

UC Office of the President

ITS reports

Title

Automated Vehicle Safety: Developing a Human-Autonomy Team Model — Findings from Driving Simulator Study

Permalink

<https://escholarship.org/uc/item/87x5w38t>

Authors

Correa-Jullian, Camila, PhD

D'Agostino, Mollie C., MS

Mosleh, Ali, PhD

et al.

Publication Date

2026-10-01

DOI

10.7922/G2VH5M7Q

Automated Vehicle Safety: Developing a Human-Autonomy Team Model – Findings from Driving Simulator Study

Camila Correa-Jullian, Ph.D., Postdoctoral Scholar, The B. John Garrick Institute for the Risk Sciences, University of California, Los Angeles

Mollie Cohen D'Agostino, Executive Director, Mobility Science, Automation, and Inclusion Center, University of California, Davis

Ali Mosleh, Ph.D., Director, The B. John Garrick Institute for the Risk Sciences, University of California, Los Angeles

Jiaqi Ma, Ph.D., Associate Professor, Mobility Lab, University of California, Los Angeles

October 2026



California
Resilient and Innovative
Mobility Initiative

Technical Report Documentation Page

1. Report No. UC-ITS-2024-27-5L	2. Government Accession No. N/A	3. Recipient's Catalog No. N/A	
4. Title and Subtitle Automated Vehicle Safety: Developing a Human-Autonomy Team Model – Findings from Driving Simulator Study		5. Report Date October 2026	
		6. Performing Organization Code ITS-Davis, UCLA ITS	
7. Author(s) Camila Correa-Jullian, Ph.D., https://orcid.org/0000-0002-4622-4064 Mollie Cohen D'Agostino, M.S., https://orcid.org/0000-0002-3689-9471 Ali Mosleh, Ph.D., https://orcid.org/0000-0002-4628-1050 Jiaqi Ma, Ph.D., https://orcid.org/0000-0002-8184-5157		8. Performing Organization Report No. N/A	
		10. Work Unit No. N/A	
9. Performing Organization Name and Address Institute of Transportation Studies, Davis 1605 Tilia Street Davis, CA 95616 The B. John Garrick Institute for the Risk Sciences, UCLA 404 Westwood Plaza Los Angeles, CA		11. Contract or Grant No. UC-ITS-2024-27-5L	
		13. Type of Report and Period Covered Final Report (October 2023 – October 2025)	
12. Sponsoring Agency Name and Address The University of California Institute of Transportation Studies www.ucits.org		14. Sponsoring Agency Code UC ITS	
15. Supplementary Notes DOI: 10.7922/G2VH5M7Q			
16. Abstract To improve the safety of automated vehicles, this study developed a mathematical model to indicate the degree to which certain traffic conditions, driver characteristics, and behaviors of an automated driving system (ADS) affect the probability of a driver safely taking over control in a level 3 ADS to avoid a potential crash. (A level 3 ADS requires the driver to intervene when the ADS cannot safely handle a situation.) Experiments in which 50 participants used a driving simulator informed the model development. The presence of a warning signal from the ADS telling the driver to intervene significantly reduced collision probability, improved driver reaction time, and increased the safety margin between the study vehicle and the vehicle in front of it. Traffic complexity—the number of nearby vehicles—increased the probability of hard-braking, tailgating, and steering variability, and it increased the drivers' perceived workload and distrust of automation. The absence of a driver warning from the ADS interacted with higher traffic complexity, so that their combination greatly increased collision probability. These findings indicate that driver warnings provide measurable safety benefits and will be used to refine the model and examine the effects of the design and timing of warnings.			
17. Key Words Automated vehicle control, Level 3 driving automation, Autonomous vehicle handover, Driving simulators, Crash avoidance systems		18. Distribution Statement No restrictions.	
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 107	22. Price N/A

Form Dot F 1700.7 (8-72)

Reproduction of completed page authorized

About the UC Institute of Transportation Studies

The University of California Institute of Transportation Studies (UC ITS) is a network of faculty, research and administrative staff, and students dedicated to advancing the state of the art in transportation engineering, planning, and policy for the people of California. Established by the Legislature in 1947, ITS has branches at UC Berkeley, UC Davis, UC Irvine, and UCLA.

The California Resilient and Innovative Mobility Initiative

The California Resilient and Innovative Mobility Initiative (RIMI) serves as a living laboratory – bringing together university experts from across the four UC ITS campuses, policymakers, public agencies, industry stakeholders, and community leaders – to inform the state transportation system’s immediate COVID-19 response and recovery needs, while establishing a long-term vision and pathway for directing innovative mobility to develop sustainable and resilient transportation in California. RIMI is organized around three core research pillars: Carbon Neutral Transportation, Emerging Transportation Technology, and Public Transit and Shared Mobility. Equity and high-road jobs serve as cross-cutting themes that are integrated across the three pillars.

Acknowledgments

This study was made possible with funding received by the University of California Institute of Transportation Studies from the State of California through a one-time General Fund allocation in the 2021 State Budget Act for the Resilient and Innovative Mobility Initiative. The authors would like to thank the State of California for its support of university-based research, and especially for the funding received for this project.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the State of California in the interest of information exchange. The State of California assumes no liability for the contents or use thereof. Nor does the content necessarily reflect the official views or policies of the State of California. This report does not constitute a standard, specification, or regulation.

Automated Vehicle Safety: Developing a Human-Autonomy Team Model – Findings from Driving Simulator Study

Camila Correa-Jullian, Ph.D., Postdoctoral Scholar, The B. John Garrick Institute for the Risk Sciences, University of California, Los Angeles

Mollie Cohen D'Agostino, Executive Director, Mobility Science, Automation, and Inclusion Center, University of California, Davis
Ali Mosleh, Ph.D., Director, The B. John Garrick Institute for the Risk Sciences, University of California, Los Angeles

Jiaqi Ma, Ph.D., Associate Professor, Mobility Lab, University of California, Los Angeles

October 2026

Table

of

Contents

Table of Contents

Acronyms	xi
Executive Summary	1
Introduction	3
Developing a Human-System Interaction Model: Driving Simulation Study.....	4
Report Organization	5
Human-System Interaction Model.....	6
Safety Driver	7
Selected Performance Shaping Factors	9
Experimental Study Design	17
Proposed Driver-ADS Team Model	17
Experiment Setup and Protocol	20
Data Collection	23
Data Processing and Analysis.....	25
Experimental Results	28
Driver Profile: Initial Survey.....	29
Scenario Perception: Test Survey	37
Simulation Data: Continuous Variables	42
Simulation Data: Categorical Variables	48
Building Driver-ADS PSF Model.....	52
Concluding Remarks	55
Initial Findings	55
Next Steps	56
Appendix	57
Experimental Design.....	57
Experimental Results	74
References	86

List of Tables

Table 1: Proposed PSF Classification for Use in HATs Involved in ADS Operations.....	10
Table 2: Equivalent Agent-Based PSFs for Safety Drivers and ADS-Equipped Vehicles.....	12
Table 3: Equivalent Scenario-Based PSFs for Safety Drivers and ADS-Equipped Vehicles.....	13
Table 4: Equivalent Team-Based PSFs for Safety Drivers and ADS-Equipped Vehicles.....	15
Table 5: Failure Modes by Information, Design, and Action Stage for Experimental Study.....	18
Table 6: Treatment Description by Factor (A, B) and Level (0,1).....	21
Table 7: Objective Measures Collected from CARLA Simulator.....	24
Table 8: Total collected dataset.	28
Table 9: Main and Interaction Effect Analysis for Categorical Variables.	51
Table 10: Subjective Measures and Factors Collected from the Initial Questionnaire.	63
Table 11: Subjective Measures and Factors Collected from Test Questionnaire.....	67
Table 12: Subjective Measures and Factors Collected from Test Questionnaire – Open Questions.	68
Table 13: Driver Profile Response Mapping.....	70
Table 14: Raw Data Retrieved from the Simulation.....	72
Table 15: Scenario Perception Survey Response Mapping.....	74
Table 16: Summary of Test Survey Results; Main and Interaction Effects Analysis.	74
Table 17: Summary of Simulation Variables, Main and Interaction Effects.	76
Table 18: Summary of Categorical Variables per Treatment.....	80
Table 19: Proposed PSFs for Driver-ADS Team Model.....	81

List of Figures

Figure 1: Functional Breakdown of On-Board Driver and ADS-Equipped Vehicle.....	8
Figure 2: PSF Model for Team Coordination. Expanded Upon from Azarkhil and Mosleh (2014).	14
Figure 3: Proposed Driver-ADS BBN Model for Experiment Design.	19
Figure 4: The Driving Simulator Setup	23
Figure 5: Histogram of (a) participant age and (b) years since driver's license was obtained.	30
Figure 6: Participant Pool: Primary Mode of Transportation.	30
Figure 7: Participant Pool (a) Average vehicle use and (b) Estimated miles driven (last 12 months).....	31
Figure 8: Reported ADAS feature familiarity (a) number of known features and (b) percentage of mentions.....	32
Figure 9: Initial Survey - Bias and Responsibility-Related Questions.....	33
Figure 10: Initial Survey - Risk Tolerance-Related Questions.	34
Figure 11: Initial Survey - Trust in Driving Automation.....	36
Figure 12: Summary of Test Survey Results.....	37
Figure 13: Summary of Workload Ratings.....	38
Figure 14: Average Scores for Scenario Perception by Treatment.	40
Figure 15: Average Scores for Workload by Treatment.....	41
Figure 16: Driving Simulator Experiment Sample Data.	43
Figure 17: Interaction Plot: Time to First Warning (left) and Total Time Warnings Were Active (right).	45
Figure 18: Interaction Plot: Time to First Takeover (left) and Total Time Drivers Were in Control of the Vehicle (right).	45
Figure 19: Interaction Plot: Mean Speed During the Scenario (left) and When Drivers Were in Control of the Vehicle (right).	46
Figure 20: Interaction Plot: Driver Throttle Inputs – Mean (left) and Standard Deviation (right).	47
Figure 21: Interaction Plot: Driver Brake Inputs – Mean (left) and Standard Deviation (right).	47
Figure 22: Interaction Plot: Mean TTC During the Scenario (left) and When Drivers Were in Control of the Vehicle (right).	48
Figure 23: Interaction Plot: Driver Reaction Time to Warnings (left) and Tailgating Events (right).	48
Figure 24: Predicted Probability of Collision per Treatment.....	49
Figure 25: Predicted Number of Hard-Braking Events per Treatment.	50
Figure 26: Predicted Number of Driver Takeovers per Treatment.	50

Figure 27: Predicted Number of Tailgating Events per Treatment.	51
Figure 28: Overview of Data Analysis and Model Construction Pipeline.	53
Figure 29: Experimental Session Design.....	57
Figure 30: Experimental Setup - View from the Driver’s Seat.....	58
Figure 31: Example of Driver Interface, (a) Warning Alert and (b) Collision Alert.	59
Figure 32: Driving Scenario (a) B0: Simple (low traffic) and (b) B1: Complex (high traffic, additional vehicles represented in red).	61

Executive Summary

Acronyms

ADAS	advanced driver assistance systems
ADS	automated driving system
BBN	Bayesian belief network
DDT	dynamic driving task
HAT	human-autonomy team
HRA	human reliability analysis
HSI	human-system interface
ODD	operational design domain
OEDR	object and event detection and response
PIF	performance influencing factor
PSF	performance shaping factor
SAE	Society of Automotive Engineers
TOR	Takeover request
TTC	time-to-collision

Executive Summary

As research and development on automated driving systems (ADSs) continues to advance, the full scope of safety challenges associated with automated vehicles (AVs) remains uncertain, particularly regarding human-system interaction. As more ADS-equipped systems move towards deployment on public roads, understanding their operational safety remains an urgent priority.

This report presents the development and initial findings supporting the empirical validation of a driver-system model through controlled driving simulator experiments involving drivers performing takeover scenarios in Level 3 ADS. The foundations of this model lie in team performance models developed within the context of Human Reliability Analysis (HRA) in complex socio-technical systems. HRA provides key input into quantitative and qualitative risk assessments, and, by extending these concepts to human-system interactions in ADS operations, it offers a path to extract and quantify the impacts that ADS system and operation design have on operational safety.

Current literature and ongoing research exploring human-autonomy teams (HATs) in ADS operations are used to establish and refine relationships within human-system team models. From the perspective of HATs, a causal model is constructed using Bayesian networks to represent the relationship between identified performance shaping factors that contribute to team performance (Correa-Jullian et al. 2024a; Correa-Jullian et al. 2025, Correa-Jullian et al. 2026b). Driving simulation-based experiments inform and complement the model's development and quantification, with a focus on operation and takeover performance. This report provides an overview of the experimental design and analysis of the data collected from the driving simulator-based experiments. Given that time-constrained control transitions between drivers and the ADS have been identified as a critical source of safety hazards, the experiments are designed to assess the impact of silent failures – when a system failure that should have triggered a driver-facing alert has failed to do so – (Factor A: Takeover Request / Warning Availability) and environmental complexity (Factor B: Traffic Level) on operation and takeover performance—represented through a causal Bayesian belief network (BBN) structure.

The experiments were conducted in a medium-fidelity physical simulator and leveraged the OpenCDA open-source platform to design the driving scenarios (Xu et al. 2023). A total of 72 volunteers participated in the research study conducted from February to May 2025. Questionnaires were used to extract the participants' driver profile, trust in driving automation, and workload perception of the simulated driving scenarios. These subjective ratings were combined with simulation data, including vehicle kinematics, reaction times, and recorded incidents, to derive the impact of agent, scenario, and team factors on safety performance. Appropriate statistical tests were applied to continuous, count, and categorical (ordinal and nominal) data to determine the significance of main and interaction effects, i.e., the impact of warning availability and traffic levels. Future work will focus on the final causal model developed and calibration to link the participant's profile and scenario complexities (e.g., traffic, warnings) to safety outcomes, i.e., probability of collision or near-miss hard-braking events.

Contents

Introduction

Automated driving systems (ADSs) and their associated technologies possess significant implications for transportation, commercial sectors, and legal frameworks. Their successful adoption depends on rigorous identification and mitigation of associated risks, particularly those arising from human-system interactions.

The integration of ADS into heavy-duty commercial transport operations is driven by the objectives of reducing traffic collisions and incidents attributable to human errors, while enhancing efficiency and extending operational hours (Bhoopalam et al. 2023). As automated driving technology is increasingly used on public roads, understanding the relationship between drivers and ADS-equipped vehicles is a first step towards establishing robust evidence-based policies in commercial applications.

To date, the Society of Automotive Engineers (SAE) categorizes automation on a scale from 0 to 5, based on the division of the dynamic driving task (DDT) between human and machine (SAE On-Road Automated Vehicle Standards Committee 2021). At Levels 0 to 2, the human remains responsible for performing the DDT, with automated features supporting when certain conditions are met. At Level 3, the ADS controls the DDT within a prescribed operational design domain (ODD) but may require human fallback. Level 4 systems can operate without in-vehicle fallback but remain ODD-limited, whereas Level 5 automation performs all driving tasks in any environment. Currently, there are multiple prospective models of heavy-duty autonomous vehicles operations spanning Levels 2-4 depending on the application, underscoring the need to examine how human operators supervise, intervene, or collaborate with automation.

Automation poses many challenges regarding human-system interaction, such as complacency, human-out-of-loop, sudden increases in cognitive workload, distraction, and skill degradation. However, human supervision also serves as the ultimate safety barrier to prevent or mitigate risks during operation. Therefore, systems must be designed to ensure safe human intervention when needed. This analysis incorporates methods from human reliability analysis (HRA) and human factor research to characterize and quantify risks arising from human-ADS interactions.

This work examines operational safety across two dimensions. A previous report (Correa-Jullian, et al. 2026) established a structured hazard identification methodology to derive safety metrics emphasizing the role of human-system interactions in heavy-duty autonomous vehicle operations. Building on these findings, the present report develops a theoretical interaction model informed by driving simulator experiments to partially validate causal relationships.

Developing a Human-System Interaction Model: Driving Simulation Study

To advance the research on human-system interactions in the context of ADS operations, this research adapts team performance frameworks from HRA to the transportation domain. The HRA discipline has long supported quantitative and qualitative risk assessments of complex socio-technical systems. Extending these concepts to ADS operations provides a path to quantify operational safety impacts of system and mission design.

Recent research emphasizes collaborative dynamics between humans and automation, including task allocation strategies, trust calibration, attention management, and explainability challenges (Janssen et al. 2019; Lyons et al. 2021). The concept of human-autonomy teams (HATs) offers a framework for analyzing these interactions from a collaborative perspective. These HATs are defined as teams in which at least one member meets the definition of an *autonomous machine agent* – expressing degrees of agency, interdependence, and proactivity – working alongside human team members to achieve a collective goal (O'Neill et al. 2023; Mosier et al. 2017; O'Neill et al. 2022; McNeese et al. 2018). This hybrid team concept makes it possible to apply decades of research on the impact of human team dynamics to system safety and explore emerging behaviors.

With rising automation across industries, HRA provides structured approaches to modeling human performance under evolving operational demands (Porthin, Liinasuo, and Kling 2020). Many HRA methods rely on performance shaping factors or performance influencing factors (PSF/PIF) to express factors impacting human performance and estimate human error probabilities (Ekanem, Mosleh, and Shen 2016; Boring and Blackman 2007; Sundaramurthi and Smidts 2013). PSFs are aspects of individual characteristics, the environment, scenario development, or the organization in charge that increase or decrease the likelihood of human error (Groth and Mosleh 2012). The team-centered information, decision, and action in crew (Tc-IDAC) model (Azarkhil, Mosleh, and Ramos 2025b; 2025a) further established connections between individual-level and team-level tasks, error modes, and factors. This work focused on expanding individual, team, and communication related factors, and proposed causal models for the five team-related factors (Azarkhil and Mosleh 2014). Multiple methods have been used to estimate human error probabilities, including the use of Bayesian belief networks to represent the causal relationships between human factors, tasks, and system safety (Groth, Smith, and Moradi 2019). However, a persistent limitation is the lack of benchmark datasets to support parametrization. As a result, many models rely on expert judgement to estimate values of nominal human error probabilities and performance shaping factors (Burns and Bonaceto 2020). In ADS contexts, existing experimental studies typically measure physiological and behavioral responses under various interface designs and operational scenarios. These studies provide an opportunity to align existing HRA theory with empirical evidence.

Accordingly, this report reviews existing PSF taxonomies and adapts them for ADS and safety-driver contexts. We then selected PSFs to construct a driver–system interaction model focused on takeover performance based on the findings from the safety metric derivation methods study (Correa-Jullian et al. 2026a), combined with

relevant literature on human factors and human reliability analysis. We conducted driving simulations with human subjects to refine and parameterize this model, where risk scenarios involving human or interaction failures were implemented in a virtual driving simulator environment. Given that time-constrained control transitions between ADS and safety drivers are a critical hazard source, experiments focused on how workload and silent failures (e.g., a warning fails to appear) affect driver takeover performance.

Report Organization

This report is structured as follows: (1) review of human–system interaction literature, (2) selection of PSFs relevant to ADS safety drivers, (3) experimental design of the driving simulator study, and (4) overview of key results.

Human-System Interaction Model

Heavy-duty autonomous vehicles are expected to operate at SAE Level 4 with no human intervention. However, a variety of operational concepts have emerged in this space, such that different operator roles and levels of human intervention may be required at certain operational periods (Correa-Jullian et al. 2026a). Many elements from human-system interactions studies can provide insights to guide the system and operation design of these systems, so that when human intervention is required (in the event of an emergency), it is timely and contributes to prevent or mitigate risks.

The SAE levels of driving automation define “Driver Support Systems” (Levels 0-2) and “Automated Driving Systems” (Level 3-5). Commercially available driver support systems – commonly referred to as advanced driver assistance systems (ADASs) – have expanded over the past decade, offering assistance at low speeds and on highways under specified conditions (Aukema et al. 2023; Yu et al. 2021). Up to SAE Level 2, the driver needs to continuously monitor the vehicle’s actions, respond to vehicle alerts or warnings, and intervene to correct inadequate responses as the automated driving technology does not perform all object and event detection and response (OEDR) functions. Driver warning design depends on the vehicles’ manufacturer, and can span different modes, including visual (e.g., dashboards alerts), audible (alarms), and/or tactile (e.g., steering wheel, seat belt or seat vibrations). These warnings are activated even when there is not a risk of a crash—i.e., beyond “crash-imminent scenarios” (Mueller et al. 2020; Yue et al. 2021; Zhu, Wang, and Hu 2020). At Level 3, the ADS performs all OEDR functions under specific operational conditions, i.e., it can perform all DDTs within its ODD; but the human remains as the fallback-ready user, intervening upon the request of the ADS or in response to unsafe driving conditions (Boggs, Arvin, and Khattak 2020; Favarò, Eurich, and Nader 2018). Shared-control, task-allocation, and control-transition designs depend on vehicle specifications and operational envelopes (Vanderhaegen 2021; Warg, Skoglund, and Sassman 2021; Marcano et al. 2020).

A core challenge at Level 3 is resolving conflicting assessments between system and driver during automation failures or when the driver’s situational awareness is limited (Vanderhaegen 2021). Takeover performance in shared autonomy is influenced by time pressure, stress, sleep deprivation, and the longer-term effects of automation on attention and cognitive load during routine and emergency disengagements (Bruck, Haycock, and Emadi 2021; Rendon-Velez et al. 2016; Gouribhatla and Pulugurtha 2022; Ramakrishnan, Balasingam, and Biondi 2021; Paschalidis, Choudhury, and Hess 2018; Wu et al. 2019). Human-system interface design, non-driving-related tasks, and alarm characteristics also shape situational awareness and takeover quality (Samani, Mishra, and Dey 2022; DeGuzman et al. 2021).

At Level 4, the ADS is expected to execute all DDTs independently, including DDT fallbacks and achieving a minimal risk condition, without requiring human driver intervention, provided the system operates within its predefined operational design domain (ODD). Hence, the user is not expected to monitor or intervene in the vehicle’s actions (SAE On-Road Automated Vehicle Standards Committee 2021). Yet, remote human assistance

is expected to remain necessary for supporting vehicle and passenger operations in the near future (Kettwich et al. 2021) or may remain indefinitely as part of their concept of operations, especially in middle-mile operations (automation is engaged between first-mile and last-mile delivery, the latter being operated in manual driving mode by a safety driver or operator). Many planned heavy-duty autonomous vehicle concepts involve interactions among SAE Levels 2–4 systems and remote operators, safety drivers, or safety operators who support tasks outside the ODD or serve as fallback-ready users in emergencies (Bhoopalam et al. 2023). Thus, evaluating safety drivers and the feasibility of remote interventions—determining which tasks are appropriate and when—requires comprehensive modeling of factors that influence system-level safety and performance. This report centers on onboard safety drivers and their interactions with ADS-equipped heavy-duty vehicles. Details on remote operators can be found in Correa-Jullian et al. 2024b; Ma et al. 2024; and Correa-Jullian et al. 2025.

Safety Driver

This case considers a trained and certified safety driver responsible for supervision, responding to takeover requests, initiating disengagements, and interacting with the ADS, while remote fleet operators provide limited service-related support. Figure 1 depicts the functions and communications between the safety driver and the ADS-equipped vehicle.

When the ADS is engaged, the onboard driver serves as a fallback-ready user, independent of whether the ADS can nominally perform the DDT fallback. When the ADS is not engaged, the driver performs the DDT and may request ADS engagement when appropriate. The driver can initiate control transitions (e.g., manual input) or accept takeover requests (e.g., steering or braking interventions). Although expected to monitor the environment continuously, the driver may fail to maintain situational awareness (e.g., because reduced attention or engagement in non-driving-related tasks).

To sustain engagement, the ADS employs noninvasive driver-monitoring tools. Alert messages include those referring to safety (e.g., multi-modal audio, visual, and haptic alerts), driver monitoring (e.g., front-facing gaze or hand-on steering wheel), vehicle health (e.g., a vehicle failure has been detected), driver-initiated control transfer status (e.g., success or failure), and system-initiated control transfer (i.e., take-over requests). The ADS may transmit regular vehicle warnings or suggested actions based on internal health metrics (e.g., low battery or tire pressure). Additionally, some information is communicated passively, such as general health metrics, as well as infotainment and navigation applications. Even with certification, the safety driver may have limited understanding of heavy-duty autonomous vehicle functions and may be unable to diagnose system misbehavior. Identifying factors that affect performance is therefore a first step toward systematic frameworks for assessing system-level safety impacts.

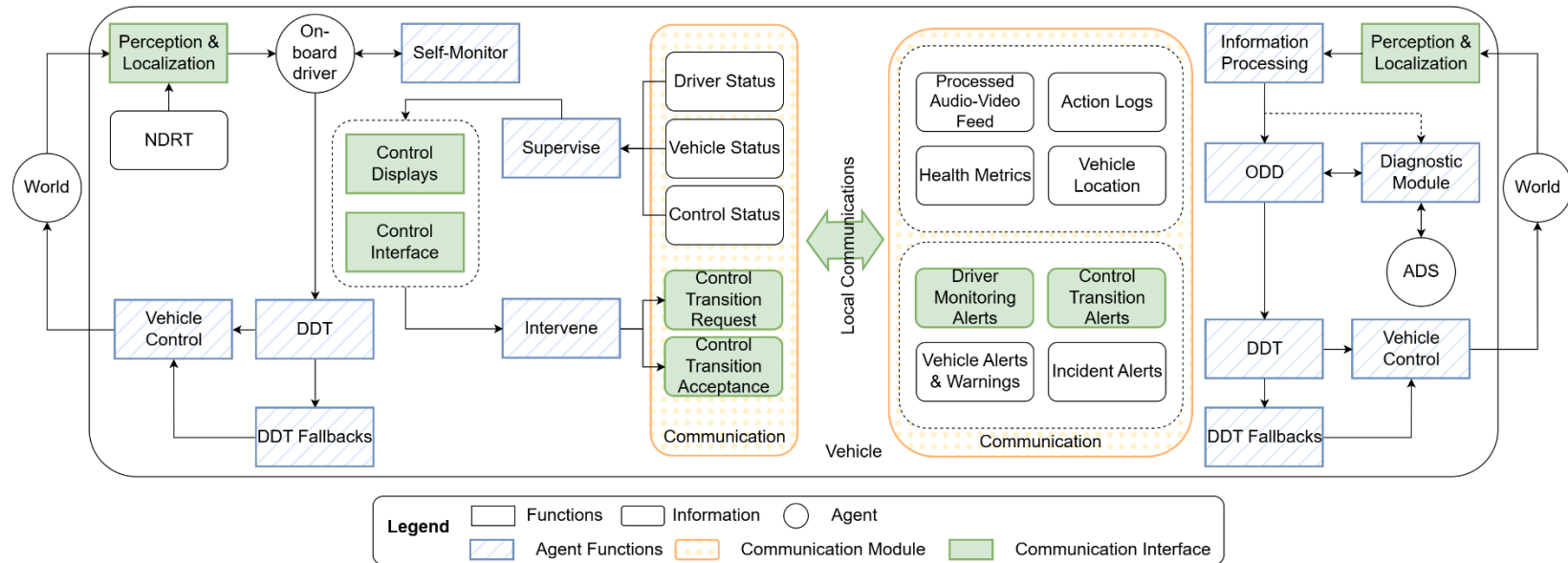


Figure 1: Functional Breakdown of On-Board Driver and ADS-Equipped Vehicle.

Selected Performance Shaping Factors

Developing a human-autonomy team (HAT) performance model requires analyzing, adapting, and extending existing PSF nomenclatures to represent both human and automated machine agents. While the underlying decision-making mechanisms of an autonomous machine agent fundamentally differ from a human, the effect of both internal and external factors, e.g., scenario and design elements, can be traced to observable ADS behaviors.

Building on the hierarchy developed in Groth and Mosleh (Groth and Mosleh 2012), this section applies PSF concepts to HATs at comparable levels of detail across agents, considering organization, agent, scenario, and team factors, and identifying data sources for design-time and operation-time characterization. Team-behavior insights from team-centered information, decision, and action in crew (Tc-IDAC) complement the taxonomy (Groth and Mosleh 2012; Ekanem, Mosleh, and Shen 2016), adding detail on how individual-level factors shape team-level cohesion, role awareness, and leadership. Effective error detection, indication, and correction require shared mental models, mutual monitoring, backup behavior, and adaptability (Mosier et al. 2017), offering greater transparency into system-level decisions when interpreted through team interactions (Azarkhil, Mosleh, and Ramos 2025b; 2025a). A summary of the proposed taxonomy is presented in Table 1, with additions highlighted in bold.

This section contains extracts from Correa-Jullian et al. (Correa-Jullian et al. 2024a; Correa-Jullian et al. 2025; Correa-Jullian et al. 2026b) as well as additions focused on safety drivers.

