 10.1109/TIE.2023.3288182.

[51] K. K. Ravindran, C. D. Monica, G. Atzori, D. Lambert, V. Revell, and D.-J. Dijk, "Evaluating the empatica e4 derived heart rate and heart rate variability measures in older men and women," en, in *2022 44th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, Glasgow, Scotland, United Kingdom: IEEE, Jul. 2022, pp. 3370–3373, ISBN: 978-1-72812-782-8. DOI: 10.1109/EMBC48229.2022.9871559. [Online]. Available: https://ieeexplore.ieee.org/document/9871559/.

[52] H. Stuyck, L. Dalla Costa, A. Cleeremans, and E. Van den Bussche, "Validity of the empatica e4 wristband to estimate resting-state heart rate variability in a lab-based context," *International Journal of Psychophysiology*, vol. 182, pp. 105–118, Dec. 2022, ISSN: 0167-8760. DOI: 10.1016/j.ijpsycho.2022.10.003.

[53] *Vi-dashboard — vi-grade*, en. [Online]. Available: https://www.vi-grade.com/en/products/vi-dashboard//.

[54] *Super cruise for select vehicles*, en-US. [Online]. Available: https://www.chevrolet.com/index/experience-chevrolet/connected-services-nav/supercruise-overview.html.

[55] C. Purucker, F. Naujoks, A. Prill, and A. Neukum, "Evaluating distraction of in-vehicle information systems while driving by predicting total eyes-off-road times with keystroke level modeling," *Applied Ergonomics*, vol. 58, pp. 543–554, Jan. 2017, ISSN: 0003-6870. DOI: 10.1016/j.apergo.2016.04.012.

[56] E. E. Wiese and J. D. Lee, "Auditory alerts for in-vehicle information systems: The effects of temporal conflict and sound parameters on driver attitudes and performance," *Ergonomics*, vol. 47, no. 9, pp. 965–986, Jul. 2004, ISSN: 0014-0139. DOI: 10.1080/00140130410001686294.

[57] J. Radlmayr, K. Brüch, K. Schmidt, C. Solbeck, and T. Wehner, "Peripheral monitoring of traffic in conditionally automated driving," en, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 62, no. 1, pp. 1828–1832, Sep. 2018, ISSN: 1071-1813. DOI: 10.1177/1541931218621416.

[58] F. E. Block, L. Nuutinen, and B. Ballast, "Optimization of alarms: A study on alarm limits, alarm sounds, and false alarms, intended to reduce annoyance," en, *Journal of Clinical Monitoring and Computing*, vol. 15, no. 2, pp. 75–83, Feb. 1999, ISSN: 1573-2614. DOI: 10.1023/A:1009992830942.

[59] J. P. Bliss and S. A. Acton, "Alarm mistrust in automobiles: How collision alarm reliability affects driving," en, *Applied Ergonomics*, vol. 34, no. 6, pp. 499–509, Nov. 2003, ISSN: 00036870. DOI: 10.1016/j.apergo.2003.07.003.

[60] F. Naujoks, C. Mai, and A. Neukum, "The effect of urgency of take-over requests during highly automated driving under distraction conditions," en, p. 8, 2021.

[61] F. A. Kingdom and N. Prins, "Signal detection measures," en, in *Psychophysics*. Elsevier, 2016, pp. 149–188, ISBN: 978-0-12-407156-8. DOI: 10.1016/B978-0-12-407156-8.00006-2. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/B9780124071568000062.

[62] K. A. Hoff and M. Bashir, "Trust in automation: Integrating empirical evidence on factors that influence trust," *Human Factors*, vol. 57, no. 3, pp. 407–434, May 2015, ISSN: 0018-7208. DOI: 10.1177/0018720814547570.

[63] R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," en, *Human factors*, vol. 39, no. 2, pp. 230–253, 1997, ISSN: 0018-7208. DOI: 10.1518/001872097778543886. [Online]. Available: http://dx.doi.org/10.1518/001872097778543886.

