

MASTERARBEIT | MASTER'S THESIS

Titel | Title

A Predictive Processing Account of the Immediate and Sustained Effects of TOR-Assertiveness in Automated Driving

verfasst von | submitted by

Dipl.-Ing. Vanessa Kraut B.A.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree
programme code as it appears on the
student record sheet:

UA 066 013

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Joint-Degree Middle European
interdisc. master prog. in Cognitive Science

Betreut von | Supervisor:

Ronald Sladky Bakk. MSc PhD

Acknowledgements

I would like to thank those who supported me throughout the completion of this thesis. First, I would like to thank my supervisor Ronald Sladky, for his insightful feedback and expertise, which were essential for this work. I am grateful to the Austrian Institute of Technology and the employees at the Technology Experience Unit for providing me with access to their state-of-the-art research equipment and priceless experience. I would also like to extend my thanks to project supervisor Peter Fröhlich for his mentorship throughout this project and researcher Lenart Motnikar, whose significant contributions in the study design, data analysis, and interpretation were decisive to the success of this research. Lastly, I am deeply grateful to my family, my close friends and Daniel for their endless encouragement and support throughout this process.

Abstract

As Advanced Driver Assistance Systems (ADAS) continue to evolve, the interaction between human drivers and automated systems becomes increasingly safety-critical, particularly during the transition of control from the system to the driver. This thesis investigates the immediate and sustained effects of assertiveness in takeover requests (TOR) on driver performance and stress. Through an experimental study, the impact of assertive TORs on drivers was researched quantitatively and qualitatively. A combination of classical user research methodologies and the predictive processing framework allowed it to perform a more comprehensive analysis.

Previous driving automation studies suggest that assertive takeover requests, for instance, those delivered through loud, urgent, or disruptive auditory signals, can prompt faster driver reactions. However, these signals may increase stress, leading to a suboptimal driving performance. The key concepts of the predictive processing framework – while still underutilized in human factors research – offer a unique perspective on how drivers cope with fluctuating mental and physical demands in automated driving. By conceptualizing the mind as constantly generating and updating internal models in response to sensory inputs, this framework allows to understand drivers' reflexive actions as well as their longer-term decision-making processes.

A driving simulation experiment was conducted, in which assertive and non-assertive takeover requests were administered to 31 participants. Physiological responses, driving performance metrics, and self-reported stress were collected and analyzed. The results confirm that highly assertive TORs lead to faster reaction times and more forceful braking. Contrary to expectations, they did not significantly affect driver performance during the sustained driving period after the takeover.

Finally, this thesis bridges methodologies of human-computer interaction and cognitive science, contributing to the body of research on ADAS interface design for enhancing safety and user experience in semi-automated vehicles.

Kurzfassung

Durch die fortschreitende Entwicklung von Advanced Driving Assistance Systems (ADAS) wird die Interaktion von System und Lenker*innen zunehmend zu einem Sicherheitsfaktor. Der Autopilot in einem teil-automatisierten Fahrzeug ist zwar in der Lage alle Fahraufgaben selbständig auszuführen, jedoch obliegt den menschlichen Fahrer*innen die Verantwortung über die Systemüberwachung. Auf Aufforderung des Systems hin, müssen Lenker*innen zu jedem Zeitpunkt eine sichere Kontrollübernahme über das Fahrzeug gewährleisten können. In der vorliegenden Arbeit wurden die unmittelbaren und längerfristigen Auswirkungen von besonders nachdrücklichen Übernahmeaufforderungen (Englisch: Takeover Requests, abgekürzt: TORs) auf Fahrperformance und -stress untersucht. Zur Durchführung einer umfassenden Analyse werden im Zuge dieser Studie klassische User Research Methoden mit dem Kognitions-wissenschaftlichen Predictive Processing Framework kombiniert.

Besonders nachdrückliche TORs, etwa laute, abrupte oder disruptive Töne, führen zu schnelleren Kontrollübernahmen, wie frühere Studien zeigen. Gleichzeitig können diese Art von Signalen erhöhten Stress und suboptimale Fahrperformance bedingen. In Ergänzung zur üblichen User Research Forschung, bietet das Predictive Processing Framework eine einzigartige Perspektive, wie Fahrer*innen mit den fluktuierenden mentalen und physischen Anforderungen in teil-automatisierten Fahrzeugen umgehen. Das Framework betrachtet das Gehirn als eine statistische Maschine, die innere Modelle mit Vorhersagen auf Basis von sensorischen Reizen kontinuierlich aktualisiert. Dies ermöglicht die Interpretation von sowohl reflexiven Reaktionen wie auch längerfristigen Entscheidungsprozessen von Fahrer*innen.

Für diese Arbeit wurde ein Fahrsimulationsexperiment mit 31 Teilnehmer*innen durchgeführt, bei dem nachdrückliche und neutrale TORs präsentiert wurden. Im Zuge dessen wurden physiologische Reaktionen, Metriken für Fahrperformance und subjektive Stressbewertungen erfasst und analysiert. Die Ergebnisse zeigten, dass nachdrückliche Übernahmeaufforderungen zu schnelleren Reaktionszeiten und stärkeren Bremsreaktionen führen. Entgegen den Erwartungen wurden jedoch keine signifikanten Auswirkungen im Zeitraum nach der Übernahme beobachtet. Durch die Verknüpfung von Human-Computer Interaction und Kognitionswissenschaft leistet diese Arbeit einen Beitrag zur ADAS-Forschung, die das Ziel verfolgt, eine möglichst hohe Sicherheit und Benutzerfreundlichkeit in teilautomatisierten Fahrzeugen sicherzustellen.

Contents

Acknowledgements.....	i
Abstract.....	ii
Kurzfassung.....	iii
Contents.....	iv
1. Introduction	1
1.1 Research Objective	2
1.2 Thesis Structure	5
1.3 Remarks on Publications	5
2. Literature Review.....	6
2.1 Overview on Advanced Driving Automation Assistance	6
2.2 Design of Takeover Requests (TORs)	9
2.3. Stress and Cognitive Load in Driving Automation	12
2.4. Predictive Processing and Cognitive Science Paradigms.....	14
2.4.1 Central Concepts	15
2.4.2 Predictive Processing in Driving Automation Studies	18
3. Methodology	20
3.1 Pilot Study	20
3.2 Main Study Design.....	25
3.3 Participants	27
3.3 Apparatus and Equipment	27
3.3 Procedure	30
3.6 Data Collection	34
3.7 Data Analysis	37
4. Experimental Results	39
4.1 Quantitative Results.....	39
4.1.1 Immediate Effects of Assertive Takeover Requests.....	40
4.1.2 Sustained Effects of Assertive Takeover Requests.....	40
4.1.3 Subjective Ratings on Stress and Performance	41
4.2 Qualitative Results	43
4.3 Additional Findings.....	46
4.3.1 Adaptation Effects across Laps	46
4.3.2 Correlations	49
5. Discussion	50

5.1 Interpretation of Experimental Findings.....	50
5.1.1 Significant Effects for Immediate Measures	50
5.1.2 No Impacts on Subsequent Measures Observed	51
5.2 Predictive Processing Insights	52
5.2.1 Internal Model Updating when Driving with ADAS	52
5.2.2 Prediction Error Minimization in the Immediate Period	56
5.2.3 Prediction Error Minimization in the Sustained Period.....	60
5.3 Study Limitations	62
6. Conclusion.....	64
Bibliography.....	67
List of Figures.....	78
List of Tables.....	79
Appendix A: Pilot Study Procedure	80
Appendix B: Forms and Questionnaires.....	81
B.1 NASA Task Load Index.....	81
B.2 German 10-item Perceived Stress Scale	82
B.3 English 10-item Perceived Stress Scale.....	83

1. Introduction

As the capabilities of Advanced Driver Assistance Systems (ADAS) are continuously developed, these systems reshape driving dynamics, and the cognitive demands placed on drivers. The Society of Automotive Engineers (SAE) defines six levels of automation. The currently widely distributed Level 3 systems allow full control over safety-critical functions but require immediate driver re-engagement in critical instances (Agrawal & Peeta, 2021; Du et al., 2020). Assistance systems have been found to reduce crashes per mile, but any mismatch between driver expectations and automation capabilities can lead to fatal incidents (Luetge, 2017). Therefore, it is crucial for interface designers to understand how the system and human drivers can interact optimally.

Vehicular automation induces inadvertent secondary effects: the reduction of driving effort can make drivers more vulnerable to decreased situational awareness. A decisive point of interaction is the safe transition of control from the autopilot to the human driver, facilitated by the takeover request (TOR). Previous research observed that auditory TORs, which were particularly loud, annoying, and assertive, led to faster reaction times. However, they may also induce stress, compromising control quality shortly after the takeover (Politis et al., 2015; Wong et al., 2019). Moreover, driving performance is further impaired when a driver experiences cognitive overload and excessive emotional arousal in response to stressful, unexpected alerts (Agrawal & Peeta, 2021; World Health Organization, 2022). Even up to a minute after the takeover, drivers may exhibit overload through high variability in accelerations or steering (Biondi et al., 2014; Roche & Brandenburg, 2018).

To ensure adequate safety with ADAS, it is essential to design interfaces that align with human capabilities, as well as human error (Engström et al., 2018; Wei et al., 2022). Besides extensive research on ADAS in human-computer interaction literature, a small amount of research has explored the application of the predictive processing framework to TOR-design. This comprehensive paradigm, based in cognitive science, depicts the brain

as a statistical machine that continuously aims to predict future states by generating internal models of the world and updating these based on experiences and assumptions. Predictive processing provides a single principle which may account for perception, action and the essential contribution of the body and environment in the structure of sensorimotor interactions (Seth, 2014). Extending beyond classical human-computer interaction research, predictive processing provides a theoretical approach with a plausible physiological foundation that has the potential to describe driver behavior and states during takeovers more comprehensively.

Previous research on assertive TORs strongly focuses on the initial takeover process, even though the post-takeover performance has been identified as a crucial research avenue (Tan & Zhang, 2024). This thesis aims to explore the sustained effects of assertiveness on driver performance and emotional arousal. In addition, experimental results are discussed from the perspective of predictive processing and the possibility of a combined research strategy for a more comprehensive view of driver responses. This thesis argues that assertive auditory human-computer interaction elements can improve performance but may induce stress, negating the benefits. By correlating perceived stress with psychophysiological measurements and examining effects on cognitive functions, this research aims to contribute a better understanding of the interaction between assertive TORs, driver behavior, stress and mental models using the predictive processing framework. This balance is necessary to ensure takeover requests enhance safety without compromising driver state and situational awareness (Agrawal & Peeta, 2021).

1.1 Research Objective

Automated driving systems change drivers' responsibilities and cognitive workloads, potentially leading to decreased situational awareness due to underload or distractions from secondary tasks. While assertive signals can grab attention and prompt quick reactions, they may also induce stress, compromising takeover quality. This thesis explores the primary and secondary effects of these signals over time and theorizes about the cognitive processes involved using the predictive processing framework. The first research question was determined as such:

RQ1: How does TOR-assertiveness affect driver performance and stress during a takeover and a subsequent driving task?

The hypotheses address relevant driving performance measures in addition to measures of stress and self-reported ratings on both driving performance and stress. Findings from neuroscientific research allow to draw conclusions about stress due to cognitive overload from heart rate variability measures. Additionally, reliable insights on stress levels can be gathered from self-reported ratings (Schneider, 2020).

During the immediate time period after a takeover request, assertive takeover requests result in...

- H1: ... a faster reaction time.
- H2: ... more forceful braking.
- H3: ... a higher risk of collisions.
- H4: ... higher average speed.
- H5: ... a higher heart rate ratio.
- H6: ... a higher SDNN ratio.
- H7: ... a lower RMSSD ratio.

During a subsequent driving task of a minimum 40 second duration, assertive takeover requests result in...

- H8: ... higher average speed.
- H9: ... higher speed variability.
- H10: ... higher steering wheel variability.
- H11: ... a higher heart rate ratio.
- H12: ... a higher SDNN ratio.
- H13: ... a higher RMSSD ratio.

After completion of a takeover and manual driving section, assertive takeover requests result in...

H14: ... higher self-reported stress on the Perceived Stress Scale.

H15: ... lower self-reported driving performance on a Likert scale.

Research Question One (RQ1) is based in human-computer interaction, focusing on driver responses within the context of interaction design research. To better understand the cognitive processes underlying this behavior, Research Question Two (RQ2) is rooted in cognitive science. Prediction and expectancy have been emphasized in existing human factors theories and models (e.g. Endsley 1995, Wickens 2003, Rasmussen 1985), however only a small percentage of applied human factors research address specific concepts offered by the predictive processing framework (Engström et al., 2018; Nozari et al., 2023; Wei et al., 2022). The predictive processing paradigm is used to examine driver mental models concerning the immediate and sustained effects of TOR-assertiveness. Therefore, research question two is formulated as:

RQ2: How does TOR-assertiveness affect prediction error during a takeover and subsequent driving task?

Several guiding questions are used to provide a predictive processing account of the study results.

1. How might drivers' internal models update and optimize when operating a vehicle with an Advanced Driving Assisted System?
2. How is prediction error minimization observable in driver responses immediately following a highly assertive TOR compared to a less assertive TOR?
3. How does TOR-assertiveness influence predictions error minimization during the subsequent driving task?

Overall, this master thesis aims to test a specific TOR-characteristic in a simulated environment and conceptualize the results using a theoretical framework with plausible physiological basis. These insights, combining classical HCI-research with cognitive science, can potentially enhance the design of TORs for improved safety in advanced driver assistance systems.

1.2 Thesis Structure

This thesis explores assertiveness as a characteristic for takeover requests (TORs) with the help of a human-computer interaction (HCI) experiment and through the lens of the predictive processing perspective. The literature review in Section 2 provides state-of-the-art insights on the development and testing of TORs, along with an overview of the predictive processing framework and its application to ADAS-research.

"Methods and Materials" in Section 3 details the experimental setup and procedure used to explore Research Question 1 (RQ1). This includes the pilot study which determined the experimental variables. Section 4, titled "Experimental Results," presents the findings related to the hypotheses of RQ1.

The "Discussion"-section first provides an interpretation of the RQ1 results, before addressing the guided questions formulated in RQ2 and outlining future research possibilities. In Section 4, "Conclusion", the thesis summarizes its contributions to integrating ADAS research and cognitive science.

1.3 Remarks on Publications

The empirical research for this thesis was supported by the Austrian Institute of Technology. During the writing of this master's thesis, I also contributed to a conference paper titled "Assertive Takeover Requests: Immediate and Sustained Effects on Stress and Performance" (Kraut et al., 2023). This paper was accepted by the 15th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, held in September 2023 in Ingolstadt, Germany. Many of the tables and figures in Section 4 "Experimental Results", as well as Section 5 "Discussion" were created concurrently with the preparation of the conference paper and are appropriately referenced. Additionally, the research plan was first presented at the May Conference 2022 in Zagreb, Croatia, where it was featured in a short presentation and an abstract published in the conference proceedings (Kraut, 2022).

2. Literature Review

The following section reviews state-of-the-art literature on the key components of this research: Advanced Driving Automation Assistance (ADAS), design of takeover requests (TOR), the influence of stress in automated driving, and the predictive processing framework. The review reveals a research gap for long-term effects of TORs, despite previous findings that safe driving extends beyond the moment of takeover and mental load fluctuations continuously affect drivers in automated vehicles. Additionally, while human factors research includes concepts like attention, relatively few studies integrate predictive processing. This highlights the need for an interpretative framework to evaluate the interaction between assistance systems and drivers, while explaining the interactions of mental load with driving performance over time.

2.1 Overview on Advanced Driving Automation Assistance

Driving automation is designed to support both the safety and comfort of drivers by assisting in or fully automating aspects of the driving task. In order to increase safety, reduce serious injuries and eliminate fatal crashes, the World Health Organization advocates to accommodate human error as much as possible. Globally, road accidents account for more than 1.3 million deaths each year, with causes ranging from speeding, distracted driving, or driving under the influence to infrastructural issues like poor law enforcement and vehicle safety standards (World Health Organization, 2022).

As driving automation systems continue to advance, their potential to mitigate human error and reduce accident risks becomes increasingly significant. Recognizing the critical role of automation, several countries, including the United Kingdom, the United States, and Japan, have proposed strategic plans for the adoption of fully automated vehicles (Wu et al., 2022). The currently most widely recognized framework for automation levels was created by the

American Society of Automotive Engineers (SAE) (SAE International, 2021). These levels range from Level 0 (no automation) to Level 5 (full automation) and reflect the gradual shift of driving responsibilities from the human driver to the automated system (SAE International, 2021). A summary is provided in Table 1.

Table 1 Levels of driving automation assistance: Descriptions of automation levels by the SAE International specifications, based on the latest taxonomy and definitions document (SAE International, 2021, p. 25).

SAE Level of Automation	Description
0: No Automation	All aspects of the dynamic driving task are done by the human driver.
1: Driver Assistance	The driver assistance system (DAS) controls either steering or acceleration, while the human driver carries out the remainder of the dynamic driving task.
2: Partial Automation	One or more DAS perform both steering and acceleration, while the human driver carries out the remainder of the dynamic driving task.
3: Conditional Automation	The advanced automated driving system (ADAS) performs all aspects of the dynamic driving task, although the human driver is expected to intervene when prompted.
4: High Automation	The ADAS performs all aspects of the driving task, even when a human driver does not intervene upon requests.
5: Full Automation	The ADAS performs all aspects of the driving task under any roadway and environmental conditions.

SAE Level 3 automation represents a significant advancement since the system performs all driving tasks but relies on the human driver to intervene when prompted. The system monitors the environment and issues takeover requests, when encountering situations beyond its operational design domain, such as poor weather conditions or complex traffic scenarios. In this case, the role of the human driver shifts to one of a fallback-ready user (SAE International, 2021).

Takeover Requests in Semi-Automated Vehicles

Takeover requests (TORs) are crucial for ensuring safe transitions between automated and manual driving in SAE Level 3 vehicles (Jeon, 2019). As a prerequisite the system must properly identify limitations in novel situations or potentially dangerous environments. The Society of Automotive Engineers defines three causes triggering takeovers. Firstly, in the case of a vehicular failure, the driver is required to be fallback ready and recognize the need to intervene. Secondly, should the assistance system itself fail, a prompt is issued to the

driver to take on the dynamical driving task. Thirdly, boundaries can be derived from usage specifications, limiting assistance to geographical or environmental conditions (SAE International, 2021).

In the process of assuming manual control over a vehicle, the driver completes multiple steps, some of which in parallel (Ma et al., 2021):

- Situational awareness: Shifting attention to the road and away from non-driving related tasks.
- Motor readiness: Restoring mechanical control by placing feet on pedals and hands on steering wheel.
- Taking action: Decision making and assuming longitudinal and lateral control of the vehicle.

The takeover process strongly depends on a driver's awareness to the traffic environment (Li et al., 2017). Since the responsibilities of a driver change drastically in an ADAS-supported vehicle, mental and physical demands on the driver change as well (de Winter et al., 2014).

Non-driving Related Tasks

The use of in-vehicle entertainment systems and mobile devices are major sources of distractions at any level of vehicle automation (Politis et al., 2015). However, drivers are increasingly more likely to engage with non-driving related tasks (NDRTs) at higher levels of automation (de Winter et al., 2014). These NDRTs negatively affect the cognitive processing time necessary for situational awareness and factor into both takeover readiness and quality (Alrefaie et al., 2019; Wong et al., 2019). For example, drivers occupied with their phones demonstrate considerable worse braking reaction, reactions to traffic signals and lane keeping (World Health Organization, 2022). Even so, limiting a driver's responsibility to only monitoring the automation for prolonged periods also degrades takeover readiness (de Winter et al., 2014).