Organization-Based and System-Based Factors

Organization-based and system-based factors (previously machine-based factors) are those under the control of the organization operating or designing the system. In the context of ADS, the boundary between system- and organization-based factors becomes less distinct due to the involvement of multiple organizations, such as vehicle manufacturers (OEMs [original equipment manufacturers]), technology suppliers, ADAS-ADS developers, and fleet operators. The main applicable factors are:

- **Safety culture:** Each organization's safety culture is reflected in risk-management policies and responsibility allocation across system and operations design. The OEM-ADS developer relationship and over-the-air updates further complicate distinctions between organization- and machine-based factors, embedding safety culture in design, testing, validation, maintenance, updates, and incident reporting. Safety culture can be used as a proxy for elements likely to be codified by standards or law (e.g., corrective-action programs, task-allocation requirements).
- **Training programs:** As in other complex systems, ADS fleets will rely on the availability and quality of decision-support procedures. Safety driver training and certification may be essential depending on task allocation and scheduling (e.g., ODD-breach detection, DDT fallbacks, post-incident management,

communications with remote operators/first responders) (Automated Vehicle Safety Consortium 2024).

- Corrective action programs: Although many ADS design attributes are fixed (robustness, verification and validation, testing), operational software/hardware elements may be updated periodically. The availability and quality of corrective-action programs determine how operational experience informs system improvements, potentially altering task allocation, procedures, and tools for both humans and automation (Zhao et al. 2020).

Table 1: Proposed PSF Classification for Use in HATs Involved in ADS Operations.

Organization/ System-based	Team-based	Agent-based	Scenario-based
Training program Corrective action program Safety culture Management activities • Task Allocation • Task Scheduling Workplace adequacy Resources Procedures Tools Necessary information	Team Cohesion • Shared Goals • Mutual Trust Role Awareness Direct Supervision • Leadership Team Communication • Human-system interface • Control Input/Output • Control transitions • Non-safety related Team Coordination • Responsiveness • Engagement	Attention/Resource Allocation Physical & psychological abilities • Alertness/Resource Availability • Fatigue • Impairment • Physical attributes Knowledge/Experience/System Maturity Skills Bias Familiarity with situation Risk Tolerance	External environment Conditioning events Task load Time load Other loads • Non-task • Passive information Task complexity Stress Perceived situation • Severity • Urgency Perceived decision responsibility

Agent-Based Factors

Agent- or individual-level factors include attention, experience, and risk tolerance. Historically these human-focused and often unobservable factors (requiring behavioral markers) can be extended to autonomous agents to capture how ADS software and architecture shape problem-solving. We have introduced four new factors to express ADS behaviors to act as parallels to human-based factors, as depicted in Table 2. These are:

- **Resource Allocation:** Human attention maps to the ADS's resource allocation strategies, as shaped by its electrical/electronic architecture design. This is usually developed at early design stages, focusing on complying with safety and non-safety requirements, such as timing, redundancy, and freedom from interference (Askaripoor, Mueller, and Knoll 2023).
- **Resource Availability:** Physical and psychological abilities refer to the internal resources of the agent available to perform their tasks, including the level of alertness, fatigue, impairment, or physical attributes (relevant to impaired/distracted driving and remote operators) (Ayas et al. 2023; McKerral, Pammer, and Gauld 2023). For the ADS, resource availability captures system degradation (including elements of sensor or software reliability and calibration) and computational resource depletion (energy, memory, and processing capacity) during operation (Y. Zhang et al. 2023).
- **System Maturity:** Safety drivers and remote operators are expected to fulfill training requirements and potential future certification requirements (Automated Vehicle Safety Consortium 2019; 2023). These factors are significantly shaped by design choices, verification and validation processes, and software maintenance of the ADS. Consequently, the factor system maturity is proposed to account for knowledge/experience comprising (1) data exposure (e.g., dynamic environmental conditions, high-definition map accuracy, and driving behavior model sophistication), (2) testing performance (e.g., real-world scenarios, closed tracks, simulation), and (3) the extent of the certified ODD.
- **Risk Tolerance:** This factor encapsulates the agent's concern for safety and attitude towards risk, central to rule-based decision making, traffic law compliance, and prioritizing safety based on contextual cues. For the driver, this may be a defining factor in calibrating time budgets for takeovers and other emergency situations (Tanshi and Soffker 2022). Software architecture design can directly impact risk tolerance—for example in the selection of safety margins or confidence thresholds, or the design of trigger automated fallbacks or sensor fusion schemes impacting OEDR tasks (Lodhi, Kumar, and Pandey 2023). Fleet-level consistency depends on maintenance and update policies (Correa-Jullian et al. 2024b; 2024a).

Bias, skill, and perceived familiarity apply to human and automated machine agents. Human cognitive biases (expectation, confirmation, belief, recency) affect decision-making, while the ADS is subject to bias from the training data quality (data-based bias), decision-making algorithm sophistication (algorithm-based bias), or sensor limitations under varying conditions (sensor-based bias). The skill factor represents the levels of perception (e.g., hazard detection, localization), decision-making (e.g., path and maneuvers planning, risk assessment), and control skills (vehicle control precision and responsiveness) of human and automated machine agents (Zangi et al. 2022; Di and Shi 2021). Perceived familiarity captures how recent experiences shape decisions: memory of recent events for drivers; training/testing completeness for ADS, including incorporation of operational experience via updates (Zhao et al. 2020; Warg, Skoglund, and Sassman 2021). For the driver, long-term performance over multiple exposures to takeover or intervention events is also relevant (Samani, Mishra, and Dey 2022).

Table 2: Equivalent Agent-Based PSFs for Safety Drivers and ADS-Equipped Vehicles (new factors highlighted with *).

PSF		Safety Driver	ADS
Attention	Task	Effect of driving automation	Resource Allocation*
	Environment		
Physical & Psychological Abilities	Alertness	Driver behavior and state	Resource Availability*
	Fatigue		
	Impairment		
	Physical Attributes	Relevant for slip errors	
Knowledge/ Experience		Extent of formal training	System Maturity*
Skills		<i>Perception, decision-making, control & communication</i>	
Bias		Expectation, confirmation, belief, etc.	<i>Data, algorithm, & sensor</i>
Perceived Familiarity		e.g., multiple exposure to takeover events	<i>Data exposure & software update policy</i>
Risk Tolerance*		Trust & time-budget calibration	<i>Safety margins</i>

Scenario-Based Factors

Scenario-based factors capture how scenario development influences performance. They include observable characteristics (e.g., weather, road geometry, traffic) and perception-related elements (primarily for humans). Table 3 provides a summary of applicable factors.

External environment refers to the overall complexity of driving conditions (e.g., weather, road geometry, traffic density) and interactions with other road users (e.g., vehicles, cyclists, pedestrians). Conditioning events encompass latent failures and alert-triggering events that influence scenario evolution. These include ODD breaches, vehicle malfunctions and connectivity failures (which may require DDT fallback responses), and takeover actions from the driver (Mishler and Chen 2023), either initiated by the driver or prompted by the ADS.

Task complexity reflects cognitive/physical demands, required knowledge/procedures, ambiguity, and task-switching impacts (Stapel, Mullakkal-Babu, and Happee 2019; Heikoop et al. 2019). Task load is an individual-level representation of the overall system design, accounting for the effective task allocation schemes implemented and resulting task demand assigned to each agent (Khastgir et al. 2019). For the driver, this includes takeover/handover procedures and system reliance on driver monitoring; for ADS, it includes computational demands within the ODD or in response to an ODD breach, and other contextual information increasing vehicle maneuver difficulty (e.g., road geometry, traffic density, weather conditions). Non-task-related loads denote additional tasks beyond DDTs. For the driver, this refers to the non-driving related tasks (McKerral, Pammer, and Gauld 2023); for the ADS, it is directly related to the resource allocation factor. Information load represents the information and cues presented by elements external to the agent, including their team members and the external environment. The design and calibration of safety alarms critically shape the information load affecting the driver or remote operator, especially in emergency situations (Louw et al. 2019). Time load—the ratio of available to required time—constrains latency, task, and alarm design for both agents, bounded by system and physical limits (B. Zhang et al. 2019; Tanshi and Soffker 2022; Marberger et al. 2018).

Table 3: Equivalent Scenario-Based PSFs for Safety Drivers and ADS-Equipped Vehicles (new factors highlighted with *).

PSF	Safety Driver	ADS
Conditioning Events	Fallback triggers, e.g., ODD breaches, vehicle failures, driver alerts	
External Environment	Driving conditions and traffic interactions	
Information Load	Driver assistance warnings and requests	Resource Availability*
Non-task Load	<u>Non-driving related tasks</u>	Resource Allocation*
Task Load	Dynamic Driving Tasks and Control Transitions	
Task Complexity	Scenario complexity (e.g., within/outside ODD)	
Time Load	Physical and design constraints	
Stress	Mixed results in driver simulator/on-road testing	n/a
Perceived Severity & Urgency		Risk Tolerance*
Perceived Decision Responsibility	Interactions with other road users	

Perception-related factors (stress; perceived urgency, severity, responsibility) are primarily human-focused: stress is influenced by attitudes toward driving and automation (Heikoop et al. 2019); both severity and urgency play an important role for the driver (Becker, Brandenburg, and Thüring 2022); and perceived decision responsibility impacts interactions with other road users, reflecting the remote operator's or driver's perceived responsibility to intervene in a scenario if the ADS or the driver or operator fail to prevent or mitigate a hazardous scenario (Azevedo-Sa et al. 2021). For ADS, these constructs are better represented by risk tolerance and resource availability (e.g., algorithmic capacity limits that can induce latency or unsafe actions), providing proxies for “perceived” urgency and information and time constraints (Marcano et al. 2020).

Team-Based Factors

A team perspective increases problem-solving capacity and provides layered safety, but introduces coordination challenges hinging on cohesion, role awareness, communication, and authority gradients (Groth and Mosleh 2012; Ham, Jung, and Park 2021). In turn, these elements are shaped by the technical and non-technical skills of each member, as well as mutual trust and understanding (Xing et al. 2021). Team-centered information, decision, and action in crew (Tc-IDAC) expands individual factors and causal pathways to team-level outcomes as shown in Figure 2 (cohesion, role awareness, and leadership).

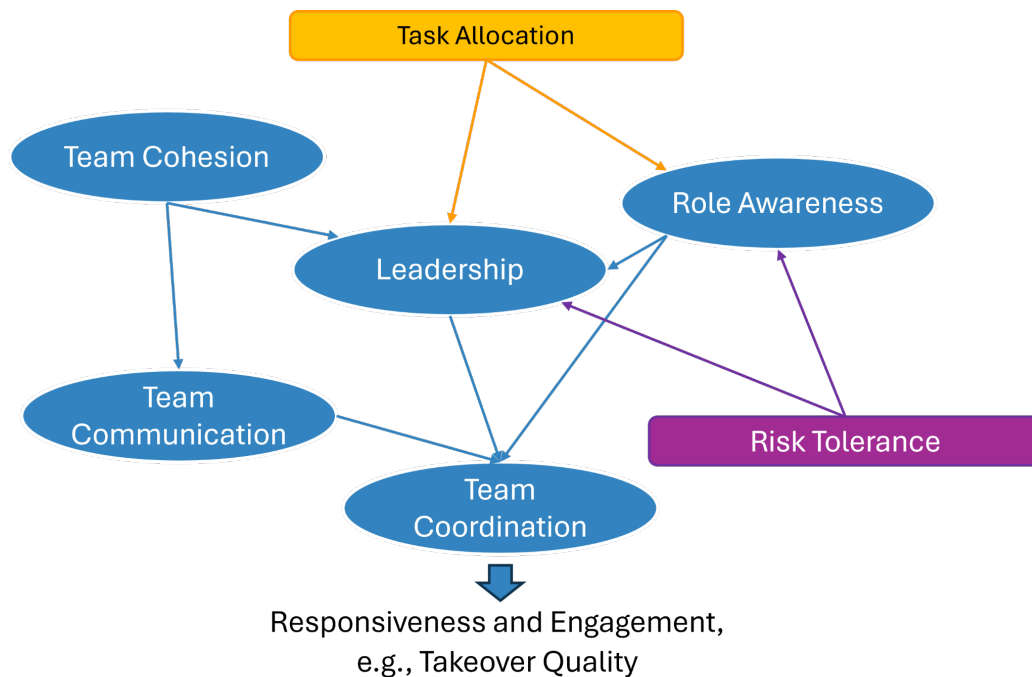


Figure 2: PSF Model for Team Coordination. Expanded Upon from Azarkhil and Mosleh (2014).

The applicability of these factors is summarized in Table 4.

- **Team cohesion:** Includes interactions among members (task/social cohesion), diversity, mutual trust, shared goals, authority gradients, and procedural compliance (Manchon, Bueno, and Navarro 2022). For the driver, this relates to personal experience, the belief in shared goals (i.e., safety) and trust towards the autonomous agent (Pan et al. 2023; Payre et al. 2022). For the ADS, cohesion-analogous attributes are set by design and training (e.g., risk tolerance). Driver monitoring system relevance to system-initiated takeovers suggests the need for calibrated detection and communication as driver experience grows (Wang, Zhang, and Zhou 2022).

Table 4: Equivalent Team-Based PSFs for Safety Drivers and ADS-Equipped Vehicles (new factors highlighted with *).

PSF	Safety Driver	ADS
Team Cohesion	Belief in shared goals, compliance, and mutual trust	Risk Tolerance
Role Awareness	Knowledge of roles and responsibilities	Task Allocation
Direct Supervision / Leadership	Authority gradients, supervision, adherence to procedures	Task Allocation*/ Risk Tolerance*
Team Communication	Human-system interface Design & Availability (Quality & Effectiveness)	
Team Coordination	Responsiveness and Engagement, e.g., Takeover Quality	

- **Role awareness:** Refers to how each team member perceives their role within the team, including the knowledge and perception of the team's goals, roles, and norms, their training and experience as a unit, and individual concern for safety and quality. In automated driving, misunderstanding ADS limitations or driver responsibilities can impair emergency performance (Khastgir et al. 2019). For the ADS, task-related cohesion elements may be sufficiently encoded via design-time task allocation (organization-based).
- **Leadership:** Driver-ADS relationships are perceived as a vertical hierarchy, where the driver should have complete veto power over the ADS functions (Lyons et al. 2021). However, shared control paradigms and safety mechanisms against driver impairment raise questions on whether the ADS can implement control transition actions without the driver's consent (Correa-Jullian et al. 2024b). Design-time task allocation (organization-based) and risk tolerance (agent-based) may best describe the ADS decision-

making hierarchy. Further discussion on authority and error management dynamics within HATs is required (Vanderhaegen 2021).

- Team communication: The capacity to exchange information via verbal and device-based channels, influenced by device familiarity, accessibility, and usability. For driver-ADS HATs, human-system interface design and availability warrant explicit consideration as team-communication factors, given the criticality of interface familiarity, warnings, alarm triggers, and control-transition procedures in emergencies (Tinga et al. 2023). For ADS, modality and content are design-determined (system-based factors); hence, we only consider aspects related to communication quality and availability (e.g., timely warnings) in this category.
- Team coordination: As shown in Figure 2, the preceding factors strongly influence coordination and the probability of meeting safety and mission goals (Ham, Jung, and Park 2021). Coordination captures task interdependence between ADS and driver—e.g., engagement rate, responsiveness, and quality of shared tasks such as control transitions (Xing et al. 2021)—and integrates planning, scheduling, and action implementation (Azarkhil, Mosleh, and Ramos 2025a; 2025b).

Experimental Study Design

We used human subject experiments to study the impact of environment complexity and silent failures on control transitions, i.e., driver takeovers, in time-constrained collision avoidance scenarios. These experiments focused on how participants, acting as drivers, transition between manual and automated driving modes when experiencing automation failure, in this case, represented by an onset of tailgating behavior by the ADS. The simulated driving scenarios involve events requiring participants' intervention, such as silent failures or explicit takeover requests, allowing researchers to assess the ability of the drivers to detect issues (i.e., tailgating behavior), react appropriately, and resume manual control effectively. The specific objective of the experiments is to assess the effect of scenario complexity ("External Environment") and the availability of early warnings ("Conditioning Events") over driver performance in takeover scenarios, i.e., when the driver is expected to switch from automated to manual driving.

This section provides an overview of the methods, tools, and approaches used to design, conduct, and analyze the driving simulator experiments.

Proposed Driver-ADS Team Model

Based on the foundations described in the previous section, we identified key high-level tasks and errors – at individual and team level – of the Driver-ADS HAT that are potential candidates for experimental assessment. The failure modes depicted in Table 5, linked to specific Information, Design, and Action (I-D-A) stages, are mapped to identified factors contributing to each agent's performance. Failures to detect, assess risk, and intervene correctly can be interpreted as stemming from individual tasks, i.e., monitoring, planning, and executing the DDTs (as well as vehicle control). Failures to indicate or to takeover are linked to the concept of error correction (as described by team-centered information, decision, and action in crew [Tc-IDAC]), related to the tasks of requesting support and initiating control transitions.

Table 5: Failure Modes by Information, Design, and Action Stage for Experimental Study.

Failure Modes	Type	Context	IDA Stage
Failure to detect	Subjective	Silent failures	Information
Failure to assess risk	Subjective	Explicit failures	Decision
Failure to indicate	Objective	Silent failures	Action
Failure to takeover	Objective	Silent/Explicit failures	Action
Failure to intervene correctly	Objective	Silent/Explicit failures	Decision/Action

Given the analysis of the driver's and ADS's PSFs, the following model in Figure 3 is proposed. For simplicity, this model focuses on the driver, while the interactions with the ADS are concentrated in the scenario-based factors: conditioning events (silent failure of driver warnings) and external environment (level of traffic), and their impact on team-based factors: compliance with procedures, engagement, and responsiveness.

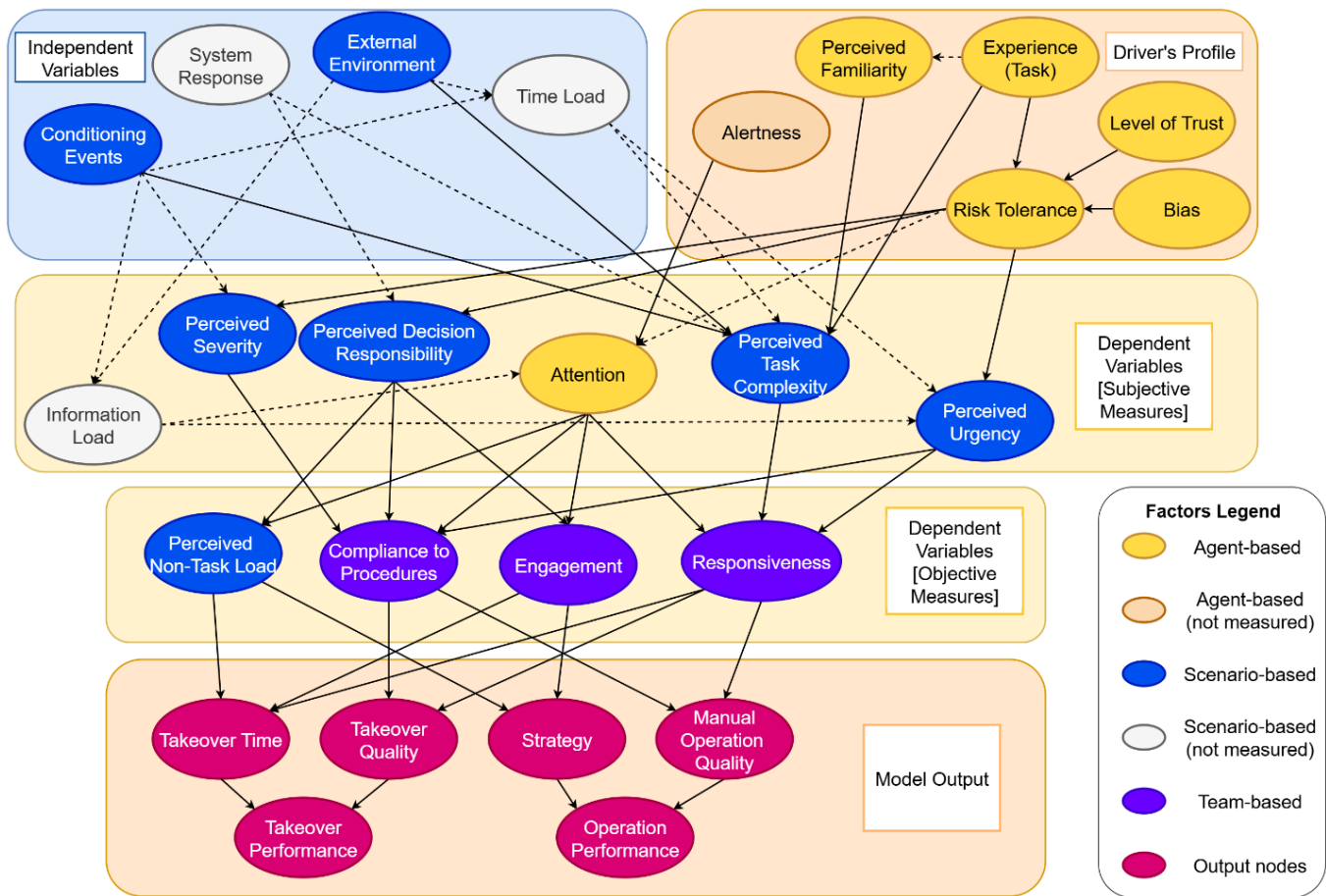


Figure 3: Proposed Driver-ADS BBN Model for Experiment Design.

The choices of which PSFs and associated metrics to include in the model are made through a combination of expert opinion and validated through a systematic process based on quantitative results (Groth and Mosleh 2012). The selected PSFs to study include:

- Agent-based (Driver): Perceived Familiarity, Experience, Bias, Level of Trust, Risk Tolerance, and Attention (assessed dynamically).
- Scenario-based: Conditioning events, External Environment, Perceived Severity, Urgency, and Decision Responsibility, Task Load/Complexity, and Non-Task Load.
- Team-based: Responsiveness, Engagement, and Compliance to Procedures as proxies for Team Communication, Team Coordination, and Team Cohesion, respectively.

In addition, output nodes are meant to represent safety outcomes related to the identified high-level failures, including failures to detect and assess risk as well as failures to takeover and intervene correctly. These are:

- **Takeover Performance:** Represented through indicators for Takeover Time and Quality. These factors are meant to depict the impact of driver reaction times to warnings or tailgating events, as well as the difference between driver/ADS responses to the simulated risks.
- **Operation Performance:** Represented through indicators for Driving Strategy and Manual Driving Quality. These factors compare the driving quality (e.g., speeding, traffic lane violation, hard-braking events) for the driver and the ADS.

Figure 3 also includes factors that were not captured by the experimental design, but whose connections (arcs) aim to provide additional context, such as System Response, Time Load, and Information Load.

Experiment Setup and Protocol

The driving simulator experiments were designed to support the refinement and quantification of the model presented in Figure 3. Participants were required to attend an in-person appointment that consisted of the following main stages: (1) an initial questionnaire to characterize each participant, (2) a training session to introduce the simulator and purpose of the study, (3) a test session consisting of four scenarios, and (4) a final questionnaire leading to the end of the study (5).

Participants played the role of a driver in an ADS-equipped vehicle and could interact with the driving simulation environment through a physical setup of medium fidelity characteristics. Each participant experienced four test scenarios, each treatment assessing a combination of two independent variables: the availability of early warning triggered by a tailgating event (A) and scenario complexity represented as the traffic level (B). As shown in Table 6, two levels are selected for the early warnings (A0: Warning, A1: No Warning), and two levels of driving conditions (B0: low traffic, B1: high traffic). A counterbalanced design was used for this study, such that the order of the scenarios for each participant was randomized. This resulted in 24 unique experiment orders, for which the design considered at least three iterations. Additional details about the experiment protocol can be found in the Appendix.

Table 6: Treatment Description by Factor (A, B) and Level (0,1).

Treatment	TOR	Complexity	Description
A	A0 - Warning	B0 - Low Traffic	Simple - Explicit TOR
B	A1 - No Warning	B0 - Low Traffic	Simple - Silent Failure
C	A0 - Warning	B1 - High Traffic	Complex - Explicit TOR
D	A1 - No Warning	B1 - High Traffic	Complex - Silent Failure

We used data collected from questionnaires and simulation data, combining subjective perception of the participants and the objective system measurements, such as vehicle kinematics, time-to-collision (TTC), and reaction times. Simulation data was processed from single-timestamp files to aggregated indicators per scenario by subject, including continuous and categorical counts and binary data. Questionnaire data describing the subject's profile and trust in automation, as well as scenario perception, mostly categorical ordinal data, were organized by subject and per treatment by subject. Collected data was analyzed using descriptive statistics and the appropriate statistical tests to determine the main and interaction effects.

Experimental Setup

Driving simulator environments have played a major role in both industry and academic research for studying aspects of driving, including vehicle functions, human factors, and the development of autonomous driving technologies. Simulators offer a controlled, relatively low-cost, and safe environment that allows researchers and engineers to explore different scenarios, test new algorithms, and analyze human responses without the risks associated with real-world driving (Khastgir et al. 2021; Bruck, Haycock, and Emadi 2021).

OpenCDA is a comprehensive open-source ecosystem that includes a collection of pre-trained models, a range of simulators for driving and traffic at different levels of detail, benchmark datasets for training and testing, and a scenario database and generator for Cooperative Driving Automation (CDA) research (Xu et al. 2023; 2021). OpenCDA integrates CARLA for realistic environment rendering (Dosovitskiy et al. 2017), SUMO (Simulation of Urban Mobility) for realistic background traffic (Lopez et al. 2018), and DI-Engine to provide a reinforcement learning environment¹. In collaboration with U.S.DOT CDA Research and the FHWA CARMA Program (Lochrane, Dailey, and Tucker 2020), OpenCDA, as an open-source project, is designed and built to support early-stage fundamental research for CDA research and development.

¹ DI-engine Contributors, "DI-engine: OpenDILab decision intelligence engine," <https://github.com/opendilab/DI-engine>, 2021.

This study is based on the new OpenCDA-HDI (human driver integration) comprehensive human-in-the-loop simulation framework developed upon the OpenCDA platform (Han et al. 2025; Xu et al. 2021), in collaboration with the UCLA Mobility Lab.² The framework is designed to support the study of human behavior in automated driving contexts, integrating existing OpenCDA capabilities with high-fidelity traffic simulation using CARLA (Dosovitskiy et al. 2017) and SUMO (Lopez et al. 2018). OpenCDA-HDI allows researchers to construct customizable driving scenarios that reflect a wide range of operational complexities and risk conditions.

A key feature of the platform is the shared control of the ego vehicle between a human driver and the ADS agent, through the integration of a haptic steering wheel and pedal setup, with automatically prioritized control to the human upon engagement. The platform supports real-time, high-resolution data logging mechanisms to capture human and ADS control inputs. By combining ADS functions, traffic simulation, and direct human input, OpenCDA-HDI enables reproducible and ecologically valid experiments. It provides a flexible and robust testbed for systematically studying human-automation interaction in shared-control environments. Additional details about the simulation environment will be published in a separate journal article (Han, Correa-Jullian, and Ma, under development).

The experimental setup consists of a medium-fidelity driving simulator based on OpenCDA's virtual simulation platform, and physical controls and hardware, including a Logitech G923 steering wheel and pedal bundle, a Next Level Racing F-GT simulator cockpit and seat, and a triple 32-inch monitor configuration (Figure 4). To conduct the driving simulator study, we developed a new module dedicated to human-system interaction modeling, based on the existing functionalities of both OpenCDA and CARLA platforms.³

² UCLA Mobility Lab, <https://mobility-lab.seas.ucla.edu/>

³ Additional details will be available in future publications.



Figure 4: The Driving Simulator Setup, Consisting of the Physical Steering Wheel, Cockpit, Seat, and Pedals, the Triple Monitor Configuration, and the Virtual Driving Environment OpenCDA.

Data Collection

Driving Simulation Environment

The data collection process consisted of subjective and objective measurements. The testing sessions were recorded through multiple methods.

The OpenCDA platform allows the selection of multiple scenario parameters, including weather and lighting conditions, and the behavior of the surrounding traffic and that of the ego-vehicle, i.e., TTC safety margins. For every scenario, the platform records various objective driving performance metrics, including reaction times, vehicle speeds, steering wheel and pedal positions, and adherence to traffic rules. Additionally, video and audio recordings of the steering wheel and pedals are captured to complement the simulator's data (not capturing the participant's face). Table 7 shows a summary of the variables that were collected during the experiments.

Table 7: Objective Measures Collected from CARLA Simulator.

Variable Type	Variable	Values	Source
Takeover Rate	System-initiated	Count	Simulation
	Driver-initiated	Count	Simulation
Reaction Times	Visual	sec	Recording
	Hands-on	sec	Recording
	Braking	sec	Simulation/Recording
	Steering	sec	Simulation/Recording
Vehicle dynamics	Max. Acceleration (Lat)	m/s^2	Simulation
	Max. Acceleration (Long)	m/s^2	Simulation
	Velocity	km/h	Simulation
	Throttle	%	Simulation
	Braking	%	Simulation
	Steering Angle	deg	Simulation
	Lateral Deviation	%	Simulation
Scene	Vehicle Gap/TTC	m or sec	Simulation
	Traffic Violations – Lane Invasion	Count	Simulation
	Traffic Violations – Lights/Signs	Count	CARLA
	Collisions	#	CARLA/Recording
	Stops	#	Recording
	Non-driving-related task engagement	%	Recording

Questionnaires

The study also captures subjective data from the participants through three questionnaires with single and multiple-choice questions, as well as open-ended questions. Pre-experiment questionnaires assess participants' characteristics, such as driving experience, risk tolerance, and baseline trust in automated driving systems. Post-experiment questionnaires assess the participants' experience with the simulator, such as their perceived workload, stress, and attention. Open-ended questions were used to clarify or explain answers on the questionnaire. A final questionnaire, repeating trust-related questions from the initial survey, was also administered.