[64] T. Helldin, G. Falkman, M. Riveiro, and S. Davidsson, "Presenting system uncertainty in automotive uis for supporting trust calibration in autonomous driving," en, in *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13*, Eindhoven, Netherlands: ACM Press, 2013, pp. 210–217, ISBN: 978-1-4503-2478-6. DOI: 10.1145/2516540.2516554. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2516540.2516554.

[65] F. Ekman, M. Johansson, and J. Sochor, "Creating appropriate trust in automated vehicle systems: A framework for hmi design," en, *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 1, pp. 95–101, Feb. 2018, ISSN: 2168-2291, 2168-2305. DOI: 10.1109/THMS.2017.2776209.

[66] B. M. Muir, "Trust between humans and machines, and the design of decision aids," en, *International Journal of Man-Machine Studies*, vol. 27, no. 5–6, pp. 527–539, Nov. 1987, ISSN: 00207373. DOI: 10.1016/S0020-7373(87)80013-5.

[67] L. Tijerina, M. Blommer, R. Curry, R. Swaminathan, D. S. Kochhar, and W. Talamonti, "An exploratory study of driver response to reduced system confidence notifications in automated driving," en, *IEEE Transactions on Intelligent Vehicles*, vol. 1, no. 4, pp. 325–334, Dec. 2016, ISSN: 2379-8904, 2379-8858. DOI: 10.1109/TIV.2017.2691158.

[68] M. N. Lees and J. D. Lee, "The influence of distraction and driving context on driver response to imperfect collision warning systems," en, *Ergonomics*, vol. 50, no. 8, pp. 1264–1286, 2007, ISSN: 0014-0139. DOI: 10.1080/00140130701318749. [Online]. Available: http://dx.doi.org/10.1080/00140130701318749.

[69] F. Naujoks, A. Kiesel, and A. Neukum, "Cooperative warning systems: The impact of false and unnecessary alarms on drivers' compliance," en, *Accident Analysis Prevention*, vol. 97, pp. 162–175, Dec. 2016, ISSN: 00014575. DOI: 10.1016/j.aap.2016.09.009.

[70] B. A. Lewis, J. L. Eisert, D. M. Roberts, and C. L. Baldwin, "Designing unambiguous auditory crash warning systems," en, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, no. 1, pp. 2078–2082, Sep. 2014, ISSN: 2169-5067, 1071-1813. DOI: 10.1177/1541931214581437.

[71] D. Park, W. C. Yoon, and U. Lee, "Cognitive states matter: Design guidelines for driving situation awareness in smart vehicles," en, *Sensors*, vol. 20, no. 1010, p. 2978, Jan. 2020, ISSN: 1424-8220. DOI: 10.3390/s20102978.

[72] *Typing speed test - master - apps on google play*, en. [Online]. Available: https://play.google.com/store/apps/details?id=com.aswdc_typingspeed&hl=en.

APPENDICES

Appendix A

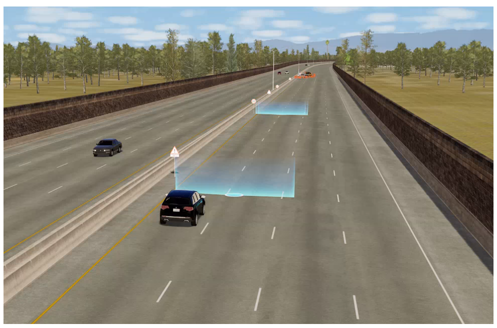
Experiment

A.1 Apparatus

A.1.1 Simulation Hardware

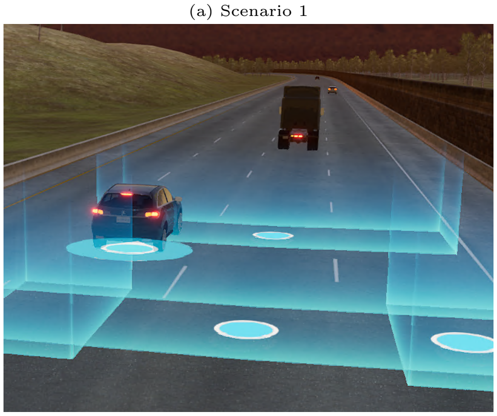
This study utilized a VI-Grade STATIC driving simulator with vehicle shell and equipment based on a 2018 Chevrolet Traverse with screens as mirrors, shown in A.1. The Adaptive Cruise Control (ACC) and autopilot logic were based on VI-WorldSim Offline example Simulink logic with modifications to integrate with the online simulation system.

