Results of the MEDIATOR driving simulator evaluation studies

Deliverable D3.3 – WP3 – Public

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 814735.

Results of the MEDIATOR driving simulator evaluation studies

Work package 3, Deliverable D3.3

Please refer to this report as follows:
Borowsky, A., Schwarz-Chassidim, H., Hollander, C., Rauh, N., Enhuber, S., Oron-Gilad, T., Beggiato, M. (2023). Results of the MEDIATOR driving simulator evaluation studies, Deliverable D3.3 of the H2020 project MEDIATOR.

Project details:	
Project start date:	01/05/2019
Duration:	48 months
Project name:	MEDIATOR – MEdiating between Driver and Intelligent Automated Transport systems on Our Roads
Coordinator:	Prof.dr Nicole van Nes | SWOV – Institute for Road Safety Research Bezuidenhoutseweg 62, 2594 AW, The Hague, The Netherlands
	This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 814735.

Deliverable details:	
Version:	Final
Dissemination level:	PU (Public)
Due date:	31/12/2022
Submission date:	30/03/2023

Lead contractor for this deliverable:
Avinoam Borowsky – Ben-Gurion University of the Negev

Report Author(s):	
	Borowsky, A., Schwarz-Chassidim, H., Oron-Gilad, T. - Ben-Gurion University of the Negev (BGU), Israel Hollander, C., Rauh, N., Enhuber, S., Beggiato, M. - Chemnitz University of Technology (TUC), Germany

Revision history

Date	Version	Reviewer	Description
09/12/2022	Integrated Draft 1 - studies TUC and BGU fatigue	Christer Ahlström - VTI National Road and Transport Research Institute, Sweden	Internal review
12/03/2023	Integrated Draft 2 – all studies	Christer Ahlström - VTI National Road and Transport Research Institute, Sweden	Internal review
20/03/2023	Final draft	Matthias Beggiato (WP 3 leader) – TU Chemnitz, Germany, Ernst Verschragen (QA Officer) – SWOV Institute for Road Safety Research, The Netherlands	Final editing and checks
30/03/2023	Final deliverable	Michiel Christoph (Technical project coordinator) – SWOV Institute for Road Safety Research, The Netherlands	Approved

Legal Disclaimer

All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user, therefore, uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission and INEA has no liability in respect of this document, which is merely representing the authors' view.

© 2023 by MEDIATOR Consortium

Table of contents

List of abbreviations

ADAS	Advanced Driver Assistance Systems
AOI	Area Of Interest
ATI	Affinity for Technology Interaction
BGU	Ben-Gurion University of the Negev
CI	Confidence Interval
CM	Continuous Mediation (automation level)
ECG	Electrocardiogram
EM	Estimated Mean
GLMM	Generalized Linear Mixed Models
HMI	Human Machine Interface
HRV	Heart Rate Variability
KSS	Karolinska Sleepiness Scale
LMM	Linear Mixed Models
M	Mean value
NDRT	Non-Driving Related Tasks
NHTSA	National Highway Traffic Safety Administration
PAV	Partially Automated Vehicle
rmANOVA	repeated measures Analysis Of Variance
RMSSD	Root Mean Square of Successive Differences
RQ	Research Question
SAE	Society of Automotive Engineers
SB	Driver Standby (automation level)
SD	Standard Deviation
SDNN	Standard Deviation of the interbeat intervals of Normal sinus heart beats
SE	Standard Error
SUS	System Usability Scale
TOC	Transition Of Control
TOR	Take Over Request
TUC	Technische Universität Chemnitz / Chemnitz University of Technology
UC	Use Case
WP	Work Package

List of tables

List of figures

Mediator

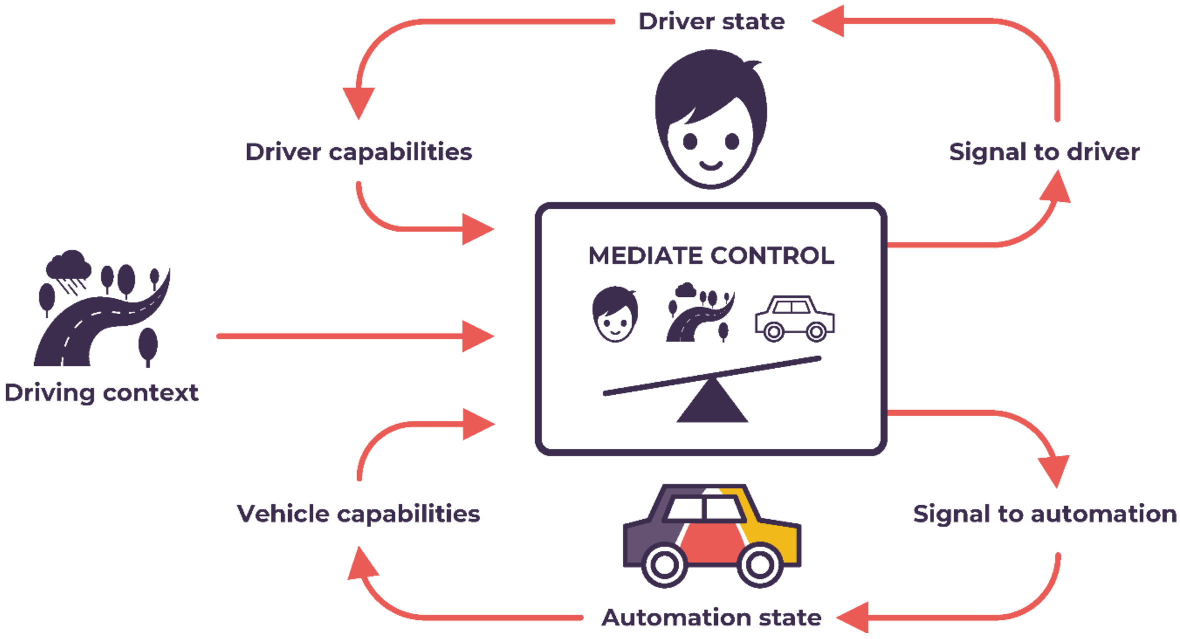
About MEDIATOR

MEDIATOR, a 4-year project led by SWOV, started on May 1, 2019.

MEDIATOR will develop a mediating system for drivers in semi-automated and highly automated vehicles, resulting in safe, real-time switching between the human driver and automated system based on who is most fit to drive. MEDIATOR pursues a paradigm shift away from a view that prioritises either the driver or the automation, instead integrating the best of both.

Vision

Automated transport technology is developing rapidly for all transport modes, with huge safety potential. The transition to full automation, however, brings new risks, such as mode confusion, overreliance, reduced situational awareness and misuse. The driving task changes to a more supervisory role, reducing the task load and potentially leading to degraded human performance. Similarly, the automated system may not (yet) function in all situations. The objective of the mediator system is to intelligently assess the strengths and weaknesses of both the driver and the automation and mediate between them, while also taking into account the driving context.



The MEDIATOR system will constantly weigh driving context, driver state and vehicle automation status, while personalising its technology to the drivers' general competence, characteristics, and preferences.

MEDIATOR will optimise the safety potential of vehicle automation during the transition to full (level 5) automation. It will reduce risks, such as those caused by driver fatigue or inattention, or on the automation side imperfect automated driving technology. MEDIATOR will facilitate market exploitation by actively involving the automotive industry during the development process.
To accomplish the development of this support system MEDIATOR will integrate and enhance existing knowledge of human factors, taking advantage of the of expertise in other transport modes (aviation, rail and maritime). It will develop and adapt available technologies for real-time data collection, storage and analysis and incorporate the latest artificial intelligence techniques, such as deep learning.

Partners

MEDIATOR will be carried out by a consortium of highly qualified research and industry experts, representing a balanced mix of top universities and research organisations as well as several OEMs and suppliers. The consortium, supported by an international Industrial Advisory Board and a Scientific Advisory Board, will also represent all transport modes, maximising input from, and transferring results to, aviation, maritime and rail (with mode-specific adaptations).

Executive summary

The Mediator system is an intelligent mediating system that enables safe switching between the human driver and automated system. The main aim of Work Package 3 (WP3) "Testing & Evaluation" within the MEDIATOR project was to evaluate the functionality of the Mediator system using both driving simulators and field studies with real drivers. This functionality evaluation includes assessing the system's performance, reliability, and effects on driving safety, as well as the acceptance, trustworthiness, perceived safety, and user-friendliness of different groups of users. This deliverable is focused on the driving simulator studies conducted in the framework of WP3. Two studies were conducted at Ben-Gurion University of the Negev (BGU) and one at Chemnitz University of Technology (TUC). The main objectives of the driving simulator studies reported in this deliverable were:

1. To evaluate the effects of preventive and corrective mediation in mitigating the adverse effects of passive fatigue and distraction on driving performance and hazard perception. Preventive and corrective mediation is a Mediator-related concept for preventing drivers from becoming fatigued or distracted or correcting drivers' performance if they are already in a state of fatigue or distraction.
2. To evaluate the longer-term effects of the Mediator Human Machine Interface (HMI) concept by inviting participants for a second driving session a week after their first experience with the system.
3. To study users' evaluation of the Mediator system / HMI concept in comfort related use cases with regard to acceptance, trust, usability, comfort and transitions of control (TOC) while addressing driver characteristics.

Summary driving simulator studies on fatigue and distraction

The two driving simulator studies at BGU focused on Continuous Mediation (CM) automated driving conditions investigating drivers' states of fatigue and distraction. CM is similar to the SAE L2 definition (i.e., partially automated driving). In this driving condition, the driver must continuously monitor the road and the automation. One study focused on fatigue, and one on distraction. Both studies investigated the effectiveness of the Mediator HMI in mitigating the driver-state-related adverse effects in the short term and the long term (a week after the first experience with the system). An audio-visual Trivia game was used as a preventive mediation.

Main results showed that in the fatigue study, using an audio-visual Trivia game as a preventive mediation helped in maintaining situation awareness compared to drivers who were not playing Trivia, despite the fact that there were no fatigue-related differences in physiological and questionnaire data. As for the hazard notification system, it was found helpful in attracting drivers' attention (regardless of whether they played Trivia or not) to typical on-road hazards and improved safety. Concerning whether these positive effects persist over time, there was consistent evidence that the Mediator HMI was effective in both driving sessions of the fatigue study. Participants who experienced Trivia and hazard notifications as preventive mediation perceived the interface as more convenient, usable, acceptable, and safe than the Non-Trivia group in both drives. In addition, while for the Non-Trivia group, there were some questionnaire dimensions where their scoring decreased between the first and second driving sessions, for the Trivia group, it remained unchanged. Similarly, in the distraction study (where the Trivia game was absent), participants perceived the Mediator HMI as useful, safe, and trusted. Participants also indicated their willingness to adopt this type of system.

Summary driving simulator study on comfort-related use cases

The driving simulator study at Chemnitz University of Technology focussed on comfort use cases with related TOC, simulated automation degradation and related TOC by the human driver, as well as driver characteristics. The study aimed primarily on the user's evaluation of the Mediator system / HMI (i.e. acceptance, trust, usability, comfort, experience of TOC), including the analysis of driver characteristics effects such as age, gender and driving experience. For this purpose, a balanced sample regarding age (19 to 75 years) and gender (49% female) was recruited. All participants experienced Mediator with its HMI elements in four different conditions. Next to CM, also the Driver Standby (SB) automated condition was tested which is comparable to SAE level 3 (conditional automation). Several standardized questionnaires, self-designed items and interview questions were used to examine drivers' evaluation of Mediator for different use cases.

Main results revealed very high intention to use Mediator in future cars (88%) as well as very high driving comfort (90%), easiness to use (95%) and usability ratings (SUS score 85 out of 100) when driving with Mediator. Additionally, drivers' expressed high acceptance (*M* = 1.21 on a scale ranging from -2 to +2) and trust (*M* = 5.5 on a scale ranging from 1 to 7). Further, 86% of the participants expected an increase in road safety when driving with Mediator. The analyses on driver characteristics did not show systematic differences with regard to gender and driving experience, and only small differences between age groups (i.e. more positive evaluations of older age groups) in the balanced sample. This gives a first indication that Mediator follows the "Design for All" principle, i.e. does not systematically discriminate certain groups. Driving with Mediator and in automated driving mode was preferred compared to manual driving without Mediator. In addition, driving in SB was preferred compared to driving in CM. Specifically, when participants experienced a traffic jam scenario under various levels of driving automation, they felt this was an appropriate situation where automation should take over and free the driver to engage with other tasks.

1. Introduction

The Mediator system is an intelligent mediating system that enables safe switching between the human driver and automated system. The main aim of Work Package 3 (WP3) “Testing & Evaluation” within the MEDIATOR project was to evaluate the functionality of the Mediator system using both driving simulators and field studies with real drivers. This functionality evaluation includes assessing the system’s performance, reliability, and effects on driving safety, as well as the acceptance, trustworthiness, perceived safety, and user-friendliness of different groups of users. This deliverable is focused on the driving simulator studies conducted in the framework of WP3. The field studies conducted in WP3 are reported in MEDIATOR deliverable D3.4 (Fiorentino et al., 2023).

The following section describes essential aspects from previous MEDIATOR work that provided the foundations for the driving simulator testing and evaluation experiments conducted in WP3.

1.1. General evaluation principles

The MEDIATOR testing and evaluation strategy in WP3 consequently took into account the specific features of the different evaluation methods (computer simulation study, driving simulator studies, on-road studies). Thus, on the one hand a **complementary approach** was followed, taking advantage of the specific methodological assets in combination with corresponding UCs and Mediator components/prototypes. Driving simulator studies allow for high experimental control and a variety of focuses. They can be replicated and also introduce potentially critical scenarios without harm. Simulators provide extensive options for integrating additional components (e.g., driver state monitoring systems, HMI components, additional tasks and measurements). Research designs can flexibly adapt to study requirements, including various user and technical evaluation options. Major drawbacks are the limited degree of realism due to several just simulated aspects, connected with potentially reduced external validity of results. On the other hand, **comparability** between all MEDIATOR evaluation studies was maximised by using the same UCs, the same components (e.g. HMI) and, whenever possible, the same measurements such as questionnaires. These two basic principles were agreed in the consortium and led to a specific work organisation:

Complementary approach: Each evaluation study was planned by the respective study team, considering study-specific UCs, research questions and components of the Mediator system. It allowed for exploiting the maximum benefit of each evaluation platform/method. Thus, each study was planned by each study team with close connections to the teams working on the technical realisation of Mediator components. Dedicated “platform subteams” were formed to best align technical developments with experimental evaluation designs for each study. The complementary approach is also reflected in the main chapter structure of this deliverable, presenting each study by the respective study team BGU or TUC.

Comparability of results: In parallel to the complementary approach, maximum comparability of study results shall be achieved. One pillar for ensuring comparability were the UCs defined in deliverable D1.4 (Cleij et al., 2020), which ensure a common set of driving contexts, driver states, automation states and Mediator functionalities across the studies. A second important aspect of comparability was the common design concept for the HMI. Finally, common questionnaires and interview guidelines as dependent variables of the user evaluations were used. The list of questionnaires was communicated and coordinated among the partners to ensure maximum comparability between the empirical studies. The questionnaires used in each driving simulator study are described in Table 2.2 for BGU and in Table 3.4 for TUC.

1.2. Essential aspects from prior work in MEDIATOR

The evaluation and testing of the Mediator system reported in this deliverable rely on two essential aspects of previous MEDIATOR work. (1) Mediator UCs including levels of automation CM and SB defined in deliverable D1.4 (Cleij et al., 2020) and (2) the general Mediator HMI concept (Grondelle et al., 2021).

In D1.4 (Cleij et al., 2020) a total of ten UCs were developed to define the scope of the MEDIATOR project. The Mediator system was designed to work in ten possible UCs, and the driving simulator experiments addressed a subset of these. Next, we describe the relevant UCs for the driving simulator studies and their implications in terms of HMI requirements.

Preventive and corrective mediation use cases

Two studies at BGU focused on fatigue and distraction under CM. The main goal of these studies was to prevent drivers from becoming passively fatigued or distracted (preventive mediation) or, in case they are already in such a state, to correct their degraded performance (corrective mediation). For this purpose, three use cases UC4, UC7, and UC8 were evaluated:

- **Use Case 4 (corrective SB):** If the driver monitoring system detects that the driver is drowsy or distracted, the HMI will try to get the driver back into the loop.
- **Use Case 7 (preventive CM):** The HMI's role is to keep the driver in the loop (e.g., active settings, active monitoring), preventing the driver from becoming fatigued or distracted.
- **Use Case 8 (corrective CM):** The Mediator role is to assist the driver if they are already fatigued or distracted and bring them back into the loop.

Comfort-related use cases

The third driving simulator study at TUC focused on evaluating users' preferences for the Mediator HMI concept under CM and SB driving conditions where the automated vehicle applied different driving behaviours when approaching a traffic jam. This study examined the following five use cases: UC2, UC3, UC5, UC6, and UC9. For UCs 3, 5, and 6 only one of the two possible sub-use cases was examined (i.e., UC 3b, UC 5a, UC 6b). The drivers experienced four conditions (1 – baseline, 2 – CM, 3 – SB, 4 – SB + close approach) of the same driving scenario, each covering several UCs. The examined UCs were:

- **Use Case 2 - Driver takes back control:** The driver uses the HMI to indicate a desire to take back vehicle control. The Mediator system reacts by confirming that the driver is fit enough to drive and guide the takeover.
- **Use Case 3b - Comfort takeover (human to automation):** Comfort takeover from human to automation initiated by the Mediator system: The Mediator system detects an event, such as receiving a text message or an upcoming traffic jam, from which it concludes that the driver comfort could be improved. The system reacts by suggesting a takeover to automation.
- **Use Case 5a - Mediator-initiated takeover (automation to human):** Planned takeover from automation to human initiated by the Mediator system: The automation indicates that the current route leads to automation unfitness as it will leave its operational design domain. The

Mediator system reacts by preparing the driver for and guiding the driver through a non-urgent takeover.

- **Use Case 6b - Comfort CM switch on:** The Mediator system switches on driving in CM due to comfort-related reasons. The Mediator system detects sufficient fitness for driving in CM from which it concludes that the driver comfort could be improved, and reacts by suggesting to switch on CM.
- **Use Case 9 - CM shuts off instantly:** While driving in CM, the automation fitness degrades, and automation can no longer perform its driving task. The Mediator system reacts by communicating to the driver that CM is switching off.

General HMI design concept

The Mediator HMI concept (details in Grondelle et al., 2021) was based on the insights gained by extensive research in three main knowledge gaps: TOC, transparency and information overload, and keeping the driver in the loop. A strategy of research-by-design was implemented, serving as the basis for the HMI concept design.

The recommendations for HMI design consist of adapting Mediator's intervention to the dynamic situation of the triangle: driver, vehicle, and context. An essential condition is that the driver should understand the automation state fully and intuitively. In addition, the HMI evaluated and balanced between driver needs and available technology. Related challenges include trust, comfort, mode awareness, fatigue and distraction, information load, user acceptance, industry acceptance, and learning. The Mediator HMI embraced a holistic approach that attempted to deal with multiple challenges simultaneously rather than individually. This holistic approach provides a solution for interactions of challenges where one challenge may have adverse side effects when dealing with other challenges, requiring evidence-based trade-offs.

In general, it is recommended that the interaction with the driver should be conducted via more than one modality. For the visual inputs, the HMI design must use appropriate and effective colours, referring to established techniques in graphical HMI. The frequency of the interaction and the number of intervention modalities depend on the situation's immediacy. Another principle to be considered is the information content that should encourage the driver to adopt a behaviour that may decrease the risk of an accident. The use cases with corresponding HMI were coordinated over all evaluation studies to align with a similar-common concept as much as possible. Two main communication channels were used: visual and auditory, and the locus of interactions were the dashboard, the infotainment screens, and the steering wheel and ambient lighting. In addition, all evaluation studies used similar icons and colours. Detailed explanation about the implementation of each part and the adaptation to the specific study appears under each study method section.

1.3. Main aims and objectives of the driving simulator studies

The main objectives of the driving simulator studies were:

1. To evaluate the effects of preventive and corrective mediation in mitigating the adverse effects of passive fatigue and distraction on driving performance and hazard perception. Preventive and corrective mediation is a Mediator-related concept for preventing drivers from becoming fatigued or distracted or correcting drivers' performance if they are already in a state of fatigue or distraction.
2. To evaluate the longer-term effects of the Mediator HMI concept by inviting participants for a second driving session a week after their first experience with the system.

3. To study users’ evaluation of the Mediator system/HMI concept in comfort related use cases with regard to acceptance, trust, usability, comfort and TOC while addressing driver characteristics.

Two studies were conducted at BGU and one at TUC to achieve the above objectives. The studies at BGU focused on CM automated driving conditions investigating drivers' states of fatigue and distraction. CM is similar to the SAE L2 definition (i.e., partially automated driving). In these driving conditions, the driver must continuously monitor the road and the automation. One study focused on fatigue, and one on distraction. Both studies investigated the effectiveness of the Mediator HMI in mitigating the driver-state-related adverse effects in the short term and the long term (a week after the first experience with the system).

The driving simulator study at TUC focussed on comfort use cases with related TOC, simulated automation degradation and related TOC by the human driver, as well as driver characteristics. The study aimed primarily on the user’s evaluation of the Mediator system / HMI (i.e. acceptance, trust, usability, comfort, experience of TOC), including the analysis of driver characteristics effects such as age, gender and driving experience. For this purpose, participants experienced Mediator with its HMI elements in four different conditions. Next to CM, also SB was tested which is comparable to SAE level 3 (conditional automation). Several standardized questionnaires, self-designed items and interview questions were used to examine drivers’ evaluation of Mediator for different use cases.

2. Driving simulator studies on fatigue and distraction

This part reports the results, discussion, and conclusions of the driving simulator experiments conducted at BGU (two studies) focused on the driver state and the Mediator HMI concept. One study focused on fatigue, and one on distraction. There are similarities between the two studies regarding their goals and methods. We present the similar/common parts first, then deliver the results of each experiment separately, followed by a short discussion of each experiment individually, then a general discussion of both studies.

The main objectives of the two studies were:

1. To evaluate the effects of preventive and corrective mediation (as implemented in the Mediator HMI concept) in mitigating fatigue and distraction effects on driving performance under CM driving conditions.
2. To evaluate the longer-term effects of the Mediator HMI concept, including preventive and corrective mediation.

To address these two objectives, we developed an in-house concept of preventive and corrective mediation HMI for fatigue and distraction while utilizing a hazard perception paradigm (Borowsky & Oron-Gilad, 2013) as the primary proxy for measuring driving performance degradation. In addition, we used complementary objective and subjective measures such as heart rate variability (HRV), automation disengagement, and standardized questionnaires. To test the HMI effect in the long term, we asked participants to drive in the simulator a week after their first experience with the HMI.

2.1. Introduction to the fatigue and distraction studies

The increased prevalence of automation functions in partially automated vehicles (PAVs) relieves drivers from vehicle control tasks but not from their role as supervisors of the automated system or the driving task (Fisher et al., 2016). This supervision requires continuous mode awareness and passive monitoring, which are monotonous and tedious, resulting in passive fatigue, decreased vigilance, and even compromising the driver's ability to react to a critical event (May & Baldwin, 2009, Körber et al., 2015, Merat et al., 2019). While the primary countermeasure for active fatigue is rest, monotony rupture is the primary countermeasure for passive fatigue. Monotony rupture is not necessarily best achieved through a rest break. If fatigue is due to underload, then activation rather than rest may function as a countermeasure (Oron-Gilad, Ronen, & Shinar, 2008; Gershon et al., 2009). A study conducted at the BMW group laboratories used a motion-based driving simulator to show that participants who engaged only with the driving monitoring task demonstrated higher percentage of eye closure than participants who also engaged in an activating quiz task under partially automated driving (Jarosch, Bellem, & Bengler, 2019). In another study, participants indicated significantly lower fatigue levels when they had to deal with a quiz task for 15 minutes (Schömig et al., 2015). According to an additional study by Lee, Hirano and Itoh (2020), a casual discussion with another passenger while driving under CM significantly reduced driver weariness. Therefore, the authors suggested that developing an in-vehicle system that can chat like a human may help mitigate drivers' fatigue during monotonous driving on highways. These findings show that non-driving related tasks (NDRTs) during tedious or long drives may help fatigued drivers maintain alertness and vigilance.