Data Processing and Analysis

To analyze the data from the experiments, corresponding codes were developed in Python and R, leveraging the main statistical analysis packages. Given the variety of data collected, these are grouped into categories: continuous variables and categorical ordinal, count, and binary variables. The appropriate statistical models were used to analyze each of them:

- Continuous variables: Refers to data obtained through OpenCDA, e.g., vehicle kinematics.
- Count variables: Refers to processed data obtained through OpenCDA, e.g., the number of warnings triggered per scenario, per subject.
- Binary variables: Refers to processed data obtained through OpenCDA, e.g., whether a collision occurred per scenario, per subject.
- Ordinal variables: Refers to questionnaire data in Likert-scale form. Data from the initial and final surveys characterize each subject, while the test survey provides the participant's responses per scenario.

Data Preprocessing

Driver Profile

Responses from the initial and final questionnaires were mapped to numerical values, as summarized in Table 13 (Appendix). Some of the answers which were not inherently quantitative had to be coded, such as the preferred mode of transport or the participant's gender.

The ordering of the most frequent transport mode (question 1.4) was adapted to include e-scooters (a new addition to this study) (Bazilinskyy et al. 2018). Descriptive statistics—i.e., mean, standard deviation, and quartiles—were used to characterize each variable.

Test Scenario Data

The data collected during each simulated scenario was compiled by subject and by treatment. The data collected was processed and condensed into one file per scenario, such that each subject was linked to seven files: training scenarios 1-2-3 and test treatments A-B-C-D. Only the test scenarios were processed to study the

participants' behavior and performance. The raw data retrieved from each scenario is summarized in Table 14 (Appendix).

The raw data was processed to retrieve the following variables of interest:

- Vehicle kinematics, described by the mean, standard deviation, and the max/min values. Vehicle speed and TTCs were retrieved by subject, treatment, and control status (i.e., when the vehicle was controlled by OpenCDA or the driver). Throttle, brake, and steer input values were retrieved from the driver.
- The timestamp on different events, such as triggered warnings, TTC violations, takeovers, and handovers. This enabled the calculation of reaction times to warnings (A, C) or to tailgating events (TTC < 5 seconds) (B, D). If the driver took control of the vehicle prior to the warning or tailgating event, the reaction time was calculated as a negative value to the nearest triggering event, hence, it was marked as a driver-initiated control transition (as opposed to a system-initiated control transition). In addition, the number of warnings, tailgating events, and takeovers were collected per scenario.
- Safety-related incidents were collected based on their timestamps. In addition, traffic lane violations and collision events were collected by control status.
- Relevant time intervals collected include the total time the driver was in control of the vehicle, the total time the vehicle was in a lane violation, and the total time a warning was active.

The test survey data was mapped to numerical values, as summarized in Table 15 (Appendix). Descriptive statistics, i.e., mean, standard deviation, and quartiles, were used to characterize each variable.

Statistical Models and Tests

Various statistical models and tests were applied to the collected data to determine the impact of the two independent variables (warning availability and traffic conditions). A priori, models with random intercepts were considered to account for intra-subject variability; however, model complexity was reduced to fixed effects when necessary, based on model fit or convergence criteria. The methods and their application are described in the following sections. The methods included the following.

- Repeated measures analysis of variance (RM ANOVA) was used to analyze continuous outcome variables from the simulation data. Partial eta-squared was computed as a measure of effect size, and post-hoc pairwise comparisons were conducted using Bonferroni correction. This method was also used to analyze performance shaping factors (PSFs) derived from simulation measures.
- Generalized linear models (GLMs) were used to analyze binary (binomial distribution) and count (Poisson distribution) response variables from the simulation data. Post-hoc pairwise comparisons were conducted using Bonferroni correction.
- Cumulative link mixed models were used to analyze ordinal Likert-type responses from the test survey. These models included random intercepts to account for subject-level variability. Bonferroni-corrected post-hoc comparisons were used to examine differences across conditions.
- Linear mixed-effects models (LMMs) were used to analyze scenario-level PSFs constructed from aggregated Likert-type data. Normality of residuals was assessed using the Shapiro-Wilk test, and

homoscedasticity was evaluated with the Breusch-Pagan test. Bonferroni-corrected pairwise comparisons were also conducted where relevant.

Experimental Results

This section provides an overview of the experimental results, covering the driver profile extracted from the initial survey, the scenario perception extracted from the test survey, and both categorical and variables extracted from the simulation.

A total of 72 volunteers participated in the research study conducted during February-May 2025. Each participant experienced the four treatments or scenarios depicted in Table 6. After each experiment, the validity of the session was evaluated based on vehicle trajectory, warning triggers, and simulation stability. Abnormal scenarios were omitted, e.g., OpenCDA simulation was interrupted, or re-engaging the ADS resulted in failure modes out of the scope of the experiment. In the case of treatments A-C, the participants were required to see at least one warning during the experiment to be counted as valid. Given that the warnings are triggered once the TTC falls below 5 seconds, on many occasions participants disengaged the ADS before it was triggered. While these scenarios are valid to extract other relevant information, these subjects were omitted from the main and interaction effects analysis. After filtering out invalid experiments, the participant pool was reduced to 50 entries, ensuring at least two participants experienced the same treatment order. While these 50 valid entries were considered for the effects analysis section of this work, 61 out of the 72 participants may be revisited for further use to quantify relationships.

The dataset dimension is summarized in Table 8, consisting of 107 survey questions covering the initial, test, and final questionnaire per subject, and 41 variables extracted from the multiple continuous data streams from the simulation (e.g., mean, standard deviation, etc.).

Table 8: Total collected dataset.

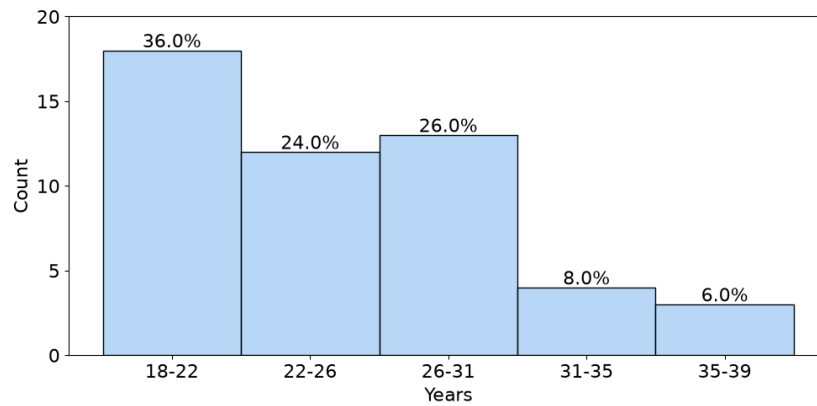
Element	Items	Repetitions	Total (Subject)	Total (Experiment Population)	Total (Valid Sample)
Initial Survey	35	1	35	2,520	1,750
Test Survey	15	4	60	4,320	3,000
Final Survey	12	1	12	864	600
Simulation	12 (41)	4	48 (164)	3,456 (11,808)	2,400 (8,200)
Total	74 (103)		155 (271)	11,160 (19,512)	7,750 (13,550)

For the construction of the Bayesian belief network (BBN), an estimated 57.2% of the data was used from the valid sample of 50 participants. In total, more than 90 hours of audiovisual footage were collected and will be analyzed in future publications. The remainder of this section contains the initial statistical description for surveys focused on driver profile and scenario perception, as well as the variables retrieved from the simulation environment. Main and effect analysis results are also provided for scenario and simulation-based variables.

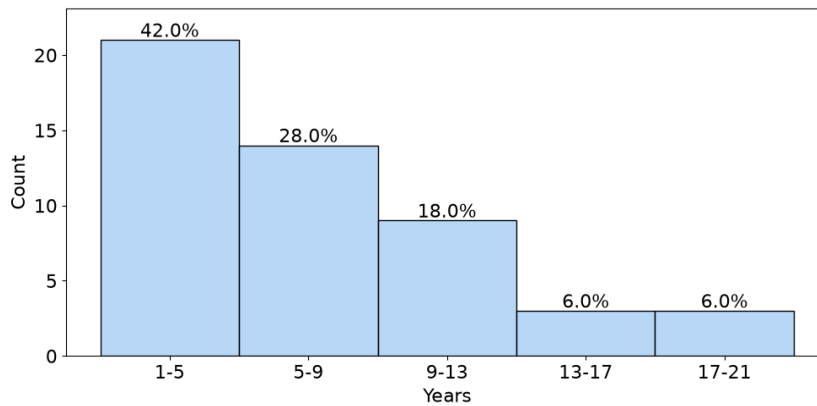
Driver Profile: Initial Survey

The participants' profile was extracted from their responses to the initial survey. The sample of 50 participants contained 32 male (64%), 17 female (34%), and 1 who preferred not to state gender (2%). The age and years since the participants obtained their driver's license are shown in Figure 5a and Figure 5b. The age range of the participants was 19 to 39 years, with a mean age of 25.68 years and a standard deviation of 5.14 years. The years since the driver's license was obtained ranged from 2 to 21 years, with a mean of 7.72 years and a standard deviation of 4.82 years.

The participant pool was mostly divided between using their private vehicle (46.8%) and public transportation (38.7%) as their main mode of transport (Figure 6). The majority reported using a private vehicle multiple times a week (54%), followed by once a month to once a week (36%). Heavy users represented only 6% of the pool, followed by 4% who reported driving less than once a month (Figure 7a). Regarding vehicle use, 36% of the participants reported an estimated 1,001-5,000 miles driven in the past year, while 22% reported less than 1,000 miles, 20% between 5,001-10,000 miles (Figure 7b), and 20% above 10,000 miles.



(a) Participant Pool: Age



(b) Participant Pool: Driver's License.

Figure 5: Histogram of (a) participant age and (b) years since driver's license was obtained.

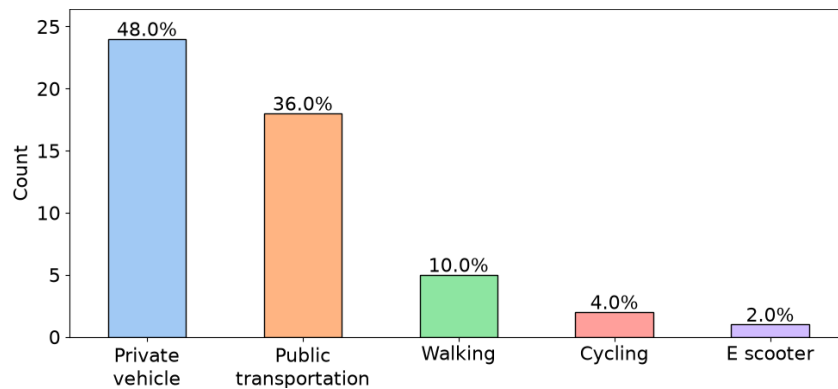


Figure 6: Participant Pool: Primary Mode of Transportation.

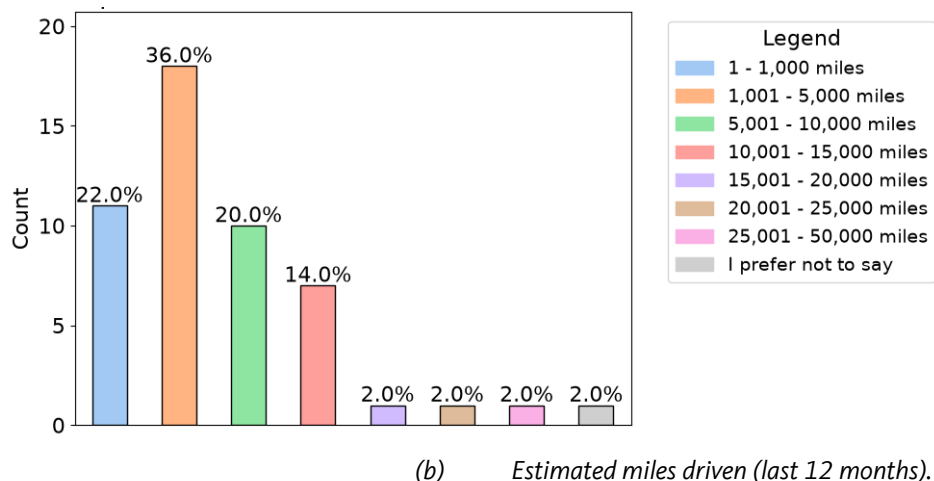
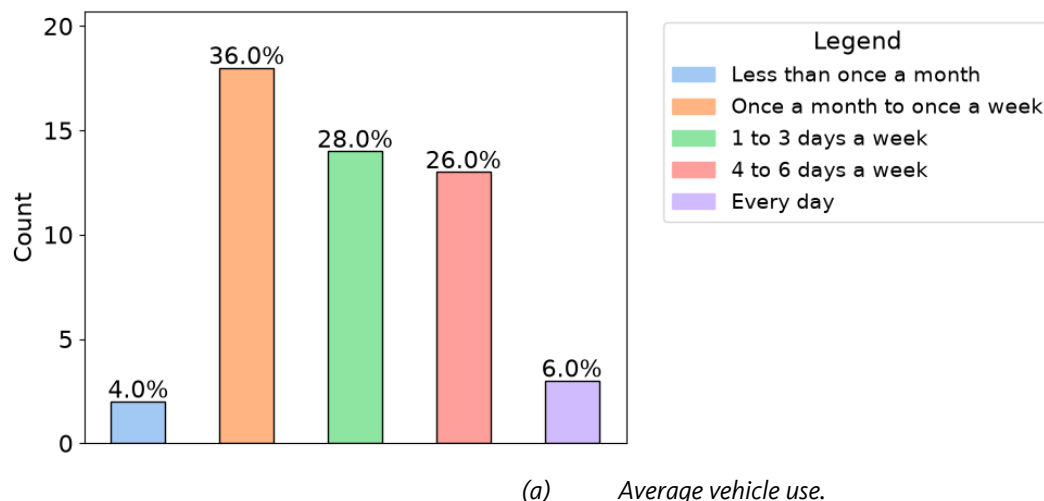
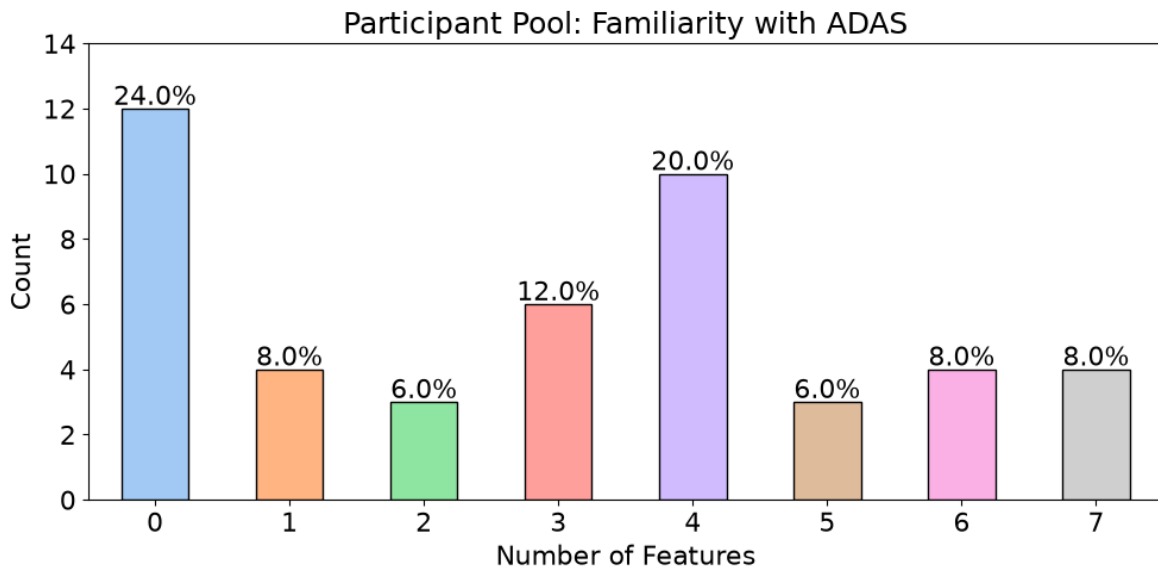
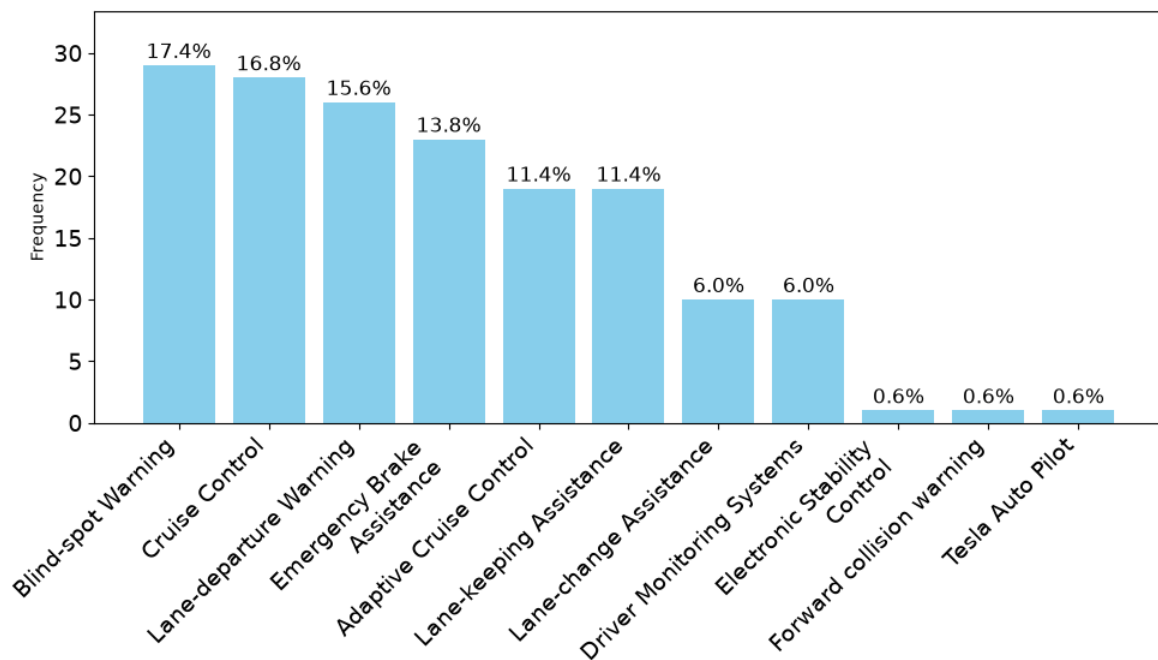


Figure 7: Participant Pool (a) Average vehicle use and (b) Estimated miles driven (last 12 months).

As shown in Figure 8a and Figure 8b, 76% of the participants reported familiarity with at least one ADAS feature, the two most commonly mentioned features being blind-spot warning (by 17% of participants) and cruise control (by 16.8%). Participants reported familiarity with other types of warnings, such as lane-departure warnings (15.6%) and forward collision warnings (0.6%), as well as automated responses, such as automated emergency brake (AEB) systems (13.8%). Lane-keeping and lane-change assistance were mentioned by 11.4% and 6.0% of respondents. Regarding more advanced features, adaptive cruise control was mentioned by 11.4% and driver monitoring systems by 6.0%. Notably, Tesla Autopilot (Level 2) was mentioned by one participant (0.6%).



(a) Number of reported known ADAS features.



(b) Histogram of reported ADAS features.

Figure 8: Reported ADAS feature familiarity (a) number of known features and (b) percentage of mentions.

The next section of questions in the initial survey aimed to characterize the participants' perception of their enjoyment of driving, of interacting with ADS-equipped vehicles (level of driving automation unspecified), and their attitude towards driving incidents (i.e., external loci of responsibility) (Choi and Ji 2015). Figure 9 displays the distribution of responses on a 5-point Likert scale (from “Strongly Disagree” to “Strongly Agree”) for each question (1.8-1.11). The first six items (1.8.1-1.8.6) are interpreted as the participant's effective evaluation of manual driving versus autonomous vehicle (AV) usage. The first three (1.8.1-1.8.3) overall show a strong positive tendency from the sample, with over 70% of participants agreeing or strongly agreeing to enjoy manual driving. The last three elements (1.8.4-1.8.6) repeat the questions but are aimed at automated driving. Here, most of the participants (around 60–64%) express agreement, but this is less intense than for manual driving, and a small fraction (8–10%) disagree.

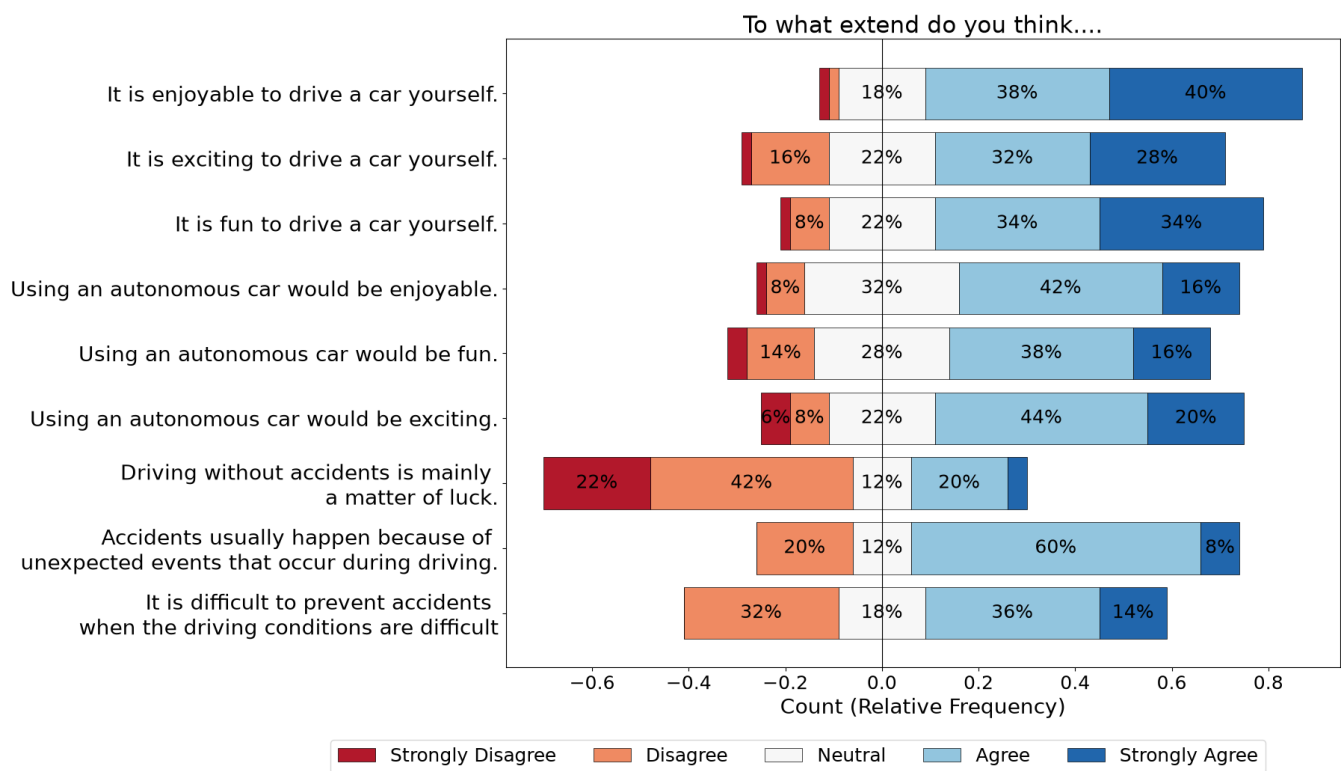


Figure 9: Initial Survey - Bias and Responsibility-Related Questions.

Overall, this suggests that while participants may see some potential enjoyment in AVs (positive bias towards ADS), they still prefer the experience of driving themselves (negative bias towards ADS). The last three items refer to the participant's locus of control in accident causation, which potentially can shape how they evaluate both their own and the ADS's responsibility in ensuring driving safety. In general, these responses indicate that participants perceive an internal locus of control rather than luck but recognize the role of unexpected events and complex driving conditions.

Figure 10 presents an overview of the participants' responses to Question 1.9 of the initial survey, which asks how frequently they engage in certain driving behaviors, represented on a 5-point Likert scale indicating frequency, from 'Never' to 'Always'. These questions are related to errors and violations (DBQ) as well as the driver's attention to common events (De Winter 2013). These questions are interpreted as being related to the participant's risk tolerance, i.e., the level of risk seeking or risk adverse behavioral tendencies. For items 1.9.1-1.9.4, a higher frequency implies higher tolerance for driving errors and violations. Regarding driver violations, responses related to speeding are mostly mixed (34% very often, 34% rarely), but indicate moderate tolerance for speeding, while tailgating is reported as a less common violation (54% rarely, 34% never). Regarding driver errors, wrong lane selection or curb collisions are moderately uncommon (36% rarely, 36-50% never, respectively). Items related to driver distraction indicate a stronger risk averse tendency, with frequent awareness of other road users (50%-52% always) and of traffic lights (40% never, 42% rarely miss traffic lights).

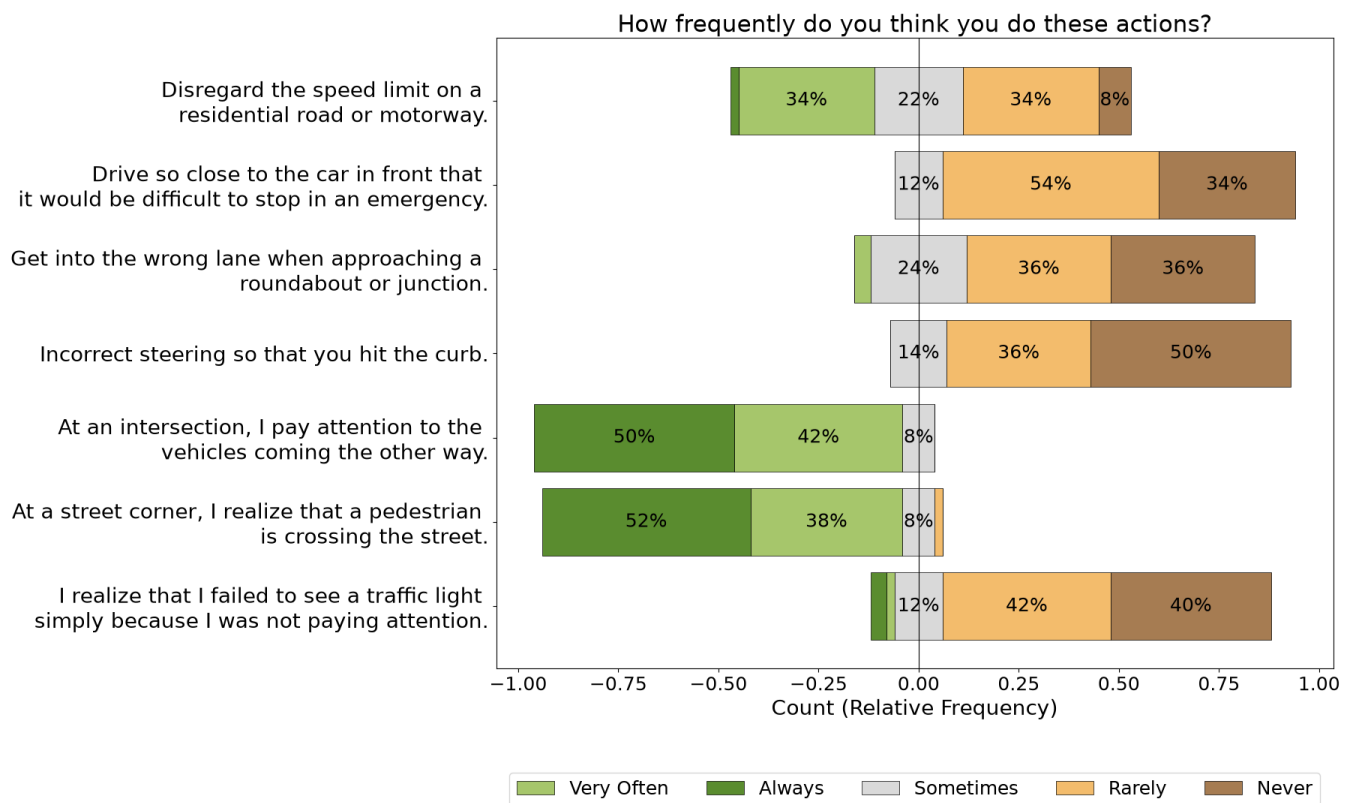


Figure 10: Initial Survey - Risk Tolerance-Related Questions.