Since NDRTs are a part of drivers' habits, the administration of secondary tasks is common in driving simulation studies. Examples include: Inputting calculations (Agrawal & Peeta, 2021), games on a smartphone (Wong et al., 2019), variations of the Surrogate Reference Task (SuRT; ISO/TS 14198) (Gold et al., 2013), or the n-back memory task (Heikooop et al., 2018).

2.2 Design of Takeover Requests (TORs)

Inadequate warnings to drivers to resume control have been shown to be the predominant cause of safety issues in deadly, high-level automation accidents (Roche et al., 2019). If drivers are unsure of the underlying message of a warning, trust in the system decreases and the ADAS will be under-utilized (Agrawal & Peeta, 2021). Driver behavior can be influenced by design choices such as the modality of the alert, its timing, and characteristics like the assertiveness of the TOR.

Effectiveness of TOR Modalities

Takeover requests can be administered via visual, auditory, or haptic signals, or a combination thereof. Auditory TORs, particularly in urgent situation, provoke faster responses, such as braking (Gold et al., 2013; Politis et al., 2015). Auditory cues can either consist of speech, speech-based sounds, or non-speech-based sounds, e.g., beeps. Auditory alerts, whether speech-based or non-speech-based, prompt quicker responses than visual alerts, Jeon et al. (2019) found. In addition, auditory cues remain perceivable even without focused attention. Visual cues are usually administered on heads-up displays or additional screens near the dashboard (Merat et al., 2014). This means that drivers need to shift their gaze frequently to process visual TORs (Fagerlönner, 2010). Often, TORs are administered across multiple modalities. Most common is a combination of auditory and visual representations (Pakdamanian et al., 2022). In general, multimodal warnings result in faster reaction times over unimodal alerts (Politis et al., 2014).

Takeover requests may be presented in multiple stages to prime a driver before delivering a warning. However, there exists conflicting evidence for the effectiveness of multi-modal TORs. Agrawal and Peeta (2021) observed performance decrements, when an “uncertainty beep” preceded the actual TOR. However, Ma et al. (2021) observed increased situational awareness, less physiological stress, and longer time-to-collision (TTC). In this study, the phrases “attention please” and “takeover please” were played at a ten and seven second lead time and the sequence was counterbalanced across all trials (Ma et al., 2021). A promising field of research includes adaptive in-vehicle-alert systems. Driver demographic data and personal preferences could be used to administer individualized and highly effective alerts (Du et al., 2020).

Lead Times

Lead time describes the interval between when a TOR is issued and when the system disengages automated assistance. Its length must allow the driver to safely transition to manual control before a collision occurs (NHTSA, 2013). In general, drivers are able to achieve motor readiness within 1.2 and 1.8 seconds, but additional time is needed to regain situational awareness as well as physical and mental control (Gold et al., 2018; Zeeb et al., 2015). The most commonly used lead time in simulation studies are seven seconds, a duration first determined by (Gold et al., 2013) to be a safe window that enhances user acceptance of the automated system.

Reaction time may increase when drivers are aware of a long TOR lead time. They might consciously exploit the surplus of time with restoring situational awareness (Gold et al., 2013; Wu et al., 2022; Zeeb et al., 2015). Harder braking can be expected for a shorter lead time, which may be an attempt by drivers to gain more time for decision-making (Gold et al. 2013).

Assertiveness and Urgency in Takeovers

Apart from shortening lead times, driving simulation studies may opt to enhance the sense of urgency for participants by administering more intense takeover requests or designing critical traffic situations. The perceived criticality of a takeover scenario has been found to significantly shorten takeover times (Zhang et al., 2019). Although the urgency of a situation seems to be a less deciding factor compared to TORs, which were designed to be urgent through volume, intensity or multi-modality. (Politis et al., 2014).

Auditory warnings, especially signals designed for high urgency, have been shown to provoke faster and more forceful responses. Fagerlönner et al. (2010) found that drivers receiving high-urgency non-speech-based warnings braked harder and later compared to those receiving low-urgency alerts. However, participants also reported high-urgency warnings to be more startling and annoying. Biondi et al. (2014) further showed that a TOR in the form of an abrupt beep, caused rapid motor reactions like releasing the gas pedal and minor steering errors, affecting lateral and longitudinal control.

Assertive voice commands were also found to improve reaction times. Wong et al. (2019) tested assertive wording, such as "Danger" instead of "Please", in combination with action descriptions (e.g. "Take over control"). The authors found that assertive warnings led to faster takeover responses, even when drivers were engaged in distracting tasks. This is supported by an earlier study where the voice tones designed to be assertive and urgent,

grabbed driver attention more effectively (Politis et al., 2015). However, trust in the automated system does not increase with assertive voice commands or overly urgent alerts (Biondi et al., 2017; Wong et al., 2019).

Urgency can be enhanced by multimodal alerts, combining visual, auditory, or tactile signals. Politis et al. (2014) observed that TORs designed with high urgency across multiple modalities consistently led to faster reactions. To test this, the authors designed multiple combinations of colors for visual alerts, inter-pulse intervals for tactile alerts and frequency of sounds for auditory alerts. While assertive TORs can improve driver response times significantly, signals should be carefully designed to not impair driving performance. Overly abrupt signals could disrupt the drivers for up to 45 seconds after TOR onset (Biondi et al., 2014).

Measuring Driving Performance for Takeovers

Reaction time: Reaction time, often referred to as takeover time, is defined as the time from the onset of a TOR until the driver takes action. As a meaningful temporal metric, it helps to better understand a driver's ability to intervene (Gold et al., 2018). It is common practice to record it as a driver's first response through either steering or braking. Steering reaction time can be measured at an absolute steering wheel velocity of $>0.075^\circ$ and an absolute steering wheel angle acceleration of $>5^\circ/s^2$ (Zeeb et al., 2015), or at an relative steering wheel angle greater than $\pm 2^\circ$ (Gold et al., 2018). Braking reaction can be recorded, when the brake pedal travel exceeds 10% (Zeeb et al., 2015).

Braking force: Another critical measure, closely related to drivers' situation awareness, is braking force (Li et al., 2017). As one of the most fundamental driving maneuvers, braking behavior can indicate whether drivers act with foresight or reactively. In an analysis spanning 500 brake events, Li et al. (2007) identified reactive hard driving by drivers with low situational awareness with recorded brake intensities reaching upwards of $11m/s^2$.

Collisions: A key advantage of driving simulation studies is the possibility of testing safety-critical situations without putting participants in danger. Inadequate or delayed reactions can increase the likelihood of collisions, measured by the number of incidents or time-to-collision (TTC), which is defined as the proximity of the ego vehicle to other traffic participants. This variable can be used as a quality measure and predict safe driving behavior (Winner et al., 2015).

Speed variability: Speed variability reflects how a driver maintains or adjusts their speed. Excessive speed variability can indicate stressed driving, uncertainty or poor situational awareness, with heightened risk of accidents (Biondi et al., 2017).

Steering variability and lateral stabilization: In the event of an unexpected handover of control, drivers may take up to 40 seconds to stabilize lateral control, Merat et al. (2014) found. The authors noted that limited manual driving led to reduced attention, in turn increasing steering variability. When drivers were actively expecting automation to switch off, they adjusted the steering wheel less frequently and were more capable at regaining control over the vehicle (Merat et al., 2014). Moreover, a high steering variability has been found to coincide with driver stress (Paredes et al., 2018).

2.3. Stress and Cognitive Load in Driving Automation

Understanding the relationship between workload, stress, and driving performance is essential for improving safety in both manual and automated driving contexts. This subsection explores how workload, stress and driving performance interact, along with best practices for measuring driver stress.

First, some key terms are clarified. In this thesis, the term “stress” does not refer to its clinical definition, but as a proxy for emotional arousal. Emotional arousal can be measured physiologically, e.g. through heart rate or skin conductance (Alrefaie et al., 2019). It has been shown to increase and decrease with workload (DeGuzman et al., 2022; Riener et al., 2009), which is composed of all physical and mental demands of a task. Workload is highly individual, depending on the resources a driver can and chooses to allocate to a task (Riener et al., 2009). The subjective experience of these demands is referred to as mental workload (Charles & Nixon, 2019).

Relationship between Workload and Performance

The Yerkes-Dodson law can be used as a helpful frame of reference to describe the interaction between cognitive load and performance in the context of driving (Riener 2022). According to this law, performance follows an inverted U-shaped curve relative to emotional arousal levels. Optimal performance occurs when arousal is moderate. Both under-stimulation, evident for example in boredom, and over-arousal, e.g. during a period of excessive stress, result in performance decrements (Yerkes & Dodson, 1908). The driving context places fluctuating demands on the driver through the driving task and non-driving

tasks, affecting cognitive load, arousal, and, in turn, driving performance (DeGuzman et al., 2022). This effect is more pronounced in highly automated driving, where drivers are more likely to engage in secondary tasks or experience boredom due to reduced involvement in the driving task (de Winter et al., 2014; Heikoop et al., 2018).

It is important to note that stress is not inherently detrimental. Moderate, acute stress can enhance alertness and improve drivers' reaction times. However, chronic, or extreme stress levels can impair decision-making, reduce reaction times, and degrade overall driving performance (Kontogiannis, 2006). In automated driving scenarios, the transition from automated to manual control is particularly stressful. The differences in workload between automated and manual driving are even observable through physiological measures. Automated driving may lead to decreased heart rate variability, narrowed gaze dispersion, and less time spent scanning the road. (de Winter et al., 2014; Du et al., 2020).

Measuring Stress in Driving Automation Studies

Arguably, physiological signals are sensitive to latency and require careful processing to match recorded peaks to their causes. While self-reports on stress are highly subjective, they can enhance quantitative measures and provide valuable, immediate insights. Researchers may use the Visual Analogue Scale or variations to allow participants to quickly indicate stress levels (Cohen et al., 1983; Wijsman et al., 2010). However, sensors enable continuous monitoring without being-invasive to the participant or interfering with the driving task. Commonly used measures are summarized below.

Heart Rate (HR) and Heart Rate Variability (HRV): Heart rate (HR) and heart rate variability (HRV) are two primary physiological indicators used in driving studies. As cognitive load rises, HR increases, reflecting the body's physiological response to task demands. HRV, which measures variation in time intervals between heartbeats, is an indirect indicator of the sympathetic nervous system's activation. A high heart rate and low HRV are typically associated with heightened mental stress (Agrawal & Peeta, 2021). HR is considered as a more straightforward, nearly real-time indicator of workload changes. HRV can be analyzed for different time frames, the shortest recommended duration being 40 seconds (Jiang et al., 2017). Other researchers suggest a time window of at least two minutes to avoid confounding effects from the respiratory cycle (Shaffer et al., 2016).

Physiological indices like HR and HRV might not always effectively discriminate between different mental workload conditions, as these measures are highly interrelated (Heikoop et al. 2018). For example, HR is often strongly correlated with HRV's time-domain measure,

the standard deviation of the N-N interval (SDNN), and frequency-domain measure, the low-frequency to high-frequency (LF/HF) ratio, suggesting that these cardiovascular metrics might not independently reflect different levels of cognitive load.

Electrodermal Activity (EDA): Electrodermal activity (EDA), also known as galvanic skin response (GSR), measures the skin's conductivity, which increases with the activation of sweat glands under stress. EDA is typically used to assess emotional arousal and can provide continuous feedback on driver stress without interfering with driving behavior (Lohani et al., 2019).

Gaze Behavior and Eye-Tracking: Observing gaze allows researchers to understand how drivers allocate visual attention and assess the impact on situational awareness. Visual scanning patterns can be used to categorize drivers into risk-levels resulting in periods of inattention and slower reaction times (Zeeb et al., 2015).

Other Measures: Another less frequently used physiological indicator is electromyography (EMG), which tracks muscle tension and has been shown to correlate with stress levels during driving tasks (Wijsman et al.). An emerging avenue for driver state analysis is emotion detection using machine learning. This includes methods such as visual recognition of facial expressions and speech recognition for both verbal and non-verbal utterances (Khair Ooi et al., 2016).

2.4. Predictive Processing and Cognitive Science Paradigms

According to predictive processing, the brain is continuously involved in generating hypotheses about sensory input and updating these predictions through perception and action. The framework can potentially advance existing human factors research on driver behavior by linking perception, action, and attention within a dynamic process of error minimization (Engström et al., 2018). This subsection first provides an overview of the framework, before exploring applications in driving automation studies.

The predictive processing framework (PP) has been described as a radical framework, unifying processes of perception, action and attention (Clark, 2013). It might even provide a unified account of sub-fields in cognitive science that, traditionally, have often been studied in isolation. Clark (2013) views it as a systematic bridge between research methodologies, such as cognitive neuroscience, computational modelling, and probabilistic Bayesian approaches to dealing with evidence and uncertainty. For this

reason, researchers have also referred to PP as a “unified theory of the brain” (Friston 2010) or a package deal for cognitive science (Clark, 2013, 2017).

In contrast to the goal-and-reward framework employed in standard computationalism, PP argues that cognition is driven by experience-dependent expectations (Friston, 2010). Moreover, the framework’s central mechanisms can be applied to not only descriptions on the personal level (e.g. processes behind reasoning, phenomenological reports), but on sub-personal levels (e.g. neurobiological) of analysis. The authors Wiese & Metzinger (2017) believe that this versatility opens a new philosophical research field and define three areas where predictive processing could have a significant impact:

- “1. it may provide the means to build new conceptual bridges between theoretical and empirical work on cognition and consciousness,*
- 2. it may reveal unexpected relationships between seemingly disparate phenomena, and’*
- 3. it may unify to some extent different theoretical approaches.”*

(Wiese & Metzinger, 2017, p. 2)

Since the predictive processing framework is continuously evolving, different accounts may emphasize different features or interpretations. The following explanations focus on the most widely accepted attributes of PP.

2.4.1 Central Concepts

Hierarchical Prediction Error Minimization

The brain constantly processes sensory information to learn about the outside world and its own physical, physiological, and emotional states. Here, perception is a constructive rather than passive process, which closely interacts with action. Perception is largely driven by top-down processing, while the brain compares sensory inputs with its own predictions based on prior knowledge. However, a mismatch between the predicted and actual input results in a prediction error. The predictions operate within hierarchical models: Errors at the sensory level are propagated upwards a hierarchy of levels, meeting and updating predictions at each level. At the higher levels, abstract concepts, goals, or experiences shape the downward-flowing predictions. This process refines the internal models by minimizing the sizes of the prediction errors with each update. It also explains how an agent learns about new environments or efficiently moves through familiar ones (Clark, 2013; Engström et al., 2018; Wiese & Metzinger, 2017). For instance, a driver’s prior

experience with a vehicle will shape their driving behavior: Predictions might concern the braking force necessary to attain a safe full stop, based on the expected inertia from the vehicle's deceleration or anticipated changes in the visual field. However, the same actions in an unfamiliar vehicle can result in various prediction errors and prompt the driver to adjust accordingly.

Bayesian Inference

Bayesian inference models uncertainty in a probabilistic format, where probability is assigned to a finite number of causes, or represented in a probability density function for continuous values. New information can be used to reduce uncertainty: The a priori distribution (short: prior), containing the existing knowledge, is combined with the likelihood, a measure of how the new data relates to the prior. The resulting a posteriori distribution (posterior) often has a reduced variance and thus is more precise than the prior (Friston, 2010; Goekoop & de Kleijn, 2021; Wiese & Metzinger, 2017).

Even though Bayesian inference is not tied to prediction error minimization, combining the two can result in better approximations. Since sensory signals can arise from an indefinite number of external events, an agent often remains uncertain about the hidden states in the environment. Bayesian inference might allow to perform a “probabilistic inverse mapping,” linking sensory inputs to all possible hidden causes according to their probabilities. However, in complex cases, updates of Bayesian inference can be computationally challenging. Prediction error minimization provides a more efficient approximation, leading to more precise models on perceptual causes (Wiese & Metzinger, 2017).

The Free Energy Principle and Active Inference

Prediction error reduction and Friston (2010)'s Free Energy Principle (FEP), share some commonalities. An agent needs to sustain herself in the world based on sensory inputs, of which she cannot know the real causes, but can gain understanding of by learning and expecting their effects. In the event of a takeover, an agent experiences large surprisal: The internally coded recognition distribution diverges from the posterior distribution of true hidden states. The agent can minimize free energy by either actively sampling repeatedly or by increasing the internal model evidence. This engages the agent in a constant trial-and-error procedure, which hinges on the causal connections between agent and environment. Perceptual and active inference are thereby closely linked and not temporally separated: Perceptual discrepancies might elicit an agent to sample more information through

movement, while moving about the world changes an agent's perceptual inferences informing the recognition distribution (Friston, 2010; Friston et al., 2012).

Active inference involves the reduction of prediction errors by either adjusting internal models or taking action to alter sensory input (Friston et al., 2012). Perception and action thus are deeply intertwined: Predictions about the environment guide actions, and sensory feedback refines those predictions. This interplay leads to sensory attenuation, where the brain suppresses sensory feedback from self-generated movements, allowing for easier action execution. However, this can also lead to perceptual distortions, as demonstrated in the force-matching illusion, where insufficient sensory attenuation leads agents to misattribute sensory signals (Brown et al., 2013). However, active inference might be key to managing the perception-action-loop. Organisms not only adjust internal models but also act to align sensory inputs with predictions, ultimately resulting in better adaptations to exist in fluctuating environments (Parr et al., 2022).

Goal-driven Attention and Precision

Central motivators in predictive processing are the resolution of uncertainty and expected information gain (Parr et al., 2022). Actions which are expected to reduce uncertainty are prioritized. These opportunities for action stand out as affordances, they draw attention and are salient for attaining goals. Moreover, attention is additionally tuned by the so-called precision of predictions (Clark, 2012). When an agent has high confidence in a prediction, strong prediction errors arise for deviating outcomes. High-precision prediction errors capture attention and prompt more observation, while low-precision errors may be ignored. The violation of a high precision prediction will draw more attention and may lead the agent to sample for more observations (Deterding et al., 2022). In any situation, the brain may identify a number of desirable actions and compare them based on the desirability of predicted outcomes. Goal-driven acting is thus a navigation in an affordance landscape (Pezzulo & Cisek, 2016). The choice of a successful action sequence can furthermore elicit positive affect, caused by a faster-than-expected reduction of the prediction error. On the other hand, slower error reduction rates can lead to negative moods and contribute to a cautious decision-making process (Hesp et al., 2021; Kiverstein et al., 2019).

Estimated states and predictions are also affected by interoceptive feedback. In addition to controlling vital parameters like blood sugar or blood oxygenation, interoceptive predictive coding may constitute subjective feeling states by continually updating

prediction of the causes of interoceptive input (Seth, 2014; Seth & Friston, 2016). Predictions are shaped by generative models informed by “efference copies” of visceral, autonomic, and motor control signals. These are generated, compared and updated within a salience network anchored on the anterior insular and anterior cingulate cortices that engage brainstem regions as targets for visceromotor control and relays of afferent interoceptive signals (Seth, 2014).

2.4.2 Predictive Processing in Driving Automation Studies

Engström et al. (2018) were one of the first researchers to connect predictive processing to driving automation through previous, related human factors models that include prediction and expectancy. The authors introduced the framework for driving studies by focusing on three key mechanisms: active inference, precision, and model tuning.

In semi-automated driving, drivers transition from active control to monitoring the system, in which they might experience passive inference from proprioceptive senses. Engström et al. (2018) argue that drivers in automated driving are no longer actively involved in any tactical and operational decision-making. Consequently, the driver’s engagement with the task is altered, leading to reduced situational awareness and slower reaction times. Drivers who do not fully trust automation, might increase their monitoring frequency, while those who overly trust the system may become complacent with reduced attention (Engström et al., 2018). Studies show that a lack of engagement during passive monitoring can delay drivers' ability to respond effectively in critical takeover scenarios (de Winter et al., 2014). Engström’s research helps explain why drivers often struggle to quickly re-establish their mental models during the takeover process (Engström et al., 2018).