Conversely, engaging in NDRTs while driving under CM driving conditions may impair the hazard perception performance (Zangi et al., 2022). The tradeoffs between the benefits of using an NDRT as a fatigue countermeasure and the potential distraction it may cause received little attention. A limited number of studies have found support for improvements in driving performance from engaging with an NDRT. For example, Song et al. (2017) have found that providing alertness and a trivia tasks to young drivers led to significantly lower fatigue levels than driving without such tasks. Atchley and Chan (2011) used a monotonous manual drive in a driving simulator to investigate the effects of a concurrent verbal task on drivers' vigilance. Results showed that when engaged in a verbal NDRT task, drivers improved their lane-keeping performance and steering control when drivers' vigilance was lowest (Atchley & Chan, 2011). Further, Oron-Gilad and Hancock (2005) had already raised this trade-off dilemma, see Figure 2.1. They argued that models like the compensatory control model (Hockey, 1997) imply that the system is typically biased toward maintaining high-priority goals under stress at the expense of neglecting others (like passive fatigue and boredom that frequently occur under normal driving conditions in vehicles).

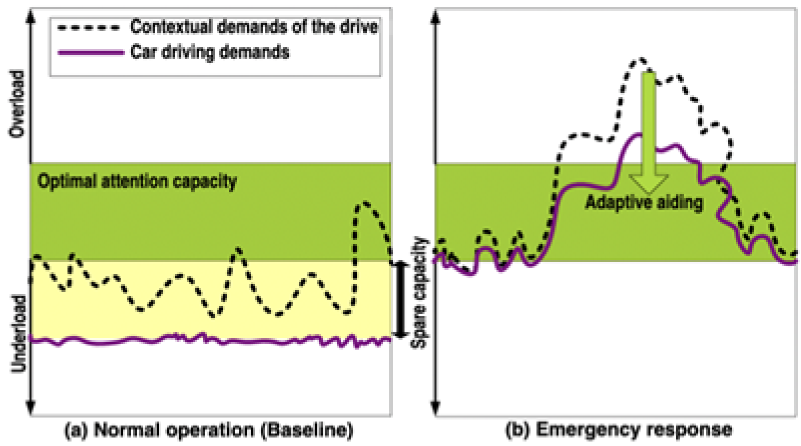


Figure 2.1 Balancing drivers' attention capacity and driving demands (adapted from Oron-Gilad and Hancock, 2005).

The routine operation of a manual car is already very close to underload, while emergency or critical events require immediate attention and exceed the optimal driver's attention level. This driving pattern highlights the challenge of attending to passive fatigue versus causing a distraction that may affect the driver's ability to respond to critical events.

The increasing penetration of automated vehicles into our roads enhances the dilemma of whether to engage with an NDRT to counteract fatigue. In the new era of partially and conditionally automated vehicles, drivers are relieved from vehicle control-related tasks. Consequently, drivers in CM and SB conditions have additional attentional resources to spare. Evidence shows that drivers under CM driving conditions tend to engage more often with NDRTs, interact more frequently with NDRTs, and have a higher average duration of glancing away from the forward roadway than manual driving (Solís-Marcos et al., 2018). Furthermore, when drivers engage with NDRTs under CM driving conditions, their hazard perception ability is impaired (Zangi et al., 2022). On the other hand, recent studies show that driving under CM and SB conditions enhances the adverse effects of fatigue and sleepiness on driver performance (Arefnezhad, Eichberger and Koglbauer, 2022) and facilitate sleepiness progression during night-time driving (Ahlström et al., 2021).

The engagement with an NDRT dilemma during CM and SB driving should also consider drivers' adaptation and regulation strategies while driving. Regarding regulation strategies, Schömig and Metz (2013) proposed a three hierarchical levels model to describe how drivers regulate their engagement with an NDRT while driving. The levels are planning, decision and control. On the planning level, drivers plan a priori where and when it would be most appropriate to engage with an NDRT along their route. On the decision level, drivers decide in real-time whether to begin engagement with an NDRT or postpone it based on the current situation. Finally, on the control level, drivers who have already initiated engagement with an NDRT will shift their attention between the driving and the NDRTs by applying various divided attention strategies. Naujoks et al. (2016) conducted a field study and found that under CM driving conditions, drivers adjusted their level of engagement with NDRT to the traffic situation (decision-making level). As their vehicle's velocity increased, drivers focused more on the forward roadway and less on the NDRT. This decrement in NDRT engagement due to traffic conditions can be considered **situation-adaptive behaviour** since it may reduce the perceived safety or increase the perceived workload during partially automated driving. Recently, Zangi et al. (2022) showed that when drivers are asked to engage with a time-based NDRT (control level) under simulated CM conditions, engagement with an NDRT impaired driving performance in two primary aspects. First, drivers were less aware of road hazards, and second, their mental workload was higher when they engaged with an NDRT. In addition, the findings reveal that for drivers engaged with an NDRT, the attentional time-sharing strategy between the NDRT and the roadway monitoring task affected the probability of identifying a hazard.

Considering behavioural adaptation in road safety and automated vehicles, Robertson et al. (2017) argued that it refers to how drivers modify their driving habits in response to new traffic and vehicle safety measures influencing crash risk. E.g., drivers may afford to drive faster because they believe the automatic braking system can help them stop more quickly. Due to the rapid evolution of in-vehicle technology, a lack of knowledge exists regarding how drivers adapt their behaviour to cope with this in-vehicle automation. However, some evidence shows, for example, that drivers with previous experience with Adaptive Cruise Control tended to increase their interaction with NDRT under CM driving conditions compared to drivers who were unfamiliar with it (Naujoks et al., 2016). In another on-road study, Kraft et al. (2018) asked 32 participants to drive in a driving simulator three times in one of six possible conditions (no HMI, full HMI, and reduced HMI, each under L0 and L2 conditions). The study found that the number of distractions by the HMI (calculated as the percentage of the time the subject looked at the HMI) decreased with each repetition in partially automated and manual driving.

Recently, a panel of traffic research specialists discussed various crucial topics that call for future research in automated vehicles (Shupsky et al., 2020). Two of these interrelated topics dealt with the following questions: (1) ***how should we balance situations of overload and passive fatigue in automated driving***, (2) ***what the most effective method for ensuring that drivers maintain adequate vigilance in the driving task is***? Research addressing these questions will provide a further understanding of how to resolve the abovementioned dilemma appropriately.

In light of this, we have designed two driving simulator experiments to answer the above questions while considering the complexities related to the interaction between drivers and partially and conditionally automated vehicles. Each experiment refers to a different driver state: fatigue and distraction. Both experiments rely on BGU's previous experiments conducted in MEDIATOR WP1 (Borowsky et al., 2020), where we laid the foundation (e.g., scenarios, measurements) and knowledge needed for the Mediator HMI evaluation. Thus, the two experiments we will present in the following sections aimed to assess the Mediator HMI concept developed in the project in mitigating the effects of fatigue and distraction in the short and long term (behavioural adaptation).

2.2. Research questions and motivation for both studies

In our previous work in WP1, we conducted two studies, one on fatigue and one on distraction (Borowsky et al., 2020). The fatigue study aimed to evaluate how passive fatigue progresses during partially automated driving when drivers are either engaged or not with an NDRT (Simon Game) and how drivers' hazard perception abilities are affected when such an NDRT is involved. Notably, the NDRTs never appeared concurrently with a hazardous situation but always between scenarios. In this study, we showed that fatigue progression was slower under CM when drivers engaged with the NDRT than when they were not. In addition, a marginally significant trend showed that participants who engaged with an NDRT had a higher probability of identifying a hazard than participants who did not engage with an NDRT during CM driving conditions. Although these findings were encouraging regarding NDRTs potential to mitigate the effects of fatigue, we did not investigate whether such benefits persist over time and experience with the system.

Similarly, the distraction study evaluated how engagement with a visual-manual NDRT affects drivers' hazard perception performance under CM driving conditions. The findings showed that, first, drivers were less aware of road hazards, and second, their mental workload was higher when they engaged with an NDRT. In this study, we triggered the NDRTs in proximity to the hazardous situation to examine how drivers regulate their attention between the hazardous situation and the task.

The current study aimed to elaborate upon the previous research and has the following two objectives:(1) To evaluate the effects of preventive and corrective mediation (as implemented in the Mediator HMI concept) in mitigating fatigue and distraction effects on driving performance under CM driving conditions. (2) To evaluate the longer-term effects of the Mediator HMI concept, including preventive and corrective mediation. The concept of preventive and corrective mediation is further explained in the HMI section. In short, preventive mediation refers to actions/alerts taken by the HMI to prevent drivers from being distracted or fatigued before they occur. Corrective mediation relates to actions taken by the HMI to correct a driver's aberrant behaviour, given that distraction or fatigue had already happened. These two objectives are translated into the following research questions.

- **RQ1.** Will the Mediator HMI preventive and corrective mediation concept assist in mitigating the adverse effects of distraction and fatigue? To answer this question, we followed the Mediator HMI general guidelines and developed HMI features of preventive and corrective mediation. In brief, preventing mediation features included an auditory and visual Trivia game offered to the participant at certain points along the drive and a hazard notification system, notifying the driver of an upcoming hazard. Corrective mediation used the same HMI features, but their trigger was only when the driver was in a state of distraction or fatigue. To answer this question, we examined whether including such features improves drivers' hazard perception abilities.
- **RQ2.** Will the benefits of preventive and corrective mediation persist in the longer term? To address this question, we invited participants to the lab for two consecutive sessions one week apart. Participants experienced the same HMI and its associated features for the second time. We examined whether participants' hazard performance changed from one session to the next.

- **RQ3.** Will participants accept, trust, and perceive the Mediator system as useful in the short and longer term? To answer this question, we administered various questionnaires several times during the experiments and analysed their results.

In the following sections, we describe the methods and results of both studies. Both studies focused on use cases 4, 7, and 8 (see introduction section 1).

2.3. Materials and methods (both experiments)

2.3.1. HMI Design concept

The Mediator HMI design had two primary functions. The first function aimed to convey the current driving mode to the driver (CM, SB, or manual) to increase mode awareness and prevent mode confusion. The second function sought to prevent drivers from becoming fatigued or distracted (preventive mediation) or, in case they are already in such a state, to correct their degraded performance (corrective mediation). We will elaborate on each function in the following paragraphs.

The HMI used visual and auditory stimuli as the primary communication channels. The locus of interaction was the driver dashboard (Figure 2.2 left rectangle) and the infotainment screen (Figure 2.2 right rectangle). The overall design and specific elements chosen for each interaction followed the general principles developed in the MEDIATOR project.

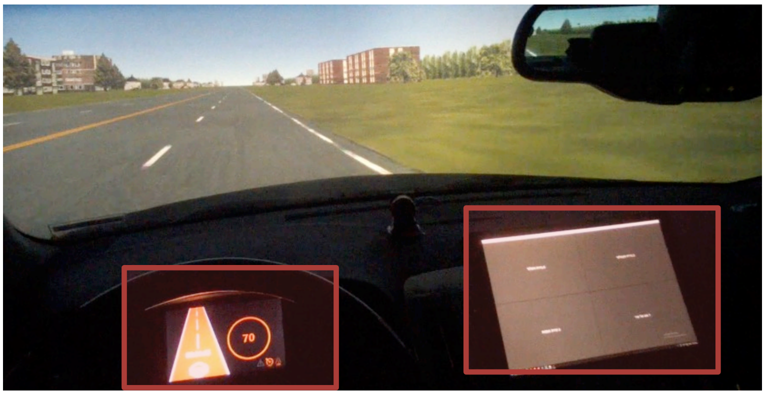


Figure 2.2 Dashboard and infotainment displays (left and right rectangles, respectively).

2.3.1.1. Automated driving mode indication

Drivers' ability to follow the current automated driving mode and be aware of it depends on how the HMI design minimizes mode confusion and errors. Mode confusion is a kind of automation surprise where the system fails to behave according to user expectations. Consequently, users lose track of the currently active system (Kurpiers et al., 2020). Multiple modes could contribute to mode confusion or errors (Sarter & Woods, 1995).

The function of conveying the current driving mode to the driver to minimize mode confusion consists of several parameters. Regarding colour, following the Mediator's guidelines, grey indicated manual driving mode, and amber indicated CM driving mode. Only one colour indicated each driving mode. Next, in the current experiments, drivers mainly drove under CM driving conditions. Still, there was the option to shift back to manual driving and forth to partial automation if needed. Thus, to ensure the driver is aware of a mode change during transitions between automation and manual modes, we used an auditory stimulus (a pre-determined tone) and changed the screen's colour and the relevant icons' colours.

In addition, in cases where the driver was driving in manual mode, we initiated a voice message telling him to activate automation as soon as he felt safe. Figure 2.3 shows the screens presented on the driver's dashboard in each driving mode.

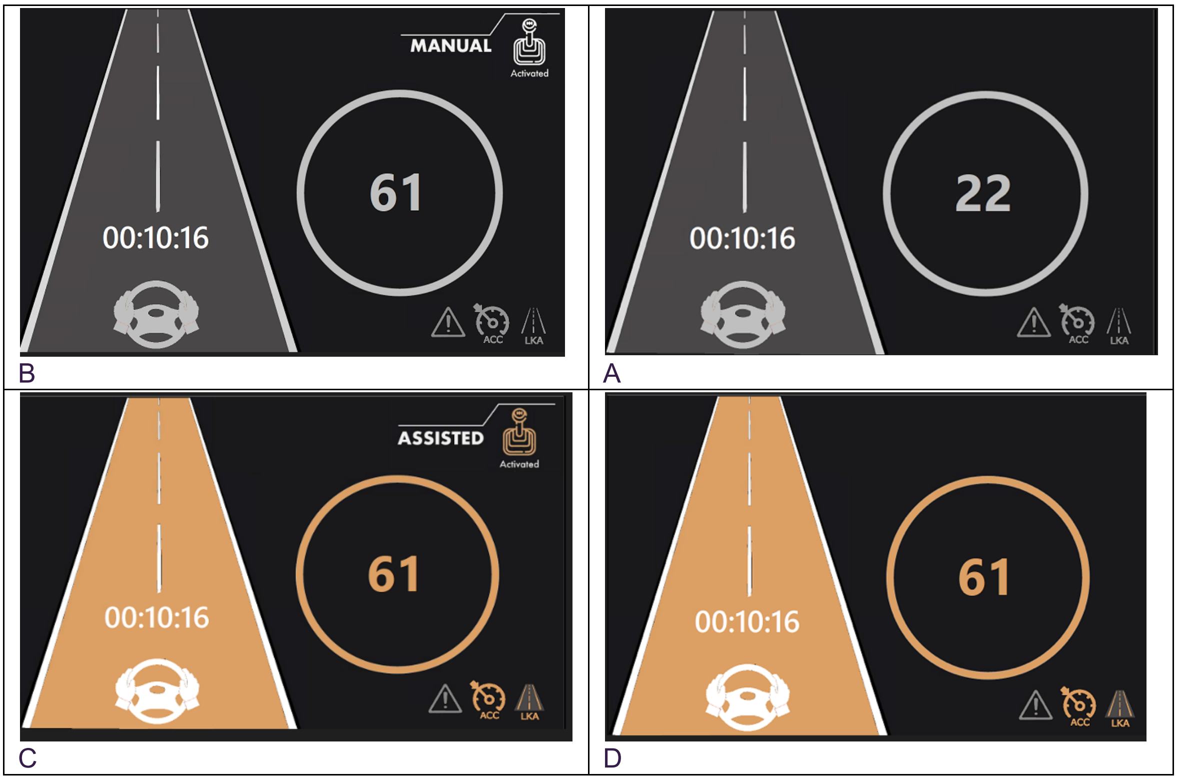


Figure 2.3 Colours and icons reflect the automation level and current automation driving mode.

The amber indicates that the current driving mode is CM, and the grey indicates that the current driving mode is manual. The icons ACC (Adaptive Cruise Control) and LKA (Lane Keeping Assistance) in the bottom right indicate whether the lateral control (LKA) and longitudinal control (ACC) are currently activated. The gear shift icon + the text "Assisted" for the activation of manual (a transition from CM to manual, see Figure 2.3B) and vice versa (see Figure 2.3C).

2.3.1.2. Preventive and corrective mediation

The HMI's second function sought to prevent drivers from becoming passively fatigued or distracted (preventive mediation) or, in case they are already in such a state, to correct their degraded performance (corrective mediation). To achieve this goal, we developed HMI preventive and corrective mediation concepts. Figure 2.4 describes the general preventive and corrective mediation concept and the driver states to which this concept was targeted.

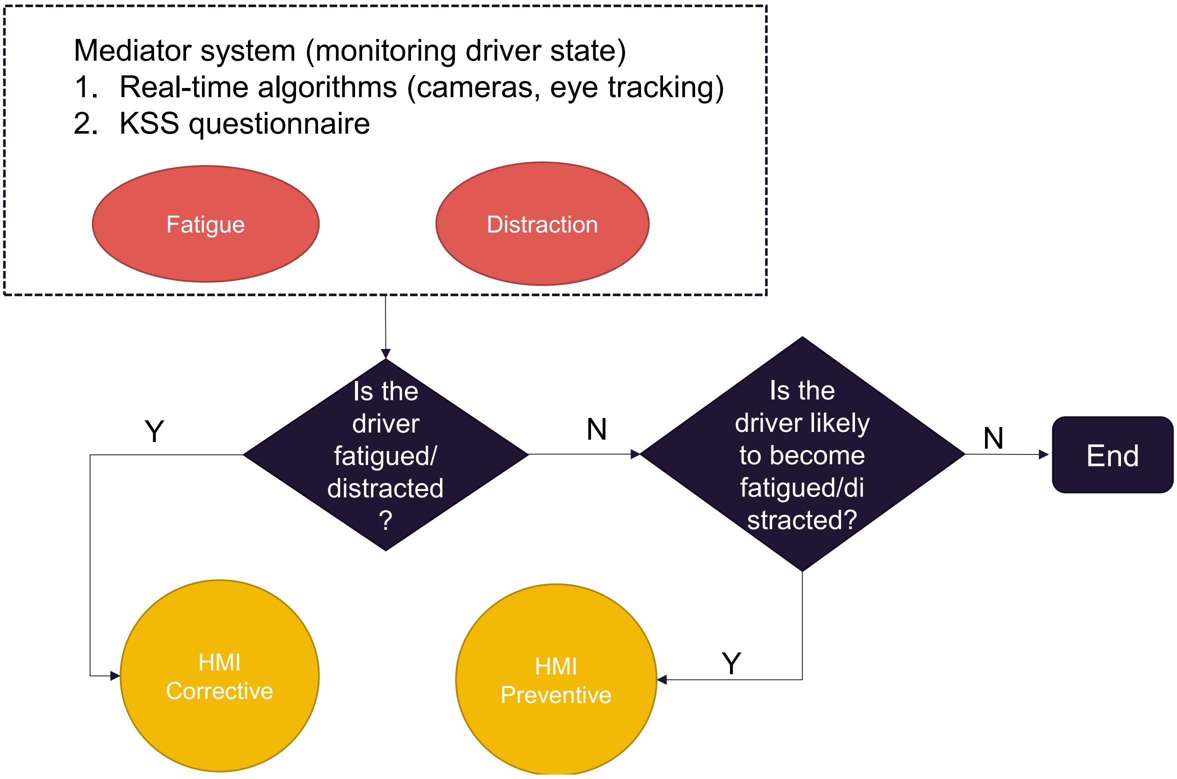


Figure 2.4 Mediator corrective and preventive mediation concept.

Figure 2.4 shows that preventive and corrective mediation, as implemented in the HMI, was intended to mitigate the effects of fatigue and distraction. Nevertheless, due to the immaturity of the online algorithms for detecting fatigue and distraction, the corrective mediation option was not implemented in the distraction study. Although we did not use any fatigue detection real-time algorithm, the Karolinska Sleepiness Scale (KSS) ratings (given in real-time along the drive) were used to detect fatigue in real-time.

We built two HMI components within the preventive and corrective mediation function: (1) notifications of upcoming hazardous situations and (2) a Trivia game. The first component is preventive mediation aimed at notifying drivers of upcoming hazards and raising drivers' awareness of these hazards. This component can help drivers draw attention to on-road hazards in case of fatigue or distraction. This component was implemented in the fatigue and distraction study. The second component was implemented only in the fatigue study. We used this component as a preventive mediation (independent of KSS scores) and corrective mediation (dependent on KSS scores along the drive).

1) Notifications of upcoming hazards

The first sub-function of the Mediator HMI was to notify drivers of upcoming hazards and raise drivers' awareness of these hazards. We assumed that the hazardous situations we used in our studies could be detected in real-world situations using a fusion of various existing sensors. As our driving simulator studies focus on CM, this assistive automation was done at the perception level and not at the execution or response selection levels. Thus, we implemented pre-hazard alerts such that a 'Hazard ahead' notification appears when entering a launch zone, which is a time point along a scenario when a driver should begin glancing at the hazard (see the detailed description of the eight hazardous scenarios in Table 2.1). The notification appears visually on both screens (driver and infotainment), accompanied by an audio signal. The icon is a general alert symbol independent of the specific hazardous situation it precedes (Figure 2.5). The reason for using visual and auditory signals is that multimodal HMIs typically evokes reaction times faster than unimodal HMIs (Petermeijer et al., 2017). In addition, an interface providing multimodal warnings is perceived as delivering more urgent cues than unimodality (Van Erp et al., 2015). Finally, as drivers under CM driving conditions may engage with NDRTs, it was decided to add an auditory signal to the visual 'hazard ahead' notification.

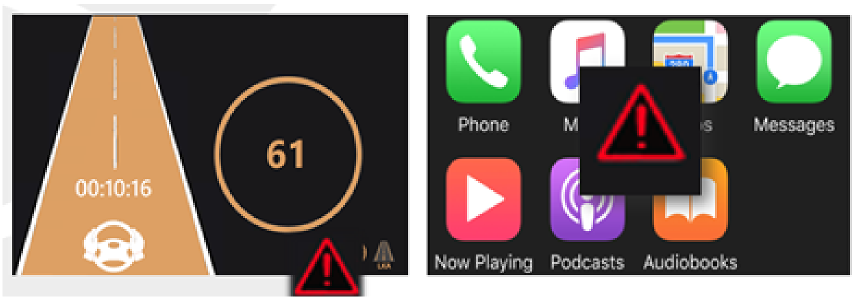


Figure 2.5 The hazard alert - visual and vocal alerts of hazards ahead.

2) Trivia

Disengagement from driving-related activities for prolonged durations can cause passive fatigue, particularly when drivers are focused on the road, monitor the automation, and are not engaged with NDRTs to maintain a suitable arousal level (Naujoks, et al., 2016). Some evidence shows that engaging with an NDRT, such as Trivia, can mitigate these harmful effects (e.g., Oron-Gilad et al., 2008). On the other hand, when a driver engages with NDRTs under partially automated driving, it is necessary to ensure that drivers continuously allocate visual attention to the roadway and do not glance inside the vehicle for more than two seconds (and for a total time of 12 seconds per task; NHTSA, 2016). In addition, it is essential to ensure that the mental demands required by the NDRT are not impairing drivers' ability to take control safely whenever necessary.