Figure A.1: Hardware setup of VI-Grade STATIC DriveSim system, from outside vehicle (left), and inside the cockpit (right)



(a) Scenario 1

(b) Scenario 2



(c) Scenario 3

(d) Scenario 4

Figure A.2: Example of VI-WorldSim Scenarios before hazard in each scenario with ego vehicle (black SUV)

A.1.2 Simulation Software

All scenarios used in this study were created using VI-WorldSim Studio with trigger logic to simulate road hazards in each scenario. An example of the hazard in each scenario can be identified in Figure A.2. The dynamic takeover incidents in Scenario 2, 3, and 4 were based on the trigger scripts activated by the ego vehicle while static hazard was used in Scenario 1. All TOR alerts were based on the trigger script activated by the ego vehicle.

A.1.3 Smartphone for the Typing Task

In this study, the NDRT was depicted as a typing task on the phone to simulate distracted driving conditions. The device used was the Samsung Galaxy S10 (SM-G973F) with the Android 12 operating system as shown in Figure 3.1a, equipped with dark-theme Gboard touchscreen keyboard in English (US) QWERTY layout with haptic and popup feedback on keypress. Participants were asked to perform the task based on the Word Practice module in the Typing Speed Test mobile app (v7.9) by ASWDC, Computer Engineering Department at Darshan University [72] as shown in Figure 3.1b with varied word lengths. The hardware was also attached with an eye-tracking marker to its right.

Following is the example of 50-word paragraph from the Word Practice in the Typing Speed Test app:

"wrangle barrage especially silent ceramic drove outlove institute throve above pummels learn warranty advocate column connected roaring engineer immerges mammey recherche humble portrait quantity carbonic checker correct fritz kiaugh climax audio indicate wetproof wheeze kinase retrieval northern matchup thirty kurgan kvases kickshaw landing putoff theory unlikely triac clarion separate somatic knickers wicket from rimmed"

A.2 Questionnaires

A.2.1 Pre-Study Questionnaire

Pre-study questionnaires consist of questions regarding experience and questions regarding driving behavior.

- How familiar are you with driver assistance or autonomous driving technology?

- I am not familiar with it.
- I'm aware of it, but I haven't utilized any driver assistance features before.
- I have utilized driver assistance features in the past, but I don't use them regularly.
- I occasionally make use of driver assistance features while driving.
- I regularly utilize driver assistance features whenever possible.

- Which driver-assistive technologies have you encountered or used?
  - Adaptive cruise control (ACC)
  - Lane-keeping assist (LKA)
  - Automatic emergency braking (AEB)
  - Parking Assist
  - Automatic parking assist
  - Blind-spot monitoring
  - Collision warning system
  - Traffic sign recognition
  - Driver attention monitoring
  - Cross-traffic alert
  - Self-driving cars (Autonomous Level 2 or above)

A.2.2 Motion Sickness Questionnaire (Short-MSSQ)

MMSQ was used to evaluate the likelihood of a participant getting motion sickness during the study so that the researcher could provide any notice or suggestions before the experiment session. Two sections of questions ask about their experience, before the age of 12 and in the last 10 years, if they ever felt sick or nauseated in the following conditions:

- Cars
- Buses
- Yrains

- Sircrafts
- Small boats
- Ships
- Swings or roundabouts in playgrounds
- Big Dippers or Funfair Rides
- Immersive simulator, e.g., 3D/4D, driving/flight simulator

With 4 scales ranging from never felt sick, rarely felt sick, sometimes felt sick, and frequently felt sick. An option was also given if the condition was not applicable or never experienced.