The final set of questions in the initial survey aimed at characterizing participants' trust in automated driving (Manchon, Bueno, and Navarro 2022; Payre, Cestac, and Delhomme 2014; 2016). Figure 11 shows the distribution of responses on a 7-point Likert scale ('Strongly Disagree' to 'Strongly Agree'). Items indicating preconceived trust in ADS correspond to 1.12.1-1.12.2 and 1.12.4-1.12.9, which are summarized as follows:

there is a moderate to high baseline trust that ADS are safe (42% Agree, 18% Slightly Agree, 14% Strongly Agree), but less indication that it is safer than manual driving (24% Agree, 20% Slightly Agree, 14% Strongly Agree).

There is a general tendency to trust the ADS decision-making process (64% Agree or Slightly Agree) and participants were willing to delegate control under certain conditions for convenience: 40% Agree/30% Slightly Agree if driving was boring; 50% Agree, 22% Slightly Agree if driver was tired. However, participants were more conditional about their trust in complex driving situations (44% Agree/Slightly Disagree, 44% Disagree/Slightly Disagree) or adverse weather conditions (40% Agree, 32% Disagree).

Items potentially indicating mistrust or preference for manual control (Table 10: 1.12.3 and 1.12.10-1.12.12, Appendix) indicate a strong preference for retaining authority to intervene and take control of the vehicle if the situation would require it (52% Strongly Agree, 46% Agree/Slightly Agree). The participants were divided regarding their preference for manual control when driving themselves or passengers (50% prefer manual control on every occasion, 54% if there were other passengers on board) but strongly trusted their capacity to regain control effectively if required, indicating their belief that human intervention may be still needed (38% Strongly Agree, 28% Agree, 24% Slightly Agree). In general, participants exhibited a conditional view of the ADS's capacities, potentially affecting their takeover performance in a fallback scenario.

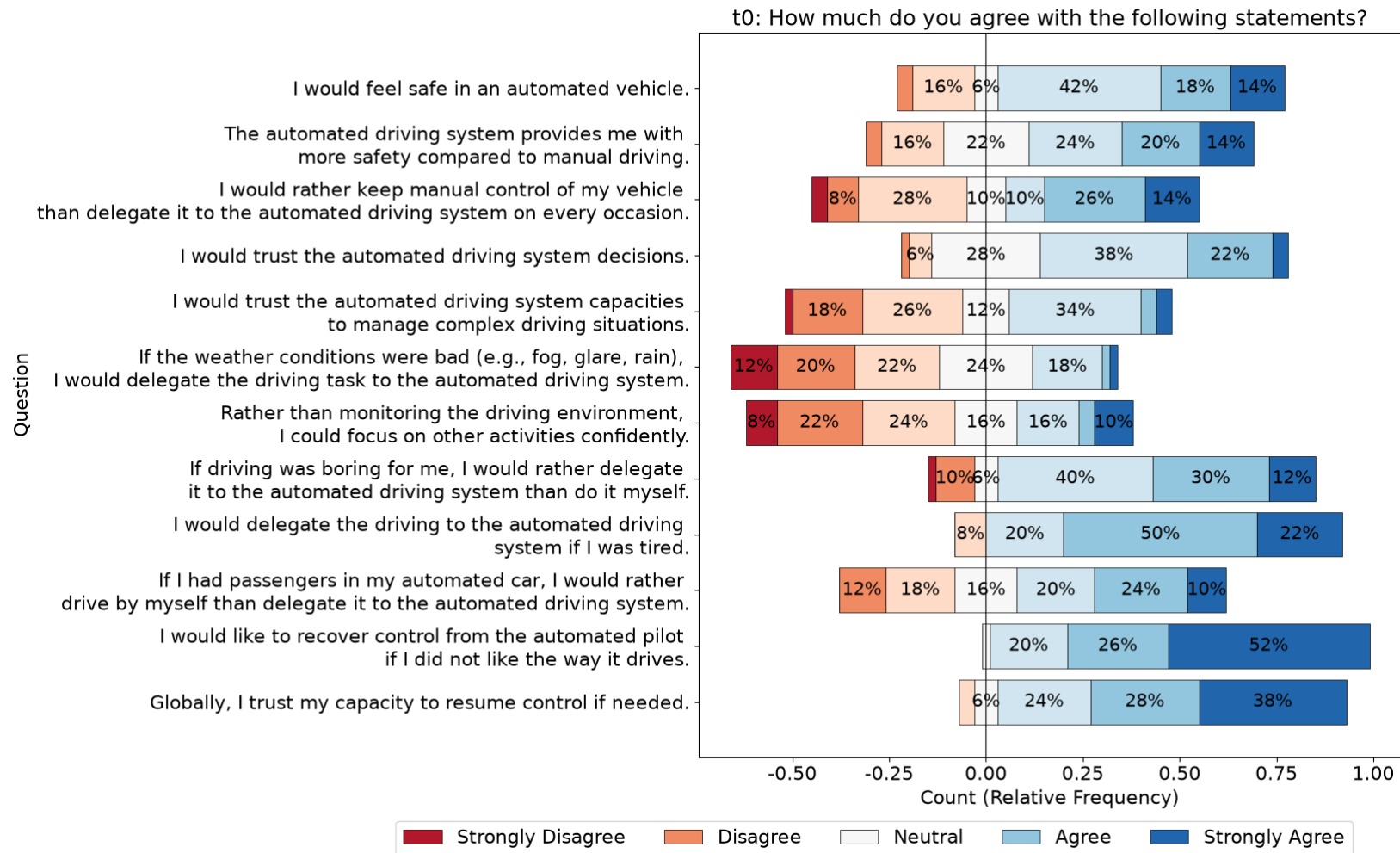


Figure 11: Initial Survey - Trust in Driving Automation.

Scenario Perception: Test Survey

Four test surveys were collected during the test scenarios for each subject, one per treatment. Questions 2.1 to 2.20 were used to characterize the driver's perception of the scenario. Figure 12 and Figure 13 show the aggregated results for the test survey sample across all scenarios (A, B, C, D) and all participants, providing a general impression of how participants perceived the system's behavior, their own role in events, and their trust and understanding of the automated system. Values below 5% are excluded from the figure labels. Figure 12 focuses on questions 2.1-2.9, while Figure 13 provides the distribution of responses for questions 2.10.1-2.10.6 referring to workload.

As shown in Figure 12, participants reported a high perceived transparency of the ADS (Question 2.1, 24% Strongly Agree, 42% Agree) even if they did not agree with its behavior in the scenario (Question 2.2, 50% Disagree or Neutral), suggesting a general misalignment with user expectations. While agreement varied somewhat across scenarios, reported understanding was mostly constant. Notably, after experiencing the scenarios, the perception on whether the ADS was trustworthy (Question 2.3) was skewed towards disagreement (22%-30% Disagree/Slightly Disagree, 19%-10% Slightly Agree/Agree) – particularly in the high traffic scenarios (B, D).

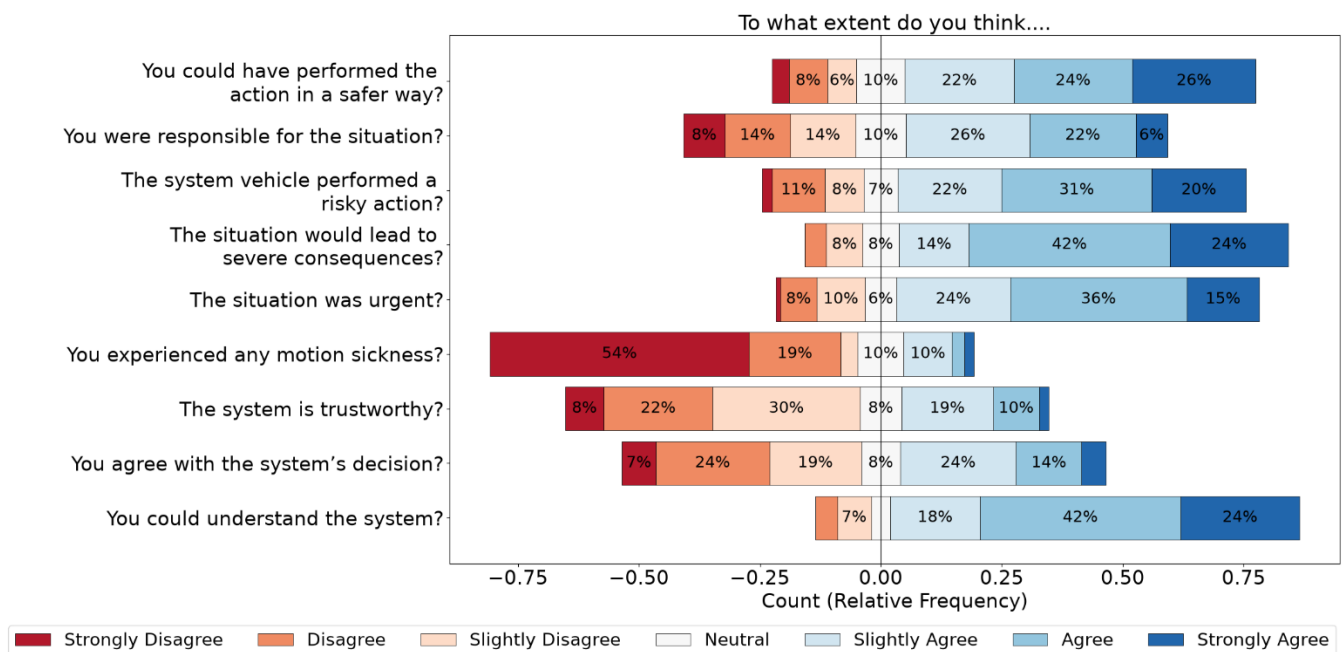


Figure 12: Summary of Test Survey Results.

Regarding the participants' perception of the simulated driving scenario, most participants agreed to some extent that the system performed a risky action during the scenario (Question 2.7). Higher risk awareness was reported in scenario B, followed by C-D. Potential consequences were generally perceived as severe (Question 2.6, 42% Agree, 24% Strongly Agree) and reported a high sense of urgency (Question 2.5, 36% Agree, 15% Strongly Agree), suggesting that the experimental setup was sufficient to engage participants. Motion sickness (Question 2.4) was not considered for future analysis. In relation to the participant's self-perception, mixed results were obtained regarding internal *loci* of control and responsibility in the situation's outcome (Question 2.8, 22% Agree, 26% Slightly Agree, 26% Disagree); however, most participants agreed that they could have performed the action in a safer way (Question 2.9, over 70% Agree).

Figure 13 presents the aggregated NASA-TLX workload ratings across all participants and scenarios, assessing six core dimensions of workload on a 9-point Likert scale. Mental demand was the workload dimension with the highest rating, with over 70% of participants reporting high levels (26% Slightly High, 22% Moderately High, 11% High, 10% Very High). As shown in Table 16 (Appendix), Mental Load presented the highest mean values of the workload dimensions, followed by Performance, Efforts, Temporal, Frustration, and finally Physical Load. Scenario C presented the highest mental demand, potentially because of the combination of high traffic and pressure that participants felt to respond to the warnings (as opposed to scenario D, no warnings), while scenario A (warning, low traffic) presented the lowest mental demand. Independent of the specific scenario tested, this corroborates results usually found in literature, indicating a higher mental workload for drivers interacting with systems in shared autonomy regimes.

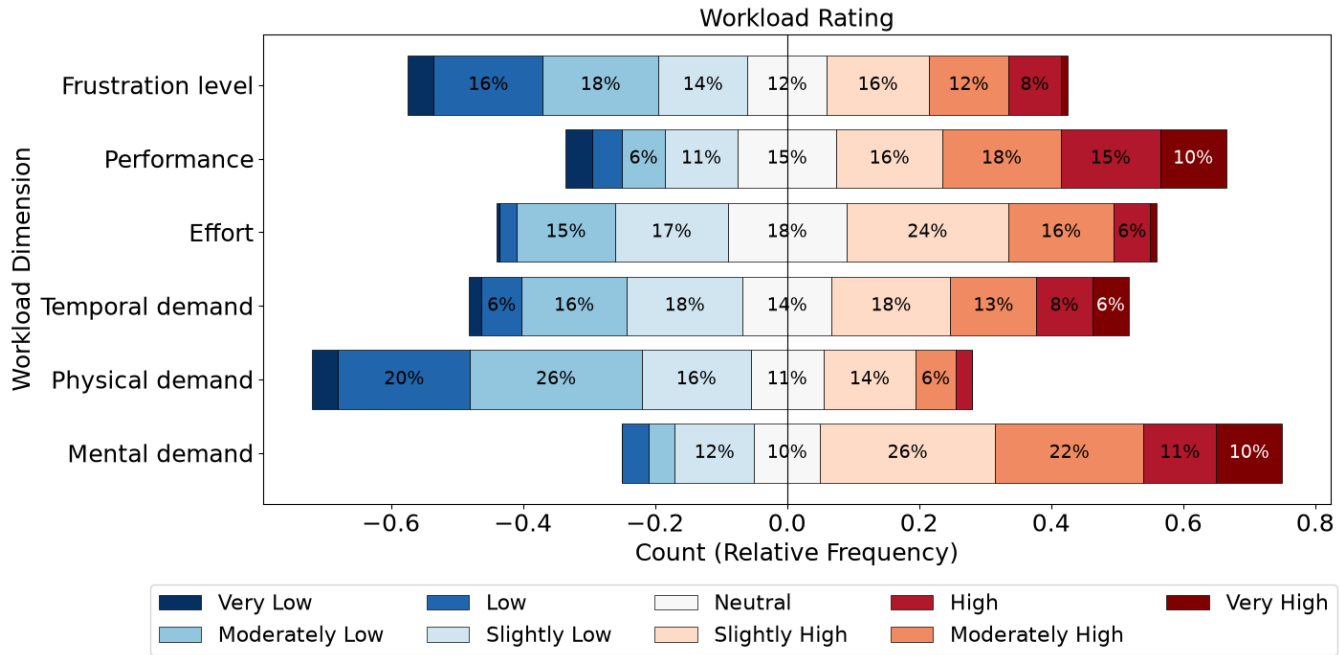


Figure 13: Summary of Workload Ratings.

Results from testing at L2-L3 indicate higher cognitive load discerning the safety risks of the driving environment and of the behavior of the ADAS or ADS-equipped vehicle (Stapel, Mullakkal-Babu, and Happee 2019). In contrast, physical demand is the lowest dimension rated, with over 65% of participants reporting low demand (4% Very Low, 20% Low, 26% Moderately Low, 16% Slightly Low). Both scenarios in high traffic (B, D) presented higher physical demand values. Although during the experiments several participants reported difficulties in braking (due mostly to self-reported miscalculations on how much force was required or timing), participants did not necessarily attribute this to physical demand.

Regarding temporal demand, participants showed a more mixed response. Around 45% of participants reported high time pressure (6% Very High, 8% High, 13% Moderately High, and 18% Slightly High), while around 40% reported lower temporal demand. Notably, the temporal demand increased in the high traffic scenarios (C, D). Effort, described as the combination of mental and physical demand, exhibited a similar distribution but tended towards neutral values (24% Slightly High, 18% Neutral, 17% Slightly Low), and was relatively stable across scenarios. While participants rated performance as being a fundamental aspect of the workload (with higher values reported for the low traffic scenarios A-B), frustration had a tendency towards slightly lower values. However, around 36% of participants reported moderate to high frustration, possibly due to mismatched expectations with the system's behavior, in particular with scenario D (no warning, high traffic).

In summary, even though participants were not physically engaged, the mental and attentional burden remained high, suggesting that automation did not reduce workload, but impacted the nature of it. The low performance ratings indicate a potential gap between participant expectations and system outcomes, aligning with earlier trust data. Moderate frustration and temporal pressure suggest that scenarios were realistically demanding, but may need interface, communication, or control improvements. Scenario A (warning and low traffic) consistently was associated with lower workload, across all dimensions.

Main Effects and Interaction Effects

A key aspect of the experiments was to determine the difference between warning (Factor A) and traffic complexity conditions (Factor B). The distribution of responses by treatment conditions is found in Figure 14 for all scenario variables, except for workload dimensions which are pictured in Figure 15.

The analysis of main and interaction effects focused on evaluating the effects of two experimental factors on the Likert-scale variables categorized into two groups: L7 (1–7 scale) and L9 (1–9 scale) through cumulative link mixed models. Table 16 (Appendix) describes the Likert responses for the test survey, as well as the p-values for statistical significance of the main and interaction effects of Warning availability (Factor A) and complexity of driving environment (Factor B). Notably, there are significant interaction effects between warning availability and the complexity of the driving environment for multiple variables.

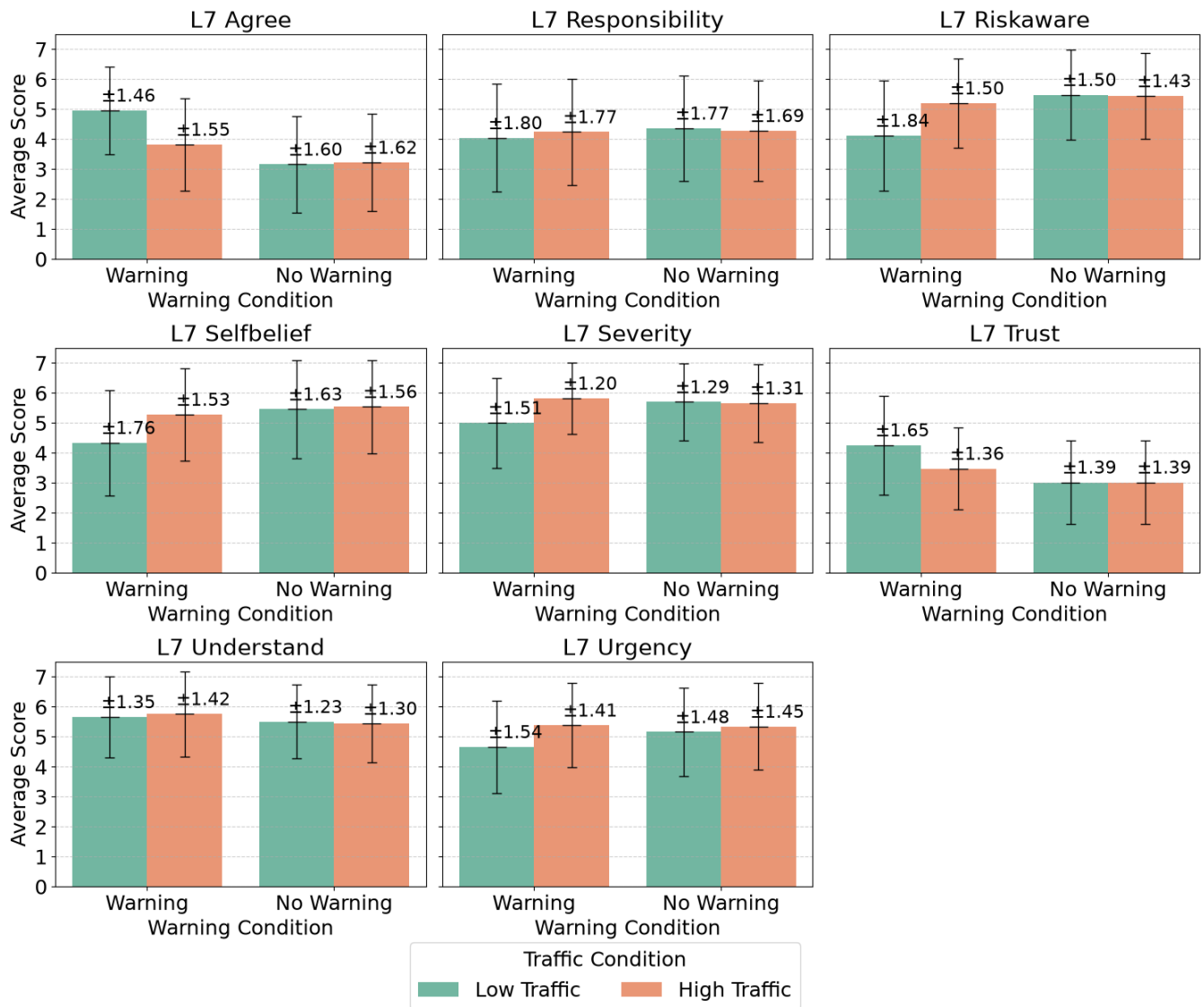


Figure 14: Average Scores for Scenario Perception by Treatment.

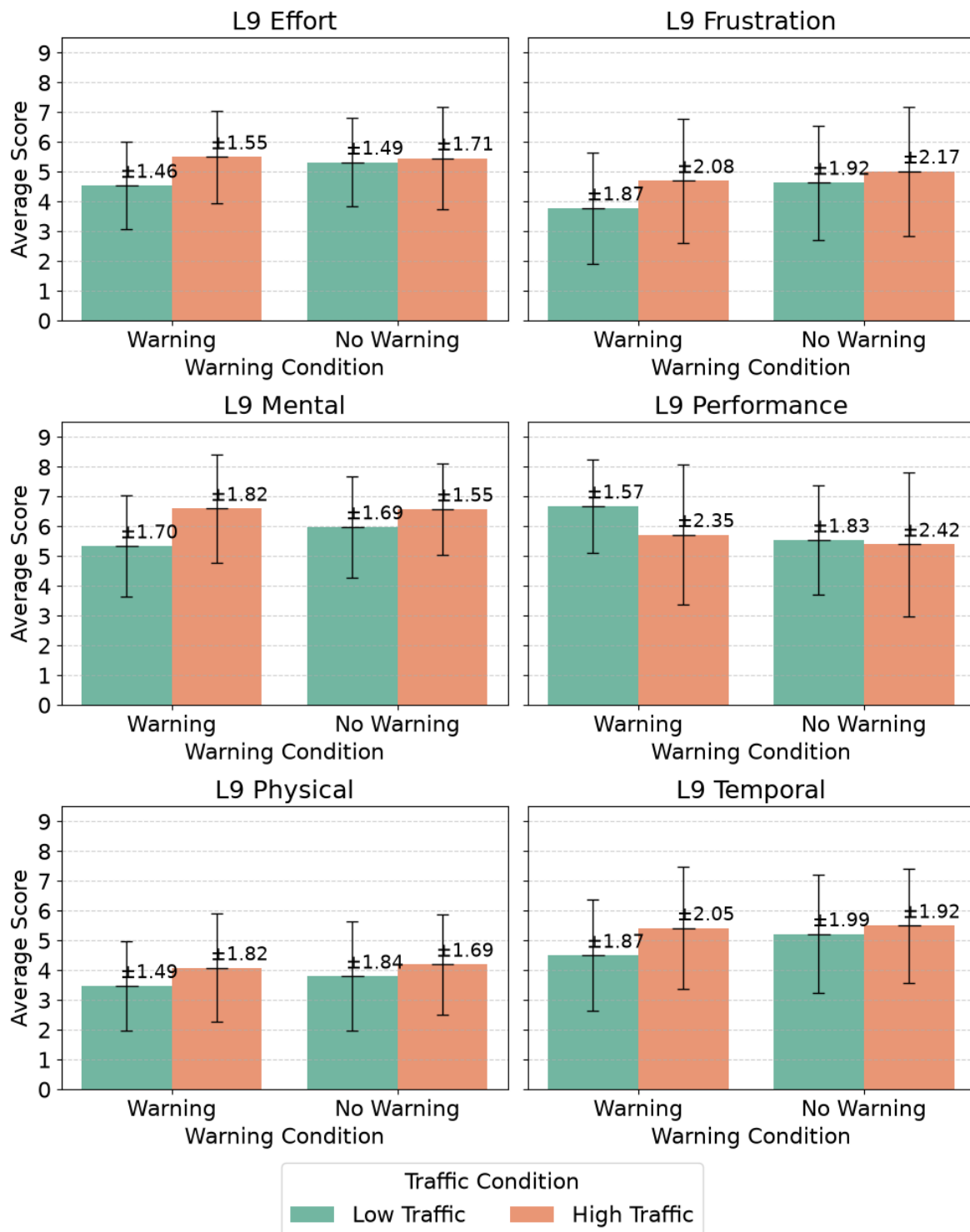
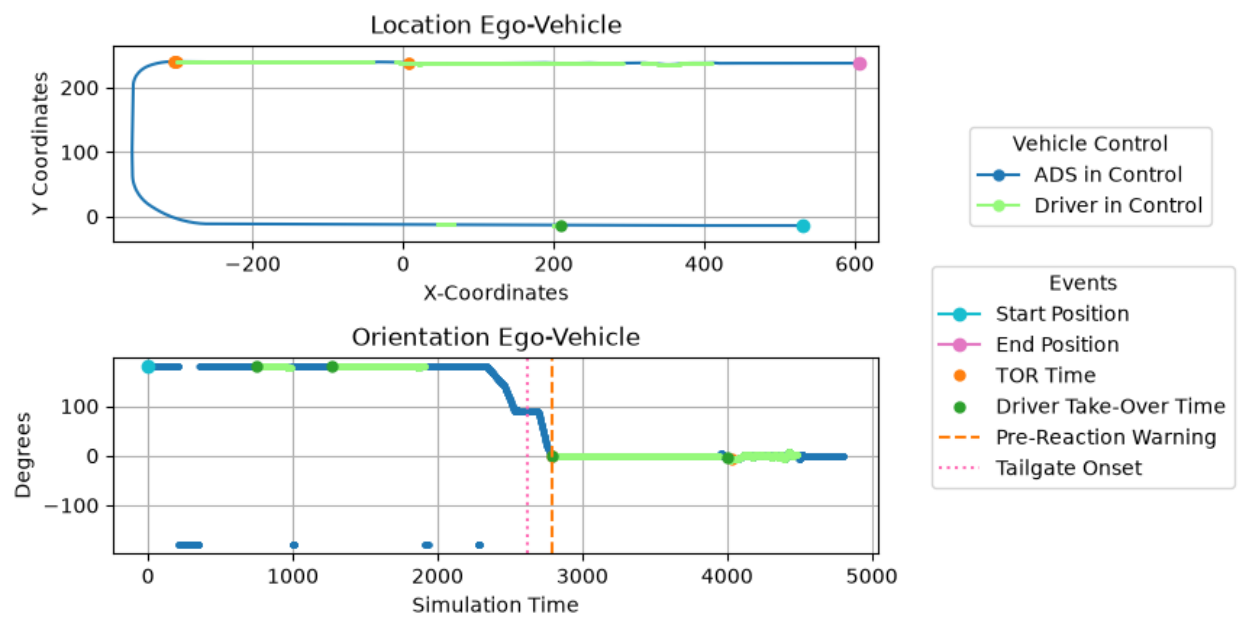


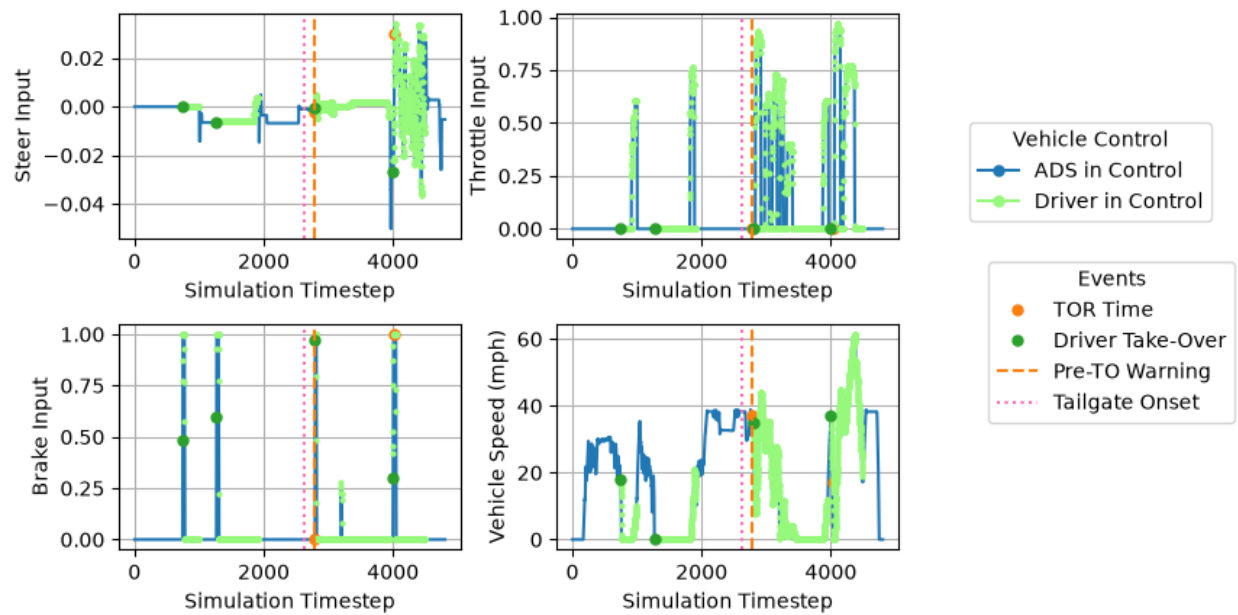
Figure 15: Average Scores for Workload by Treatment.