In the current fatigue study, we used audio-visual Trivia as an NDRT that may suspend the progression of passive fatigue (see Figure 2.6). Previous studies showed that under prolonged manual driving sessions, auditory Trivia helped to prevent driving performance deterioration, increased alertness (measured by standardized HRV), and was preferable by drivers over the other two common Alertness Maintaining Tasks (Oron-Gilad et al., 2008). In the current study, the invitation to play appeared on the infotainment display as a visual message. When the invitation was independent of the KSS scores, it was considered preventive, and when it followed KSS scores higher than a certain threshold, it was considered corrective.

Each time a participant accepts an invitation to play Trivia, this HMI component triggers 11 Trivia questions on various topics heard via in-car speakers, including four answer choices. Each query (and the possible answers) also appears on the infotainment display. The driver should select the correct answer by pressing it on the infotainment display.

Figure 2.6 The invitation to play Trivia (on the right) and an example of a question.

We used the following alert escalation strategy when suggesting a driver play Trivia due to corrective mediation. We allowed three levels of escalation (Figure 2.7).

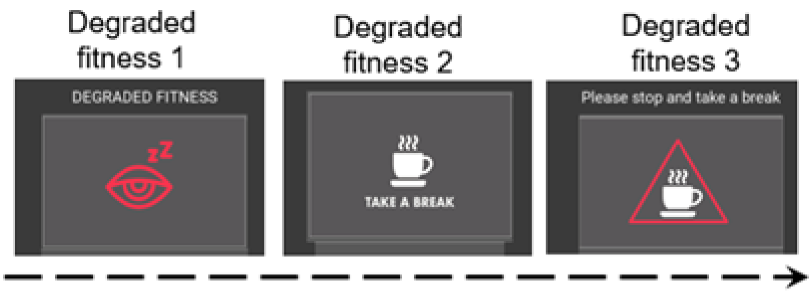


Figure 2.7 The Mediator escalation policy.

Once a driver provides a KSS score higher than 6 (this number follows the Mediator general principles) for the first time (degraded fitness 1), an eye icon appears on both the dashboard and infotainment displays, notifying the driver his performance degrades due to fatigue. If the driver is not playing Trivia at that moment, the system suggests playing Trivia to counteract the effects of fatigue. The driver could either accept or reject the offer. Then, in the following KSS instance, given the driver provides KSS scores higher than 6 for the second time (degraded fitness 2), the system suggests taking a break (coffee) and, in the meanwhile, until the driver decides to stop for a coffee, the system also suggests playing Trivia to counteract fatigue. The driver could accept or reject the offer to play Trivia. The difference between degraded fitness levels 1 and 2 is in the icon presented on the dashboard and a recommendation to take a break. Lastly, in the next KSS instance, if the driver continues providing a KSS score higher than 6 for the third consecutive time (degraded fitness 3), the previous notification escalates and asks the driver more assertively to stop and take a break (before the emergency takeover is activated). Each visual announcement was accompanied by a sound where alert sound type 1 was embedded in the first escalation level, and alert sound type 3 was embedded in the second and third escalation levels to indicate escalation.

2.3.2. Driving simulator, in-vehicle displays, and vehicle automation

Both studies used an RTI (Realtime Technologies, Inc.) non-moving base driving simulator. The driving simulator consists of an engineless Cadillac-STS sedan and a 7 m diameter curved screen (2.4 m × 6.1 m), creating a visual angle of 165 degrees of the virtual world. The curved screen is located about 1 m in front of the Cadillac). Three laser projectors displayed the virtual world on the curved screen, and a designated software (Wrapalizer, Inc.) did the edge blending. A rear projector and a screen at the back of the simulator presented the virtual environment through the in-vehicle rear-view mirror. In addition, each physical side mirror included a 7″ LCD showing the respective views of the virtual environment.

The core modelling of the driving simulator is SimCreator, and the virtual world and scenarios were designed in SimVista. In addition, we used the Altia Designer software (Altia Designer, Inc.) to design the in-vehicle display visuals. The driving simulator provides various vehicle dynamics measures such as driving speed, steering angles, acceleration, and braking at a rate of 30Hz.

Concerning audio, the driving simulator has two stereophonic audio systems. One is located below the curved screen outside the vehicle and provides engine sounds and external noises from the road environment. Another audio system is located inside the car in the back seat. This system allows communication with the driver via audio.

The HMI in the driving simulator includes two interfaces. The first interface is located at the instrument panel. It consists of a 7" LCD visual display presenting various icons and the speedometer (Figure 2.8 left). The second interface is located at the central stack between the driver and the passenger. It includes a 10" touchscreen connected to a PC. This interface displays information to the driver, presents NDRTs, and receives inputs from the driver (Figure 2.8 right).

Figure 2.8 Real-time technologies high fidelity driving simulator and an in-vehicle touchscreen display.

The simulated vehicle was designed in two modes: manual driving mode and CM driving mode. The participant controlled all aspects of the driving task in the manual driving mode. In the CM driving mode, the PAV controlled the car’s speed, acceleration, time headways from lead vehicles, and lane position (i.e., longitudinal and lateral control). Activating the PAV was done by the driver using a designated button located inside the car. The participant could disengage from the CM and switch back to manual driving by pressing the brake pedal.

2.3.3. Eye tracker and electrocardiogram device

Eye tracker: We recorded participants’ eye movements throughout the experiment via Tobii Pro Glasses 2 head-mounted Eye Tracking System. The glasses weigh 45 grams and include four eye cameras, a full HD wide-angle scene camera, a gyroscope, and accelerometer sensors. The eye tracker samples gaze position at a rate of 50 Hz with an accuracy of 0.5 degrees of visual angle.

The Tobii Pro Lab software (ver. 1.142) was utilized to analyse glance patterns. The main output of the data includes videos with gaze positions superimposed.

Electrocardiogram (ECG): A BioPac ECG system (MP150) was utilized to measure participants' heart rates at 2000 Hz. The system consists of a matched transmitter and receiver module, which emulates a "wired" connection from the subject to the computer. Connecting the electrodes to the participant was done by pasting three stickers on the participant's chest. These electrodes interface with two parts: the data acquisition and analysis platform and the AcqKnowledge software. The ECG and the driving simulator were synchronized.

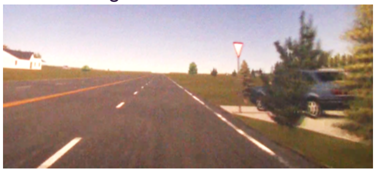
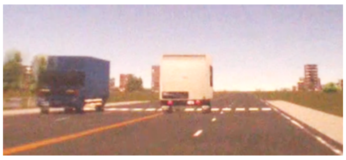
2.3.4. In-vehicle cameras

Four 3.2 MP FLIR cameras were installed in the simulator's vehicle. Three cameras are in colour; one is monochrome and has a near-infrared sensor. All cameras work at a high speed of 117 fps. One colour camera captured the driver's face and was in front of the driver above the instrument panel. The second colour camera captured the simulated environment from the car's roof. The third was positioned over the participant's shoulder to grasp his hands and activities with the infotainment display. Finally, the fourth monochrome camera captured the participant's body posture from the right. These cameras are not part of the original experiment plan, and the data will be post-processed later in collaboration with MEDIATOR partners.

2.3.5. Hazardous scenarios

The virtual drive in the fatigue and distraction studies included highways and urban environments in daylight. The roads in the urban environment had 2-lanes bi-directional roads, except for one scenario with one lane in each travel direction. During the drive, participants navigated eight latent (unmaterialized) hazard scenarios. These hazards rely on previous studies (Zangi et al., 2022) with some modifications. The scenarios were distributed differently throughout the virtual drive in the fatigue and distraction studies; thus, details about their distribution in each study will be described in the relevant sections. Table 2.1 presents a detailed description of the eight scenarios.

Table 2.1 A detailed description of the eight hazardous scenarios used in the fatigue and distraction study.

Drive set	ID	Scenario name and picture	Description
Set A	1	Exit from a gas station 	The participant's PAV drives at 60 km/h on the right lane of a 4-lane highway (two lanes in each travel direction) in a suburban environment. The participant's PAV passes by a gas station with an entrance and exit from and to the main road from the right. One hundred and fifty meters before passing by the gas station's entrance, a car in front of the participant's PAV turns into the gas station, and the participant's PAV keeps driving straight. The gas station's exit is located approximately 200 meters after the entrance, partially obscured by bushes. Approximately 50 meters before the exit, there is a relatively large yellow sign saying "Hidden passageway." When the participant's PAV passes by the exit, another vehicle is at the exit, waiting to merge into the road. **Hazard:** The vehicle on the right is waiting to merge into traffic. **AOI (Area Of Interest):** The vehicle at the gas station's exit. **Cue:** The vehicle ahead enters the gas station, indicating the possibility of other vehicles exiting the gas station further down the road.
	2	SUVs obscure the view of a midblock cross-walk 	The participant's PAV drives 60 km/h on the right lane of a 4-lane road in a suburban environment. Two trucks, one on the left lane of the driving direction as the PAV and the other on the rightest lane of the opposite direction, obscure the left side of a crosswalk. **Hazard:** Trucks obscure possible pedestrians crossing the road from the left side. **AOI:** The right edge in front of both trucks obscures the crosswalk and the curb on the opposite side. **Cue:** Crosswalk, stopping SUVs.
	3	Pulling out from a line of parked vehicles  	The participant's PAV drives at 60 km/h on the right lane of a two-lane road in a suburban environment as it approaches a three-way intersection with a left turn. Five cars are parked in a line on the right curb, and one of them is signalling left, indicating its intention to merge into the driver's lane. Simultaneously, A car approaches the intersection from the left and stops at a stop sign waiting to enter traffic, serving as a distractor as the driver approaches the line of parked vehicles. **Hazard:** The vehicle intending to pull out and the car on the left waiting to merge into the main road. **AOI:** The parked vehicle with the signal light, the car approaching from the left. **Cue:** Line of parked vehicles, the car approaching from the left.

Drive set	ID	Scenario name and picture	Description
	4	'Goral' intersection - SUVs obscure the view of cars merging from the left 	The participant's PAV drives at 70 km/h on the right lane of a 4-lane highway in a suburban environment and approaches a four-way intersection. The right turn is blocked. Two SUVs were in the left lane of the driving direction as the PAV stopped by the intersection. When the participant's PAV is approximately 100 meters before the intersection, a car turns from the left side to the main road. Another vehicle from the left is waiting to turn to the main road, and the SUVs obscure it on the left lane. **Hazard:** The obscured car on the left of the intersection can turn left and merge into the participant's road. **AOI:** Right edge in front of the SUV closest to the intersection, car waiting on the left. **Cue:** Car turning from the left side of the intersection into the main road.
Set B	5	Obscured car stuck on the right 	The participant's PAV drives 60 km/h on the right lane of a 2-lane highway in a suburban environment. A large truck that is driving in front obscures a car that is stuck on the right-side curb. The truck suddenly slows down and makes a sharp manoeuvre to overtake the obscured stuck car. **Hazard:** A car stuck on the right-side curb, partially blocking the road's right-hand shoulder. **AOI:** Car stuck on the right-side curb. **Cue:** Truck driving in front.
	6	Partially obscured view in a curve 	The participant's PAV drives 60 km/h on the right lane of a 2-lane highway in a forested suburban environment and approaches a long, sharp right curve partially obscured by trees. Approximately 150 meters before entering the curve, another car can be spotted behind the curve's apex at the other end, indicating the possible existence of other vehicles currently obscured along the curve. **Hazard:** Possible obscured vehicles or other dangers along the curve. **AIO:** Right edge of the road and glances behind the trees at the side. **Cue:** Car at the other end across the curve.
	7	A car approaches an intersection from the left-hand side. 	The participant's PAV drives 70 km/h on the right lane of a 4-lane highway in a suburban environment and approaches a four-way stop-controlled intersection (the participant has the right of way). Approximately 400 meters before the participant's PAV arrives at the intersection; another car can be spotted approaching on the left lane from the left-hand side road of the intersection. Approximately 300 meters before the participant's PAV arrives at the intersection, a truck stopped on the right lane on the left-hand side of the intersection obscures the car when it reaches the stop line in front of the intersection.

Drive set	ID	Scenario name and picture	Description
			Hazard: The obscured car on the left of the intersection can turn left and merge into the participant's road. **AOI:** Area in front of the truck. **Cue:** The car approaches from the left-hand side of the intersection.
	8	Bus on the opposing lane obscures a midblock crosswalk 	The participant's PAV drives 100 km/h on the right lane of a 4-lane highway in a suburban environment. A bus stops at a bus station on the right lane in the opposite direction. The rear part of the bus obscures the left side of a crosswalk behind the bus. **Hazard:** The bus obscures the view of possible pedestrians that may cross the road on the crosswalk behind the bus. **AOI:** Area behind the bus at the crosswalk. **Cue:** Bus station, Stopping bus.

The "cue" field describes the hazard-related visual cues that preceded each hazardous situation and cued the participant to the upcoming hazard. For each of the eight latent hazards, we defined a time window consisting of the period during which the participant should identify the visual cue(s) indicating the developing hazard and the period during which the driver should glance at the hazard instigator to avoid a potential crash. The start of the second period is called a launch zone, a pre-defined area of each driving environment where the hazard becomes visible, and the driver must begin glancing toward the target zone. A Target zone is a visual area of interest (AOI) of a latent hazard (Krishnan et al., 2019). The time window ended when the detection of the latent hazard would have been too late to avert a crash (Vlakveld et al., 2018). However, none of the eight latent hazards ever materialized. The eight-time windows ranged from 18 to 38 s. The second period (launch zone) lasted between 1 and 3 s.

In the fatigue and distraction studies, each experiment included two driving sessions with four hazards each. Thus, we divided the eight hazards into two sets of four hazards (A and B). We verified that each set contains similar hazards with the same level of difficulty and environmental characteristics. Within each set, we generated four hazards' combinations such that each hazard appeared once in each possible location (1, 2, 3, and 4). In each study (fatigue and distraction), half of the participants did the first driving session with set A and the second with set B and vice versa. Each participant received a different combination of hazards within each set.

2.3.6. Questionnaires

The current studies applied a series of questionnaires dealing with trust, usability, familiarity, safety, experience with automation, fatigue, and mental workload. Most questionnaires are based on standardized/established methods and were also used in other MEDIATOR study teams. When needed, questionnaires were translated into Hebrew and validated using forwards and back translations and a small-scale pretest ($N >= 5$). Table 2.2 summarizes the questionnaires implemented in the simulator studies, including the time of the measurements, resources, and related hypotheses and results. The three instances of filling out the questionnaires were: '*Pre-questionnaire'* refers to the first time right after signing the consent form and before the training drive. '*Before*' was administrated right after the training drive, and '*post'* was administrated at the end of the driving session. The times for each questionnaire were explicitly defined, as shown in Table 2.2.

Table 2.2 Summary of the questionnaires implemented in the BGU driving simulator study

	Time of administration	Questionnaires from literature / previous projects	Topics
Demographics	Pre-questionnaire	• Education level, employment status (L3 Pilot project; Metz et al., 2019) see Appendix 5.1.1	Items related to age, gender, and employment status.
Mobility / ADAS	Pre-questionnaire/	• ADAS availability and usage (ADAS&Me: Pereira Cocron et al., 2019; L3 Pilot: Metz et al., 2019) see Appendix 5.1.2	Items related to annual mileage driven, license possession, use of automation.
Adoption &Trust Attitude toward vehicle automation	Three times: Pre-questionnaire+ 'post' drive	• Choi and Ji (2015) • SUaaVE pre-trial questionnaire (Post et al., 2020) See Appendix **Fout! Verwijzingsbron niet gevonden.**	Automation adaptation and trust, acceptance, and safety.
Acknowledgment of CM	Two times: Before each driving session	• Familiarity with CM definitions (see Appendix 5.1.4)	
Acknowledgment of Mediator system indication functionality	Two times: 'post' drive	• Understanding of the Mediator HMI concept, see Appendix 5.1.5	
Mental workload-NASA-TLX	Two times: 'post' drive	• (Hart & Staveland, 1988; see Appendices 5.1.6), administrated after each experimental	
Fatigue	Throughout the drive, every ~10 min, four times in total (see the design of the drive Figure 2.10)	• Karolinska Sleepiness Scale (KSS) (Shahid et al., 2011; Åkerstedt, 1990)	
Usability & Acceptance of Mediator	Two times: 'post' drive	• System Usability Scale (SUS; Brooke, 1996) • Acceptance scale (Van der Laan et al., 1997) See Appendix 5.1.7	Usability-related items and self-designed.

Driving demographics and Familiarity with automation questionnaires were based on Mediator Guidelines for design and methodological approach of the evaluation studies. Some of the text translation to Hebrew was inspired by the Israeli Ministry of Transportation Procedure No. 0313 - Safety systems in vehicles of types M1 and N1, 2020 (shown in Appendix 5.1.1).

The fatigue questionnaire was based on the KSS (Åkerstedt, 1990; Shahid et al., 2011). It was implemented in the fatigue study using a digital survey presented on the infotainment display, as illustrated in Figure 2.9. Participants evaluated their current level of fatigue every ten minutes throughout both experimental driving sessions.

Figure 2.9 Digital KSS survey, presented on the infotainment display.

The SUS questionnaire data were processed prior to the analysis according to their contribution. Every item contribution score ranged between 1 (strongly disagree) to 5 (strongly agree). For questions 1, 3, 5, 7, and 9, their contribution score is the position of scale subtracted by 1. For questions 2, 4, 6, 8, and 10, their contribution score was five subtracted by the position of the scale. The total contribution score was multiplied by 2.5 to acquire the overall value of system usability. The SUS scores ranged between 0 and 100.

2.3.7. Dependent measures

Table 2.3 summarizes the dependent variables and their interpretation.

Table 2.3 Summary of the dependent variables

Measurement	Definition	Source
Cue identification before the alert onset	Hazard-related visual cues that preceded each hazardous situation. A cue could be dynamic or static (see Table 2.1 for the visual cues preceding the hazards). For each of the eight latent hazards, we defined a time window consisting of the period during which the participant should identify the cue. Identification of a cue is binary, indicating whether the driver had at least one glance longer than 100 msec at the cue ('1') or not ('0') before the hazard notification onset.	Eye tracker
Number of glances at a hazard	This variable counts the number of glances longer than 100 msec at a hazard during the launch zone.	Eye tracker
Hazard identification	Whether a participant had at least one glance of at least 100 msec at the hazards within the launch zone (after the alert appeared)	Eye tracker
Automation deactivation	Did the participant deactivate the automation mode? (yes/no)	Driving simulator
KSS score difference (Karolinska sleepiness scale)	The difference between the values of two consecutive instances of KSS. Given there were four instances in each drive (*n* = 4). (KSS score at n+1)- (KSS score at n)	Digital KSS questionnaire
HRV-RMSSD (root mean square of successive differences	$\text{logQuotientRMSS} = \frac{Recorded\ RMSSD}{Recorded\ rest\ RMSSD\ baseline}$	ECG

Measurement	Definition	Source
between normal heartbeats), Bartels and Peçanha (2020)		
HRV-SDNN The standard deviation of the interbeat intervals of normal sinus beats (SDNN), Bartels and Peçanha (2020)	$\text{logQuotientSDNN} = \frac{RecordedSDNN}{Recorded\ rest\ SDNN\ baseline}$	ECG
Mental workload scores	The average scores of all six questions in the NASA-TLX mental workload questionnaire	NASA-TLX questionnaire

2.4. Method (fatigue)

The fatigue study aimed to elaborate upon the previous research by addressing the following questions:

1. Will the positive effects of engagement with NDRTs persist over time under CM driving conditions of underload?
2. Will the Mediator HMI concept of preventive and corrective mediation help mitigate fatigue's adverse effects?

To address these research questions, we asked participants to arrive at the driving simulator twice one week apart. Participants drove for about 40 minutes in each session under CM driving conditions and experienced four latent hazardous situations. Underload was induced by embedding 8 minutes road sections with sparse traffic before each hazardous event. These road segments did not include any situation requiring the driver’s intervention. During the driving session, half of the participants were offered to play Trivia. The following sections describe the method relevant only to the fatigue study.

2.4.1. Participants

Twenty-four participants were recruited for the study as paid volunteers. All participants were undergraduate students from Ben-Gurion University of the Negev in Beer-Sheva city in Israel. Half of the participants were males (mean age = 27.3 years, *SD* = 7.03; mean driving experience = 9.63 years, *SD* = 7.26), and half were females (mean age = 25.25 years, *SD* = 2.83; mean driving experience = 7.25, *SD* = 2.7). Participants received monetary compensation for their participation (~$40). All participants had normal or corrected-to-normal visual acuity (contact lenses only), normal contrast sensitivity, normal colour vision, and no background of heart problems. Vision tests included a Snellen chart for visual acuity and Ginsburg's functional acuity contrast test (FACT) chart for contrast sensitivity evaluation (Ginsburg, 2002). Participants had a valid driver's license for at least five years without any previous experience with adaptive cruise control and lane keeping assistance. This study received the BGU ethical committee's approval.

2.4.2. Tools

The tools used in the fatigue study are described in detail in section 2.3.

2.4.3. Experimental design

The experiment was a 2-by-2 mixed factorial design. Participants were randomly assigned to one out of two experimental conditions (half males and half females in each experimental condition): (1) CM driving with Trivia and notifications of upcoming hazards or (2) CM driving with notifications of upcoming hazards only. This between-subjects independent variable aimed to examine the benefits of using Trivia to mitigate fatigue development. Participants of each experimental condition were invited to drive in the driving simulator twice, one week apart. Each driving session lasted about 40 minutes. This independent variable was within subjects and examined the long-term effects of the Mediator HMI.

Each driving session included a pre-defined set of four un-materialized hazardous scenarios (eight in total), four KSS instances, three Trivia (independent of the KSS score; preventive mediation), and a maximum of additional two Trivia instances depending on the KSS score (above 6; corrective mediation). Figure 2.10 presents the structure of a driving session for the Trivia experimental group. This structure was identical for the second experimental group except for the Trivia, which was absent.

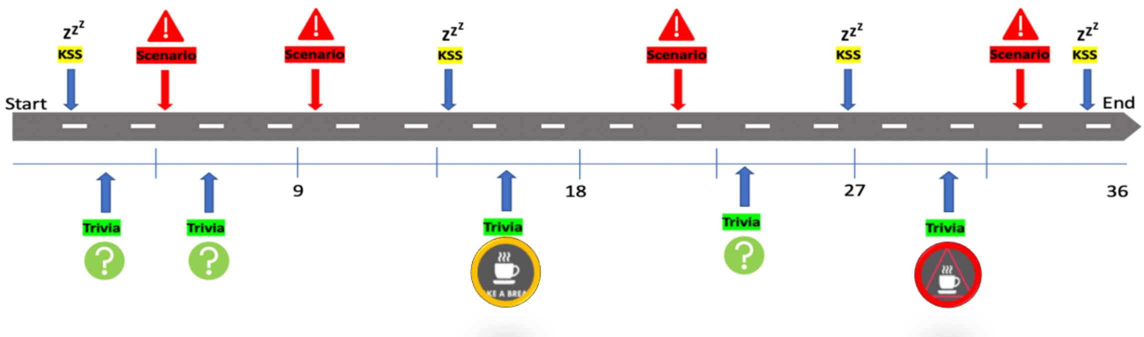


Figure 2.10 The structure of a 40-minute driving session.