Wei et al. (2020) extended this perspective by examining how drivers react to visual looming, a measure of visually perceiving an object to be rapidly approaching. To do so, the authors modelled driver behavior with a computational, active inference model. Their study found that prolonged use of automation impairs drivers’ ability to predict and react to fast-approaching hazards, in turn compromising safe driving behavior. When drivers are “out of the loop”, their ability to judge speeds and initiate timely braking decreases, resulting in delayed responses during critical situations (Wei et al., 2022).

The researchers Kujala and Lappi (2021) focused on how drivers manage uncertainty, providing further insights into the role of attention in automated driving. Their model suggests that drivers adjust their attention and behavior based on the level of uncertainty

they perceive in the driving environment. When uncertainties are high, such as in complex traffic situations or during automation failures, drivers increase their information-gathering behavior. For instance, drivers may glance at the road or user interfaces in the vehicle more frequently, in order to reduce uncertainty in the form of prediction error minimization. On the other hand, overconfidence can lead to reduced vigilance and slower responses when intervention would be required. This balance between overconfidence and cautiousness is crucial for driver safety, especially in scenarios where automation has not adequately accounted for unexpected events or hazards (Kujala & Lappi, 2021).

In conclusion, the predictive processing framework offers a comprehensive way to understand the complex dynamics of driver mental models and behaviors. While research has explored predictive processing for automated driving studies from multiple perspectives, the ongoing advancements in both fields will open new avenues for future research.

3. Methodology

The experimental design and user interface were developed and refined using a two-step approach, similar to the methodology of Biondi et al. (2014). The pilot study aimed to empirically identify the most assertive and least assertive takeover request sounds, based on those used in previous research. However, the pilot study proved to be crucial in optimizing the overall test setup for the main study. This section first describes the pilot study, including its design and findings. It then presents the participants, materials, and procedures used in the main study.

3.1 Pilot Study

The aim of the pilot study was to assess the stress associated with verbal and non-verbal auditory alerts, a common practice in TOR-design studies (Biondi et al., 2014; Marshall et al., 2007; Politis et al., 2015; Wong et al., 2019). A within-subject study was chosen to evaluate a total of six designed TORs, both in the context of a driving task in the simulator and outside of the simulator.

The participant group consisted of six usability experts (three female, three male) from the Austrian Institute of Technology. Their areas of expertise included Virtual Reality design, stress research, teleoperation research, and collaborative technology design.

In a single-user test format, each participant first completed a preliminary sound rating to evaluate the designed cues based on perceived urgency, annoyance, and alerting effectiveness, using a 10-point Likert scale. Previous research (Engström et al., 2018) found that alerts were perceived to be less annoying in a driving scenario compared to out-of-context evaluations. Therefore, the pilot study included a driving task in the simulator, during which participants were again presented with the highly assertive cues.

Participants first became familiar with an n-back-task, manual driving, and automated driving in a learning scenario. The driving scenario in the actual pilot study consisted of four

laps, each including sections of automatic driving, takeover tasks, manual maneuvers around parked vehicles and obstacles to induce a safety-critical event (Ma et al., 2021), and driving on an unobstructed road segment. Fog was added to the simulation, so that participants could not see any obstructions ahead and possibly expect the takeover, before hearing the request. After each lap, an edited version of the Perceived Stress Scale was carried out, asking participants to indicate their stress levels on a scale of 1 to 10. At the end of the drive, participants filled in a NASA-TLX questionnaire to assess their workload. A final interview with open questions led to qualitative feedback on the visual and auditory user interface, the quality of the sounds, feedback on the stressfulness of sounds, scenario, and study set-up. An itemized procedure of the pilot study can be found in Appendix A.

Evaluated Alerts

The study considered both verbal and non-verbal cues, exploring the impact of information density on perceived assertiveness. While most studies have relied on non-verbal cues (Agrawal & Peeta, 2021; Biondi et al., 2014; Fagerlön, 2010), Wong et al. (2019) found that perceived assertiveness can also arise from verbal commands. Based on this insight, the pilot study explored verbal takeover requests with low and high levels of information density, see Table 2.

Table 2 Alerts tested in the pilot study: All sounds evaluated in the pilot study were based on TORs used in previous research and a real-world example.

Nr.	Type	Description
<i>Low-assertive takeover requests</i>		
1	Non-verbal	Low-pitched chimes based on BMW alarm for 5s (BMW Group, 2022).
2	Non-verbal	Low-pitched beeps at 179 Hz for 4.5 Seconds, see (Fagerlön, 2010)
<i>High-assertive takeover requests</i>		
3	Non-verbal	High-pitched beep at 2000Hz for 4.5 Seconds, see (Biondi et al., 2014)
4	Non-verbal	Dissonance of six tones (B4, C5, D5, C6, B6) for 4.5 Seconds, see (Fagerlön, 2010)
5	Verbal	High information density plus high-pitched Beep plus Dissonance for 7s: "Achtung, Unfall voraus! Bitte übernehmen Sie die Kontrolle!", see (Wong et al., 2019)
6	Verbal	Low information density plus high-pitched Beep plus Dissonance for 7s.: "Achtung! Übernehmen Sie die Kontrolle", see (Wong et al., 2019)

Non-assertive warning sounds in vehicle user interfaces are less discussed by previous research. To address this gap, the pilot study included one non-assertive sound derived from literature as a counterpart to a high assertive sound (both from Fagerlönner et al., 2010). Additionally, a BMW alarm sound was used to represent real-life auditory interfaces (BMW Group, 2022). This choice aimed to provide a real-world context to the study's findings, to ensure that both assertive and non-assertive sounds were represented.

Expected Results

Although the initial experiment was largely exploratory, several hypotheses for the pilot study were formulated to guide the evaluation:

- HP1: High assertive sounds (3, 4, 5, 6) will be rated as more stressful, annoying, and urgent than low assertive sounds (1, 2).
- HP2: The sound with low information density (Sound 6) will be perceived as more stressful than the sound with high information density (Sound 5).
- HP3: The sound with high information density (Sound 5) will show the highest perceived urgency.
- HP4: Biondi's high-pitched beep (Sound 3) will receive the highest annoyance ratings.

Quantitative Results

The pilot study rejected HP1, as sound Nr. 4, which was intended to be highly assertive, was rated less assertive than the low assertive sounds Nr. 1 and Nr. 2 ($M=2$ on a scale 1-10). However, all other high assertive sounds were perceived to as more urgent and stressful than any of the low assertive sounds. Sound Nr. 1, the BMW-chime, scored particularly low for mean urgency, stress and annoyance ratings at $M=1.94$ ($SD = 0.94$), while sound Nr. 2, the tone used in Fagerlönner et al. (2010), had a slightly higher mean score of 2.06 ($SD = 1.11$) with most ratings below 3 points.

HP2 was also rejected, with no significant difference between the two verbal takeover requests found in a Wilcoxon rank-sum test ($p = 0.93$). HP3 was confirmed, as sound Nr. 6 was rated the highest for perceived urgency ($M=8.33$, $SD=2.07$), exceeding the ratings for the other verbal cue in sound Nr. 5 ($M=7.33$, $SD=2.42$). Lastly, HP4 was rejected: Regarding the feature annoyance, sound Nr. 3 ($M=7.5$, $SD=2.345$) received lower ratings than the verbal commands (Nr. 5: $M=9.17$, $SD=1.6$; Nr. 6: $M=9.17$, $SD=1.33$).

Higher Stress Ratings in Preliminary Part

Contrary to expectations, sound Nr. 5 (high information density with high beep and dissonance) was rated more stressful than its low information density counterpart (Nr. 6). It received the highest stress (M=9), urgency (M=10), and annoyance (M=10) ratings. All ratings are visualized in Figure 1.

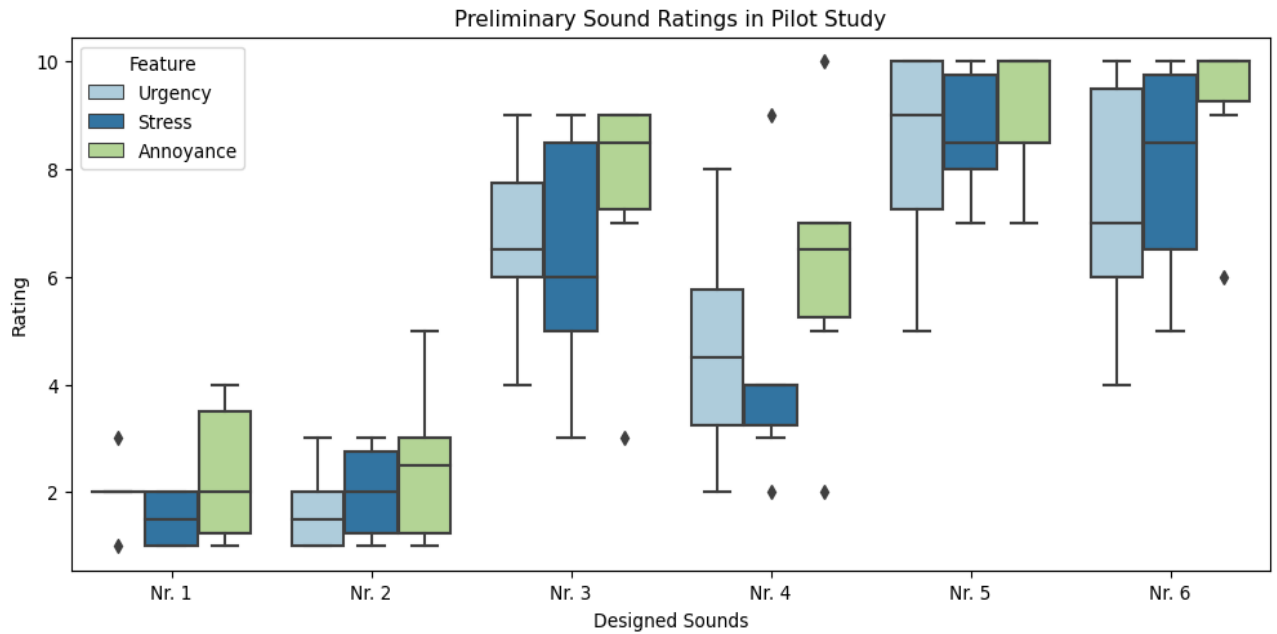


Figure 1 Ratings on perceived assertiveness in pilot testing: As expected, low assertive cues were perceived as less annoying, stressful and urgent than the high assertive cues.

Moreover, a paired t-test revealed a significant difference ($p = 0.047$) in high assertive sound ratings between preliminary and simulator contexts. Ratings for annoyance, stress, and urgency were higher in the preliminary rating ($M=6.83$) compared to the driving task ($M=5.33$). During the driving task, participants rated takeovers prompted by sound Nr. 3 as most stressful ($MD=7$).

Perceived Stress and Time Pressure Correlate Positively

The NASA workload assessment administered in the pilot study revealed a high variability in perceived stress ($M=4.5$, $SD=4.5$ on a 0-20 scale). Participants felt confident about success of task completion ($M=18$, $SD=1.9$) with low perceived combined mental and physical effort ($M = 7$, $SD = 5.29$). Ratings show that mental effort ($M=10.83$, $SD=4.26$) exceeded physical effort ($M=3$, $SD=1.79$). Time pressure varied widely among participants ($M=8$, $SD=6.29$), with four participants rating it as low (less than 8) and two as high (14 and 17). Moreover, positive correlations between time pressure and stress ($r = 0.56$) and between mental effort and stress ($r = 0.58$) could be identified.

Qualitative Results

Interviews revealed that participants' preferences and subjective perceptions varied widely. The verbal alerts for sounds Nr 5 and Nr. 6 were rated as quite stressful in preliminary evaluations ($M=8.667$ and $M=8$), but some participants reported that they were useful and reassuring during the driving task (stress rating: $M=4.667$ and $M=5.5$). Moreover, verbal signals contributed to a sense of security, particularly in the test setting with an unfamiliar user interface. However, some participants felt verbal signals were confusing or redundant. One participant was unable to differentiate between sounds 5 and 6, leading to increased annoyance. While four participants found the high-pitched beep (Nr. 3) irritating, exhausting, or disruptive, one viewed it as less forceful than verbal signals. The dissonant sound (Nr. 4), although perceived as alerting, was easier tolerable overall. In general, the highly assertive sounds were viewed as annoying but effective in capturing attention.

Requests for more Challenging Obstacle Courses

The pilot study's course, with predictable alerts and increasing familiarity, led to reduced stress and immersion. "The element of surprise was not there," one participant noted. Participants reported stress spikes during the moment of takeover, followed by a steady decrease. To make the course more challenging, they suggested placing obstacles closer together, introducing dynamic elements, and adding auditory distractions such as background sounds or aggressive music. Additional recommendations included varying driving conditions (e.g., rural roads, night driving) and incorporating more complex scenarios like traffic jams. To heighten the sense of urgency, participants could be encouraged to complete the course as quickly as possible. Extending the automated driving phase and varying the intervals between automated and manual driving were also suggested to improve immersion.

Visual Elements Lacking Realism

Participants reported a lack of realism due to the static GPS map on screen and the missing rearview mirror in the interface. These omissions contributed to participants' nervousness during the task.

Adjustments

The pilot study highlighted multiple areas of improvement for the design of the obstacle courses. On two occasions, participants swerved off the road to avoid the intended

obstacle course completely. Interviews also showed that the predictability needed to be reduced, for instance varying levels of difficulty of the courses. Key improvements included adding an auditory indicator to signal when the autopilot resumed control after each lap and making the intended route through the obstacle course more intuitive, while making the courses more challenging. Among the sounds tested, the high-pitched beep was deemed the most effective for high-assertive signaling due to its strong urgency and annoyance ratings, coupled with lower variability compared to verbal commands. Methodological changes were made, including clearer task phrasing (e.g. "Please navigate through the course" instead of "around") and key user interface elements were integrated into the learning period without revealing too many details of the main study.

3.2 Main Study Design

The effects of TOR-assertiveness were explored with a driving simulation study. This method was chosen because it can produce a sense of immersion similar to real-world driving experiences without exposing participants to the inherent risks of real-world driving environments. Insights from the pilot study highlighted the need for more challenging driving scenarios, leading to the development of six obstacle courses with varying levels of difficulty. In addition, the TORs were issued at irregular intervals, ranging from thirty to ninety seconds of driving with automation.

Given that emotional arousal and stress can significantly impact driving performance, the experiment collected both objective and subjective measures, including physiological stress responses, perceived stress ratings, driving performance metrics, and perceived performance ratings. Additionally, a semi-structured interview was conducted to gather participants' subjective evaluations of TOR-assertiveness, aiming to bridge the gap between quantitative user data and subjective perceptions.

Independent Variables

Following the results of the pilot study, two levels of assertiveness for the TOR sound were determined for the main study:

- *Low-Assertive Sound:* This was determined as sound Nr. 1, the BMW chime, which was selected due to its low ratings in perceived annoyance, urgency, and stress.
- *High-Assertive Sound:* This was determined as sound Nr. 3, a high-pitched beep, selected for its significant ratings in both the preliminary and simulator evaluation.

Dependent Variables

The study focused on several dependent variables to capture the effects of TOR-assertiveness for both an immediate and sustained period after takeover. Descriptions of the variables and corresponding hypotheses can be found in Table 3.

Table 3 Dependent variables with descriptions and units: For physiological measures, baselines were defined as average values of the preceding auto-drive period. Hypotheses 1 – 7 contain measures for the immediate period, while hypotheses 8 – 13 pertain to the sustained period, and hypotheses 14 and 15 address subjective measures after each lap. Adapted from Kraut et al. (2023, p.7).

Measure	Hypotheses	Description
<i>Performance Measure</i>		
Reaction time	H1	Time (s) between TOR signal and first reaction (steering >2°, braking >0.5 N, or pressing TOR button)
Braking force	H2	Maximum brake pedal force (N) in the 5 seconds after TOR onset
Collision rate	H3	Ratio between hit and all obstacles
Average speed	H4, H8	Average speed (km/h) during selected period
Speed variability	H9	Std. deviation of speed (km/h)
Steering variability	H10	Std. deviation of steering wheel angle (°)
<i>Heart rate measures of stress</i>		
Heart rate ratio	H5, H11	Ratio of average heart rate to baseline
SDNN ratio	H6, H12	Ratio of std. deviation of the inter-beat interval to baseline
RMSSD ratio	H7, H13	Ratio of root mean square of successive differences between beats to baseline
<i>Subjective measures</i>		
Stress	H14	Perceived stress (VAS 1 – 3)
Performance	H15	Perceived performance (VAS 1 – 3)

3.3 Participants

A total of 32 participants completed the experiment in a single-user setting. All participants had normal or corrected-to-normal vision and a valid driver's license. The group included five employees from the Technology Experience unit at AIT, while the remaining 27 participants were recruited from AIT's user tester pool or through friends and family. The initial distribution included 16 female and 16 male participants between the ages of 24 and 57 ($M=36.9$, $SD=10.99$). Due to incomplete data recordings, one participant was excluded, resulting in a final sample of 31 participants (16 male, 15 female) with ages ranging from 24 to 57 ($M=39.2$, $SD=10.75$).

3.3 Apparatus and Equipment

Driving Simulator Set-up

Heads-Up Display

The study featured a heads-up display (HUD) that provided participants with pre-programmed visual information. It displayed ADAS activation status, the vehicle position on a map, speed, gear, and battery charge. During takeover requests, a red light illuminated part of the screen to further emphasize the TOR. All graphical user interface elements for the HUD were created using Figma, a web-based design tool (Figma, 2024). The figures below illustrate the various user interfaces used throughout the study.

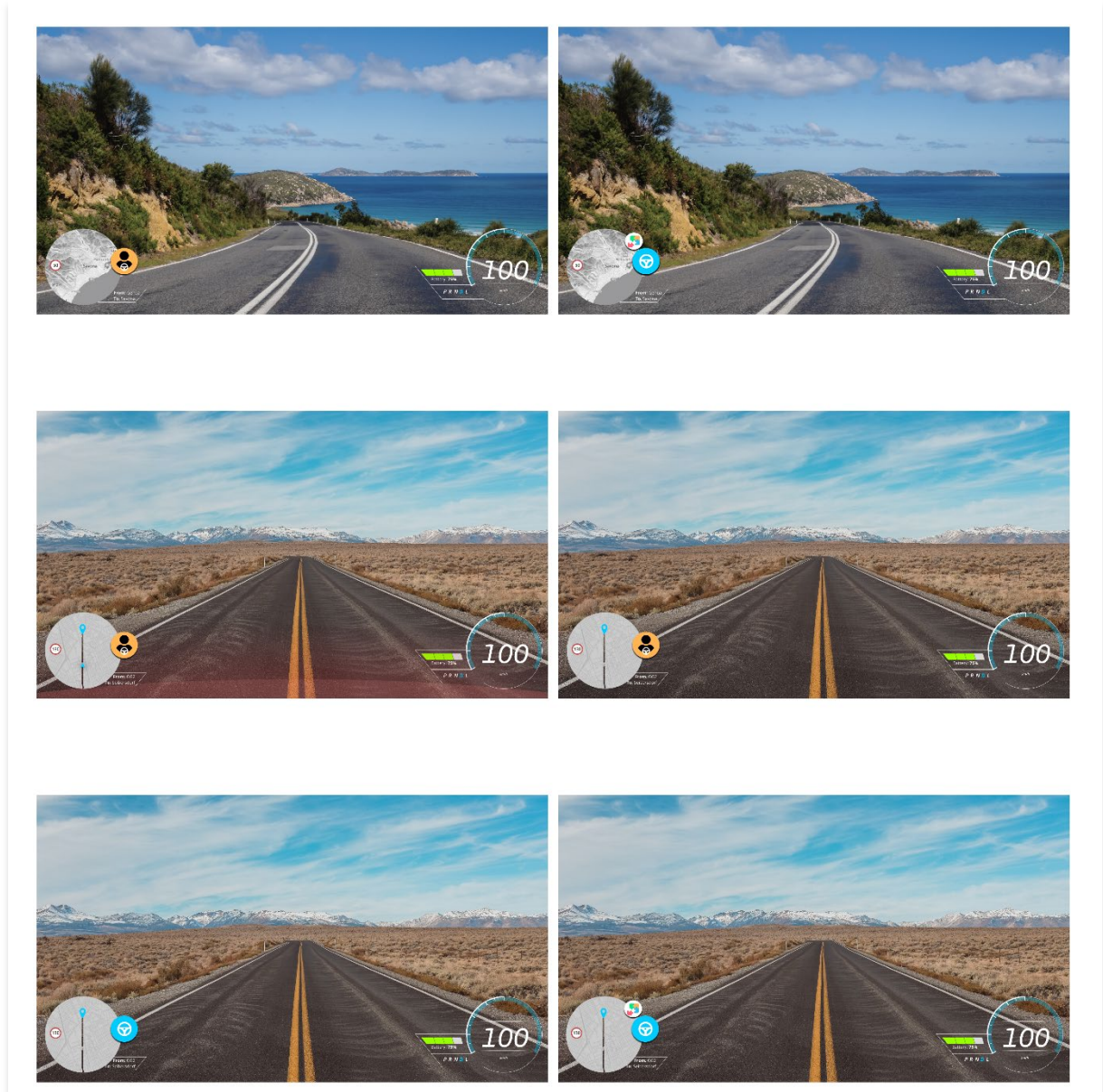


Figure 2 Final versions of the visual user interface for the learning and test settings: Top row was used in the learning period (left: Manual driving, right: automated system driving, reminder to play game). The middle row shows the interfaces used in the experiment for manual driving during a takeover request (left), and for the rest of the manual driving period (right). The bottom row contains the heads-up display indicating automated driving (left) and automated driving with a reminder for the n-back-game (right).

