Evaluation of Driver Situational Awareness in ADAS Speed Limit Recognition

Insights from Eye Tracking and Behavioural Analysis

Thesis Project

Yining Chen

Delft University of Technology

TU Delft

Evaluation of Driver Situational Awareness in ADAS Speed Limit Recognition

Insights from Eye Tracking and Behavioural Analysis

by

Yining Chen

Student Name	Student Number
Yining Chen	5722896

Chair:	Dr. ir. Haneen Farah
Daily Supervisor:	Dr. Yan Feng
Company Supervisor:	ir. Maria Oskina
External Examiner:	Dr.ir. Yke Bauke Eisma
Project Duration:	March 2024 - March 2025
Faculty:	Civil Engineering and Geosciences (CEG)

Preface

The completion of this master's project would not have been possible without the support and help of many people. This project was conducted as part of my internship at Royal HaskoningDHV, and I am deeply grateful for the opportunities and resources provided by the company. Additionally, I sincerely appreciate the support from TU Delft, which has been instrumental throughout my academic journey. I would like to take this opportunity to express my deepest gratitude to everyone who has contributed to this project.

First and foremost, I would like to extend my heartfelt thanks to Maria, my company supervisor, who has selflessly provided immense support throughout this journey. I still vividly remember how she was always present at the experimental site ahead of time, ensuring everything was well-prepared. She kept me on track with timely reminders and provided valuable feedback during our regular meetings. Her enthusiasm and kindness have been instrumental in driving my project forward smoothly.

I am also deeply grateful to my academic supervisor, Yan, whose guidance has been invaluable throughout my two years at TU Delft. From my research internship to this master's thesis, her support has been indispensable. Reflecting on my time at the MXR lab, I cherish the numerous opportunities I had to participate in lab exhibitions, which were invaluable experiences for me. I sincerely appreciate all her efforts and encouragement.

My gratitude also goes to Haneen, the chair of my thesis committee. She has always been there to offer help whenever I needed it. I fondly remember our collaboration during the 183rd anniversary of TU Delft, where we worked together to present and demonstrate the driving simulator. Working with her has always been an enjoyable and rewarding experience.

A special thanks to my colleagues at Royal HaskoningDHV, especially Yvo, Evert, and Sander, along with all the colleagues who participated in our experiments. Their time and effort have greatly contributed to the success of this research, and I am truly grateful for their support and involvement.

Finally, I want to thank my friends and family. Yixin and Honghao, every moment we shared during our academic journey at TU Delft is a treasured memory—filled with hard work, occasional struggles, but most importantly, joy and growth. Zhaoyu, Yiman, and Haodong, although we have had fewer interactions after our second year, the time spent with you remains incredibly precious to me. Zhenlin, your support in VR-related discussions has been immensely helpful and greatly appreciated. To my parents and my younger brother, who are thousands of miles away yet have always been my silent pillars of support, I am forever grateful. Their unwavering belief in me has been my driving force.

With all my gratitude, I will continue moving forward. Because of all of you, my efforts have never felt lonely. I am truly fortunate to have met such wonderful people along the way.

Yining Chen
Delft, March 2025

Summary

Introduction

Advanced Driver Assistance Systems (ADAS) have evolved significantly over the years, offering a variety of features designed to enhance road safety and elevate driver comfort. These systems, incorporating technologies such as adaptive cruise control, automatic emergency braking, and lane departure warning, providing critical support to drivers in managing complex driving environments. Despite their potential, challenges such as standardization, driver trust, and over-reliance on automation continue to hinder their widespread adoption. Among these challenges, ADAS warning signals come in various designs, each playing a pivotal role in ensuring effective communication between the system and the driver, ultimately influencing situational awareness (SA) and driving performance. SA consists of three levels: perception, comprehension, and projection. It serves as a key framework for evaluating driver interaction with ADAS signals, providing a structured approach to assessing their effectiveness in real-world driving scenarios.

Research Gap

While previous research has explored various aspects of ADAS functionality, there remains a gap in understanding how different ADAS signal designs impact SA across its three levels. Existing studies often focus on individual components of SA, neglecting the interplay between perception, comprehension, and projection. Additionally, the majority of studies lack a comprehensive evaluation framework that integrates subjective assessments, physiological measurements, and task-based performance metrics to capture the full spectrum of driver awareness. Furthermore, driver variability in responding to ADAS signals, influenced by factors such as experience, cognitive capacity, and environmental conditions, remains unexplored. Addressing these gaps is crucial for developing ADAS warning systems that effectively support driver decision-making and enhance road safety.

Research Questions

The study defined the primary research question and its associated sub-questions as follows:

- How does driver situational awareness differ when responding to existing road speed limit alerts generated by different ADAS systems?
    1. How do different ADAS design-related features influence driver situational awareness across its three levels: perception, comprehension, and projection?
    2. How do internal factors, such as age, gender, and other demographic characteristics, contribute to variations in driver situational awareness under different ADAS signal designs?
    3. Based on the results, what design-related features might explain the differences in driver situational awareness and guide the development of more effective ADAS signal designs?

Research Methods

Data collection was conducted through a real-world driving task in Amersfoort, covering a route of 4.3 km. The methods were designed to align with the specific characteristics of each level of SA to ensure a comprehensive and accurate analysis. The independent variable in this study was the ADAS system, which corresponded to different signal designs, while the dependent variables were the indicators representing each SA level. To compare the differences between ADAS systems, we used two different brands, ADAS System A and ADAS System B. The most significant difference in signal design between the two systems is that System A provides more information, displaying both the current road segment's speed limit and the speed limit of the upcoming segment, as shown in the figures.

(a) ADAS System A Speed Limit Alert Design

(b) ADAS System B Speed Limit Alert Design

Figure 1: Comparison of ADAS Speed Limit Alerts for System A and System B

For situation awareness level 1 (SA1), visual attention indicators were measured using image processing techniques. Specifically, the YOLOv8 object detection algorithm was employed to identify and track eye fixations on ADAS signals, providing insights into drivers' attention allocation. For situation awareness level 2 (SA2), self-reported questionnaire responses were aggregated and averaged to obtain an overall comprehension score for each participant, serving as an indicator of their understanding of the ADAS signals. Regarding situation awareness level 3 (SA3), speed-related indicators were analyzed using computational methods to detect deviations and patterns in driving speed data. These indicators reflected drivers' compliance with speed limits and their responses to ADAS alerts.

The analysis methods applied to the processed data included descriptive statistics, significance tests (such as t-tests and Wilcoxon rank-sum tests), and correlation analysis to explore the relationships between different SA levels. Linear Mixed Models (LMM) were utilized in two ways: first, to analyse the effect of the ADAS system (System A vs. System B) on SA1 and SA3 indicators by treating the system as a fixed effect while accounting for participant, order, and turn as random effects. Here, order refers to the sequence in which participants experienced the two ADAS systems, allowing us to account for potential learning effects. Turn represents the four key observation points along the route where speed limit changes occurred, ensuring that system effects were evaluated consistently at predefined locations.

For SA2, given the ordinal nature of the questionnaire data, an ordered logistic regression model was applied to evaluate the impact of internal factors and system design on driver comprehension.

Results

However, specific indicators reveal notable differences. In SA1, *Fixation Count per Turn* varies between the systems, as indicated by the results of t-tests and Wilcoxon rank-sum tests. In SA2, *Scores*, which are derived from the self-reported questionnaire and represent drivers' comprehension of the situation, also show significant differences based on these statistical tests. These findings suggest that System A may help drivers reduce cognitive load and better understand the provided alerts.

Additionally, internal factors, particularly familiarity with ADAS technology and driving experience, were found to have a significant influence on SA. Participants with more driving experience exhibited shorter fixation durations, indicating reduced cognitive load in SA1, while familiarity with ADAS systems was associated with faster response times and more efficient visual processing. Familiarity also showed a marginal influence on SA2 comprehension, suggesting that more familiar participants may evaluate ADAS alerts more critically. Age displayed a near-significant effect, with younger participants tending to react more quickly. Another key finding is the weak or non-existent correlation between SA levels, which suggests that drivers rely not only on ADAS signals but also on external factors such as road signs and personal driving experience when forming SA.

The differences observed between the results obtained from significance tests and those from LMMs highlight the importance of adopting multiple analytical approaches to achieve a more comprehensive understanding of driver behaviour. Using a combination of methods allows for a more nuanced interpretation of the data, capturing trends that might be overlooked when relying on a single analytical technique.

Conclusion

This study provides valuable insights into how different ADAS signal designs influence driver SA. While most indicators showed no statistically significant differences between System A and System B, System A consistently demonstrated better performance across several key aspects, such as visual attention and comprehension. This suggests that the design of System A's alerts provides clearer information, potentially making it easier for drivers to process speed limit changes accurately and efficiently.

Despite these findings, several limitations should be acknowledged. First, the sample size in this study was relatively small, which may limit the generalisability of the results. A larger sample would not only allow for more statistically robust conclusions but also provide a better representation of the natural variations in driver behaviour, experience levels, and cognitive abilities.

Second, the selection of SA indicators, while covering key aspects of driver awareness, could be expanded to include additional measures. A broader set of indicators may help identify further meaningful effects of ADAS signals on driver SA and provide deeper insights into how these systems influence situational processing.

Third, this study was conducted in a real-world driving environment rather than a more controlled setting. While this approach enhances ecological validity, it also introduces external factors that were not directly controlled or analysed in this study. Elements such as roadside speed limit signs and environmental cues may have influenced driver SA independently of ADAS signals, making it challenging to isolate the precise effects of different ADAS designs. Future studies conducted in controlled environments could help mitigate these influences and allow for a more precise evaluation of ADAS signal effectiveness.

In addition to addressing the limitations mentioned above, future research could explore how repeated exposure to different ADAS signal designs influences SA over time, providing insights into drivers' long-term adaptation and behavioural changes. Expanding the study to incorporate additional ADAS functionalities, such as lane-keeping assistance and adaptive cruise control, would help assess how multiple systems interact to support SA. Furthermore, conducting experiments in controlled environments alongside real-world studies could help isolate the effects of ADAS signals more precisely, offering a clearer understanding of their role in enhancing driver awareness and decision-making.

The findings of this study offer practical implications for ADAS design and implementation. The results suggest that providing more comprehensive speed limit information can enhance driver situation awareness. Specifically, anticipatory speed limit information may help drivers better prepare for speed changes in advance, reducing cognitive load and improving response efficiency.

Beyond the implications for ADAS design, this study also highlights the value of a layered SA-based research approach. By structuring the evaluation of driver interaction with ADAS signals across three levels of SA, this method offers a systematic way to assess how different signal designs influence attention allocation, comprehension, and decision-making. Future research on ADAS effectiveness and driver support systems can benefit from this structured framework, ensuring a more comprehensive understanding of how these systems contribute to driving performance and road safety.

Contents

List of Figures

List of Tables

1
Introduction

Overview

This chapter provides an introduction to the research by outlining the context, scope, and key objectives of the study. It begins with the research background (Section 1.1), which highlights the advancements in Advanced Driver Assistance Systems (ADAS) and the challenges related to their adoption, particularly in the context of warning signal design and driver situational awareness (SA). Section 1.2 defines the scope of the study, focusing on the evaluation of ADAS warning signals and their impact on SA. Next, the research problems in (Section 1.3) are identified, emphasising the critical gaps in understanding how signal design influences driver performance. Finally, the research questions (Section 1.4) are presented to guide the study, targeting the differences in driver responses to warning signals and the design features contributing to these variations. Together, these sections set the foundation for the methodological framework and analyses discussed in the subsequent chapters.

1.1. Research Background

The evolution of Advanced Driver Assistance Systems (ADAS) represents a significant technological advancement in the field of transportation safety and vehicular automation. The development of ADAS can be traced back to the mid-20th century, with the introduction of basic driver support features such as cruise control, which allowed vehicles to maintain a constant speed without manual throttle adjustment. Over the decades, advancements in sensor technology, computer vision, and machine learning have enabled the development of more sophisticated ADAS features, including adaptive cruise control (ACC), automatic emergency braking (AEB), lane departure warning (LDW), and blind-spot detection (BSD) [1]. These advancements were driven by growing concerns over traffic safety, increased demand for driver comfort, and the automotive industry's shift toward vehicle electrification and automation. Key milestones include the integration of radar and lidar sensors for enhanced environmental perception, the use of real-time video analysis for lane-keeping assistance, and the development of predictive algorithms that enable proactive interventions. By the 2010s, many of these features were integrated into commercial vehicles, with major manufacturers such as Tesla, BMW, and Toyota incorporating Level 2 ADAS in their flagship models [2]. The ultimate goal of ADAS development has been to improve road safety, reduce driver workload, and pave the way for higher levels of vehicular autonomy.

Despite the rapid progress in ADAS development, its widespread adoption faces several challenges that hinder its full potential [3]. One of the most pressing issues is the lack of standardization in ADAS design and functionality across different manufacturers. While systems like adaptive cruise control and lane-keeping assistance are broadly available, their operational logic, signal design, and driver interaction models vary significantly. This lack of uniformity increases the cognitive load on drivers, especially those who switch between vehicles with different ADAS implementations. Another challenge is driver acceptance and trust. Many drivers remain sceptical of ADAS systems due to concerns over their reliability and the fear of relinquishing control to an automated system. Incidents involving high-profile failures of semi-autonomous systems have further furled public distrust. Building driver confidence requires not only technological reliability but also effective user education and system transparency. A third critical challenge is driver over-reliance on ADAS. Research indicates that some drivers become overly dependent on ADAS features, leading to inattentiveness or "automation complacency." This over-reliance can result in slower reaction times during emergencies, especially in situations where manual intervention is required. Therefore, balancing automation with driver engagement remains a critical design challenge for system developers. Legal and regulatory issues also pose a challenge to the adoption of ADAS. Different countries have different regulations governing the use of autonomous and semi-autonomous systems, which complicates the certification and deployment of ADAS-equipped vehicles in multiple markets. Inconsistent safety assessment protocols and liability questions in case of system failure further complicate regulatory acceptance. Lastly, cost and affordability play a significant role. The advanced sensors (such as lidar and high-definition cameras) and computational systems required for ADAS significantly increase production costs. These costs are often transferred to consumers, making ADAS-equipped vehicles less accessible to budget-conscious buyers. Addressing these challenges requires a holistic approach that considers technological development, regulatory alignment, and consumer education.

The design of ADAS warning signals plays a crucial role in ensuring effective human-machine interaction and, ultimately, traffic safety. These signals act as the primary medium of communication between the system and the driver, ensuring that critical information is conveyed efficiently. Different ADAS systems utilize varying designs for warning signals, including visual, auditory, and haptic feedback, each with unique advantages and limitations. Such variability in design profoundly impacts driver SA and reaction times, both of which are critical for avoiding potential hazards [4]. Auditory signals, such as beeps, chimes, or voice alerts, are commonly employed to draw attention but are susceptible to issues like alarm fatigue or ambiguity in interpretation. Haptic feedback, such as seat vibrations or steering wheel alerts, provides tactile prompts that are particularly useful in noisy environments where auditory signals may be ineffective. Meanwhile, visual signals, including flashing icons or text on dashboards, remain the most widely utilized and intuitive form of communication [5]. However, poorly designed warning signals—regardless of modality—can lead to confusion, delayed reactions, or even reduced trust in the system. Understanding the impact of these designs on driver behavior and evaluating their effectiveness under varying conditions are therefore essential for optimizing ADAS functionality. This necessity forms the foundation of the current research, which emphasizes the importance of systematically comparing different ADAS warning signal designs through the lens of SA as a key metric.

SA has become a critical focus in understanding driver behaviour and performance within complex systems like ADAS, particularly as drivers interact with increasingly automated technologies. In the context of ADAS, SA reflects the driver's ability to perceive critical information from the system, comprehend its implications, and anticipate future developments. This cognitive process is directly shaped by the design of ADAS warning signals, which serve as the primary interface for communicating system-generated alerts. Unlike traditional usability metrics that focus on superficial attributes such as colour, shape, or duration, SA provides a more comprehensive framework for evaluating the cognitive demands placed on drivers. Effective ADAS signals should enhance a driver's ability to quickly recognize and prioritize critical information, fostering a heightened state of awareness in dynamic and high-pressure scenarios. By analysing how different signal designs influence a driver's SA, this research aims to assess their effectiveness in supporting real-world decision-making and aligning with cognitive processes. This approach provides a structured and insightful basis for comparing ADAS signals, ensuring they are intuitive, impactful, and conducive to safer driving experiences.

Measuring SA is a crucial step in understanding driver behaviour and assessing the effectiveness of ADAS warning signal designs. SA measurement methods are generally classified into three main categories: subjective assessment, physiological measurement, and task-based performance metrics. Each of these methods offers unique insights into driver cognitive states and provides a distinct perspective on how well ADAS designs support SA.

Subjective assessment methods rely on self-reported data from drivers, typically collected through questionnaires or interviews after a driving task. These methods allow participants to reflect on their awareness and cognitive workload, providing qualitative insights into their perception of the ADAS signals. subjective measures are widely used for their simplicity and cost-effectiveness, but they may be influenced by memory recall biases or individual differences in self-reporting accuracy.

Physiological measurement methods capture changes in physiological signals, such as heart rate, eye movement (eye-tracking), pupil dilation, and brain activity (EEG). These indicators provide real-time, objective data on a driver's cognitive load and attentional focus. For instance, eye-tracking can reveal where the driver is focusing their attention, while pupil dilation can indicate cognitive workload or stress levels. These methods are valued for their objectivity and granularity, but they often require sophisticated equipment and may be intrusive to drivers.

Task-based performance metrics assess SA through direct observation of driver actions, reaction times, or task completion rates. By evaluating how quickly and accurately a driver responds to ADAS signals or unexpected changes in the driving environment, researchers can infer the level of SA. These measures are considered practical and reflective of real-world driving performance, but they often require the design of controlled experiments or simulated driving scenarios.

In our research, we adopt a comprehensive approach that incorporates a combination of these measurement methods to capture a holistic view of driver SA. Each method targets specific layers of SA, ensuring a more granular and precise assessment. for instance, subjective assessment provides insight into the driver's perception and comprehension of the system's alerts, while physiological measurements offer real-time tracking of cognitive load and attentional focus. task-based performance metrics, on the other hand, reflect how effectively drivers can act on the information presented by adas warning signals. by integrating these approaches, we aim to measure multiple layers of sa, from the perception of information to the driver's ability to predict and react to changes in the environment. This multifaceted measurement strategy allows for a more robust assessment of adas signal design, supporting a deeper understanding of how signal presentation impacts driver cognitive states, decision-making, and driving performance.

1.2. Research Scope

This research focuses on evaluating how ADAS speed warning signals influence driver SA and provides insights grounded in real-world applications. By selecting two ADAS systems from actual commercial brands, the study ensures that its findings are not only theoretically significant but also practically relevant to current ADAS technologies. This approach allows for an exploration of real-world signal designs and their impact on driver cognition, offering actionable insights for system developers and industry stakeholders.

The primary focus of this study is to compare the differences in driver SA elicited by two distinct ADAS warning signal designs. Using SA as the evaluation framework, the research emphasizes understanding how signal design influences the levels of perception, comprehension, and projection. The comparison highlights the nuanced ways in which different warning signals affect driver awareness, with particular attention to the cognitive processes underlying these effects.

To maintain a controlled and focused investigation, this study is specifically designed around the speed limit alert function of ADAS systems, which falls under the broader category of driver assistance and safety alert functions. This targeted approach allows for the isolation of the impact of speed warning signals on driver SA, avoiding the complexities introduced by analysing multiple ADAS functionalities simultaneously.

By adopting this scope, the research not only sheds light on the variability in driver SA across different signal designs but also identifies potential design factors that contribute to these differences. The practical significance of the study lies in its ability to bridge theoretical understanding with real-world applications, providing guidance for the future development and optimization of ADAS warning systems.

1.3. Research Problems

Despite the widespread adoption of ADAS, the methods for accurately understanding and measuring drivers' SA under different ADAS warning signals remain uncertain. While existing studies have explored various approaches to SA evaluation, there is a lack of clarity on how to comprehensively assess SA across all three levels: perception, comprehension, and projection. This uncertainty presents a significant challenge in developing a robust framework for evaluating the cognitive state of drivers interacting with ADAS.

Another critical issue is the variability in SA among drivers when exposed to the same ADAS signals. Drivers may exhibit significant differences in their SA depending on factors such as individual cognitive capacity, experience, and driving conditions. These variations underscore the importance of understanding not only how SA differs across individuals but also how these differences influence their interaction with ADAS systems and overall driving performance.

Finally, the underlying reasons behind these differences in SA require further investigation. Potential contributing factors may include the design characteristics of ADAS warning signals, environmental and situational variables, and driver-specific attributes. Identifying these causal relationships is essential for advancing the design of ADAS systems that are adaptable to diverse driver needs and capable of enhancing SA consistently across varied contexts.

1.4. Research Questions

This study is guided by the following primary research question and its associated sub-questions:

- How does driver situational awareness differ when responding to current road speed limit alerts generated by different ADAS systems?
    1. How do different ADAS signal designs influence driver situational awareness across its three levels: perception, comprehension, and projection?
    2. How do internal factors, such as age, gender, and other demographic characteristics, contribute to variations in driver situational awareness under different ADAS signal designs?
    3. Based on the results, what design-related features might explain the differences in driver situational awareness and guide the development of more effective ADAS signal designs?

The research questions aim to investigate the differences in driver SA under two ADAS designs and identify the internal and design-related factors contributing to these differences.

2
Literature Review

Overview

This chapter explores the impact of Advanced Driver Assistance Systems (ADAS) warning signal designs on driver situational awareness (SA) and aims to provide actionable insights for system improvements. Section 2.1 examines the functionalities of ADAS, with particular attention to speed limit transition warnings, a critical feature for enhancing SA and road safety. Section 2.2 introduces the concept of SA, outlining its three-level framework—perception, comprehension, and projection—and emphasising its relevance to dynamic driving environments. Section 2.3 discusses current methods for evaluating SA in transportation, categorising them into physiological measurements, memory probe methods, task performance indicators, and subjective evaluations. Section 2.4 provides a detailed review of the specific SA evaluation methods used in this study, focusing on eye-tracking for level 1 SA, memory probe techniques for level 2 SA, and task performance metrics for level 3 SA. Section 2.5 identifies key research gaps, including the need for systematic SA measurement frameworks across all levels and the lack of standardised metrics for evaluating ADAS effectiveness, which form the basis for this study's research objectives. Finally, Section 2.6 outlines the research objectives, detailing how this study addresses these gaps and contributes to the advancement of SA assessment and ADAS signal design.

2.1. Advanced Driver Assistance Systems (ADAS)

2.1.1. Overview of ADAS

ADAS have made significant strides due to rapid technological advancements utilizing radar, lidar, and camera technologies. These systems range from simple alerts that use visual, auditory, or haptic signals to more complex mechanisms that actively intervene during critical driving situations. As car manufacturers increasingly integrate ADAS, these systems, though generally classified under Level 2 assisted driving systems, exhibit considerable variability in their operational design and functionality [6].

The classification of ADAS functionalities can be based on their intended purposes, such as safety improvement or driver convenience [7]. Information assistance systems provide drivers with critical data to support decision-making without intervening in vehicle control. Examples include traffic sign recognition (TSR), navigation aids, and speed limit warnings. These systems primarily aim to enhance drivers' SA by improving their ability to perceive and comprehend the driving environment. Warning and alert systems, on the other hand, detect potential hazards and notify the driver, with features such as lane departure warnings, forward collision alerts, and blind spot monitoring being prominent examples. These systems aim to reduce accidents by prompting timely driver responses.

Active control systems take intervention a step further by actively mitigating risks or preventing collisions. Functions such as automatic emergency braking (AEB), adaptive cruise control (ACC), and lane-keeping assistance fall into this category [7]. While safety remains a priority, some ADAS features focus on driver comfort, such as automated parking, adaptive headlights, and traffic jam assist, which help reduce driver workload in complex situations. Finally, integrated systems combine multiple functionalities, such as forward collision warnings paired with automatic braking, to deliver comprehensive safety and comfort solutions.

The design of HMI within these systems is particularly diverse, especially in how emergency warning signals are delivered to drivers. Despite the proliferation of studies on ADAS, much of the research primarily centres on how different drivers react to the same ADAS system, variations in system performance across different vehicles, and how these systems behave under various environmental conditions. For instance, some studies have highlighted that while certain ADAS technologies improve driver safety, their effectiveness can vary significantly based on the design of the interface and the driver's ability to understand and respond to the warnings.

2.1.2. Speed Limit Transition Warnings

Speed limit transition warnings are a sub-category of information assistance systems that notify drivers of changes in speed limits when transitioning between road segments. Given their importance in ensuring road safety, particularly in areas where speed limits change frequently—such as school zones, construction sites, and urban-rural boundaries—this study specifically focuses on their impact on driver SA. By facilitating timely recognition and comprehension of these changes, speed limit transition warnings directly affect the first two levels of SA—perception and comprehension—and play a pivotal role in driver decision-making.