The KSS instances in Figure 2.10 show their time-based appearance (minutes) during the driving session. The red hazard sign indicates the location of the upcoming hazard notification during the driving session. The green question marks represent the location of the Trivia offers along the driving session for the first experimental condition (preventive mediation).

Both experimental groups interacted with the HMI system and were asked to evaluate their subjective fatigue level. They received visual and vocal alerts of hazards ahead, automated driving mode status, and alerts in case of fatigue.

2.4.4. Procedure

Before the participants arrived in the lab for the first time, they filled out the following forms and questionnaires (see Table 2.2 for more information):

1. Registration form containing details and conditions for participating in the experiment
2. Consent form
3. Demographics questionnaire
4. Experience with partial automation driving (Mobility/ADAS questionnaire)
5. Adoption and trust in automated vehicles (Trust & SUaaVE questionnaires)

6. Familiarity with automation (acknowledge of CM questionnaire), filled out after the training just before the first driving session.
7. Acceptance and usability (Acceptance scale and SUS questionnaires). Participants filled out this questionnaire thrice, once upon their arrival at the lab for the first time and after each driving session.

When arriving at the lab for the first time, participants gave the experimenter the forms they filled out at home and conducted visual acuity and contrast sensitivity tests. Qualified participants received written instructions regarding the simulator automation capabilities, the simulated environment, the HMI system, and the remaining procedure of the session. In the written instructions, participants were informed that they carry full responsibility for the driving task and safety and should monitor the automation continuously despite the automation capabilities. The participants could assume manual control whenever necessary by pressing the brake pedal. Participants were instructed to drive as they usually would in similar real-world situations.
At the beginning of both sessions, participants were connected to the ECG using electrodes attached to their chest and asked to sit and relax for 5 minutes while reading a magazine. Meanwhile, a baseline ECG, including R-R interval measurement, was recorded. Next, the participants entered the vehicle, wore eye-tracking glasses, and underwent a short calibration process.

Then, after calibration, participants drove a 5-minutes training session to acclimate to the simulator, the measurement apparatus, and the HMI system. Then, the participants filled out a short knowledge-test questionnaire regarding the simulated CM vehicle's operation.
Depending on their experimental condition group, participants started their 40-minute driving session after the training session. At the end of their driving session, participants completed the post-drive questionnaire consisting of a NASA-TLX questionnaire, usability evaluation of the HMI questionnaire, acknowledgment of HMI indication functionality questionnaire, and adoption and trust in automation questionnaire (see Appendix 5.1).

A week later, each participant returned to the lab approximately at the same hour as the first session and underwent the same procedure as the first session. After completing two driving sessions and questionnaires, the participant was thanked for her participation and received monetary compensation. Figure 2.11 presents a schematic description of the procedure.

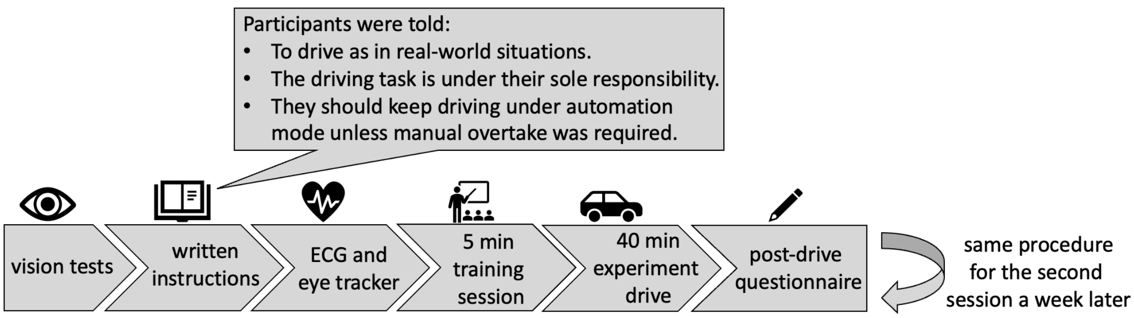


Figure 2.11 Overview of the experimental procedure.

2.5. Results (fatigue)

We used various measurements to evaluate the effectiveness of an auditory Trivia game in mitigating fatigue progression and its negative impact on driving performance under CM driving conditions. We evaluated Trivia's effectiveness as a fatigue counteracting measure in the short and long term. This section describes the analyses of these measures. The descriptions of the measurements and their computations appear in Table 2.3. Generally speaking, this section contains three categories of analysis: hazard perception performance, passive fatigue development and mental workload, and usability and acceptance of the Mediator HMI.

Recall that there were two experimental groups. One group experienced Trivia as preventive and corrective mediation, and one group did not use Trivia at all (Control). Corrective mediation is intended to be carried out if reported fatigue levels exceed 6. However, our findings show that the max average KSS score was 4.3 out of 9, indicating that participants did not feel fatigued throughout the driving session regardless of whether they engaged with the Trivia. Among the 24 drives of the Trivia group (12 participants with two driving sessions each), there were only five drives where the score of the 4th KSS instance was 7. The corrective mediation occurred after the fourth KSS instance for these five cases (four participants), and they all played the Trivia when offered. Since, however, the fourth KSS instance always appeared towards the driving session's end, participants accepting the invitation were unable to play much. Based on our findings, corrective mediation did not come into play in this study, and further analyses refer to preventive mediation only.

All statistical analyses were carried out at a significance level of 0.05. For dependent variables that are binary distributed, we used a logistic regression model with a logit link function within the Generalized Linear Mixed Models (GLMM) framework. These variables included cue identification, hazard identification, and deactivation of the automation mode. Cue and hazard identification are binary variables that describe whether a participant had at least one fixation of at least 100 msec at the hazard instigator or the cue ("1″) or not ("0″). Automation deactivation is a binary variable that describes whether a participant deactivated the automation mode during a scenario before or at the launch zone. For the Poisson distributed dependent variable, the total number of fixations at the hazard, we used a Poisson regression model with a log link function within the GLMM framework. Next, we used a linear regression model within the Linear Mixed Models (LMM) framework for dependent variables with a Normal or Log-normal distribution. These variables included the mental workload score, usefulness and satisfaction scores of the Mediator systems, perceived acceptance scores, convenience, safety, and trust in automation, KSS questionnaire score, logQuotientRMSSD and logQuotientSDNN measured preceding each KSS instance.
The initial model in each analysis included the following fixed effects: drive number (1/2), experimental group (Trivia/ Non-Trivia; i.e., Control), gender (M/F), the chronological order of the relevant events along the drive (KSS [1,2,3,4]; scenario[1,2,3,4]; Trivia[1,2,3]), and the first and second-order interactions between the variables. In addition, participants were included as the random effect in all models.

The final model of each analysis was achieved via a backward elimination procedure where all non-significant interaction effects were removed from the model. For significant fixed effects with more than two levels, post hoc pairwise contrasts comparisons analysis was applied. The Tukey HSD procedure was used to correct alpha for multiple comparisons. All statistical analyses were conducted in R software.

2.5.1. RQ1, RQ2: Hazard perception performance

Probability of identifying the cue before the alert onset. In this analysis, we modelled the probability of identifying the visual cue preceding the hazard **before** triggering the notification of an

upcoming hazard. This analysis examined whether the Trivia group was more likely to identify the visual cue before the alert than the None-Trivia group (a proxy of vigilance).
The final Logistic regression model included one marginally significant main effect of engagement with the Trivia task ($X_1^2 = 3.65$, $p = 0.056$). The estimated probability of identifying the cue was significantly higher for the Trivia group (*EM* probability = 0.86; *SE* = 0.04) than the Non-Trivia group (*EM* probability = 0.75; *SE* = 0.04). There was no effect of the scenario sequence, meaning that the probability does not change as driving progresses.

Hazard identification following the hazard notification onset. Following the start of the hazard notification in the HMI, we wanted to examine if there would be hazard identification differences between the Trivia and Non-Trivia groups. The final Logistic regression model yielded no significant effects. i.e., The Trivia and Non-Trivia groups had a similar estimated probability of identifying the hazard **after** the HMI alerted of an upcoming hazard (*EM* = 0.771, *SE* = 0.042, and *EM* = 0.724 *SE* = 0.0445 correspondingly). The main effect of the experimental group was non-significant ($X_1^2 = 0.638$, $p = .42$).

Number of glances at a hazard. In this analysis, we examined the number of glances drivers allocated to the hazard instigator, given that it was detected. The final Logistic regression model of the number of glances at the hazard yielded one marginally significant main effect of the experimental group ($X_1^2 = 3.135$, $p = .077$). Participants who engaged in the Trivia game had more glances at the hazard (*EM* = 3.26; *SE* = 0.37) than participants who were not allowed to play the Trivia (*EM* = 2.45; *SE* = 0.29).

Automation deactivation. This analysis examined possible differences between the two experimental groups in the probabilities of deactivating automation when approaching hazardous situations. The final Logistic regression included one significant main effect of the alert timing ($X_1^2 = 12.19$, $p < .01$). This variable is factorial, with two levels indicating whether the alert notification appeared late (after the hazard had passed). The probability of disabling automation before or at the launch zone in a scenario where the alert came late (curve scenario) was significantly lower (*EM* = 0.01; *SE* = 0.01) than when the alert appeared on time (*EM* = 0.53; *SE* = 0.11).

2.5.2. Development of passive fatigue and mental workload

This section includes analyses related to passive fatigue development. We hypothesized that participants who had the opportunity to play Trivia would be slower in developing passive fatigue and, thus, be more vigilant than drivers who had no such opportunity (control group).

We used the KSS scores across the four instances to measure drivers' passive fatigue progression. In addition, we used LogQuotientRMSSD and LogQuotientSDNN as proxies of HRV to objectively evaluate passive fatigue progression.

KSS scores analysis. The final Linear regression model of KSS scores yielded a significant main effect of KSS instance ($X_3^2 = 172.5$, $p < .01$). Post-hoc pairwise comparisons analysis revealed that the estimated mean of KSS scores in the first instance (*EM* = 1.98, *SE* = 0.25) was significantly lower than in the second (*EM* = 2.9, *SE* = 0.25; $p_{adj} < .01$). In addition, the estimated mean of KSS scores in the second KSS instance was significantly smaller than in the third (*EM* = 3.88, *SE* = 0.25; $p_{adj} < .01$). There was no significant difference between the third and fourth KSS instances.

The model also included two significant interactions between the KSS instance and gender ($X_1^2 = 21.28$, $p < .01$, see Figure 2.12 left panel) and between the drive number and gender ($X_1^2 = 5.24$, $p < .05$, see Figure 2.12 right panel).

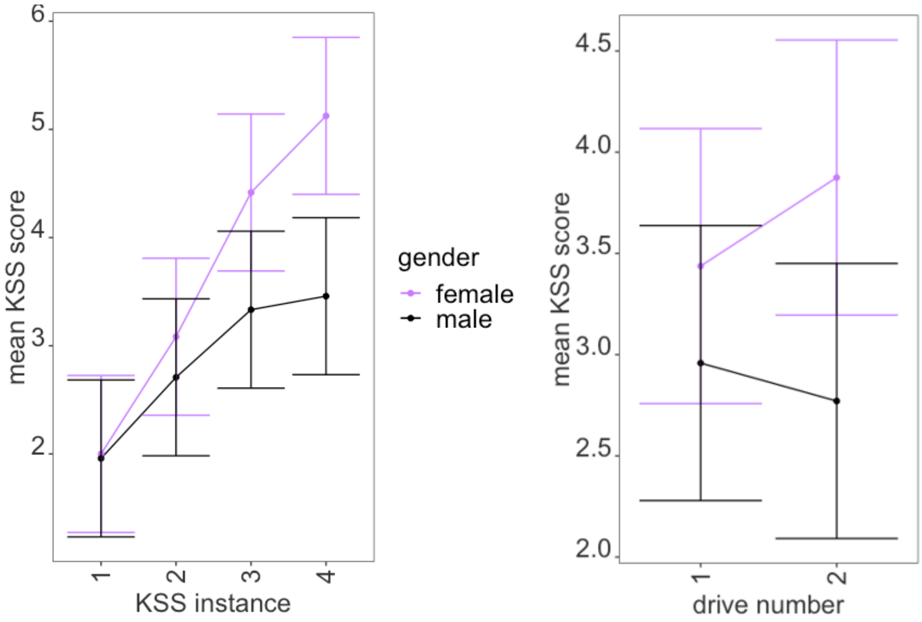


Figure 2.12 The interaction between KSS instance and gender (left panel) and between drive number and gender (right panel).

Concerning the first interaction between KSS instance and gender, posthoc contrasts pairwise comparisons analysis revealed that the estimated mean KSS score of female participants in the first instance (*EM* = 2, *SE* = 0.36) was significantly lower than in the second instance (*EM* = 3.08, *SE* = 0.36; $p < .01$). In addition, their estimated mean KSS score in the second KSS instance was significantly smaller than in the third instance (*EM* = 4.42, *SE* = 0.36; p < .01), which was significantly lower than in the fourth instance (*EM* = 5.12, *SE* = 0.36; p < .05). The estimated mean KSS score of the male participants in the first instance (*EM* = 1.96, *SE* = 0.36) was significantly smaller than in the second (*EM* = 2.71, *SE* = 0.36; p < .05). Furthermore, the estimated mean KSS score in the second KSS instance was not significantly different than in the third (*EM* = 3.33, *SE* = 0.36; $p_{adj} = .11$), which was not significantly different than in the fourth (*EM* = 3.46, *SE* = 0.36; $p_{adj} = .97$).

Regarding the second interaction between the drive number and gender, posthoc pairwise comparisons analysis revealed that the estimated mean KSS score of female participants in the first driving session (*EM* = 3.44, *SE* = 0.33) was significantly lower than in the second (*EM* = 3.88, *SE* = 0.33; $p < .05$). The estimated mean KSS score of male participants in the first driving session (EM = 2.96, SE = 0.33) was not significantly different from the second driving session (*EM* = 2.77, *SE* = 0.33; $p = .33$). In addition, the main effect of the experimental group was non-significant ($X_1^2 = 0.02$, $p = .89$).

HRV preceding a KSS instance

logQuotientRMSSD and logQuotientSDNN (HRV). As mentioned above, these measures are used as a proxy for HRV. We computed these two measures in two ways: in a segment of 45 seconds preceding the onset of the KSS instances (four times) and for the overall driving session (see Table 2.3). Since the second way yielded no significant main effect, we focus on the RMSSD and SDNN measurements as computed in the first way (segments of 45 seconds preceding a KSS instance). In these analyses, we compare the subjective perception measure (i.e., KSS) and the physiology measurements (i.e., RMSSD and SDNN). Each measure was modelled separately using Linear regression. The following section describes the significant effects of RMSSD and SDNN measured before the onset of the KSS instance.

logQuotientRMSSD. As mentioned above, the interaction between the experimental group and the driving session was non-significant in the final RMSSD model ($X_1^2 = 0.16$, $p = .7$; $X_1^2 = 0.08$, $p = .78$). However, the interaction between the experimental group and the driving session was significant ($X_1^2 = 6.06$, $p < .05$). Post-hoc pairwise contrasts comparisons analysis revealed a non-significant difference between the driving sessions for the control group ($p = .12$). For the control group, the measure for the first driving session was (*EM* = -0.098, *SE* = 0.102), and for the second driving session (*EM* = -0.232, *SE* = 0.102). There was a marginally significant difference between the driving sessions for the Trivia group ($p = .055$). For the Trivia group, the measure for the first drive was (*EM* = -0.197, *SE* = 0.102), and for the second drive (*EM* = -0.029, *SE* = 0.102) shown in Figure 2.13 on the left.

logQuetientSDNN. The final SDNN model included the interaction between driving sessions and the KSS instance ($X_3^2 = 8.6$, $p < .05$). A post-hoc pairwise comparisons analysis revealed a significant difference between the driving sessions only for the 1st KSS instance ($p < .05$). For the 1st instance, in the first driving session, the measure was (*EM* = -0.332, SE = 0.102), and in the second driving session, it was (*EM* = -0.022, *SE* = 0.102). For the other KSS instances, the difference between the driving sessions was non-significant, as shown in Figure 2.13 on the right (2nd instance: $p = .19$, 3rd instance: $p = .75$, 4th instance: $p = .12$).

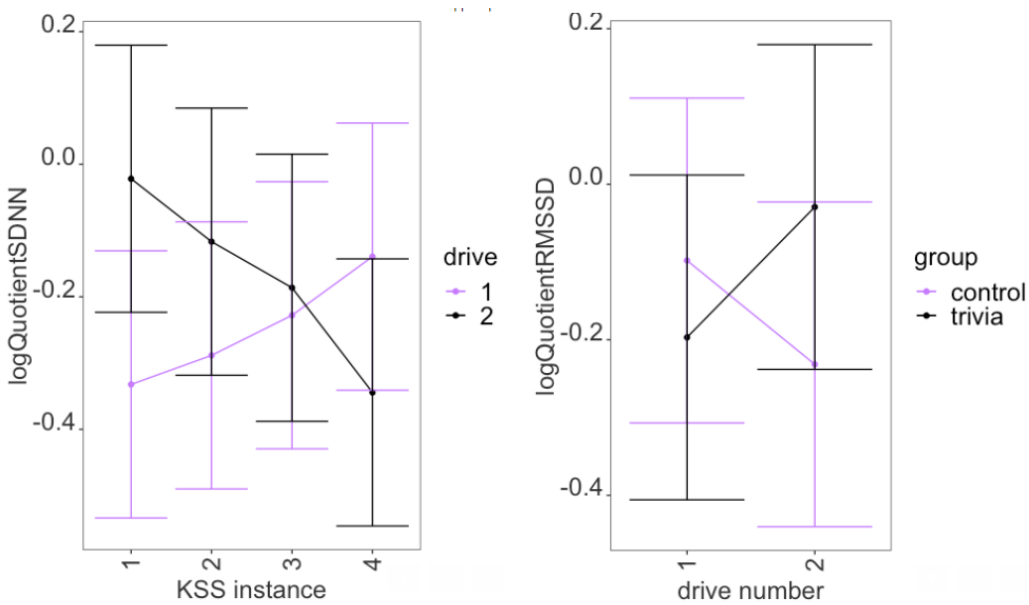


Figure 2.13 Fatigue progression as a function of KSS instance (on the left) and over time (on the right).

Mental workload using NASA-TLX

A linear regression model was applied to analyse mental workload rating scores. We analysed four dependent variables representing mental workload constructs: mental demand, physical demand, effort, and frustration. The mental workload ratings reported at the end of each driving session averaged across all questions of each construct. The final models yielded one common significant main effect of engagement with the Trivia ($X_1^2 = 4.61$, $p < .05$; $X_1^2 = 4.6$, $p < .05$; $X_1^2 = 4.38$, $p < .05$; $X_1^2 = 7.98$, $p < .01$ for mental demand, physical demand, effort, and frustration respectively). The engagement with the Trivia significantly decreased the mental workload compared to the control condition, as shown in Figure 2.14. In addition, the final model of mental demand yielded a significant drive number main effect ($X_1^2 = 3.77$, $p = .05$). Thus, the reported mental demands in the second driving session (*EM* = 3.02; *SE* = 0.275) were higher than in the first (*EM* = 2.64; *SE* = 0.243).

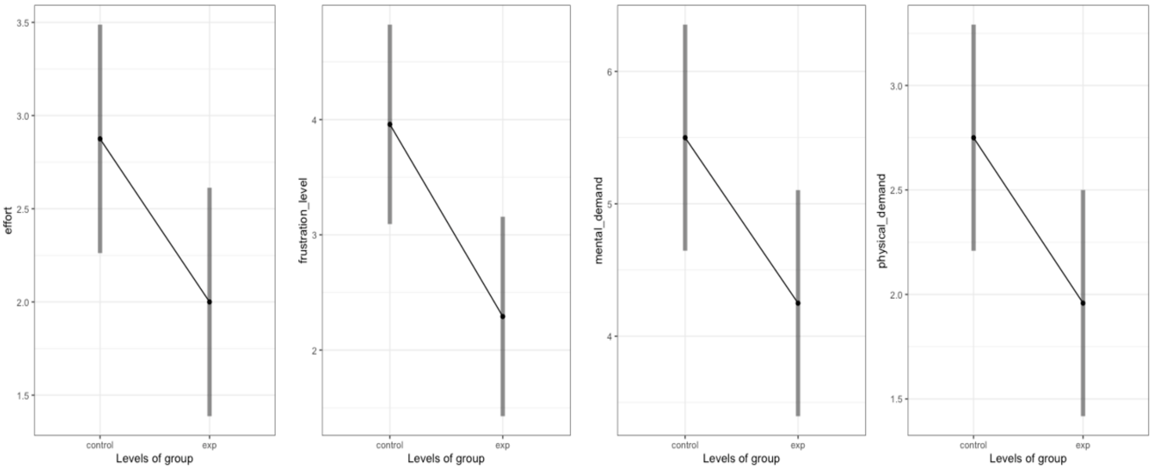


Figure 2.14 Mental workload constructs.

2.5.3. RQ3, RQ2: Adaptation and Trust in the Mediator HMI

In this section, we report the results of participants' perceptions toward adoption and trust in the short and long term. We assessed participants' acceptance and trust in automated vehicles using the SUaaVE questionnaire thrice (pre-driving and twice, once after each driving session; see Table 2.2). Participants rated statements related to four constructs: acceptability, trust, convenience, and safety, on a 7-point Likert Scale. Generally, participants had a positive attitude toward vehicle automation. The final linear regression model yielded a significant main effect of the driving session ($X^2_1 = 10.84$, $p < .001$). That is, participants, rated all trust constructs (except acceptance) significantly higher in the second driving session (*EM* = 5.16, *SE* = 0.17) than in the first (*EM* = 5.07, *SE* = 0.17)

In addition, the final linear model yielded a significant main effect of gender ($X^2_1 = 19.12$, $p < .001$). This effect showed that males provided higher ratings for the adoption & Trust constructs (*EM* = 5.93, *SE* = 0.23) than females (*EM* = 4.90, *SE* = 0.23). A marginally significant effect was found among participants who engaged with Trivia and rated adoption & Trust constructs significantly higher ($X^2_1 = 3.41$, $p = .06$) than the Non-Trivia group after both drives (*EM* = 5.51 *SE* = 0.229 and *EM* = 4.90 *SE* = 0.229). There is no significant change between drives in any questionnaire dimension for the Trivia group. For the control group, there was a significant decrease in the average score between the registration and the first driving session. This decrease was significant across all four dimensions (Figure 2.15). For the control group, there were no significant differences between the first and second driving sessions.

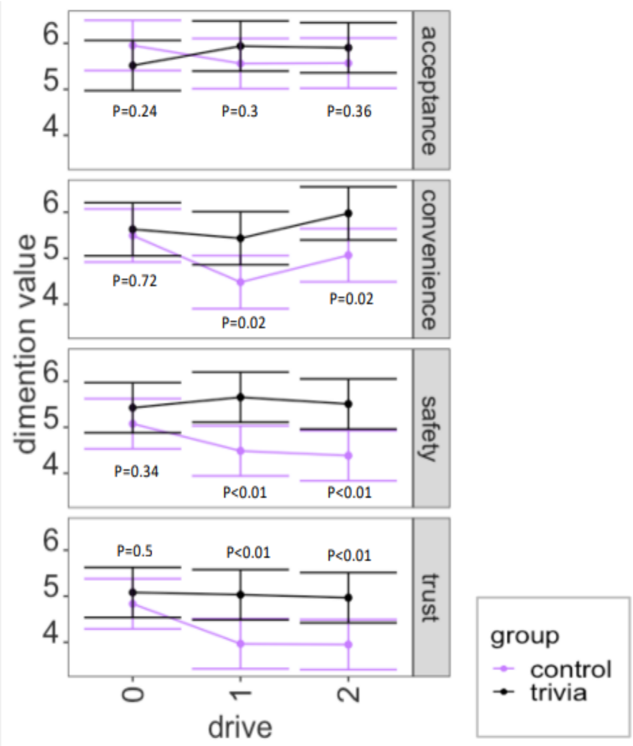


Figure 2.15 Trust Convenience, acceptance, and safety perception.