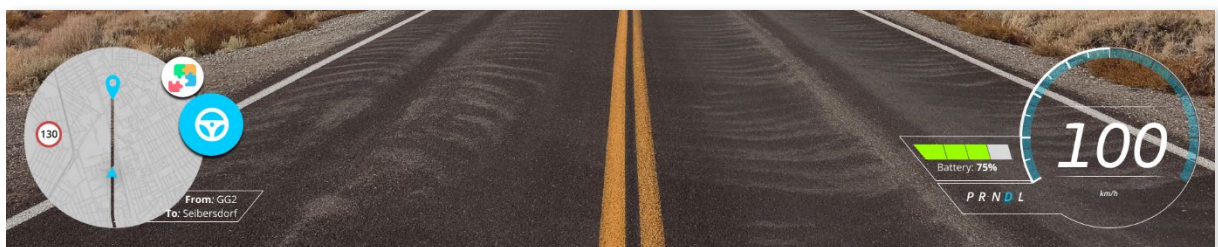


Figure 3 Close-up of the mock-up for the visual user interface in autopilot: The blue icon indicated automated driving, while the puzzle-icon reminded participants of playing the secondary task. A moving blue arrow marked the current position on the route.



Figure 4 Close-up of the heads-up display during manual driving during a takeover request: A red tint at the bottom of the screen served to draw attention to the heads-up display.

Tablet with Non-driving Related Task (NDRT)

Both the pilot and main studies included a Non-driving Related Task (NDRT), consisting of a game with a n-back task on a hand-held tablet. This game was chosen to simulate real-world distractions from devices, a common issue for driving with automated assistance. In the pilot study, a visual n-back task at difficulty level 1 was used but found to be too simple to effectively engage participants. Consequently, the main study included a more challenging audio-visual 2-back task.

This later n-back task involved a 4x4 grid where grey squares were sequentially highlighted in blue, following a random order. Participants were required to monitor this sequence and press a button every time a square turned blue in the same position as it had been two steps earlier. Simultaneously, random letters were played aloud, prompting a different button press when a letter repeated. This setup created a multi-sensory challenge with visual, auditory, and cognitive components, simulating the multiple demands that may arise from interacting with devices or conversing with passengers in a semi-automated driving scenario.

3.3 Procedure

The study contained six rounds of manual takeovers whereby the high assertive and non-assertive sound were administered three times each. Participants completed the experiment in single-user settings. Below, the procedure, scenario, and obstacle course types are explained in detail.

Preparation

Introduction and Consent: Participants were provided with information about the process of the study and signed consent forms, authorizing the anonymized use of their data.

Baseline Measurements: Afterwards, participants were outfitted with physiological sensors. For the physiological baseline measurement, participants stood looking in front of a grey wall for two minutes. Additionally, they completed the 10-item Perceived Stress Scale.

Learning Period: Participants completed a test drive in a separate scenario to familiarize themselves with the vehicle's functions, user interface, and takeover procedures. Instead of the TOR-sounds, participants were prompted to take over control by the study investigator. They also practiced the dual n-back game as a secondary task and switching between the hand-held tablet and steering. The main study began only after participants verbally confirmed their comfort with all these tasks.

Main Driving Task

Scenario Description: Participants drove on a three-lane highway with fog and activated ADAS. The automated speed was set to 120km/h and periods of automated driving ranged from thirty to ninety seconds to reduce predictability of the TOR. Traffic density was simulated with 450 vehicles per hour, or approximately of 15 vehicles per kilometer on the road (see (Gold et al., 2013)). The other vehicles produced background noise when they approached and overtook the vehicle during automated driving. Participants had been instructed to always engage with the secondary task, up until they heard a TOR.

Takeover Requests and Obstacle Courses: In total, each participant heard the high assertive and low assertive sound three times each. The six laps contained the same structure, a schematic of one lap is shown in Figure 5 below. Upon hearing a TOR, participants received a five second lead time to transition to manual driving, before navigating through the obstacle course. The obstacle courses, described in the "Obstacle Courses" subsection, varied in difficulty. After completing the obstacle course,

participants drove 1km in the middle lane. At that point, the lap concluded, and the simulation halted to present the Visual Analogue Scales.

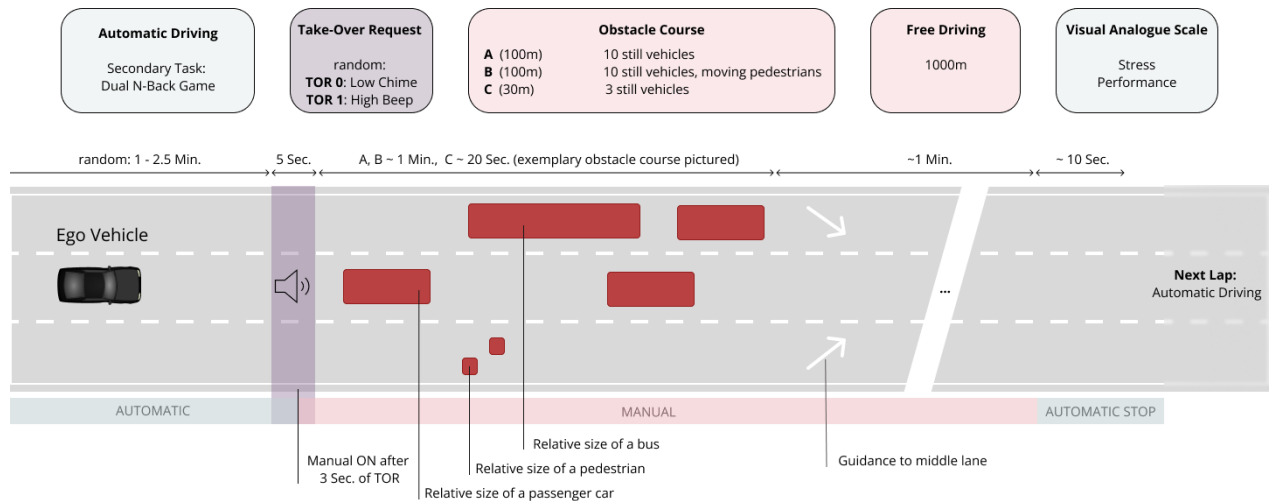


Figure 5 Illustration of a lap with time estimates. Adapted from Kraut et al. (2023, p. 6). The lap begins with a period of automated driving, during which the participant plays a game. Three seconds after receiving a TOR,, the ADAS disengages, and the participant must manually navigate the vehicle through a critical situation and continue driving on an unobstructed road.

Post-Lap Measure: After each lap, participants were presented with two Visual Analogue Scales (VAS) on the monitor to rate their stress and performance. Following these ratings, the simulation proceeded to the next lap. The simulation re-started at the position, where the vehicle halted earlier, and the re-engaged ADAS accelerated the vehicle back to a speed of 120km/h. This cycle was repeated for a total of six laps.

Ending

Questionnaires and interview: Sensors were removed and cleaned. Participants then completed the NASA TLX questionnaire, an adapted Perceived Stress Scale focusing on self-efficacy related to the driving task, a final Visual Analogue Scale on stress, and a demographic questionnaire. The study concluded with a semi-structured study interview containing seven questions.

Compensation: Participants completed a form to confirm their attendance and qualify for compensation from the Austrian Institute of Technology.

Post-Study Procedure: After each user test, all generated datasheets, forms, and recordings were exported and labeled with anonymized identifiers. Demographic information and compensation forms were scanned and filed separately to prevent any links between the study data and personal information.

Further Scenario Details

Several points are relevant to mention. The setting was based on previous research, since highways with a limited field of view is a common setup for simulator studies (Agrawal & Peeta, 2021; Heikoop et al., 2018; Ma et al., 2021). The straight road, lack of intersections and constant speed limit allow for a controlled environment. Other vehicles were simulated to create more immersion in the driving scenario (Gold et al., 2013). However, traffic density was turned to zero shortly before the moment of a TOR, to avoid unplanned distractions at the takeover or during the obstacle course.

Additionally, varying obstacle courses were created in response to the results of the pilot study. The surveyed experts requested a bigger challenge, resulting in more course variability, less distance between obstacles and a reduced TOR-lead time – from seven seconds in the pilot study to five seconds in the main study. Previous research showed that shorter lead times have been found to effectively induce higher workload (Du et al., 2020; Gold et al., 2013).

Each participant received the high and low assertive TORs in random order. This aimed to mitigate possible side effects from the increased course type variability. The reaction time for a takeover was recorded, once participants either pressed a dedicated button on the steering wheel, turned the steering wheel by more than two degrees or pressed down on the brake.

Obstacle Courses

Three types of obstacle courses with distinct levels of difficulty were designed for the main study. Screenshots from the simulation software SCANeR application can be found in Figure 6.

Course A: Moderately difficult, featuring ten static obstacles (vehicles and objects) spaced 10 meters apart.

Course B: Most difficult, containing ten static obstacles and pedestrians walking on the sidelines. The dynamic element was added after collecting feedback on course type design in the pilot study.

Course C: Easiest course, with three obstacles requiring minimal maneuvering. This level of difficulty was added in response to the pilot study.

The succession of obstacle courses was as follows:

- Lap 1: Course A
- Lap 2: Course A mirrored on y-axis for reduced predictability
- Lap 3: Course C
- Lap 4: Course B
- Lap 5: Course C mirrored on y-axis
- Lap 6: Course B mirrored on y-axis

Obstacle Course A



Obstacle Course A-mirrored



Obstacle Course B



Obstacle Course B-mirrored



Obstacle Course C



Obstacle Course C-mirrored



Figure 6 Obstacle course layouts: Pictured are snapshots of the obstacles courses as designed in the SCANeR software. Course A was moderately difficult, while Course B was highly difficult and safety critical due to moving pedestrians, type C was the easiest category with least required.

3.6 Data Collection

Driving performance data from SCANeR driving simulation software

The driving simulation was created and monitored with the software SCANeR. Performance metrics were tracked at twenty Hertz, recording every 0.05 seconds. A custom script was developed to define specific variables and extract relevant measures aligned with the study's hypotheses. The driving data for each participant were exported into individual Excel spreadsheets, allowing for an organized and participant-specific data analysis.

Instead of recording standard time, SCANeR used an internal clock that restarted with each recording session. This presented a challenge in synchronizing data with the physiological signals. An additional custom script was developed to write real-time timestamps according to Coordinated Universal Time (UTC) at key points during the SCANeR recording.

iMotions for Video Recordings, Questionnaires, and Synchronized Physiological Data

The iMotions platform (iMotions Inc., 2021) was utilized to collect physiological data and to structure each user test. This platform allowed to connect sensors via Bluetooth for real-time monitoring and enabled the synchronization of physiological data with screen recordings of the driving simulation. It also allowed it to create distinct segments within the user study, facilitating a structured data analysis along several specified segments:

- **Baseline Measurements:** A two-minute-long recording to capture baseline physiological data.
- **Learning Period:** At first, participants became familiar with automated and manual driving, as well as the takeover process and n-back game in a dedicated learning scenario. The recording of the learning period was later excluded from analysis.
- **Laps 1 – 6:** In the main study, the sensor recordings were synchronized with screen recordings of the SCANeR software. Each lap was recorded in an individual segment and post-experiment, key periods were annotated manually in the recordings. These include: the moment a TOR was issued, the start of the obstacle course and the start of the free driving period.
- **Visual Analogue Scales (VAS):** The VAS were displayed on screen after each lap segment, allowing participants to select their responses from the driver seat and

using a computer mouse. Subjective stress ratings were assessed this way before and after the main experiment, additionally.

This setup resulted in easily accessible video recordings for each participant, along with exportable datasheets containing physiological sensor recordings, recorded questionnaire responses and the manually defined key moments in the study.

Heart Rate Measures with the Polar Heart Rate Monitor

Physiological signals were collected using the Polar H10 heart rate monitor, which was attached to a chest strap (*Polar H10* | *Polar USA*, 2024). Following the manufacturer's guidelines, the conductive section of the strap was dampened before use to enhance signal quality and was positioned at the base of the sternum. To ensure hygiene, the chest straps were washed and disinfected after each use.

Extended Perceived Stress Scale

The Perceived Stress Scale (PSS) is a ten-item questionnaire which produces a score based on questions related to perceived self-efficacy and perceived helplessness (Cohen et al., 1983). At the beginning of the trial, non-native German participants received the original PSS, while native German speakers used a translation provided by (Schneider, 2020). The pre-task administration of the PSS assessed how participants were affected by general life stressors. During the study, participants filled out a modified version of the PSS, to assess their self-efficacy in managing the task stressors. The complete PSS can be found Appendix B.2 in German and in Appendix B.3 in the original English version.

Visual Analogue Scales

Visual Analogue Scales are a common method for assessing mental stress in simulator studies. (Du et al., 2020; Wijsman et al., 2010). After completing each lap, participants rated their stress levels and perceived success in performing the task by moving a slider on the scales (Figure 7). These forms were created within iMotions and participants were administered with either the German or English version, according to their native language.

Figure 7 Visual Analogue Scales for self-reported ratings. The Figure shows a screenshot for a German VAS from the experiment. Participants placed the slider along the scale to indicate the extent of perceived stress and performance.

NASA Task Load Index (TLX)

The NASA Task Load Index (NASA TLX) was used to measure perceived workload during the driving task. This multi-dimensional rating scale assesses workload across six dimensions: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. It is widely used in human factors research to evaluate the subjective workload, making it a valuable tool for understanding the mental and physical effort required by the participants in the study (Hart & Field, 2006; *NASA TLX: Task Load Index*, 2024).

Semi-Structured Interview

The experiment concluded with an interview consisting of seven guiding questions, allowing participants to share their experiences, feedback, and final thoughts. This final step provided valuable insights into the participants' perceptions of the tasks and the user interface, as well as possible discrepancies between the intended design and the actual user experience. According to the individual participant's native language, questions were either posed in English or in German. The interview guideline was as follows:

- Could you briefly describe how you felt today while performing the navigation tasks?
- Could you summarize the tasks you were required to complete?
- How did the system signal you to take over control?
 - Can you describe key features of the acoustic signal(s)?

- Please listen to the following two audios and confirm if you remember hearing one, both, or none of these signals during the experiment.
- Which acoustic or visual signals from the interface did you find particularly helpful during the takeover process?
- Was there any information you would have needed during the takeover process that was not provided?
- Were there any signals that you would have preferred to be presented differently?
- Do you have any concluding thoughts, comments, or questions?

3.7 Data Analysis

During the initial examination of the data sets, it was found that one participant's physiological recordings were incomplete. After excluding this dataset, data by thirty-one participants were analyzed (16 male, 15 female, mean age = 39.2 years, SD = 10.75 years). To ensure anonymization of the participants, each participant was assigned a test number, and any identifying information was removed.

Quantitative Data Preparation

The first step in the data preparation involved assessing quantitative data from multiple sources: SCANeR datasheets, iMotions exports, and accompanying text files for time stamps. Using the UTC timestamp-export, the driving performance data and iMotions data were synchronized for each participant with the help of a Python script (Python TM, 2024). Key points for each lap of the experiment were identified to calculate scores for specific periods of interest, such as the takeover process, obstacle course, free drive section, and automated driving section. It was found that the physiological baseline data recorded before the start of the experiment were not suitable, as the majority of stress measures were higher than those recorded during the experiment. This suggests that participants were more nervous before sitting down in the driver seat than during the experiment. Therefore, physiological baselines were determined from the automatic driving periods during the main study.

Next, composite data frames were created to integrate relevant variables from both the physiological and driving performance metrics. Additional columns were added for the NASA Task Load Index (TLX) scores and all Visual Analogue Scale ratings, which contain subjective measures on workload, perceived task performance and stress. For easier

interpretation, the Visual Analogue Scales were normalized to a scale from 0 to 1 (previously 1 – 3). The Python script was automated to efficiently pre-process each of the participant's data. The result was thirty-one data frames, exported into separate CSV-files, each containing all relevant quantitative data for any individual participant.

Statistical Analysis of Quantitative Data

After pre-processing all data sources, all individual datasheets were combined in full data frame in a second Python script. This script was developed to apply statistics to the data. A within-subject comparison of all variables, averaged on TOR-assertiveness was performed. The paired aggregated scores for the low- and high-assertive TOR were further compared with the Wilcoxon signed-rank test, which is suitable for paired non-parametric data. Rosenthal's formula was then used to calculate effect sizes, to determine the extent of the observed differences.

Since these aggregate scores cannot very well capture dynamic changes throughout the whole experiment, linear regressions on the number of laps were conducted for all relevant variables. This made it possible to observe potential learning effects and to discuss adaptation effects.

Qualitative Insights

At the end of the study, participants provided valuable insights during concluding interviews. Seven open-ended questions were posed, resulting in interviews of five to ten minutes length. Mayring's techniques for systematic content analysis were applied. In this process, the source material is being summarized in a series of steps until only significant content remains, from which a corpus of categories can be abstracted (Mayring, 2015).

Initially, all questions from each participant were paraphrased individually. Then, the responses were summarized for each interview question to identify recurring themes. The key themes and main points of the post-study interviews were subsequently integrated into the results section. This analysis allowed to better understand participants' subjective experiences and perceptions, which helped to contextualize the quantitative data in the discussion.

4. Experimental Results

During the main experiment, participants completed six laps, each containing three types of obstacle courses presented in the same sequence. In total, 186 laps were analyzed across 31 participants, with each TOR type (high-assertive and non-assertive) administered 93 times. In this section, each of the hypotheses of Research Question 1 is examined through driving data, physiological data, and questionnaire responses. Additionally, insights into potential learning effects and other findings are discussed, followed by a summary of the feedback obtained from interviews.

4.1 Quantitative Results

The first research question was as follows: “How does TOR-assertiveness affect driver performance and stress during a takeover and a subsequent driving task?”

To explore this question, a total of 14 hypotheses were formulated. These were further categorized into three distinct segments:

- The immediate period, which included the takeover and the drivers’ performance in a critical situation in a simulator study.
- The sustained period, addressing the manual driving over a longer period.
- Participants’ retrospective self-reports on the task demands with a focus on stress and success of performance.