Investigating speed limit transition warnings provides a targeted analysis of how signal design affects drivers' cognitive processes. Existing literature suggests that effective ADAS designs should enhance all three levels of SA but notes that many systems fall short, particularly in helping drivers project future states based on current information [8]. Furthermore, variations in the design of these warnings, such as auditory signals, visual displays, or multimodal combinations, may significantly impact their effectiveness. By focusing on this functionality, the research aims to provide insights into how HMI design influences SA, contributing to the development of more effective ADAS solutions.

2.2. Understanding Situational Awareness

Various conceptualizations of SA have been proposed in academic literature, addressing SA at the individual [9], team [10], and sociotechnical system levels [11]. The formal definition of SA is often described as three ascending levels [12]:

- Perception of the elements in the environment

- Comprehension or understanding of the situation
- Projection of future status

SA serves as a crucial metric for evaluating drivers' performance, particularly in navigating complex and dynamic environments safely [13]. It encompasses drivers' ability to accurately perceive their surroundings and adjust their interactions with distracting elements [14]. Enhanced SA enables drivers to detect potential hazards more effectively during their journeys [15]. In analysing driving behaviour, researchers often explore its correlation with the likelihood of being involved in traffic accidents, considering a range of driver actions and decisions behind the wheel [16].

2.3. Current Methods for Evaluating SA in Transportation

Differences in the research orientations and fields of SA lead to significant variations in measurement methods, with multiple researchers proposing different classifications [17]. However, measurements of SA can generally be classified into four main categories: (1) Physiological Measurements, (2) Memory Probe Measurements, (3) Task Performance Measurements, and (4) Subjective Measurements [18]. In this section, we will highlight several measurement methodologies that may offer valuable insights for the present study.

2.3.1. Physiological Measurements

Physiological measurements provide a direct quantification of bodily responses, offering insights into an operator's cognitive and emotional state which are indicative of their SA. Techniques such as heart rate variability, electroencephalography (EEG), and eye-tracking are commonly utilized. Eye-tracking, for example, measures where and how long a driver or pilot focuses, which can be critical in understanding attention allocation in dynamic environments. Studies have shown that such physiological indicators can effectively predict SA under various operational conditions, highlighting their relevance for safety-critical tasks in transportation [19, 20, 21].

2.3.2. Memory Probe Measurements

Memory probe techniques assess SA by querying operators about specific elements of their environment or task at random intervals. This method assumes that accurate recall of information correlates with high SA. The Situation Present Assessment Method (SPAM) is one such technique where participants are intermittently asked to report the status of relevant variables or conditions in their environment. These snapshots of situational data can then be analysed to gauge the accuracy and speed of their perceptions, which are indicative of their SA [22]. The Situation Awareness Global Assessment Technique (SAGAT) method, widely used in the field, is another excellent example of Memory Probe Measurements [23]. This technique leverages the capability to pause current tasks in a simulated environment. At randomly determined intervals, which are commonly referred to as task breaks, participants are prompted with questions. During these intervals, the simulation temporarily hides information about the current environment from the participant. The accuracy of the participants' responses to these questions serves as the basis for evaluating their SA. This method effectively isolates the participant's memory and awareness of situational elements, thus providing a direct measure of SA without the interference of ongoing task performance.

2.3.3. Task Performance Measurements

Task Performance Measurements offer a practical and indirect method to evaluate SA by analysing how operators handle specific tasks that mirror real-world responsibilities. These measurements, often conducted through simulator-based assessments, focus on the operator's ability to respond to predefined scenarios that replicate critical decision-making situations, such as pilots managing unexpected events during simulations—a robust indicator of their SA [24]. The strength of this approach lies in its objectivity and non-intrusiveness. By relying on quantifiable task outcomes such as accuracy and completion time, it provides a reliable proxy for assessing an operator's SA without disrupting their regular duties. This method integrates smoothly into the natural workflow, utilizing existing simulations and operational environments for regular assessments. Such seamless integration is particularly valuable in high-stakes fields like aviation and driving, where continuous evaluation of SA is crucial [25, 26]. Ultimately, Task Performance Measurements facilitate an authentic and efficient assessment process, making them a cornerstone for ongoing safety and performance evaluations in dynamic operational settings [27, 28].

2.3.4. Subjective Measurements

Subjective methods involve self-assessment scales and questionnaires where individuals report their perception of their own SA. Although subjective, these measures are invaluable for capturing personal insights into cognitive processes and the perceived effectiveness of SA-supporting systems. The Situational Awareness Rating Technique (SART) is one widely used tool that provides a comprehensive measure of awareness based on factors like complexity, variability, and familiarity of the task [29, 30].

2.4. Detailed Measurement Approaches for SA at Three Levels

2.4.1. Level 1 Measurements: Eye-Tracking

Given that the initial stage of SA relies on perception, and considering that most cockpit information is visual, analysing eye movements is a viable method for assessing SA. This approach is supported by the assumption that monitoring an operator's eye movements can effectively gauge their SA [31]. Xu et al. [32] specifically evaluated the first stage of SA through the distribution of visual attention. Eye tracking enables researchers to directly observe how visual information is processed and attended to during tasks, providing an objective way to connect gaze patterns with perception-related behaviours.

This approach provides the unique benefit of allowing real-time assessments with minimal interference [33]. Unlike retrospective methods that rely on participants' memory or verbal explanations, eye tracking captures immediate responses, reducing biases caused by memory decay or interpretation errors. It addresses issues like participants relying on long-term memory, a problem often seen with freeze-probe methods, to explain subtasks or provide detailed information about specific tasks [28]. Furthermore, the continuous nature of eye-tracking data enables researchers to assess dynamic changes in attention allocation, making it well-suited for environments where task demands fluctuate over time [34, 35].

To date, eye tracking is one of the most commonly used techniques in SA research [36]. Its popularity stems from its versatility in capturing perceptual data across a variety of domains, including aviation, driving, and control room operations [37, 38, 39]. The ability to correlate gaze direction, fixation duration, and scan patterns with cognitive processes has proven invaluable for understanding how individuals interact with their environment [40, 41]. Additionally, technological advancements in wearable eye-tracking systems have made it possible to gather high-fidelity data even in naturalistic settings [42, 43].

Even though there are a few examples of it being used to measure level 3 SA [44, 45, 46], eye tracking is typically used to measure level 1 SA. The reason lies in its direct alignment with perception, which forms the foundation for higher SA levels. Therefore, in this research, we employ eye tracking to assess the first stage of SA, focusing on how operators perceive and attend to critical visual information during interaction with ADAS systems.

2.4.2. Level 2 Measurements: Memory Probe

Post-assessment evaluations and questionnaires are frequently employed for accessing SA. These memory probe measurements often facilitate in-depth insights into the operator's SA, providing detailed information on various aspects.

The SAGAT stands out as the predominant method utilized for measuring SA through questionnaires [47]. This approach freezes work tasks at a designated moment, prompting individuals to respond to situation-specific questionnaires. Discrepancies between the responses and the actual situation indicate the individual's level of SA [48]. It can address the SA issues of three levels, thereby determining the extent of SA at three levels for individuals. However, this intrusive method is considered to have a certain impact on the subjects. Moreover, this method is not suitable for field testing because it requires the task to be frozen, and therefore it is usually cited in simulated environments.

Therefore, in our project, we aim to design a scale similar to SAGAT to measure the second level of SA (perception) of the subjects, but this scale will be provided to the subjects after the experiment ends, instead of freezing the task during the experiment. Doing so has the advantage of not interfering with and affecting the field tests, ensuring safety. Additionally, setting questions for the second level SA can enhance efficiency and yield more accurate results.

2.4.3. Level 3 Measurements: Task Performance

The Task Performance method is typically non-invasive, as it does not interfere with ongoing tasks and can be automatically documented, making it highly suitable for real-time and naturalistic settings [49]. This method is widely recognized for its ability to provide objective and quantifiable data on task outcomes, such as response accuracy, reaction times, and task efficiency [50]. By avoiding subjective biases often associated with self-reporting methods, task performance measures allow for a more direct assessment of cognitive processes related to SA.

In this research, we aim to utilize task performance indicators to assess subjects' Level 3 SA, which relates to their ability to project future states and take proactive actions based on the information available. Specifically, we focus on driving performance metrics such as the average speed, the deviation from the designated speed limit, and the percentage of time the vehicle remains within the speed limit. These indicators provide valuable insights into whether participants notice and respond appropriately to speed limit changes, reflecting their anticipation and decision-making capabilities. This approach ensures a more comprehensive evaluation of Level 3 SA in real-world driving scenarios.

Additionally, the use of task performance data allows for seamless integration with other SA measurement methods, such as eye-tracking for level 1 SA and questionnaires for level 2 SA, to form a comprehensive evaluation framework. This combined approach ensures that we capture not only the subjects' actions but also the underlying cognitive processes driving these actions, offering a holistic understanding of SA. The response time metric, in particular, has been validated in prior research as an effective measure of action readiness in driving and other high-stakes operational environments [51, 52].

By focusing on task performance as an objective and scalable method, this research contributes to the broader effort of refining SA measurement techniques and improving our understanding of driver-ADAS interactions.

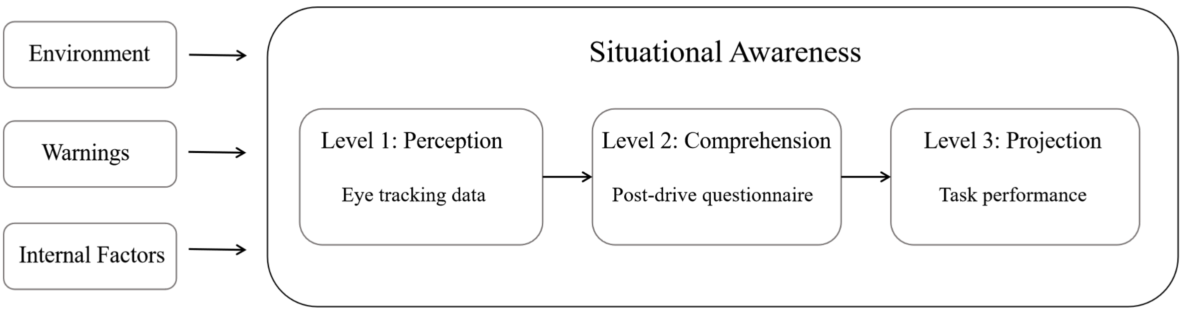


Figure 2.1: Situational Awareness Measurements

2.5. Research Gaps

In light of the findings discussed above, several research gaps can be identified. First, while there are numerous methods available for monitoring SA, there is a notable lack of research focusing on the systematic exploration of each SA level—perception, comprehension, and projection—individually. This gap often leads to incomplete evaluations of SA, making it difficult to fully understand how different design factors influence each cognitive layer of awareness during driving tasks.

Second, although ADAS have been widely implemented, there remains no standardized framework for evaluating their effectiveness. This absence of clear evaluation metrics hinders the identification of specific design elements that contribute to their success or failure. SA offers a measurable and structured approach to assess the impact of ADAS on driver cognition, providing a potential pathway for understanding and improving system design.

To address these gaps, we propose a framework that integrates multiple measurement techniques to evaluate each level of SA independently within the context of driving tasks. By doing so, this research aims to provide a comprehensive assessment of how drivers interact with ADAS and how different SA levels are affected. This approach will enable us to identify key differences in system designs and propose specific directions for improvement based on SA outcomes.

Furthermore, this study seeks to explore the contribution of internal factors, such as age and gender, to variations in SA during ADAS-assisted driving tasks. Understanding how these demographic characteristics influence SA under different ADAS conditions can provide valuable insights into personalizing system designs to accommodate diverse driver populations. Ultimately, this research aims to bridge the gap between theoretical understanding of SA and practical advancements in ADAS, contributing to the development of safer and more effective driver assistance systems.

2.6. Research Objectives

The primary objective of this research is to analyse how different ADAS warning signal designs influence driver SA and to identify design-related factors contributing to observed differences. By conducting real-world experimental analyses of two existing ADAS systems, this study offers a unique and practical exploration of their warning signal designs and their impacts on driver SA.

First, the research aims to measure and evaluate driver SA across its three levels: perception, comprehension, and projection. This study employs specific methods for each level, including eye-tracking metrics for perception, questionnaires for comprehension, and task-based performance measures for projection, to provide a comprehensive assessment of SA under different ADAS signal designs.

Second, the research seeks to investigate the variability in driver SA between the two ADAS signal designs. This analysis highlights the unique cognitive and behavioral impacts of each signal, offering insights into the effectiveness of their design features in promoting driver SA.

Finally, the study explores potential design-related factors that contribute to differences in driver SA. By focusing on the characteristics of the two ADAS signal designs, the research identifies critical elements that influence SA, providing valuable guidance for optimizing future warning systems.

3

Experiment Design

Overview

This chapter provides a detailed description of the experiment design, structured into six sections. Section 3.1 outlines the experimental objectives and variables, including the independent variable, dependent variables, and moderating variables. Section 3.2 discusses participant recruitment and demographics, highlighting the sample composition and recruitment process. Section 3.3 describes the experimental route and environment, focusing on route design, safety measures, and the structure of the questionnaires. Section 3.4 presents the equipment used for data collection, including eye-tracking glasses, GPS devices, and electronic questionnaires, and explains their integration into the experiment. Section 3.5 details the experimental procedure, covering pre-test preparation, the driving task, and post-test assessment. Finally, Section 3.6 addresses ethical considerations, ensuring compliance with research standards.

3.1. Experimental Objectives and Variables

This study investigates how different ADAS system designs influence driver SA across its three hierarchical levels: perception (Level 1), comprehension (Level 2), and projection (Level 3). To achieve this, a field test was conducted on a predefined route, focusing on four road segments where speed limits changed, referred to as *observation areas*. Each *observation area* corresponds to a section of road following a speed limit transition, where drivers needed to adapt their behaviour to the new speed regulation.

Data collection and analysis were conducted specifically within these *observation areas*, capturing speed variations, eye-tracking data, and questionnaire responses. By examining driver responses across these segments, the study aims to assess the effectiveness of different ADAS signal designs in supporting SA at all three levels.

3.1.1. Independent Variable

The independent variable in this study is the ADAS system design, represented by two distinct systems (System A and System B). These systems differ in their warning signal designs and information delivery methods, and they are hypothesized to affect SA differently across its three levels.

3.1.2. Dependent Variables

The dependent variables correspond to the evaluation of SA at its three hierarchical levels. Each level is assessed through specific indicators derived from experimental data.

Level 1 SA (Perception) Level 1 SA reflects the driver's ability to detect and focus on relevant environmental information. The study uses eye-tracking technology to measure visual attention as participants approach and enter *observation areas*. The key indicators for this level include:

- Fixation count: The number of times the participant's gaze fixates on relevant areas (e.g., speed limit signs) within an *observation area*. Data are recorded for each *observation area* and aggregated across all areas.
- Fixation duration: The cumulative duration of all fixations on relevant areas within an *observation area*, measured in seconds. Data are recorded for each *observation area* and aggregated across all areas. In this study, the influence of current driving speed on fixation duration was not considered. Since the driver's speed may vary within a road segment, accounting for speed-dependent variations in fixation duration would significantly increase computational complexity.
- Time to first fixation: The time taken for the participant to first fixate on relevant areas after entering an *observation area*, measured in seconds. The moment of entering an *observation area* is defined as the time when the speed limit transition signals are displayed. Data are recorded for each *observation area*.

Level 2 SA (Comprehension) Level 2 SA pertains to the driver's understanding of the situation and its implications. To assess this level, a questionnaire-based approach is used, with participants providing subjective scores after completing the driving task. The primary indicator is:

- Questionnaire scores: Aggregated subjective scores reflecting the participant's comprehension of the situation across all *observation areas*.

Level 3 SA (Projection) Level 3 SA involves the driver's ability to predict future states based on current information. GPS data are used to analyse driving behaviour through the following indicators:

- Average speed relative to the speed limit: The deviation of the participant's average speed from the designated speed limit, recorded for each *observation area*.
- Speed compliance percentage: The percentage of time the vehicle's speed remains within the speed limit, recorded for each *observation area*.

3.1.3. Moderating Variables

To account for individual differences, internal factors such as age, gender, driving experience, ADAS familiarity, and the order of participation in ADAS system trials (System A first or System B first) are included as moderating variables. These factors are hypothesized to influence SA outcomes under different ADAS designs, providing insights into how demographic characteristics, prior exposure to ADAS technology, and test order affect driver-ADAS interactions.

3.1.4. Controlled Conditions

All data were collected within a single day to minimize the effects of temporal variability, such as differences in weather, lighting, or road conditions. This approach ensured that participants experienced consistent external conditions throughout the experiment, reducing confounding variables and enhancing the reliability of the results.

Table 3.1: Variables and Indicators

Variable Type	Variable Name	Indicators / Measurement
Independent Variable	ADAS Systems (System A & System B)	N/A (Two ADAS systems compared)
Dependent Variable	Level 1 SA (Perception)	Time to first fixation (per area)
		Fixation count (per area and total)
		Fixation duration (per area and total)
	Level 2 SA (Comprehension)	Questionnaire scores (total only)
	Level 3 SA (Projection)	Average speed relative to the speed limit (per area)
		Speed compliance percentage (per area)
Moderating Variable	Internal Factors	Age
		Gender
		Driving Experience
		ADAS Familiarity
		Order

Table 3.2: Indicators with Descriptions and Units

Indicator	Description	Unit
Time to first fixation (per area)	Time taken for the participant to first fixate on relevant areas after entering an observation area	Seconds (s)
Fixation count (per area and total)	Total number of fixations on relevant areas (per observation area and aggregated)	Count
Fixation duration (per area and total)	Total duration of fixations on relevant areas (per observation area and aggregated)	Seconds (s)
Questionnaire scores (total only)	Subjective scores representing participants' comprehension of situations across all observation areas	Subjective Score
Average speed relative to the speed limit (per area)	Deviation of average speed from the designated speed limit (per observation area)	km/h (Speed difference)
Speed compliance percentage (per area)	Percentage of time the vehicle's speed remains within the speed limit (per observation area)	Percentage (%)

3.2. Participants and Recruitment

This study involved the recruitment of participants to evaluate the influence of different ADAS systems on driver SA across its three hierarchical levels. The target sample size was 15 participants; however, due to the complexity of the study design and logistical constraints, a total of 11 participants were successfully recruited.

The participants were diverse in terms of age and gender, with a nearly equal gender distribution, ensuring balanced representation. Specific selection criteria included the following:

- Participants were required to possess a valid driver's license and have at least one year of driving experience to ensure familiarity with real-world driving scenarios.
- No prior requirement for ADAS experience was imposed, as the study aimed to include drivers with varying levels of familiarity to reflect diverse user groups.
- Participants needed to be available for a single-day session to ensure consistency in experimental conditions.

Despite the relatively small sample size, it is deemed sufficient for this exploratory study. The study employed a within-subjects design, where each participant interacted with both ADAS systems. This design enhances the statistical power by reducing variability across participants and allows for a more controlled comparison of the two systems.

Prior to the experiment, all participants were provided with detailed information about the study and signed informed consent forms. Ethical considerations were prioritised throughout the study, and multiple measures were implemented to ensure participant safety. All participants were required to have at least one year of driving experience to minimise risks associated with novice drivers. Before the experiment, they were fully informed about the details of the driving task, ensuring they understood the procedure and potential challenges. The experiment was conducted on roads with relatively low traffic volume to reduce external hazards, and participants were given the option to stop the experiment at any time if they felt uncomfortable or unsafe. Additionally, a researcher was present in the back seat throughout the drive, actively monitoring the situation to identify potential issues and assist in terminating the experiment if necessary. These precautions collectively ensured a safe testing environment for all participants.

3.3. Experimental Route and Environment

3.3.1. Route Design

The driving route for this experiment was carefully designed to evaluate the influence of different ADAS systems on driver SA across varying speed limit conditions. The total length of the route is 4.3 kilometres, starting and ending at the Royal HaskoningDHV premises in Amersfoort, Utrecht Province, Netherlands. The route is divided into two sections: a 0.8 kilometre familiarity segment and a 3.5 kilometre formal experimental segment. The figure below provides an annotated illustration of the experimental route, which is based on Google Maps.

Familiarity Segment The familiarity segment starts at Point A (Royal HaskoningDHV) and ends at Point B, covering a distance of 0.8 kilometres. This section allows participants to get accustomed to the vehicle, ADAS systems, and experimental setup. The speed limit for this segment is 60 km/h. Before the driving task begins, participants also receive an explanation of the ADAS systems and a briefing on the driving task. Further details on this process are provided in Section 3.5.

Formal Experimental Segment The formal test recording begins at Point B and follows the sequence described below, covering four distinct speed limit transition points. Each segment between two points represents an *observation area*, where data on speed adaptation and situational awareness were collected:

- **B to C:** 1.2 kilometres with a speed limit of 50 km/h (*Observation Area 1*).
- **C to D:** 0.8 kilometres with a speed limit of 30 km/h (*Observation Area 2*).
- **D to E:** 0.4 kilometres with a speed limit of 50 km/h (*Observation Area 3*).
- **E to F:** 0.25 kilometres with a speed limit of 30 km/h (*Observation Area 4*).

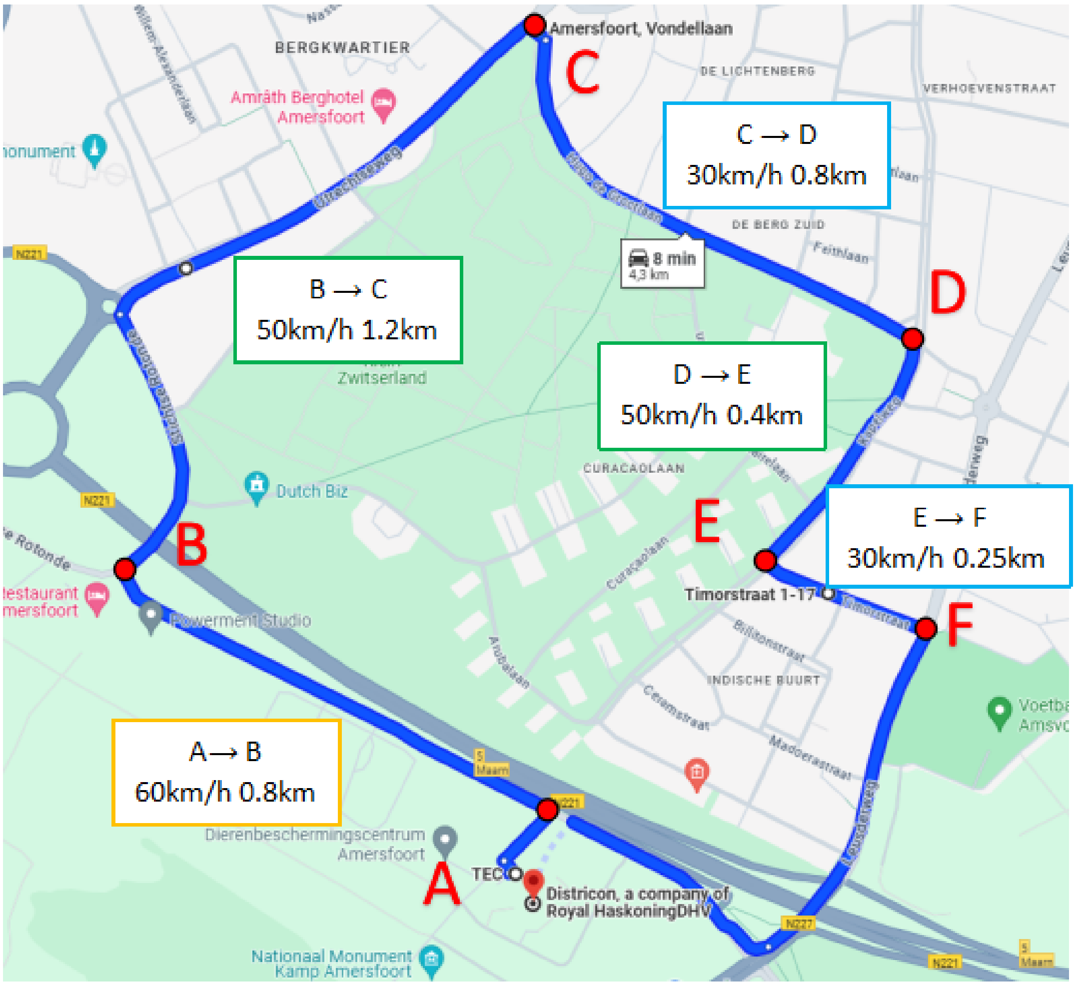


Figure 3.1: Route Design Illustration

This structured route provides opportunities to assess drivers' responses to speed limit transitions under real-world conditions. Each speed limit change point represents an observation area where data on SA indicators are collected.

Table 3.3: Route Segments and Lengths

Section	Route	Length (km)
Familiarity Segment	A-B	0.8
Formal Experimental Segment	B-C	1.2
	C-D	0.8
	D-E	0.4
	E-F	0.25

The experiment was conducted on urban roads in Amersfoort, Utrecht Province, Netherlands. The selected route included typical urban road characteristics such as clear signage, moderate traffic, and varying speed limits. While the weather was not strictly controlled, the experiment was scheduled and conducted under stable and favourable conditions to minimize potential disruptions caused by rain, fog, or strong winds.