Simulation Data: Continuous Variables

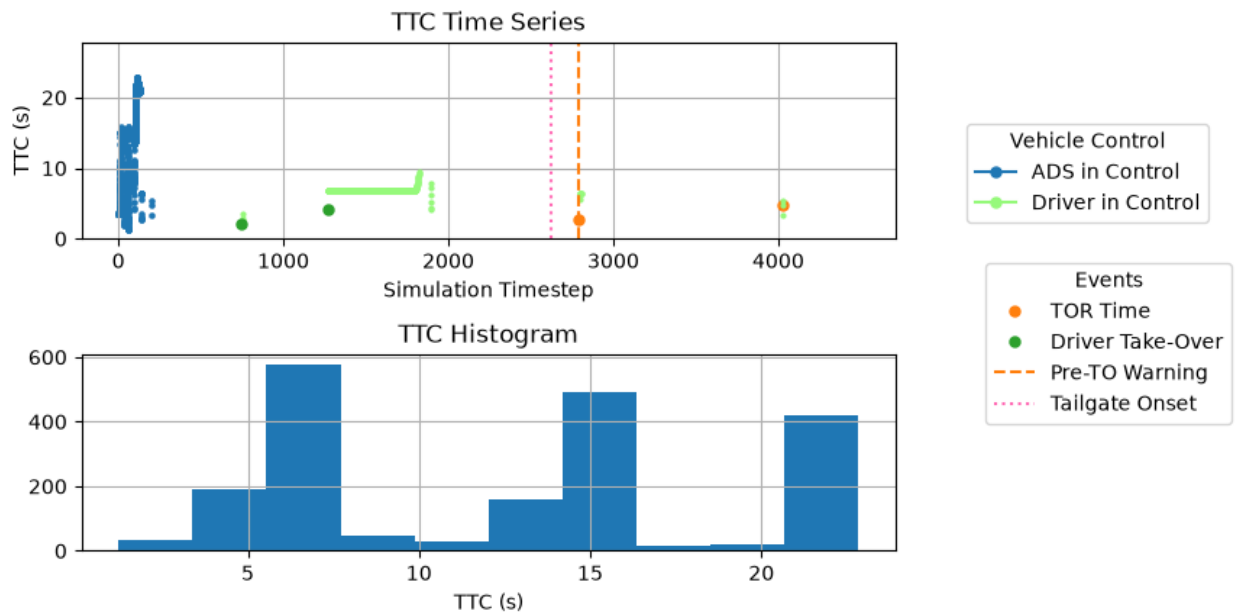
Data was processed to extract key variables and corresponding timestamps. A sample scenario for participant P41 is depicted in Figure 16a, Figure 16b, and Figure 16c, where the recorded vehicle’s trajectory, driver input, and time-to-collision (TTC) is presented, along with key events (e.g., takeover request time, tailgate failure onset) that impact scenario evolution.



(a) Vehicle Trajectory



(b) Recorded Driver Input



(c) Recorded TTC Distribution

Figure 16: Driving Simulator Experiment Sample Data.

The continuous simulation variables are grouped into five main categories:

- 1) Scenario Timing and Collision Metrics
 - End Time, Coll. Time: Indicate when the scenario ended or a collision occurred, serving as high-level indicators of scenario outcomes.
 - Warning Time, Warning Total: Measure the activation time and duration of automated warnings.
- 2) Driver Intervention Metrics
 - Driver Time, Driver Total: Capture when and how long a human driver took over control from the ADS.
 - Driver Warning Reaction Time, Driver Tailgating Reaction Time: Quantify driver responsiveness to warnings or tailgating events.
- 3) Lane Keeping and Vehicle Positioning
 - Lane Inv. Time, Lane Inv. Time Driver: Measure the duration of lane boundary violations, with the latter restricted to periods under human control.
- 4) Vehicle Dynamics and Control Inputs
 - Veh. Mean Speed, Std, Max: Capture speed patterns overall and during driver control.
 - Steering, Throttle, and Brake (Mean, STD, Max): Reflect driver control input variability, with implications for stability and maneuvering behavior.
- 5) Time-to-Collision (TTC) Metrics
 - TTC Mean, TTC Std, TTC Min: Estimate proximity to potential collisions and risk exposure overall and during driver control.

Main Effects and Interaction Effects

Interaction plots were produced to visualize and assess the impact of warnings (A) and traffic complexity (B) and complemented with the RM ANOVA results summarized in Table 17 (Appendix). This table summarizes scenario-level metrics collected for each treatment and provides the means by factor (A: Warning, B: Traffic) and the results from the RM ANOVA tests.

Regarding scenario duration and collision outcomes, scenarios without warnings (A1) and lower traffic complexity (B0) tended to end sooner, suggesting either quicker scenario progression or earlier failures (End Time). In addition, collisions occurred less frequently in scenarios with warnings, indicating that these conditions supported safer outcomes (Coll. Time is closer to End Time). As expected, both metrics related to warning activation (Warning Time and Warning Total) were significantly higher in the warning condition by design (as reflected in Figure 17). Traffic complexity did not apply to these variables, as warnings were not triggered in certain scenarios (e.g., B and D).

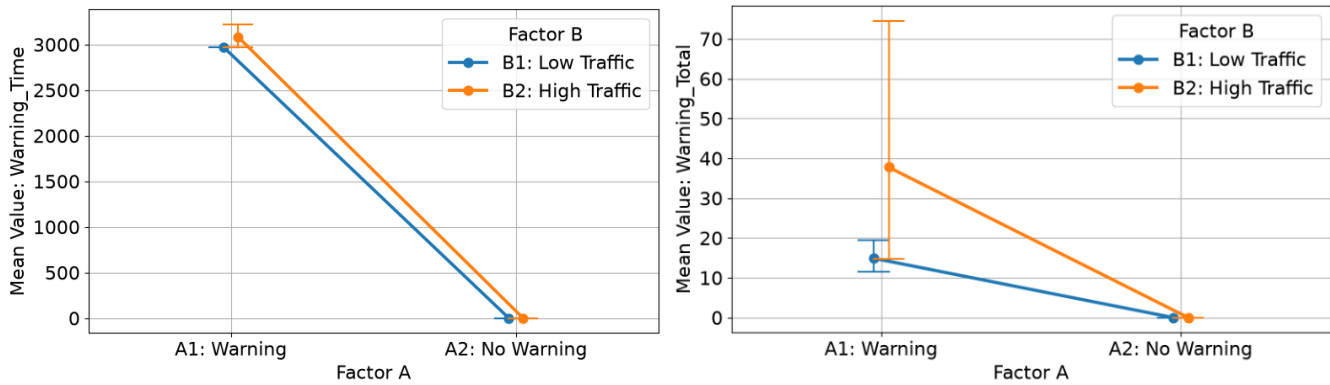


Figure 17: Interaction Plot: Time to First Warning (left) and Total Time Warnings Were Active (right).

In terms of driver intervention, drivers took over more frequently and remained in control longer in high complexity scenarios (refer to Figure 18). The presence of warnings did not significantly affect the likelihood or duration of driver intervention (Driver Time, Driver Total). In general, lane violation behavior was not impacted by warning availability. However, this behavior tended to occur more frequently in scenarios with low traffic complexity (e.g., scenarios A, B) potentially due to a higher variability in the sparse traffic.

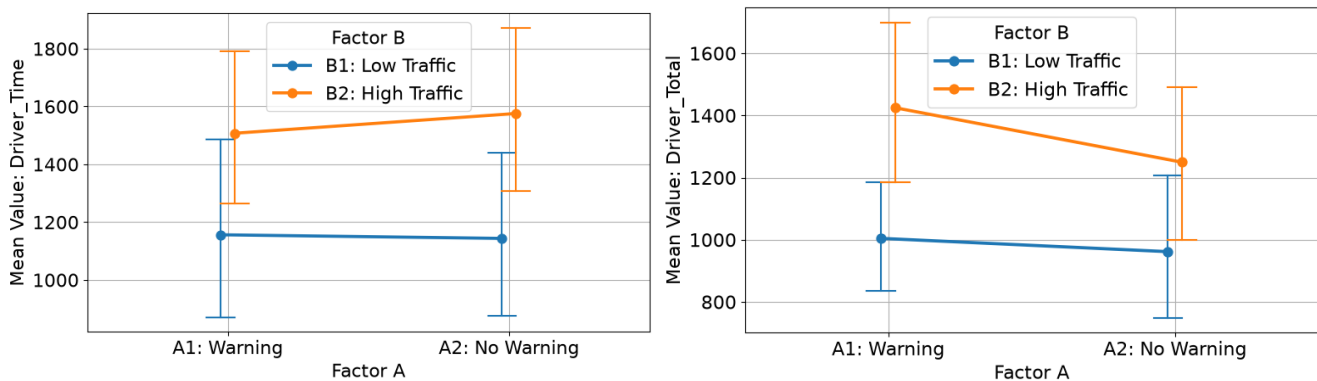


Figure 18: Interaction Plot: Time to First Takeover (left) and Total Time Drivers Were in Control of the Vehicle (right).

The vehicle kinematic variables (mean, standard deviation, and max vehicle speed) were collected to compare the driver's behavior to the default ADS control. A significant interaction effect showed that scenarios with warnings and low complexity led to higher mean speeds. Drivers tended to drive faster when warnings were available, possibly because they took over later or felt more confident post-takeover (Figure 19). Speed variability increased in high-complexity scenarios, which is consistent with the need for more reactive or adaptive speed modulation. Maximum vehicle speed did not significantly change across conditions, indicating upper speed limits were stable regardless of scenario.

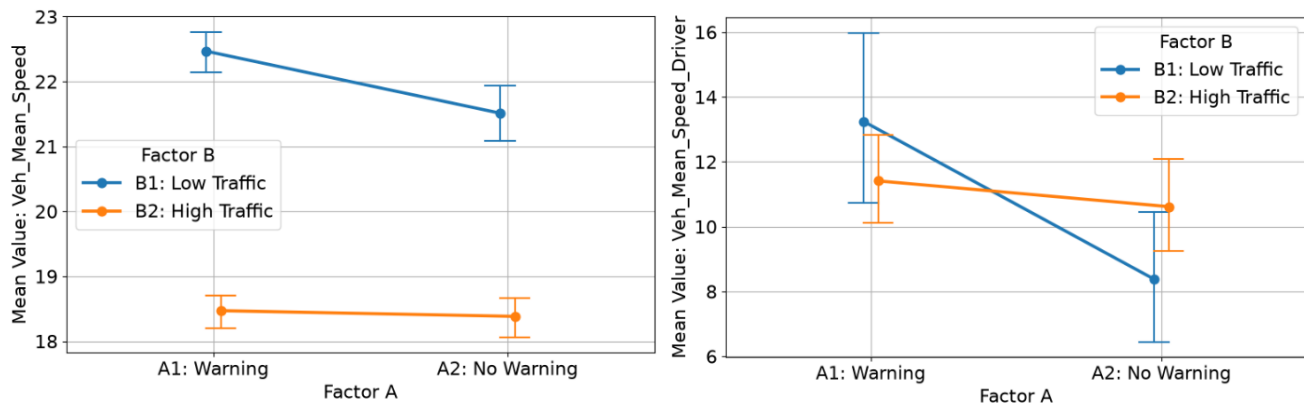


Figure 19: Interaction Plot: Mean Speed During the Scenario (left) and When Drivers Were in Control of the Vehicle (right).

Other driver behavior indicators, such as steering, throttle, and brake inputs were analyzed. Steering behavior under high-complexity conditions had sharper (high mean values) and more frequent (higher standard deviation values) directional changes. Maximum steering input remained unchanged across conditions, likely reflecting the nature of the selected vehicle trajectory. Mean throttle inputs showed a significant interaction effect between warning and complexity, indicating context-dependent differences in acceleration (Figure 20). However, no significant effects were observed for maximum throttle nor throttle variability. Despite indications provided by Figure 21, braking behavior was not significantly affected by different driving conditions, only showing higher variability in braking in the no-warning condition, suggesting less consistent braking patterns.

In general, considering the entire vehicle trajectory (under the driver's and the system's control) warnings led to significantly higher TTC mean, indicating safer distances, although the TTC standard deviation and TTC minimum were not significantly different. However, when only considering when the driver was in control of the vehicle, only TTC standard deviation (Std.) was affected by warning availability. Higher traffic scenarios significantly reduced TTC mean, standard deviation, and minimum values, pointing to generally more constrained following distances. Indeed, TTC varied more in low-complexity scenarios, possibly due to higher speeds and more spacing fluctuations.

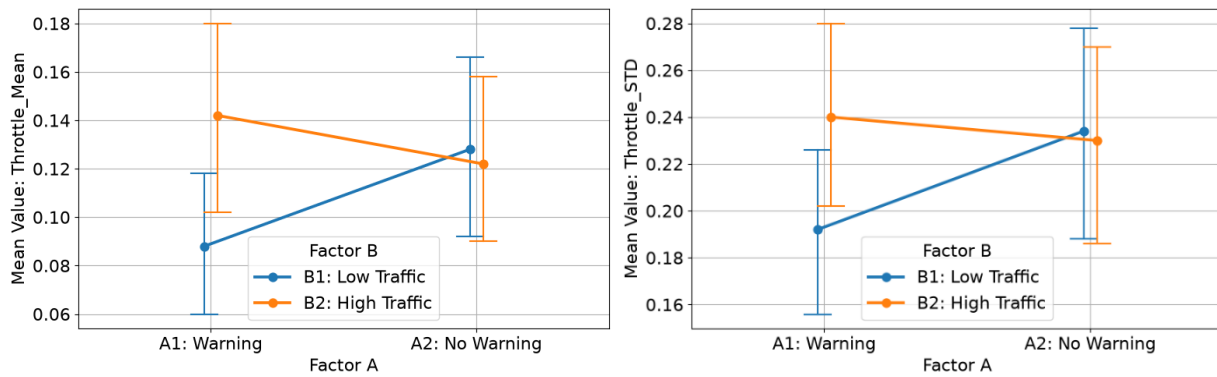


Figure 20: Interaction Plot: Driver Throttle Inputs – Mean (left) and Standard Deviation (right).

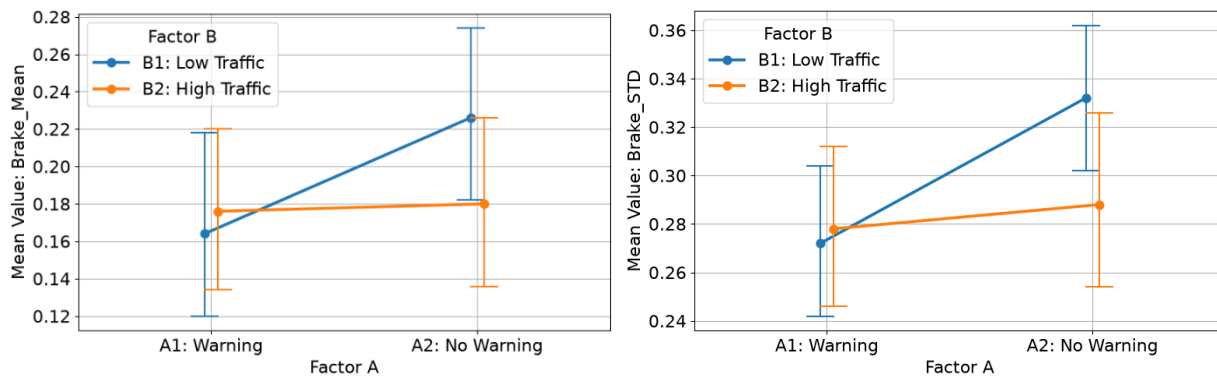


Figure 21: Interaction Plot: Driver Brake Inputs – Mean (left) and Standard Deviation (right).

Driver TTC metrics, however, remained largely stable across conditions, except for TTC Std Driver, which showed more variation in warning scenarios, suggesting that warnings may induce more reactive behavior (Figure 22). Warnings significantly affected Driver Warning Reaction Time and Tailgating Reaction Time, indicating faster or more anticipatory responses under warnings (Figure 23).

In summary, Warnings (Factor A – A0) generally improved safety-related outcomes: later collision times, longer TTC, and more anticipatory driver reactions. High Traffic (Factor B – B1) increased control complexity, with more human intervention, reduced following distances, and more steering variation. Interaction effects were limited but suggest that in some cases—in relation to vehicle speed and minimum safety distances—the combination of warning presence and traffic density leads to non-additive effects on driving behavior and safety.

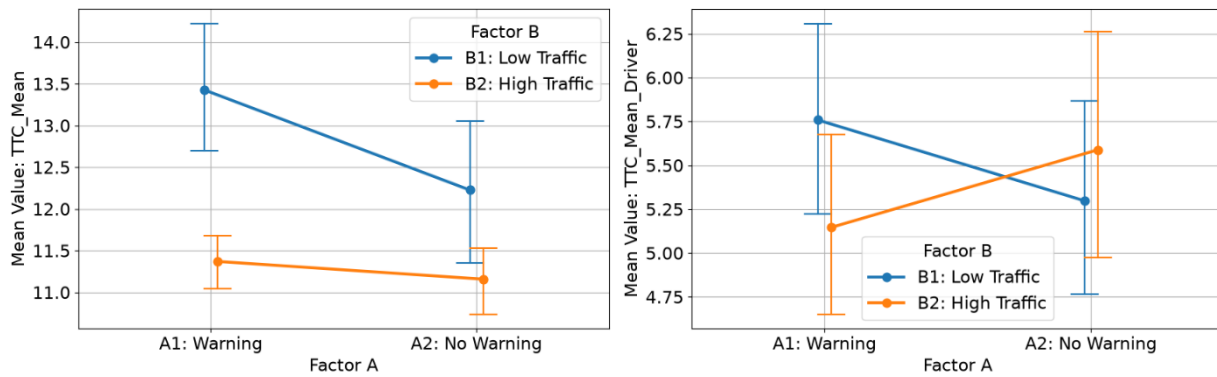


Figure 22: Interaction Plot: Mean TTC During the Scenario (left) and When Drivers Were in Control of the Vehicle (right).

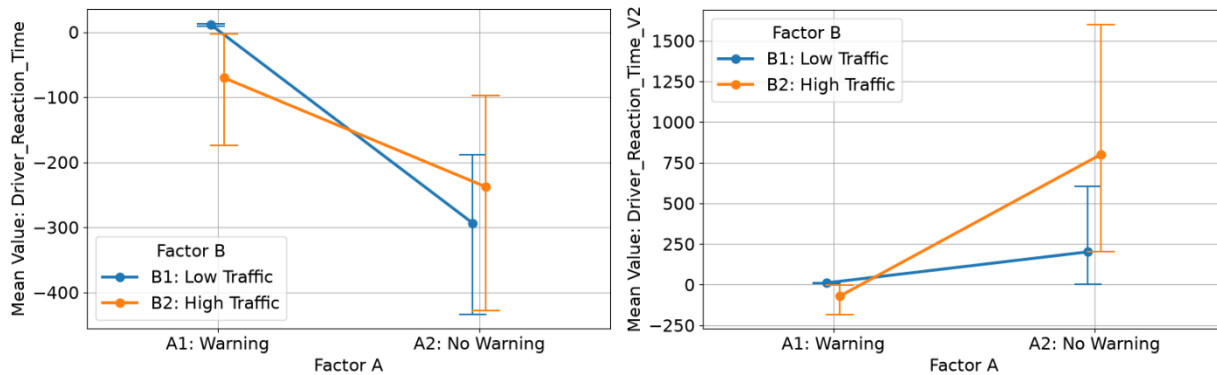


Figure 23: Interaction Plot: Driver Reaction Time to Warnings (left) and Tailgating Events (right).

Simulation Data: Categorical Variables

Key safety-related events were collected from the continuous simulation variables and tested for main and interaction effects. These consist of:

- Binary variables: whether a collision occurred during a scenario, modeled by a Binomial GLM. This is translated into the probability of having a collision during a scenario depending on the treatment.
- Categorical count variables: the number of events occurring during a scenario depending on the treatment, modeled by Poisson GLM. This included the number of hard-braking events (as a proxy for near-miss events), driver takeovers, and tailgating events.
 - Hard-braking events are defined as timestamps where the driver brake input was above 80% of the value range [0, 1].

- Driver takeover events were defined by the number of times the vehicle experienced control transitions from the ADS to the driver.
- Tailgating events are defined as the timestamps where $TTC < 5s$ safety threshold, counted starting from the onset of tailgating behavior, i.e., after the ADS experiences the induced failure. These events were collected in all scenarios, independent of whether a warning was triggered.

A summary of these variables is provided in Table 18 (Appendix), reporting the mean and standard error per scenario across all subjects. This table also includes the estimated upper and lower boundaries (5%, 95%). In relation to the probability of collision, Figure 24 shows the predicted values grouped by treatment. Treatment A (Warning, Low Traffic) exhibits the lowest collision rate at 0.06 (0.02, 0.17). However, the impact of no warnings (B, D) and high traffic (C, D) increase the probability of collision to similar values. While treatment D (No Warning, High Traffic) results in the highest probability 0.52 (0.38, 0.65), its overlap in confidence intervals overlap with those of scenarios B and C. This suggests important interaction effects between warnings and traffic level.

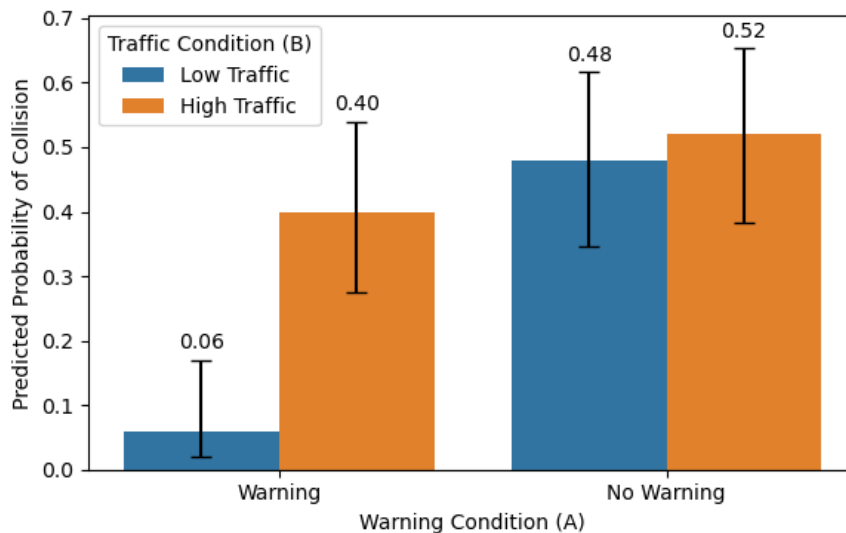


Figure 24: Predicted Probability of Collision per Treatment.

Similarly, Figure 25 shows the predicted count of hard-braking events for each treatment, showing a progressive increase across treatments. In this case, the traffic condition factor seems to have the most significant impact on the frequency of near-miss events, with similar values estimated when warnings were active (A: 3.4, C: 4.07 ~ 4 hard-braking events per trial) and inactive (B: 5.58, D: 5.88 ~ 6 hard-braking events per trial).

The number of driver takeovers, however, presented a slightly different trend (Figure 26). While traffic conditions also appear to have the greatest effect on whether the subjects took over control of the vehicle, the number increased in scenarios where the warnings were active (A, C). The number of takeovers and hard-

braking events was similar in scenario A (between 3.36-3.4 per trial), while for all the other scenarios, there were more driver takeovers than hard-braking events (by at least one event).

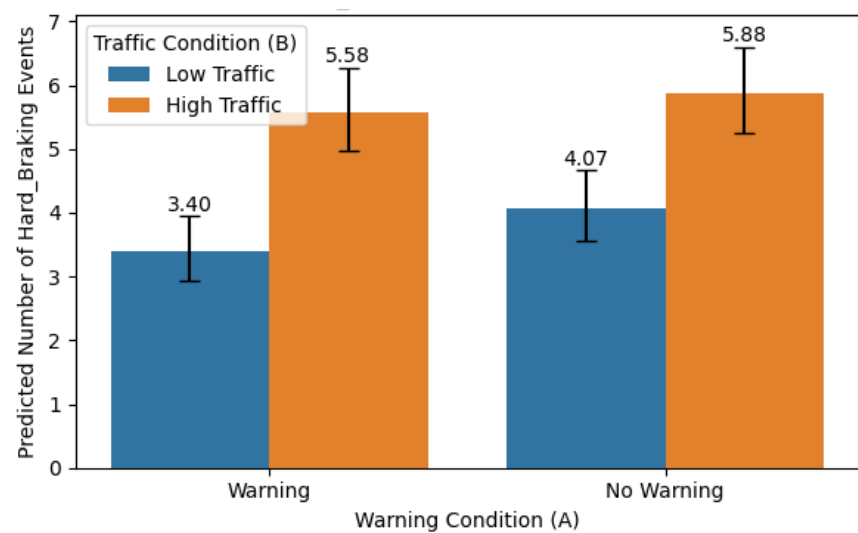


Figure 25: Predicted Number of Hard-Braking Events per Treatment.

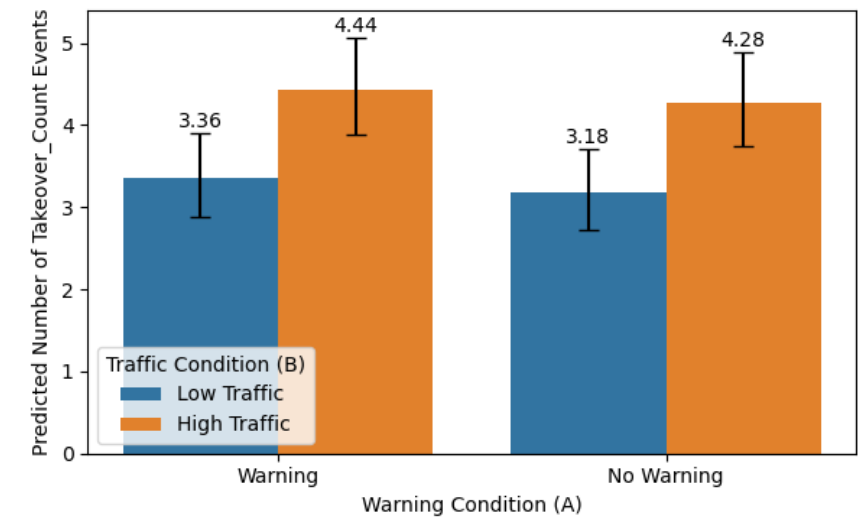


Figure 26: Predicted Number of Driver Takeovers per Treatment.

Tailgating Events (Figure 27) were significantly less frequent in scenarios A and B (means = 0.38, 0.62), but dramatically increased under treatments C and D, reaching a mean of 3.00 and 3.54 events, respectively. The number of driver takeovers is much higher than the number of tailgating events after the ADS's induced failure, suggesting that the following distance under nominal conditions was not one that subjects were comfortable with. The frequencies of driver takeovers and tailgating would-be conflicts are more similar under high traffic conditions than in lower traffic conditions. As in the case of hard-braking and driver takeover events, this may suggest the impact of the higher traffic conditions.

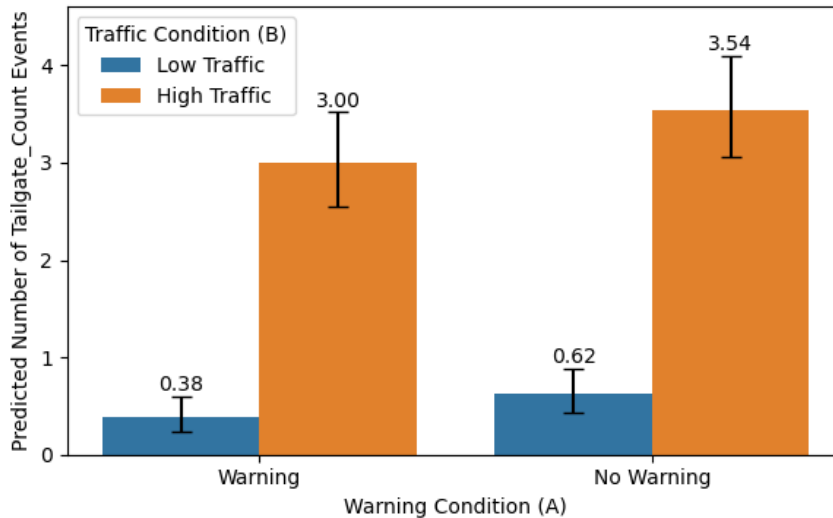


Figure 27: Predicted Number of Tailgating Events per Treatment.

To test the statistical significance of these differences, the main and interaction effects were formally assessed. Table 9 provides a summary of the p-values reported for each factor (A, B) and interaction effects, where values below the 5% significance levels are highlighted. These values confirm the trends identified in Table 18 (Appendix).

Table 9: Main and Interaction Effect Analysis for Categorical Variables.