The following quantitative results are presented according to these segments. Notably, significant results were found for the immediate timeframe, while no significant effects were observed in the other segments.

4.1.1 Immediate Effects of Assertive Takeover Requests

The immediate post-TOR measures include the hypotheses 1 – 7, containing performance metrics such as reaction time, braking force, collision rate, and average speed as well as physiological measures including heart rate (HR) ratio, standard deviation of NN intervals (SDNN) ratio, and root mean square of successive differences (RMSSD) ratio. Table 4 contains the aggregate scores for both non-assertive and assertive TORs calculated for this time frame.

TOR-assertiveness showed a significant influence on reaction time, with assertive TORs prompting participants to intervene by more than half a second earlier, on average (effect size $r = 0.48$, $p < 0.01$). In addition, drivers applied approximately 15% more braking force (effect size $r = 0.31$, $p = 0.04$).

However, no significant effects of TOR-assertiveness were found for the other driving performance measures or any of the heart rate measures. The minimal relative changes to the baseline indicate no meaningful changes in physiological stress or efforts between TOR-types.

Table 4 Statistical results for immediate post-TOR measures. ES r represents effect sizes and probability was adjusted ($p < 0.5$). Significant results are highlighted. Table adapted from Kraut et al. (2023, p.8).

Hyp.	Measures	TOR type (M, SD)		ES <i>r</i>	Adj <i>p</i>
		Non-assertive	Assertive		
	<i>Performance Measure</i>				
H1	Reaction time (s)	3.49 (0.93)	2.92 (0.88)	0.48	<0.01
H2	Braking force (N)	21.92 (10.95)	26.06 (11.54)	0.31	0.04
H3	Collision rate	0.05 (0.10)	0.04 (0.07)	0.17	0.37
H4	Average speed (km/h)	12.04 (5.98)	11.77 (4.76)	0.04	0.77
	<i>Physiological Measures</i>				
H5	HR ratio	1.03 (0.04)	1.03 (0.04)	0.04	1
H6	SDNN ratio	1.02 (0.28)	1.06 (0.40)	0.03	1
H7	RMSSD ratio	1.04 (0.20)	1.20 (0.95)	0.00	1

4.1.2 Sustained Effects of Assertive Takeover Requests

The hypotheses H8 to H13 assessed the sustained effects of assertive takeover requests. Driving behavior was evaluated using measures such as average speed, variability in steering, and speed variability. Physiological responses were examined using measures of heart rate, SDNN and RMSSD.

Table 5 shows that no significant impact of TOR assertiveness on driving performance or physiological measures during the subsequent manual driving periods was observed.

Table 5 Statistical results for sustained post-TOR measures. *ES r* represents effect sizes, and *p*-values were adjusted ($p < 0.05$). Table adapted from Kraut et al. (2023, p.9).

Hyp.	Measures	TOR type (M, SD)		ES <i>r</i>	Adj <i>p</i>
		Non-assertive	Assertive		
<i>Performance Measure</i>					
H8	Average speed (km/h)	101.97 (10.52)	99.91 (11.37)	0.19	0.42
H9	Speed variability (km/h)	26.37 (4.9)	26.51 (4.84)	0.01	1
H10	Steering variability (°)	3.97 (1.7)	3.64 (1.25)	0.07	1
<i>Physiological Measures</i>					
H11	HR ratio	1 (0.03)	0.99 (0.02)	0.03	0.99
H12	SDNN ratio	0.97 (0.14)	1.06 (0.40)	0.18	0.42
H13	RMSSD ratio	0.99 (0.12)	1.04 (0.08)	0.19	0.42

4.1.3 Subjective Ratings on Stress and Performance

This subsection presents all subjective evaluations, including the self-reported measures on stress and performance in VAS, Perceived Stress Test results with baseline and post-study measurements, and the NASA Task Load Index findings.

The average subjective ratings of stress and performance were not significantly affected by TOR-assertiveness. Table 6 illustrates the results for these measures according to hypotheses H13 and H14. For this analysis, the Visual Analogue Scale was normalized to a range from 0 to 1, where 0 represents very low stress or very low success in driving performance, and 1 represents very high stress or very high success in driving performance.

Table 6 Statistical results for subjective ratings on stress and performance, depending on TOR-assertiveness. The data was normalized to a scale ranging from 0 to 1. Table adapted from Kraut et al. (2023, p.9).

Hyp.	Measures	TOR type (M, SD)		ES <i>r</i>	Adj <i>p</i>
		Non-assertive	Assertive		
H14	Stress rating	0.53 (0.12)	0.52 (0.14)	0.06	0.6
H15	Performance rating	0.77 (0.16)	0.78 (0.15)	0.13	0.6

Perceived Stress Test

Although a separate perceived stress test questionnaire (PSS) was administered to participants, the data collected from this questionnaire were not included in the analysis of the primary hypotheses. This decision was made to maintain consistency with the other subjective measures used in this study. However, the PSS provides a metric to evaluate how much participants are affected by stress in their daily lives and their attitude towards resolving stressors.

The evaluation of the PSS before the experiment revealed that 21 participants experienced low stress in their daily lives (score less than 14 out of 40 points), and 10 participants experienced moderate stress (score of 14-26 points). No participant was found to experience high stress, which would have produced a score ranging from 27 to 40 points.

A two-sample t-test comparing the one-item baseline stress ratings with the post-study ratings found no significant difference ($p = 0.368$). This suggests that the participants' perceived stress levels were not significantly influenced by the demands of the experiment.

Workload Assessment with the NASA Task Load Index

The NASA Task Load Index (NASA-TLX) was administered without the weighting step, which reduced the time required to complete the form. This simplified approach is also known as Raw TLX (RTLX), as defined by the creators of the NASA-TLX (Hart & Field, 2006). Figure 8 shows that in the experiment, high success rates ($M=14.81$, $SD=4.2$) were associated with medium levels of mental workload ($M=8.72$, $SD=9.7$) and overall effort ($M=9.2$, $SD=4.53$). This might be an indicator for an adequate level of task difficulty. The moderately high ratings for temporal demands ($M=8.6$, $SD=5.4$) provide further support for this interpretation. In contrast, high task performance coupled with low workload scores would suggest that the task was not challenging enough (Hart & Field, 2006). As expected in a driving simulation study, the recorded physical loads were low ($M=8.73$, $SD=5.14$). In

addition, ratings on frustration ($M=5.67$, $SD=4.32$) were also rated fairly low, consistent with findings from Table 6 on low self-reported stress during the experiment.

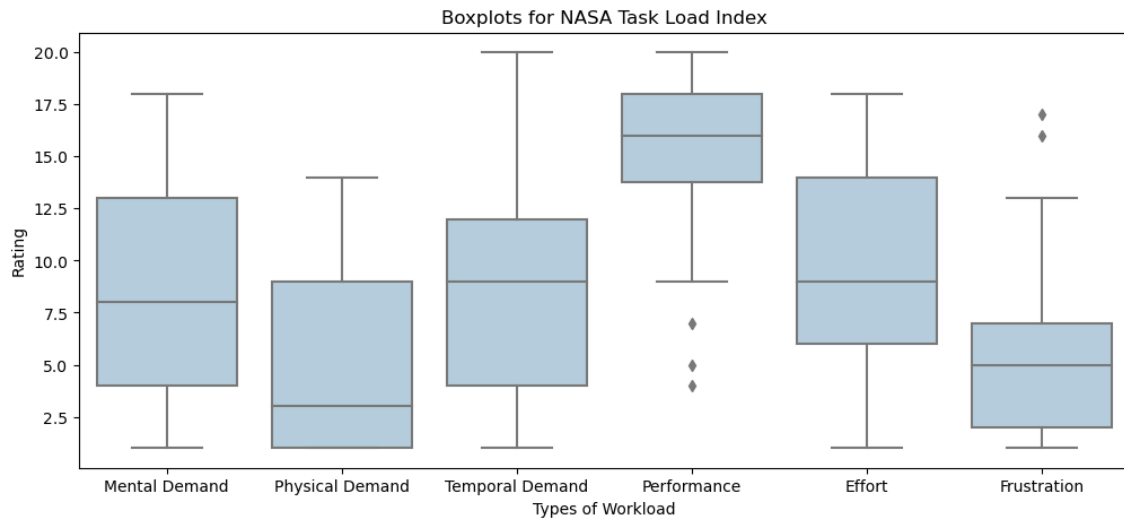


Figure 8 Workload assessment in NASA TLX: Participants experienced high measures of success combined with low physical efforts and stress.

4.2 Qualitative Results

At the end of their experiment, participants reflected on their experiences in a concluding interview. All participant feedback was first paraphrased alongside each interview question, before recurring themes were identified and summarized below.

Recall of TOR-Signals

Differences in subjective perceptions could be observed, when participants were asked about defining features of the takeover requests. Three participants were unable to describe any characteristics of the takeover alerts post-experiment. 13 participants remembered only the high-assertive TOR, while eight recalled only the low-assertive TOR, and seven had a good recollection of both. The high-assertive sound was often described as alarming or like a siren, while the low-assertive sound was seen as pleasant or dull. Despite no haptic signals being administered, three participants reportedly felt vibration feedback in the steering wheel. The visual information, including the icon for manual driving and red tint in the HUD indicating the takeover, were recalled less frequently.

Most Helpful Features

The TOR sounds were deemed most helpful by 22 participants, with visual signals being less important. The majority of participants reported focusing on the road rather than the visual information given in the heads-up display. Among the information presented on the display, the speedometer was sometimes noted as useful, while only a few participants checked the icon indicating manual driving for additional confirmation.

Requested Design Changes

Participants suggested several enhancements to the user interface: eight wanted verbal instructions or information on the distance to a possible accident, two preferred earlier TOR notifications, and six participants requested more prominent visual alerts. Vibration alerts and vehicle-to-vehicle communication were mentioned by some, while one person proposed a two-stage warning system. Frequent requests were made for longer, louder, or even more aggressive takeover requests. Regarding the full setup, two participants proposed to mount the tablet for the secondary task on a fixture or to administer alerts on any connected devices, respectively. Eleven participants were content with the interface elements as they were and did not request any changes.

Progression of Perceived Stress throughout the Experiment

For most participants, the level of perceived stress decreased with practice and the reported learning effect. One participant became more irritated as the experiment went on, and two participants did not report feeling stressed at any time. Participants who were more stressed in the beginning explained that the cause was the unfamiliarity of the obstacle course or the sounds, or feeling overwhelmed in the situation. Participants claimed to feel most comfortable after the third round. One participant said that having an accident in the first obstacle course was stressful and it took until the free driving period to calm down again. Two participants felt more hectic or nervous in the first obstacle course, but their perceived stress decreased over time. One person stated that their stress level did not change at all. Participants used different strategies to deal with the task: Some developed trust in the system and focused their concentration on the secondary task, while others anticipated the TOR-signals strongly and concentrated more on the system's sounds and the road environment.

Task Comprehension and Subjective Experience

The interviews showed that all participants had a proper understanding of the task and their responsibilities. Participants mostly enjoyed the experiment, although some noted differences from real-life driving. For some, issues arose from an unrealistic resistance of the steering wheel or difficulty with estimating the vehicle's dimensions and speed correctly. While some found the task more stressful or less game-like than expected, others felt it resembled a game optics too much. Several participants expressed discomfort with the demands of the secondary task, suggesting improvements like device fixtures or automatic braking before TOR issuance. However, anticipatory actions were common, with several participants keeping their foot on the brake during the secondary task or listening out for the TOR signals. Eight participants directly mentioned improvements with practice, reporting on feeling more comfortable and successful over time.

Thoughts about Automated Driving

The ADAS shifted the role of a driver to a passenger, some participants noted. One participant reported very little trust in the system, leading to a more cautious driving behavior than usual. When asked for final thoughts after the experiment, several participants voiced concerns about driving automation. In particular, participants reported a lack of trust in ADAS for complex environments, such as narrow city streets or steep mountain roads, as well as in dynamic situations requiring quick decision-making. Other concerns focused on potential misuse of automation by either the manufacturers or other drivers who might neglect their responsibilities to takeover control when necessary. Despite these issues, others saw driving automation as an opportunity for increased mobility of persons without a driver's license. Three participants suggested that fully automated vehicles could potentially serve as efficient taxis, reducing the need for personal vehicle ownership. However, their attitudes varied, as one feared the loss of the car as a status symbol, while the others were optimistic about more mobility and a future with less traffic jams.

4.3 Additional Findings

The above presented quantitative results examined aggregate scores for assertive and non-assertive TORs across all laps. To better capture changes among the sequence of laps, linear regressions were calculated in subsection 4.3.1. Additionally, subsection 4.3.2 highlights correlations between variables to provide further insights into potential connections among all independent and dependent measures.

4.3.1 Adaptation Effects across Laps

Table 7 shows weak to moderate tendencies for several measures throughout the experiment. Notably, significant changes were found in performance measures such as speed in the subsequent period, reaction time, number of collisions, and braking force. Although no significant changes were detected in physiological measures, both self-reported stress levels and subjective assessments of performance showed weak trends.

Table 7 Linear regression statistics of all measured variables. The table contains the correlation coefficient r and the adjusted p -value. Significant results are marked in bold. Hypotheses 1 – 7 contain measures for the immediate period, while hypotheses 8 – 13 pertain to the sustained period, and hypotheses 14 and 15 involve the post-lap measures. Table adapted from Kraut et al. (2023, p. 10).

Hyp.	Measures	r	Adj p	Hyp.	Measures	r	Adj p
<i>Performance Measures</i>							
H1	Reaction time	-0.38	<0.01	H8	Average speed	0.44	<0.01
H2	Braking force	0.25	<0.01	H9	Speed variability	-0.01	1
H3	Collisions	-0.31	<0.01	H10	Steering variability	-0.18	0.11
H4	Average speed	-0.12	0.68				
<i>Heart rate measures of stress</i>							
H5	HR ratio	-0.15	0.36	H11	HR ratio	-0.05	1
H6	SDNN ratio	-0.02	1	H12	SDNN ratio	-0.09	1
H7	RMSSD ratio	0.06	1	H13	RMSSD ratio	-0.12	0.68
<i>Subjective Ratings</i>							
H14	Stress	-0.22	0.02	H15	Performance	0.34	<0.01

Trends for Immediate Effects

The quantitative results uncovered significant differences for reaction time and braking force for TOR-assertiveness. When visualizing regressions on lap number, it can be observed that assertive TORs produced quicker or more forceful responses on average at every lap, even though an adaptation can be recognized. However, considerable variation between participants is visible.

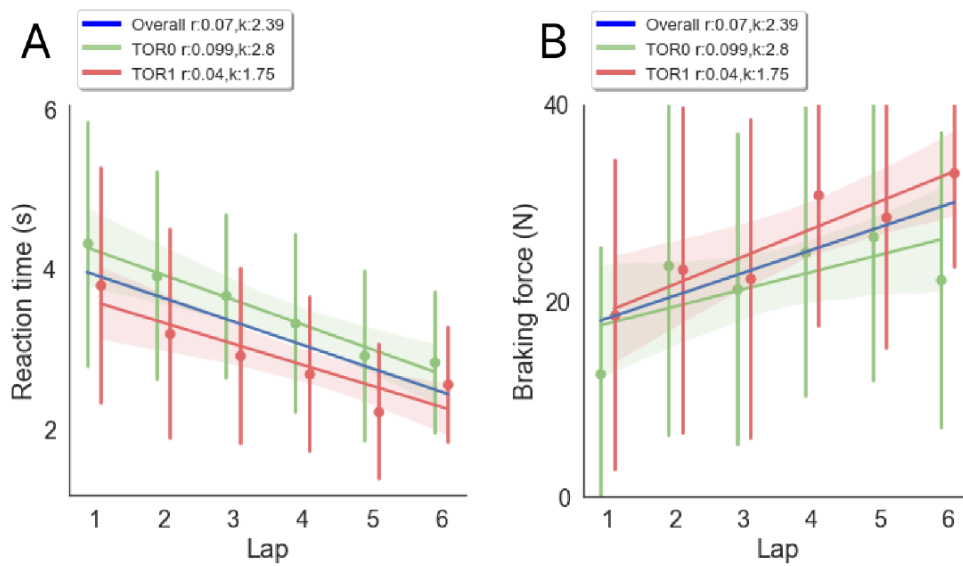


Figure 9 Adaptation effects for reaction time (A) and braking force (B). The two measures were regressed on the number of laps for the assertive TOR (red), the non-assertive TOR (green) and both TORs combined (blue). The whiskers correspond to the SD and the bands indicate confidence intervals.

No significant differences in the number of collisions were found when comparing high and low assertive takeover requests. 13 of the 31 participants recorded collision events, with collisions happening most frequently in the first two obstacle courses (Lap 1: 37, Lap 2: 26, Lap 3: 3, Lap 4: 9, Lap 5: 0, Lap 6: 0).

Trends for Subsequent Effects

Most notably, participants experienced adaptation effects in the sustained driving period for average speed. A moderate effect can be observed for increasing speeding as the experiment progressed.

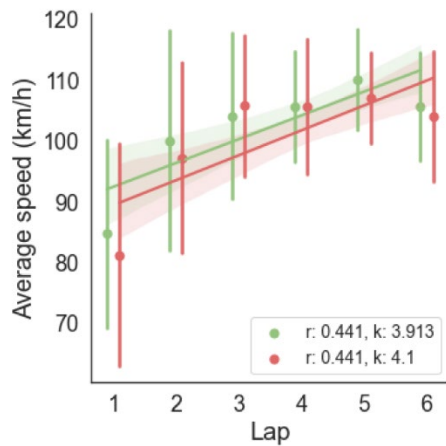


Figure 10 Adaptation effects for average speed in the sustained period after the takeover. Participants became more comfortable with accelerating in the free driving period throughout the course of the experiment.

Trends for Self-reported Measures

The performance and stress ratings in the laps were normalized between 0 and 1, further labelled as “very high” and “very low”. Figure 11 shows that independent of TOR-type, self-reported stress decreased, while performance increased over the course of the experiment.

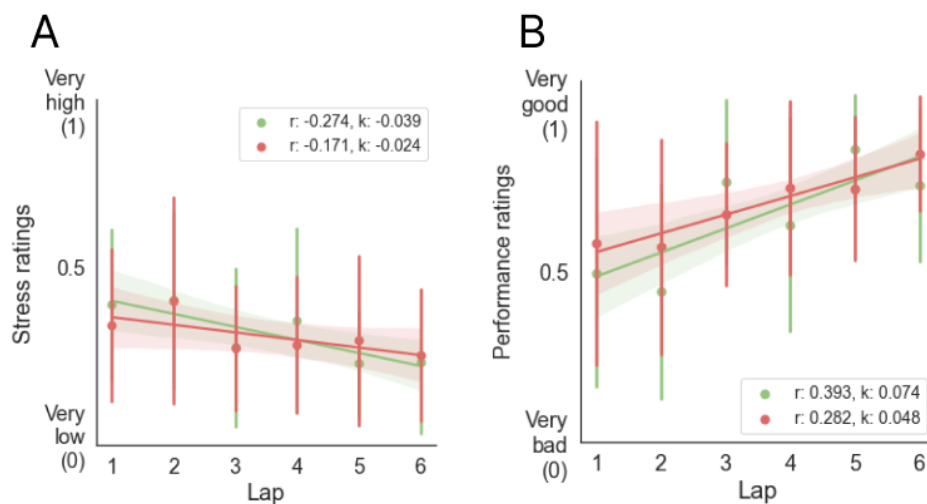


Figure 11 Adaptation effects for stress ratings (A) and performance ratings (B) regressed on the number of laps. Participants rated their performance increasingly better, while reporting lower perceived stress.