Prior to the formal experiment, the responsible researcher and an additional licensed driver conducted pre-tests using both ADAS systems under similar environmental conditions. This pre-test ensured that the vehicles and systems operated as intended and that the route was appropriate for data collection. The results confirmed the ADAS systems' functionality, the clarity of speed limit transitions, and the suitability of the selected indicators.

To minimise interference with participants' driving behaviour, navigation was entirely conducted using Google Maps, which provided both audio and visual directions. A trained researcher, the same as mentioned earlier, was seated in the rear seat, responsible for unobtrusively recording data and providing assistance only when necessary. This setup preserved the naturalistic driving conditions essential for the study while ensuring real-time monitoring.

Each driving task lasted approximately ten minutes per vehicle, resulting in a total driving duration of around 20 minutes for both vehicles. In addition, the pre-drive preparation and questionnaire took approximately ten minutes, and the post-drive questionnaire required another ten minutes. Thus, the total experiment time for each participant was approximately 40 minutes.

All driving sessions were completed within a single day to maintain consistent lighting, traffic, and environmental conditions. The use of a closed circular path starting and ending at the Royal HaskoningDHV premises ensured logistical convenience for participants and minimised external variability.

3.3.2. Experimental Safety Measures

Safety was a primary concern throughout the experiment. The responsible researcher in the rear seat monitored all driving tasks and was prepared to intervene if necessary, though their role was primarily observational to avoid distracting the driver. Participants were required to wear eye-tracking glasses during the experiment, and data collection was conducted unobtrusively to maintain natural driving behaviour.

Participants were instructed to drive as they normally would while adhering to traffic rules. Breaks were provided between the two driving sessions to prevent fatigue and maintain focus. These measures, combined with the results of the pre-tests, ensured a safe and controlled experimental environment.

3.3.3. Questionnaire Design

The questionnaire was divided into two parts: a pre-drive questionnaire and a post-drive questionnaire. Its purpose was to collect basic demographic and contextual information and to evaluate Level 2 SA (Comprehension). All subjective assessment questions were structured using a five-point Likert scale to facilitate subsequent statistical analysis [53]. The complete questionnaire was preserved digitally and is available for review upon request.

Pre-Drive Questionnaire The pre-drive questionnaire collected demographic information and baseline data on participants, including:

- Age group.
- Gender.
- Driving experience (in years).
- Familiarity with ADAS systems.

In addition to collecting key demographic and baseline information, the pre-test questionnaire included auxiliary questions to better understand participants' familiarity with and attitudes toward ADAS technologies. These questions explored their general understanding of various ADAS functionalities, such as speed limit alerts, lane departure warnings, and adaptive cruise control, without revealing which specific functionality was the focus of the study to avoid biasing their responses. The questionnaire also assessed participants' willingness to adopt and use ADAS in their personal vehicles, as well as their overall openness to integrating new technologies into their driving routines.

These additional questions provided valuable context about participants' prior experiences and perceptions of ADAS, which could influence their interaction with the systems during the experiment.

Post-Drive Questionnaire The post-drive questionnaire was administered after each driving session and focused on evaluating participants' comprehension of the driving scenarios. The questions were structured to capture:

- Recognition of speed limit changes and the sequence of transitions along the route.
- Understanding and interpretation of ADAS warning signals in both driving sessions.
- Self-assessment of situational awareness, specifically the ability to understand the current driving environment during each session.

In addition to the main Level 2-related questions, the post-drive questionnaire included supplementary questions that addressed other aspects of situational awareness across its hierarchical levels. These questions explored whether participants noticed speed limit changes and ADAS warning signals during each session, as well as their confidence in predicting the impact of ADAS alerts on future driving behaviour. This approach ensured that the questionnaire considered all three levels of SA in a coherent and structured manner.

The questionnaire also asked participants about their preferences for one of the two ADAS systems and their perceived overall helpfulness of these systems in supporting driving tasks. These auxiliary questions were included to provide additional insights into participants' experiences with the ADAS systems and their usability.

Both questionnaires were presented in electronic format, allowing participants to complete them on a tablet or smartphone immediately after each driving session. This approach ensured timely and accurate data collection while minimizing recall bias.

3.4. Equipment

This section outlines the equipment used in the experiment for data collection across the three levels of SA. The tools were selected to ensure high data accuracy and minimal interference with participants' natural driving behaviour.

3.4.1. ADAS Systems

As shown in the figure, the two vehicles differed in both display layout and signal design. System A follows a more traditional layout, with the primary display positioned behind the steering wheel, while a separate screen on the right (centrally located between the driver and front passenger) provides navigation and additional information. Its speed limit signals present both the current road segment's speed limit and the speed limit of the upcoming segment.

In contrast, System B features a more innovative design, where the display behind the steering wheel seamlessly extends into the central screen, creating a unified interface. Unlike System A, its speed limit signals only show the current road segment's speed limit without providing information about the next segment.

(a) ADAS System A

(b) ADAS System B

Figure 3.2: Illustration of ADAS System Designs

3.4.2. Eye-Tracking Glasses

Level 1 SA (Perception) was measured using the Tobii Pro Glasses 3, a widely used eye-tracking system known for its reliability and precision. These lightweight, unobtrusive glasses recorded gaze data throughout the experiment, providing detailed information on visual attention, including fixation counts, fixation duration, and time to first fixation. The Tobii Pro Glasses 3 were chosen based on their demonstrated improved accuracy compared to earlier models, ensuring robust data quality [54]. The system recorded gaze data at 30 frames per second (fps), meaning data points were collected approximately every 0.03 seconds. All recordings were processed and analysed using the official software provided by Tobii for post-experiment data extraction.

Figure 3.3: Tobii Pro Glasses 3

3.4.3. Questionnaire

An electronic questionnaire system, implemented through Microsoft Forms, was used to collect data related to Level 2 SA (Comprehension) as well as participants' basic demographic information. The digital format minimized manual entry errors and facilitated seamless data integration into the analysis framework.

Once data collection was completed, all responses were consolidated into a unified dataset for analysis. To ensure participant confidentiality, unique identifiers were assigned to each participant, and all data were anonymized during processing. This approach safeguarded sensitive information while maintaining the integrity of the dataset for systematic evaluation.

3.4.4. GPS-Based Speed Tracking

For Level 3 SA (Projection), the GoPro Hero 7 camera was used to record GPS data, including vehicle position and speed. The device collected data every 0.055 seconds, ensuring a continuous record of vehicle dynamics. This data was synchronized with predefined *observation areas* to calculate indicators such as average speed relative to the speed limit and speed compliance percentage. The entry into each *observation area* was determined based on the recorded GPS coordinates, ensuring precise alignment of speed data with the respective road segments. The GoPro device was mounted securely inside the vehicle to ensure stable and accurate data collection.

All equipment was carefully configured to minimize interference with participants' natural driving behaviour, ensuring a seamless and safe experimental experience.

Figure 3.4: GoPro Hero 7 Camera

3.5. Experimental Procedure

This section outlines the step-by-step procedure followed during the experiment, from pre-test preparation to post-experiment procedures. The procedure was designed to ensure consistency across participants and accurate data collection for evaluating SA. In addition to participant procedures, this section also highlights the researchers' responsibilities before, during, and after the experiment.

3.5.1. Pre-Test Preparation

Before the driving task, the research team first ensured that all equipment and vehicles were properly prepared for the experiment. This involved verifying the functionality of the ADAS systems in both vehicles, calibrating the systems, and confirming that all experimental equipment, including the Tobii Pro Glasses 3 and GoPro Hero 7 cameras, was correctly configured and operational. These checks were conducted to ensure the smooth execution of the driving sessions and to prevent any technical issues during the experiment.

Once the equipment and vehicles were confirmed to be in proper working condition, participants were gathered in a designated waiting area. They were provided with the Subject Information Statement and Informed Consent Form. A researcher explained the purpose of the study, the experimental procedure, and participants' roles in detail, ensuring they fully understood the tasks involved. Participants were given the opportunity to ask questions before completing the consent form and pre-drive questionnaire.

After the pre-test preparation, participants were guided to the experimental vehicle. Inside the vehicle, the responsible researcher assisted each participant in wearing the Tobii Pro Glasses 3 eye-tracking system. The glasses were calibrated individually using a standard Tobii calibration process, during which participants were instructed to fixate on a central target point to ensure accurate gaze data collection. Simultaneously, the GoPro Hero 7 camera was activated to begin recording GPS data.

Participants were also given time to familiarize themselves with the vehicle in the parking lot at the starting point before beginning the formal experiment. This ensured they were comfortable with the vehicle and its controls, minimizing potential distractions during the driving task.

3.5.2. Driving Task

Each participant was required to complete two driving sessions, corresponding to the two ADAS systems (System A and System B). The order of the systems was alternated among participants, with an even 1:1 distribution between those starting with System A and those starting with System B. This ensured balanced testing while recording the actual sequence for analysis.

Participants drove along the predefined route described in the experimental design section, which included four observation areas (speed limit transition points). The driving performance in the segments between these observation areas served as the primary focus for data collection and analysis. Participants were instructed to drive as they normally would, following general traffic regulations and adhering to their usual driving habits.

During the driving task, the Tobii Pro Glasses 3 recorded eye-tracking data, while the GoPro Hero 7 logged GPS data. The researcher in the rear seat ensured that all devices functioned correctly while providing minimal assistance to reduce interference with participants' natural driving behaviour.

3.5.3. Post-Test Assessment

After completing each driving session, participants were asked to complete a post-drive questionnaire, which captured their comprehension of speed limit transitions and understanding of ADAS warnings. In addition to evaluating comprehension, the questionnaire also included auxiliary questions exploring their overall impressions of the systems, preferences between the two ADAS designs, and perceived helpfulness of ADAS in driving tasks. The electronic format of the questionnaire minimized errors and allowed for direct integration into the analysis framework.

3.5.4. Post-Experiment Procedures

At the conclusion of the experiment, all collected data, including eye-tracking recordings, GPS logs, and questionnaire responses, were securely stored and organized for analysis.

The Tobii Pro Glasses 3 and GoPro Hero 7 cameras were returned to Royal HaskoningDHV, and the ADAS-equipped vehicles were returned to their respective owners. The research team conducted a final review to verify that all data and equipment were accounted for, ensuring a smooth transition to the data analysis phase.

3.6. Ethical Considerations

This study was conducted in accordance with ethical guidelines and was approved by the Human Research Ethics Committee (HREC) of TU Delft. The ethical approval process included a comprehensive review of the study's methodology, participant involvement, and data management practices, ensuring compliance with institutional and international standards for research involving human subjects. A detailed data management plan was also developed to safeguard participant information and ensure data integrity.

Prior to participation, all participants were provided with an Informed Consent Form, which outlined the scope of the study, the types of data being collected, and the purposes for which the data would be used. The form also explained the anonymization procedures and assured participants of their right to withdraw from the study at any time without penalty. This process ensured that participants were fully informed and voluntarily agreed to participate in the research.

To protect participants' privacy, each participant was assigned a unique identifier based on their experimental group and sequence. For instance, participants were labeled as "Brand_A_1" or "Brand_B_2" to reflect their assigned ADAS system and the order of participation. This anonymization process ensured that no personally identifiable information was associated with the final dataset, and individual participants could not be identified during analysis or dissemination of results.

By adhering to these ethical and data management practices, and through transparent communication with participants via the Informed Consent Form, the study ensured participant confidentiality and compliance with all relevant ethical standards.

4

Data Processing and Analysis

Overview

This chapter focuses on the methodologies used for processing and analysing the data collected during the experiment. It begins with an introduction to the raw data in Section 4.1, providing an overview of the datasets gathered from the eye-tracking system, GPS logs, and questionnaire responses. Section 4.2 describes the data processing procedures, organized into subsections for each level of situational awareness (SA), detailing the methods used to extract meaningful indicators from the raw data. Section 4.3 presents the analytical approaches applied to the processed data, including descriptive statistics, significance testing to compare indicators between the two ADAS systems, and linear mixed models to explore the effects of internal factors on SA outcomes. This chapter serves as the foundation for the discussions in Chapter 5.

4.1. Introduction to Raw Data

This section focuses on the raw data obtained from three primary sources: eye-tracking data, GPS data, and questionnaire responses, all of which were recorded and stored in formats suitable for subsequent processing and analysis.

The eye-tracking data, collected using the Tobii Pro Glasses 3, included MP4 video recordings and corresponding .gz files containing detailed gaze-related information. The GPS data, recorded using the GoPro Hero 7 camera, were processed using the GoPro official Telemetry Extractor tool to extract time-stamped vehicle coordinates and speed data synchronized with the observation areas. The questionnaire responses were stored in electronic format via Microsoft Forms, capturing participants' answers and submission timestamps.

The eye-tracking data and GPS-based speed data were analyzed separately to evaluate SA at different levels. Since these analyses were conducted independently, frame-by-frame synchronization was not required. However, both data sources included timestamps, allowing us to select data points with the same starting timestamp to ensure temporal alignment. This ensured that all analyses were based on data collected simultaneously, providing consistency across driving sessions.

These datasets form the foundation for the subsequent processing and analysis steps, ensuring comprehensive coverage of all three levels of SA and enabling a robust evaluation of the research objectives.

4.2. Data Processing

4.2.1. Processing Level 1 SA (Eye-Tracking Data)

Dynamic AOI Identification in Video Frames

The raw data collected using the Tobii Pro Glasses 3 were imported into the Tobii Pro Lab software for processing. Using this software, gaze data were overlaid onto video recordings, generating annotated videos where red circles represented fixation points, and red lines indicated the movement trajectories of fixations (see Figure 4.1). While the fixation trajectories provide additional information, our study focused exclusively on fixation-related indicators.

Figure 4.1: Annotated video frame showing fixation points (red circle) and gaze trajectories (red lines).

At this stage, the system accurately recorded fixation coordinates and displayed them in the video with no additional errors beyond those inherent to the equipment itself. However, to proceed with the analysis of fixation-related indicators, it became necessary to identify the AOIs relevant to the study.

A critical challenge arises because the Tobii Pro Glasses 3 are worn on the participant's head, and natural head movements cause the AOI positions on the dashboard display to shift dynamically. Unlike a fixed-coordinate video recording, the glasses record from a perspective that moves along with the participant's head, resulting in continuous changes in the apparent position of AOIs on the dashboard display, as illustrated in Figures 4.2a and 4.2b. These figures demonstrate how the same AOI appears in different locations within the recorded video due to head movements.

(a) First View of the Dashboard Display

(b) Second View of the Dashboard Display

Figure 4.2: Examples of AOI Position Shifts on the Dashboard Display due to Head Movements

Initially, the AOI in our study was specifically defined as the speed limit sign displayed on the vehicle's dashboard. However, during the study, we decided to expand the AOI to include the entire dashboard display for the following reasons:

- The dashboard display integrates multiple critical elements, such as speed indicators, ADAS alerts, and navigation information, which collectively contribute to driver SA.
- The AOI corresponding to the speed limit sign occupies a very small area on the dashboard display, making precise fixation measurements challenging and increasing the risk of inaccuracies. By expanding the AOI to encompass the entire dashboard display, we ensured more reliable data extraction while capturing broader interactions with the display.

One possible method to address the dynamic AOI positions would involve recording the initial position of the glasses and tracking their relative movements throughout the experiment. However, implementing this would require highly precise instrumentation capable of dynamically capturing relative positional changes in real-time, which was not feasible for this study. Alternatively, interpolation techniques could be used to approximate AOI positions based on the initial coordinates and estimated movements over time. However, this approach is imprecise due to the complex, non-linear nature of head movements, variations in viewing angles, and differences in the dashboard's visual structure across frames.

YOLOv8 Image Recognition Model

To overcome these limitations, we implemented a machine learning-based image recognition method for identifying AOIs within the video frames. This approach provides several advantages:

- It enables precise identification of AOI regions, such as the dashboard display, despite variations caused by head movements.
- By extracting the video into individual frames, it allows for accurate temporal mapping of AOIs, with a time resolution of 1/30 seconds due to the video's 30 frames-per-second (fps) recording rate.
- The machine learning model generates confidence scores, which help quantify uncertainties and provide a measure of reliability in AOI detection.

The machine learning model employed for image recognition in this study was Ultralytics YOLOv8, a state-of-the-art object detection framework known for its accuracy and efficiency [55]. Specifically, the YOLOv8n.pt pre-trained base model was used, fine-tuned for identifying AOI within the vehicle's dashboard display and monitoring the fixation points represented by red circles in the processed videos.

This dual detection approach was essential for tracking when and how participants' gaze landed on the dashboard.

The YOLOv8n model comprises 23 layers arranged sequentially, including convolutional layers, bottleneck structures, and detection-specific components. The network begins with convolutional and batch normalization layers for low-level feature extraction. These are followed by C2f blocks (Cross-Stage Partial Networks) that integrate bottleneck layers for efficient feature reuse. Detection-specific layers, including upsampling, concatenation, and spatial pyramid pooling, refine features and enhance detection precision. The final detection layer outputs bounding boxes, object classifications, and confidence scores.

YOLOv8 incorporates the following key features [55]:

1. Advanced backbone and neck architectures, resulting in improved feature extraction and object detection performance.
2. An anchor-free split Ultralytics head, which contributes to better accuracy and a more efficient detection process compared to anchor-based approaches.
3. Optimized accuracy-speed tradeoff, maintaining a balance between accuracy and speed, suitable for real-time object detection tasks in diverse applications.
4. A variety of pre-trained models, catering to different tasks and performance requirements, simplifying the process of finding the right model for specific use cases.

In this study, YOLOv8n was employed to detect two distinct classes in the video data:

- The dashboard display, which served as the AOI.
- The fixation point (red circle, see Figure 4.1), which allowed precise temporal mapping of when the gaze intersected with the dashboard display.

Labelling and Training Results

Before training the YOLOv8 model, the labelling process for gaze points and dashboard displays was conducted using the LabelImg tool [56]. This tool allowed for the precise annotation of bounding boxes in video frames. Figure 4.3 illustrates an example of the labelling process, highlighting both the gaze points and the dashboard display.

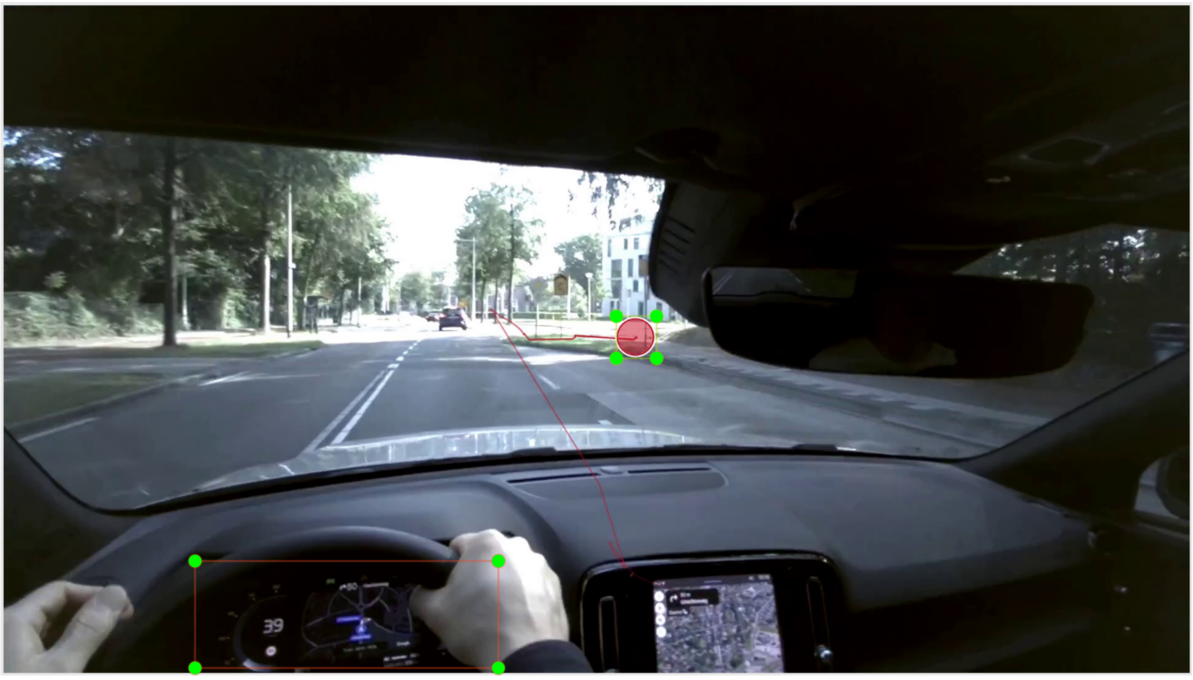
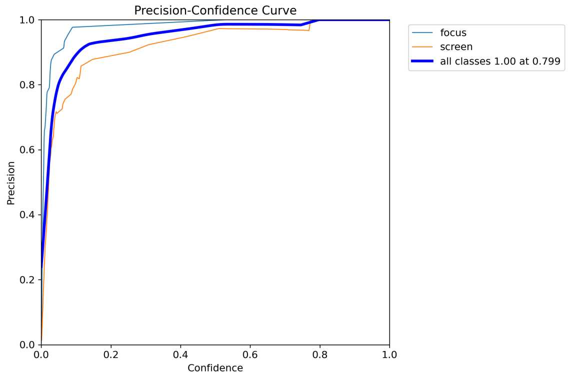


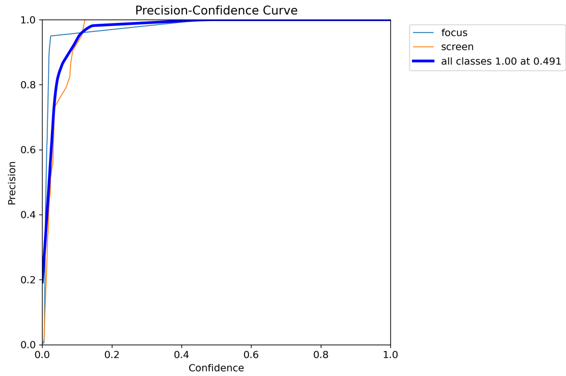
Figure 4.3: Example of the labelling

The dataset for training the YOLOv8 model was created by extracting video frames from the four observation areas of all participants. A total of 100 frames were randomly sampled from this dataset to serve as the training set, and 50 additional frames were sampled as the testing set. These frames included instances of both gaze points and dashboard displays, ensuring balanced coverage of the two target classes.

Separate training was performed for the two ADAS brands (Brand A and Brand B) using the YOLOv8n model. Figures 4.4, 4.5, and 4.6 present the results for three key metrics: precision, recall, and F1-score. Each figure consists of two sub-figures, (a) and (b), corresponding to the results for ADAS Brand A and Brand B, respectively.

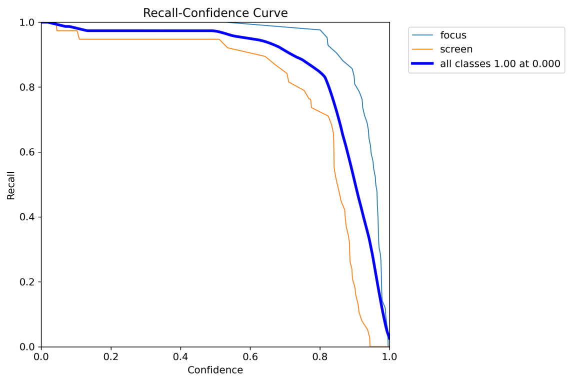


(a) Precision-Confidence Curve for ADAS Brand A.

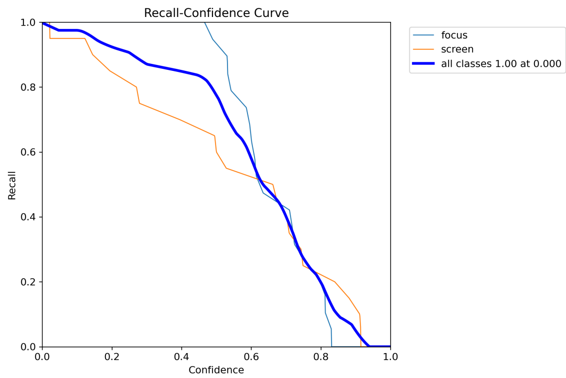


(b) Precision-Confidence Curve for ADAS Brand B.

Figure 4.4: Precision-Confidence Curves

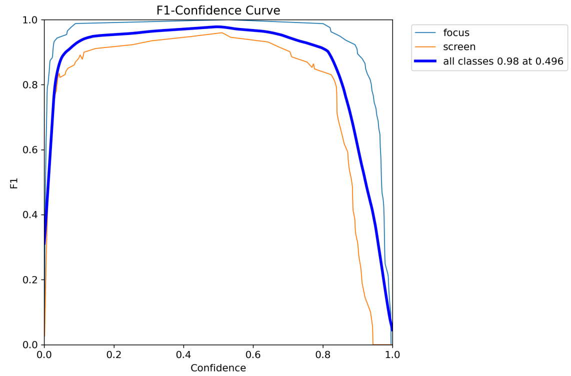


(a) Recall-Confidence Curve for ADAS Brand A.