Variable	Type	Intercept	Factor A	Factor B	Interaction AxB
Collision Probability	Binary	3.83E-06	5.08E-05	3.92E-04	4.71E-03
Hard-Braking Events	Count	2.58E-57	8.26E-02	3.51E-07	3.37E-01
Driver Takeovers	Count	1.32E-55	6.19E-01	6.42E-03	9.00E-01
Tailgating Events	Count	2.47E-05	9.29E-02	2.16E-17	2.99E-01

In summary:

- The difference in the probability of collision is statistically significant between treatment A vs B, C, and D. However, no statistically significant differences were found between B, C, and D.
- The number of hard-braking and tailgating events was statistically different when the traffic level was different, i.e., A vs. C, A vs. D, B vs. C, and B vs. D. However, no differences were identified when the traffic conditions were the same, independent of whether driver warnings were active or not (A vs. B, C vs. D).
- Driver takeovers were statistically different between scenarios A vs. C, B vs. C, and B vs. D. While the difference between treatment A vs. D is moderately significant (p-value ~ 0.11), there is no clear trend between warning and traffic level impacts.

Building Driver-ADS PSF Model

The next stage of this focused on leveraging experimental results to refine the selection of PSFs and the structure of the BBN model.

As we did for the main and interaction effect analysis, we used correlation analysis (Pearson and Spearman) to determine the relationships between different dependent PSFs to justify or refine the structure presented in Figure 3. This was performed in different stages to: (1) analyze the driver's profile, (2) analyze the driver's perception of the test scenarios, (3) connect the driver profile and perception layer, (4) analyze the scenario data extracted from the simulation, (5) connect the driver's perception to the simulation layer, and finally (5) connect the PSFs to the corresponding safety outcomes. Based on analysis per treatment and aggregated by subject, the most statistically significant results were incorporated into the BBN models' structure.

For the quantification of the BBN model, we estimated conditional probabilities (CPTs) based on the relative distribution of each participant's response and performance with respect to the entire population sample. Depending on the PSF node, two or three states were defined based on specific thresholds. For instance, the node Perceived Severity was defined as a three-state node: "high," "medium," and "low" to represent values above the 75% quartile ($>Q3$), around the median ($Q1 < Q2 < Q3$), and below the 25% quartile ($<Q1$), respectively. Other binary nodes were selected to compare the driver's performance to that of the ADS, as in the case of the Takeover Time (TOT), where the states were defined as "driver-initiated takeover" if $TOT < 0$ and "system-initiated takeover" if $TOT > 0$. The structure and CPT values were iteratively assessed and compared with literature, the theoretical model, and observational findings from the experiments to determine the final BBN model. Table 19 (Appendix) presents the proposed measures for each factor based on the model introduced in Figure 3.

In summary, Figure 28 provides an overview of the data analysis and model construction pipeline.

- Step 1, the data collection, focused on three types of sources: questionnaires, simulations, and recordings. The latter, including the study of reaction times and the codification of verbal survey answers, were not codified, but serve as important sources for future research. Questionnaires and Simulation data included the initial, test, and final survey, as well as extracted variables from vehicle dynamics, control transitions, and incident events. These are divided into categorical and continuous variables, as applicable.
- Step 2 focused on estimating the significance of the main and interaction effects over these variables, according to appropriate statistical tests. In addition, the longitudinal trust in driving automation analysis was also performed.
- Step 3 combined these variables to construct PSF and KPI values. As shown in the figure, the initial survey was used to quantify PSFs related to the driver profile, i.e., Agent PSFs, while the test survey and simulation data were used to quantify the Scenario PSF and related KPI.
- Step 4 repeated the main and interaction effect analysis to ensure that the statistical relations were maintained through the constructed PSF and KPI.
- Step 5 consisted of refining the model structure through correlation analysis and calibrating each node based on the discretized states and the cross-tabulation of the PSF distributions.

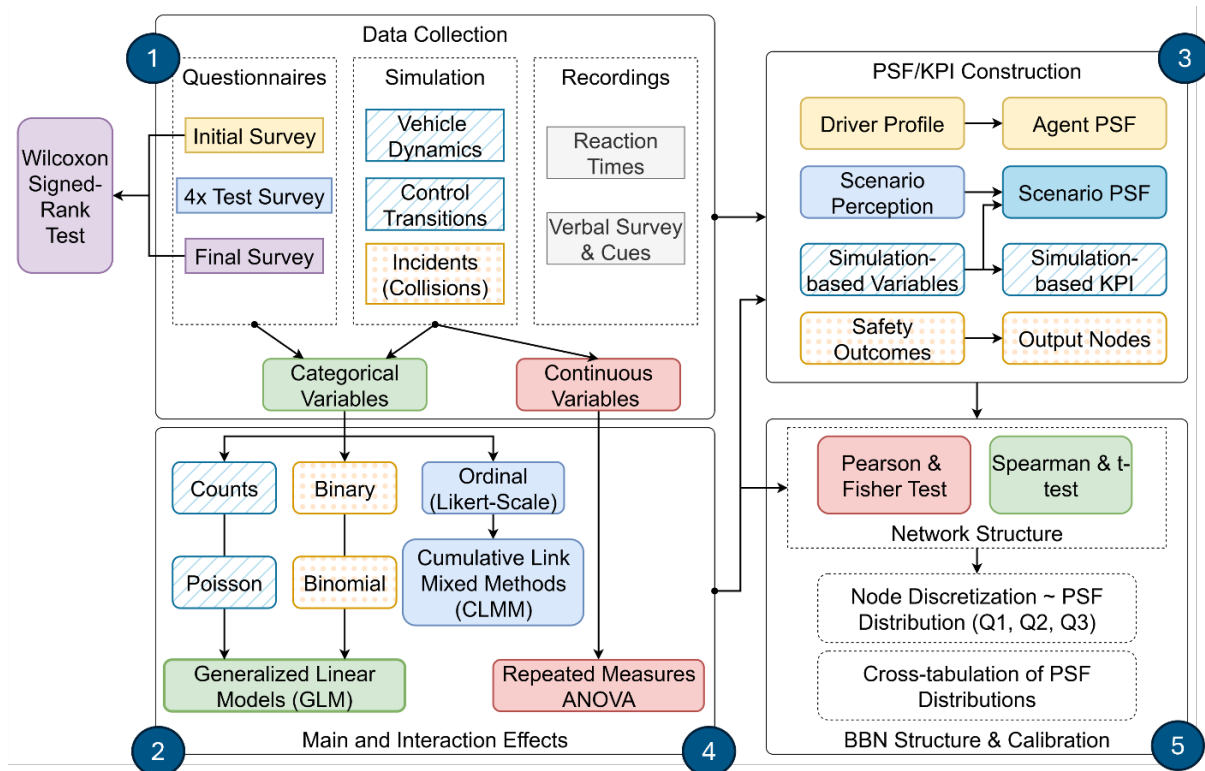


Figure 28: Overview of Data Analysis and Model Construction Pipeline.

The BBN model was implemented in the academic version of GeNIe 5.0 (Graphical Network Interface) Modeler⁴, a tool developed to model AI and ML applications with Bayesian networks and other types of graphical probabilistic models. Built on SMILE (Structural Modeling, Inference, and Learning Engine), GeNIe is a widely-used modeling environment that allows building BBNs and other graphical probabilistic and decision-theoretic models (Guo and Utne 2022; Mamdikar, Kumar, and Singh 2022; Zarei, Khan, and Abbassi 2022; Chaal et al. 2022; Basnet et al. 2023).

Final results for Steps 3-4 will be published at a later date. Further details are available in the report by Correa Jullian (Correa Jullian 2025).

⁴ Bayes Fusion, LLC. GeNIe Modeler (Version 5.0): Complete Modeling Freedom. Available [online] at: <https://www.bayesfusion.com/genie/>

Concluding Remarks

Adopting a team-based perspective for human-ADS relationships introduces a wide breadth of tools, literature, and technical language that enhance system interpretability under shared-control paradigms. This approach highlights the impact of ADS design on the overall system's safety and the value of human interventions, as opposed to characterizing operators only as technology users (Walliser et al. 2019).

This work extends principles from HATs to the domain of HRA, aiming to quantify how driver behavior and system design jointly affect operational safety in Level 3 ADS scenarios. The proposed model is grounded in causal reasoning and captures the influence of key PSFs across multiple layers of interaction, from driver profile characteristics to safety outcomes. Extending PSFs to account for the role of automation is a task currently addressed in many industries (Cheng, Ramos, and Mosleh 2024; Veitch et al. 2024). However, most of these efforts have focused on the one-sided effect of automation on human operators. In contrast, this work seeks to explore the applicability of PSFs at the system or HAT level, constructing a team model from the combined interdependent effects of individual, scenario, organization, team, and system PSFs. Aimed at providing input to qualitative and quantitative risk assessments, these concepts can then be used to develop risk reduction measures needed to support the safe operation of ADSs, such as training, industry standards, and best practices.

The model is informed by data collected through human-subject experiments conducted with a medium-fidelity driving simulator. The experiments focused on takeover scenarios triggered by silent failures of the ADS, under warning availability (Factor A: Warning Availability) and varying traffic conditions (Factor B: Traffic Level). Both objective driving performance data (vehicle kinematics, reaction times, and incident rates) and subjective survey responses (trust, perceived urgency, and perceived responsibility) were collected and analyzed using appropriate statistical tests to determine the impact of the independent variables. From these data, agent, team, and scenario-level PSFs were derived to inform the structure and quantification of a BBN model. The final calibration and findings of this BBN HAT model will be published at a later date.

Initial Findings

Across 72 participants (50 valid sessions after quality filtering), the study combined surveys with simulation telemetry to characterize driver profiles, scenario perception, workload, and behavior under four treatments varying warning availability (A) and traffic complexity (B). Participants were young adult drivers with moderate driving exposure and broad familiarity with common ADAS features. Baseline attitudes showed higher enjoyment of manual driving than ADS, and conditional trust in automation (willing to delegate under boredom/fatigue but cautious in complex or adverse conditions). In-scenario perceptions indicated high transparency of ADS decisions yet frequent disagreement with its behavior and a post-exposure shift toward

lower trust—most pronounced in high-traffic scenarios. Workload ratings (NASA-TLX) were dominated by mental demand (over 70% reporting high levels), with temporal demand and frustration increasing in high traffic; scenario A (warning, low traffic) consistently produced the lowest workload.

Behavioral and performance outcomes showed a clear protective effect of warnings and a strong risk-elevating effect of traffic complexity. Warnings increased mean time-to-collision (TTC), reduced collision likelihood, and improved driver responsiveness (faster reactions to warnings and tailgating), while not reliably changing the likelihood/duration of driver intervention. Traffic complexity increased human interventions, reduced TTC (mean/SD/min), and elevated steering variability, hard-braking, and tailgating counts. Categorical analyses and repeated-measures tests confirmed: (i) collision probability was significantly lower in treatment A (0.06; 95% CI 0.02–0.17) and highest in D (0.52; 0.38–0.65), with significant A×B interaction; (ii) hard-braking and tailgating were driven primarily by traffic level (significant for A vs. C/D and B vs. C/D, but not within the same traffic condition regardless of warnings); and (iii) driver takeovers increased with traffic and were mixed with respect to warnings. Notably, warnings in low-complexity traffic coincided with higher mean speeds (a non-additive interaction), suggesting confidence or later takeovers. Overall, warnings conferred measurable safety benefits, but high traffic complexity dominated risk, highlighting the need to calibrate warning design, timing, and the human-system interface to mitigate cognitive load and expectation mismatches in shared-autonomy regimes.

Next Steps

The next stage of this research will focus on leveraging experimental results to refine the selection of PSFs and the structure of the BBN model. Future implementations can adapt the experimental setup to:

- Perform longer experimental sessions to capture data related to long-haul commercial operations;
- Study V2X-enabled driver warnings for collision avoidance and behavior changes, expanding upon the case of the end-user driver; and
- Study CDA-related applications supporting remote driver assistance tasks, expanding upon remote operator supervising multiple vehicles.

Appendix

Experimental Design

Recruitment and Eligibility

Participants were recruited through flyers and social media posts directing them to complete a brief survey to determine eligibility (adults between 18 and 55 years of age, with a valid driver's license for more than 2 years). This online survey contained the study information sheet, eligibility questions, and a scheduling application. If eligible, the participants had to select an available time slot for the in-person appointment. Due to time constraints, participants were recruited on the UCLA campus.

Experiment Protocol

Participants were required to attend an in-person appointment with an approximate duration of 90 minutes and consisted of the following main stages (Figure 29): (1) an initial questionnaire to characterize each participant, (2) a training session to introduce the simulator and purpose of the study, (3) a test session consisting of four scenarios, and (4) a final questionnaire leading to the end of the study (5).

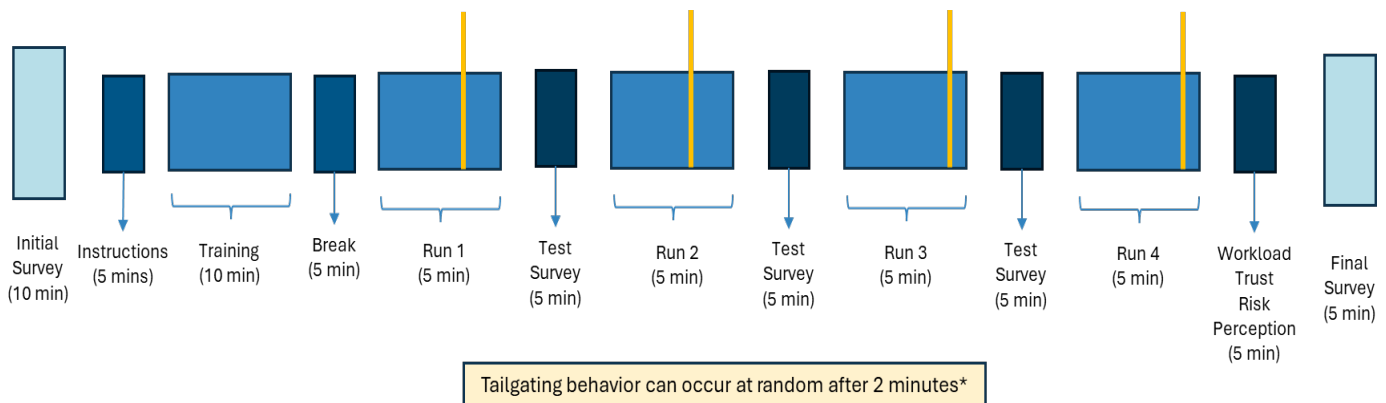


Figure 29: Experimental Session Design.

Each stage consisted of the following:

- 1) Initial questionnaire. Participants completed a questionnaire about their personal driving experience, habits, and opinions on automated driving technology. No personally identifying information was collected.
- 2) Simulator training. Participants received instructions on how to operate the driving simulator and the characteristics of the simulated vehicle's capabilities. The experiment stages were briefly explained to

the participants. This included a description of the levels of driving automation, focusing on the differences between levels 2, 3, and 4.

- 3) Testing scenarios. Participants were then given instructions to complete short driving scenarios where they could switch between manual and automated driving controls.
- 4) Final questionnaire. After the test session, participants are asked to complete one final short questionnaire, repeating a section of the initial survey.
- 5) The session ends. The subjects are thanked for their participation and are provided with their compensation (\$15 Amazon gift card).

Training Scenarios

Participants went through a short training session to become familiar with the simulator's visuals and controls. The monitors showed a virtual driving environment representing a small town with residential and highway sections (CARLA Town 06), with traffic signs and other vehicles following independent trajectories. The monitors presented an approximate view from the driver's seat, indicating the automation state (i.e., OpenCDA or Manual) and the vehicle's speedometer. Participants triggered the start of each scenario by pressing space on the keyboard (Figure 30). All scenarios were initiated in automated driving mode, and participants were instructed to brake or use the steering wheel to take control when desired. If any collision occurred during the training scenarios, the researcher manually terminated the simulation.

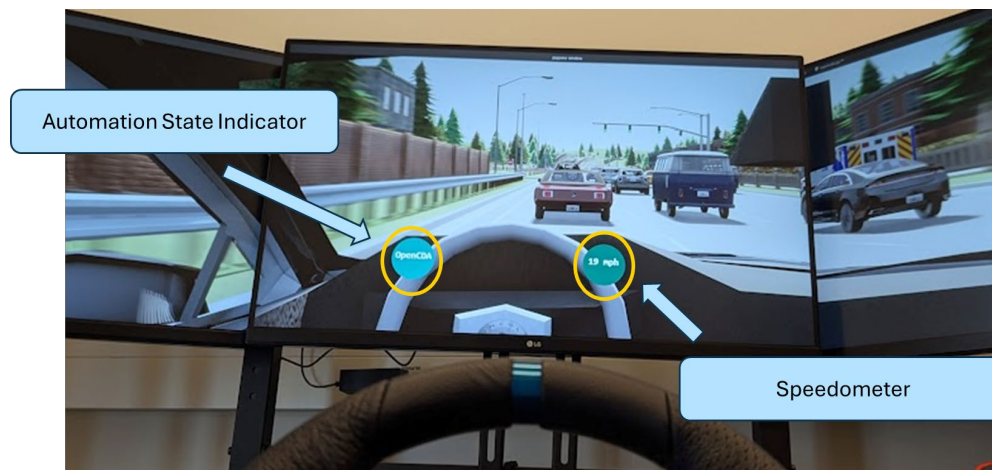


Figure 30: Experimental Setup - View from the Driver's Seat.

This training session was divided into three short tasks:

- Training scenario 1: Manual control of the vehicle. The participants are asked to disengage the automated driving system by applying the brake or steering and then continue to drive as they would in a typical vehicle down a road, interacting with other vehicles and traffic signs along three intersections. The

participant was instructed to follow the road, which included two curved sections, without turning onto the side streets, until a sign appeared on the screen indicating the route had been completed.

- Training scenario 2: Observe automated driving mode. Starting from the same position as training scenario 1, the participants were asked to initiate the scenario and watch how the vehicle performed the same route while maintaining a minimum safety distance of $TTC = 7$ seconds. Participants were reminded that while they were free to do other activities, they were encouraged to monitor the vehicle's actions at this time to familiarize themselves with its expected behavior.
- Training scenario 3: Respond to a takeover request. Beginning at the starting position, the participants were asked to initiate the scenario as in the previous scenario. However, participants were instructed that at some point in the track, an example of a takeover request will be shown on the screen if the TTC goes below the threshold of 5 seconds. This is a visual alert that appears on the central monitor with the following text: "Warning: Please Brake" (Figure 31a). Participants are instructed to use the steering wheel and/or pedals to regain control of the vehicle and bring the vehicle to a safe stop. If a collision occurred, a visual alert appeared with the following text: "Warning: Collision Occurred" (Figure 31b) and the simulation was terminated.

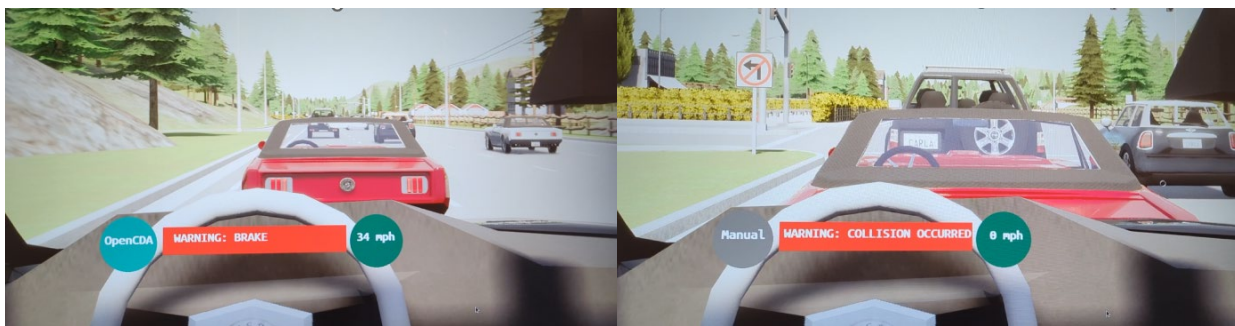


Figure 31: Example of Driver Interface, (a) Warning Alert and (b) Collision Alert.

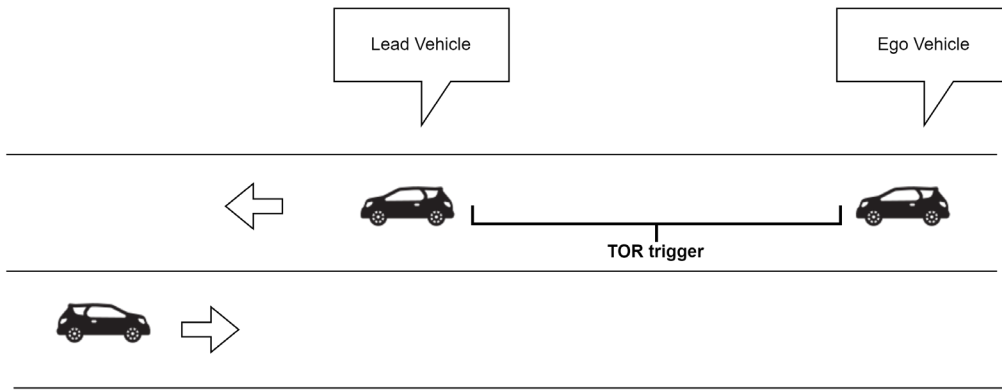
Testing Scenarios

The testing session consisted of the following:

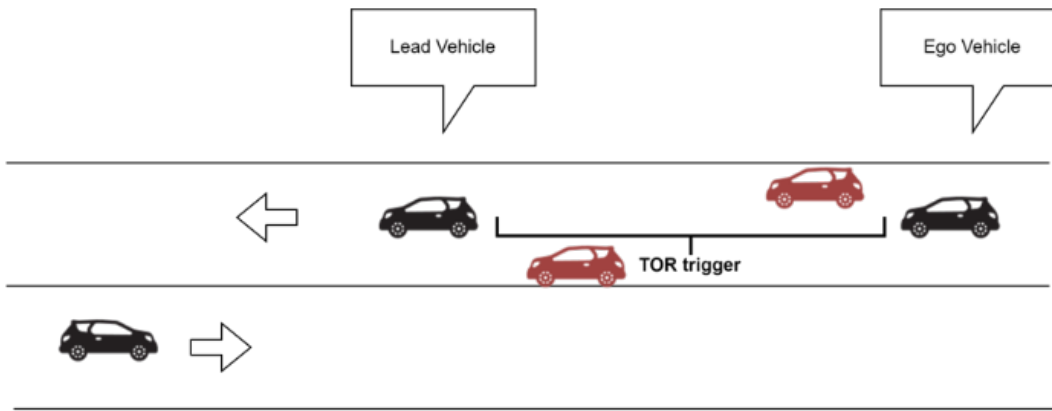
- Four trials per session with breaks of 5 minutes between trials (while test surveys were completed).
- Each trial started with the vehicle driving in automated mode and participants were told they could take control if desired or when requested by the system. Participants could take over control at any prior point in time if desired, although they were instructed that the automated driving mode was preferred.
- Each trial assessed two independent variables: the availability of early warning (A) and scenario complexity (B). Two levels were selected for the early warnings (A0 available, A1 not available), and two levels of complexity were selected to represent different driving conditions (B0 low traffic, B1 high traffic). Each trial covers a combination of these independent variables (B0-A0, B0-A1, B1-A0, B1-A1). Figure 32 depicts the two complexity levels (B0, B1).

- Each trial had a duration of approximately 5 minutes. The driving scenario consisted of the ego-vehicle (the car that can be controlled by the participant) following a lead vehicle and maintaining a safe distance at an approximate speed of 35-38 mph (56-61 km/h). This safety distance is defined as a time to collision (TTC) of more than 7 seconds.
- At a random time between minute [2-4], the safety distance of the ego vehicle was artificially reduced to 3.5 seconds. Therefore, when the lead vehicle reduced its speed (e.g., at an intersection), it would invade the safety envelope of the ego vehicle. When the TTC was below 5 seconds, this triggered an alarm under the A0 condition (but not the A1 condition), defined as a Takeover Request (TOR), shown to the participant.
- If the participant took control of the vehicle, they were expected to slow or to stop the ego-vehicle in a safe spot (as instructed earlier). If no warning was given (A1) or the participant did not take over control, the ego vehicle would attempt to maintain the 3.5-second safety distance. The scenario was terminated when either the ego vehicle collided with an object after the driver took control or after the ego vehicle completed the track (manually, or after re-engaging the system) without a collision.
- After the scenario ended, the participants were given a 5-minute break. During this break, they were asked to complete a questionnaire about their experience and follow-up, open-ended questions based on their actions during the experiment
- The simulation was reset, and the next trial with the corresponding treatment was started. This process was until four trials were completed.

Note: If participants expressed any sign of motion sickness, the trials were terminated early. While participants were encouraged to express this during the tests, a survey question was also included after each trial to verify their state.



(a) B0: Low Traffic



(b) B1: High Traffic

Figure 32: Driving Scenario (a) B0: Simple (low traffic) and (b) B1: Complex (high traffic, additional vehicles represented in red).

Questionnaires

Initial Survey

This survey covered questions related to personal driving experience, habits, and opinions on automated driving technology (Table 10).

- Questions 1.1-1.6 were adapted from P. Bazilinskyy et al. (2018) (Bazilinskyy et al. 2018).
- Question 1.7 was created for this study.
- Question 1.8 items were adapted from Ernst and Reinelt (2017) (Ernst and Reinelt 2017).

- Question 1.9 items were adapted from the short version of the Driver Behavior Questionnaire presented in De Winter (2013) (De Winter 2013), originally developed by J. Reason et al. (1990) (Reason et al. 1990).
- Question 1.10 items were adapted from Ledesma et al. (2015) (Ledesma et al. 2015).
- Question 1.11 items were adapted from Choi and Ji (2015) (Choi and Ji 2015).
- Question 1.12 items were adapted from the Trust in Automated Driving (TiAD) Questionnaire used by Payre, Cestac, and Delhomme (2014, 2016) (Payre, Cestac, and Delhomme 2016; 2014) and Manchon, Bueno, and Navarro (2022) (Manchon, Bueno, and Navarro 2022).

Table 10: Subjective Measures and Factors Collected from the Initial Questionnaire.

Num.	Question	Outcome	Factor	Source
1.1	Gender: <i>Please indicate your gender</i>	Gender [category]	Bias	P. Bazilinskyy et al. (2018)
1.2	Age: <i>What is your age?</i>	Age [value]	Bias	
1.3	Driving Experience: <i>At what age did you obtain your first license for driving a car or motorcycle?</i>	Years [value]	Experience	
1.4	Primary mode of transport: <i>What is your primary mode of transportation?</i>	Mode [category]	Experience	
1.5	Driving Frequency: <i>On average, how often did you drive a vehicle in the last 12 months?</i>	Frequency [category]	Experience	
1.6	Driving Mileage: <i>About how many miles did you drive in the last 12 months?</i>	Mileage [category]	Experience	
1.7	ADAS Experience: <i>Have you had any experience with the following features?</i>	Count [yes/no]	Familiarity	Original
1.8.1- 1.8.3	Driving Enjoyment: 3 items <i>It is enjoyable to drive a car yourself</i> <i>It is exciting to drive a car yourself</i> <i>It is fun to drive a car yourself</i>	Scale Average [1-5]: How much do you agree with the following statements?	Bias	Ernst and Reinelt (2017)

Num.	Question	Outcome	Factor	Source
1.8.4- 1.8.6	Preconceived ADS Enjoyment: 3 items • <i>Using an autonomous car would be enjoyable</i> • <i>Using an autonomous car would be fun</i> • <i>Using an autonomous car would be exciting</i>	Scale Average [1-5]: How much do you agree with the following statements?	Bias	
1.9.1- 1.9.2	DBQ Violations: 2 items • <i>Disregard the speed limit on a residential road or motorway</i> • <i>Drive so close to the car in front that it would be difficult to stop in an emergency</i>	Scale Average [1-5]: How frequently do you think you do these actions?	Risk Tolerance	De Winter (2013)
1.9.3- 1.9.4	DBQ Errors: 2 items • <i>Get into the wrong lane when approaching a roundabout or junction</i> • <i>Incorrect steering so that you hit the curb</i>	Scale Average [1-5]: How frequently do you think you do these actions?	Risk Tolerance	De Winter (2013)
1.10.1- 1.10.3	Distracted Driving: 3 items • <i>At an intersection, I pay attention to the vehicles coming the other way.</i> • <i>At a street corner, I realize that a pedestrian is crossing the street.</i> • <i>I realize that I failed to see a traffic light simply because I was not paying attention.</i>	Scale Average [1-5]: How frequently do you think you do these actions?	Risk Tolerance	Ledesma (2015)

Num.	Question	Outcome	Factor	Source
1.11.1- 1.11.3	Perceived Responsibility: 3 items • <i>Driving without accidents is mainly a matter of luck.</i> • <i>Accidents usually happen because of unexpected events that occur during driving.</i> • <i>It is difficult to prevent accidents when the driving conditions are difficult, such as darkness, rain, a narrow road with many turns.</i>	Scale Average [1-5]: How much do you agree with the following statements?	Responsibility	Choi and Ji (2015)

Num.	Question	Outcome	Factor	Source
1.12.1- 1.12.12	Trust in Automated Driving: 12 items • <i>I would feel safe in an automated vehicle.</i> • <i>The automated driving system provides me with more safety compared to manual driving.</i> • <i>I would rather keep manual control of my vehicle than delegate it to the automated driving system on every occasion.</i> • <i>I would trust the automated driving system decisions.</i> • <i>I would trust the automated driving system's capacities to manage complex driving situations.</i> • <i>If the weather conditions were bad (e.g., fog, glare, rain), I would delegate the driving task to the automated driving system.</i> • <i>Rather than monitoring the driving environment, I could focus on other activities confidently.</i> • <i>If driving were boring for me, I would rather delegate it to the automated driving system than do it myself.</i> • <i>I would delegate the driving to the automated driving system if I was tired.</i> • <i>If I had passengers in my automated car, I would rather drive by myself than delegate it to the automated driving system.</i> • <i>I would like to recover control from the automated pilot if I did not like the way it drives.</i> • <i>Globally, I trust my capacity to resume control if needed.</i>	Scale Average [1-7]; How much do you agree with the following statements?	Trust	Manchon, Bueno, and Navarro (2022), Payre, Cestac, and Delhomme (2014, 2016)

Test Survey

This survey covered the participants' experience in the simulator. This survey was administered four times, after each scenario. This survey consisted of a section of multiple-choice questions and then verbal, open-ended questions. The multiple-choice section of this survey covered the participant's perception of the scenario, including their assessment of the ADS's behavior and workload (Table 11).