2.5.4. RQ2, RQ3: Usability of the Mediator HMI

Acceptance and Usability. The acceptance and usability questionnaire was administered twice after each driving session (see Table 2.2). We applied a linear regression model to analyse the Mediator HMI's acceptance and usability in each of the two displays: the driver dashboard and the infotainment panel.

The perceived usability of the dashboard and infotainment components was rated significantly higher by the Trivia group than by the Non-Trivia group (see Figure 2.16). The final linear regression model of system perceived usability yielded a significant main effect of the experimental group ($X_1^2 = 6.25$, $p < .05$). Post-hoc pairwise comparisons analysis revealed that the estimated mean of the dashboard component perceived usability score was higher for the Trivia group ($EM = 90.1$, $SE = 2.97$) than for the Non-Trivia group ($EM = 80.4$, $SE = 2.97$; $p_{adj} < .05$). The estimated mean of the infotainment component perceived usability score was also higher for the Trivia group ($EM = 89.3$, $SE = 2.97$) than for the Non-Trivia group ($EM = 79.2$, $SE = 2.97$; $p_{adj} < .05$). In addition, scores for both displays for the trivia group were above 80 and considered excellent (Bangor et al., 2009).

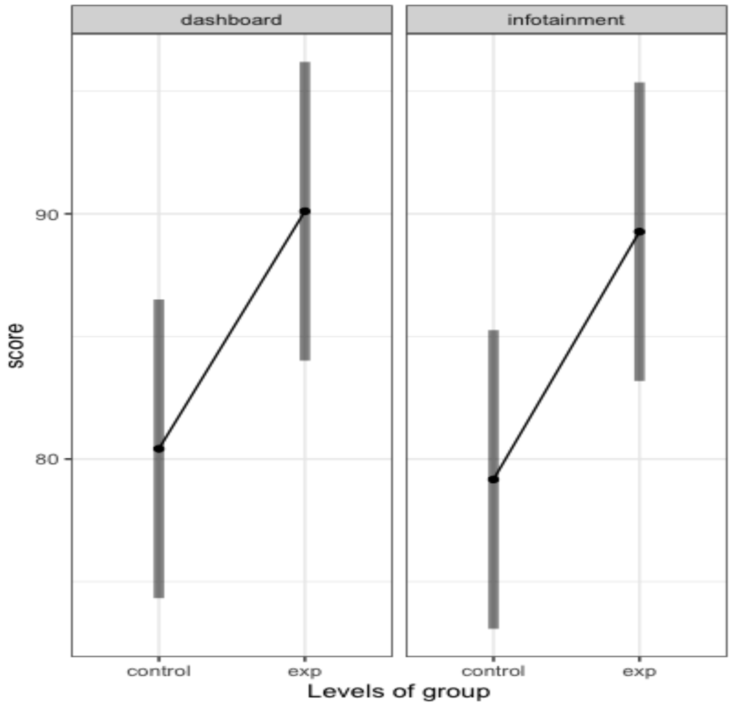


Figure 2.16 Usability ratings (SUS).

2.6. Summary and conclusions (fatigue)

The current study aimed to elaborate upon the previous research by addressing the following questions:

- **RQ1.** Will the Mediator HMI concept of preventive and corrective mediation help to mitigate the adverse effects of passive fatigue development under CM driving conditions? We hypothesized that passive fatigue would develop slower among drivers who play an activating Trivia task (preventive and corrective mediation) and will be more vigilant and aware of the traffic environment.
- **RQ2.** Will the positive effects of engagement with an activating NDRT, such as Trivia under CM driving conditions, persist over time? We hypothesized that the positive effects would last even when participants were familiar with the Mediator HMI and Trivia (second driving session)
- **RQ3.** Will participants accept, trust, and perceive the Mediator system as useful in the short and long(er) term?

Concerning the first research question, the results showed that fatigue progression was evident in KSS but not in HRV measures. There were no differences between the Trivia and Non-Trivia groups, and they both showed a significant increase in KSS scores from the first instance to the second and from the second to the third. Nevertheless, the average KSS scores never went above 5, suggesting that the increase in KSS levels was relatively moderate. The evidence that KSS levels increased and HRV remained the same indicates that drivers applied psychological adaptation but not physiological adaptation.

Consistent with Hancock and Warm’s (1989) model (Figure 2.17), as hypo-stress (underload) increases, humans begin first to adapt psychologically, and only when a certain level of hypo-stress is reached physiological adaptation begins.

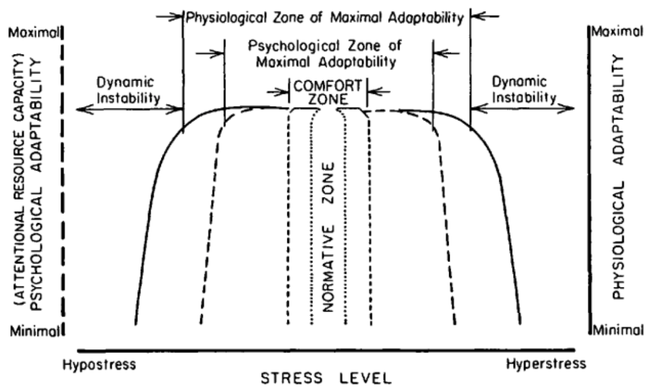


Figure 2.17 The stress adaptation model of Hancock and Warm (1989).

The fact that we did not find evidence for Trivia mitigating fatigue progression may be because the driving sessions in the experiment were not long enough to yield physiological adaptation. It could be that if driving sessions were longer and resulted in psychological adaptation as well, then the Trivia effects would become more salient. Another possible reason we did not find evidence of fatigue progression differences between the Trivia and Non-Trivia groups is that the notifications of upcoming hazards' component served as an alerting mechanism that helped both experimental conditions. Thus, the Non-Trivia group's vigilance was also preserved to some extent.

Although we did not find fatigue-related differences between the Trivia and Non-Trivia groups, we discovered that Trivia helped maintain situation awareness compared to the Non-Trivia group. Recall that all instances of Trivia occurred between scenarios and never within scenarios. Nevertheless, the results showed that before the appearance of the hazard notification in the HMI, the Trivia group had a significantly higher probability of identifying the cue preceding the hazard than the Non-Trivia group. In addition, given that a hazard was identified, we found that the Trivia group had a significantly higher number of fixations on the hazard than the Non-Trivia group. These results are consistent with Atchley and Chan (2011) study showing that passively fatigued drivers who engaged in verbal Trivia demonstrated better lane-keeping and steering control performance and higher vigilance than drivers who were not engaged with a task.

Concerning whether these positive effects persist over time (RQ2 and RQ3) we found consistent evidence that the Mediator HMI was effective in both driving sessions. We also found that participants who experienced Trivia and hazard notifications as preventive mediation perceived the interface as more convenient, usable, acceptable, and safe than the Non-Trivia group in both drives. In addition, while for the Non-Trivia group, there were some questionnaire dimensions where their scoring decreased between the first and second driving sessions, for the Trivia group, it remained unchanged.

To summarize, the fatigue study has shown the benefits of using Trivia as an activating task. The study showed that these benefits also persist long-term and improve the perception of the HMI's usability, trust, safety, and acceptance. Future studies could be focused on studying whether such positive effects persist over longer periods, such as months or years.

2.7. Method (distraction)

The distraction study aimed to elaborate upon the previous research by addressing the research questions in section 2.2. To address these research questions, we asked participants to arrive at the driving simulator twice one week apart. Participants drove for about 20 minutes in each session under CM driving conditions and experienced four latent hazardous situations. Participants were asked to handle NDRTs during each driving session as explained next.

2.7.1. Participants

Twenty-six participants were randomly assigned to two CM driving sessions, including NDRTs. All participants were undergraduate students from Ben-Gurion University of the Negev in Beer-Sheva city in Israel, participating as paid volunteers. Half of the participants were males (mean age = 26.08 years, *SD* = 1.49; mean driving experience = 8.5 years, *SD* = 2.25), and half were females (mean age = 24.14 years, *SD* = 1.06; mean driving experience = 6.57, *SD* = 1.24). Participants received monetary compensation for their participation (~$40). All participants had normal or corrected-to-normal visual acuity (contact lenses only), normal contrast sensitivity, normal colour vision, and no background of heart problems. Vision tests included a Snellen chart for visual acuity and Ginsburg's functional acuity contrast test (FACT) chart for contrast sensitivity evaluation (Ginsburg, 2002). Participants had a valid driver's license for at least five years without any previous experience with adaptive cruise control and lane-keeping assistance. This study received the BGU ethical committee's approval.

2.7.2. Tools and apparatus

The tools used in the distraction study (except for the Trivia game) are described in detail in section 2.3. For the distraction study, we adopted a WhatsApp-like visual-manual task (Zangi et al., 2022) that was presented on the infotainment display.

2.7.2.1. Non-Driving related task

The high-level decision logic is expected to initiate or respond with an appropriate action when disengagement from driving-related activities may encourage a driver to engage with an NDRT. However, the engagement with a manual-visual NDRT under CM driving conditions impairs drivers' ability to monitor the driving task and automation and their ability to safely assume manual control (Hecht et al., 2022, Reagan et al., 2021). In the current study, a WhatsApp-like task was adopted from Zangi et al. (2022) study and used as a distracting NDRT. The NDRT was displayed visually on the infotainment display, and answering the questions was done by choosing the correct answer using the touch screen. The invitation to play was accompanied by a visual message. If the participant accepted the invitation, the task started. The WhatsApp message included objective details on several topics (e.g., changes in the train schedule), and included several informative elements. Two questions followed the message presented one after the other. The participant had to answer each question by selecting the correct answer from four options and pressing the relevant area on the touchscreen. If the participant did not choose an answer within 10s, the next question appeared (see Figure 2.18). Each WhatsApp dialog screen appeared for a maximum of 10s.

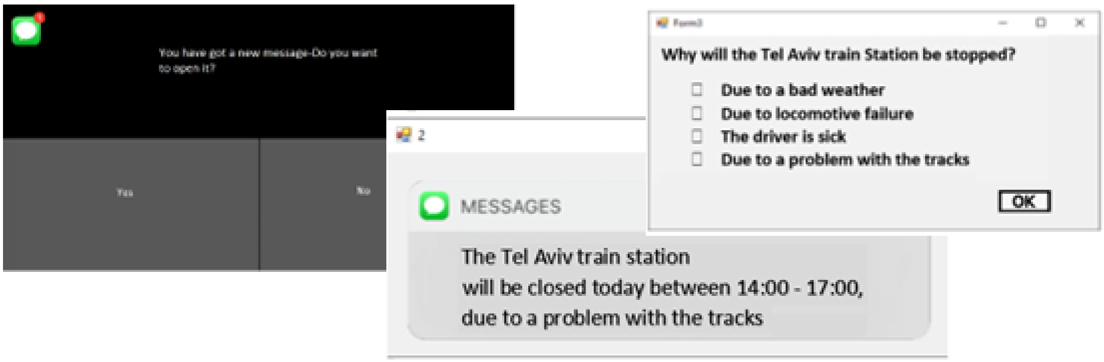


Figure 2.18 invitation to play WhatsApp (on the left) and an example of a question and corresponding answers.

2.7.3. Experimental design

The experiment was a within-subjects factorial design. Participants were invited to drive in the driving simulator twice, one week apart. Each driving session lasted about 20 minutes. This independent variable was within subjects and examined the long-term effects of the Mediator HMI concepts. In this study, only the notifications of upcoming hazards were used for preventive mediation. Trivia was unavailable in this experiment. In addition, corrective mediation was unavailable.

Each driving session included a pre-defined set of four unmaterialized hazardous scenarios and seven WhatsApp tasks (Figure 2.19).

Figure 2.19 Schematic description of an experimental drive in the distraction study.

The green icon in Figure 2.19 represents a WhatsApp (NDRT) instance and the black-red icon represents an upcoming hazard notification. Four WhatsApp tasks were each assigned to a hazardous scenario and triggered between 3 and 6 seconds before the notification of an upcoming hazard. The other three WhatsApp tasks were located sporadically before the first hazard, between the first and second hazards, and between the third and fourth hazards, respectively. The driving environment was similar to that in the fatigue study, only here, the road segments connecting between hazards were shorter (about 4 minutes each; see general method section 2.3). The analyses of this study focused on the questionnaires in Table 2.3.

2.7.4. Procedure

Before the participants arrived in the lab for the first time, they filled out the following forms and questionnaires:

1. Registration form containing details and conditions for participating in the experiment
2. Consent form
3. Demographics questionnaire
4. Participant's driving demographics
5. Familiarity with the automation questionnaire
6. Adoption and trust in automated vehicles questionnaire

When arriving at the lab for the first time, participants gave the experimenter the forms they filled out at home and conducted visual acuity and contrast sensitivity tests. Qualified participants received written instructions regarding the simulator automation capabilities, the simulated environment, the HMI system, and the remaining procedure of the session. In the written instructions, participants were informed that they carry full responsibility for the driving task and safety and should monitor the automation continuously despite the automation capabilities. The participants could assume manual control whenever necessary by pressing the brake or gas pedals. Participants were instructed to drive as they usually would in similar real-world situations.

At the beginning of both sessions, participants were connected to the ECG device using electrodes attached to their chest and asked to sit and relax for 5 minutes while reading a magazine. Meanwhile, a baseline ECG, including R-R interval measurement, was recorded. Next, the participants entered the vehicle, wore eye-tracking glasses, and underwent a short calibration process. Then, the participants drove a 5-minutes training session to familiarize themselves with the simulator, the measurement apparatus, and the HMI system. The training drive was identical to the one used in the fatigue study regarding the road environment and cultural objects. Then, the participants filled out a short knowledge questionnaire regarding the simulated CM vehicle operation.
Immediately after training, the participants began the experimental driving session. At the end of the driving session, participants completed a post-drive questionnaire consisting of a NASA-TLX questionnaire, usability evaluation of the HMI questionnaire, acknowledgment of the HMI indication functionality questionnaire, and adoption and trust of automation questionnaire (see Appendix 5.1). A week later, each participant returned to the lab approximately at the same hour as the first session and underwent the same procedure as the first session. After completion, the participant was thanked for her participation and received monetary compensation. Figure 2.20 presents a schematic description of the experimental procedure.

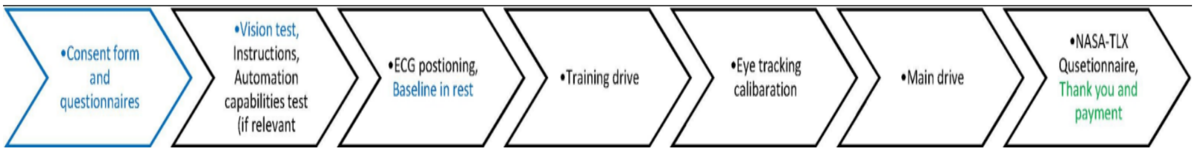


Figure 2.20 Schematic description of the experimental procedure.

2.8. Results (distraction)

The results will focus on the questionnaires administered to the participants during the experiment. We were highly interested to see how participants' preferences and attitudes impact the use and acceptance of the mediator's HMI and how participants adapt to this novel concept in the longer term. The eye-tracking data related to hazard perception performance are not provided in this deliverable as they do not differ from that reported in our previous study (Zangi et al., 2022; see section 2.1). All statistical analyses were carried out in R software at a significance level of 0.05. We used a linear regression model within the Linear Mixed Models (LMM) framework for the

dependent variables that were normally or log-normally distributed. These variables included the scores of perceived acceptance, convenience, safety, and trust in automation; mental workload scores; usefulness, satisfaction, and usability scores of the Mediator systems.
The initial model in each analysis included the following fixed effects: drive number (1/2), gender (M/F), and the first and second-order interactions between the variables. In addition, the random effects of the initial model included the participants. Table 2.4 contains additional information about each analysis. The final model of each analysis was achieved via a backward elimination procedure where all non-significant effects were removed from the model. For significant fixed effects with more than two levels, a post-hoc pairwise contrasts comparisons analysis was applied, and the Tukey HSD procedure was used for alpha correction.

Table 2.4 Summary of the LMM models used in the study

Measurement	Measured variables	Family	Random effects	Fixed effects in the initial model
Mental Workload	Nasa-TLX	Linear regression	Participant	Experimental group, drive number, gender
PAV adaptation	Acceptance	Linear regression	Participant	Experimental group, drive number, gender
	Trust			
	Convenience			
	Safety			
Usability of Mediator HMI	Dashboard Usefulness	Linear regression	Participant	Experimental group, drive number, gender
	Infotainment Usefulness			
	Dashboard Satisfaction			
	Infotainment Satisfaction			
	Dashboard SUS			
	Infotainment SUS			
	Glance duration on task ratio			

2.8.1. RQ2, RQ3: Perception of automated driving

In this section, we report the results of participants' perceptions toward adoption and trust in the short and long term. We assessed participants' acceptance and trust in automated vehicles using the SUaaVE questionnaire thrice (pre-driving, i.e., drive 0) and once after each driving session (one and two; see Table 2.2). Participants rated statements related to four constructs: acceptability, trust, convenience, and safety, on a 7-point Likert Scale. Each construct was analysed separately. The final Linear regression models yielded a significant main effect of gender for all of the constructs – trust (X_1^2 = 8.07, p < .01); acceptance (X_1^2 = 36.79, p < .001); convenience (X_1^2 = 9.68, p < .01); and safety (X_1^2 = 47.38, p < .001). Males rated higher than females in all four constructs – trust (males: *EM* = 5.36, *SE* = 0.158; females: *EM* = 3.86, *SE* = 0.137); acceptance (males: *EM* = 5.89, *SE* = 0.143; females: *EM* = 4.59, *SE* = 0.151); convenience (males: *EM* = 4.38, *SE* = 0.125; females: *EM* = 3.7, SE = 0.161); and safety (males: *EM* = 5.62, *SE* = .108; females: *EM* = 4.49, *SE* = .119).
Overall, males tended to trust and adopt the automated system to a greater extent than females, although both males and females demonstrated positive attitudes, especially for acceptance and safety.

2.8.2. RQ2, RQ3: Usability of the Mediator HMI

The Usability and Acceptance questionnaire was administered once after each driving session (one and two; see Table 2.2). We analysed the Mediator HMI's usability, usefulness, and satisfaction in each of the two HMI display components: the instrument panel dashboard and the infotainment displays. To analyse usability, usefulness, and satisfaction, we used two questionnaires: the Acceptance assessment (Van der Laan et al., 1997) and System Usability Scale (SUS; Brooke 1996), separately for each HMI display component (four models altogether). The initial model in each analysis included the following fixed effects: drive number (1/2), and gender (M/F). In addition, the random effects of the initial model included the participant and the question. The final Linear regression models yielded a significant main effect of drive number for SUS and acceptance assessment models for both HMI display components. For the Instrument panel dashboard's display, usability ($X_1^2 = 4.65$, $p < .05$) and Acceptance assessment ($X_1^2 = 6.59$, $p < .01$). For the infotainment display usability ($X_1^2 = 5.3$, $p < .001$) and Acceptance assessment ($X_1^2 = 9.58$, $p < .01$). Post-hoc pairwise comparisons analysis showed that second drive was rated higher than the first drive for instrument panel dashboard (SUS *EM* = 81.1, *SE* = 1.46 and *EM* = 84.43, *SE* = 1.36; Acceptance assessment *EM* = 6.2, *SE* = .145 and *EM* = 6.59, *SE* = 0.16) and for the infotainment display (SUS *EM* = 79.7, *SE* = 1.7 and *EM* = 85.0, *SE* = 1.3; Acceptance assessment *EM* = 5.98, *SE* = .162 and *EM* = 6.47, *SE* = 0.18) as shown in Figure 2.20.

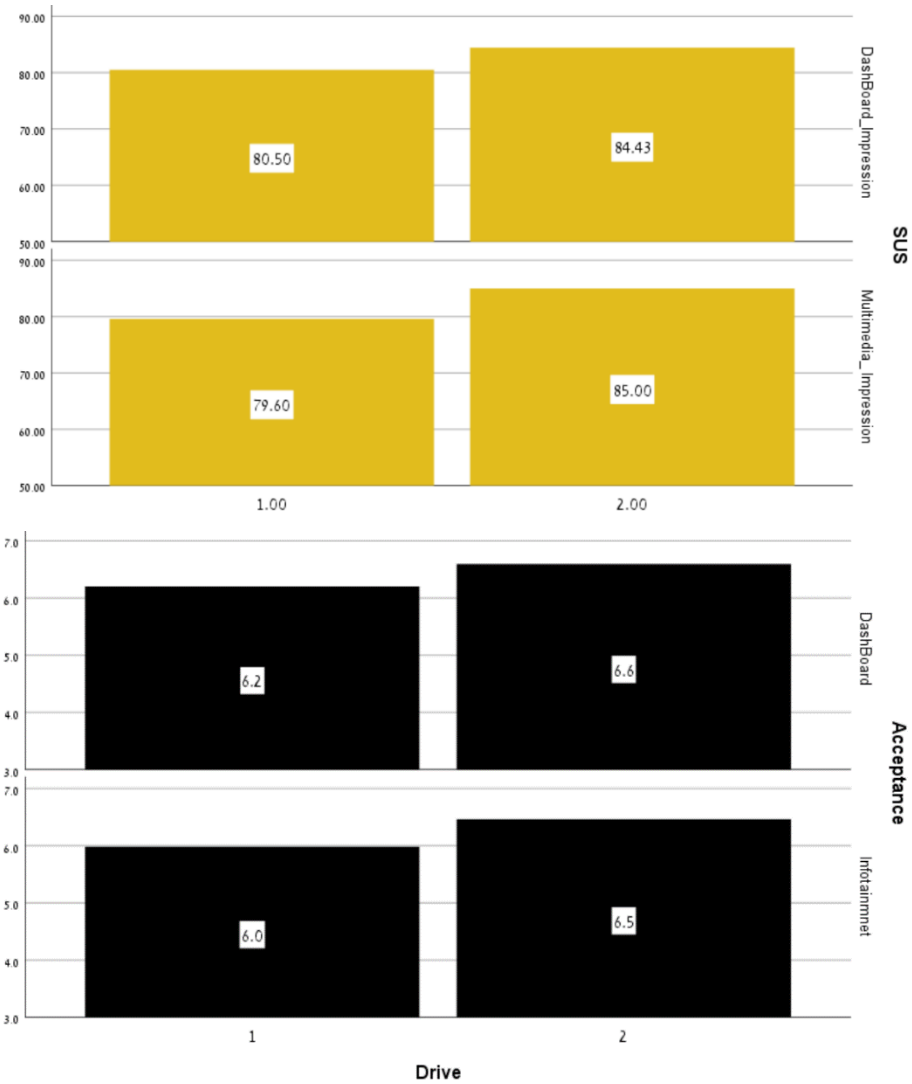


Figure 2.21 Usability (impression) and Acceptance assessment for the dashboard and infotainment displays systems.