4.3.2 Correlations

The Spearman's rank correlation coefficient was computed for all available variables, including demographic data, immediate and sustained measures and self-reported ratings. Below, insights from several strong monotonic relationships are highlighted.

Correlations of Stress Measures with Age

Overall, age showed negative monotonic relationships with SDNN during the automatic driving ($r = -.55$), obstacle courses ($r = -.53$), and the free driving periods ($r = -.63$). This effect is most visible in the age group of the 30 – 40-year-olds for SDNN during automatic driving ($r = -.75$), in the obstacle course ($r = -.7$), and in the free drive ($r = -.73$). Since higher SDNN indicates higher subjective stress, it can be assumed that participants in this age range were less comfortable with the task. No strong correlations between any of the physiological measurements and any of the driving performance indicators were found.

Correlations of Driving Performance Measures with Subjective Performance Ratings

Self-ratings of participants correlate negatively with the number of collisions ($r = -.58$), as well as the steering variability during the free driving period ($r = -.63$). A negative monotonic relationship between self-reported performance and steering wheel variability in the free drive period was found ($r = -.63$). As increased steering wheel variability indicates increased nervousness on the part of the driver, this indicates that participants provided lower performance rating when they expressed nervous steering behavior in the free driving period. Moreover, performance ratings were lower for laps, in which participants experienced collisions ($r = -.58$).

5. Discussion

This master thesis addresses two research questions: RQ1 was formulated to answer whether TOR-assertiveness has immediate and sustained effects on drivers through an empirical study. RQ2 investigates the cognitive processes involved in automated driving by interpreting the experimental data through the lens of the predictive processing framework. Following the analysis of the quantitative findings in subsection 5.1, the current section addresses RQ2's guiding questions in subsection 5.2. This section concludes with a discussion of the study's limitations and opportunities for future research.

5.1 Interpretation of Experimental Findings

The data indicate that TOR-assertiveness leads to faster reaction times and stronger initial braking force. However, no significant effects were found on subsequent driving periods or stress measures. This suggests that the assertive TOR did not impair driving ability by evoking high levels of cognitive overload.

5.1.1 Significant Effects for Immediate Measures

The thesis started out with fifteen hypotheses regarding the impact of TOR-assertiveness on stress and performance measures. Significant results were found only for the assumptions that the signal assertiveness affected driver characteristics immediately after the TOR. Participants took over control faster and exerted greater force on the brake pedal.

5.1.2 No Impacts on Subsequent Measures Observed

No significant effects were found during the subsequent driving period for measures of stress, performance, or subjective ratings. These observations disproved the hypotheses that assertiveness could negatively impact the driver's cognitive state and, consequently, driving ability. Instead, the TOR assertiveness tested in the study indicates a safety benefit, as participants responded an average of half a second earlier without experiencing high levels of stress. At a speed of 120 km/h, this half-second head start translates to a distance of 17 meters, or roughly the length of four cars.

5.1.3 Subjective Feedback and Interviews

One potential drawback of assertive signals, as noted in previous research by Fagerlönner et al. (2010), is the perception of annoyance. However, interviews in this study did not indicate that assertive takeover requests (TORs) were perceived as disruptive or annoying. In fact, several participants did not even notice the presence of multiple signals. Some participants even requested that the assertive TORs be more pronounced.

Several interviewees stated that using SAE Level 3 automation in their daily lives would lead to a perceived loss of control over their vehicles, a concern also noted in prior research (de Winter et al., 2014; Howard & Dai, 2014). However, the increase in comfort and safety appealed to participants, with positive feedback outweighing negative. While the driving task has traditionally been the responsibility of a human driver, societal and legal challenges may prove more difficult to resolve than technological advancements (Kyriakidis et al., 2019).

Contrary to the hypothesis, no increased stress responses were recorded for the assertive TOR, which was rated as stressful in the pilot study. A possible explanation could be that the driving task context influenced participants' interpretation of the signal. Fagerlönner et al. (2010) found that high-urgency signals were rated significantly less annoying during driving simulations compared to pre-study evaluations. This effect was not observed for low-urgency signals. Fagerlönner suggests that professional drivers accustomed to handling critical situations were less affected during the actual task. A similar effect might have occurred in the present study.

5.2 Predictive Processing Insights

Even though prediction and expectation are common themes in human factors research, the application of predictive processing concepts has been less explored within the discipline. This framework offers a unifying model of perception and action, based on the principle of prediction error minimization. The current section addresses the three guiding questions of RQ2 and demonstrates how predictive processing can be used to better understand the relationships between workload, stress and driving performance in the context of driving automation research.

5.2.1 Internal Model Updating when Driving with ADAS

Driving automation at SAE Level 3 not only impacts a driver's role, but also drastically reduces their workload. However, takeover events or distractions from secondary tasks introduce fluctuations in mental and physical demands for drivers. This subsection examines these dynamics through the lens of predictive processing, addressing RQ2's guiding question 1: "How might drivers' internal models update and optimize when operating a vehicle with an Advanced Driving Assisted System?".

Following the Yerkes-Dodson law, optimal driving performance requires moderate levels of workload and arousal, whereas the relationship between these variables can be described with an inverted U-shape (DeGuzman et al., 2022). Performance decrements can be expected for too high a workload as well as for too little workload – which has been a recognized risk factor for driving with automation. Since the sole task of road monitoring might induce underload and sub-par takeover performance, research argues for the appropriate usage of secondary tasks (de Winter et al., 2014) - even if they come at a risk of reduced situational awareness for events on the road. In response drivers might take longer to take over control of a vehicle following a TOR, or brake harder to gain more time for decision-making (Kujala & Lappi, 2021).

Figure 12 illustrates how a driver's mental model updates when interacting with the environment, vehicle, secondary tasks and ADAS. The figure emphasizes the role of prediction error minimization and incorporates theoretical insights from previous research, particularly from Engström et al. (2018), Kujala & Lappi (2021) and Seth (2014). It consists of three interconnected parts: at the bottom is a schema of the dynamic external world, the

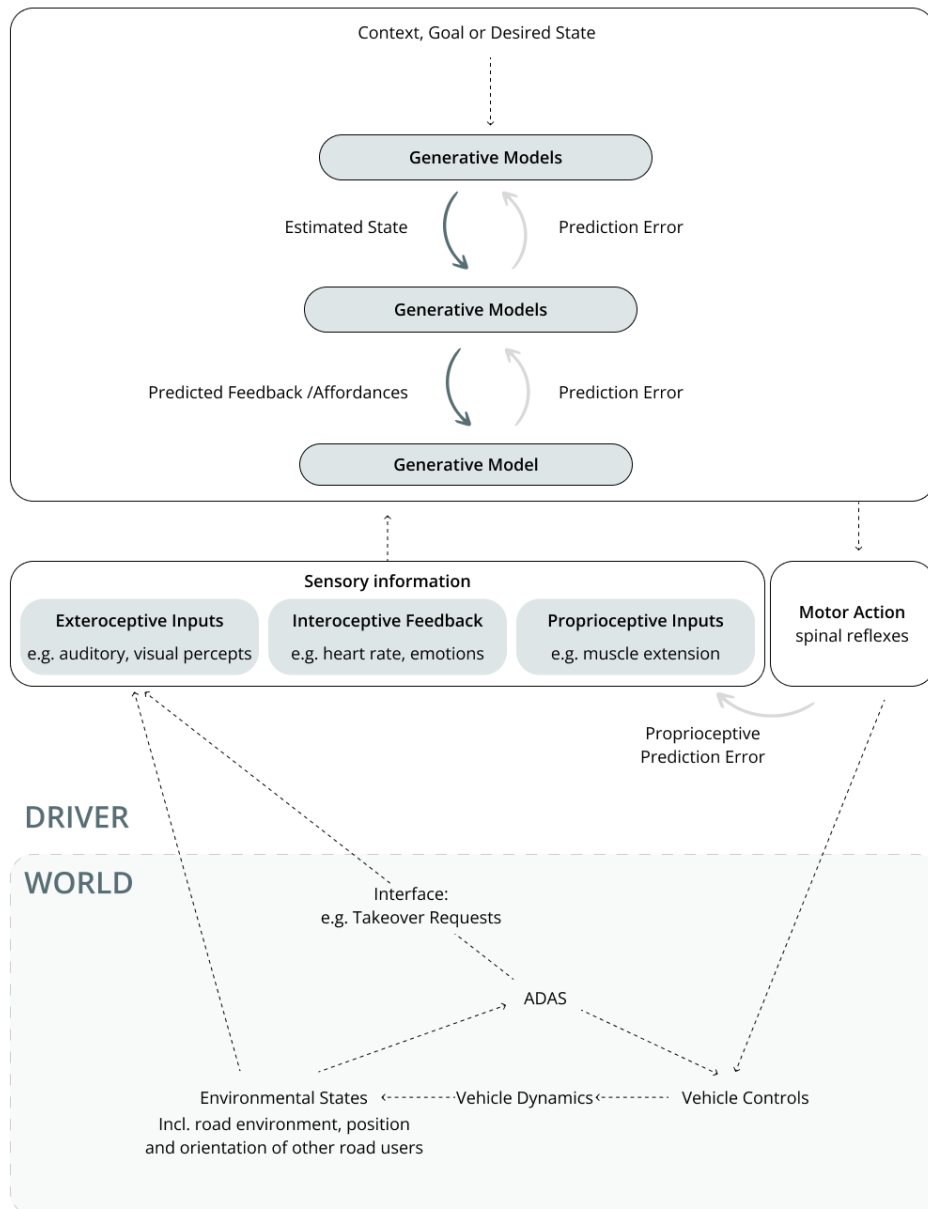


Figure 12 Conceptual illustration of driver interaction with an ADAS from a predictive processing perspective. Bodily and environmental states lead to exteroceptive, interoceptive and proprioceptive sensory input signals. Knowledge of the world and body is inferred based on this sensory input and updated by matching predictions with actual sensory input. Mismatches in the form of prediction errors flow upwards a hierarchical generative model and are met with predictions on estimated states at each level. The system updates itself through minimizing the global prediction error and generating predictions about the newest state. Own representation based on illustrations by (Engström et al., 2018; Kujala & Lappi, 2021; Seth, 2014).

middle depicts the sensory inputs from both environmental and internal states, which feed into the hierarchy of generative, internal models at the top of the figure. Together, these components depict how a driver might respond to automated or manual driving scenarios.

The external world in Figure 12 contains changing environmental states, influenced by vehicle dynamics. These are in turn controlled by either the ADAS or the human driver.

Changes in the environment produce sensory signals, as do the messages by the system, in the form of takeover requests, for instance.

Sensory information is then processed at the bottom of an internal hierarchy of generative models, allowing the driver to continuously create and update predictions about estimated states and goals with the help of prediction errors. Inputs at the lowest levels can be categorized into exteroceptive, proprioceptive and interoceptive signals. Exteroceptive inputs serve as an interface with the external world: The perception of the vehicle, its interfaces, other traffic participants or engaging with NDRTs generates visual, auditory, or haptic sensory signals for a driver. This consistent stream and processing of inputs lets researchers believe that drivers do still engage in some passive inference, even if they are involved in secondary tasks on their smartphones or other devices.

For example, drivers might rely on the body's somatosensory or proprioceptive systems for sensing the body's orientation in the vehicle, positioning of joints, or muscle extensions (Engström et al., 2018). In addition, a driver experiences a constant stream of interoceptive feedback. Interoceptive predictive coding is hypothesized to control the body's vital parameters, such as blood oxygenation or blood sugar. Moreover, subjective feeling states may arise from interoceptive processes (Seth, 2014; Seth & Friston, 2016). In this context, interoceptive predictions are generated, updated, and compared with the help of efference copies of visceral, autonomic, and motor control signals within a salience network (Seth, 2014).

While the lower levels are responsible for collecting sensory information, the representative meaning behind inputs is generated at higher levels. As the prediction error travels upwards through the hierarchy, it encounters increasingly abstract predictions at every level. For instance, signals may diverge significantly from a prediction on a lower level, but the driver might only assign the signal to be "stressful," when the prediction error has reached a higher level. When assessing driving performance, affect and emotions should not be disregarded. Previous research found that nervousness or mental stress can be deducted from driving behavior alone (Ongul & Ongun, 2020; Sugiono et al., 2018). Moreover, stressed drivers may be biased toward perceiving danger in the world more strongly and might experience reduced performance due to overly high levels of emotional arousal. Heightened stress or anxiety further determine prior beliefs and through high precision predictions, inhibit safety learning mechanisms and result in biased perceptions about risks (Sladky et al., 2024).

Mental model updating after automated driving

When a driver attends to a secondary task during automated driving, low-level perceptions of the somatosensory or proprioceptive systems may not be salient enough to command a driver's full attention. The event of a takeover request then initiates interesting changes within a driver's mental model. When the system detects safety bounds, it usually decelerates before issuing a TOR. This highly unexpected change in acceleration can produce heavily weighted proprioceptive prediction errors. Since these errors propagate upwards and do not cancel out the predictions about the expected acceleration and state on the road, the driver is more likely to actively sample further information. Ideally, the TOR signal itself is salient enough to demand a driver's attention. The size of the resulting prediction error corresponds to a driver's prior beliefs. This results in larger errors if a driver has high trust in the system and operates under the assumption that a TOR is unlikely. Only once the prediction errors propagate to a higher level within the hierarchy of generative models, the driver can become aware of the new situational context and can change her goals to address the manual driving task.

Moreover, the size of the prediction error is linked to increased repeated sampling of the environment (Kujala & Lappi, 2021). The continuous minimization of the prediction error through repeated sampling, such as frequent glances on the road, further increases a driver's confidence that a good representation of hidden causes was found (Kujala & Lappi, 2021; Wiese & Metzinger, 2017). When a driver decides to intervene and take over control, the appropriate actions will be assessed through a top-down competition between linked sets of hypotheses (Clark, 2013; Pezzulo & Cisek, 2016). The current winning hypothesis on the optimal action emerges, when the prediction errors associated with the elements of the driving sensory signal are selectively suppressed. The errors associated with remaining elements of the driving signal are left untouched and propagate upwards. This top-down suppression may only change, when the overall interpretation of the situation changes (Clark, 2013). This mechanism enables a driver to decide whether braking, steering or no intervention is the ideal choice of action. Any prediction about a motor action will thereby be cancelled out by the proprioceptive and exteroceptive prediction errors. The driver then gathers feedback if the action was successful and brought about the desired effects.

Following an action-oriented account of predictive processing proposed by Friston, error signals may self-suppress in motor systems whereby motor actions change the bottom-up proprioceptive and sensory input. In this way, proprioceptive predictions might directly act as motor commands. The author argues that active inference is not directed by desired

consequences, rather than prior expectations are guiding perceptual inference and action, shaping an experience-dependent learning process (Friston, 2010). Some researchers critique the impact of such a view on the organization of cognitive economy in contrast to the traditional appeals of cost and reward functions (Clark, 2013). However, previous research appears to agree that high layers do not communicate a complete action plan. Rather, a top-down flow of predictions on makes it possible to identify affordances for appropriate interventions, whereas the lower levels retain some autonomy (Pezzulo & Cisek, 2016).

As mentioned before, the overall success of a takeover depends on a driver's trust in the ADAS. An excess of trust can lead to increasing risks and decrements to the takeover performance, which are especially problematic given the responsibilities of a driver in SAE Level 3. Engström et al. (2018) view trust in a driving automation as the precision associated with the generative model's predictions regarding the sensory consequences of the system (Engström et al., 2018). A prior assumption about a system's infallible safety might result in high precision predictions, effectively overriding contradictory prediction errors. A possible consequence could be that drivers might not even glance on the road during autopilot, since road monitoring does not contribute to minimizing a relevant prediction error anymore.

These insights underscore the importance of designing takeover requests that align with drivers' expectations and mental models to enhance safety. A TOR must be easily understandable, to ensure that an appropriate prediction error is generated without causing heightened stress. To optimize the effectiveness of TORs, drivers should be thoroughly informed about their necessity and key characteristics. This instruction is vital because the ability to form accurate predictions regarding sensory effects, and to compare these with the actual signals, is significantly improved when drivers possess strong prior beliefs about the signals and a clear understanding of their altered role.

5.2.2 Prediction Error Minimization in the Immediate Period

Guiding question 2 asked: How is prediction error minimization observable in driver responses immediately following a highly assertive TOR compared to a less assertive TOR? This subsection aims to answer this question, by assessing the experimental procedure,

how participants might have categorized the assertive TOR and adapted their takeover response throughout the laps.

Braking force and reaction time showed consistent, significant results for TOR-assertiveness, establishing them as proxies for observing prediction error minimization. This assumption drives the discussion on how TOR-assertiveness influenced prediction error size, contributed to a gradual model tuning and to participants' internal category construction on the meaning of the TORs.

Assertive TORs Elicit Larger Prediction Errors and Stronger Responses

While every TOR was issued at an unexpected interval during driving automation, the assertive TOR caused a greater disruption to the secondary task. Participants responded more quickly and exhibited a more forceful braking reaction. It is likely that the assertive nature of the alert generated a larger prediction error at the low perceptual levels for hearing and vision.

When these prediction errors were fed into the first generative model, a strong difference from the predicted exteroceptive input was detected, allowing the errors to be propagated to the next higher level. At each level, these errors encountered expected states encoding increasingly abstract features. Once the errors had propagated to a high level, the overall goal setting or understanding of the context could be updated. This triggered a top-down cascade of predictions based on the new information, which were then met with new prediction errors from continuous perceptual sampling. Ultimately, this process enabled the driver to intervene in the takeover.

Generative models and observations are ever-changing and uncertain, which the system is aware of and is trying to adapt to in a rational way (Kujala & Lappi, 2021). The assertiveness of the TOR might be perceived as an indicator for imminent danger, which in turn skews the perceived desirability of possible actions. Braking faster and harder might appear as a more desirable choice than waiting to intervene. In comparison, longer takeover time budgets have been found to result in longer reaction times (Gold et al., 2013). This course of action might be more desirable, in order to gain situational awareness without time pressure. In the present study, less assertive TORs prompted participants to spend more time transitioning from the secondary task to driving, while they possibly increased their visual sampling and prioritized situational awareness.

Subjective Categories of TOR-Assertiveness

The interviews revealed that several participants were unaware of differences in the assertiveness of TORs. However, they still showed faster reaction times to the more assertive TOR, proving the robustness of the study design.

The brain naturally seeks to create categories to make sense of sensations, grouping similar events or objects that serve the same goal into the same category (Barrett, 2017). For those participants who did not consciously pay attention to the assertiveness level of the TOR, all warning signals might have been processed as equivalent - belonging to the same category. The process of category formation might also be context-dependent, with the brain adapting its categorization strategy based on the situational demands. Drivers tend to show higher acceptance of otherwise annoying signals in demanding driving situations, as previous research has found (Fagerlönn, 2010; Roche & Brandenburg, 2018).

Participants frequently reported having learned to anticipate both the TOR sound and the subsequent challenge of navigating the obstacle course. Since any TOR required the same manual intervention, participants possibly also learned to disregard the differentiating characteristic of assertiveness. This behavior aligns with the assumption that agents act to actively seek or generate sensory consequences that their brains expect (Friston, 2009; Friston, 2010). If a participant had the strong a priori expectation that any alert requires intervention, the takeover situation might automatically activate lower layer actions that were afforded with sufficient precision (Seth & Friston, 2016). This means that participants could take actions without needing to evaluate a TOR on its level of assertiveness.

Furthermore, when strong prior expectations are present, model parameters may not be immediately updated to reflect new sensory data. However, even among participants who did not consciously differentiate TOR assertiveness levels, a gradual tuning of relevant parameters was observed throughout the experiment. This finding is illustrated by a quote from participant 22: "You learn to react to any stimuli without listening to the sound." This suggests that, while explicit recognition of TOR assertiveness may not have been present, there were still processes of implicit learning and adaptation taking place.