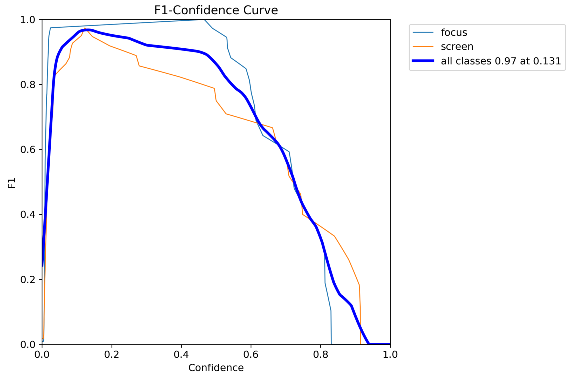


(b) Recall-Confidence Curve for ADAS Brand B.

Figure 4.5: Recall-Confidence Curves



(a) F1-Confidence Curve for ADAS Brand A.



(b) F1-Confidence Curve for ADAS Brand B.

Figure 4.6: F1-Confidence Curves

The final training results, summarized in Table 4.1, highlight strong performance across all evaluated metrics.

Table 4.1: YOLOv8 Model Training Results for ADAS Brands A and B

Metric	Description	Brand A	Brand B
Precision (P)	The proportion of true positives among all positive predictions	0.981	0.979
Recall (R)	The proportion of true positives among all actual positives	0.974	0.975
F1-Score	The harmonic mean of precision and recall	0.977	0.972
mAP@50	Mean Average Precision at a 50% IoU threshold	0.987	0.979
mAP@50-95	Mean Average Precision across IoU thresholds (50%-95%)	0.727	0.693

Validation of YOLOv8

The performance of the YOLOv8 model was further evaluated on the independent test set to validate its robustness. Figures 4.7 and 4.8 illustrate the testing results for ADAS Brand A and Brand B, respectively. Each figure includes two sub-figures: (a) the manually annotated bounding boxes, and (b) the model's detected results for comparison.

(a) Manual Annotations **(b)** Model-Detected Results

Figure 4.7: Testing Results for ADAS Brand A: (a) manually annotated bounding boxes and (b) YOLOv8-detected results

Table 4.2 summarizes the key metrics obtained from the test set for both ADAS brands. These results further demonstrate the model's reliability in detecting gaze points and dashboard displays under diverse conditions.

Table 4.2: YOLOv8 Model Testing Results for ADAS Brands A and B

Metric	Brand A	Brand B
Precision (P)	0.955	0.949
Recall (R)	1.000	1.000
F1-Score	0.978	0.975
mAP@50	0.998	0.995
mAP@50-95	0.680	0.651

While both models achieved high accuracy, the performance of the machine learning model on ADAS B was slightly lower than that of ADAS A. This difference may be attributed to varying levels of image training and adaptation across the two systems. However, this does not impact the validity of the analysis, as both models reached a high level of precision and recall, ensuring their reliability for eye-tracking data interpretation in this study.

Calculation of Indicators Based on Detected Results

The indicators described above were derived using the YOLOv8 model as discussed in the previous sections. For each participant, video frames from the four observation areas were processed by the model to detect AOIs (dashboard display) and gaze points. The model's output provides bounding box coordinates for these regions. By analysing whether the gaze point's bounding box intersects or is fully contained within the AOI's bounding box, we identified instances where the gaze point fell within or overlapped with the AOI.

To determine the timing information, the model's detection results included frame numbers. Using the video frame rate of 30 frames per second (fps), we converted frame indices to time values in seconds. This allowed us to compute various fixation-related indicators.

The process for deriving these indicators is as follows:

(a) Manual Annotations

(b) Model-Detected Results

Figure 4.8: Testing Results for ADAS Brand B: (a) manually annotated bounding boxes and (b) YOLOv8-detected results

1. Detection: Run the YOLOv8 model on each frame to detect gaze points and AOIs.
2. Overlap Check: For each frame, check whether the gaze point overlaps with the AOI.
3. Time Conversion: Use the frame index and frame rate to calculate precise timing.

An example of the code used to achieve this is included in the Appendix (A).

Below is a detailed breakdown of the derived indicators:

Table 4.3: Indicators Derived from Eye-Tracking Data Processing

Indicator	**Calculation Methodology**	**Unit**
Time to First Fixation_A (per Turn)	Time taken for the participant to first fixate on the AOI in Brand A after entering an observation area	Seconds (s)
Time to First Fixation_B (per Turn)	Time taken for the participant to first fixate on the AOI in Brand B after entering an observation area	Seconds (s)
Fixation Count_A (per Turn and Total)	Total count of fixation behaviours on the AOI in Brand A (per turn and aggregated)	Count
Fixation Count_B (per Turn and Total)	Total count of fixation behaviours on the AOI in Brand B (per turn and aggregated)	Count
Fixation Duration_A (per Turn and Total)	Total time during which gaze points overlapped with the AOI in Brand A	Seconds (s)
Fixation Duration_B (per Turn and Total)	Total time during which gaze points overlapped with the AOI in Brand B	Seconds (s)

These indicators provide a comprehensive view of participants' gaze behaviour and interaction with the AOI. This represents the final result of Level 1 SA data processing, which will be analysed alongside the indicators derived from the other two levels in Section 4.3.

4.2.2. Processing Level 2 SA (Questionnaire Data)

The data for Level 2 SA (Comprehension) were obtained from participants' questionnaire responses. Each response was scored according to a predefined scoring system. The detailed scoring methodology for each question, along with the full set of questions, can be found in Appendix (B) and Appendix (C).

For each participant, individual scores were calculated separately for ADAS Brand A and ADAS Brand B. The final score for ADAS Brand A was computed as the sum of scores from Questions 1 through 5, while the final score for ADAS Brand B was based on the sum of scores from Questions 1 and 6 through 9.

Since each question was equally weighted, the total score directly reflects participants' performance across all relevant questionnaire items for each brand. This represents the final result of Level 2 SA data processing, which will be analysed alongside the indicators derived from the other two levels in Section 4.3.

4.2.3. Processing Level 3 SA (GPS Data)

In this section, we process GPS data obtained during the driving tasks. The raw GPS data included time, latitude, longitude, and calculated instantaneous speed. However, speed limit information was manually added to the dataset based on predefined road segments. The structure of the raw data, along with sample entries, is shown in Table 4.4.

Table 4.4: Sample Entries of Raw GPS Data with Speed Limit Information

Date	GPS Coordinates (Lat.) [deg]	GPS Coordinates (Long.) [deg]	GPS (2D Speed) [m/s]	Speed Limit [km/h]
11:38:56.105Z	49.0288035	6.4585172	7.177	60
11:38:56.160Z	49.0288025	6.4585223	7.177	60
11:38:56.270Z	49.0287878	6.4585331	7.166	60
11:38:56.325Z	49.0287945	6.4585324	7.154	60
11:38:56.380Z	49.0287934	6.4585376	7.154	60
11:38:56.490Z	49.0288189	6.4585369	8.889	60
11:38:56.655Z	49.0288552	6.4585492	25.619	60
11:38:56.710Z	49.0288679	6.4585479	25.619	60

To calculate meaningful indicators, such as the average speed difference relative to the speed limit and the percentage of time spent within the speed limit, we first processed the raw data using a rolling window technique. This approach ensures smoother transitions in instantaneous speed data by taking an average over a set window of time. Furthermore, a threshold was applied to exclude outliers, which typically arise from GPS errors or rapid, unrealistic speed variations. In this study, a threshold of 20 km/h was used to filter out data points with excessive speed differences.

To visualize the effect of the rolling window and the removal of outliers, Figure 4.9 and Figure 4.10 compare the GPS data before and after rolling. The colour coding in the map legend indicates the magnitude of the speed difference:

- Green: Speed is below or matches the limit.
- Orange: Speed is above the limit by up to 10 km/h.
- Red: Speed exceeds the limit by more than 10 km/h.

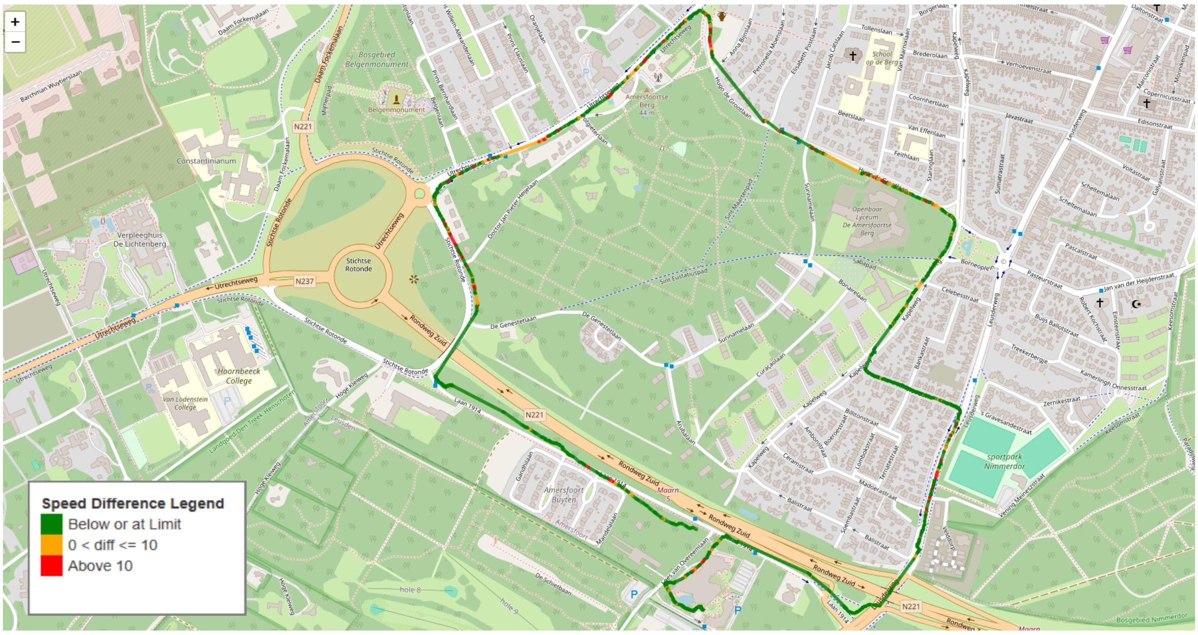


Figure 4.9: GPS Data Before Rolling Window Processing

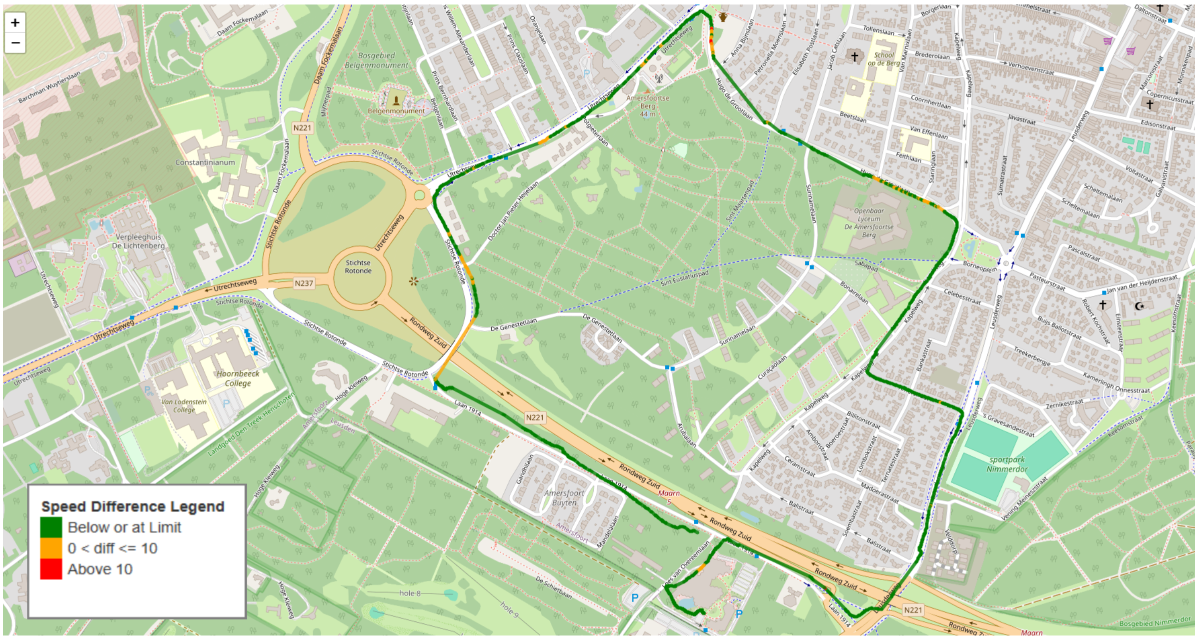


Figure 4.10: GPS Data After Rolling Window Processing

In Figure 4.9, red segments indicating significant deviations above the speed limit are relatively more frequent compared to Figure 4.10, particularly near speed limit transition areas. However, red segments remain less frequent overall when compared to the green and yellow segments. In Figure 4.10, these red segments are further reduced, demonstrating the effectiveness of the rolling window technique in minimizing anomalies and providing smoother, more reliable speed data.

After processing the data, we calculated two key metrics:

- Average Speed Difference: The difference between the mean calculated speed and the speed limit for each road segment.

- Speed Compliance Percentage: The percentage of time the vehicle speed remained within the speed limit during each observation area.

This processed dataset forms the basis for deriving Level 3 SA indicators. The final implementation code for this data processing is included in the appendix. This ensures reproducibility of the results and provides a reference for further analyses.

4.3. Data Analysis

4.3.1. Descriptive Statistics

The descriptive statistics analysis provides an overview of the data by summarizing key variables across the three SA levels.

First, basic demographic and moderating variables such as age, gender, driving experience, and familiarity with ADAS (as outlined in Table 3.1) are summarized using pie charts and other graphical methods to represent the composition of the participant group.

Next, for each SA level, descriptive statistics (e.g., mean, standard deviation, and range) are calculated for all dependent indicators listed in Table 3.1. These summary statistics will provide a foundation for identifying patterns and variations in the data, setting the stage for further statistical analyses.

4.3.2. Significance Testing for Indicators

This section represents a critical part of the analysis, as the primary objective of this study is to explore the differences in drivers' SA under varying ADAS systems. Significance testing is therefore essential to identify whether these differences across the three SA levels (Level 1 SA, Level 2 SA, and Level 3 SA) are statistically meaningful, providing insights into how different systems influence SA.

To test the significance of the dependent indicators identified in Sections 4.2.1, 4.2.2, and 4.2.3 for Level 1 SA (Eye-Tracking Data), Level 2 SA (Questionnaire Data), and Level 3 SA (GPS Data), respectively, the following methods will be applied:

Level 1 SA (Eye-Tracking Data): Paired t-tests will be conducted on indicators derived from Level 1 SA, such as gaze fixation metrics and time to first fixation. These tests are appropriate for comparing data with interval or ratio scales across the two ADAS systems, given the within-subject experimental design.

Level 2 SA (Questionnaire Data): For Level 2 SA, which consists of subjective questionnaire responses measured on ordinal scales, non-parametric tests will be applied. Specifically, the Wilcoxon signed-rank test will be used to compare participants' ratings under the two ADAS systems [57]. This approach is robust for ordinal data and accounts for any non-normality in the distribution. Moreover, some studies suggest that the Wilcoxon signed-rank test is particularly suitable for small sample sizes, further justifying its use in this context [58].

Level 3 SA (GPS Data): Indicators from Level 3 SA, such as speed compliance percentage and average speed deviation, will be analysed using paired t-tests. These metrics, measured on interval or ratio scales, are suitable for t-tests given the within-subject comparisons of the two ADAS systems.

By employing these significance testing methods for each SA level, this analysis rigorously evaluates whether the observed differences in SA indicators are statistically significant.

4.3.3. Correlation Analysis Between SA Levels

A correlation coefficient matrix will be employed to analyze the relationships between the indicators of SA1 (eye-tracking metrics), SA2 (questionnaire scores), and SA3 (driving behaviour metrics). This matrix provides a comprehensive overview of pairwise correlations across all indicators identified in Section 4.2, highlighting potential interconnections and shared patterns between SA levels. The analysis aims to reveal how different aspects of SA interact under varying ADAS systems.

4.3.4. Linear Mixed Models (LMMs) Analysis

Methodological Considerations in Model Selection

Several statistical methods were considered before selecting the final approach. GEE (Generalized Estimating Equations) was initially explored due to its suitability for repeated measures data; however, it does not account for individual differences. Given that fixation count and other key metrics exhibit significant variability between individuals, explicitly modeling these differences became essential. Bayesian methods, while advantageous for small sample sizes, posed challenges due to their computational complexity, making them less practical for our dataset. The Permutation Test was also considered because it is well-suited for small samples, but its inability to capture the effects of multiple variables made it unsuitable for our study, which aims to examine factors such as age, gender, and driving experience.

Ultimately, we selected LMM as the most appropriate approach. Despite this choice, certain limitations remain. The small sample size, while unavoidable, is acknowledged as a study limitation. Additionally, the *Fixation Count* variable in SA1 metrics is not inherently continuous; however, treating it as approximately continuous (starting from zero) ensures methodological consistency.

Given these considerations, LMM provides the necessary flexibility to account for both individual variability and contextual influences, making it the most suitable method for our study.

Modelling Systematic Differences Using LMM

LMMs were applied independently to analyse systematic differences in SA indicators for each SA level across ADAS systems. This modelling approach accounts for individual variability and contextual factors while isolating the fixed effect of the ADAS system.

For SA1 (Eye-Tracking Data) and SA3 (GPS Data), the model structure is defined as follows:

$$\text{Indicator} \sim \text{System} + \text{Turn} + (1|\text{Participant})$$

Where:

- System: a fixed effect representing the impact of the ADAS system (System A or System B) on the specific SA indicator.
- Turn: a fixed effect accounting for variability across different instances of participant interaction with ADAS systems, capturing the contextual influence on SA indicators.
- Participant: a random effect accounting for variability across individual participants, summarizing individual differences and reducing model complexity.

Since the *Turn* variable is incorporated as a fixed effect, only the per-turn data for SA1 indicators were utilized in the modeling process. The total indicators for SA1, which represent aggregated values across all four turns, were excluded from separate LMM modeling, as the influence of turns is explicitly modeled as a fixed effect. This approach allows for a more direct examination of how ADAS systems impact participants' SA indicators across different turns while ensuring that contextual variability is explicitly accounted for.

For SA2 (Questionnaire Data), we applied ordered logistic regression due to the ordinal nature of the self-reported scores, which range from 1 to 5. Unlike SA1 and SA3, where mixed-effects models accounted for random variability, ordered logistic regression does not currently support random effects unless significant computational complexity is introduced.

Thus, in the analysis of SA2, we focused on the fixed effect of the ADAS Systems A and B. The model is specified as:

$$\text{Score}_{\text{SA2}} \sim \text{System}$$

The structures facilitate a tailored evaluation of systematic differences in SA indicators for each SA level, accounting for their unique data characteristics. By modelling each level independently, the analysis ensures that the differences between System A and System B are evaluated with appropriate controls for variability at the participant, observation area, and sequence levels.

Exploring Internal Factors Using LMM

To explore the influence of internal factors such as age, gender, driving experience, and ADAS familiarity on SA indicators, separate LMMs were applied for each SA level. These models were structured as follows:

$$\text{Indicator} \sim \text{Turn} + \text{Age} + \text{Gender} + \text{Driving Experience} + \text{ADAS Familiarity} + (\text{System}|\text{Participant})$$

Where:

- Turn: A fixed effect accounting for variability across different turns, capturing contextual influences on SA indicators.
- Age, Gender, Driving Experience, ADAS Familiarity: Fixed effects hypothesized to influence the specific SA indicator.
- System | Participant: A random slope effect, allowing the effect of the ADAS system to vary across participants, thereby accounting for individual differences in system impact.

For variables not explicitly provided with values in previous sections, their encoding for use in the LMM model is defined in Table 4.5 below. This ensures consistency and clarity in the analysis of demographic and experiential factors within the models.

Table 4.5: Encoding of Variables for LMM Analysis

Variable	Description	Encoding Value and Meaning
Gender	Male or Female	0 = Male
		1 = Female
Age Group	Participant's age range	1 = Under 18
		2 = 18-24
		3 = 25-34
		4 = 35-44
		5 = 45-54
		6 = 55-64
		7 = 65 or older
Driving Experience	Years of driving experience	1 = Less than 1 year
		2 = 1-3 years
		3 = 4-7 years
		4 = 8-15 years
		5 = More than 15 years
ADAS Familiarity	Familiarity with ADAS systems	1 = Not familiar at all
		2 = Slightly familiar
		3 = Neutral
		4 = Familiar
		5 = Very familiar
Order	ADAS trial order (System A or System B first)	0 = System A first
		1 = System B first

The variable Order represents whether a participant first drove ADAS System A or System B. Although data for this variable were collected, it was not included in the final model after screening. This decision was made because the learning effect that Order represents is inherently captured within the Participant variable, which accounts for individual differences in adaptation and performance across trials.

This independent modelling approach quantifies the influence of demographic and experiential factors on SA performance across different levels. By examining how internal factors, such as age, gender, driving experience, and ADAS familiarity, interact with SA indicators, the analysis provides a comprehensive understanding of their impact on various outcomes. The results offer valuable insights into the nuanced interplay between driver characteristics and ADAS system performance, highlighting the systematic ways in which internal factors shape SA outcomes.

4.3.5. Integration Across Data Sources

In this final stage of analysis, we integrate data from SA1 (eye-tracking metrics), SA2 (questionnaire responses), and SA3 (driving behaviour metrics) to investigate their interrelationships and infer the implications of these findings for ADAS speed limit signal design. This involves analysing the cross-level relationships to identify how indicators from different SA levels align or diverge in capturing participants' SA. Beyond identifying differences, the integration aims to provide plausible explanations for these patterns.

This integrative analysis serves as the ultimate objective of the study. By synthesizing findings across levels of SA, the research seeks to uncover potential strengths and weaknesses in ADAS system designs, offering insights to inform future improvements and enhance overall driver-system interaction.

5

Results

overview

The purpose of this chapter is to present the results derived from the data processing and analysis methods outlined in Chapter 4. It systematically follows the framework established in the Data Analysis section, reporting key findings for each situational awareness (SA) level. Section 5.1 provides descriptive statistics for the indicators, summarizing their means, standard deviations, and distributions. Section 5.2 details the results of significance testing, comparing SA indicators across ADAS systems using the appropriate statistical tests for each level. Section 5.2.4 explores the relationships between SA levels through correlation analysis, identifying key patterns and connections. Finally, Section 5.2.4 presents the outcomes of the Linear Mixed Models (LMM) analysis, examining systematic differences in SA indicators and the influence of demographic and experiential factors. The results will be synthesized to address the research questions, providing a cohesive and detailed understanding of the findings.

5.1. Descriptive Statistics

This section summarizes the basic statistical characteristics of the key indicators across the three SA levels, including means, standard deviations, and distributions. Additionally, it provides a graphical illustration of the moderating variables used in the study to describe the participant demographics and experimental design.

5.1.1. Graphical Representation of Moderating and Experimental Variables

To provide an overview of the participant demographics, experimental setup, and other influencing factors, the following figures illustrate the distributions of key variables. These include age, gender, driving experience, familiarity with ADAS systems, and participants' preferences for ADAS systems. Although trial sequence (the order of using System A or System B) is not strictly a result, it is included here as a moderating variable, similar to the others, to facilitate later explanations and analyses. These graphical representations provide a comprehensive understanding of the sample composition and experimental context.