Usability evaluation of the infotainment display showed significant effects of gender and an interaction between drive and gender (X_1^2 = 11.55, p < .01; X_1^2 = 5.25, p < .01 accordingly). Post-hoc pairwise comparisons revealed that males evaluated the usability of the infotainment display significantly higher than females (*EM* = 86.4, *SE* = 2.8; *EM* = 79.16, *SE* = 2.63). As shown in Figure 2.22, the estimated mean score of females was significantly higher in the second drive than in the first (*EM* = 77.14, *SE* = 1.6, *ME* = 83.56, *SE* = 1.3). There were no other significant differences. Overall, at the end of the second drive, all participants perceived the infotainment display as highly usable, regardless of gender.

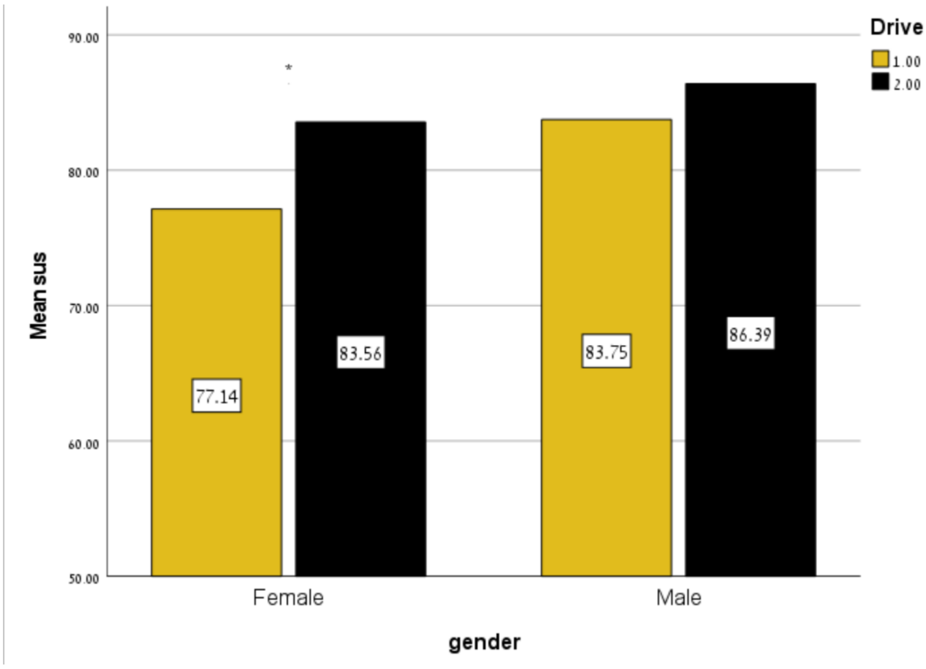


Figure 2.22 Usability evaluation for the multimedia.

2.8.3. RQ1, RQ2: Workload

The final linear regression model yielded a significant main effect of drive number ($X_1^2 = 14.69$, $p < .01$). The estimated mean rating of mental workload in the first drive was higher than in the second drive (*EM* = 3.59, *SE* = 0.33; *EM* = 2.48, *SE* = 0.34). There were no other statistical differences.

2.9. Summary and conclusions (distraction)

To summarize, the analyses of the questionnaires revealed that the Mediator HMI was accepted and trusted by the participants. Overall, participants perceived the Mediator HMI as safe and trusted and tended to favour its adoption. These attitudes were enhanced in the second drive, particularly among females. Furthermore, at the end of the second drive, males and females perceived the Mediator HMI as highly usable.

In terms of mental workload, we found that participants perceived the Mediator HMI as less mentally demanding during the second drive than the first, suggesting positive longer-term effects.

2.10. Discussion summary of the fatigue and distraction studies

The results of the fatigue and distraction studies can be concluded as follows.

First, concerning RQ1, will the Mediator HMI preventive and corrective mediation concept assist in mitigating the adverse effects of distraction and fatigue? In the fatigue study, using an audio-visual Trivia game as a preventive mediation helped maintain situation awareness compared to drivers who were not playing Trivia, despite the fact we did not find fatigue-related differences (HRV or KSS). As for the hazard notification system, it was found helpful in attracting drivers' attention (regardless of whether they played Trivia or not) to typical on-road hazards and improved safety.

Second, concerning whether these positive effects persist over time (RQ2 and RQ3), we found consistent evidence that the Mediator HMI was effective in both driving sessions of the fatigue study. We also found that participants who experienced Trivia and hazard notifications as preventive mediation perceived the interface as more convenient, usable, acceptable, and safe than the Non-Trivia group in both drives. In addition, while for the Non-Trivia group, there were some questionnaire dimensions where their scoring decreased between the first and second driving sessions, for the Trivia group, it remained unchanged. Similarly, in the distraction study (where the Trivia game was absent), participants perceived the Mediator HMI as useful, safe, and trusted. Participants also indicated their willingness to adopt this type of system. While the males' attitudes and perceptions were positive during both driving sessions (short and longer term), females tended to be more suspicious during the first driving session but similar to the males in the second driving session. Gender related attitudes and perceptions toward automated vehicles and advanced driver assistance systems show a complex picture and are typically examined with other demographic details such as age and personality traits. Some survey related studies show, for example, that female drivers are more concerned about safety and reliability of automated vehicles than males (Greenwood et al., 2022) and that male drivers are more willing to relinquish driving control than females (Charness et al., 2018). Other studies, on the other hand, provide evidence of more complex patterns (e.g., Golbabaei et al., 2020).

2.10.1. Mediator concept of preventive and corrective mediation

The Mediator HMI embedded a novel concept called preventive and corrective mediation (Figure 2.4). This concept aims to ensure that drivers in CM and SB driving situations maintain their fitness to drive whether they are currently in control of the driving task (human control) or the automated driver is. The difference between preventive and corrective mediation is that in preventive mediation, the driver's fitness has not degraded yet due to fatigue or distraction. Still, the Mediator HMI anticipates that such performance degradation can happen in the future (e.g., approaching a long monotonous road, a long traffic jam, etc.). In corrective meditation, however, drivers' fitness due to distraction or fatigue has already occurred, and the HMI initiates actions to stall further degradation.

BGU has implemented two preventive/corrective mediation components that resulted in improved safety and positive attitudes and perception toward the HMI.

The first component is the Trivia game (section 2.3.1.2). We showed that when Trivia was triggered as preventive mediation, drivers who played the Trivia game during continuous mediation driving conditions demonstrated increased alertness than participants who did not play Trivia. Drivers who played Trivia also rated the HMI as more usable and trustworthy than drivers who did not play.

The second component is notifications of upcoming hazardous situations. In the fatigue study, it served as preventive mediation, and in the distraction study, it served as corrective mediation. As preventive mediation, this component contributed to the system's safety, showing that even the Non-Trivia group, who were less vigilant than the other group, could identify hazards at the same probability as the Trivia group. As a comparison, Zangi et al. (2022) did not use such a hazard notification system. There, we showed that participants who were distracted by an NDRT were less likely to identify hazards than drivers who were not distracted under CM driving conditions (differences ranged from 10% for imminent hazards to over 30% for potential hazards). TUC used a traffic jam scenario as a preventive mediation component to offer drivers relief from driving in a traffic jam that can facilitate fatigue and mental workload. The results have shown that drivers perceived the benefits of using automation in such situations as positive and valuable.

2.10.2. Aspects of adaptation to the Mediator HMI (prolonged use)

Humans who interact with automated systems tend to modify their behavior, attitudes, and perceptions as experience with the automated system accumulates. BGU studies evaluated how drivers' performance, perceptions, and attitudes change between the first few experiences with the HMI (first driving session) and the last few experiences with it (second driving session). Overall, the results have shown that in the **fatigue study**, participants' attitudes and perceptions of the HMI were more positive in the second drive than in the first, particularly for drivers who could play Trivia. Regarding safety, the results show that the positive effects of Trivia decreased from the first to the second driving sessions. These results may imply that the novelty of the Trivia game subsided, and drivers perceived the task as less engaging and interesting. Notably, the topics in the Trivia were general, and the drivers could not choose topics according to their preferences. Older studies in manual driving (e.g., Gershon et al., 2009) emphasize the importance of using different themes such as movies, sports, and general knowledge and incorporating mechanisms to increase the task's difficulty level. Since the fatigue study's hazard notification system was 100% reliable, we did not see any adaptation from driving sessions 1 to 2.

Confirming the positive attitudes towards the HMI concept, the **distraction study** results showed that drivers' perceptions and attitudes of trust, safety, usability, and acceptance toward the Mediator HMI were highly positive for both males and females. For females who, during their first drive, were slightly more skeptical than males, there was a significant increase in their ratings from their first to the second drive. There were no significant differences between males and females in the second drive regarding subjective ratings.

3. Driving simulator study on comfort-related use cases

This section describes the simulator study conducted at Chemnitz University of Technology. The study focussed on comfort use cases with related TOC, simulated automation degradation and related TOC by the human driver, as well as driver characteristics. The study aimed primarily on the user's evaluation of the Mediator system / HMI (i.e. acceptance, trust, usability, comfort, experience of TOC), including the analysis of driver characteristics effects such as age, gender and driving experience.

For this purpose, participants experienced the Mediator system with its HMI elements in four different conditions. Several standardized questionnaires, self-designed items and interview questions were used to examine drivers' evaluation of Mediator for different use cases.

3.1. Introduction

As described in MEDIATOR deliverable D1.1 (Christoph et al., 2019), driver characteristics can play an important role for drivers' experience and behaviour in traffic. Especially age, driving experience and partly also gender can influence driver behaviour. Hence, driver characteristics need to be examined as well when introducing new systems for (partly) automated driving. Additionally, the deliverable described the importance of trust as a key element in human-technology relationships, which can affect safety, performance and usage behaviour (Schaefer et al., 2016). When driving with a system like Mediator, providing suggestions and support for the driver, trust plays an important role. Drivers' can actively decide to follow Mediator's suggestions (e.g., to activate automated driving) or not. One factor influencing usage frequency of technical systems such as Mediator is trust (Hoff & Bashir, 2015). Further, unnecessary interventions by the driver due to reduced trust in the system's performance can have an impact on driving safety. According to Hoff and Bashir (2015), trust can be distinguished in *trust prior to an interaction* (e.g., based on former experiences, individual's overall tendency to trust automation and perceived risks) and *trust formation during an interaction*. Trust during an interaction is formed based on the performance of the technical system. System performance includes factors like reliability, predictability, validity, dependability, difficulty of errors, timing as well as usefulness. To create trustworthy automation, several design recommendations were formulated like providing continuous feedback to increase transparency, using an adequate communication style, adjusting the automation level to user's preferences and simplifying the ease of use as much as possible (Hoff & Bashir, 2015). Therefore, factors like drivers' experience with ADAS and general opinion on automated driving (e.g., initial trust), Mediator's reliability, transparency, usability as well as drivers' trust in Mediator and further usage intention need to be examined in more detail.

The MEDIATOR Deliverable D1.2 (Borowsky et al., 2020) pointed out that user acceptance of automated vehicles is crucial to achieve expected benefits of automated driving (ERTRAC, 2019) like an increase of traffic efficiency and safety (Smith et al., 2018). Comfortable and positive driving experiences are considered as important prerequisites for acceptance and usage of automated vehicles (Bellem et al., 2018). Experienced comfort during automated driving can also have safety impacts. Reduced driving comfort or unexpected system performance can trigger unnecessary interventions by the driver (e.g., if apparent safety is perceived as compromised) leading to potentially safety-critical takeover situations (Hergeth et al., 2016; Techer et al., 2019). Hence, knowledge about drivers' comfort could allow for adapting automation features and/or information presentation to prevent disengagement of automation or dangerous and unnecessary takeover situations. Hence, drivers' acceptance of the system, perceived safety and comfort should be examined when introducing a system like Mediator to the drivers.

3.2. Objectives and research questions

The driving simulator study at Chemnitz University of Technology focused mainly on comfort TOC from manual to automated driving, simulated automation degradation and related TOC by the human driver, comfort critical situations (i.e., close approach to the rear-end of a traffic jam) as well as driver characteristics. Hence,

- comfort-related TOC: Take Over Requests (TOR) from Mediator due to, for instance, an upcoming traffic jam or incoming text messages and respective transitions from manual driving to (partly) automated driving (i.e., SB or CM),
- safety critical TOC: Planned or unplanned transitions from automation to manual driving due to degraded automation fitness,
- close approach scenarios: Uncomfortable driving manoeuvres for driving comfort estimation, and
- the influence of driver characteristics (e.g., age, gender, driving experience)

were examined.

The study focused strongly on drivers' evaluation of Mediator and its functionalities (e.g., acceptance, trust, usability, comfort, experience of TOC, evaluation of HMI concept) and not on the technical parameters (e.g., performance of developed algorithms and technical components). Therefore, it was important that all drivers examined the same scenarios with the same reaction of Mediator. Thus, no online detection and decision making was involved in the simulator study to ensure that drivers experienced an "optimal working" Mediator (e.g., no unplanned HMI reactions, no false positives or false negatives). Impaired driver states like fatigue as well as corrective and preventive actions were not in the focus of this driving simulator study.

The research questions that were investigated in this study were the following:

1. How is Mediator and its functionalities accepted by the drivers?
2. How is Mediator's potential to increase drivers' comfort?
3. How is the developed HMI experienced by the drivers (e.g., usability, transparency, comprehensibility, predictability)?
4. How much do the drivers trust Mediator and its functionalities?
5. How is the perceived safety rated by the drivers?
6. How are the comfort-related TOC from manual to automated driving experienced by the drivers?
7. How are the planned and unplanned TOC from automated to manual driving experienced by the drivers?
8. Which observable driver features are indicators for uncomfortable driving situations?
9. How good can eyes-off road time be detected by a camera system?
10. What is the safety impact of Mediator?

Additionally, different driver characteristics (e.g., age, gender, driving experience in terms of yearly driven mileage) were included in the analysis to examine the role of driver characteristics in answering the above-mentioned research questions.

3.3. Driving scenarios

Five use cases (UC 2, UC 3, UC 5, UC 6, UC 9) were covered by the driving simulator study. Participants experienced four different conditions of the same driving scenario in a fixed order (1 – baseline, 2 – CM, 3 – SB, 4 – SB + close approach). The driving scenario was characterized by the following features:

- highway with two lanes in each direction
- average driving speed: 80 km/h
- moderate surrounding traffic that did not provoke overtaking manoeuvres
- good weather (clear view, blue sky, no rain), except condition 2 – CM

In all conditions, drivers started in manual mode at a rest stop and, afterwards, drove in either manual mode or switched to an automated driving mode (i.e., SB or CM) depending on the condition. After several kilometres, they approached a white truck at the rear end of a traffic jam caused by a construction zone. The drivers came to a full stop and, afterwards, drove some kilometres in the traffic jam (in either manual or automated mode). After reaching the end of the traffic jam, they continued the trip on the highway at 80 km/h. The four conditions are described in more detail in below.

Condition 1 (baseline). The first condition was a baseline condition. Two kilometres before approaching the traffic jam, the HMI informed the drivers that a traffic jam was detected ahead. Drivers approached and drove through the traffic jam manually. At the end of the construction zone, the traffic jam dissolved and the drivers again drove on a free highway at 80 km/h. The drivers ended the simulation at a highway exit.

Condition 2 (CM). In this condition, it was foggy on the track. Two kilometres before approaching the traffic jam Mediator informed the drivers that a traffic jam was detected ahead. One kilometre before approaching the traffic jam, Mediator informed the drivers that automated driving was available. In this condition, only CM was available due to the fog (UC 6b). Drivers were instructed to follow the system’s instructions and were driven in CM. The vehicle approached the truck at the rear end of the traffic jam smoothly (i.e., slowed down well in advance, kept enough distance). Drivers were driven through the construction zone in CM until it shuts off after a few kilometres due to increasing fog. The drivers needed to take back vehicle control immediately (UC 9). At the end of the construction zone, the traffic jam dissolved and the drivers again drove on a free highway at 80 km/h. The drivers ended the simulation at a highway exit.

Condition 3 (SB). Two kilometres before approaching the traffic jam, Mediator informed the drivers that a traffic jam was detected ahead. One kilometre before approaching the traffic jam, Mediator informed the drivers that they received a new email and that automated driving (SB) was available (UC 3b). Drivers followed the system’s instructions and were driven in SB. At the same time, drivers engaged in a NDRT, here reading and answering an email. The vehicle approached the truck at the rear end of the traffic jam smoothly (i.e., slowed down well in advance, kept enough distance). Drivers were driven through the construction zone in SB. At the end of the construction zone, the traffic jam dissolved and the drivers were driven on a free highway at 80 km/h. After a few kilometres, the drivers received the message from Mediator that a highway exit approached (i.e., end of operational design domain will be reached) and that a TOC was necessary (UC 5a). Drivers followed the system’s instructions and took over vehicle control (UC 2). The drivers ended the simulation by taking the highway exit.

Condition 4 (SB + close approach). Two kilometres before approaching the traffic jam Mediator informed the drivers that a traffic jam was detected ahead. One kilometre before approaching the traffic jam, Mediator informed the drivers that automated driving (SB) was available (UC 3b). Drivers followed the system's instructions and were driven in SB. At the same time, drivers engaged in the NDRT, here reading and continuing in answering the email from before. When approaching the traffic jam, the approach was quite fast and the vehicle came to a stop quite close to the white truck at the end of the traffic jam (considered as an uncomfortable driving manoeuvre). Drivers had the opportunity to intervene (i.e., brake) leading to a shut-off of the automated driving mode. Afterwards, drivers could activate SB again while driving in the traffic jam at any time (in case some of the participants chose this option, we also covered UC 3a). Drivers then either were driven through the construction zone in SB or drove manually in case they took over vehicle control by intervening during the close approach event and not turning on automated driving again. After a few kilometres, the drivers received the message from Mediator that a highway exit approached (i.e., end of operational design domain will be reached) and that a TOC was necessary (UC 5a). Drivers followed the system's instructions and took over vehicle control (UC 2). The drivers ended the simulation by taking the highway exit.

3.4. Method

3.4.1. Participants

3.4.1.1. Gender, age and education

A total of N = 81 participants took part in the simulator study. They were recruited via an announcement on the website of the Chemnitz Technical University and in the local media (i.e. radio, newspapers). Seven participants could not finish the study due to simulator sickness (n = 6) or technical problems (n = 1), resulting in a sample of N = 74 participants who took part in the entire study.

In the following, results are rounded for better readability.

The participants' age ranged from 19 to 75 years (M = 40 years, SD = 17). It was intended to balance the participants regarding gender and age (according to the four age groups of the National Highway Traffic Safety Administration (NHTSA), 2016). Table 3.1 shows the achieved age and gender distribution of the participants.

Table 3.1 Distribution across NHTSA age groups and gender.

Age groups	N_{Male}	N_{Female}	N_{Total}
18-24 years	9	11	20
25-39 years	11	9	20
40-54 years	8	8	16
55 years and older	10	8	18
Total	**38**	**36**	**74**

Because the study materials as well as the HMI messages were all provided in German language, the sample included only German participants or non-German participants with sufficient German language skills.

See Appendix 5.2.2 for further information about the participants (i.e. educational degrees).

3.4.1.2. Driving experience

All participants held a valid driver's license, which was a prerequisite for taking part in the study. Time of license possession ranged from two to 53 years (*M* = 21 years, *SD* = 15). The annual mileage of the participants ranged from 1,000 to 60,000 km with an average annual mileage of *M* = 14,000 km (SD = 11,300). For further analyses, participants were divided into the following four groups based on their annual mileage: very low annual mileage (1,000-7,000km), low annual mileage (7,500-12,000km), high annual mileage (12,500-15,000km) and very high annual mileage (17,600-60,000km). Seven participants had especially high annual mileages due to their occupation (e.g., truck or bus drivers etc.).

Table 3.2 provides an overview of the relationships between the participants' characteristics (i.e., age, gender) as well as their driving experience. As expected, participants' age correlates positively with driver's license possession (large effect size), because most of the young drivers receive their drivers' license as soon as they are legally allowed to drive a car. The participants' gender did only correlate with the annual mileage (medium effect size) indicating that male participants had higher annual mileages than female participants did. Having a closer look at drivers who indicated to drive more due to their profession (e.g., bus driver, sales representative), data show that 6 out of 7 drivers in this group were male.

Overall, it can be seen that there are strong relationships between the participants' age and their driving experience (i.e. annual mileage and duration of license possession). The older and more experienced the drivers are (i.e. longer driver's license possession), the higher their annual mileage is. For further analyses, the annual mileage will be used as indicator of driving experience.

See Appendix 5.2.2 for further information about the participants' car availability, usage and mobility behaviour.

Table 3.2 Correlations of driver characteristics and driving experience.

	Age	Gender*	Annual mileage	Duration driver's license possession
Age	/	$r = -.12$ $p = .322$	$r = .22$ $p = .058$	$r = .95$ $p < .001$
Gender	$r = -.12$ $p = .322$	/	$r = -.22$ $p = .065$	$r = -.19$ $p = .117$
Annual mileage	$r = .22$ $p = .058$	$r = -.22$ $p = .065$	/	$r = .25$ $p = .037$
Duration driver's license possession	$r = .95$ $p < .001$	$r = -.19$ $p = .117$	$r = .25$ $p = .037$	/

* Coding of gender: 1 = male, 2 = female, 3 = diverse

Participants were asked about the availability of eight advanced driver assistance systems (ADAS) within the car they use as well as whether or not they use the available systems.
Figure 3.1 provides an overview with respect to the availability and usage of the eight ADAS. For the majority of ADAS (6 out of 8), more than two-thirds of the participants did not have the system within the vehicle they use, while less than a quarter of the participants had and used these systems (e.g., driver alert – sleepiness warning, lane keeping assistance etc.). Only with respect to two ADAS, cruise control (51%) and parking assist system (65%), the majority of participants

reported that they had the system in their vehicle and used it as well. Hence, no considerable influence of prior experience with ADAS are expected on drivers' evaluation of Mediator.

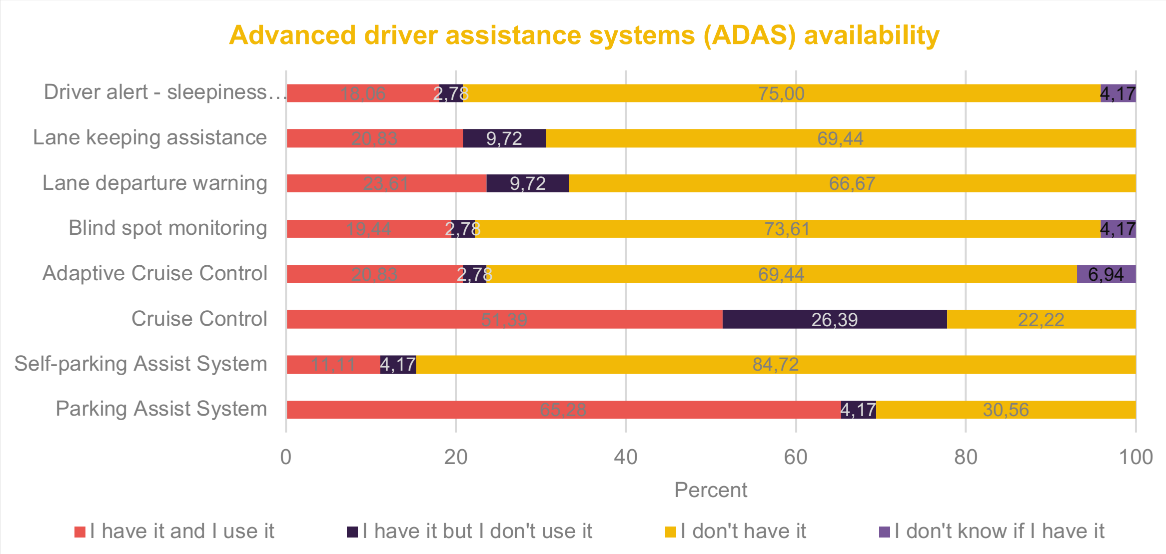


Figure 3.1 Availability of advanced driver assistance systems (ADAS).