A.2.3 Post-Scenario Questionnaire

Immediately after each scenario, the participants were provided with the post-scenario questionnaire. Questions can be found in A.4, where the alert perception section only appeared after the late scenario.

Cognizplay Experiment Feedback Form - Introduction

Thank you for participating in the Cognizplay experiment on alert design for takeover requests in autonomous driving. Your feedback is crucial for improving the safety and effectiveness of autonomous vehicle systems. Please take a few minutes to share your thoughts and experiences.

Please enter the assigned participant ID (Given by the researcher)

Please verify if you are human.

☐ I'm not a robot — reCAPTCHA Privacy - Terms

1st

Which scenario did you just completed?

○ 1st Scenario ○ 2nd Scenario ○ 3rd Scenario ○ 4th Scenario

Section 1: General

How mentally demanding did you find the secondary task during the simulation?

○ I did not complete the secondary task ○ Extremely Easy ○ Easy ○ Moderate ○ Demanding ○ Extremely Demanding

On a scale of 1 to 5, how well did the visual (message on the dashboard) and sound alerts successfully captured your attention?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

Were the alerts presented in a way that was easy to understand?

○ Strongly disagree ○ Somewhat disagree ○ Neither agree nor disagree ○ Somewhat agree ○ Strongly agree

Did the alert design enhance your ability to quickly comprehend the situation and respond accordingly?

○ Strongly agree ○ Somewhat agree ○ Neither agree nor disagree ○ Somewhat disagree ○ Strongly disagree

How satisfied were you with the timing of the alerts?

○ Too early ○ Slightly early ○ Just right ○ Slightly late ○ Too late

Section 2: Situation Awareness

Did the alerts provide you with timely information about the need to take over control?

○ Strongly disagree ○ Somewhat disagree ○ Neither agree nor disagree ○ Somewhat agree ○ Strongly agree

How aware were you of the surrounding traffic conditions while using the automated driving system?

○ Not aware at all ○ Somewhat aware ○ Moderately aware ○ Very aware ○ Extremely aware

To what extent did the alerts contribute to your overall awareness of the driving situation?

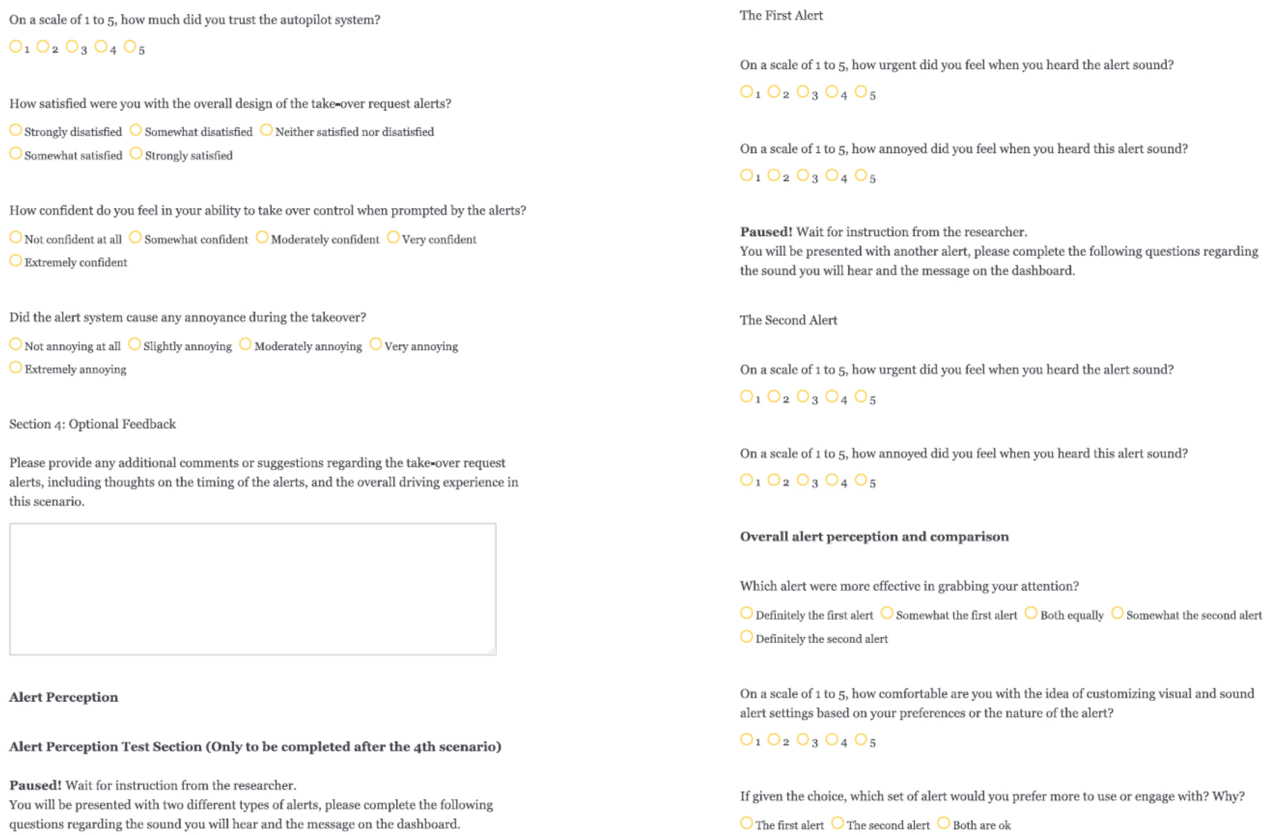
○ Not at all ○ Slightly ○ Moderately ○ Very much ○ Completely