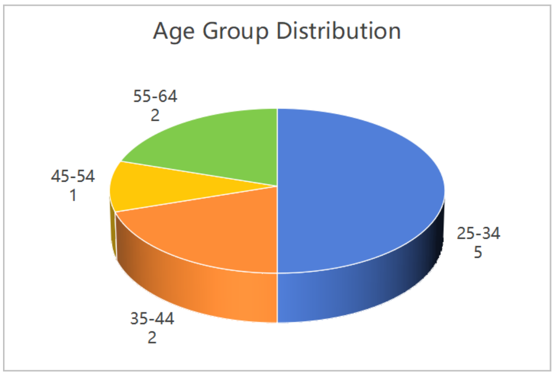


Figure 5.1: Age Group Distribution

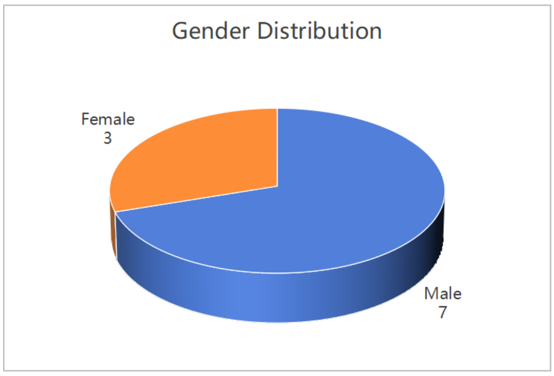


Figure 5.2: Gender Distribution

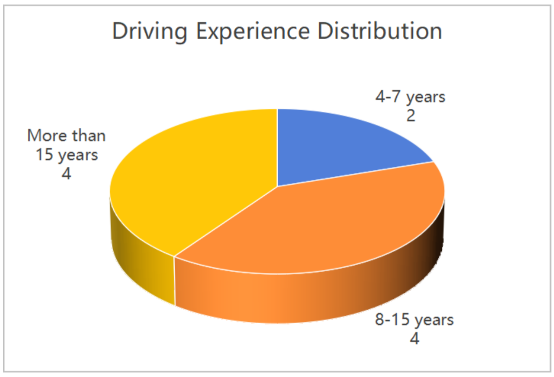


Figure 5.3: Driving Experience Distribution

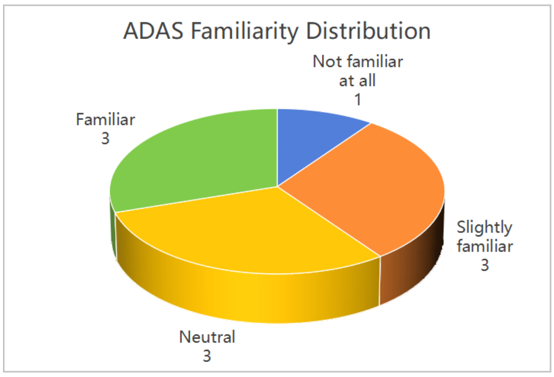


Figure 5.4: ADAS Familiarity Level Distribution

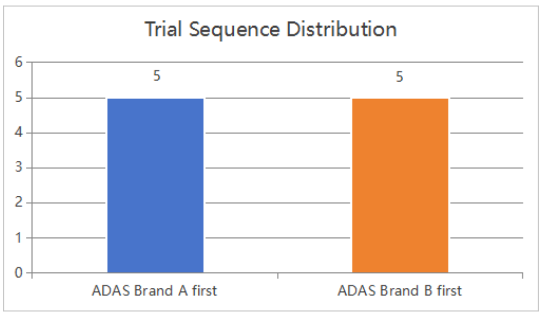


Figure 5.5: Trial Sequence Distribution

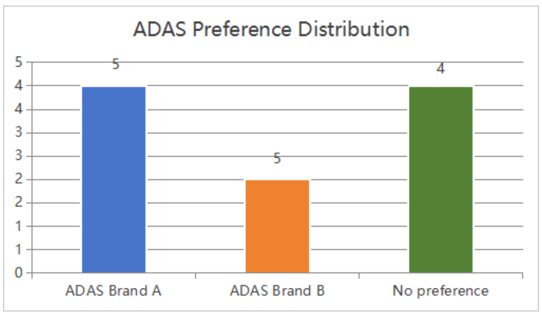


Figure 5.6: ADAS Preference Distribution

The graphical representations provide an overview of participant demographics, familiarity with ADAS systems, and experimental setup. Figure 5.1 illustrates the age group distribution, with the majority of participants falling into the 25–34 age range (50%), followed by the 35–44 group (20%) and the 55–64 group (20%). Figure 5.2 highlights the gender distribution, showing that the sample is predominantly male (70%), with females comprising 30% of the participants. Figure 5.3 indicates that most participants have significant driving experience, with 40% having driven for more than 15 years and another 40% falling into the 8–15 years category. Figure 5.4 reveals that ADAS familiarity is evenly distributed, with 30% of participants reporting familiarity, 30% neutral, and 30% slightly familiar, while only 10% reported no familiarity at all.

Figure 5.5 shows the trial sequence distribution, evenly split between participants who experienced ADAS Brand A first (50%) and those who experienced ADAS Brand B first (50%). Finally, Figure 5.6 captures participants' ADAS preferences, where 40% preferred Brand A, 20% preferred Brand B, and the remaining 40% expressed no specific preference. These distributions highlight the diversity of the participant group and provide useful context for interpreting the experimental results.

5.1.2. Descriptive Summary of SA Indicators

The descriptive statistics of SA indicators (Table 5.1) highlight key differences and similarities across the three SA levels. These observations provide a foundation for understanding the data and identifying areas for deeper analysis.

Table 5.1: Descriptive Statistics for SA Indicators

SA Level	Indicator	Count	Mean	Std	Min	Max
SA1	Fixation Count per Turn_A	40	1.00	1.13	0.00	4.00
	Fixation Count per Turn_B	40	1.48	1.60	0.00	5.00
	Time to First Fixation_A	40	20.00	10.06	0.10	30.00
	Time to First Fixation_B	40	20.29	9.69	0.13	30.00
	Fixation Duration per Turn_A	40	0.15	0.22	0.00	0.77
	Fixation Duration per Turn_B	40	0.18	0.21	0.00	0.63
	Fixation Count in Total_A	10	4.00	3.23	0.00	9.00
	Fixation Count in Total_B	10	5.90	5.22	0.00	15.00
	Fixation Duration in Total_A	10	0.59	0.54	0.00	1.50
	Fixation Duration in Total_B	10	0.70	0.63	0.00	1.54
SA2	SA2_A	10	2.82	0.48	2.00	3.60
	SA2_B	10	2.46	0.57	1.40	3.40
SA3	Speed Difference_A	40	12.77	4.53	3.38	19.01
	Speed Difference_B	40	12.71	4.73	4.57	22.41
	Speed Compliance Percentage_A	40	78.26	12.27	52.70	99.33
	Speed Compliance Percentage_B	40	76.33	11.46	56.40	98.65

For SA1 (Eye-tracking Data), notable differences were observed between System A and System B. *Fixation Count per Turn* had a higher mean for System B (mean = 1.48) compared to System A (mean = 1.00), along with greater variability (std = 1.60 vs. 1.13). Similarly, *Fixation Duration per Turn* showed a slightly higher mean and variability for System B (mean = 0.18, std = 0.21) than System A (mean = 0.15, std = 0.22). In the aggregated indicators, *Fixation Count in Total_B* was substantially higher (mean = 5.90, std = 5.22) compared to *Fixation Count in Total_A* (mean = 4.00, std = 3.23), reflecting potential differences in attention allocation. These findings suggest distinct gaze behaviour patterns across the two systems.

For SA2 (Questionnaire Data), the mean self-reported comprehension scores for System A (mean = 2.82) were marginally higher than those for System B (mean = 2.46). The variability was slightly greater for System B (std = 0.57) than for System A (std = 0.48), suggesting more diverse subjective evaluations of System B's interface.

For SA3 (GPS Data), *Speed Compliance Percentage* showed a slight advantage for System A (mean = 78.26) compared to System B (mean = 76.33). However, variability was similar across systems (std =

12.27 for System A, std = 11.46 for System B). *Speed Difference* metrics were nearly identical between the two systems, with only a marginally higher variability for System B (std = 4.73) compared to System A (std = 4.53).

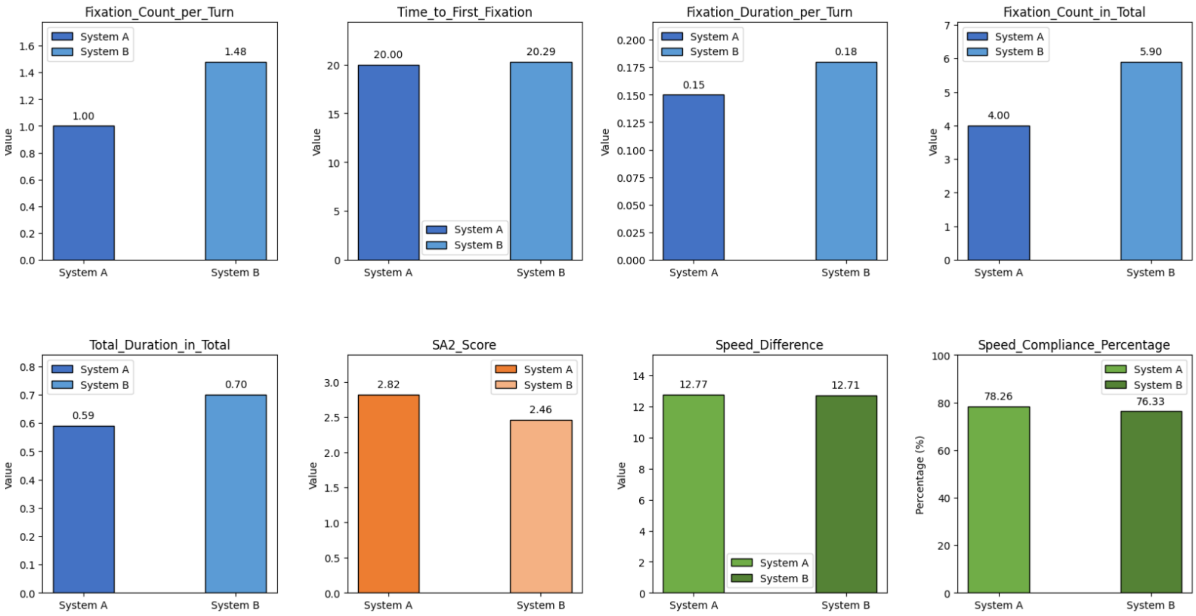


Figure 5.7: Comparison of System A and B Across Indicators

In summary, certain SA indicators appear to show more pronounced differences between the two systems. These observations are preliminary and will be further examined through significance testing to validate and deepen the understanding of these differences.

5.2. Significance Testing for Indicators

The significance testing for the indicators was conducted to assess differences between ADAS System A and System B. For SA1 and SA3, Shapiro-Wilk tests were used to evaluate normality [59], and Levene's tests were conducted to check homogeneity of variances [60]. Based on these results, the statistical methods were determined accordingly. Specifically, the Wilcoxon signed-rank test was employed for indicators that did not meet the assumptions of normality or homogeneity. This non-parametric test is advantageous over the paired t-test for data with non-normal distributions [58]. Conversely, the paired t-test was applied to indicators that satisfied the assumptions, enabling robust evaluation of differences between the two systems.

For SA1, the per-turn indicators (*Fixation Count per Turn*, *Time to First Fixation*, *Fixation Duration per Turn*) did not meet the assumptions for t-tests in terms of normality or homogeneity. Consequently, Wilcoxon signed-rank tests were applied. Although the total indicators (*Fixation Count in Total*, *Fixation Duration in Total*) aggregated data across four turns and met the assumptions for t-tests, we chose to maintain consistency by also applying Wilcoxon signed-rank tests to these aggregated measures, considering the characteristics of the per-turn indicators.

For SA2, as the indicators represent ordinal self-assessment scores, Wilcoxon signed-rank tests were applied without conducting normality or homogeneity checks.

For SA3, all indicators (*Speed Difference*, *Speed Compliance Percentage*) passed the normality and homogeneity tests, allowing paired t-tests to be used to evaluate the differences between System A and System B.

In this study, we adopted a standard threshold for p-value interpretation: results with $p < 0.05$ were considered statistically significant, while those with $0.05 \le p < 0.1$ were regarded as marginally significant.

The statistical tests applied to each indicator and their respective results are summarised in Table 5.2.

Table 5.2: Statistical Test Methods and Results for SA Indicators

SA Level	Indicator	Statistical Test	W/t-Statistic	p-value
SA1	Fixation Count per Turn A/B	Wilcoxon signed-rank test	45.000	0.0409
	Time to First Fixation A/B	Wilcoxon signed-rank test	171.000	0.0900
	Fixation Duration per Turn A/B	Wilcoxon signed-rank test	114.000	0.1918
	Fixation Count in Total A/B	Wilcoxon signed-rank test	4.000	0.0881
	Fixation Duration in Total A/B	Wilcoxon signed-rank test	7.000	0.1235
SA2	SA2 A/B	Wilcoxon signed-rank test	0.0000	0.0679
SA3	Speed Difference A/B	Paired t-test	0.1498	0.8817
	Speed Compliance Percentage A/B	Paired t-test	0.7865	0.4363

5.2.1. Results of Wilcoxon Signed-Rank Tests for SA Level 1

The Wilcoxon signed-rank tests for SA1 (Eye-Tracking Data) revealed the following findings:

First, there was a statistically significant difference in *Fixation Count per Turn* (p-value = 0.0409, W-statistic = 45.000). System B showed higher fixation times on average compared to System A. This difference indicates that the two ADAS systems influence drivers' fixation behaviour during turns differently, suggesting that System B may require more sustained visual attention or impose greater cognitive workload.

Second, indicators with marginal significance included *Time to First Fixation* (p-value = 0.0900, W-statistic = 171.000) and *Fixation Count in Total* (p-value = 0.0881, W-statistic = 4.000). These results show potential trends but do not meet the threshold for strong statistical significance.

Finally, for the other indicators, including *Fixation Duration per Turn* (p-value = 0.1918, Statistic = 114.000) and *Fixation Duration in Total* (p-value = 0.1235, W-statistic = 7.000), no statistically significant differences were observed. These findings suggest that, apart from per-turn fixation behaviour, other aspects of gaze behaviour remain largely consistent between the two ADAS systems.

5.2.2. Results of Wilcoxon Signed-Rank Tests for SA Level 2

The results of the Wilcoxon signed-rank test for the SA2 indicator reveal a p-value of 0.0679, indicating a marginally significant effect ($p < 0.1$ but slightly greater than 0.05). Similar to some indicators in SA1, this demonstrates a marginal correlation, suggesting potential differences in participants' self-reported SA between the two ADAS systems.

5.2.3. Results of t-Tests for SA Level 3

The paired t-tests for SA3 indicators revealed no significant differences between ADAS System A and System B. For *Speed Difference*, the p-value was 0.8817 (t-statistic = 0.1498), and for *Speed Compliance Percentage*, the p-value was 0.4363 (t-statistic = 0.7865). These results indicate that both systems had similar effects on drivers' speed-related behaviours, suggesting comparable performance in terms of speed management.

As with the results in Section 5.2.1 and 5.2.2, these findings provide a valuable basis for the final integration and comprehensive analysis of SA across levels.

5.2.4. Correlation Analysis Between SA Levels

The correlation matrix presented in Figure 5.8 provides an overview of the relationships between the indicators across the three SA levels. The results highlight several important patterns:

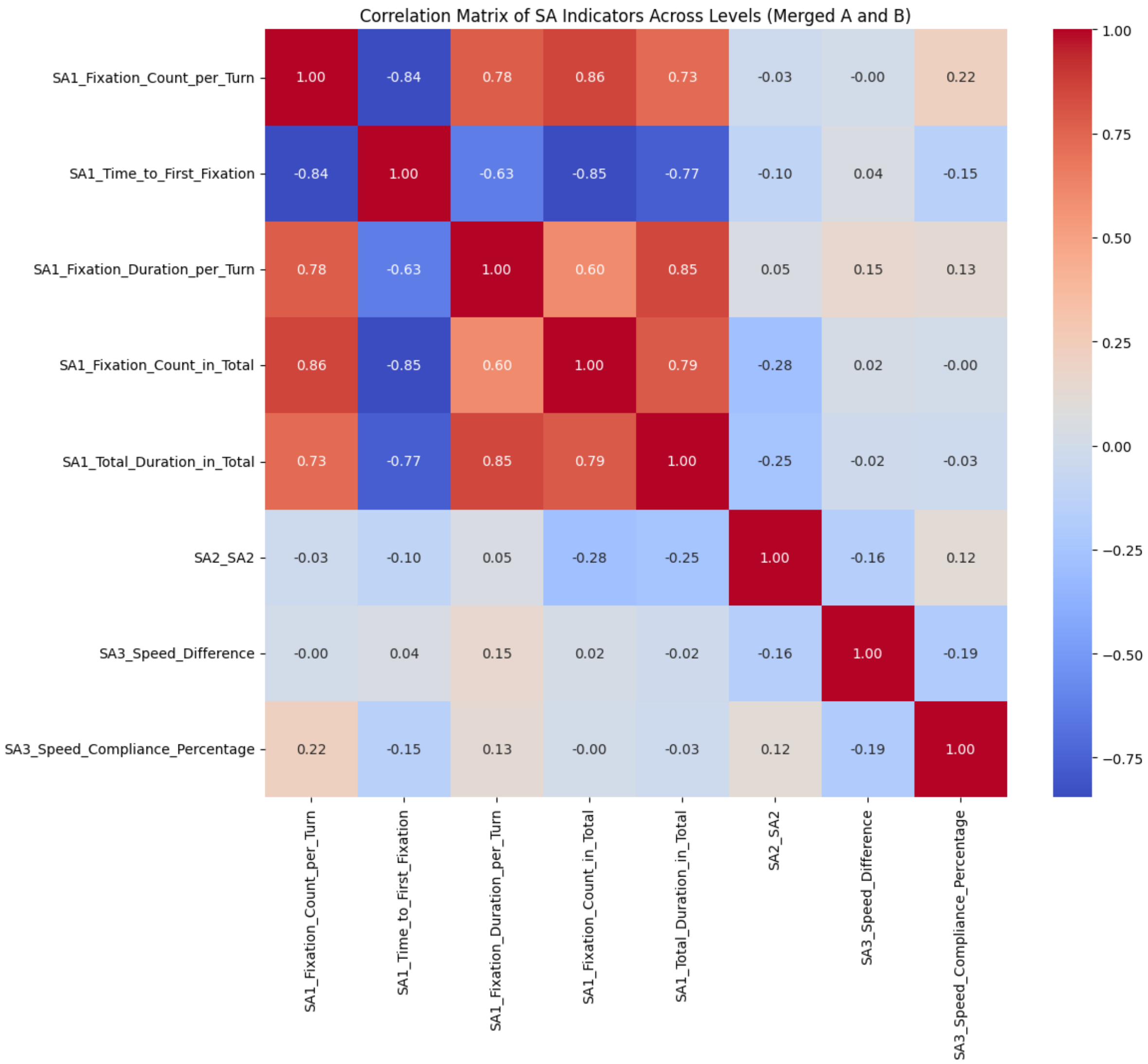


Figure 5.8: Correlation Analysis Between Indicators

For SA1 (Eye-Tracking Data), the relationship between its indicators and those in SA2 and SA3 appears minimal. The highest correlation observed is between *Fixation Count per Turn* and *Speed Compliance Percentage*, with a value of 0.22. This weak correlation suggests that eye-tracking metrics are largely independent of speed compliance behaviours. Similarly, correlations between other SA1 indicators, such as *Total Duration in Total*, and *Speed Difference* or SA2 indicators, are consistently below 0.1, indicating a lack of meaningful relationships. Overall, SA1 indicators demonstrate very limited cross-level correlations with SA2 and SA3 metrics.

For SA2 (Self-Reported Data), its association with SA3 (GPS Data) indicators is similarly weak. The correlation between SA2 and *Speed Difference* is -0.16, while the correlation with *Speed Compliance Percentage* is 0.12. These values suggest no substantial alignment between participants' self-reported SA and their speed-related behaviours. This indicates that subjective perceptions of system effectiveness do not strongly align with objective speed regulation metrics.

In summary, the indicators from SA1, SA2, and SA3 levels exhibit minimal cross-level correlations, reinforcing the notion that these dimensions of SA are largely independent and capture distinct aspects of driver interaction with ADAS systems. This underscores the importance of analysing each level separately to fully understand the drivers' SA.

5.3. Linear Mixed Models and Ordered Logistic Regression Results

This section presents and analyses the results of the LMM model. Note that only the most important data are presented here, and more results can be found in Appendix (D).

5.3.1. LMM Results for SA Level 1

Based on the significance analysis in Section 4.2, we determined that the SA1 data did not follow a normal distribution. To address this, we applied the Box-Cox transformation, which stabilizes variance and ensures residuals approximate a normal distribution. This allowed us to evaluate systematic differences between ADAS systems while accounting for participant-level random effects. Overall, the analysis did not reveal statistically significant differences between System A and System B across the SA1 indicators.

Table 5.3: Mixed Linear Model Results for SA Level 1

Indicator	Predictor	Estimate (β)	Std. Err.	z-value	p-value	95% CI
Fixation Count	Intercept	-1.177	0.675	-1.744	0.081	[-2.500, 0.146]
	System B	0.305	0.346	0.881	0.378	[-0.373, 0.983]
	Turn 2	-0.092	0.489	-0.189	0.850	[-1.051, 0.867]
	Turn 3	-0.361	0.489	-0.737	0.461	[-1.320, 0.599]
	Turn 4	-0.470	0.489	-0.961	0.336	[-1.430, 0.489]
Time to First Fixation	Intercept	16.398	2.116	7.750	<0.001	[12.251, 20.545]
	System B	0.225	1.305	0.172	0.863	[-2.333, 2.782]
	Turn 2	-1.917	1.845	-1.039	0.299	[-5.534, 1.700]
	Turn 3	-0.706	1.845	-0.383	0.702	[-3.923, 2.511]
	Turn 4	-2.645	1.845	-1.433	0.152	[-6.262, 0.973]
Fixation Duration	Intercept	-3.260	0.650	-5.014	<0.001	[-4.535, -1.986]
	System B	0.306	0.360	0.849	0.396	[-0.400, 1.012]
	Turn 2	-0.545	0.509	-1.071	0.284	[-1.544, 0.453]
	Turn 3	-0.327	0.509	-0.642	0.521	[-1.325, 0.671]
	Turn 4	-0.800	0.509	-1.570	0.116	[-1.798, 0.199]

For *Fixation Count* per turn, System B showed a slightly higher *Fixation Count* than System A (coefficient = 0.305). However, this effect was not statistically significant (p = 0.378), suggesting that differences in visual attention investment between the two systems are not strong enough to be conclusive.

For *Time to First Fixation*, participants using System B identified relevant areas slightly faster than those using System A (coefficient = 0.225). However, this difference was also not statistically significant (p = 0.863), indicating that system design did not substantially alter the time taken to locate key areas.

For *Fixation Duration* per turn, System B was associated with a longer *Fixation Duration* than System A (coefficient = 0.306), but this effect also failed to reach statistical significance (p = 0.396). This suggests that any increase in visual attention demand imposed by System B was not sufficiently large to be distinguished from normal variability.

The random effect of *Participant* showed variance estimates of *3.060* (*Fixation Count*), *23.484* (*Time to First Fixation*), and *2.607* (*Fixation Duration*). This highlights the importance of modelling individual differences in gaze behaviour.

Overall, although System B showed slightly higher cognitive demand than System A across all three indicators, the differences did not reach statistical significance.

5.3.2. Ordered Logistic Regression Results for Level 2 SA

The ordered logistic regression results for SA2 reveal differences in participants' self-reported scores for ADAS System A and B. Specifically, the coefficient for ADAS System B (-1.0223) indicates that participants tend to assign lower scores to System B compared to System A. This finding suggests that participants may perceive System B less favourably in terms of its usability or effectiveness, aligning with previous observations from SA1.

Table 5.4: Ordered Model Results for SA Level 2

Indicator	Predictor	Estimate (β)	Std. Err.	z-value	p-value	95% CI
Score	System B	-1.0223	0.843	-1.213	0.225	[-2.674, 0.629]
Thresholds	1.4/1.8	-3.5558	1.164	-3.855	0.002	[-5.837, -1.275]
	1.8/2.0	-0.2538	0.987	-0.257	0.797	[-2.189, 1.681]
	2.0/2.6	-0.1659	0.683	-0.244	0.888	[-1.505, 1.173]
	2.6/2.8	0.7251	0.294	2.468	0.014	[0.149, 1.301]
	2.8/3.2	-0.6925	0.676	-1.024	0.306	[-2.018, 0.633]
	3.2/3.4	-1.1701	0.974	-1.201	0.230	[-3.080, 0.740]
	3.4/3.6	-1.0221	0.977	-1.046	0.296	[-2.937, 0.893]

The thresholds in Table 5.4 represent the latent boundaries where participants transition between adjacent score levels on a continuous underlying evaluation scale. These boundaries do not correspond directly to integer scores (e.g., 1, 2, 3) but rather indicate the points at which participants' ratings shift more distinctly. The results reveal a clear differentiation pattern: the threshold between 1.4 and 1.8 (β = -3.5558, p = 0.002) was significantly negative, suggesting heightened sensitivity to lower scores, likely reflecting stronger perceptions of system shortcomings. Similarly, the threshold at 2.6/2.8 (β = 0.7251, p = 0.014) was significant, though the positive coefficient indicates a tendency toward higher ratings. In contrast, thresholds at the upper end of the scale (e.g., 3.4/3.6, p = 0.296) lacked significance, implying weaker differentiation between adjacent high scores. This pattern of sharper distinctions at lower ratings and greater ambiguity at higher ratings suggests that participants were more consistent in identifying system deficiencies, while their positive evaluations were more subjective. These findings align with the trend that System B received consistently lower ratings, reinforcing potential usability concerns.

While the observed differences in scores for ADAS System B compared to System A did not reach statistical significance, the consistently lower scores for System B reflect potential underlying issues in user experience or satisfaction.

5.3.3. LMM Results for SA Level 3

The results for SA3 metrics, *Speed Difference* and *Speed Compliance Percentage*, suggest minimal observable differences between ADAS System A and B.

For *Speed Difference*, the coefficient for System B is -0.058, indicating a negligible reduction compared to System A. The p-value (p = 0.923) confirms that this effect is not statistically significant. While Turn 3 and Turn 4 exhibit larger effects ($p < 0.05$), the lack of significant differences between systems suggests that speed variation is more influenced by road context rather than the ADAS system itself.

For *Speed Compliance Percentage*, the coefficient for System B is -1.925, suggesting a slight decrease in compliance under System B compared to System A. However, this difference is also not statistically significant (p = 0.447). Similarly, none of the turn-based effects reach statistical significance, indicating that variations in speed compliance are not systematically linked to the ADAS system.