- Questions 2.1-2.3 items were adapted from Muslim, Leung, and Itoh (2022) (Muslim, Kiu Leung, and Itoh 2022).
- Question 2.4-2.9 items were created for this study.
- Questions 2.10-2.11 items were adapted from the NASA TLX Questionnaire to assess workload by Hart and Staveland (1988) (Hart and Staveland 1988; Hart 2006).

The open-ended questions section of this survey covers the participant's perception of the scenario, including their assessment of the ADS's behavior, their attention, and perceived workload (Table 12). The questions were based on the occurrence of certain events during the scenario. These questions were created specifically for this study.

Table 11: Subjective Measures and Factors Collected from Test Questionnaire.

Num.	Question	Outcome	Factor	Source
2.1	Understand the ADS: <i>You could understand the system?</i>	Scale Average [1-7]: To what extent do you think...?	Trust	Muslim, Leung, and Itoh (2022)
2.2	Agree with the ADS: <i>You agree with the system's decision?</i>	Scale Average [1-7]: To what extent do you think...?	Trust	Muslim, Leung, and Itoh (2022)
2.3	Trust the system: <i>The system is trustworthy?</i>	Scale Average [1-7]: To what extent do you think...?	Trust	Muslim, Leung, and Itoh (2022)
2.4	Motion sickness: <i>You experienced any motion sickness?</i>	Scale Average [1-7]: To what extent do you think...?	n/a	Original

Num.	Question	Outcome	Factor	Source
2.5	Urgency: <i>The situation was urgent?</i>	Scale Average [1-7]: To what extent do you think...?	Perceived Urgency	Original
2.6	Severity: <i>The situation would lead to severe consequences?</i>	Scale Average [1-7]: To what extent do you think...?	Perceived Severity	Original
2.7	Risk Perception <i>The system performed a risky action?</i>	Scale Average [1-7]: To what extent do you think...?	Perceived Risk	Original
2.8	Perceived Responsibility <i>You were responsible for the situation?</i>	Scale Average [1-7]: To what extent do you think...?	Perceived Responsibility	Original
2.9	Perceived Complexity <i>You could have performed the action in a safer way?</i>	Scale Average [1-7]: To what extent do you think...?	Perceived Responsibility	Original
2.10.1- 2.10.6	Workload 1. Mental Demand 2. Physical demand 3. Temporal demand 4. Effort 5. Performance 6. Frustration level	Scale Average [1-9]: Please rate these workload dimensions from 1 (Lowest) to 9 (Highest).	Perceived Task Complexity	Hart and Staveland (1988)
2.11	Workload: Pairwise comparison of workload dimensions.	Binary: What was more important to your experience?	Perceived Task Complexity	Hart and Staveland (1988)

Table 12: Subjective Measures and Factors Collected from Test Questionnaire – Open Questions.

Participant	Open Question	Factor
All	What was the hardest task?	Perceived Task Complexity
	How engaged were you?	Alertness
	What made you pay/not pay attention?	Attention
	What made you trust/not trust the system?	Trust
Takeover w/TOR (A1)	Why did you take over?	Engagement/Strategy
	What do you think was the cause of the takeover request?	Attention
Takeover w/o TOR (A2)	Why did you take over?	Engagement/Strategy
	Do you think the system failed?	Attention/Strategy
No Takeover w/TOR (A1)	Why didn't you take over?	Engagement/Strategy
	What would you have done differently?	Perceived responsibility/Strategy
No Takeover w/o TOR (A2)	Why didn't you take over?	Engagement/Strategy
	Do you think the system should have issued a TOR?	Perceived severity/Strategy
Early Takeover	Why did you take over?	Engagement/Strategy
	Do you think the system should have issued a TOR?	System response
If the participant collided while in control of the vehicle.	What was the main issue that led you to the crash?	Compliance with procedures/ Manual Operation
	What would you have done differently?	Perceived responsibility/Strategy

Final Survey

This survey covers questions related to the participants' opinions on automated driving technology, after they have completed the four experiment scenarios. These questions consist of the items 1.12.1-1.12.12 presented in Table 10, referred to as items 3.12.1-3.12.12.

Data Processing

Driver Profile

Table 13: Driver Profile Response Mapping.

Question	Variable	Numerical Mapping
1.1	Gender (Numeric)	Male=-1, Female=1, NA=0
1.2	Age	n/a
1.3	Driver's License (Exp_Lic)	License Age = Age - Year license obtained
1.4	Mode (Numeric)	Private vehicle = 6, Public Transportation = 5, Motorcycle = 4, Cycling = 3, E-Scooter = 2, Walking = 1, NA = 0
1.5	Driving Freq. (Numeric)	Range 1-6 (Never - Every day), NA=0
1.6	Driving Freq. (Miles)	Range 1-9 (0 miles - more than 50,000 miles), NA = 0
1.7	ADAS Count	Range 1-7 per subject.
1.8.1	Driving Enjoy	5-point Likert Scale, 1: "Strongly Disagree", 2: "Disagree", 3: "Neutral", 4: "Agree", 5: "Strongly Agree"
1.8.2	Driving Exciting	
1.8.3	Driving Fun	
1.8.4	ADS Enjoy	
1.8.5	ADS Exciting	
1.8.6	ADS Fun	
1.9.1	Quality Speed	

Question	Variable	Numerical Mapping
1.9.2	Quality Tailgate	5-point Likert Scale, 1: “Never”, 2: “Rarely”, 3: “Sometimes”, 4: “Very Often”, 5: “Always”
1.9.3	Quality Lanes	
1.9.4	Quality Curb	
1.10.1	Quality Crossing	
1.10.2	Quality Vulnerable Road User	
1.10.3	Quality Signs	
1.11.1	Accidents Luck	5-point Likert Scale, 1: “Strongly Disagree”, 2: “Disagree”, 3: “Neutral”, 4: “Agree”, 5: “Strongly Agree”
1.11.2	Accidents Unexpected	
1.11.3	Accidents Complex	
1.12.1-3.12.1	Trust Safe	7-point Likert Scale, 1: “Strongly Disagree”, 2: “Disagree”, 3: “Slightly Disagree”, 4: “Neutral”, 5: “Slightly Agree”, 6: “Agree”, 7: “Strongly Agree”
1.12.2-3.12.2	Trust ADS	
1.12.3-3.12.3	Trust Manual	
1.12.4-3.12.4	Trust ADS Decision	
1.12.5-3.12.5	Trust ADS Capacity	
1.12.6-3.12.6	Trust Weather	
1.12.7-3.12.7	Trust Attention	
1.12.8-3.12.8	Trust Boring	
1.12.9-3.12.9	Trust Tired	
1.12.10-3.12.10	Trust Passengers	

Question	Variable	Numerical Mapping
1.12.11- 3.12.11	Trust TOR	
1.12.12- 3.12.12	Trust TOQ	

Test Scenario Data

Table 14: Raw Data Retrieved from the Simulation.

Variable	Values	Comment
Subject	P<x>	Added to the files when processing to identify the subject number.
Scenario Name	A-D	Added to the files when processing to identify the scenario type (A: Warning, Low Traffic, B: No Warning, Low traffic, C: Warning, High Traffic, and D: No Warning, High Traffic)
Timestep	Numeric	The simulation's timestep, each interval equal to 0.05 seconds, i.e., 20 Hz sampling rate.
Sim. End	True/False	If the simulation had ended at that timestamp. The value would be "False" unless the vehicle reached the endpoint on the track, i.e., a collision occurred.
Collision	True/False	If the vehicle had collided at that timestamp. The value would be "True" since a collision occurred until the scenario was terminated.
TTC	0-1000	Value estimated by OpenCDA. If above a certain threshold, the default value is 1000.
Warning	True/False	If a warning was present at that timestamp. The value would be "True" only for the duration of the warning, based on the TTC condition and a default inertia.
Lane Inv.	NAN/Text	If the vehicle was involved in a lane violation at that timestamp. If true, the value would be "Crossed line 'Broken'", else, it returns an empty cell.

Variable	Values	Comment
Brake Contr.	Numeric	The value of the brake input at that timestamp. This only records the driver's physical input, not the ADS control.
Handbrake Contr.	True/False	Whether the handbrake was activated at that timestamp.
Driver TO	True/False	Whether the driver was in control of the vehicle at the timestamp. Default is "False", as the simulation initiates under ADS control.
Reverse Contr.	True/False	Whether the vehicle was placed in 'reverse' gear at that timestamp.
Steer Contr.	Numeric	The value of the steering input at that timestamp. This records both the driver's physical input and the ADS control.
Throttle Contr.	Numeric	The value of the throttle input at that timestamp. This only records the driver's physical input, not the ADS control.
Veh. Speed	Numeric	The vehicle's speed at that timestamp. This records both the driver and the ADS control.
X Pos.	Coordinates	The coordinates (X, Y, Z) for the vehicle's center.
Y Pos.	Coordinates	The coordinates (X, Y, Z) for the vehicle's center.
Z Pos.	Coordinates	The coordinates (X, Y, Z) for the vehicle's center.
Roll Deg.	Angle	The rotation angle (roll, yaw, pitch) for the vehicle's center.
Yaw Deg.	Angle	The rotation angle (roll, yaw, pitch) for the vehicle's center.
Pitch Deg.	Angle	The rotation angle (roll, yaw, pitch) for the vehicle's center.

Table 15: Scenario Perception Survey Response Mapping

Question	Variable	Tag	Numerical Mapping
2.1	ADS - Understand	L7_understand	7-point Likert Scale, 1: “Strongly Disagree”, 2: “Disagree”, 3: “Slightly Disagree”, 4: “Neutral”, 5: “Slightly Agree”, 6: “Agree”, 7: “Strongly Agree”
2.2	ADS - Agree	L7_agree	
2.3	ADS - Trust	L7_trust	
2.4	Motion Sickness	L7_motionsickness	
2.5	Situation Urgency	L7_urgency	
2.6	Situation Severity	L7_severity	
2.7	Risk Awareness	L7_riskaware	
2.8	Responsibility	L7_responsibility	
2.9	Situation Complexity	L7_selfbelief	
2.10.1	Mental Load	L9_mental	9-point Likert Scale, 1: “Very Low”, 2: “Low”, 3: “Moderately Low”, 4: “Slightly Low”, 5: “Neutral”, 6: “Slightly High”, 7: “Moderately High”, 8: “High”, 9: “Very High”
2.10.2	Physical Load	L9_physical	
2.10.3	Temporal Load	L9_temporal	
2.10.4	Effort Load	L9_effort	
2.10.5	Performance Load	L9_performance	
2.10.6	Frustration Load	L9_frustration	

Experimental Results

Scenario Perception

Table 16: Summary of Test Survey Results; Main and Interaction Effects Analysis.

Question	Variable	Max	Min	Mean	Std	Median	25%	75%	p-value: Main Effect A	p-value: Main Effect B	p-value: Interaction Effects
2.1	ADS - Understand	7	3	5.59	1.06	6	5	6.25	7.54E-01	6.18E-01	2.36E-01
2.2	ADS - Agree	5.75	1.75	3.79	0.95	3.75	3.25	4.44	2.04E-08	1.68E-04	5.06E-03
2.3	ADS - Trust	5.75	1	3.44	1.12	3.62	2.5	4.25	6.59E-07	1.41E-03	2.33E-02
2.4	Motion Sickness	7	1	2.19	1.48	1.5	1	3.19	8.54E-01	2.83E-01	5.15E-01
2.5	Situation Urgency	6.75	1.75	5.14	1.05	5.25	4.56	6	4.71E-02	1.39E-03	1.13E-01
2.6	Situation Severity	7	3.5	5.54	0.83	5.62	5	6.19	7.62E-03	4.93E-04	1.61E-02
2.7	Risk Awareness	6.75	2.5	5.06	0.93	5.25	4.5	5.75	2.69E-05	6.63E-04	9.81E-03
2.8	Responsibility	7	1.25	4.23	1.37	4.38	3.06	5.44	4.10E-01	4.83E-01	5.02E-01
2.9	Situation Complexity	7	1	5.16	1.32	5.25	4.75	6	1.31E-05	1.56E-04	2.42E-02
2.10.1	Mental Load	9	2.25	6.12	1.31	6.25	5.56	6.94	1.05E-02	6.39E-06	9.85E-02
2.10.2	Physical Load	7.5	1	3.89	1.5	3.75	3	4.69	2.53E-01	7.00E-03	7.77E-01
2.10.3	Temporal Load	8.25	2	5.16	1.56	5	4.06	6.5	1.60E-02	8.29E-04	1.43E-01
2.10.4	Effort Load	8	3	5.2	1.18	5.38	4.31	6	1.40E-03	9.80E-05	2.84E-02
2.10.5	Performance Load	9	3.25	5.84	1.29	5.75	5	6.69	4.75E-03	1.12E-02	9.68E-02
2.10.6	Frustration Load	7.75	2	4.54	1.61	4.25	3.31	5.94	6.34E-03	2.34E-03	2.59E-01

Question	Variable	Max	Min	Mean	Std	Median	25%	75%	p-value: Main Effect A	p-value: Main Effect B	p-value: Interaction Effects
Legend:		< 5%	< 10%	< 15%	< 20%	< 25%					

Simulation Variables

Table 17: Summary of Simulation Variables, Main and Interaction Effects.

Variable	Definition	Factor A Mean			Factor B Mean			Interaction Effect (S/NS) *
		A0: Warning	A1: No Warning	Effect (S/NS) *	B0: Low Traffic	B1: High Traffic	Effect (S/NS) *	
End Time	The timestamp corresponding to the end of the scenario. If there was a collision, then it is the same as Coll. Time.	4144.64	3911.29	S	3785.68	4270.25	S	NS
Coll. Time	The timestamp corresponding to the collision in the scenario. If no collision occurred, then it is the same as End Time.	4124.43	3866.32	S	3762.68	4228.07	S	NS
Warning Time	The timestamp the first warning was triggered. If no warning was triggered (such as in scenario B, D), then the value is null.	3025.98	0	S	1484.2	1541.78	NS	NS

Variable	Definition	Factor A Mean			Factor B Mean			Interaction Effect (S/NS) *
		A0: Warning	A1: No Warning	Effect (S/NS) *	B0: Low Traffic	B1: High Traffic	Effect (S/NS) *	
Warning Total	The total time that warnings were active during the scenario.	26.34	0	S	7.45	18.89	NS	NS
Driver Time	The timestamp of the first driver takeover. If none, then the value is null.	1331.95	1359.8	NS	1150.03	1541.72	S	NS
Driver Total	The total length that the driver was in control of the vehicle.	1214.51	1105.48	NS	983.04	1336.95	S	NS
Lane Inv. Time	The total time that there was a lane violation.	3636.43	3405.66	S	3349.23	3692.86	S	NS
Lane Inv. Time Driver	The total time that there was a lane violation when the driver was in control of the vehicle.	1210.41	1102.49	NS	981.73	1331.17	S	NS
Veh. Mean Speed	Vehicle mean speed during the scenario.	20.469	19.947	S	21.99	18.426	S	S
Veh. Std Speed	Standard deviation of vehicle speed during the scenario.	15.949	16.095	NS	16.771	15.273	S	NS
Veh. Max Speed	Vehicle max speed during the scenario.	44.914	44.832	NS	45.237	44.509	NS	NS

Variable	Definition	Factor A Mean			Factor B Mean			Interaction Effect (S/NS) *
		A0: Warning	A1: No Warning	Effect (S/NS) *	B0: Low Traffic	B1: High Traffic	Effect (S/NS) *	
Veh. Mean Speed Driver	Vehicle mean speed when the driver is in control of the vehicle.	12.328	9.497	S	10.812	11.013	NS	S
Veh. Std Speed Driver	Standard deviation of vehicle speed when the driver was in control of the vehicle.	12.854	12.194	NS	12.108	12.94	NS	NS
Veh. Max Speed Driver	Vehicle max speed when the driver was in control of the vehicle.	44.167	42.711	NS	43.396	43.482	NS	NS
Steering Mean	Mean steering input.	-8.126	-8.307	NS	-4.93	-11.503	S	NS
Steering STD	Standard deviation of the steering input.	61.163	58.013	NS	43.462	75.714	S	NS
Steering Max	Max steering input.	219.861	167.32	NS	164.635	222.546	NS	NS
Throttle Mean	Mean throttle input.	0.115	0.125	NS	0.108	0.132	NS	S
Throttle STD	Standard deviation of the throttle input.	0.216	0.232	NS	0.213	0.235	NS	NS
Throttle Max	Max throttle input.	0.801	0.78	NS	0.743	0.838	NS	NS
Brake Mean	Mean brake input.	0.17	0.203	NS	0.195	0.178	NS	NS
Brake STD	Standard deviation of the brake input.	0.275	0.31	S	0.302	0.283	NS	NS

Variable	Definition	Factor A Mean			Factor B Mean			Interaction Effect (S/NS) *
		A0: Warning	A1: No Warning	Effect (S/NS) *	B0: Low Traffic	B1: High Traffic	Effect (S/NS) *	
Brake Max	Max brake input.	0.951	0.967	NS	0.961	0.957	NS	NS
TTC Mean	Mean TTC during the scenario.	12.398	11.692	S	12.826	11.264	S	NS
TTC Std	Standard deviation of the TTC during the scenario.	5.736	5.57	NS	6.182	5.124	S	NS
TTC Min	Min TTC during the scenario.	1.212	1.168	NS	1.375	1.005	S	NS
TTC Mean Driver	Mean TTC when the driver was in control of the vehicle.	5.452	5.443	NS	5.528	5.367	NS	NS
TTC Std Driver	Standard deviation of the TTC when the driver was in control of the vehicle.	2.094	1.713	S	1.801	2.006	NS	NS
TTC Min Driver	Min TTC when the driver was in control of the vehicle.	1.55	1.684	NS	1.73	1.504	NS	S
Driver Warning Reaction Time	The reaction time to a warning (negative if driver-initiated)	-29.82	-265.84	S	-141.43	-154.23	NS	NS
Driver Tailgating Reaction Time	The reaction time to a tailgating event (negative if driver-initiated)	-29.82	501.43	S	106.16	365.45	NS	NS

*S: Significant, NS: Not Significant, using p-value <0.05 as reference.

Table 18: Summary of Categorical Variables per Treatment.

Variable	Treatment	Mean	Standard Error	Lower Bound (5%)	Upper Bound (95%)
Collision Probability	A	0.06	0.03	0.02	0.17
	B	0.48	0.07	0.35	0.62
	C	0.40	0.07	0.27	0.54
	D	0.52	0.07	0.38	0.65
Hard-Braking Events	A	3.40	0.26	2.93	3.95
	B	4.07	0.29	3.55	4.67
	C	5.58	0.33	4.96	6.28
	D	5.88	0.34	5.25	6.59
Driver Takeovers	A	3.36	0.26	2.89	3.91
	B	3.18	0.25	2.72	3.71
	C	4.44	0.30	3.89	5.06
	D	4.28	0.29	3.74	4.89
Tailgating Events	A	0.38	0.09	0.24	0.60
	B	0.62	0.11	0.44	0.88
	C	3.00	0.24	2.56	3.52
	D	3.54	0.27	3.06	4.10

Proposed PSF Mapping

Table 19: Proposed PSFs for Driver-ADS Team Model.

Type	PSF	Measure	States
Agent	Experience	Driving experience: Years since the driver's license was obtained, frequency of approximate mileage per year, etc.	Categories “high”, “medium”, “low” defined based on mean and standard deviation of sampled population.
	Risk Tolerance	The attitude of the subject with respect to driving-related risk: risk seeking, neutral, or risk averse based on reported driving behavior.	Categories “high”, “neutral”, “low” defined on Q25, Q50, and Q75 of combined risk tolerance indicator, per state of experience.
	Perceived Familiarity	Experience with ADAS features as proxy to familiarity with tested ADS features.	Categories “above average”, “average”, “below average” defined on Q25, Q50, and Q75 of the number of ADAS features referenced by the sampled population.
	Bias	The subject’s bias towards or against the ADS based on their driving preferences and preconceived ideas on automated driving.	Categories “positive towards ADS”, “neutral towards ADS”, “negative towards ADS” defined on Q25, Q50, and Q75 of combined bias indicator, per state of perceived familiarity.
	Decision Responsibility	The subject’s perception of where responsibility lies in driving incidents (external vs. internal <i>loci</i> of control)	Categories “high control”, “neutral control”, “low control” defined on Q25, Q50, and Q75 of combined responsibility indicator, per state of risk tolerance and bias.
	Level of Trust (Role Awareness)	The subject’s reported preconceived trust in driving automation.	Categories “high”, “neutral”, “low” defined on Q25, Q50, and Q75 of combined trust indicator, per state of experience and bias.
	Alertness / Attention	Not Implemented. Data may be extracted from the video recordings for each test session.	N/A

Type	PSF	Measure	States
Scenario	External environment	Impact of surrounding traffic complexity.	Independent variable, referred to as factor B: level of traffic (B0: low traffic, B1: high traffic). Established as 50%-50% baseline for each state.
	Conditioning events	Impact of silent failures on the driver warning system.	Independent variable, referred to as factor A: warning availability (A0: warning, A1: no warning). Established as 50%-50% baseline for each state.
	Perceived trust	This node represents the evolution of the subject's trust in the ADS during the scenario.	Categories "high", "neutral", "low" defined on Q25, Q50, and Q75 of the combined trust indicator, per state of bias, responsibility, conditioning events (warning availability), and external environment (traffic level).
	Perceived urgency	Subject's perception of the urgency to intervene in the scenario, measured per treatment.	Categories "high", "neutral", "low" defined on Q25, Q50, and Q75 of the urgency indicator, per state of level of trust, risk tolerance, and external environment (traffic level).
	Perceived severity	Subject's perception of the scenario severity, measured per treatment.	Categories "high", "neutral", "low" defined on Q25, Q50, and Q75 of the severity indicator, per state of perceived familiarity, risk tolerance, and external environment (traffic level).
	Perceived risk	This node represents the evolution of the subject's risk awareness in the ADS during the scenario.	Categories "high", "neutral", "low" defined on Q25, Q50, and Q75 of the risk awareness indicator, per state of level of trust, conditioning events (warning availability), and external environment (traffic level).

Type	PSF	Measure	States
	Perceived decision responsibility	Subject's perception of their responsibility to intervene in the scenario, measured per treatment.	Categories "high", "neutral", "low" defined on Q25, Q50, and Q75 of combined responsibility indicator, per state risk tolerance, conditioning events (warning availability), and external environment (traffic level).
	Task Complexity / Task Load	Subject's perception of the task load and complexity as measured by workload tool NASA TLX.	Categories "high", "neutral", "low" defined on Q25, Q50, and Q75 of the combined workload indicator, per state of level of trust, conditioning events (warning availability), and external environment (traffic level).
	Non-task load	Not Implemented. Data may be extracted from the video recordings for each test session.	N/A
Team	Responsiveness (Team Communication)	Estimated as the rate of performed takeovers given the number of required takeovers (warnings or tailgating events). Ideal values are close to 0.5.	Categories "high", "average", "low" per state of decision responsibility, perceived trust, urgency, conditioning events (warning availability), and external environment (traffic level).
	Compliance with Procedures (Team Cohesion)	Not Implemented. The ratio between the lane invasion by the driver and the ADS was too closely correlated to Engagement.	N/A
	Engagement (Team Coordination)	Estimated as the total driver's engagement time over the total duration of the scenario. No inherent value is associated with higher or lower numbers.	Categories "high", "average", "low" per state of task complexity, perceived risk, severity, and external environment (traffic level).
Key Perfor	Manual Driving Quality (TTC)	Estimated as the driver's tendency to maintain adequate safety	Categories "above ADS" and "below ADS" per state of responsiveness and

Type	PSF	Measure	States
		distances, i.e., the ratio between the driver's mean TTC and the vehicle's mean TTC during the scenario. Higher ratios have a positive value.	conditioning events (warning availability).
	Driver Strategy (Speeding)	Estimated as the driver's tendency to speed, i.e., ratio between the driver's mean speed and the vehicle's mean speed during the scenario. No inherent value is associated with higher or lower numbers.	Categories "below ADS" and "above ADS" per state of responsiveness, engagement, conditioning events (warning availability), and external environment (traffic level).
	Takeover Time (TOT)	The driver's reaction time to a warning or tailgating event, where positive indicates system-initiated takeover and negative, a driver-initiated takeover. No inherent value is associated with higher or lower numbers.	Categories "ADS initiated" and "driver initiated" per state of engagement, responsiveness, conditioning events (warning availability), and external environment (traffic level).
	Takeover Quality (TOQ)	Estimated as the ratio between tailgating events before and after tailgating onset. Lower ratios have a positive value.	Categories "high" and "low" per state of responsiveness, and external environment (traffic level).
	Takeover Performance (Collision)	Whether there was a collision, per scenario. Mapped as -1 for no collision, 1 for collision. Lower numbers have a positive value.	Categories "collision" and "no collision" per state of TOT, TOQ, conditioning events (warning availability), and external environment (traffic level).
	Operation Performance (Hard-Braking)	Number of hard-braking events, per scenario. A priori, lower numbers have a positive value (when paired with lower collision rates).	Categories "high", "neutral", "low" defined on Q25, Q50, and Q75 of the number of hard-braking events, per state of driver strategy, manual driving operation, and external environment (traffic level).

References

- Askaripoor, Hadi, Thilo Mueller, and Alois Knoll. 2023. "E/E Designer: A Framework to Design and Synthesize Vehicle E/E Architecture." *IEEE Transactions on Intelligent Vehicles*.
<https://doi.org/10.1109/TIV.2023.3324617>.
- Aukema, Amy, Kate Berman, Travis Gaydos, Ted Sienknecht, Chou-Lin Chen, Chris Wiacek, Tim Czapp, and Schuyler St Lawrence. 2023. "Real-World Effectiveness of Model Year 2015–2020 Advanced Driver Assistance Systems."
- Automated Vehicle Safety Consortium. 2019. "AVSC Best Practice for In-Vehicle Fallback Test Driver Selection, Training, and Oversight Procedures for Automated Vehicles Under Test." *SAE Industry Technologies Consortium*. <https://www.sae.org/standards/content/avsc00001201911/>.
- . 2023. "AVSC Best Practice for ADS Remote Assistance Use Case." *SAE Industry Technologies Consortia*.
- . 2024. "Revision of Best Practice for First Responder Interactions with Fleet-Managed Automated Driving System-Dedicated Vehicles (ADS-DVs)." *SAE Industry Technologies Consortium*.
- Ayas, Suzan, Birsen Donmez, Xing Tang, and Corresponding Author. 2023. "Drowsiness Mitigation Through Driver State Monitoring Systems: A Scoping Review." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 0 (0): 1–26. <https://doi.org/10.1177/00187208231208523>.
- Azarkhil, Mandana, and Ali Mosleh. 2014. "Dynamic Behavior of Operating Crew in Complex Systems." *Proceedings - Annual Reliability and Maintainability Symposium*.
<https://doi.org/10.1109/RAMS.2014.6798446>.
- Azarkhil, Mandana, Ali Mosleh, and Marilia Ramos. 2025a. "Team-Centered IDAC: Modeling and Simulation of Operating Crew in Complex Systems - Part 2: Simulation Aspects and Application." *Reliability Engineering & System Safety* 254 (February):110529. <https://doi.org/10.1016/j.ress.2024.110529>.
- . 2025b. "Team-Centered IDAC: Modeling and Simulation of Operating Crew in Complex Systems – Part 1: Team Model and Fundamentals." *Reliability Engineering & System Safety* 254 (February):110541. <https://www.sciencedirect.com/science/article/pii/S0951832024006136?via%3Dihub>.
- Azevedo-Sa, Hebert, Huajing Zhao, Connor Esterwood, X. Jessie Yang, Dawn M. Tilbury, and Lionel P. Robert. 2021. "How Internal and External Risks Affect the Relationships between Trust and Driver Behavior in Automated Driving Systems." *Transportation Research Part C: Emerging Technologies* 123 (February):102973. <https://doi.org/10.1016/j.trc.2021.102973>.
- Basnet, Sunil, Ahmad BahooToroody, Meriam Chaal, Janne Lahtinen, Victor Bolbot, and Osiris A. Valdez Banda. 2023. "Risk Analysis Methodology Using STPA-Based Bayesian Network- Applied to Remote Pilotage Operation." *Ocean Engineering* 270 (February):113569.
<https://doi.org/10.1016/j.oceaneng.2022.113569>.