3.4.1.3. Affinity for technology, attitude towards and acceptance of vehicle automation

The affinity for technology interaction (ATI) scale (Franke et al., 2017) revealed an average affinity for technology of $M = 4.45$ ($SD = .86$) on a 6-point Likert Scale. Compared to a reference sample (M = 3.58; SD = 1.09; Franke et al., 2019), the current sample is more affine towards technology compared to the public.

In addition to their affinity for technology, the participants were asked regarding their opinion about and their acceptance of vehicle automation in general. With respect to their general opinion, the participants were asked to answer the following question "What is your general opinion about functions in the vehicle that can automate parts or the entire driving task?" and their answers could range from "*1 – very negative*" to "5 – *positive*". On average ($M = 4.07$, $SD = .93$), participants seemed to be *somewhat positive* in their opinion towards functions that can automate parts or the entire driving task. Moreover, the participants were also asked to rate their general acceptance of vehicle automation. Here, the Van der Laan Acceptance scale (Van der Laan et al., 1997) was used. The participants' mean values for the total scale ($M = 0.98$, $SD = 0.55$) were within the positive range (>0 to 2), which relates to the positive attributes of the nine semantic differentials (i.e., desirable, likeable, assisting etc.). Hence, the participants tend to accept vehicle automation in general.

The relationships between the ATI scores, the general opinion about and acceptance of vehicle automation were analysed as well (Table 3.3). ATI, hence participants' affinity and interest in technology, correlated positively with participants' acceptance of vehicle automation (medium effect size). ATI also correlated positively with the general opinion about vehicle automation (medium effect size). Lastly, the acceptance of vehicle automation also correlated positively with the general opinion about vehicle automation (large effect size).

Table 3.3 Correlations of affinity for technology (ATI), general opinion about and acceptance of vehicle automation.

	ATI	General opinion about vehicle automation	Acceptance of vehicle automation
ATI	/	r = .22 p = .066	r = .34 p = .003
General opinion about vehicle automation	r = .22 p = .066	/	r = .80 p < .001
Acceptance of vehicle automation	r = .34 p = .003	r = .80 p < .001	/

To further assess participants' attitude towards vehicle automation in general, the SUaaVE questionnaire was used (Post et al., 2020). Participants rated several statements on a 7-point Likert Scale ranging from "*1 - completely disagree*" to "*7 – completely agree*". Figure 3.2 shows the means for the 12 items of the SUaaVE questionnaire. In addition, the items are color-coded to show the four subscales: acceptability (dark purple), trust (red), perceived convenience (gold) and perceived safety (purple). Participants seemed to be especially positive regarding the vehicle automation's acceptability. The subscale consisting of three items revealed an average score of M = 5.93 (SD = .95) indicating *agreement* to the statements. Hence, participants accepted the usage of automated vehicles as part of current and future traffic. The subscale averages for perceived convenience (M = 5.02, SD = 1.12) and perceived safety (M = 4.95, SD = 1.15) both indicated *slight agreement* to the respective statements. Hence, participants agreed that vehicle automation would be rather safe and convenient. With regard to the subscale trust, the average (M = 4.55, SD = 1.22) was the lowest. Especially regarding the statement "I trust the computer systems of automated vehicles cannot get hacked." participants slightly disagreed. Hence, participants trust the vehicle to behave as intended and correctly detect other road users, but they have concerns that the computer systems will be hacked. Nonetheless, the participants overall had a quite positive attitude towards vehicle automation.

Figure 3.2 SUaaVE questionnaire (scale ranging from 1 – completely disagree to 7 – completely agree) with the four colour-coded subscales: Acceptability (dark purple), trust (red), perceived convenience (gold) and perceived safety (purple). Error Bars: 95th-CI.

3.4.2. Material and apparatus

Additional detailed information regarding the hardware-setup as well as the integration of HMI components can be found in MEDIATOR deliverable D2.9 (Bakker et al., 2022).

3.4.2.1. Driving simulator

The study took place in a fixed-base driving simulator with a mock-up front vehicle section that included the front seats, seatbelts, a physical steering wheel, pedals, an automatic gear and two integrated screens, one located behind the steering wheel and one in the centre console. The simulation was projected on a curved screen with a 180° field of view and with additional projections for the back and side mirrors (see Figure 3.3). The simulation used the SILAB 7.0 software.

Figure 3.3 Driving simulator at Chemnitz University of Technology.

3.4.2.2. Mediator components

HMI module. The main focus of the study was on the drivers' perception of Mediator and its functionalities (e.g., detection of events related to driving comfort like upcoming traffic jams, comfort-related and safety-critical TOR, TOC from manual to automated driving and vice versa). The interface that displayed the status of Mediator as well as communicated with the driver played an important role. For this matter, a sophisticated HMI was implemented into the simulator. The HMI concept was designed by the colleagues from Delft University of Technology and adapted to the use cases, driving scenarios and spatial conditions of the driving simulator. The HMI concept included the following components (see Figure 3.4):

- A screen behind the steering wheel displaying driving relevant data (e.g., speedometer, driving mode, time to automation unfitness, available driving mode for the upcoming road section), instructions to the drivers (e.g., keep hands on steering wheel, TOR) and information about the context (e.g., detection of traffic jam)
- LED strips on the dashboard and the steering wheel as well as ambient lighting indicating the current driving mode, the intended or necessary changes in the driving mode as well as remaining time by, for instance, different colours, brightness and pulsation frequencies
- Sound system for sound alerts and spoken messages
- A laptop mounted in front of the centre console to run the NDRT

Figure 3.4 HMI concept for the TUC lab prototype in Mediator SB mode.

The different HMI components were triggered directly by specific events in the driving simulation (e.g., distance to the programmed traffic jam) or by drivers' actions (e.g., initiate takeover). The display was implemented in the software Labview (see Figure 3.5).

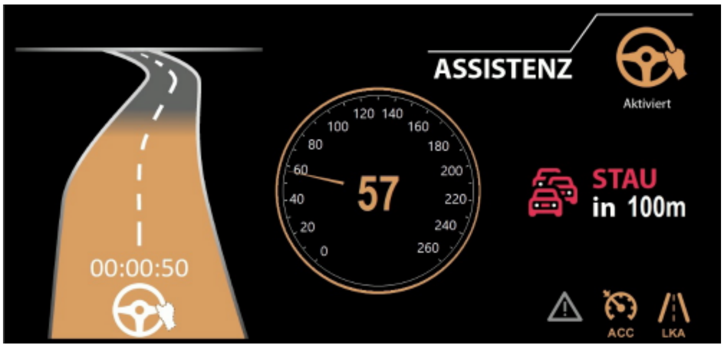


Figure 3.5 Example of the TUC lab prototype system in Mediator CM mode.

Driver monitoring module. The driver monitoring was relatively simple. Several cameras were implemented to record data for post-hoc analysis of driver comfort detection and driver distraction detection (i.e., eyes-off-road time). No online detection was implemented to avoid, for instance, unintended false alarms for the participant. Four cameras (FLIR, Blackfly S USB3) were installed:

- monochrome (NIR-sensitive) driver face camera for post-hoc facial action unit feature analysis for discomfort estimation as well as gaze estimation of where the driver is looking for distraction analysis
- upper body view camera for driver activity recognition
- outward-looking forward-facing camera
- over the shoulder camera mounted on the internal roof seeing the whole dashboard that the driver sees for post-hoc analysis of the HMI actions and the driver's reaction

Automation module and decision logic module. The automated driving (i.e., CM, SB) and degraded automation fitness (e.g., due to an upcoming highway exit), the outcomes of the driver monitoring module (e.g., degraded driving comfort) as well as the outcomes of the decision logic (e.g., TOR due to degraded automation fitness; proposal of takeover due to expected decrease of driving comfort) were simulated. The simulated actions of the decision logic module were based on predefined scenarios / tables (e.g., knowledge about surrounding traffic or incoming messages, knowledge about the moment when fitness of the automation system decreased due to an upcoming highway exit) and literature based expectations regarding comfort and discomfort investigated in WP1 (e.g., driving in a traffic jam). For more details, see section 4 in deliverable

D1.2 (Borowsky et al., 2020). The actions of the HMI components were triggered directly in the simulation.

3.4.2.3. Questionnaires

Table 3.4 provides an overview of the topics (e.g., driver characteristics, usability, trust, acceptance), the related research question, if applicable, the used questionnaires from the literature or comparable EU-projects as well as whether self-designed items were included or not. Moreover, a reference to the particular section in which the results will be discussed is included.

As can be seen, most of the questionnaires were based on established and standardized methods, including the Van der Laan Acceptance scale (Van der Laan et al., 1997), ATI scale (Franke et al., 2019), Trust in Automation scale (Jian et al., 2000) and the System Usability Scale (SUS; Brooke, 1996). In addition, questionnaires from recent, comparable EU-projects were used and adapted to MEDIATOR, such as L3 Pilot (Metz et al., 2019), ADAS&ME (Pereira Cocron et al., 2019), SUaaVE (Post et al., 2020) or AdaptIVe (Rodarius et al., 2015). Moreover, the standardized methods were enriched by self-designed items. These items were especially tailored to the purpose of the current simulator study and will be discussed in the corresponding sections. The used questionnaires are attached (see Appendix 5.2.4). Anonymized questionnaire raw data is publically available as open dataset at Zenodo under https://doi.org/10.5281/zenodo.7638299 (MEDIATOR, 2023).

The measurement timings of the questionnaires and self-designed items are described in

Table 3.5 and in more detail in the particular result sections as well. The questionnaires were always administered as a whole, only the results are later separated thematically, as depicted in Table 3.4. For instance, all acceptance-related items from the questionnaire adapted from L3 Pilot will be discussed in the corresponding section about acceptance, while the comfort-related items from the same questionnaire are discussed in the section about comfort.

Table 3.4 Questionnaires implemented in the driving simulator study.

Topic	RQ	Questionnaires from literature / previous projects	Self-designed items	Results discussed in section
Demographics	/	• Education level, employment status (L3 Pilot project; Metz et al., 2019)	• Items related to age, gender etc.	3.4.1.1; 5.2.2
Mobility / ADAS	/	• Car availability and usage (L3 Pilot project; Metz et al., 2019) • ADAS availability and usage (ADAS&Me: Pereira Cocron et al., 2019; L3 Pilot: Metz et al., 2019)	• Items related to annual mileage, license possession etc.	3.4.1.2; 5.2.2
Attitude towards vehicle automation	/	• General opinion about vehicle automation (ADAS&Me: Pereira Cocron et al., 2019) • SUaaVE pre-trial questionnaire (Post et al., 2020)	/	3.4.1.3
Affinity for Technology	/	• Affinity for technology Interaction (ATI; Franke et al., 2019)	/	3.4.1.3
Acceptance of Mediator	1	• Van der Laan Acceptance scale (Van der Laan et al., 1997)	/	3.5.1; 5.2.1

Topic	RQ	Questionnaires from literature / previous projects	Self-designed items	Results discussed in section
		• Questionnaire adapted from L3 Pilot (Metz et al., 2019)		
Intention to buy and to use Mediator	(1)	• Questionnaire adapted from L3 Pilot (Metz et al., 2019) • Comparison Mediator vs. no Mediator (adapted from L3 Pilot: Metz et al., 2019 / ADAPTIVE: Rodarius et al., 2015)	/	3.5.1
Comfort	2	• Comfort questionnaire on automated driving style (L3 Pilot: Metz et al., 2019) • Takeover questionnaire adapted from L3 Pilot (Metz et al., 2019) • Questionnaire adapted from L3 Pilot (Metz et al., 2019) • Comparison Mediator vs. no Mediator (adapted from L3 Pilot: Metz et al., 2019 / ADAPTIVE: Rodarius et al., 2015)	• Comfort-related items	3.5.2
Usability	3	• System Usability Scale (SUS; Brooke, 1996)	• Usability-related items	3.5.3
Trust	4	• Trust in Automation (Jian et al., 2000) • Questionnaire adapted from L3 Pilot (Metz et al., 2019)	/	3.5.4
Perceived safety	5	• Comparison Mediator vs. no Mediator (adapted from L3 Pilot: Metz et al., 2019 / ADAPTIVE: Rodarius et al., 2015) • Questionnaire adapted from L3 Pilot (Metz et al., 2019)	• Safety-related items	3.5.5
Comfort-related transfers from human to machine	6	• Questionnaire adapted from L3 Pilot (Metz et al., 2019)	• Comfort-transfer-related items	3.5.6
(Un-)planned transfers from machine to human	7	• Questionnaire adapted from L3 Pilot (Metz et al., 2019)	• (Un-)planned transfer-related items	3.5.7

Table 3.5 Overview of the questionnaires and corresponding measurement timings in the driving simulator study.

	Topics	Measurement timing
Pre-questionnaire	• Demographics • Mobility/ADAS • Attitude towards vehicle automation • Affinity for technology • Acceptance of automated driving	• After signing the informed consent, but before the first drive in the simulator
Post-drive 1 questionnaire	• Self-designed items regarding the traffic jam	• After the first drive, i.e. condition 1 (baseline)
Pre-drive 2 questionnaire	• Trust in Mediator (expected) • Acceptance of Mediator (expected)	• After receiving information about Mediator, but before the first drive with Mediator, i.e. condition 2 (CM)
Post-drive 2-4 questionnaire	• Comfort regarding automated driving style (only post-drive 2)	• After each of the three drives with Mediator, hence after condition 2 (CM),

	Topics	Measurement timing
	• Comfort-related transfers from manual to automated driving • (Un-) planned transfers from automated to manual driving • Trust in Mediator • Acceptance of Mediator • Self-designed items (usability, safety, acceptance etc.)	condition 3 (SB) and condition 4 (SB + close approach)
Post-questionnaire	• Acceptance of Mediator (across all three drives) • Trust in Mediator (across all three drives) • Usability • Expected differences when using Mediator vs. not using it	• After the four drives

3.4.2.4. Interviews

Several structured interviews were conducted throughout the study. Firstly, after each of the four drives, a short interview was conducted. Participants were asked whether they experienced uncomfortable or critical situations, when and how these situations occurred, and how they were solved either by themselves (condition 1) or by the system (condition 2 to 4).

Secondly, after one of the three drives with the Mediator system (condition 2 to 4), a slightly longer, more in-depth interview was conducted. For each participant, it was randomly assigned after which drive the interview took place. In order to not prolong the study time, this in-depth interview was not conducted after each drive for each participant. The participants were asked to respond to the following three questions:

- *"Mediator offered to take over the driving task for you while you were driving. Can you briefly explain when that was and why you think the proposal was sent by Mediator?"*
- *"While driving, Mediator asked you to take over the driving task again. Can you briefly explain when that was and why you think the request was sent by Mediator?"*
- *"Please think back to the first time you drove the entire route yourself. In a similar everyday situation, would you prefer driving with or without Mediator? Why?"*

After experiencing all four drives, a final, highly detailed interview was conducted. During this interview, the participants were asked with regard to

- their experience with Mediator, including (dis-)advantages of the system;
- how they felt about being driven through the traffic jam and about reading and answering emails during automated driving;
- other situations, in which they would like to have the support provided by Mediator;
- their trust in Mediator, including in what they trusted, why they trusted the system, what would help them to trust Mediator more and what situations might reduce trust in Mediator;
- their experience during the four drives, including a ranking of the drives from most comfortable to least comfortable
- the HMI concept:
    - including questions with regard to the different elements of the HMI concept, i.e. the light concept, clarity of the task distribution between system and driver, comprehensibility of messages, amount of information provided by Mediator

    - in-depth questions with respect to the information presented in the display behind the steering wheel, i.e. comprehensibility of the colour concept, comprehensibility of the time budget etc.
- their ideas on how to improve (or change) Mediator.

The German interview guidelines are attached to this deliverable (see Appendix 5.2.5).

3.4.3. Experimental design and procedure

The study design was a within subject design. All participants experienced all four conditions (in a fixed order) representing the different use cases and functionalities of Mediator. The fixed order was implemented to ensure that

- (1) participants could experience Mediator's functionalities gradually (manual driving without Mediator as baseline and a scenario participants are most familiar with; driving assistance, i.e. CM, introducing Mediator but requiring hands-on and monitoring; autopilot, i.e. SB, allowing for NDRT execution), and
- (2) participants would have the most positive experience possible with Mediator (a strength of the highly controlled simulator environment compared to on-road studies with the Mediator prototype) before encountering the close approach scenario, which was expected to represent an uncomfortable driving scenario (see section 4 in deliverable D1.2, Borowsky et al., 2020).

Interested potential participants filled out an online recruitment questionnaire examining the variables relevant for the selection of the final sample (i.e., age, gender, possession of drivers' license, annual mileage). Selected participants were contacted and received an appointment confirmation, the data privacy statement, data recording consent, informed consent as well as first information regarding the driving simulator study in advance via email.

Participants were welcomed at the TUC driving simulator, introduced to the study's purpose, procedure, hygiene concepts and had to fill in the necessary forms (i.e. the data privacy statement, data recording consent, informed consent). Afterwards, participants filled out the first questionnaire (pre-questionnaire, see

Table 3.5), asking for demographic data, mobility behaviour, experience with ADAS and relevant attitudes and personality traits.

In the main part of the study, participants experienced all four conditions in a fixed order (i.e., condition 1 to 4). Firstly, the participants experienced the baseline condition (i.e., condition 1), which was also used to familiarize the participants with the driving simulator. Afterwards, a short interview (see section 3.4.2.4) was executed, followed by a short questionnaire (i.e., post-drive1 questionnaire, see

Table 3.5). Afterwards, Mediator with all its components (HMI elements, decision logic, driver state detection and automated driving system) was explained to the participants, followed by a short questionnaire on the participants' expectations regarding Mediator (i.e., pre-drive2 questionnaire, see

Table 3.5). The participants then experienced the three conditions with the Mediator system (i.e., condition 2 to 4), each followed by a short interview (see section 3.4.2.4). After one of the three drives, an additional, more in-depth interview was conducted with the participants (see section 3.4.2.4). Additionally, after each of the three drives, participants filled out a questionnaire concerning, for instance, acceptance of Mediator, its functionalities, trust in Mediator etc. (i.e., post-drive2-4 questionnaire, see

Table 3.5). After all drives (including the interviews and questionnaires) were finished, participants filled out a final questionnaire to evaluate Mediator in general in terms of, for instance, trust, acceptance as well as the system's usability (i.e., post-questionnaire, see

Table 3.5). Afterwards, an in-depth interview was conducted to examine the drivers' experience of Mediator, its functions, and their opinion regarding the HMI etc. (see section 3.4.2.4). The study lasted approx. 2.5 h and participants received a compensation of 10 € / h (i.e. 25€).

3.4.4. Data logging and processing

All driving data (e.g., driving speed, lateral and vertical position, distances to other vehicles – 614 logged variables overall) were collected with a frequency of 60 Hz. The Labview code contained a data logging component where all Mediator HMI actions were recorded at 60Hz and saved to a csv file synchronized to the driving data from the simulator software. Video data (h265 format) were transcoded, synchronized and checked for completeness.

Datasets of all 74 participants were processed in the data analysis framework of TUC based on the relational database system PostgreSQL. For more details, see deliverable 2.9 (Bakker et al., 2022). Within the database framework, several standardized test scripts were executed to check, for instance, for delays in the simulation as well as completeness and plausibility of all data. Visual checks of the video data synchronized with all other data using NI Diadem made sure that all data were correctly synchronized over the full period of recording.

Questionnaires were implemented in Lime Survey 3. Participants' answers were logged and exported to SPSS 28 for further processing. Data were checked for outliers, standard scales were calculated according to the standard procedures and prerequisites were checked for all statistical analysis that were conducted.

Interview data were analysed using the inductive category development methodology according to Mayring (2004). First, all recorded answers were transcribed word by word. Afterwards, all of the participants' statements were coded by two independent coders, and a system of categories was developed based on the answers (bottom-up approach) and structured by the research questions (top-down approach). Over the course of the coding process, including regularly discussions between the two coders, the system of categories was refined until a sufficiently condensed categorical structure was obtained for describing how participants experienced Mediator. It needs to be taken into account that, although all participants were interviewed using the same guidelines, not all participants gave clear answers to all questions. Sometimes, they referred to other topics than the one they were asked about, and the conversation developed in another direction. Sometimes, they gave only vague answers that could not be categorized as agreement or disagreement to a certain statement. Therefore, the interview results are not complementary. Hence, when 20% of participants mentioned a certain HMI element as helpful it does not mean that the other 80% found it less helpful. These 80% just didn't mention the respective HMI element. If a considerable amount of participants evaluated a HMI element positively and another considerable amount of participants evaluated the HMI element negatively, both categories are mentioned.

Table 3.6 provides the interpretations of the effect sizes in terms of small, medium and large effect sizes of the different analytical methods, i.e. correlations, t-tests and repeated measure analyses of variance (rmANOVAs).

Table 3.6 Effect sizes and their interpretation.

	Person's correlation coefficient (*r*, Correlations, *t*-Test)	**Partial eta squared (η_p^2, rmANOVAs)**
Small effect	r = .10	η_p^2 = .01
Medium effect	r = .30	η_p^2 = .06
Large effect	r = .50	η_p^2 = .14

3.5. Results

3.5.1. RQ1: Acceptance of and intention to buy / use Mediator

To answer the first research question (RQ1) "*How is Mediator and its functionalities accepted by the drivers?*" questionnaire as well as interview answers were analysed.
Firstly, the Van der Laan Acceptance scale (Van der Laan et al., 1997) was used which was administered five times during the simulator study:

- (1) after participants received a detailed explanation of Mediator and its functionalities but before the first actual drive with the system (i.e., Drive 2) to assess participants' expected acceptance of Mediator,
- (2) once after each drive with Mediator (i.e., Drive 2-4) to assess participants' acceptance of Mediator in different use cases (e.g., CM vs. SB, close approach), and
- (3) in the post evaluation to assess participants' overall acceptance of Mediator after all drives.

Additionally, acceptance-related items from a questionnaire adapted from the L3 Pilot project (Metz et al., 2019) were analysed. This questionnaire will be referred to as L3 Pilot questionnaire in the following sections and was administered once in the post evaluation. In the following subsection, only the results regarding the items assessing the participants' acceptance will be discussed. Items relating to other subjects (i.e., trust or comfort) will be discussed in the respective sections.

In addition, participants' intention to buy and to use Mediator was identified as a particularly important subtopic of acceptance for the MEDIATOR project and further work packages. The results regarding respective items from the L3 Pilot questionnaire as well as from the questionnaire comparison Mediator vs. no Mediator adapted from the L3 Pilot and ADAPTIVE projects (Metz et al., 2019; Rodarius et al., 2015) were analysed. These results will be enriched by the results from the interviews.

Summary of results. On average, participants evaluated Mediator positively (*M* = 1.21 on a 5-point Likert scale ranging from -2 to +2). Based solely on the information provided about Mediator at the beginning, participants already indicated quite high acceptance of the system. Acceptance ratings increased after experiencing Mediator in actual driving scenarios but decreased again after drive 4, when participants experienced the close approach. Almost 88% of the participants would want to use Mediator in their car. About 70% of participants expect the possibility to drive longer trips, to drive when the driver is not capacitated and drive for more years, while even 90% expect to do other tasks when having Mediator in their car. No considerable effects of driver characteristics on acceptance were revealed. Higher affinity for technology and a more positive general opinion about driving automation is positively related to acceptance of and intention to use Mediator.