Table 5.5: Mixed Linear Model Results for SA Level 3

Indicator	Predictor	Estimate (β)	Std. Err.	z-value	p-value	95% CI
Speed Difference	Intercept	14.574	0.961	15.167	<0.001	[12.691, 16.458]
	System B	-0.058	0.594	-0.097	0.923	[-1.233, 1.107]
	Turn 2	-5.209	0.841	-6.196	<0.001	[-6.856, -3.561]
	Turn 3	2.429	0.841	2.889	0.004	[0.781, 4.076]
	Turn 4	-4.450	0.841	-5.293	<0.001	[-6.097, -2.802]
Speed Compliance Percentage	Intercept	77.907	3.137	24.838	<0.001	[71.759, 84.054]
	System B	-1.925	2.533	-0.760	0.447	[-6.890, 3.041]
	Turn 2	0.584	3.583	0.163	0.871	[-6.438, 7.606]
	Turn 3	1.598	3.583	0.446	0.656	[-5.424, 8.620]
	Turn 4	-0.774	3.583	-0.216	0.829	[-7.796, 6.248]

In conclusion, the observed differences between the two systems in SA3 metrics are minimal. While *Speed Difference* shows larger variability across turns, no significant effect of ADAS system type was found. Similarly, *Speed Compliance Percentage* exhibited a slight but statistically insignificant decrease under System B. These findings suggest that factors beyond the ADAS system itself, such as road conditions and driver variability, may play a more dominant role in shaping speed regulation behaviours.

5.3.4. LMM and Ordered Logistic Regression Results for Internal Factors

To explore the potential influence of internal factors, including age, gender, driving experience, and ADAS familiarity, on the SA indicators, we incorporated these factors as fixed effects in LMMs. Before proceeding with the modelling, we calculated the Variance Inflation Factor (VIF) for these variables to assess potential multicollinearity. VIF is commonly used to identify collinearity issues among predictor variables in regression analysis. High VIF values indicate multicollinearity, which could distort model estimates.

The results of the VIF analysis are presented in Table 5.6. While *age* and *driving experience* exhibit high VIF values (37.79 and 43.11, respectively), which suggests multicollinearity, this is expected as these variables are logically related. Nevertheless, we decided to retain both variables in the models to allow for a more granular understanding of their individual contributions to the SA indicators. The other two variables, *gender* and *ADAS familiarity*, show acceptable VIF values (1.68 and 5.28, respectively), indicating limited multicollinearity concerns for these predictors.

Table 5.6: Variance Inflation Factor (VIF) Analysis for Internal Factors

Variable	VIF
Age	37.79
Gender	1.68
Driving Experience	43.11
ADAS Familiarity	5.28

Following this analysis, we proceeded to fit separate LMMs for each SA level's indicators. Each model includes internal factors as fixed effects and combines other random effects. The results for each level's indicators are illustrated as follows.

Based on the results, we first analysed SA1 gaze-related metrics and found significant effects for driving experience and familiarity, as well as a near-significant effect for age.

For *Time to First Fixation*, age was nearly significant ($p = 0.051$), suggesting that younger participants reacted slightly faster. Additionally, familiarity was strongly associated with faster reaction times ($p = 0.003$), indicating that drivers more accustomed to the ADAS system processed visual information more efficiently.

For *Fixation Duration*, both driving experience ($p = 0.030$) and familiarity ($p = 0.004$) had significant effects. More experienced drivers spent less time fixating, likely due to better anticipation, while familiar drivers processed visual information more efficiently. Gender did not have a significant effect on fixation patterns.

Overall, the results suggest that experience and familiarity enhance visual efficiency, and younger drivers tend to react faster.

Table 5.7: Mixed Linear Model Results for Internal Factors (SA1)

Indicator	Predictor	Estimate (β)	Std. Err.	z-value	p-value	95% CI
Fixation Count	Intercept	14.748	5.228	2.821	0.005	[4.502, 24.994]
	Turn 2	-0.092	0.488	-0.189	0.850	[-1.048, 0.863]
	Turn 3	-0.361	0.488	-0.740	0.460	[-1.316, 0.595]
	Turn 4	-0.470	0.488	-0.965	0.335	[-1.426, 0.485]
	Age	0.827	0.717	1.153	0.249	[-0.579, 2.232]
	Gender	-1.427	1.108	-1.288	0.198	[-3.598, 0.744]
	Driving Experience	-3.122	1.319	-2.366	0.018	[-5.707, -0.536]
	Familiarity	-1.927	0.639	-3.015	0.003	[-3.180, -0.674]
Time to First Fixation	Intercept	-31.570	14.094	-2.240	0.025	[-59.194, -3.947]
	Turn 2	-1.917	1.828	-1.040	0.294	[-5.500, 1.666]
	Turn 3	-0.706	1.828	-0.386	0.699	[-4.289, 2.877]
	Turn 4	-2.645	1.828	-1.447	0.148	[-6.228, 0.939]
	Age	-3.928	2.012	-1.953	0.051	[-7.871, 0.015]
	Gender	4.958	3.345	1.482	0.138	[-1.598, 11.514]
	Driving Experience	10.537	3.555	2.964	0.018	[3.569, 17.505]
	Familiarity	6.236	2.033	3.067	0.003	[2.251, 10.221]
Fixation Duration	Intercept	10.585	5.000	2.117	0.034	[0.785, 20.385]
	Turn 2	-0.545	0.507	-1.075	0.282	[-1.540, 0.449]
	Turn 3	-0.327	0.507	-0.645	0.519	[-1.321, 0.667]
	Turn 4	-0.800	0.507	-1.576	0.115	[-1.794, 0.195]
	Age	0.891	0.715	1.246	0.213	[-0.510, 2.293]
	Gender	-1.052	1.101	-0.956	0.339	[-3.210, 1.106]
	Driving Experience	-2.807	1.290	-2.176	0.030	[-5.336, -0.279]
	Familiarity	-1.791	0.618	-2.898	0.004	[-3.002, -0.580]

For SA2 ordinal data, an ordered logistic regression was conducted. None of the fixed effects were statistically significant, but familiarity ($p = 0.073$) showed a marginal effect, suggesting that participants familiar with the ADAS system might rate SA2 slightly lower.

Table 5.8: Ordered Logistic Regression Results for Internal Factors (SA2)

Indicator	Predictor	Estimate (β)	Std. Err.	z-value	p-value	95% CI
Score	Age	-0.6469	0.703	-0.921	0.357	[-2.024, 0.730]
	Gender	-1.8498	1.257	-1.472	0.141	[-4.313, 0.613]
	Driving Experience	1.5931	1.407	1.133	0.257	[-1.164, 4.350]
	Familiarity	-1.2108	0.676	-1.792	0.073	[-2.535, 0.113]
Thresholds	1.4/1.8	-3.7767	5.375	-0.703	0.482	[-14.311, 6.758]
	1.8/2.0	-0.1285	0.968	-0.133	0.894	[-2.025, 1.768]
	2.0/2.6	-0.0234	0.676	-0.035	0.972	[-1.348, 1.301]
	2.6/2.8	1.0713	0.303	3.541	0.000	[0.478, 1.664]
	2.8/3.2	-0.3320	0.689	-0.482	0.630	[-1.682, 1.018]
	3.0/3.2	-1.0321	0.978	-1.056	0.291	[-2.948, 0.884]
	3.4/3.6	-0.7951	0.968	-0.822	0.412	[-2.692, 1.102]

For SA 3 indicators *Speed Difference* and *Speed Compliance Percentage*, none of the fixed effects were statistically significant ($p > 0.05$), suggesting that age, gender, experience, and familiarity had limited influence.

Table 5.9: Mixed Linear Model Results for Internal Factors (SA3)

Indicator	Predictor	Estimate (β)	Std. Err.	z-value	p-value	95% CI
Speed Difference	Age	1.024	1.708	0.549	0.323	[-2.323, 4.371]
	Gender	0.173	1.028	0.168	0.867	[-1.843, 1.188]
	Driving Experience	-0.497	2.125	-0.234	0.815	[-4.661, 3.667]
	Familiarity	0.948	1.284	0.738	0.460	[-1.569, 3.465]
Speed Compliance Percentage	Age	-2.743	3.556	-0.771	0.441	[-9.713, 4.227]
	Gender	-1.425	2.774	-0.514	0.608	[-6.862, 4.013]
	Driving Experience	1.626	5.360	0.303	0.762	[-8.879, 12.131]
	Familiarity	-3.917	4.249	-0.922	0.357	[-12.246, 4.412]

The analysis of internal factors (age, gender, driving experience, and ADAS familiarity) revealed varying influences across SA levels. Driving experience and familiarity significantly enhanced visual efficiency in SA1, while age showed a marginal association with faster initial responses. For SA2 subjective ratings, no predictors reached significance, though familiarity approached a marginal effect. In SA3, none of the factors significantly impacted speed-related metrics. Despite collinearity between age and driving experience, retaining both variables provided nuanced insights. Overall, user characteristics played a stronger role in early-stage SA (e.g., visual processing) but diminished in later stages, underscoring context-dependent impacts on situational awareness.

6 Discussion and Conclusion

Overview

This chapter provides a comprehensive discussion of the study's results and their implications, while also addressing key research questions and situating the findings within the broader literature. Section 6.1 summarizes the main findings, highlighting the comparative effectiveness of Advanced Driver Assistance Systems (ADAS) Systems A and B in influencing driver situational awareness (SA). Section 6.2 addresses the research questions, demonstrating how specific ADAS signal designs impact SA at different levels and identifying the role of internal factors such as driving experience. Section 6.3 connects the study's findings to existing literature, emphasizing the methodological and content contributions of analyzing speed limit signals across SA levels. Then Section 6.4 discusses the limitations of the study and proposes future directions for enhancing the understanding and application of SA in ADAS contexts. Finally, some practical implications are shown in Section 6.5.

6.1. Key Findings

Based on the methodologies outlined in Chapter 4 and the analyses presented in Chapter 5, several key findings can be summarized as follows:

1. Most situational awareness indicators across different levels do not show statistically significant differences between ADAS System A and ADAS System B. The analyses of significance testing and mixed linear models (LMMs) jointly reveal that the alert signals provided by these systems during speed limit changes have comparable effects on drivers' situational awareness, indicating no substantial differences in their effectiveness. This conclusion highlights that both systems are equally reliable in the functionality emphasized in this study.
2. Certain indicators exhibit significant or marginally significant differences. For instance, as discussed in Sections 4.2.1 and 4.2.2, the *Fixation Count per Turn* in level 1 of situational awareness (SA1) and the *Scores* in level 2 of situational awareness (SA2) show variability between systems. These findings suggest that drivers' performance differs between ADAS systems at the first and second levels of situational awareness. Specifically, the *Fixation Count* in SA1 indicates that System B imposes a higher cognitive load. The marginally significant differences in SA2 *Scores* further support this observation. As shown in Figure 5.6, more participants tend to prefer ADAS System A, perceiving its assistance as clearer and more effective. These findings suggest that ADAS System A and ADAS System B differ in how they shape drivers' situational awareness, even though these differences do not ultimately impact driving performance. However, in terms of cognitive load, ADAS System A likely provides clearer and more explicit cues, helping drivers more quickly comprehend the current speed limit. Meanwhile, other indicators, including those related to level 3 of situational awareness (SA3), show no statistically significant differences between systems.
3. Observing all the indicators, ADAS System A consistently scores higher than ADAS System B across nearly all measures. From a general statistical perspective, ADAS System A appears to require lower cognitive load, receive higher understanding-level scores, and assist drivers in achieving better speed compliance rates. These results collectively suggest that ADAS System A's signal design provides clearer and more effective guidance, enhancing its utility for drivers.
4. As analysed in Section 4.2.1, the indicators of SA1, SA2, and SA3 are not significantly correlated. This is an important and somewhat unconventional finding, as conventional understanding suggests that these levels represent different layers of situational awareness and should be inherently connected. However, within the scope of this study, where these levels are separately examined, the independence across levels is both interpretable and acceptable. In the driving task, drivers' understanding of the environment (e.g., speed limit conditions) is not solely derived from ADAS signals but also from environmental observations, road signs, and their own experience and knowledge. In this study, SA1 focuses on drivers' visual attention to ADAS signals, which is only one component of situational awareness. SA2 captures drivers' understanding of these alerts, which may also stem from the aforementioned sources. Nonetheless, the lack of or weak correlations with SA3 highlights an important insight: drivers' comprehension of road information still largely relies on external tools, such as physical speed limit signs or in-car navigation apps, rather than solely on ADAS signals. A detailed comparison of all indicators further illustrates this finding (see Figure 5.7).
5. Internal factors influence certain aspects of situational awareness, particularly familiarity and driving experience. In SA1, familiarity with the ADAS system is associated with faster response times, while driving experience reduces fixation duration, suggesting that experienced and familiar drivers process visual information more efficiently. Age also shows a near-significant effect, indicating that younger participants tend to react faster. In SA2, familiarity has a marginal effect, with more familiar participants rating their understanding of ADAS alerts lower, possibly due to stricter self-evaluation. Other internal factors show no significant influence. Overall, experience and familiarity enhance visual efficiency, while the impact of internal factors on SA2 assessments and SA3 speed-related performance is limited.
6. The differences between the results of significance tests (including t-tests and Wilcoxon tests) and LMMs demonstrate that model selection has a significant impact on the outcomes. By combining

multiple methods, a more comprehensive analysis of the data can be achieved, ensuring that both simple and complex patterns are effectively captured.

6.2. Answers to the Research Questions

Based on the analysis in Chapter 4 and the findings in Section 5.1, we can now address the research question and its sub-questions.

Firstly, the sub-questions:

- How do different ADAS signal designs influence driver situational awareness across its three levels: perception, comprehension, and projection?

From the results of significance testing, the ADAS signal designs of different systems showed no statistically significant differences in shaping situational awareness levels of perception and projection. However, certain SA1 indicators exhibited significant differences. Based on these findings, we can conclude that the signal designs of the two ADAS systems underlying this study influence drivers' reception of signals, especially regarding the frequency of observing ADAS signals in different regions. While only a few indicators showed significant or marginally significant results for overall situational awareness, the collected data suggest that drivers performed better across all three levels of situational awareness when using System A. This implies that although subtle, System A exhibits a potential advantage.

- How do internal factors, such as age, gender, and other demographic characteristics, contribute to variations in driver situational awareness under different ADAS signal designs?

Internal factors influence the shaping of situational awareness, with familiarity and driving experience playing key roles. Participants with more driving experience exhibited shorter fixation durations, indicating reduced cognitive load, while familiarity with the ADAS system was associated with faster response times in SA1. Age also showed a near-significant effect, suggesting that younger participants tended to react more quickly. Other internal factors, including gender, did not exhibit strong significant effects.

- Based on the results, what design-related features might explain the differences in driver situational awareness and guide the development of more effective ADAS signal designs?

By observing the ADAS systems A and B, we can identify that ADAS System A provides both the current road speed limit and a semi-transparent indicator of the speed limit for the upcoming road section in the upper left corner. In contrast, ADAS System B uses a flashing red speed limit sign to indicate a new speed limit section. Comparatively, ADAS System A offers more comprehensive and easier-to-understand information. ADAS System B's signalling design may create potential confusion for drivers, making it unclear whether the signal refers to the current road or the upcoming road. This explains why drivers observed the area of interest more frequently in System B.

Next, we integrate the above conclusions to answer the main research question:

- How does driver situational awareness differ when responding to current road speed limit alerts generated by different ADAS systems?

In general, the two ADAS systems' speed limit signaling designs had no significant differences in their effects on driver situational awareness. However, the more comprehensive information provided by System A (current and next road speed limit information) potentially reduces drivers' cognitive load, making it easier for them to identify and comprehend. Additionally, both driving experience and familiarity with the ADAS system play critical roles in influencing a driver's ability to use these systems effectively. Experienced drivers are better at noticing and understanding ADAS prompts, while familiar drivers process ADAS information more efficiently, leading to quicker responses. The observed lack of correlations between overall situational awareness levels further indicates that drivers' final driving behaviour does not primarily rely on their perception and understanding of ADAS systems. At present, drivers may still prefer to rely on traditional tools, such as physical speed limit signs and their own experience, for driving decisions.

6.3. Integration with Existing Literature

Recent years have witnessed significant advancements in the study of situational awareness, marked by two key trends: the proliferation of measurement methods enabling more precise evaluations of situational awareness and the application of situational awareness concepts across a broader range of fields. While situational awareness research initially focused on domains such as aviation, nuclear power generation, and military systems, driving has long been a critical area of application [61]. Early studies demonstrated how situational awareness is essential for understanding the impact of technologies on driver performance by addressing spatial, temporal, goal, and system awareness [62]. This foundational work set the stage for integrating situational awareness into ADAS and human-machine interaction (HMI), which have become central to transportation research.

One key area of focus in recent research is the role of ADAS information displays in enhancing driver situational awareness. Kim et al. [63] highlighted that SA-based displays, especially in urgent situations, significantly improve driver trust, reduce cognitive workload, and enhance situational awareness, particularly at SA3 (projection). Their findings demonstrated that well-designed displays can enhance system transparency and situational trust, offering critical insights for ADAS design. Expanding on this, Biswas et al. [64] emphasized the use of eye-tracking data to capture drivers' attention and awareness dynamically. Their work underscored the value of gaze metrics, such as fixation duration, in assessing SA1. These studies collectively underscore the importance of tailored methods for evaluating situational awareness across its levels.

Additional insights into situational awareness evaluation come from Lu et al. [65], who investigated how drivers regain situational awareness after periods of inattention in automated driving. Their findings showed that spatial awareness (SA1) is typically achieved within 7–12 seconds, while assessing relative speeds requires over 20 seconds. This highlights the temporal nature of situational awareness development and the necessity of level-specific measurement approaches. Similarly, Avetisyan et al. [66] demonstrated that explanation modalities influence situational awareness differently across levels: visual explanations were most effective for SA 1 and SA2, while combining visual and auditory explanations enhanced SA3. Physiological responses provide invaluable insights into drivers' attention and cognitive processes during the takeover [67], further supporting the role of dynamic monitoring in situational awareness evaluation. In some studies, eye movement metrics such as gaze duration have even been directly utilized for the measurement and calculation of situational awareness [68]. These studies reinforce the need to select assessment methods that align with the specific cognitive demands of each level of situational awareness.

Our research builds on these insights with a methodological innovation: integrating different measurement approaches tailored to each level of situational awareness. For SA1, we utilize eye-tracking data to capture drivers' perceptual awareness. For SA2, we employ self-assessment methods with targeted questions to evaluate comprehension. For SA3, task performance metrics are used to reflect drivers' predictive capabilities and their control over driving tasks. This approach ensures that each level of situational awareness is measured using the most appropriate method, addressing the limitations of prior studies that often relied on single-measurement techniques.

From a content perspective, unlike traditional studies that either assess situational awareness as a whole or use a single method such as SAGAT to measure all three levels [69, 70, 71], our study adopts a layered analysis to examine the impact of ADAS information displays on different levels of situational awareness. We focus specifically on speed limit signals, which complement and extend existing studies on the effects of ADAS functionality on situational awareness.

In summary, recent research highlights the increasing relevance of situational awareness in ADAS and underscores the importance of tailored assessment methods. Our study builds on these developments by introducing a framework that evaluates situational awareness at all levels using level-specific methods and by examining how ADAS signals influence each level. The recency of related studies demonstrates the growing attention this field is receiving, and our work provides valuable insights to further advance research and inform the design of ADAS systems that enhance driver performance and safety.

6.4. Limitations and Future Directions

This study, while providing valuable insights into the relationship between ADAS information displays and situational awareness, has several limitations that should be acknowledged, alongside avenues for future research.

A key methodological limitation lies in the indicators selected for measuring situational awareness across different levels. Although robust methods were employed, incorporating additional indicators could enhance the comprehensiveness of the analysis. For instance, combining physiological measures such as heart rate variability with eye-tracking data for SA1 could provide a more nuanced understanding of perceptual awareness. Similarly, using behavioural observations (including eye-tracking and driving behaviour) alongside self-assessments for SA2 might unify various factors influencing comprehension into a cohesive framework. Expanding the self-assessment questions for SA2 to cover more aspects of comprehension would also improve the depth and breadth of the insights gained.

Another limitation relates to the size and diversity of the dataset. While the sample allowed for meaningful analysis, it was not sufficiently large to ensure broad generalizability. Additionally, all participants were recruited from Royal HaskoningDHV, which may have introduced sample representativeness issues. For instance, participants might have had greater familiarity with ADAS technology than the general driving population, potentially influencing their responses and interactions with the system.

The experimental design, conducted in real-world settings, presents another challenge. While field experiments offer high ecological validity, they also introduce confounding variables that are difficult to control. For instance, the presence of roadside speed limit signs or variations in traffic conditions could have influenced participants' responses to the ADAS signals. Conducting similar experiments in controlled environments, such as simulators or closed test tracks, could help minimize these external influences and offer more precise evaluations of the relationship between ADAS displays and situational awareness.

Furthermore, this study focused on speed limit signals as the primary ADAS functionality, which, while providing in-depth insights, does not capture the full range of ADAS capabilities. Future research could explore additional features, such as lane-keeping assistance, adaptive cruise control, or collision warning systems, to provide a more comprehensive understanding of how various functionalities interact with situational awareness at different levels.

Finally, the cross-sectional nature of this study limits its ability to capture changes in situational awareness over time. Longitudinal studies could examine how drivers adapt to ADAS systems and how their situational awareness evolves with prolonged use. Additionally, real-time data analysis during driving could provide dynamic feedback on how situational awareness changes in response to ADAS cues, offering richer insights into the relationship between drivers and these systems.

In addressing these limitations, future research can build on the findings of this study to deepen the understanding of situational awareness in ADAS contexts. By expanding methodological approaches, increasing dataset diversity, refining experimental designs, and broadening the scope of analysis, future work can contribute to the development of more effective and user-friendly ADAS systems, ultimately enhancing driver performance and safety.

6.5. Practical Implications

The findings of this study provide practical insights for both ADAS developers and transportation policymakers. The key difference between the two systems is that System A presents both the current road segment's speed limit and the upcoming segment's speed limit, whereas System B only displays the current limit. The results suggest that including anticipatory speed limit information enhances driver situational awareness, allowing them to prepare in advance for speed changes. This could reduce cognitive load and improve response efficiency, leading to smoother driving adjustments.

For ADAS manufacturers, these results suggest that signal design should prioritise clarity and anticipation. Displaying upcoming speed limits can help drivers make smoother transitions, potentially reducing sudden braking or unintended speeding. Future ADAS designs could integrate context-aware speed recommendations, adjusting alerts based on road type, traffic conditions, or driver behaviour.

For government agencies and transportation policymakers, these findings highlight the potential safety benefits of standardising speed limit transition warnings. While physical road signs remain essential, reinforcing these changes through ADAS displays could improve compliance, particularly in areas with frequent speed limit variations, such as school zones or construction sites. Policymakers may consider collaborating with ADAS developers to align digital alerts with official traffic regulations, ensuring consistency between in-vehicle warnings and roadside signage.

References

[1] Jaswanth Nidamanuri et al. “A progressive review: Emerging technologies for ADAS driven solutions”. In: *IEEE Transactions on Intelligent Vehicles* 7.2 (2021), pp. 326–341.

[2] Ben Nassi et al. “Phantom of the adas: Securing advanced driver-assistance systems from split-second phantom attacks”. In: *Proceedings of the 2020 ACM SIGSAC conference on computer and communications security*. 2020, pp. 293–308.

[3] John Waraniak. *Unsettled Issues on Sensor Calibration for Automotive Aftermarket Advanced Driver-Assistance Systems*. SAE International, 2021.

[4] Maria Merin Antony and Ruban Whenish. “Advanced driver assistance systems (ADAS)”. In: *Automotive Embedded Systems: Key Technologies, Innovations, and Applications*. Springer, 2021, pp. 165–181.

[5] Avinoam Borowsky, David Shinar, and Tal Oron-Gilad. “Age, skill, and hazard perception in driving”. In: *Accident analysis & prevention* 42.4 (2010), pp. 1240–1249.

[6] Alexandra S Mueller, Ian J Reagan, and Jessica B Cicchino. “Addressing driver disengagement and proper system use: human factors recommendations for level 2 driving automation design”. In: *Journal of cognitive engineering and decision making* 15.1 (2021), pp. 3–27.

[7] Maria Merin Antony and Ruban Whenish. “Advanced Driver Assistance Systems (ADAS)”. In: *Automotive Embedded Systems: Key Technologies, Innovations, and Applications*. Ed. by M. Kathiresh and R. Neelaveni. Cham: Springer International Publishing, 2021, pp. 165–181. ISBN: 978-3-030-59897-6. DOI: `10.1007/978-3-030-59897-6_9`. URL: `https://doi.org/10.1007/978-3-030-59897-6_9`.

[8] Henry Alexander Ignatious et al. “Analyzing factors influencing situation awareness in autonomous vehicles—A survey”. In: *Sensors* 23.8 (2023), p. 4075.

[9] Mica R Endsley. “Toward a theory of situation awareness in dynamic systems”. In: *Human factors* 37.1 (1995), pp. 32–64.