Learning Effects Indicate Improved Precision Estimates

Over the course of the experiment, a trend was observed towards increasingly shorter reaction times and stronger braking reactions could be observed. This suggests that participants may have adopted a strategy of braking first and monitoring later, especially

when hearing the assertive TOR. The high prediction error associated with the assertive TOR likely contributed to this trend. Since every prediction includes an estimate of the expected precision or certainty of the sensory input, prediction errors that deviate from high precision predictions generally receive a higher weighting. This mechanism serves as a means of selectively enhancing sensory input (Feldman & Friston, 2010).

A driver may selectively base corrective steering actions on aspects, which are perceived as reliable. For example, when driving in very foggy weather, a driver is more likely to use lane markings for guidance rather than the roadway in the distance (Engström et al., 2018). In the current study, participants appeared to focus their attention on the auditory component of the takeover requests, particularly those with assertive characteristics. This assumption aligns with previous research suggesting that driving with Advanced Driver Assistance Systems (ADAS) failures involves stimulus-driven precision weighting. These bottom-up processes enable drivers to detect and react to salient features in the environment (Li et al., 2017). The associated precision can be estimated as a function of obtained samples, or the inverse of sample variance. In contrast, top-down precision estimates are modulated in contexts where increases or decreases of precision are anticipated or can function as goal representations for mental action (Metzinger 2017). This would apply to goals on a less-immediate timescale, such as planning and performing a lane change over the course of several seconds.

The improved takeover performance in the study allows to infer that participants' internal model parameters gradually adjusted to support the behaviors associated with successful outcomes. The internal model eventually learns to capture statistical structures in the sensory input and infers their external causes, thereby yielding more accurate predictions with fewer errors. Two key implications follow from this: First, a driver, who is well attuned to a scenario can be expected to exhibit visual scanning patterns that reflect an optimal balance between pragmatic actions and epistemic actions. These can be characterized as exploitation (e.g., steering braking or acceleration) and exploration (e.g., glances on the road or the in-vehicle information system), respectively. This balance aims to minimize prediction error by integrating both immediate response actions and information-gathering behaviors. Second, problems can arise when the statistical regularities of the environment suddenly change, leading to what is known as an expectancy problem. In such cases, the driver's model may fail to adjust quickly enough, resulting in increased prediction error and potentially unsafe responses. This highlights the importance of designing systems that help

drivers maintain flexibility in their predictive models, enabling them to quickly adapt to changing conditions in the driving environment (Wei et al., 2022).

5.2.3 Prediction Error Minimization in the Sustained Period

No significant differences in driver stress and behavior were identified during the subsequent driving task, indicating that drivers quickly regulated their responses and maintained consistent performances. This subsection further explores the differences between the immediate and sustained period, while addressing guiding question 3: “How does TOR-assertiveness influence prediction error during the subsequent driving task?”

Short and Long-term Affordance Competition for Safe Driving

After performing the takeover process and transitioning to the subsequent driving task, participants likely experienced a contextual shift, allowing them to refocus their attention on the road and effectively adapt to the demands of manual driving. Their performance is influenced by prior beliefs and real-world driving experience, where actions are often executed with a high degree of automaticity.

The subsequent driving task presented two distinct sections: a risky segment with obstacles and a free-driving segment without them. These scenarios require decisions and actions that are organized across different timescales (Pezzulo & Cisek, 2016). Immediate responses, such as collision avoidance, rely on short-term control loops to keep a controlled variable stable. This variable may be the visual distance to a leading vehicle or object ahead. Previous research found that this so-called looming metric can influence reaction time in takeover scenarios (Ayoub et al., 2022). It is important to note that actions are context dependent and are neither a fixed response nor predefined by a higher level (Araújo et al., 2006; Pezzulo & Cisek, 2016). Interestingly, the perception of a high urgency scenario appears to create a mismatch between expected and actual looming, triggering faster interventions (Ayoub et al., 2022). If the looming variable exceeds a subjective threshold, short-term control loops engage proprioceptive reflexes, prompting braking or steering actions. Since higher layers within the hierarchical generative models do not compute a behavioral plan, but rather expectations and reference signals, lower levels retain significant autonomy in selecting optimal actions (Pezzulo & Cisek, 2016).

Control loops operating on longer timescales facilitate an affordance competition to estimate the expected broader consequences of actions. During the free driving phase, participants had the aim of driving safely with manual vehicle operation. Higher hierarchical layers might encode abstract goals, such as complying with traffic regulations and speed limits, which then induce top-down expectations and reference signals. These reference signals might include speedometer readings, looming thresholds, or expected speed assessed through proprioceptive inputs, which further constrain the recognized affordances. For instance, perceiving and remembering the highway as straight may inhibit the natural reflex to adjust the steering wheel. Observations in this study proved that participants mostly adhered to speed limits and avoided unnecessary lane changes in the free driving period. This behavior illustrates how participants intentionally navigate an affordance landscape, where the brain continuously generates predictions about future opportunities rather than merely reacting to sensory inputs (Pezzulo & Cisek, 2016).

Impact of TOR-Assertiveness on Sustained Stress Responses

Contrary to the initial hypotheses, the experiment showed no negative effects in the driving task following the more assertive TOR. Participants did not exhibit stress-related performance indicators, such as increased variability in steering or acceleration, nor was there a higher crash rate. This is further supported by the physiological measures, which also showed no significant changes. According to the Yerkes-Dodson law on performance as a function of workload and arousal (Yerkes & Dodson, 1908), the data imply that participants experienced a suitable level of workload and emotional arousal for handling the subsequent driving task. Although it is difficult to differentiate between the effects of the initial takeover and the subsequent driving period on driver stress, it can be assumed that none of the tasks resulted in a dangerous cognitive overload for participants.

In addition, participants did not develop maladapted prior assumptions about the expected risk of the driving task. A fearful driver would have exhibited consistently perceived stress ratings and might have experienced a suboptimal learning process. Maladapted prior assumptions or suboptimal learning process can lead to biased action plans and heightened anxiety. Anticipatory anxiety can impair learning and result in overly cautious or inappropriate behavior (Sladky et al., 2024). Interestingly, the experiment suggests that the assertive TOR did not lead to maladaptive responses. Participants might have effectively managed interoceptive signals and not interpreted possible heightened emotional arousal as negative (Seth & Friston, 2016). These findings confirm that participants did not perceive

the driving task as excessively risky, indicating an effective regulation of stress in all stages of the experiment.

Positive Affects from Rapid Uncertainty Reduction

Ratings of stress and performance after each round were likely influenced by a rapid reduction of uncertainty during the interval between the initial takeover and the subsequent driving task. This effect becomes more evident when considering the notable performance improvements over the six rounds of the experiment. Crash rates and reaction times decreased, while speeds within the obstacle courses increased. These effects may reflect participants' growing confidence and sense of mastery over the task. When facing a challenging task or game, agents expect to reduce uncertainty at a subjective baseline speed. When an agent minimizes prediction error faster than expected by choosing an optimal action sequence, this elicits momentary positive affective experiences. In general, this process of faster deceleration of error rates contributes to the appeal of uncertainty and novelty in games and other tasks (Deterding et al., 2022; Feldman & Friston, 2010; Kiverstein et al., 2019).

5.3 Study Limitations

While the study conducted for this thesis made it possible to gain valuable insight into the effects of TOR-assertiveness, several limitations must be acknowledged. First, the experiment included secondary tasks during the automated driving period to simulate a more realistic driving environment. However, these distractor tasks likely induced cognitive load in participants, which may have influenced takeover times, driver behavior, and physiological signals (Li et al., 2017).

Additionally, the study did not evaluate the effects of different types of obstacle courses after the takeover. The complexity of these courses might have had confounding effects on the dependent variables (Braunagel et al., 2017; Samuel et al., 2016), limiting the generalization of the results in the sustained driving period. The use of simulated traffic conditions may present another limitation since they cannot fully replicate the unpredictability and intensity of real-life traffic scenarios. While some researchers argue that simulations effectively evoke a sense of presence (Deniaud et al., 2015), others have noted that practicing takeovers in simplified, simulated environments may not adequately

prepare drivers for safety-critical situations (Agrawal & Peeta, 2021). Several participants remarked on the differences from real-world driving, such as the size and resistance of the steering wheel. This discrepancy between simulation and real-world conditions could affect the external validity of the study, as the simulated environment may not fully capture the complexities of actual driving.

Finally, the study may have been affected by a repetition effect, a common issue in driving simulation studies (Agrawal & Peeta, 2021; Gold et al., 2018; Large et al., 2019). Despite the initial learning period in the study, a learning effect was observed for certain variables. While this adaptation could have influenced results, the robustness of the methodology is supported by the significant effects for TOR-assertiveness that were observed consistently.

These limitations highlight the challenges of conducting studies with driving simulators and the potential impact of factors such as cognitive load, course complexity, and simulation realism on a study's outcomes. Future research should carefully consider these factors to enhance the applicability of findings to real-world settings.

Future Research

In this study, no immediate or persistent stress responses were observed in contradiction to the initial hypotheses. Future studies could explore alternative experimental designs under different conditions. For example, takeover scenarios could be made more challenging to enhance perceived urgency, secondary tasks could be more difficult to test participants' information processing abilities, or the duration of driving with autopilot could be extended from two minutes to a longer period. Additionally, other measures could enhance robustness or capture additional effects. For this, situational awareness could be assessed in more detail using eye-tracking techniques.

Consistent differences between assertive and non-assertive TORs were observed, despite evident learning effects. However, longer adaptation periods seem necessary to reduce potential learning effects. Moreover, the takeover scenarios within the learning and experimental tasks should be similar. Different outcomes in previous research may also be attributed to demographic factors. It is advisable to give more consideration to variables such as age, driving experience, and familiarity with the vehicle when investigating TOR effects.

6. Conclusion

The advancements of driving automation systems produce a shift in how driver and system share their responsibilities in safely commandeering a vehicle. Systems in SAE Level 3 vehicles are able to perform all driving tasks independently, while the human driver acts as a fallback-ready user. The takeover request is therefore one of the most crucial interactions, during which the system transitions control to the driver. In this process, physical and mental demands are placed on the driver, who needs to switch tasks and regain situational awareness quickly. Modality, urgency, timing, and other characteristics of the takeover request have been found to significantly affect reaction times, driving performance and stress. Even though stress has sustained negative impacts on driving behavior, previous research is largely focused on the takeover process itself to evaluate performance.

This research aimed to investigate the effects of assertiveness in takeover requests on driver performance and stress, focusing on both immediate responses and sustained performance post-takeover. The study employed central concepts from the predictive processing framework, motivated by the need for a comprehensive interpretation of performance in connection with fluctuating cognitive demands in takeover scenarios for more efficient transitions in semi-automated vehicles. By examining drivers' physiological responses, self-reported stress levels, and driving performance metrics, this thesis provides new insights into the balance between effective attention capture and induced stress in ADAS design.

The study's results showed that highly assertive takeover requests significantly improved reaction times and led to more forceful braking, as hypothesized. However, contrary to expectations, the assertive signals did not significantly affect driving performance during the subsequent manual driving phase after the takeover. The assertiveness of the TORs did not result in sustained performance decrements or heightened stress at any point of the experiment. This suggests that participants were generally not overwhelmed with the

takeover task, even though the experimental context might have led to increased baseline stress measurements.

The study's integration of the predictive processing framework allowed for a more comprehensive interpretation of possible cognitive processes behind driver behavior. From a predictive processing perspective, any agent constantly generates, and updates internal models based on incoming sensory inputs. Driven by top-down processes, predictions about oneself and the environment drive experience and perception. The assertive takeover request produces a larger than expected divergence from expectations, leading to a faster intervention. With each repetition, the participants train to react more strongly to the assertive signal, as evident by increasingly quicker reactions. Despite consistent significant results for faster takeovers after the assertive takeover request, some participants could not differentiate between the assertive and non-assertive signal. This leaves room for interpretation on reflexive reactions without having to understand the signal at a high level.

Recommendations for Combining TOR-Research with Predictive Processing Concepts

Lastly, this work may be used as a starting point to formulate recommendations for similar research, combining human-computer-research with cognitive science.

Most Effective Takeover Requests Align with Driver Expectations: Prior expectations and beliefs strongly shape not only driver behavior, but the perceived salience of certain alert features. Drivers can make more accurate predictions once their internal models have learned to recognize the consistent patterns of a takeover request. This can be achieved by providing drivers with a consistent signaling framework. In the context of real-world applications, clear explanations of the features of takeover requests are advised to ensure that drivers can internalize appropriate responses.

Different Performance Benchmarks for Short- and Long-term Timescales: One of the strengths of the predictive processing framework is its ability to examine performance over different timescales. Reference signals might be decisive for the choice of quick, instinctive decisions. While longer-term, more complex actions can be interpreted to be goal-directed and subject to learning processes. Moreover, positive, or negative attitude towards a task involving uncertainty can be attributed to the temporal rate of prediction error reduction.

Evaluating Takeover Success from a Predictive Processing View: Although prediction error minimization cannot be directly observed, it can be inferred by combining driving

performance metrics and physiological signals. According to Wei et al. (2022), an optimal balance between pragmatic actions (like steering or braking) and epistemic actions (e.g., gathering visual information) indicates that a driver is well-attuned to a takeover scenario. Observing how drivers choose to take actions and engage in information-gathering behaviors, enables researchers to quantitatively evaluate the effectiveness of prediction error minimization.

Bibliography

- Agrawal, S., & Peeta, S. (2021). Evaluating the impacts of driver's pre-warning cognitive state on takeover performance under conditional automation. *Transportation Research Part F: Traffic Psychology and Behaviour*, 83, 80–98.
<https://doi.org/10.1016/j.trf.2021.10.004>
- Alrefaie, M., Summerskill, S., & Jackson, T. W. (2019). In a heart beat: Using driver's physiological changes to determine the quality of a takeover in highly automated vehicles. *Accident; analysis and prevention*.
<https://doi.org/10.1016/J.AAP.2019.06.011>
- Araújo, D., Davids, K., & Hristovski, R. (2006). The ecological dynamics of decision making in sport. *Psychology of Sport and Exercise*, 7(6), 653–676.
<https://doi.org/10.1016/j.psychsport.2006.07.002>
- Ayoub, J., Du, N., Yang, X. J., & Zhou, F. (2022). Predicting Driver Takeover Time in Conditionally Automated Driving. *IEEE Transactions on Intelligent Transportation Systems*, 23(7), 9580–9589. <https://doi.org/10.1109/TITS.2022.3154329>
- Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, 12(1), 1–23. <https://doi.org/10.1093/scan/nsw154>
- Biondi, F., Leo, M., Gastaldi, M., Rossi, R., & Mulatti, C. (2017). How to Drive Drivers Nuts: Effect of Auditory, Vibrotactile, and Multimodal Warnings on Perceived Urgency, Annoyance, and Acceptability. *Transportation Research Record: Journal of the Transportation Research Board*, 2663(1), 34–39. <https://doi.org/10.3141/2663-05>
- Biondi, F., Rossi, R., Gastaldi, M., & Mulatti, C. (2014). Beeping ADAS: Reflexive effect on drivers' behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, 25, 27–33. <https://doi.org/10.1016/j.trf.2014.04.020>

BMW Group. (2022). *PressClub Global · Audios*.

<https://www.press.bmwgroup.com/global/audio/search/category:sound/>

Braunagel, C., Rosenstiel, W., & Kasneci, E. (2017). Ready for Take-Over? A New Driver Assistance System for an Automated Classification of Driver Take-Over Readiness. *IEEE Intelligent Transportation Systems Magazine*, 9(4), 10–22. IEEE Intelligent Transportation Systems Magazine.

<https://doi.org/10.1109/MITS.2017.2743165>

Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing*, 14(4), 411–427.

<https://doi.org/10.1007/s10339-013-0571-3>

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

<https://doi.org/10.1016/j.apergo.2018.08.028>

Clark, A. (2012). Dreaming the Whole Cat: Generative Models, Predictive Processing, and the Enactivist Conception of Perceptual Experience. *Mind*, 121(483), 753–771.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences*, 36(3), 181–204.

<https://doi.org/10.1017/S0140525X12000477>

Clark, A. (2017). Busting Out: Predictive Brains, Embodied Minds, and the Puzzle of the Evidentiary Veil. *Noûs*, 51(4), 727–753.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of Perceived Stress. *Journal of Health and Social Behavior*, 24(4), 385–396.

<https://doi.org/10.2307/2136404>

de Winter, J. C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research. Part F: Traffic*

Psychology and Behaviour, 27(Part B), 196–217.

<https://doi.org/10.1016/j.trf.2014.06.016>

DeGuzman, C. A., Kanaan, D., & Donmez, B. (2022). Attentive User Interfaces: Adaptive Interfaces that Monitor and Manage Driver Attention. In A. Riener, M. Jeon, & I. Alvarez (Hrsg.), *User Experience Design in the Era of Automated Driving* (Bd. 980, S. 305–334). Springer International Publishing. https://doi.org/10.1007/978-3-030-77726-5_12

Deniaud, C., Honnet, V., Jeanne, B., & Mestre, D. (2015). The concept of “presence” as a measure of ecological validity in driving simulators. *Journal of Interaction Science*, 3(1), 1. <https://doi.org/10.1186/s40166-015-0005-z>

Deterding, S., Andersen, M. M., Kiverstein, J., & Miller, M. (2022). Mastering uncertainty: A predictive processing account of enjoying uncertain success in video game play. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.924953>

Du, N., Yang, X. J., & Zhou, F. (2020). Psychophysiological responses to takeover requests in conditionally automated driving. *Accident; Analysis and Prevention*, 148, 105804. <https://doi.org/10.1016/j.aap.2020.105804>

Engström, J., Bärghman, J., Nilsson, D., Seppelt, B., Markkula, G., Piccinini, G. B., & Victor, T. (2018). Great expectations: A predictive processing account of automobile driving. *Theoretical Issues in Ergonomics Science*, 19(2), 156–194. <https://doi.org/10.1080/1463922X.2017.1306148>

Fagerlönn, J. (2010). *Distracting Effects of Auditory Warnings on Experienced Drivers*. 6.

Feldman, H., & Friston, K. J. (2010). Attention, Uncertainty, and Free-Energy. *Frontiers in Human Neuroscience*, 4, 215. <https://doi.org/10.3389/fnhum.2010.00215>

Figma. (2024). *Kostenloses Design-Tool für Websites & mehr*. Figma.