Did you observe any of the following while driving in the last scenario?

☐ Car Crashes
☐ An Ambulance
☐ A Fire Truck
☐ A Moose
☐ A Cyclist

Section 3: Satisfaction

Figure A.3: Post-Scenario Questionnaire (Page 1-2)



On a scale of 1 to 5, how much did you trust the autopilot system?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

How satisfied were you with the overall design of the take-over request alerts?

○ Strongly disatisfied ○ Somewhat disatisfied ○ Neither satisfied nor disatisfied ○ Somewhat satisfied ○ Strongly satisfied

How confident do you feel in your ability to take over control when prompted by the alerts?

○ Not confident at all ○ Somewhat confident ○ Moderately confident ○ Very confident ○ Extremely confident

Did the alert system cause any annoyance during the takeover?

○ Not annoying at all ○ Slightly annoying ○ Moderately annoying ○ Very annoying ○ Extremely annoying

Section 4: Optional Feedback

Please provide any additional comments or suggestions regarding the take-over request alerts, including thoughts on the timing of the alerts, and the overall driving experience in this scenario.

Alert Perception

Alert Perception Test Section (Only to be completed after the 4th scenario)

Paused! Wait for instruction from the researcher.
You will be presented with two different types of alerts, please complete the following questions regarding the sound you will hear and the message on the dashboard.

The First Alert

On a scale of 1 to 5, how urgent did you feel when you heard the alert sound?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

On a scale of 1 to 5, how annoyed did you feel when you heard this alert sound?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

Paused! Wait for instruction from the researcher.
You will be presented with another alert, please complete the following questions regarding the sound you will hear and the message on the dashboard.

The Second Alert

On a scale of 1 to 5, how urgent did you feel when you heard the alert sound?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

On a scale of 1 to 5, how annoyed did you feel when you heard this alert sound?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

Overall alert perception and comparison

Which alert were more effective in grabbing your attention?

○ Definitely the first alert ○ Somewhat the first alert ○ Both equally ○ Somewhat the second alert ○ Definitely the second alert

On a scale of 1 to 5, how comfortable are you with the idea of customizing visual and sound alert settings based on your preferences or the nature of the alert?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

If given the choice, which set of alert would you prefer more to use or engage with? Why?

○ The first alert ○ The second alert ○ Both are ok

Figure A.4: Post-Scenario Questionnaire (Page 3-4)