[10] Eduardo Salas et al. “Situation awareness in team performance: Implications for measurement and training”. In: *Human factors* 37.1 (1995), pp. 123–136.

[11] Paul M Salmon, Neville A Stanton, and Daniel P Jenkins. *Distributed situation awareness: Theory, measurement and application to teamwork*. CRC Press, 2017.

[12] Mica R Endsley, Daniel J Garland, et al. “Theoretical underpinnings of situation awareness: A critical review”. In: *Situation awareness analysis and measurement* 1.1 (2000), pp. 3–21.

[13] Arslan Munir, Alexander Aved, and Erik Blasch. “Situational awareness: techniques, challenges, and prospects”. In: *AI* 3.1 (2022), pp. 55–77.

[14] Myounghoon Jeon, Bruce N Walker, and Thomas M Gable. “The effects of social interactions with in-vehicle agents on a driver’s anger level, driving performance, situation awareness, and perceived workload”. In: *Applied ergonomics* 50 (2015), pp. 185–199.

[15] Klaus Bengler et al. “Three decades of driver assistance systems: Review and future perspectives”. In: *IEEE Intelligent transportation systems magazine* 6.4 (2014), pp. 6–22.

[16] Moataz Mohamed and Nicole F Bromfield. “Attitudes, driving behavior, and accident involvement among young male drivers in Saudi Arabia”. In: *Transportation research part F: traffic psychology and behaviour* 47 (2017), pp. 59–71.

[17] Francis T Durso et al. *Handbook of applied cognition*. John Wiley & Sons, 2007.

[18] J. Zhang and K. Zhang. “Theoretical Models, Measurement, and Application of Situational Awareness”. In: *Advances in Psychological Science* 12.5 (2004). Originally published in Chinese, pp. 725–732.

[19] Hongtao Zheng et al. “An interpretable prediction framework for multi-class situational awareness in conditionally automated driving”. In: *Advanced Engineering Informatics* 62 (2024), p. 102683.

[20] Xiaozhou Zhou et al. “Research on multilevel situation awareness changes under the cumulative effect of mental fatigue”. In: *Cognition, Technology & Work* 25.2 (2023), pp. 203–215.

[21] Peizheng Wang, Robert Houghton, and Arnab Majumdar. “Detecting and Predicting Pilot Mental Workload Using Heart Rate Variability: A Systematic Review”. In: *Sensors* 24.12 (2024), p. 3723.

[22] Mica R Endsley. “Measurement of situation awareness in dynamic systems”. In: *Human factors* 37.1 (1995), pp. 65–84.

[23] Mica R Endsley. “Situation awareness global assessment technique (SAGAT)”. In: *Proceedings of the IEEE 1988 national aerospace and electronics conference*. IEEE. 1988, pp. 789–795.

[24] David B Kaber and Mica R Endsley. “The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task”. In: *Theoretical issues in ergonomics science* 5.2 (2004), pp. 113–153.

[25] Ruiqi Ma and David B Kaber. “Situation awareness and driving performance in a simulated navigation task”. In: *Ergonomics* 50.8 (2007), pp. 1351–1364.

[26] Luke Petersen et al. “Situational awareness, drivers trust in automated driving systems and secondary task performance”. In: *arXiv preprint arXiv:1903.05251* (2019).

[27] Valerie Jane Gawron. *Human performance and situation awareness measures*. Crc Press, 2019.

[28] Joost CF De Winter et al. “Situation awareness based on eye movements in relation to the task environment”. In: *Cognition, Technology & Work* 21.1 (2019), pp. 99–111.

[29] Stephen J Selcon, Ronald M Taylor, and Eva Koritsas. “Workload or situational awareness?: TLX vs. SART for aerospace systems design evaluation”. In: *Proceedings of the Human Factors Society Annual Meeting*. Vol. 35. 2. SAGE Publications Sage CA: Los Angeles, CA. 1991, pp. 62–66.

[30] RM Taylor and SJ Selcon. “Cognitive quality and situational awareness with advanced aircraft attitude displays”. In: *Proceedings of the Human Factors Society Annual Meeting*. Vol. 34. 1. SAGE Publications Sage CA: Los Angeles, CA. 1990, pp. 26–30.

[31] Kavyaganga Kilingaru et al. “Monitoring pilot “Situation Awareness””. In: *J. Intell. Fuzzy Syst.* 24.3 (2013), pp. 457–466. ISSN: 1064-1246.

[32] Xidong Xu, Christopher D Wickens, and Esa M Rantanen. “Effects of conflict alerting system reliability and task difficulty on pilots’ conflict detection with cockpit display of traffic information”. In: *Ergonomics* 50.1 (2007), pp. 112–130.

[33] Minji Choi, Seungjun Ahn, and JoonOh Seo. “VR-Based investigation of forklift operator situation awareness for preventing collision accidents”. In: *Accident Analysis & Prevention* 136 (2020), p. 105404.

[34] Nadine Marie Moacdieh et al. “Effects of workload and workload transitions on attention allocation in a dual-task environment: Evidence from eye tracking metrics”. In: *Journal of Cognitive Engineering and Decision Making* 14.2 (2020), pp. 132–151.

[35] Chao Shi. *Eye-Tracking and Situation Awareness: Predicting The Malfunctions in Dynamic Control Tasks with a Mathematical Model for Different Age Groups and Interactive Devices*. The Pennsylvania State University, 2020.

[36] Ting Zhang et al. “Physiological measurements of situation awareness: a systematic review”. In: *Human factors* 65.5 (2023), pp. 737–758.

[37] AC Russo, MM Cardoso Junior, and E Villani. “Eye-tracking analysis to assess the mental load of unmanned aerial system operators: systematic review and future directions”. In: *The Aeronautical Journal* (2024), pp. 1–30.

[38] Taisir Alhilo and Akeel Al-Sakaa. “Eye Tracking Review: Importance, Tools, and Applications”. In: *International Conference on Emerging Trends and Applications in Artificial Intelligence*. Springer. 2023, pp. 383–394.

[39] Chandresh Sharma et al. “Eye gaze movement studies of control room operators: A novel approach to improve process safety”. In: *Computers & Chemical Engineering* 85 (2016), pp. 43–57.

[40] Ulrich J Pfeiffer, Kai Vogeley, and Leonhard Schilbach. “From gaze cueing to dual eye-tracking: novel approaches to investigate the neural correlates of gaze in social interaction”. In: *Neuroscience & Biobehavioral Reviews* 37.10 (2013), pp. 2516–2528.

[41] Ben Steichen, Cristina Conati, and Giuseppe Carenini. “Inferring visualization task properties, user performance, and user cognitive abilities from eye gaze data”. In: *ACM Transactions on Interactive Intelligent Systems (TiiS)* 4.2 (2014), pp. 1–29.

[42] Issam Tanoubi et al. “Comparing the visual perception according to the performance using the eye-tracking technology in high-fidelity simulation settings”. In: *Behavioral Sciences* 11.3 (2021), p. 31.

[43] Bryan A Wilbanks, Edwin Aroke, and Katherine M Dudding. “Using eye tracking for measuring cognitive workload during clinical simulations: literature review and synthesis”. In: *CIN: Computers, Informatics, Nursing* 39.9 (2021), pp. 499–507.

[44] Koen Van De Merwe, Henk Van Dijk, and Rolf Zon. “Eye movements as an indicator of situation awareness in a flight simulator experiment”. In: *The International Journal of Aviation Psychology* 22.1 (2012), pp. 78–95.

[45] Sylvia Peißl, Christopher D Wickens, and Rithi Baruah. “Eye-tracking measures in aviation: A selective literature review”. In: *The International Journal of Aerospace Psychology* 28.3-4 (2018), pp. 98–112.

[46] Céline M Muehlethaler and Chiara P Knecht. “Situation awareness training for general aviation pilots using eye tracking”. In: *IFAC-PapersOnLine* 49.19 (2016), pp. 66–71.

[47] Paul M Salmon et al. “Measuring Situation Awareness in complex systems: Comparison of measures study”. In: *International Journal of Industrial Ergonomics* 39.3 (2009), pp. 490–500.

[48] Mica R Endsley. “Direct measurement of situation awareness: Validity and use of SAGAT”. In: *Situational awareness*. Routledge, 2017, pp. 129–156.

[49] Jonathan Z Bakdash et al. “The validity of situation awareness for performance: a meta-analysis”. In: *Theoretical Issues in Ergonomics Science* 23.2 (2022), pp. 221–244.

[50] Tina Mioch, NJJM Smets, and Mark A Neerincx. “Predicting performance and situation awareness of robot operators in complex situations by unit task tests”. In: *the fifth international conference on advances in computer-human interactions*. 2012.

[51] Debra G Jones and Mica R Endsley. “Use of real-time probes for measuring situation awareness”. In: *The international journal of aviation psychology* 14.4 (2004), pp. 343–367.

[52] Debra G Jones and Mica R Endsley. “Can real-time probes provide a valid measure of situation awareness”. In: *Proceedings of the human performance, situation awareness and automation: user-centered design for the new millennium, Savannah, GA* (2000).

[53] Rensis Likert. “A technique for the measurement of attitudes”. In: *Archives of Psychology* (1932).

[54] V Onkhar, D Dodou, and JCF De Winter. “Evaluating the Tobii Pro Glasses 2 and 3 in static and dynamic conditions”. In: *Behavior Research Methods* 56.5 (2024), pp. 4221–4238.

[55] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. *Ultralytics YOLOv8*. Version 8.0.0. 2023. URL: `https://github.com/ultralytics/ultralytics`.

[56] Tzutalin. *LabelImg*. `https://github.com/tzutalin/labelImg`. Git code. 2015.

[57] SM Taheri and Gholamreza Hesamian. “A generalization of the Wilcoxon signed-rank test and its applications”. In: *Statistical Papers* 54 (2013), pp. 457–470.

[58] Steven T Garren and Grace H Davenport. “Using Kurtosis for Selecting One-Sample T-Test or Wilcoxon Signed-Rank Test”. In: *Current Journal of Applied Science and Technology* 41.18 (2022), pp. 46–55.

[59] Samuel Sanford Shapiro and Martin B Wilk. “An analysis of variance test for normality (complete samples)”. In: *Biometrika* 52.3-4 (1965), pp. 591–611.

[60] H Levene. *Robust tests for equality of variances*. 1960.

[61] Simon Banbury and Sebastien Tremblay. *A cognitive approach to situation awareness: Theory and application*. Routledge, 2017.

[62] Michael L Matthews et al. “Model for situation awareness and driving: Application to analysis and research for intelligent transportation systems”. In: *Transportation research record* 1779.1 (2001), pp. 26–32.

[63] Hwiseong Kim, Jeonguk Hong, and Sangwon Lee. “Investigating impact of situation awareness-based displays of semi-autonomous driving in urgent situations”. In: *Transportation research part F: traffic psychology and behaviour* 105 (2024), pp. 454–472.

[64] Abhijat Biswas et al. “Modeling Drivers’ Situational Awareness from Eye Gaze for Driving Assistance”. In: *8th Annual Conference on Robot Learning*. 2024.

[65] Zhenji Lu, Xander Coster, and Joost De Winter. “How much time do drivers need to obtain situation awareness? A laboratory-based study of automated driving”. In: *Applied ergonomics* 60 (2017), pp. 293–304.

[66] Lilit Avetisyan, Jackie Ayoub, and Feng Zhou. “Investigating explanations in conditional and highly automated driving: The effects of situation awareness and modality”. In: *Transportation research part F: traffic psychology and behaviour* 89 (2022), pp. 456–466.

[67] Gaojian Huang, Christine Petersen, and Brandon J Pitts. “The impact of mental states on semi-autonomous driving takeover performance: a systematic review”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 64. 1. SAGE Publications Sage CA: Los Angeles, CA. 2020, pp. 1372–1376.

[68] Shubham Agrawal and Srinivas Peeta. “Evaluating the impacts of situational awareness and mental stress on takeover performance under conditional automation”. In: *Transportation research part F: traffic psychology and behaviour* 83 (2021), pp. 210–225.

[69] Simon Cooper, Joanne Porter, and Linda Peach. “Measuring situation awareness in emergency setting: a systematic review of tools and outcomes”. In: *Open Access Emergency Medicine* (2014), pp. 1–7.

[70] Sabrina B Orique and Laurel Despins. “Evaluating situation awareness: an integrative review”. In: *Western journal of nursing research* 40.3 (2018), pp. 388–424.

[71] Thanh Nguyen et al. “A review of situation awareness assessment approaches in aviation environments”. In: *IEEE Systems Journal* 13.3 (2019), pp. 3590–3603.

A

Source Code

This section records the main codes used for visual detection and data processing in this study. The specific instructions and functions are described below.

```python
"""

The following code is used to train the yolov8 model and apply it to the detection task

"""

"""
    bibtex
    @software{yolov8_ultralytics,
      author = {Glenn Jocher and Ayush Chaurasia and Jing Qiu},
      title = {Ultralytics YOLOv8},
      version = {8.0.0},
      year = {2023},
      url = {https://github.com/ultralytics/ultralytics},
      orcid = {0000-0001-5950-6979, 0000-0002-7603-6750, 0000-0003-3783-7069},
      license = {AGPL-3.0}
    }
"""

from ultralytics import YOLO

model = YOLO('yolov8n.pt')
model.train(data='dateset.yaml', epochs=30, batch=10, imgsz=1080)

metrics = model.val(data='dataset.yaml')
print(metrics)

results = model.predict(source= 'video_frames')

for index, result in enumerate(results):  # Add index to track the current image number
    boxes = result.boxes
    focus_box = None
    screen_box = None
    image_name = result.path  # Assuming the result object has a 'path' attribute for image
        name

    for box in boxes:
        cls = int(box.cls[0])  # Get the class index
        name = model.names[cls]  # Use the class index to get the class name

        if name == 'focus':
            focus_box = box.xyxy[0]  # Get bounding box for 'focus'
        elif name == 'screen':
            screen_box = box.xyxy[0]  # Get bounding box for 'screen'

```

```
    # Check if both 'focus' and 'screen' are detected and determine if 'focus' is within '
        screen'
    if focus_box is not None and screen_box is not None:
        if (focus_box[0] >= screen_box[0] and  # Top-left of 'focus' is inside 'screen'
            focus_box[1] >= screen_box[1] and
            focus_box[2] <= screen_box[2] and  # Bottom-right of 'focus' is inside 'screen'
            focus_box[3] <= screen_box[3]):
            print(f"Image␣{index␣+␣1}␣({image_name}):␣'!!!focus'␣is␣inside␣'screen'")
        # else:
        #     print(f"Image {index + 1} ({image_name}): 'focus' is not inside 'screen'")
    # else:
    #     print(f"Image {index + 1} ({image_name}): 'focus' or 'screen' not detected")



"""

The following code is used to process the speed data related to SA Level3

"""

import pandas as pd
from datetime import timedelta

# Path to your Excel file
file_path = 'example.xlsx'

# Load the data from the first sheet
data = pd.read_excel(file_path)

# Convert the "Time" column to datetime objects
data['Time'] = pd.to_datetime(data['Time'], format='%H:%M:%S.%f')

# Add two hours to the "Time" column (keep it as a full datetime object)
data['Time'] = data['Time'] + timedelta(hours=2)
print(data.head())

# Set the "Time" column as the index (full datetime format)
data = data.set_index('Time')

# Perform a rolling window operation on the 'calculated_speed' column
# Rolling window of 2 seconds
rolling_window = data['calculated_speed'].rolling('2s', closed="right", center=True).mean()

# Merge the result back into the original DataFrame
result = pd.merge(data.reset_index(), pd.DataFrame(rolling_window).reset_index(), on="Time",
    suffixes=('', '_rolling_mean'))

data = result

# Rename the existing 'speed_difference' column for clarity
data = data.rename(columns={'calculated_speed_rolling_mean': 'rolling_speed'})

# Create a shifted version of the 'calculated_speed' column
data['previous_speed'] = data['calculated_speed'].shift(1)

# Calculate the mean speed between the current and previous points
data['mean_speed'] = (data['calculated_speed'] + data['previous_speed']) / 2

# Remove the first row where previous_speed is NaN due to the shift operation
data = data.dropna(subset=['mean_speed'])

# Create a new column for the difference between actual speed and rolling speed
data['rolling_actual_speed_difference'] = data['rolling_speed'] - data['calculated_speed']

# Rename the existing 'speed_difference' column for clarity
data = data.rename(columns={'speed_difference': 'speed_limit_difference'})

def filter_unreliable_data(data, threshold):
    """
```

```
    Filters out unreliable data points based on the absolute value of rolling_actual_speed_difference.

    Parameters:
    data (DataFrame): The input DataFrame containing the rolling_actual_speed_difference column.
    threshold (float): The threshold value for filtering unreliable data.

    Returns:
    DataFrame: The filtered DataFrame.
    """
    # Filter out rows where the absolute value of rolling_actual_speed_difference is greater than the threshold
    filtered_data = data[abs(data['rolling_actual_speed_difference']) <= threshold]

    return filtered_data

# Example usage with a threshold of 20 (adjust as needed)
threshold_value = 20
filtered_data = filter_unreliable_data(data, threshold_value)

import folium
from branca.element import Template, MacroElement

def df(data, speed_diff_column='calculated_speed_difference', threshold=10, file_name='colored_map.html'):
    ""
    This function takes in a DataFrame 'data', a column name for speed difference,
    a speed difference threshold, and a file name to generate a folium map.

    Parameters:
    data (DataFrame): The input DataFrame containing GPS coordinates and speed data
    speed_diff_column (str): The name of the column containing the speed difference
    threshold (int): The threshold to classify speed differences
    file_name (str): The name of the file where the map will be saved
    ""

    # Function to generate a color based on the speed difference
    def get_color(diff, threshold):
        if diff <= 0:
            return 'green'
        elif 0 < diff <= threshold:
            return 'orange'
        else:
            return 'red'

    # Initialize a map centered at the average location
    center_lat = data['GPS␣(Lat.)␣[deg]'].mean()
    center_long = data['GPS␣(Long.)␣[deg]'].mean()
    mymap = folium.Map(location=[center_lat, center_long], zoom_start=14)

    # Draw lines between consecutive GPS points
    for i in range(1, len(data)):
        prev_row = data.iloc[i-1]
        curr_row = data.iloc[i]

        # Dynamically use the speed difference column passed as a parameter
        speed_diff = curr_row[speed_diff_column]

        # Get the color for the line based on speed difference
        color = get_color(speed_diff, threshold)

        # Extract latitude and longitude for the previous and current point
        coordinates = [
            [prev_row['GPS␣(Lat.)␣[deg]'], prev_row['GPS␣(Long.)␣[deg]']],
            [curr_row['GPS␣(Lat.)␣[deg]'], curr_row['GPS␣(Long.)␣[deg]']]
        ]

        # Draw a line between the two points
        folium.PolyLine(
            locations=coordinates,
```

```
            color=color,
            weight=5,  # Line thickness
            opacity=0.8,
            popup=f"Speed difference: {speed_diff} km/h"
        ).add_to(mymap)

    # Add a legend to the map
    legend_html = """
    <div style="position: fixed;
        bottom: 50px; left: 50px; width: 200px; height: 130px;
        background-color: white; z-index:9999; font-size:14px;
        border:2px solid grey; padding: 10px;">
        <b>Speed Difference Legend</b><br>
        <i style="background:green; width:20px; height:20px; float:left; margin-right:5px"></i> Below or at Limit <br>
        <i style="background:orange; width:20px; height:20px; float:left; margin-right:5px"></i> 0 < diff <= {threshold} <br>
        <i style="background:red; width:20px; height:20px; float:left; margin-right:5px"></i> Above {threshold} <br>
    </div>
    """.format(threshold=threshold)

    # Add legend to the map as a custom HTML element
    mymap.get_root().html.add_child(folium.Element(legend_html))

    # Save the map to an HTML file
    mymap.save(file_name)
```

B

Scoring Approach for SA Level 2

This section records the main codes used for visual detection and data processing in this study. The specific instructions and functions are described below.

Question 1: Sequence of Speed Limit Changes

The first question in the post-drive questionnaire aimed to evaluate participants' awareness of ADAS alerts and their comprehension of overall speed limit transitions during the experiment. Participants were presented with multiple-choice options, including the correct sequence and some plausible alternatives, as well as the option to indicate they did not notice any speed limit change. This question was scored based on the participants' response accuracy.

The scoring method for this question is summarized in Table B.1. Participants who selected the option "I did not notice any speed limit change" were given the lowest score of 1. For those selecting one of the speed limit sequences, the score was determined based on the percentage of correct speed limits in their chosen sequence compared to the actual sequence observed during the experiment. The scoring breakdown is as follows:

Table B.1: Scoring for Question 1: Sequence of Speed Limit Changes

Option	Score
60-50-30-50-30 (Correct Sequence)	5
60-30-50-60-40 (Partially Correct)	3
60-50-60-30-50 (Partially Correct)	2
I did not notice any speed limit change	1

Question 2: Understanding of ADAS Brand A Warning Signals (a)

Question 2 in the post-drive questionnaire assessed participants' understanding of ADAS Brand A's warning signals. Specifically, participants were asked to identify the position of the speed limit warning on the display. The scoring method for this question is summarized in Table B.2. Selecting the correct position earned a score of 5, while selecting an incorrect position or indicating they did not notice the warning resulted in a score of 1.

Table B.2: Scoring for Question 2: Understanding of ADAS Brand A Warning Signals (a)

Option	Score
Upper left (Correct Answer)	5
Lower right (Incorrect Answer)	1
I didn't notice this warning (Incorrect)	1

Question 3: Understanding of ADAS Brand A Warning Signals (b)

This question assessed participants' understanding of ADAS Brand A's warning signals during speed limit transitions from high to low. The correct answer was "A flashing speed limit sign on the upper left of the present sign," which was assigned the highest score of 5. Other options received lower scores, as shown in Table B.3.

Table B.3: Scoring for Question 3: Understanding of ADAS Brand A Warning Signals (b)

Option	Score
A flashing speed limit sign on the upper left of the present sign	5
A flashing speed limit sign in red numbers	3
An arrow in the direction of down	2
I didn't notice any signals	1

Question 4: Confidence in Understanding ADAS Brand A Warnings

This question assessed participants' confidence in understanding the warnings provided by ADAS Brand A during the experiment. Participants rated their confidence levels using a five-point Likert scale ranging from "Very confident" to "Not confident at all." The scoring method is presented in Table B.4.

Table B.4: Scoring for Question 4: Confidence in Understanding ADAS Brand A Warnings

Option	Score
Very confident	5
Confident	4
Neutral	3
Slightly confident	2
Not confident at all	1

Question 5: Comfort with ADAS Brand A Assistance

This question evaluated participants' comfort levels while driving with the assistance of ADAS Brand A's speed change warnings. Similar to the previous question, participants rated their comfort using a five-point Likert scale. The scoring is summarized in Table B.5.

Table B.5: Scoring for Question 5: Comfort with ADAS Brand A Assistance

Option	Score
Very comfortable	5
Comfortable	4
Neutral	3
Slightly comfortable	2
Not comfortable at all	1

Question 6: Understanding of ADAS Brand B Warning Signals (a)

The following four questions mirror the structure and content of Questions 2–5, but focus on participants' understanding and comfort with ADAS Brand B. Each question's scoring method and options are identical to those used for ADAS Brand A, except that the context refers to Brand B. The scoring breakdown for each question is summarized below.

Table B.6: Scoring for Question 6: Understanding of ADAS Brand B Warning Signals (a)

Option	Score
Increase speed limit (Correct Answer)	5
Decrease speed limit (Incorrect Answer)	1
Maintain current speed (Incorrect Answer)	1
I didn't notice this warning (Incorrect)	1

Question 7: Understanding of ADAS Brand B Warning Signals (b)

Table B.7: Scoring for Question 7: Understanding of ADAS Brand B Warning Signals (b)

Option	Score
A flashing speed limit sign in red numbers	5
A flashing speed limit sign on the upper left of the present sign	3
An arrow in the direction of down	2
I didn't notice any signals	1

Question 8: Confidence in Understanding ADAS Brand B Warnings

Table B.8: Scoring for Question 8: Confidence in Understanding ADAS Brand B Warnings

Option	Score
Very confident	5
Confident	4
Neutral	3
Slightly confident	2
Not confident at all	1

Question 9: Comfort with ADAS Brand B Assistance

Table B.9: Scoring for Question 9: Comfort with ADAS Brand B Assistance

Option	Score
Very comfortable	5
Comfortable	4
Neutral	3
Slightly comfortable	2
Not comfortable at all	1

C
Questionnaire

This section records the questionnaire used to collect data in this study, which consists of two parts: pre-drive questionnaire and post-drive questionnaire.