<https://www.figma.com/de-de/design/>

- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138. <https://doi.org/10.1038/nrn2787>
- Friston, K., Samothrakis, S., & Montague, R. (2012). Active inference and agency: Optimal control without cost functions. *Biological Cybernetics*, 106(8), 523–541. <https://doi.org/10.1007/s00422-012-0512-8>
- Goekoop, R., & de Kleijn, R. (2021). How higher goals are constructed and collapse under stress: A hierarchical Bayesian control systems perspective. *Neuroscience & Biobehavioral Reviews*, 123, 257–285. <https://doi.org/10.1016/j.neubiorev.2020.12.021>
- Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). “Take over!” How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1), 1938–1942. <https://doi.org/10.1177/1541931213571433>
- Gold, C., Happee, R., & Bengler, K. (2018). Modeling take-over performance in level 3 conditionally automated vehicles. *Accident Analysis & Prevention*, 116, 3–13. <https://doi.org/10.1016/j.aap.2017.11.009>
- Hart, S. G., & Field, M. (2006). *NASA-TASK LOAD INDEX (NASA-TLX); 20 YEARS LATER*.
- Heikooop, D. D., de Winter, J. C. F., van Arem, B., & Stanton, N. A. (2018). Effects of mental demands on situation awareness during platooning: A driving simulator study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 58, 193–209. <https://doi.org/10.1016/j.trf.2018.04.015>
- Hesp, C., Smith, R., Parr, T., Allen, M., Friston, K. J., & Ramstead, M. J. D. (2021). Deeply Felt Affect: The Emergence of Valence in Deep Active Inference. *Neural Computation*, 33(2), 398–446. *Neural Computation*. https://doi.org/10.1162/neco_a_01341

- iMotions Inc. (2021, November 2). *Exploring Mental Workload with iMotions—iMotions*.
<https://imotions.com/blog/learning/product-guides/exploring-mental-workload-with-imotions/>
- Jeon, M. (2019). Multimodal Displays for Take-over in Level 3 Automated Vehicles while Playing a Game. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–6. <https://doi.org/10.1145/3290607.3313056>
- Jiang, M., Mieronkoski, R., Rahmani, A. M., Hagelberg, N., Salanterä, S., & Liljeberg, P. (2017). Ultra-short-term analysis of heart rate variability for real-time acute pain monitoring with wearable electronics. *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 1025–1032.
<https://doi.org/10.1109/BIBM.2017.8217798>
- Khai Ooi, J. S., Ahmad, S. A., Chong, Y. Z., Md Ali, S. H., Ai, G., & Wagatsuma, H. (2016). Driver emotion recognition framework based on electrodermal activity measurements during simulated driving conditions. *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 365–369.
<https://doi.org/10.1109/IECBES.2016.7843475>
- Kiverstein, J., Miller, M., & Rietveld, E. (2019). The feeling of grip: Novelty, error dynamics, and the predictive brain. *Synthese*, 196(7), 2847–2869.
- Kraut, V. (2022, Juni 23). Effects of User Interface Assertiveness on Acute Stress and Performance in a Takeover Request. *Proceedings of the MEi:CogSci Conference 2022*. 16th MEi:CogSci Conference, Zagreb, Croatia.
- Kraut, V., Motnikar, L., Mirnig, A., & Fröhlich, P. (2023). Assertive Takeover Requests: Immediate and Sustained Effects on Stress and Performance. *Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 103–111. <https://doi.org/10.1145/3580585.3607174>

- Kujala, T., & Lappi, O. (2021). Inattention and Uncertainty in the Predictive Brain. *Frontiers in Neuroergonomics*, 2, 718699. <https://doi.org/10.3389/fnrgo.2021.718699>
- Large, D. R., Burnett, G., Salanitri, D., Lawson, A., & Box, E. (2019). A Longitudinal Simulator Study to Explore Drivers' Behaviour in Level 3 Automated Vehicles. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 222–232. <https://doi.org/10.1145/3342197.3344519>
- Li, N., Misu, T., & Tao, F. (2017). Understand driver awareness through brake behavior analysis: Reactive versus intended hard brake. *2017 IEEE Intelligent Vehicles Symposium (IV)*, 1523–1528. <https://doi.org/10.1109/IVS.2017.7995925>
- Lohani, M., Payne, B. R., & Strayer, D. L. (2019). A Review of Psychophysiological Measures to Assess Cognitive States in Real-World Driving. *Frontiers in Human Neuroscience*, 13, 57. <https://doi.org/10.3389/fnhum.2019.00057>
- Luetge, C. (2017). The German Ethics Code for Automated and Connected Driving. *Philosophy & Technology*, 30(4), 547–558. <https://doi.org/10.1007/s13347-017-0284-0>
- Ma, S., Zhang, W., Yang, Z., Kang, C., Wu, C., Chai, C., Shi, J., Zeng, Y., & Li, H. (2021). Take over Gradually in Conditional Automated Driving: The Effect of Two-stage Warning Systems on Situation Awareness, Driving Stress, Takeover Performance, and Acceptance. *International Journal of Human–Computer Interaction*, 37(4), 352–362. <https://doi.org/10.1080/10447318.2020.1860514>
- Marshall, D. C., Lee, J. D., & Austria, P. A. (2007). Alerts for In-Vehicle Information Systems: Annoyance, Urgency, and Appropriateness. *Human Factors*, 49(1), 145–157. <https://doi.org/10.1518/001872007779598145>

- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12., überarbeitete Auflage). Basel. https://content-select.com/index.php?id=bib_view&ean=9783407293930
- Merat, N., Jamson, A. H., Lai, F. C. H., Daly, M., & Carsten, O. M. J. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, 27, 274–282. <https://doi.org/10.1016/j.trf.2014.09.005>
- NASA TLX: Task Load Index. (2024). National Aeronautics and Space Administration. <https://humansystems.arc.nasa.gov/groups/TLX/>
- NHTSA. (2013). *Automated Vehicles for Safety*. United States Department of the National Highway Traffic Safety Administration: Automated Vehicles for Safety. <https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety>
- Nozari, S., Krayani, A., Marin, P., Marcenaro, L., Martin, D., & Regazzoni, C. (2023). Autonomous Driving Based on Imitation and Active Inference. In M. Valle, D. Lehmhus, C. Gianoglio, E. Ragusa, L. Seminara, S. Bosse, A. Ibrahim, & K.-D. Thoben (Hrsg.), *Advances in System-Integrated Intelligence* (Bd. 546, S. 13–22). Springer International Publishing. https://doi.org/10.1007/978-3-031-16281-7_2
- Ongul, B. S., & Ongun, M. A. (2020). Driver Stress Level Change Detection Using Model-Based System Engineering Method. *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 1–5. <https://doi.org/10.1109/ISMSIT50672.2020.9255072>
- Pakdamanian, E., Hu, E., Sheng, S., Kraus, S., Heo, S., & Feng, L. (2022). Enjoy the Ride Consciously with CAWA: Context-Aware Advisory Warnings for Automated Driving. *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 75–85. <https://doi.org/10.1145/3543174.3546835>

- Paredes, P. E., Ordonez, F., Ju, W., & Landay, J. A. (2018). Fast & Furious: Detecting Stress with a Car Steering Wheel. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–12. <https://doi.org/10.1145/3173574.3174239>
- Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. <https://doi.org/10.7551/mitpress/12441.001.0001>
- Pezzulo, G., & Cisek, P. (2016). Navigating the Affordance Landscape: Feedback Control as a Process Model of Behavior and Cognition. *Trends in Cognitive Sciences*, 20(6), 414–424. <https://doi.org/10.1016/j.tics.2016.03.013>
- Polar H10 | Polar USA. (2024). <https://www.polar.com/us-en/sensors/h10-heart-rate-sensor>
- Politis, I., Brewster, S., & Pollick, F. (2014). Speech Tactons Improve Speech Warnings for Drivers. *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 1–8. <https://doi.org/10.1145/2667317.2667318>
- Politis, I., Brewster, S., & Pollick, F. (2015). To Beep or Not to Beep? Comparing Abstract versus Language-Based Multimodal Driver Displays. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3971–3980. <https://doi.org/10.1145/2702123.2702167>
- Python TM. (2024, August 29). *Welcome to Python.org*. Python.Org. <https://www.python.org/>
- Riener, A., Ferscha, A., & Aly, M. (2009). Heart on the road: HRV analysis for monitoring a driver's affective state. *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 99–106. <https://doi.org/10.1145/1620509.1620529>
- Roche, F., Somieski, A., & Brandenburg, S. (2019). Behavioral Changes to Repeated Takeovers in Highly Automated Driving: Effects of the Takeover-Request Design

- and the Nondriving-Related Task Modality. *Human Factors*, 61(5), 839–849.
<https://doi.org/10.1177/0018720818814963>
- SAE International. (2021, April). *J3016_202104: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles - SAE International*. https://www.sae.org/standards/content/j3016_202104/
- Samuel, S., Borowsky, A., Zilberstein, S., & Fisher, D. L. (2016). Minimum Time to Situation Awareness in Scenarios Involving Transfer of Control from an Automated Driving Suite. *Transportation Research Record*, 2602(1), 115–120.
<https://doi.org/10.3141/2602-14>
- Schneider, E. E. (2020). *Perceived Stress Scale | Klinische Psychologie und Neuropsychologie*. <https://www.klinische-psychologie-und-neuropsychologie.uni-mainz.de/perceived-stress-scale/>
- Seth, A. K. (2014). The Cybernetic Bayesian Brain. In T. Metzinger & J. M. Windt (Hrsg.), *Open MIND*. Open MIND. Frankfurt am Main: MIND Group.
<https://doi.org/10.15502/9783958570108>
- Seth, A. K., & Friston, K. J. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions: Biological Sciences*, 371(1708), 1–10.
- Shaffer, F., Shearman, S., & Meehan, Z. M. (2016). The Promise of Ultra-Short-Term (UST) Heart Rate Variability Measurements. *Biofeedback*, 44(4), 229–233.
<https://doi.org/10.5298/1081-5937-44.3.09>
- Sladky, R., Kargl, D., Haubensak, W., & Lamm, C. (2024). An active inference perspective for the amygdala complex. *Trends in Cognitive Sciences*, 28(3), 223–236.
<https://doi.org/10.1016/j.tics.2023.11.004>
- Sugiono, S., Widhayanuriyawan, D., & Andriyani, D. P. (2018). Mental Stress Evaluation of Car Driver in Different Road Complexity Using Heart Rate Variability (HRV)

- Analysis. *Proceedings of the 5th International Conference on Bioinformatics Research and Applications*, 90–94. <https://doi.org/10.1145/3309129.3309145>
- Tan, X., & Zhang, Y. (2024). A Computational Cognitive Model of Driver Response Time for Scheduled Freeway Exiting Takeovers in Conditionally Automated Vehicles. *Human Factors*, 66(5), 1583–1599. <https://doi.org/10.1177/00187208221143028>
- Wei, R., McDonald, A. D., Garcia, A., & Alambeigi, H. (2022). Modeling Driver Responses to Automation Failures With Active Inference. *IEEE Transactions on Intelligent Transportation Systems*, 23(10), 18064–18075. *IEEE Transactions on Intelligent Transportation Systems*. <https://doi.org/10.1109/TITS.2022.3155381>
- Wiese, W., & Metzinger, T. K. (2017). Vanilla PP for Philosophers: A Primer on Predictive Processing. *Vanilla PP for Philosophers: A Primer on Predictive Processing. PPP*. <https://doi.org/10.15502/9783958573024>
- Wijsman, J., Grundlehner, B., Penders, J., & Hermens, H. (2010). *Trapezius muscle EMG as predictor of mental stress*.
- Winner, H., Hakuli, S., Lotz, F., & Singer, C. (2015). *Handbuch Fahrerassistenzsysteme—Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort* (3. Aufl.). Springer Vieweg.
- Wong, P. N. Y., Brumby, D., Babu, H., & Kobayashi, K. (2019). „Watch Out!": Semi-Autonomous Vehicles Using Assertive Voices to Grab Distracted Drivers' Attention. 1–6. <https://doi.org/10.1145/3290607.3312838>
- World Health Organization. (2022, Juni 20). *Road traffic injuries*. Road Traffic Injuries. <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>
- Wu, H., Wu, C., Lyu, N., & Li, J. (2022). Does a faster takeover necessarily mean it is better? A study on the influence of urgency and takeover-request lead time on takeover performance and safety. *Accident Analysis & Prevention*, 171, 106647. <https://doi.org/10.1016/j.aap.2022.106647>

Yerkes, R. M., & Dodson, J. D. (1908). The Relation of Strength of Stimulus to Rapidity of Habit Formation. *Journal of Comparative Neurology & Psychology*, 18, 459–482.
<https://doi.org/10.1002/cne.920180503>

Zeeb, K., Buchner, A., & Schrauf, M. (2015). What determines the take-over time? An integrated model approach of driver take-over after automated driving. *Accident Analysis & Prevention*, 78, 212–221. <https://doi.org/10.1016/j.aap.2015.02.023>

Zhang, B., De Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, 64, 285–307.
<https://doi.org/10.1016/j.trf.2019.04.020>

List of Figures

Figure 1 Ratings on perceived assertiveness in pilot testing.....	23
Figure 2 Final versions of the visual user interface for the learning and test settings	28
Figure 3 Close-up of the mock-up for the visual user interface in autopilot	28
Figure 4 Close-up of the heads-up display during manual driving after an assertive TOR. 29	
Figure 5 Illustration of a lap with time estimates	31
Figure 6 Obstacle course layouts	33
Figure 7 Visual Analogue Scales for self-reported ratings	36
Figure 8 Workload assessment in NASA TLX	43
Figure 9 Adaptation effects for reaction time (A) and braking force (B)	47
Figure 10 Adaptation effects for average speed	48
Figure 11 Adaptation effects for stress ratings (A) and performance ratings (B)	48
Figure 12 Illustration of driver interaction with an ADAS from a PP perspective.	53

List of Tables

Table 1 Levels of driving automation assistance	7
Table 2 Alerts tested in the pilot study	21
Table 3 Dependent variables with descriptions and units	26
Table 4 Statistical results for immediate post-TOR measures.....	40
Table 5 Statistical results for sustained post-TOR measures	41
Table 6 Statistical results for subjective ratings on stress and performance.....	42
Table 7 Linear regression statistics of all measured variables	46

Appendix A: Pilot Study Procedure

Preparation

- Introduction and signing consent form
- Attaching heart rate sensor
- Participants learn the simple n-back task

Preliminary Sound Rating: Rating sounds to assess assertiveness

Learning scenario

- Manual driving in a scenario with
- Automatic driving
- Takeover process, including takeover via pressing a button on steering wheel, rotating steering wheel by > 2 degrees, or depressing the brake pedal with $> 0.5\text{N}$.

Main driving task: Determining the most assertive TOR

Four laps:

- Automatic driving with n-back task for 90 seconds
- TOR is presented 7s before obstacle course, ADAS turns off
- Drivers navigate through obstacles: 10 obstacles each, spaced 20m apart.
- Drive 1km on unobstructed highway
- ADAS turns back on automatically
- Perceived Stress Test on 1 -10 scale

Additional road conditions:

- Three lane highway per direction. Ego vehicle on middle lane.
- Simulated fog (35% percent concentration)
- Moderate to heavy traffic with motor sounds, 15 road users/km in both direction, see (Du et al., 2020).
- Safety-critical events were induced by inactive vehicles and obstacles placed on the lanes in the form of obstacle courses. There were two different obstacle course patterns, each of which was repeated once. These obstacle courses later formed the basis for Courses A and B in the main study.

Ending

- Detaching and cleaning sensors
- Post: Interview (5 Minutes)
- NASA TLX
- Demographic data

B.1 NASA Task Load Index

Figure 8.6

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

81

B.2 German 10-item Perceived Stress Scale

Schneider (2020) provided a German translation of the 10-item perceived stress scale, originally developed by Cohen et al. (1983).

Instructions, items and subscales of a German adaption of the 10-item Perceived Stress Scale (PSS-10)

From: Schneider, E. E., Schönfelder, S., Domke-Wolf, M., & Wessa, M. (in press). Measuring stress in clinical and nonclinical subjects using a German adaptation of the Perceived Stress Scale. International Journal of Clinical and Health Psychology.

PERCEIVED STRESS SCALE - 10

Die folgenden Fragen beschäftigen sich mit Ihren Gedanken und Gefühlen während des letzten Monats. Bitte geben Sie für jede Frage an, wie oft sie in entsprechender Art und Weise gedacht oder gefühlt haben.

		nie	Fast nie	Manchmal	Ziemlich oft	sehr oft
1	Wie oft waren Sie im letzten Monat aufgewühlt, weil etwas unerwartet passiert ist?	1	2	3	4	5
2	Wie oft hatten Sie im letzten Monat das Gefühl, nicht in der Lage zu sein, die wichtigen Dinge in Ihrem Leben kontrollieren zu können?	1	2	3	4	5
3	Wie oft haben Sie sich im letzten Monat nervös und gestresst gefühlt?	1	2	3	4	5
4	Wie oft waren Sie im letzten Monat zuversichtlich, dass Sie fähig sind, ihre persönlichen Probleme zu bewältigen?	1	2	3	4	5
5	Wie oft hatten Sie im letzten Monat das Gefühl, dass sich die Dinge zu Ihren Gunsten entwickeln?	1	2	3	4	5
6	Wie oft hatten Sie im letzten Monat den Eindruck, nicht all Ihren anstehenden Aufgaben gewachsen zu sein?	1	2	3	4	5
7	Wie oft waren Sie im letzten Monat in der Lage, ärgerliche Situationen in Ihrem Leben zu beeinflussen?	1	2	3	4	5
8	Wie oft hatten Sie im letzten Monat das Gefühl, alles im Griff zu haben?	1	2	3	4	5
9	Wie oft haben Sie sich im letzten Monat über Dinge geärgert, über die Sie keine Kontrolle hatten?	1	2	3	4	5
10	Wie oft hatten Sie im letzten Monat das Gefühl, dass sich so viele Schwierigkeiten angehäuft haben, dass Sie diese nicht überwinden konnten?	1	2	3	4	5

Skala Hilflosigkeit (H): Summe der Items 1, 2, 3, 6, 9, 10; Skala Selbstwirksamkeit (S): Summe der Items 4, 5, 7, 8. Für die Berechnung des Gesamtscores müssen die Items 4, 5, 7 und 8 der Selbstwirksamkeitsskala invertiert werden. Der Gesamtscore berechnet sich aus der Summe der Items der Hilflosigkeitsskala und der Summe der invertierten Items der Selbstwirksamkeitsskala. Höhere Werte deuten auf ein erhöhtes Stresslevel hin.

Original Items in English (Cohen et al., 1983)

The questions in this scale ask you about your feelings and thoughts during the last month. In each case, you will be asked to indicate how often you felt or thought a certain way.

- 1 PH In the last month, how often have you been upset because of something that happened unexpectedly?
- 2 PH In the last month, how often have you felt that you were unable to control the important things in your life?
- 3 PH In the last month, how often have you felt nervous and "stressed"?
- 4 PSE In the last month, how often have you felt confident about your ability to handle your personal problems?
- 5 PSE In the last month, how often have you felt that things were going your way?
- 6 PH In the last month, how often have you found that you could not cope with all the things that you had to do?
- 7 PSE In the last month, how often have you been able to control irritations in your life?
- 8 PSE In the last month, how often have you felt that you were on top of things?
- 9 PH In the last month, how often have you been angered because of things that were outside of your control?
- 10 PH In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

Answer range: 1 = never, 2 = almost never, 3 = sometimes, 4 = fairly often, 5 = very often; PH=perceived helplessness subscale; PSE=perceived self-efficacy; Items 4, 5, 7 and 8 are reverse scored for the total score. The PH subscale is computed by summing up items 1, 2, 3, 6, 9 and 10; the PSE subscale is computed by summing up items 4, 5, 7 and 8; the total score is the sum of all PH and reversed PSE items. Higher scores reflect greater levels of stress

B.3 English 10-item Perceived Stress Scale

Cohen et al. devised the perceived stress scale in 1983. Below are the instructions and the ten-item variation (Cohen et al., 1983, S. 394).

Instructions:

The questions in this scale ask you about your feelings and thoughts during the last month. In each case, you will be asked to indicate how often you felt or thought a certain way. Although some of the questions are similar, there are differences between them, and you should treat each one as a separate question. The best approach is to answer each question fairly quickly. That is, don't try to count up the number of times you felt a particular way, but rather indicate the alternative that seems like a reasonable estimate. For each question choose from the following alternatives:

- 0. never
- 1. almost never
- 2. sometimes
- 3. fairly often
- 4. very often

1. In the last month, how often have you been upset because of something that happened unexpectedly?
2. In the last month, how often have you felt that you were unable to control the important things in your life?
3. In the last month, how often have you felt nervous and "stressed"?
4. ^a In the last month, how often have you dealt successfully with irritating life hassles?
5. ^a In the last month, how often have you felt that you were effectively coping with important changes that were occurring in your life?
6. ^a In the last month, how often have you felt confident about your ability to handle your personal problems?
7. ^a In the last month, how often have you felt that things were going your way?
8. In the last month, how often have you found that you could not cope with all the things that you had to do?
9. ^a In the last month, how often have you been able to control irritations in your life?
10. ^a In the last month, how often have you felt that you were on top of thing

^a Scored in the reverse direction.