- Bazilinskyy, P., S. M. Petermeijer, V. Petrovych, D. Dodou, and J. C.F. de Winter. 2018. "Take-over Requests in Highly Automated Driving: A Crowdsourcing Survey on Auditory, Vibrotactile, and Visual Displays." *Transportation Research Part F: Traffic Psychology and Behaviour* 56:82–98. <https://doi.org/10.1016/j.trf.2018.04.001>.
- Becker, Sandra, Stefan Brandenburg, and Manfred Thüning. 2022. "Driver-Initiated Take-Overs during Critical Braking Maneuvers in Automated Driving – The Role of Time Headway, Traction Usage, and Trust in Automation." *Accident Analysis and Prevention* 174 (September):106725. <https://doi.org/10.1016/j.aap.2022.106725>.
- Bhoopalam, Anirudh Kishore, Roy van den Berg, Niels Agatz, and Caspar G. Chorus. 2023. "The Long Road to Automated Trucking: Insights from Driver Focus Groups." *Transportation Research Part C: Emerging Technologies* 156 (November):104351. <https://doi.org/10.1016/j.TRC.2023.104351>.
- Boggs, Alexandra M., Ramin Arvin, and Asad J. Khattak. 2020. "Exploring the Who, What, When, Where, and Why of Automated Vehicle Disengagements." *Accident Analysis & Prevention* 136 (March):105406. <https://doi.org/10.1016/j.aap.2019.105406>.
- Boring, Ronald L., and Harold S. Blackman. 2007. "The Origins of the SPAR-H Method's Performance Shaping Factor Multipliers." *IEEE Conference on Human Factors and Power Plants*, 177–84. <https://doi.org/10.1109/HFPP.2007.4413202>.
- Bruck, Lucas, Bruce Haycock, and Ali Emadi. 2021. "A Review of Driving Simulation Technology and Applications." *IEEE Open Journal of Vehicular Technology* 2:1–16. <https://doi.org/10.1109/OJVT.2020.3036582>.
- Burns, Kevin, and Craig Bonaceto. 2020. "An Empirically Benchmarked Human Reliability Analysis of General Aviation." *Reliability Engineering & System Safety* 194 (February):106227. <https://doi.org/10.1016/j.res.2018.07.028>.
- Chaal, Meriam, Ahmad Bahootoroody, Sunil Basnet, Osiris A. Valdez Banda, and Floris Goerlandt. 2022. "Towards System-Theoretic Risk Assessment for Future Ships: A Framework for Selecting Risk Control Options." *Ocean Engineering* 259 (September):111797. <https://doi.org/10.1016/j.OCEANENG.2022.111797>.
- Cheng, Tingting, Marilia Ramos, and Ali Mosleh. 2024. "Human Errors and Influencing Factors in Digital Control Rooms : A Review and Some Suggestions." In *17th International Conference on Probabilistic Safety Assessment and Management & Asian Symposium on Risk Assessment and Management (PSAM17&ASRAM2024)*.
- Choi, Jong Kyu, and Yong Gu Ji. 2015. "Investigating the Importance of Trust on Adopting an Autonomous Vehicle." *International Journal of Human-Computer Interaction* 31 (10): 692–702. <https://doi.org/10.1080/10447318.2015.1070549>.
- Correa-Jullian, Camila, Marilia A. Ramos, Ali Mosleh, and Jiaqi Ma. 2024a. "An STPA-Based Analysis of Automated Driving Systems Fleet Maintenance Activities." In *2024 Annual Reliability and Maintainability*

Symposium (RAMS), 1–6. IEEE. <https://doi.org/10.1109/RAMS51492.2024.10457750>.

Correa-Jullian, Camila, Marilia Ramos, Ali Mosleh, and Jiaqi Ma. 2024b. “Operational Safety Hazard Identification Methodology for Automated Driving Systems Fleets.” *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, February. <https://doi.org/10.1177/1748006X241233863>.

———. 2025. “Human-Autonomy Teams: The Case of Remote Operators and Automated Driving Systems.” In *Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference*, edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Boudier, Roger Flage, and Marja Ylönen, 1957–64. Singapore: Research Publishing. https://doi.org/10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P3614-cd.

Correa-Jullian, Camila, Marilia Ramos, Ali Mosleh, and Jiaqi Ma. 2024a. “Exploring Performance Shaping Factors for Human-Autonomy Teams in Automated Driving System Operations.” In *17th International Conference on Probabilistic Safety Assessment and Management & Asian Symposium on Risk Assessment and Management (PSAM17&ASRAM2024)*.

———. 2024b. “Task Allocation and Control Transitions in Autonomous Driving System Operations.” In *34th European Safety and Reliability Conference (ESREL 2024), Cracow, Poland, 2024*. Vol. 1.

Correa Jullian, Camila Asuncion. 2025. “Human-Autonomy Teams in Automated Driving Systems: An Integrated Operational Safety Methodology.” University of California, Los Angeles. <https://escholarship.org/uc/item/6zt6r8r0>.

Correa-Jullian, C., D'Agostino, M. C, Mosleh, A., and Ma, J. (2026a). Automated Vehicle Safety: Heavy Duty Safety Drivers and Remote Operators – Safety Metric Foundations. *UC Office of the President: University of California Institute of Transportation Studies*. <http://dx.doi.org/10.7922/G2VM49N1> Retrieved from <https://escholarship.org/uc/item/01b0s8tm>

Correa-Jullian, C., Ramos, M., Mosleh, A., and Ma, J. (2026b). Exploring Factors Influencing Human-Autonomy Teams in Automated Driving System Operations. *IEEE Transactions on Human-Machine Systems*, vol. 56, no. 4, pp. 868-876, Aug. 2026, doi: 10.1109/THMS.2026.3702567.

DeGuzman, Chelsea A., Dina Kanaan, Samantha A. Hopkins, and Birsan Donmez. 2021. “Takeover Request (TOR) Effects during Different Automated Vehicle Failures.” *Journal of Intelligent Transportation Systems*, March, 1–31. <https://doi.org/10.1080/15472450.2021.1891536>.

Di, Xuan, and Rongye Shi. 2021. “A Survey on Autonomous Vehicle Control in the Era of Mixed-Autonomy: From Physics-Based to AI-Guided Driving Policy Learning.” *Transportation Research Part C: Emerging Technologies* 125 (April):103008. <https://doi.org/10.1016/j.trc.2021.103008>.

Dosovitskiy, Alexey, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. 2017. “CARLA: An Open Urban Driving Simulator,” no. CoRL, 1–16. <http://arxiv.org/abs/1711.03938>.

Ekanem, Nsimah J., Ali Mosleh, and Song Hua Shen. 2016. “Phoenix - A Model-Based Human Reliability Analysis Methodology: Qualitative Analysis Procedure.” *Reliability Engineering and System Safety* 145

(January):301–15. <https://doi.org/10.1016/j.res.2015.07.009>.

- Ernst, Claus Peter H., and Patrick Reinelt. 2017. "Autonomous Car Acceptance: Safety vs. Personal Driving Enjoyment." In *AMCIS 2017 - America's Conference on Information Systems: A Tradition of Innovation*, 2017-Augus:257. New York University Press.
- Favarò, Francesca, Sky Eurich, and Nazanin Nader. 2018. "Autonomous Vehicles' Disengagements: Trends, Triggers, and Regulatory Limitations." *Accident Analysis & Prevention* 110 (January):136–48. <https://doi.org/10.1016/j.aap.2017.11.001>.
- Gouribhatla, Raghuveer, and Srinivas S. Pulugurtha. 2022. "Drivers' Behavior When Driving Vehicles with or without Advanced Driver Assistance Systems: A Driver Simulator-Based Study." *Transportation Research Interdisciplinary Perspectives* 13 (March):100545. <https://doi.org/10.1016/j.trip.2022.100545>.
- Groth, Katrina M., and Ali Mosleh. 2012. "A Data-Informed PIF Hierarchy for Model-Based Human Reliability Analysis." *Reliability Engineering and System Safety* 108 (December):154–74. <https://doi.org/10.1016/j.res.2012.08.006>.
- Groth, Katrina M., Reuel Smith, and Ramin Moradi. 2019. "A Hybrid Algorithm for Developing Third Generation HRA Methods Using Simulator Data, Causal Models, and Cognitive Science." *Reliability Engineering & System Safety* 191 (November):106507. <https://doi.org/10.1016/j.RESS.2019.106507>.
- Guo, Chuanqi, and Ingrid Bouwer Utne. 2022. "Development of Risk Indicators for Losing Navigational Control of Autonomous Ships." *Ocean Engineering* 266 (December):113204. <https://doi.org/10.1016/j.oceaneng.2022.113204>.
- Ham, Dong Han, Won Jun Jung, and Jinkyun Park. 2021. "Identifying Key Factors Affecting the Performance of Team Decision-Making Based on the Analysis of Investigation Reports Issued from Diverse Industries." *Reliability Engineering and System Safety* 206 (November 2020): 107304. <https://doi.org/10.1016/j.res.2020.107304>.
- Han, Xu, Zhaoliang Zheng, Zewei Zhou, Yun Zhang, Tiahui Cai, Yifan Liu, Hao Xiang, et al. 2025. "CDA.AI for OpenCDA: AI Pathways for Cooperative Driving Automation Research." *Artificial Intelligence for Transportation* 1 (July):100002. <https://doi.org/10.1016/j.ait.2025.100002>.
- Han, X., Correa-Jullian, C. and Ma, J. When Humans Take the Wheel: A Behavioral Analysis of Emergency Takeovers in Shared-Control Driving Using OpenCDA (Under development).
- Hart, Sandra G. 2006. "NASA-Task Load Index (NASA-TLX); 20 Years Later." In *Proceedings of the Human Factors and Ergonomics Society*, 904–8. SAGE PublicationsSage CA: Los Angeles, CA. <https://doi.org/10.1177/154193120605000909>.
- Hart, Sandra G., and Lowell E. Staveland. 1988. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." *Advances in Psychology* 52 (C): 139–83. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9).
- Heikoop, Daniël D., Joost C.F. de Winter, Bart van Arem, and Neville A. Stanton. 2019. "Acclimatizing to

Automation: Driver Workload and Stress during Partially Automated Car Following in Real Traffic.” *Transportation Research Part F: Traffic Psychology and Behaviour* 65 (August):503–17. <https://doi.org/10.1016/j.trf.2019.07.024>.

Janssen, Christian P., Stella F. Donker, Duncan P. Brumby, and Andrew L. Kun. 2019. “History and Future of Human-Automation Interaction.” *International Journal of Human Computer Studies* 131 (November):99–107. <https://doi.org/10.1016/j.ijhcs.2019.05.006>.

Kettwich, Carmen, Andreas Schrank, Hüseyin Huseyin Avsar, and Michael Oehl. 2021. “What If the Automation Fails? – A Classification of Scenarios in Teleoperated Driving.” In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 92–96. New York, NY, USA: ACM. <https://doi.org/10.1145/3473682.3480271>.

Khastgir, Siddartha, Stewart Birrell, Gunwant Dhadyalla, and Paul Jennings. 2019. “Effect of Knowledge of Automation Capability on Trust and Workload in an Automated Vehicle: A Driving Simulator Study.” In *Advances in Intelligent Systems and Computing*, 786:410–20. Springer Verlag. https://doi.org/10.1007/978-3-319-93885-1_37.

Khastgir, Siddartha, Simon Brewerton, John Thomas, and Paul Jennings. 2021. “Systems Approach to Creating Test Scenarios for Automated Driving Systems.” *Reliability Engineering & System Safety* 215 (November):107610. <https://doi.org/10.1016/j.ress.2021.107610>.

Ledesma, Rubén Daniel, Silvana Andrea Montes, Fernando Martín Poó, and María Fernanda López-Ramón. 2015. “Measuring Individual Differences in Driver Inattention: Further Validation of the Attention-Related Driving Errors Scale.” *Human Factors* 57 (2): 193–207. https://doi.org/10.1177/0018720814546530/ASSET/IMAGES/LARGE/10.1177_0018720814546530-FIG2.JPEG.

T. Lochrane, L. Dailey, and C. Tucker, “CarmaSM: Driving innovation,” *Public Roads*, vol. 83, no. 4, 2020. [Online]. Available: <https://its.dot.gov/cda/>

Lodhi, Shikhar Singh, Neetesh Kumar, and Pradumn Kumar Pandey. 2023. “Autonomous Vehicular Overtaking Maneuver: A Survey and Taxonomy.” *Vehicular Communications* 42 (August):100623. <https://doi.org/10.1016/J.VEHCOM.2023.100623>.

Lopez, Pablo Alvarez, Evamarie Wiessner, Michael Behrisch, Laura Bieker-Walz, Jakob Erdmann, Yun-Pang Flotterod, Robert Hilbrich, Leonhard Lucken, Johannes Rummel, and Peter Wagner. 2018. “Microscopic Traffic Simulation Using SUMO.” In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018-Novem:2575–82. IEEE. <https://doi.org/10.1109/ITSC.2018.8569938>.

Louw, Tyron, Jonny Kuo, Richard Romano, Vishnu Radhakrishnan, Michael G. Lenné, and Natasha Merat. 2019. “Engaging in NDRTs Affects Drivers’ Responses and Glance Patterns after Silent Automation Failures.” *Transportation Research Part F: Traffic Psychology and Behaviour* 62 (April):870–82. <https://doi.org/10.1016/j.trf.2019.03.020>.

Lyons, Joseph B., Katia Sycara, Michael Lewis, and August Capiola. 2021. “Human–Autonomy Teaming:

Definitions, Debates, and Directions.” *Frontiers in Psychology* 12 (May):1932.
<https://doi.org/10.3389/fpsyg.2021.589585>.

- Ma, Jiaqi, Camila Correa-Jullian, Marilia Ramos, and Xin Xia. 2024. “Risk Assessment for Remotely Operation of Level 4 Automated Driving Systems in Mobility as a Service Transport.” UC Office of the President: University of California Institute of Transportation Studies.
<https://doi.org/https://doi.org/10.7922/G23N21QC>.
- Mamdikar, Mohan Rao, Vinay Kumar, and Pooja Singh. 2022. “Dynamic Reliability Analysis Framework Using Fault Tree and Dynamic Bayesian Network: A Case Study of NPP.” *Nuclear Engineering and Technology* 54 (4): 1213–20. <https://doi.org/10.1016/J.NET.2021.09.038>.
- Manchon, J. B., Mercedes Bueno, and Jordan Navarro. 2022. “How the Initial Level of Trust in Automated Driving Impacts Drivers’ Behaviour and Early Trust Construction.” *Transportation Research Part F: Traffic Psychology and Behaviour* 86 (April):281–95. <https://doi.org/10.1016/j.trf.2022.02.006>.
- Marberger, Claus, Holger Mielenz, Frederik Naujoks, Jonas Radlmayr, Klaus Bengler, and Bernhard Wandtner. 2018. “Understanding and Applying the Concept of ‘Driver Availability’ in Automated Driving.” In *Advances in Intelligent Systems and Computing*, 597:595–605. Springer Verlag.
https://doi.org/10.1007/978-3-319-60441-1_58.
- Marcano, Mauricio, Sergio Diaz, Joshue Perez, and Eloy Irigoyen. 2020. “A Review of Shared Control for Automated Vehicles: Theory and Applications.” *IEEE Transactions on Human-Machine Systems* 50 (6): 475–91. <https://doi.org/10.1109/THMS.2020.3017748>.
- McKerral, Angus, Kristen Pammer, and Cassandra Gauld. 2023. “Supervising the Self-Driving Car: Situation Awareness and Fatigue during Highly Automated Driving.” *Accident Analysis & Prevention* 187 (July):107068. <https://doi.org/10.1016/J.AAP.2023.107068>.
- McNeese, Nathan J., Mustafa Demir, Nancy J. Cooke, and Christopher Myers. 2018. “Teaming With a Synthetic Teammate: Insights into Human-Autonomy Teaming.” *Human Factors* 60 (2): 262–73.
<https://doi.org/10.1177/0018720817743223>.
- Mishler, Scott, and Jing Chen. 2023. “Effect of Automation Failure Type on Trust Development in Driving Automation Systems.” *Applied Ergonomics* 106 (January):103913.
<https://doi.org/10.1016/J.APERGO.2022.103913>.
- Mosier, Kathleen L., Ute Fischer, Barbara K. Burian, and Janeen A. Kochan. 2017. “Autonomous, Context-Sensitive, Task Management Systems and Decision Support Tools I: Human-Autonomy Teaming Fundamentals and State of the Art,” no. NASA/TM—2017–219565, 75.
<https://ntrs.nasa.gov/search.jsp?R=20180003355>.
- Mueller, Alexandra S., Jessica B. Cicchino, Jeremiah Singer, and James W. Jenness. 2020. “Effects of Training and Display Content on Level 2 Driving Automation Interface Usability.” *Transportation Research Part F: Traffic Psychology and Behaviour* 69 (February):61–71. <https://doi.org/10.1016/j.trf.2019.12.010>.
- Muslim, Husam, Cho Kiu Leung, and Makoto Itoh. 2022. “Design and Evaluation of Cooperative Human–

Machine Interface for Changing Lanes in Conditional Driving Automation.” *Accident Analysis and Prevention* 174 (September):106719. <https://doi.org/10.1016/j.aap.2022.106719>.

O’Neill, Thomas A., Christopher Flathmann, Nathan J. McNeese, and Eduardo Salas. 2023. “Human-Autonomy Teaming: Need for a Guiding Team-Based Framework?” *Computers in Human Behavior*, April, 107762. <https://doi.org/10.1016/j.chb.2023.107762>.

O’Neill, Thomas, Nathan McNeese, Amy Barron, and Beau Schelble. 2022. “Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature.” *Human Factors* 64 (5): 904–38. <https://doi.org/10.1177/0018720820960865>.

Pan, Hengyan, Ke Xu, Yang Qin, and Yonggang Wang. 2023. “How Does Drivers’ Trust in Vehicle Automation Affect Non-Driving-Related Task Engagement, Vigilance, and Initiative Takeover Performance after Experiencing System Failure?” *Transportation Research Part F: Traffic Psychology and Behaviour* 98 (October):73–90. <https://doi.org/10.1016/j.TRF.2023.09.001>.

Paschalidis, Evangelos, Charisma F. Choudhury, and Stephane Hess. 2018. “Modelling the Effects of Stress on Gap-Acceptance Decisions Combining Data from Driving Simulator and Physiological Sensors.” *Transportation Research Part F: Traffic Psychology and Behaviour* 59 (November):418–35. <https://doi.org/10.1016/j.trf.2018.09.019>.

Payre, William, Julien Cestac, and Patricia Delhomme. 2014. “Intention to Use a Fully Automated Car: Attitudes and a Priori Acceptability.” *Transportation Research Part F: Traffic Psychology and Behaviour* 27 (PB): 252–63. <https://doi.org/10.1016/J.TRF.2014.04.009>.

———. 2016. “Fully Automated Driving: Impact of Trust and Practice on Manual Control Recovery.” *Human Factors* 58 (2): 229–41. <https://doi.org/10.1177/0018720815612319>.

Payre, William, Jaume Perello March, Giedre Sabaliauskaite, Hesamaldin Jadidbonab, Siraj Shaikh, Hoang Nga Nguyen, and Stewart Birrell. 2022. “Understanding Drivers’ Trust After Software Malfunctions and Cyber Intrusions of Digital Displays in Automated Cars.” *Human Factors in Transportation* 60:320–28. <https://doi.org/10.54941/AHFE1002463>.

Porthin, Markus, Marja Liinasuo, and Terhi Kling. 2020. “Effects of Digitalization of Nuclear Power Plant Control Rooms on Human Reliability Analysis – A Review.” *Reliability Engineering & System Safety* 194 (February):106415. <https://doi.org/10.1016/j.ress.2019.03.022>.

Ramakrishnan, P., B. Balasingam, and F. Biondi. 2021. “Cognitive Load Estimation for Adaptive Human–Machine System Automation.” In *Learning Control*, 35–58. Elsevier. <https://doi.org/10.1016/B978-0-12-822314-7.00007-9>.

Reason, James, Antony Manstead, Stradling Stephen, James Baxter, and Karen Campbell. 1990. “Errors and Violations on the Roads: A Real Distinction?” *Ergonomics* 33 (10–11): 1315–32. <https://doi.org/10.1080/00140139008925335>.

Rendon-Velez, E., P. M. van Leeuwen, R. Happee, I. Horváth, W. F. van der Vegte, and J. C.F. de Winter. 2016. “The Effects of Time Pressure on Driver Performance and Physiological Activity: A Driving Simulator

Study.” *Transportation Research Part F: Traffic Psychology and Behaviour* 41 (August):150–69. <https://doi.org/10.1016/j.trf.2016.06.013>.

- SAE On-Road Automated Vehicle Standards Committee. 2021. “Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles.” *SAE Standard J3016_202104*. 400 Commonwealth Drive, Warrendale, PA, United States: SAE International. https://doi.org/10.4271/J3016_202104.
- Samani, Ali Riahi, Sabyasachee Mishra, and Kakan Dey. 2022. “Assessing the Effect of Long-Automated Driving Operation, Repeated Take-over Requests, and Driver’s Characteristics on Commercial Motor Vehicle Drivers’ Driving Behavior and Reaction Time in Highly Automated Vehicles.” *Transportation Research Part F: Traffic Psychology and Behaviour* 84 (January):239–61. <https://doi.org/10.1016/j.trf.2021.10.015>.
- Stapel, Jork, Freddy Antony Mullaikkal-Babu, and Riender Happee. 2019. “Automated Driving Reduces Perceived Workload, but Monitoring Causes Higher Cognitive Load than Manual Driving.” *Transportation Research Part F: Traffic Psychology and Behaviour* 60:590–605. <https://doi.org/10.1016/j.trf.2018.11.006>.
- Sundaramurthi, R., and C. Smidts. 2013. “Human Reliability Modeling for the Next Generation System Code.” *Annals of Nuclear Energy* 52 (February):137–56. <https://doi.org/10.1016/j.anucene.2012.07.027>.
- Tanshi, Foghor, and Dirk Soffker. 2022. “Determination of Takeover Time Budget Based on Analysis of Driver Behavior.” *IEEE Open Journal of Intelligent Transportation Systems* 3 (November):813–24. <https://doi.org/10.1109/OJITS.2022.3224677>.
- Tinga, Angelica M., Ilse M. van Zeumeren, Michiel Christoph, Elmer van Grondelle, Diane Cleij, Anna Aldea, and Nicole van Nes. 2023. “Development and Evaluation of a Human Machine Interface to Support Mode Awareness in Different Automated Driving Modes.” *Transportation Research Part F: Traffic Psychology and Behaviour* 92 (April 2022): 238–54. <https://doi.org/10.1016/j.trf.2022.10.023>.
- Vanderhaegen, Frédéric. 2021. “Heuristic-Based Method for Conflict Discovery of Shared Control between Humans and Autonomous Systems - A Driving Automation Case Study.” *Robotics and Autonomous Systems* 146 (December):103867. <https://doi.org/10.1016/j.robot.2021.103867>.
- Veitch, Erik, Ole Andreas Alsos, Tingting Cheng, Kristin Senderud, and Ingrid Bouwer Utne. 2024. “Human Factor Influences on Supervisory Control of Remotely Operated and Autonomous Vessels.” *Ocean Engineering* 299 (May):117257. <https://doi.org/10.1016/J.OCEANENG.2024.117257>.
- Walliser, James C., Ewart J. de Visser, Eva Wiese, and Tyler H. Shaw. 2019. “Team Structure and Team Building Improve Human–Machine Teaming With Autonomous Agents.” *Journal of Cognitive Engineering and Decision Making* 13 (4): 258–78. <https://doi.org/10.1177/1555343419867563>.
- Wang, Yi, Wei Zhang, and Ronggang Zhou. 2022. “Speech-Based Takeover Requests in Conditionally Automated Driving: Effects of Different Voices on the Driver Takeover Performance.” *Applied Ergonomics* 101 (May):103695. <https://doi.org/10.1016/j.apergo.2022.103695>.
- Warg, Fredrik, Martin Skoglund, and Matthew Sassman. 2021. “Human Interaction Safety Analysis Method for Agreements with Connected Automated Vehicles.” In *2021 IEEE 94th Vehicular Technology Conference*

- (VTC2021-Fall), 2021-Sept:01–07. IEEE. <https://doi.org/10.1109/VTC2021-Fall52928.2021.9625202>.
- Winter, J. C.F. De. 2013. “Predicting Self-Reported Violations among Novice License Drivers Using Pre-License Simulator Measures.” *Accident Analysis & Prevention* 52 (March):71–79. <https://doi.org/10.1016/j.aap.2012.12.018>.
- Wu, Yanbin, Ken Kihara, Yuji Takeda, Toshihisa Sato, Motoyuki Akamatsu, and Satoshi Kitazaki. 2019. “Effects of Scheduled Manual Driving on Drowsiness and Response to Take over Request: A Simulator Study towards Understanding Drivers in Automated Driving.” *Accident Analysis and Prevention* 124 (June 2018): 202–9. <https://doi.org/10.1016/j.aap.2019.01.013>.
- Xing, Yang, Chen Lv, Dongpu Cao, and Peng Hang. 2021. “Toward Human-Vehicle Collaboration: Review and Perspectives on Human-Centered Collaborative Automated Driving.” *Transportation Research Part C: Emerging Technologies* 128 (July):103199. <https://doi.org/10.1016/j.trc.2021.103199>.
- Xu, Runsheng, Yi Guo, Xu Han, Xin Xia, Hao Xiang, and Jiaqi Ma. 2021. “OpenCDA: An Open Cooperative Driving Automation Framework Integrated with Co-Simulation.” *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC 2021-Sept*:1155–62. <https://doi.org/10.1109/ITSC48978.2021.9564825>.
- Xu, Runsheng, Hao Xiang, Xu Han, Xin Xia, Zonglin Meng, Chia-Ju Chen, Camila Correa-Jullian, and Jiaqi Ma. 2023. “The OpenCDA Open-Source Ecosystem for Cooperative Driving Automation Research.” *IEEE Transactions on Intelligent Vehicles* 8 (4): 2698–2711. <https://doi.org/10.1109/TIV.2023.3244948>.
- Yu, Bo, Shan Bao, Yuren Chen, and David J. LeBlanc. 2021. “Effects of an Integrated Collision Warning System on Risk Compensation Behavior: An Examination under Naturalistic Driving Conditions.” *Accident Analysis & Prevention* 163 (December):106450. <https://doi.org/10.1016/j.aap.2021.106450>.
- Yue, Lishengsa, Mohamed Abdel-Aty, Yina Wu, Jorge Ugan, and Cheng Yuan. 2021. “Effects of Forward Collision Warning Technology in Different Pre-Crash Scenarios.” *Transportation Research Part F: Traffic Psychology and Behaviour* 76 (January):336–52. <https://doi.org/10.1016/j.trf.2020.12.004>.
- Zangi, Noa, Rawan Srour-Zreik, Dana Ridet, Hadas Chasidim, and Avinoam Borowsky. 2022. “Driver Distraction and Its Effects on Partially Automated Driving Performance: A Driving Simulator Study among Young-Experienced Drivers.” *Accident Analysis and Prevention* 166 (March):106565. <https://doi.org/10.1016/j.aap.2022.106565>.
- Zarei, Esmaeil, Faisal Khan, and Rouzbeh Abbassi. 2022. “A Dynamic Human-Factor Risk Model to Analyze Safety in Sociotechnical Systems.” *Process Safety and Environmental Protection* 164 (August):479–98. <https://doi.org/10.1016/j.psep.2022.06.040>.
- Zhang, Bo, Joost de Winter, Silvia Varotto, Riender Happee, and Marieke Martens. 2019. “Determinants of Take-over Time from Automated Driving: A Meta-Analysis of 129 Studies.” *Transportation Research Part F: Traffic Psychology and Behaviour* 64:285–307. <https://doi.org/10.1016/j.trf.2019.04.020>.
- Zhang, Yuxiao, Alexander Carballo, Hanting Yang, and Kazuya Takeda. 2023. “Perception and Sensing for Autonomous Vehicles under Adverse Weather Conditions: A Survey.” *ISPRS Journal of Photogrammetry*

and Remote Sensing 196 (February):146–77. <https://doi.org/10.1016/J.ISPRSJPRS.2022.12.021>.

Zhao, Xingyu, Kizito Salako, Lorenzo Strigini, Valentin Robu, and David Flynn. 2020. “Assessing Safety-Critical Systems from Operational Testing: A Study on Autonomous Vehicles.” *Information and Software Technology* 128 (December):106393. <https://doi.org/10.1016/j.infsof.2020.106393>.

Zhu, Meixin, Xuesong Wang, and Jingyun Hu. 2020. “Impact on Car Following Behavior of a Forward Collision Warning System with Headway Monitoring.” *Transportation Research Part C: Emerging Technologies* 111 (February):226–44. <https://doi.org/10.1016/J.TRC.2019.12.015>.