(Pre-Drive) Characterzing Situation Awareness of Transport Operators based on Eye-Tracking

1. What is your participant ID?

2. What is your age?

- Under 18
- 18-24
- 25-34
- 35-44
- 45-54
- 55-64
- 65 or older

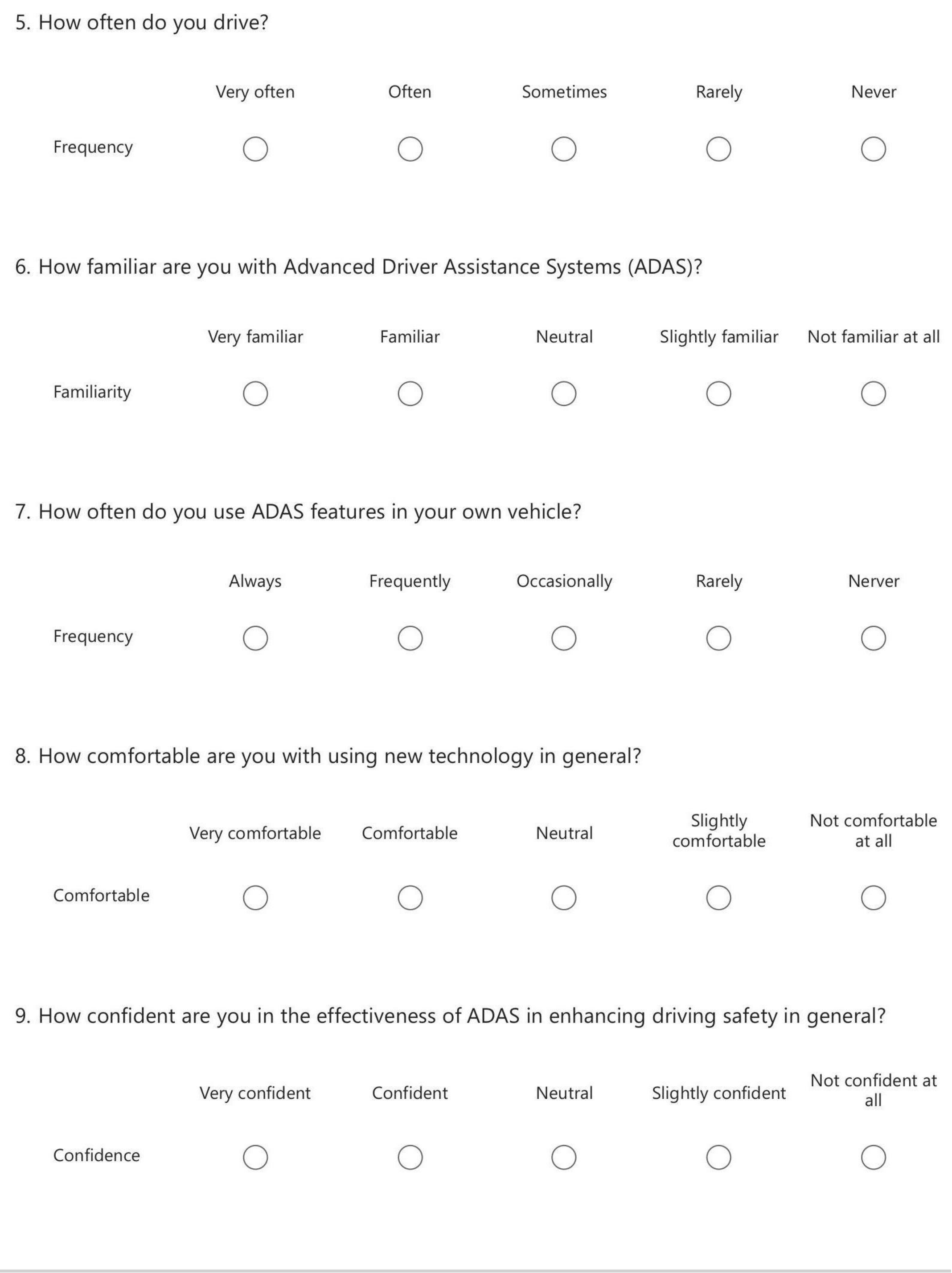
3. What is your gender?

- Male
- Female
- Non
- Non-binary
- Prefer not to say

4. Hoe many years have you been driving?

- Less than 1 year
- 1-3 years
- 4-7 years
- 8-15 years
- More than 15 years

Figure C.1: Pre-Drive Questionnaire Part 1



5. How often do you drive?

	Very often	Often	Sometimes	Rarely	Never
Frequency	○	○	○	○	○

6. How familiar are you with Advanced Driver Assistance Systems (ADAS)?

	Very familiar	Familiar	Neutral	Slightly familiar	Not familiar at all
Familiarity	○	○	○	○	○

7. How often do you use ADAS features in your own vehicle?

	Always	Frequently	Occasionally	Rarely	Nerver
Frequency	○	○	○	○	○

8. How comfortable are you with using new technology in general?

	Very comfortable	Comfortable	Neutral	Slightly comfortable	Not comfortable at all
Comfortable	○	○	○	○	○

9. How confident are you in the effectiveness of ADAS in enhancing driving safety in general?

	Very confident	Confident	Neutral	Slightly confident	Not confident at all
Confidence	○	○	○	○	○

Figure C.2: Pre-Drive Questionnaire Part 2

(Post-Drive) Characterzing Situation Awareness of Transport Operators based on Eye-Tracking

1

What is your Participant ID?

2

Which vehicle's ADAS speed limit warning did you find clearer?

- Brand A
- Brand B
- No preference

3

What was the sequence of speed limit changes during the experiment (excluding the familiarization part)?

- 60-50-30-50-30
- 60-30-50-60-40
- 60-50-60-30-50
- I did not notice any speed limit change

4

How did the traffic conditions influence your response to the Intelligent Speed Adaptation(ISAs) speed change warnings?

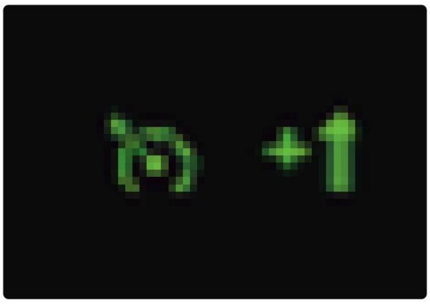
ISAs is a system which informs, warns and discourages the driver to exceed the statutory local speed limit. The in-vehicle speed limit is set automatically as a function of the speed limits indicated on the road.

	Very influential	Influential	Neutral	Slightly influential	Not influential at all
Influence	○	○	○	○	○

Figure C.3: Post-Drive Questionnaire Part 1

5

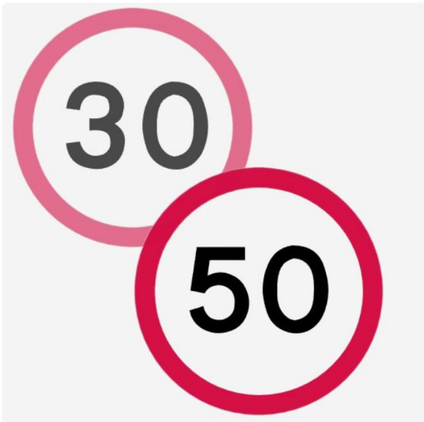
What does this speed change warning from the ADAS system indicate?



- ◯ Increase speed limit
- ◯ Decrease speed limit
- ◯ Maintain current speed
- ◯ I didn’t notice this warning

6

Which sign from the ADAS system indicates the speed limit of the present road section?



- ◯ Upper left
- ◯ Lower right
- ◯ I didn’t notice this warning

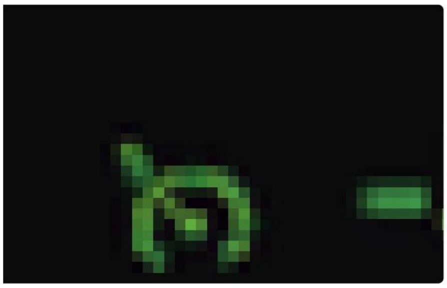
Figure C.4: Post-Drive Questionnaire Part 2

7

What warning signals did you see as the speed limit changed from high to low during the experiment when driving ADAS Brand A? Choose all that apply.

◯ A flashing speed limit sign on the upper left of the present sign.

◯ A flashing speed limit sign in red numbers.



◯ An arrow in the direction of down.

◯ I didn't notice any signals.

8

How confident were you in understanding the ISAs warnings when driving ADAS Brand A?

	Very confident	Confident	Neutral	Slightly confident	Not confident at all
Confidence	◯	◯	◯	◯	◯

9

How confident were you in predicting the changes in the driving environment based on the ISAs warnings you received when driving ADAS Brand A?

	Very confident	Confident	Neutral	Slightly confident	Not confident at all
Confidence	◯	◯	◯	◯	◯

10

How comfortable did you feel with the assistance of ADAS speed change warnings when driving ADAS Brand A?

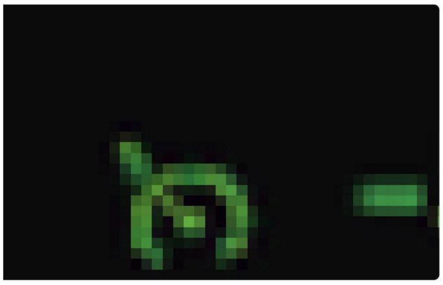
	Very comfortable	Comfortable	Neutral	Slightly comfortable	Not comfortable at all
Comfort	◯	◯	◯	◯	◯

11

What warning signals did you see as the speed limit changed from high to low during the experiment when driving ADAS Brand B? Choose all that apply.

◯ A flashing speed limit sign on the upper left of the present sign.

◯ A flashing speed limit sign in red numbers.



◯ An arrow in the direction of down.

◯ I didn't notice any signals.

12

How confident were you in understanding the ISAs warnings when driving ADAS Brand B?

	Very confident	Confident	Neutral	Slightly confident	Not confident at all
Confidence	◯	◯	◯	◯	◯

13

How confident were you in predicting the changes in the driving environment based on the ISAs warnings you received when driving ADAS Brand B?

	Very confident	Confident	Neutral	Slightly confident	Not confident at all
Confidence	◯	◯	◯	◯	◯

14

How comfortable did you feel with the assistance of ADAS speed change warnings when driving ADAS Brand B?

	Very comfortable	Comfortable	Neutral	Slightly comfortable	Not comfortable at all
Comfort	◯	◯	◯	◯	◯

D

LMM Model Results

This appendix contains the detailed results of the LMM models.

```
Fitting LMM for Indicator: BoxCox_Fixation_Count...
               Mixed Linear Model Regression Results
===================================================================
Model:              MixedLM Dependent Variable: BoxCox_Fixation_Count
No. Observations: 80      Method:             REML
No. Groups:         10      Scale:              2.3947
Min. group size:    8       Log-Likelihood:     -157.5371
Max. group size:    8       Converged:          Yes
Mean group size:    8.0
---------------------------------------------------------------------
                  Coef.    Std.Err.     z      P>|z|   [0.025   0.975]
---------------------------------------------------------------------
Intercept        -1.177      0.675   -1.744   0.081   -2.500    0.146
System[T.B]       0.305      0.346    0.881   0.378   -0.373    0.983
Turn[T.2]        -0.092      0.489   -0.189   0.850   -1.051    0.867
Turn[T.3]        -0.361      0.489   -0.737   0.461   -1.320    0.599
Turn[T.4]        -0.470      0.489   -0.961   0.336   -1.430    0.489
Group Var         3.060      1.091
===================================================================
```

Figure D.1: Mixed Linear Model Regression Results for Fixation Count per Turn

```
Fitting LMM for Indicator: BoxCox_Time_to_First...
| | | |   Mixed Linear Model Regression Results
==================================================================
Model:             MixedLM Dependent Variable: BoxCox_Time_to_First
No. Observations: 80      Method:             REML
No. Groups:        10      Scale:              34.0584
Min. group size:   8       Log-Likelihood:     -254.6470
Max. group size:   8       Converged:          Yes
Mean group size:   8.0
------------------------------------------------------------------
| | | |         Coef.   Std.Err.    z     P>|z|  [0.025   0.975]
------------------------------------------------------------------
Intercept      16.398    2.116   7.750  0.000  12.251   20.545
System[T.B]     0.225    1.305   0.172  0.863  -2.333    2.782
Turn[T.2]      -1.917    1.845  -1.039  0.299  -5.534    1.700
Turn[T.3]      -0.706    1.845  -0.383  0.702  -4.323    2.911
Turn[T.4]      -2.645    1.845  -1.433  0.152  -6.262    0.973
Group Var      23.484    2.389
==================================================================
```

Figure D.2: Mixed Linear Model Regression Results for Time to First Fixation

```
Fitting LMM for Indicator: BoxCox_Fixation_Duration...
| | | |    Mixed Linear Model Regression Results
=====================================================================
Model:             MixedLM Dependent Variable: BoxCox_Fixation_Duration
No. Observations: 80      Method:             REML
No. Groups:        10      Scale:              2.5940
Min. group size:   8       Log-Likelihood:     -159.5627
Max. group size:   8       Converged:          Yes
Mean group size:   8.0
---------------------------------------------------------------------
| | | |        Coef.    Std.Err.     z      P>|z|   [0.025    0.975]
---------------------------------------------------------------------
Intercept      -3.260     0.650   -5.014   0.000   -4.535   -1.986
System[T.B]     0.306     0.360    0.849   0.396   -0.400    1.012
Turn[T.2]      -0.545     0.509   -1.071   0.284   -1.544    0.453
Turn[T.3]      -0.327     0.509   -0.642   0.521   -1.325    0.671
Turn[T.4]      -0.800     0.509   -1.570   0.116   -1.798    0.199
Group Var       2.607     0.915
=====================================================================
```

Figure D.3: Mixed Linear Model Regression Results for Fixation Duration

```
                             OrderedModel Results
==============================================================================
Dep. Variable:                  Score   Log-Likelihood:                -35.222
Model:                   OrderedModel   AIC:                             88.44
Method:            Maximum Likelihood   BIC:                             97.41
No. Observations:                  20
Df Residuals:                      11
Df Model:                           1
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
System_B      -1.0223      0.843     -1.213      0.225      -2.674       0.629
1.4/1.8       -3.5558      1.164     -3.055      0.002      -5.837      -1.275
1.8/2.0       -0.2538      0.987     -0.257      0.797      -2.189       1.681
2.0/2.6       -0.1659      0.683     -0.243      0.808      -1.505       1.173
2.6/2.8        0.7251      0.294      2.468      0.014       0.149       1.301
2.8/3.0       -0.6925      0.676     -1.024      0.306      -2.018       0.633
3.0/3.2       -1.1701      0.974     -1.201      0.230      -3.080       0.740
3.2/3.4       -1.0221      0.977     -1.046      0.296      -2.937       0.893
3.4/3.6        0.2076      0.708      0.293      0.769      -1.180       1.595
==============================================================================
```

Figure D.4: Ordered Logistic Regression Results for SA2 Scores

```
Fitting LMM for Indicator: Speed_Difference...
            Mixed Linear Model Regression Results
==============================================================
Model:             MixedLM Dependent Variable: Speed_Difference
No. Observations: 80      Method:             REML
No. Groups:       10      Scale:              7.0662
Min. group size:  8       Log-Likelihood:     -195.6257
Max. group size:  8       Converged:          Yes
Mean group size:  8.0
--------------------------------------------------------------
                Coef.  Std.Err.    z    P>|z|  [0.025  0.975]
--------------------------------------------------------------
Intercept      14.574     0.961  15.167 0.000  12.691  16.458
System[T.B]    -0.058     0.594  -0.097 0.923  -1.223   1.107
Turn[T.2]      -5.209     0.841  -6.196 0.000  -6.856  -3.561
Turn[T.3]       2.429     0.841   2.889 0.004   0.781   4.076
Turn[T.4]      -4.450     0.841  -5.293 0.000  -6.097  -2.802
Group Var       4.818     1.078
==============================================================
```

Figure D.5: Mixed Linear Model Regression Results for Speed Difference

```
Fitting LMM for Indicator: Speed_Compliance_Percentage...
                  Mixed Linear Model Regression Results
==========================================================================
Model:            MixedLM Dependent Variable: Speed_Compliance_Percentage
No. Observations: 80      Method:             REML
No. Groups:       10      Scale:              128.3482
Min. group size:  8       Log-Likelihood:     -299.3695
Max. group size:  8       Converged:          Yes
Mean group size:  8.0
-----------------------------------------------------------------------------
                   Coef.     Std.Err.      z       P>|z|    [0.025    0.975]
-----------------------------------------------------------------------------
Intercept          77.907       3.137    24.838    0.000    71.759    84.054
System[T.B]        -1.925       2.533    -0.760    0.447    -6.890     3.041
Turn[T.2]           0.584       3.583     0.163    0.871    -6.438     7.606
Turn[T.3]           1.598       3.583     0.446    0.656    -5.424     8.620
Turn[T.4]          -0.774       3.583    -0.216    0.829    -7.796     6.248
Group Var          18.161       1.517
==========================================================================
```

Figure D.6: Mixed Linear Model Regression Results for Speed Compliance Percentage

```
Fitting LMM for Indicator: BoxCox_Fixation_Count...
               Mixed Linear Model Regression Results
==================================================================
Model:            MixedLM Dependent Variable: BoxCox_Fixation_Count
No. Observations: 80      Method:             REML
No. Groups:       10      Scale:              2.3771
Min. group size:  8       Log-Likelihood:     -150.7549
Max. group size:  8       Converged:          Yes
Mean group size:  8.0
------------------------------------------------------------------
                         Coef.  Std.Err.   z    P>|z| [0.025 0.975]
------------------------------------------------------------------
Intercept                14.748    5.228  2.821 0.005  4.502 24.994
Turn[T.2]                -0.092    0.488 -0.189 0.850 -1.048  0.863
Turn[T.3]                -0.361    0.488 -0.740 0.460 -1.316  0.595
Turn[T.4]                -0.470    0.488 -0.965 0.335 -1.426  0.485
age                       0.827    0.717  1.153 0.249 -0.579  2.232
gender                   -1.427    1.108 -1.288 0.198 -3.598  0.744
driving_experiences      -3.122    1.319 -2.366 0.018 -5.707 -0.536
familiarity              -1.927    0.639 -3.015 0.003 -3.180 -0.674
Group Var                 1.276    0.823
Group x System[T.B] Cov   0.175
System[T.B] Var           0.029
==================================================================
```

Figure D.7: LMM Results for SA1 Indicator with Internal Factors: Fixation Count

```
Fitting LMM for Indicator: BoxCox_Time_to_First...
              Mixed Linear Model Regression Results
====================================================================
Model:              MixedLM Dependent Variable: BoxCox_Time_to_First
No. Observations:   80      Method:             REML
No. Groups:         10      Scale:              33.4209
Min. group size:    8       Log-Likelihood:     -244.5281
Max. group size:    8       Converged:          Yes
Mean group size:    8.0
--------------------------------------------------------------------
                         Coef.  Std.Err.   z    P>|z|  [0.025 0.975]
--------------------------------------------------------------------
Intercept               -31.570   14.094 -2.240 0.025 -59.194 -3.947
Turn[T.2]                -1.917    1.828 -1.049 0.294  -5.500  1.666
Turn[T.3]                -0.706    1.828 -0.386 0.699  -4.289  2.877
Turn[T.4]                -2.645    1.828 -1.447 0.148  -6.228  0.939
age                      -3.928    2.012 -1.953 0.051  -7.871  0.015
gender                    4.958    3.345  1.482 0.138  -1.598 11.514
driving_experiences      10.537    3.555  2.964 0.003   3.569 17.505
familiarity               6.236    2.033  3.067 0.002   2.251 10.221
Group Var                12.305    2.300
Group x System[T.B] Cov  -2.863    1.955
System[T.B] Var           0.722    2.900
====================================================================
```

Figure D.8: LMM Results for SA1 Indicator with Internal Factors: Time to First Fixation

```
Fitting LMM for Indicator: BoxCox_Fixation_Duration...
                Mixed Linear Model Regression Results
======================================================================
Model:            MixedLM Dependent Variable: BoxCox_Fixation_Duration
No. Observations: 80      Method:             REML
No. Groups:       10      Scale:              2.5729
Min. group size:  8       Log-Likelihood:     -153.4857
Max. group size:  8       Converged:          Yes
Mean group size:  8.0
----------------------------------------------------------------------
                           Coef.  Std.Err.   z    P>|z| [0.025 0.975]
----------------------------------------------------------------------
Intercept                  10.585    5.000  2.117 0.034  0.785 20.385
Turn[T.2]                  -0.545    0.507 -1.075 0.282 -1.540  0.449
Turn[T.3]                  -0.327    0.507 -0.645 0.519 -1.321  0.667
Turn[T.4]                  -0.800    0.507 -1.576 0.115 -1.794  0.195
age                         0.891    0.715  1.246 0.213 -0.510  2.293
gender                     -1.052    1.101 -0.956 0.339 -3.210  1.106
driving_experiences        -2.807    1.290 -2.176 0.030 -5.336 -0.279
familiarity                -1.791    0.618 -2.898 0.004 -3.002 -0.580
Group Var                   1.315
Group x System[T.B] Cov     0.171
System[T.B] Var             0.029
======================================================================
```

Figure D.9: LMM Results for SA1 Indicator with Internal Factors: Fixation Duration

```
                              OrderedModel Results
==============================================================================
Dep. Variable:                  Score   Log-Likelihood:                -29.967
Model:                   OrderedModel   AIC:                             83.93
Method:            Maximum Likelihood   BIC:                             95.88
No. Observations:                  20
Df Residuals:                       8
Df Model:                           4
=======================================================================================
                          coef    std err          z      P>|z|      [0.025      0.975]
---------------------------------------------------------------------------------------
age                    -0.6469      0.703     -0.921      0.357      -2.024       0.730
gender                 -1.8498      1.257     -1.472      0.141      -4.313       0.613
driving_experiences     1.5931      1.407      1.133      0.257      -1.164       4.350
familiarity            -1.2108      0.676     -1.792      0.073      -2.535       0.113
1.4/1.8                -3.7767      5.375     -0.703      0.482     -14.311       6.758
1.8/2.0                -0.1285      0.968     -0.133      0.894      -2.025       1.768
2.0/2.6                -0.0234      0.676     -0.035      0.972      -1.348       1.301
2.6/2.8                 1.0713      0.303      3.541      0.000       0.478       1.664
2.8/3.0                -0.3320      0.689     -0.482      0.630      -1.682       1.018
3.0/3.2                -1.0321      0.978     -1.056      0.291      -2.948       0.884
3.2/3.4                -0.7951      0.968     -0.822      0.411      -2.692       1.102
3.4/3.6                 0.5420      0.681      0.796      0.426      -0.793       1.877
=======================================================================================
```

Figure D.10: Ordered Model Results for SA2 Indicators with Internal Factors

```
Fitting LMM for Indicator: Speed_Difference...
              Mixed Linear Model Regression Results
==================================================================
Model:               MixedLM   Dependent Variable:   Speed_Difference
No. Observations:    80        Method:               REML
No. Groups:          10        Scale:                6.9583
Min. group size:     8         Log-Likelihood:       -191.0530
Max. group size:     8         Converged:            Yes
Mean group size:     8.0
------------------------------------------------------------------
                        Coef.  Std.Err.   z    P>|z| [0.025 0.975]
------------------------------------------------------------------
Intercept               14.540    0.990 14.683 0.000 12.599 16.481
Turn[T.2]               -5.209    0.834 -6.244 0.000 -6.844 -3.574
Turn[T.3]                2.429    0.834  2.912 0.004  0.794  4.064
Turn[T.4]               -4.450    0.834 -5.334 0.000 -6.085 -2.815
age                      1.024    1.708  0.600 0.549 -2.323  4.371
gender                   0.173    1.028  0.168 0.867 -1.843  2.188
driving_experiences     -0.497    2.125 -0.234 0.815 -4.661  3.667
familiarity              0.948    1.284  0.738 0.460 -1.569  3.465
Group Var                6.652    2.279
Group x System[T.B] Cov -0.345    1.645
System[T.B] Var          0.018
==================================================================
```

Figure D.11: LMM Results for SA3 Indicator with Internal Factors: Speed Difference

```
Fitting LMM for Indicator: Speed_Compliance_Percentage...
                  Mixed Linear Model Regression Results
=========================================================================
Model:            MixedLM Dependent Variable: Speed_Compliance_Percentage
No. Observations: 80      Method:             REML
No. Groups:       10      Scale:              125.2098
Min. group size:  8       Log-Likelihood:     -291.9032
Max. group size:  8       Converged:          Yes
Mean group size:  8.0
--------------------------------------------------------------------------
                          Coef.  Std.Err.    z     P>|z|   [0.025  0.975]
--------------------------------------------------------------------------
Intercept                 77.230    2.874  26.872  0.000   71.597  82.863
Turn[T.2]                  0.584    3.539   0.165  0.869   -6.351   7.519
Turn[T.3]                  1.598    3.539   0.452  0.652   -5.337   8.533
Turn[T.4]                 -0.774    3.539  -0.219  0.827   -7.709   6.161
age                       -2.743    3.556  -0.771  0.441   -9.713   4.227
gender                    -1.425    2.774  -0.514  0.608   -6.862   4.013
driving_experiences        1.626    5.360   0.303  0.762   -8.879  12.131
familiarity               -3.917    4.249  -0.922  0.357  -12.246   4.412
Group Var                  6.636    2.968
Group x System[T.B] Cov    6.818    3.013
System[T.B] Var            7.005    3.421
=========================================================================
```

Figure D.12: LMM Results for SA3 Indicator with Internal Factors: Speed Compliance