3.5.1.1. Questionnaire and interview results on acceptance and intention to buy / use

Van der Laan acceptance scale. Results of the total scale are shown in Figure 3.6 separately for all data collection points. Results regarding the two subscales satisfying and usefulness can be found in Appendix 5.2.1. Results showed that participants evaluated Mediator on average positively with regard to all five data collection points. A statistical comparison of all five data collection points revealed a significant effect, $F(2.32, 169.01) = 9.09$, $p < .001$, $\eta_p^2 = .11$. As can be seen in Figure 3.6, the expected acceptance (pre drive 2) was significantly lower ($M = 0.97$, $SD = 0.49$) than acceptance ratings after experiencing Mediator in actual driving scenarios, except for drive 4 when participants experienced the close approach ($M = 1.05$, $SD = 0.82$). Moreover, acceptance was significantly lower after drive 4 compared to drive 3 ($M = 1.27$, $SD = 0.6$; same condition without the close approach). The results show that based solely on the information provided about Mediator, participants already indicated quite high acceptance of the system. After actual, incident-free system experience, i.e. during drive 2 ($M = 1.24$, $SD = 0.53$) and drive 3, the acceptance increased considerably. However, compared to drive 3, participants' acceptance of Mediator decreased significantly after drive 4 with the close approach. The acceptance level even decreased to a similar level as that of the expected acceptance (pre drive 2). The post evaluation ($M = 1.21$, $SD = 0.58$) revealed that the overall acceptance rating is significantly higher compared to the rating after drive 4, almost reaching the same acceptance level as that of drives 2 and 3.

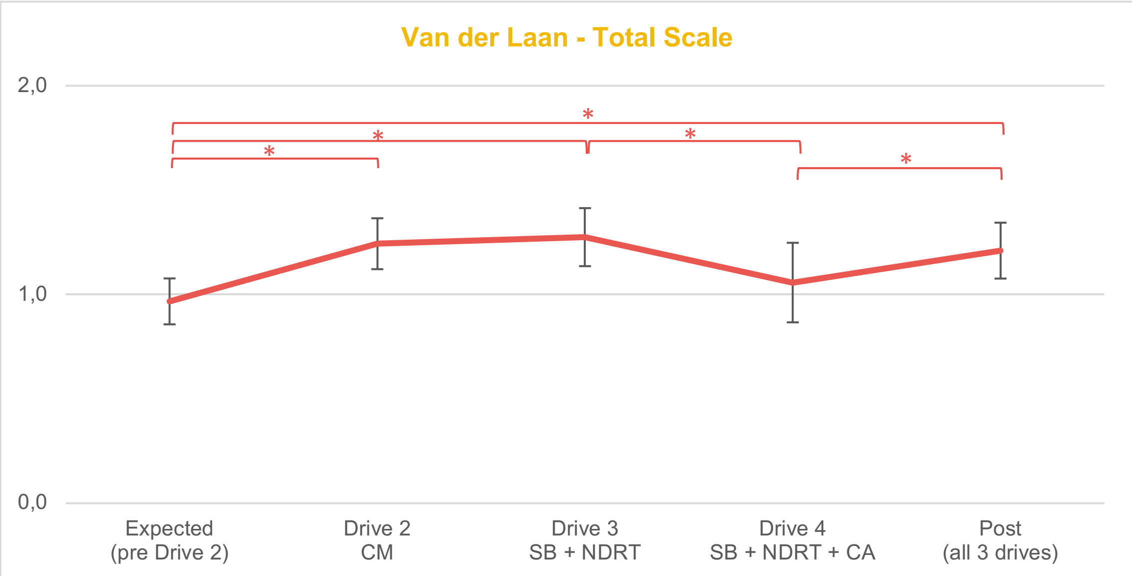


Figure 3.6 Acceptance of the Mediator System (Van der Laan – Total scale) across all five data collection points. Scale range: -2 to 2; Error Bars: 95th-CI.

In addition to the statistical comparison of the five data collection points, the mean values of the post evaluation were analysed separately for the nine individual semantic differentials of the Van der Laan scale (see Figure 3.7). The mean values for eight out of nine semantic differentials showed that the participants leaned strongly (mean values >1) towards the positive attributes (i.e., desirable, assisting, likeable etc.). Only with respect to the attributes *sleep inducing – raising alertness*, the participants evaluated Mediator considerably less positively (mean value <1), while still leaning more towards *useful* (mean values >0). With regard to the other data collection points, a similar picture arose.

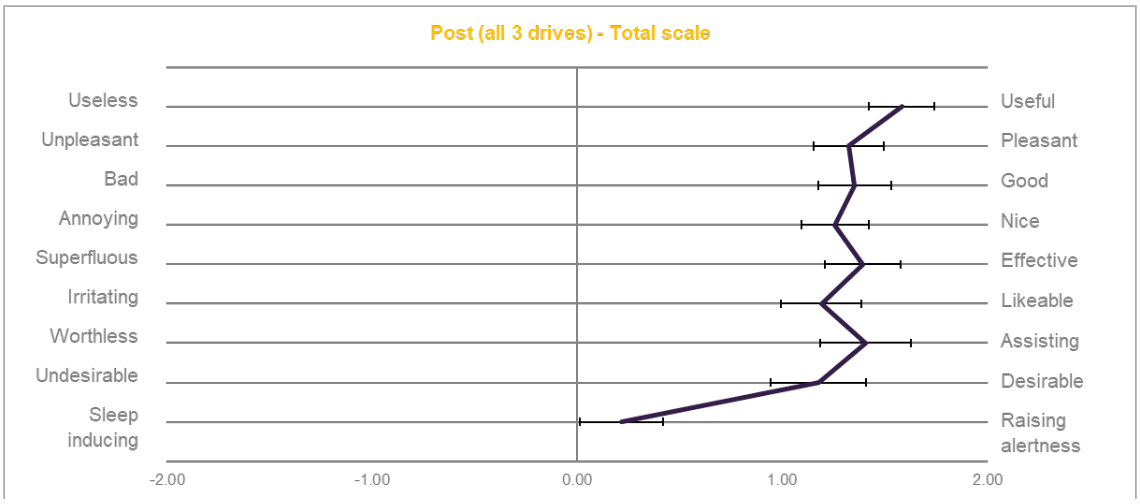


Figure 3.7 Acceptance of the Mediator System (Van der Laan – Total scale) during the post evaluation. Scale range: -2 to 2; Error Bars: 95th-CI.

L3 Pilot questionnaire (acceptance). Figure 3.8 presents the participants' mean values for all acceptance-related items of the L3 Pilot questionnaire. Answers could range from "*1 - totally disagree*" to "*5 – totally agree*", with "*3 – neutral*". Moreover, Table 3.7 provides the mean and standard deviation, as well as the share of (dis-)agreement with regard to the acceptance-related items.

Participants tended to *agree* (M = 3.88, SD = 1.10) with the statement "Mediator worked as it should work". More than 75% of the participants indicated agreement, while less than 15% disagreed. Regarding the statement "Mediator acted appropriately in all situations", participants' answers were between the two categories *neutral* and *agree* (M = 3.53, SD = 1.23). While a little more than 66% of the participants agreed that Mediator acted appropriately in all situations, more than a quarter of the participants did not agree. The statement "Sometimes Mediator behaved unexpectedly" reached even more divergence between the participants. While the mean values showed slight *disagreement* with a strong tendency towards *neutral* answers (M = 2.86, SD = 1.36), 50% of the participants disagreed and 40% agreed with the statement. Hence, while half of the participants felt that Mediator did not behave unexpectedly, 40% actually did feel that Mediator behaved unexpectedly in certain situations. Lastly, participants tended towards *agreement* regarding the last statement "I would use the time Mediator was active to do other activities" (M = 2.86, SD = 1.36). Here, more than 56% agreed with the statement, hence would potentially do other activities, while less than 33% participants were not as keen to do other activities with Mediator active.

In line with the results of the Van der Laan scale, the participants evaluated acceptance of Mediator quite positively again. While the vast majority of the participants indicated that Mediator worked as it should do, a quite large percentage of the participants (40%) also felt that Mediator behaved unexpectedly. The latter aspect might be explainable through the close approach in drive 4. Since Mediator and the automated system worked incident-free in drive 2 and 3, the close approach might have seemed as unpredictable behaviour on the account of the vehicle automation but also of Mediator as it did not inform the participants. Since a reduction in acceptance could reduce drivers' willingness to use Mediator, it is necessary to design Mediator and the vehicle automation in a way that decreases unexpected behaviour as much as possible.

Table 3.7 Acceptance-related items adapted from L3 Pilot questionnaire: Average (M; range from 1 – totally disagree to 5 – totally agree) and standard deviation (SD), share of (dis-) agreement for each statement.

Statement	*M* (*SD*)	Disagreement (%)	Neutral (%)	Agreement (%)
Mediator worked as it should work.	3.88 (1.10)	14.86	9.46	75.68
Mediator acted appropriately in all situations.	3.53 (1.23)	25.68	8.11	66.22
Sometimes Mediator behaved unexpectedly.	2.86 (1.36)	50.00	9.46	40.54
I would use the time Mediator was active to do other activities.	3.41 (1.28)	32.43	10.81	56.76

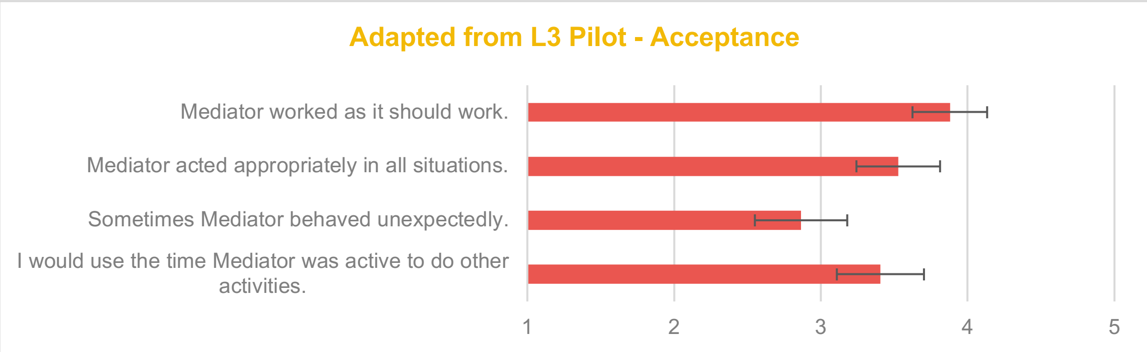


Figure 3.8 Acceptance-related items adapted from L3 Pilot questionnaire. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

L3 Pilot questionnaire (intention to buy / use). Figure 3.9 presents participants' mean values for the items related to the intention to buy and use Mediator. Answers could range from "*1 - totally disagree*" to "*5 – totally agree*", with "*3 – neutral*". Additionally, Table 3.8 provides the mean and standard deviation, as well as the share of (dis-)agreement with regard to these items. Regarding the item "I would buy Mediator", the mean value indicated a somewhat *neutral* opinion with a tendency towards *agreement* (*M* = 3.54, *SD* = 1.08). A little more than 63% indicated that they would potentially buy Mediator, while slightly less than 19% would not. Regarding the item "I would recommend Mediator to others", participants were on average a little more *agreeable* (*M* = 3.95, *SD* = 0.94). A little more than 66% indicated they would recommend Mediator to others, while a little more than a quarter of the participants would not. Participants indicated *neutral* opinions with a tendency towards *agreement* with respect to the two statements focusing on the costs (*M* = 3.39, *SD* = 1.17) and benefits from Mediator (*M* = 3.66, *SD* = 1.04). A little more than 47% of the participants indicated that the costs would be the most important aspect to consider for purchasing Mediator, while the rest of the participants were almost equally distributed across disagreement (i.e., almost 26%) or undecided (i.e., >27%). However, a larger group of participants (i.e., >62%) indicated that the benefits of Mediator would be the most important aspect to consider when buying Mediator, while only less than 15% disagreed with the statement. Hence, a larger portion of the participants seemed to think that the benefits of Mediator are more important than the costs, when thinking about purchasing it.

Participants *agreed* on average with the statement "I would use Mediator if it was in my car" (*M* = 4.20, *SD* = 0.97). Almost 88% of the participants would want to use Mediator, while less than

11% would not want to use it. When asked about using Mediator on everyday trips, participants were slightly less *agreeable* (*M* = 3.73, *SD* = 1.14). Nonetheless, almost 65% agreed with the statement, while a little less than 19% disagreed. Hence, a grant majority of the participants would want to use Mediator if it was in their car and a little less than two thirds of the participants would want to use it for their everyday trips. In contrast, participants indicated *disagreement* up to a *neutral* opinion towards the statement “I would make MORE trips if I had Mediator in my car” (*M* = 2.26, *SD* = 1.15). Almost 57% indicated disagreement, while less than 15% agreed with the statement. Hence, the majority of the participants was not inclined to make more trips with Mediator in their car. Similarly, participants indicated on average *disagreement* up to an almost *neutral* opinion towards the statement “I would select destinations further away if I had Mediator in my car” (*M* = 2.73, *SD* = 1.26). The distribution was more equally distributed. Only a little more than 40% of the participants disagreed, while a little more than 31% agreed with the statement and the rest was undecided. Hence, only a relatively small majority of the participants would not choose further destinations. A possible explanation might be that participants already do all necessary trips with their car (in terms of trip number and distances to the destination) and, hence, Mediator would not make a big difference.

Table 3.8 Buy and usage intention-related items adapted from L3 Pilot questionnaire: Average (M; range from 1 – totally disagree to 5 – totally agree) and standard deviation (SD), share of (dis-) agreement for each statement.

Statement	*M* (*SD*)	Disagreement (%)	Neutral (%)	Agreement (%)
I would buy Mediator.	3.54 (1.08)	18.92	17.57	63.51
I would recommend Mediator to others.	3.95 (0.94)	25.68	8.11	66.22
The cost of Mediator would be the most important thing I would consider before purchasing one.	3.39 (1.17)	25.68	27.03	47.30
The benefits of Mediator would be the most important thing I would consider before purchasing one.	3.66 (1.04)	14.86	22.97	62.16
I would use Mediator if it was in my car.	4.20 (0.97)	10.81	1.35	87.84
I would use Mediator during my everyday trips.	3.73 (1.14)	18.92	16.22	64.86
I would make MORE trips if I had Mediator in my car.	2.26 (1.15)	56.76	28.38	14.86
I would select destinations further away if I had Mediator in my car.	2.73 (1.26)	40.54	28.38	31.08

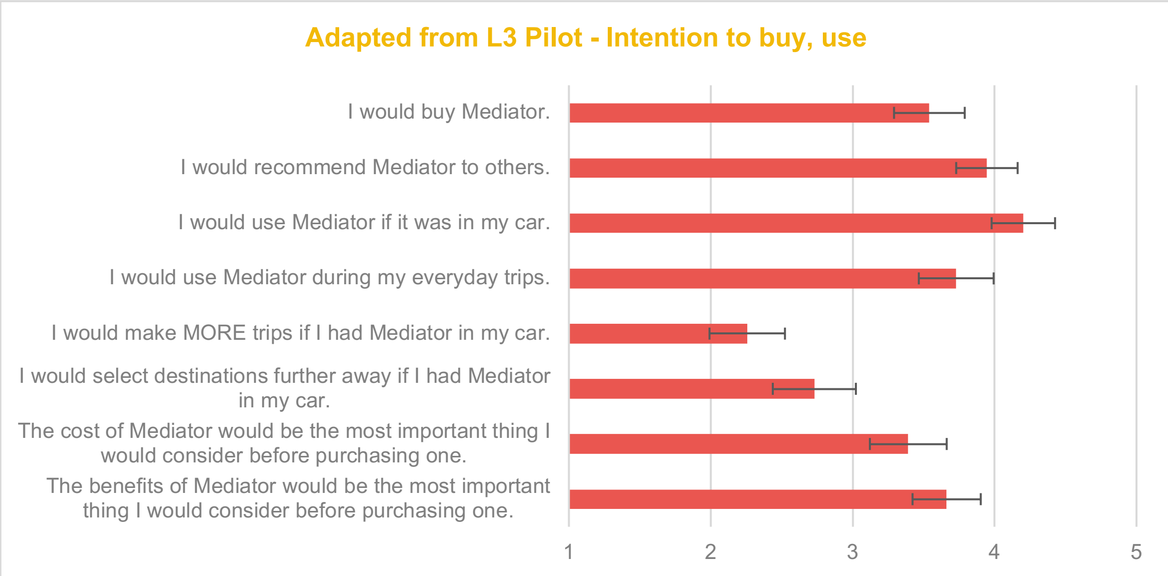


Figure 3.9 Purchase and usage-related items adapted from L3 Pilot questionnaire. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

Comparison Mediator vs. no Mediator. Table 3.9 provides the results of the usage-related items. The answers could range from "*1 – strong decrease*" to "*5 – strong increase*". Participants expected on average a *slight increase* (M = 3.92, SD = 0.81) in the "possibility to drive longer (longer trips)". Almost 72% expected an increase, while a quarter of the participants expected no differences at all. Participants expected almost exactly the same *slight increase* (M = 3.84, SD = 0.94) regarding the "possibility to drive longer (more years)". Again, a little more than 70% expected an increase, while a quarter of the participants expected no differences at all. Moreover, participants also expected on average a *slight increase* in the "possibility to drive when driver is not fully capacitated" (M = 3.70, SD = 1.06). Two thirds of the participants expected an increase, while a quarter of the participants did not expect any difference. Lastly, the participants expected on average a *slight increase* (M = 4.24, SD = 0.81) regarding the "possibility to do other tasks". More than 93% of the participants expected an increase in their possibility to do other tasks while driving with Mediator. Hence, participants seemed to see the potential of using Mediator to do other tasks with Mediator active, to drive longer either in terms of longer trips or in terms of driving at advanced age as well as to drive when not fully capacitated.

Table 3.9 Usage-related items of the questionnaire comparing Mediator vs. no Mediator [adapted from L3 Pilot / ADAPTIVE]. Average (M; range from "1 – great decrease" to "5 – great increase") and standard deviation (SD).

What differences do you expect when using the Mediator compared to driving without Mediator?	*M (SD)*	Decrease (%)	No difference (%)	Increase (%)
Possibility to drive longer (longer trips)	3.92 (0.81)	4.05	24.32	71.62
Possibility to drive longer (more years)	3.84 (0.94)	5.41	24.32	70.27
Possibility to drive when driver is not fully capacitated	3.70 (1.06)	8.11	25.68	66.22
Possibility to do other tasks	4.24 (0.81)	4.05	2.70	93.24

Results from the interviews. During the interviews, participants were asked if they would like to use Mediator in their future cars. Eighty-six percent of all participants expressed their willingness to use Mediator in the future underlining the questionnaire results. Eighteen percent of these participants justified their decision with expected improvements in driving comfort and driving safety as well as in a simplification of the driving task. Additionally, some of these participants (11%) mentioned mobility even in higher age or illness, possible time savings and the possibility to do other tasks as reasons why they would use Mediator. Nevertheless, participants mentioned certain prerequisites that need to be considered. Thirty-six percent of all participants mentioned the final price as important aspect that might influence their decision to purchase Mediator. Fifteen percent stated that reliability of the technology, sufficient safety and transparency regarding possible mistakes by the system need to be ensured. Sixteen percent will need more test drives and more experience with Mediator before they want to use the system in their future car. Eight percent of all drivers stated that they want to be in control all the time, i.e., decide if they want to drive manually or in automated driving mode. Seven percent mentioned that they would prefer using Mediator only or mostly on highways. Eight percent of all drivers stated that they won't use Mediator in the future, for instance, because they are happy with manual driving or they expect to not have the possibility of such a system in middle class vehicles.

When asked for specific use cases in which they want to have Mediator, participants mentioned long and monotonous trips (31%), highway driving (19%), city traffic (16%), traffic jams and construction zones (19%), bad weather conditions (11%) and for commuters or professional drivers (8%). Sixteen percent of the participants mentioned use cases related to NDRTs like reading, talking with other passengers, playing with children or talking on the phone.

3.5.1.2. Influence of driver characteristics

Van der Laan Acceptance scale. The influence of various driver characteristics were analysed by means of a rmANOVA which included the participants' age, gender, driving experience (i.e. annual mileage) and their affinity for technology (ATI score). No significant main effects of the driver characteristics or interaction effects of the driver characteristics with each other or the five data collection points could be found (see Appendix 5.2.3.1). Hence, participants' age, gender, driving experience and affinity for technology did not seem to influence their acceptance of Mediator.

Additionally, the relationship between the five data collection points of the Van der Laan acceptance scale, the participants' ATI scores and their general opinion about vehicle automation were analyzed (Table 3.10). The participants' ATI scores correlated positively with the participants' acceptance ratings in all five data collection points, indicating that a higher ATI score was related to higher acceptance ratings. All correlations, with exception for drive 4, were significant with a small (i.e. drive 3 and post) or medium effect size (i.e. pre drive 2 and drive 2).

The participants' general opinion about vehicle automation also correlated positively with the five data collection points of the Van der Laan acceptance scale. All correlations were significant and had medium (i.e. drive 2 and drive 4) or even large effect sizes (i.e. pre drive 2, drive 3, and post). Hence, the more positive the participants' opinion about automating parts or the entire driving task was, the higher their acceptance ratings of Mediator throughout the entire experiment was.

Table 3.10 Correlations between the Van der Laan Acceptance scale and the affinity for technology (ATI) as well as the general opinion about vehicle automation.

	Expected (pre drive 2)	Drive 2 CM	Drive 3 SB + NDRT	Drive 4 SB + NDRT + close approach	Post (all three drives)
ATI	*r* = .34 *p* = .003	*r* = .35 *p* = .002	*r* = .28 *p* = .018	*r* = .15 *p* = .210	*r* = .27 *p* = .019
General opinion about vehicle automation	*r* = .50 *p* < .001	*r* = .46 *p* < .001	*r* = .61 *p* < .001	*r* = .47 *p* < .001	*r* = .52 *p* < .001

L3 Pilot questionnaire (acceptance). For this questionnaire, the influences of age, gender and driving experience (i.e. annual mileage) were analysed descriptively. In total, the differences were relatively small. For instance, the participants' gender had no noticeable, systematic influence on the participants' answer to the four acceptance-related items. Comparisons of the four age groups indicated differences that were a little more pronounced but still relatively small. A tendency can be seen that with increasing age, the behaviour of Mediator seemed to be experienced less unexpected and more appropriate. With respect to the participants' driving experience (i.e. annual mileage), the differences were also relatively small. However, the group with the very high annual mileage (17,600-60,000km) were most inclined to use the time Mediator was active to do other activities compared to the other three groups, who were more neutral.

See Appendix 5.2.3.1 for more detailed descriptions of the influences of age, gender and driving experience (i.e. annual mileage).

In addition, the participants' answers regarding the acceptance-related items were also correlated with their ATI scores as well as their general opinion (see Table 3.11). Only two significant correlations were found. Firstly, a positive correlation with a small effect size was found for the ATI and the item "I would use the time Mediator was active to do other activities". Hence, participants with a higher ATI score seemed to be more inclined to do other activities when Mediator is active, than participants with lower ATI scores. Moreover, a negative correlation with a small effect size was found between the participants' general opinion about driving automation and the item "Sometimes Mediator behaved unexpectedly". Hence, a more positive opinion towards driving automation seemed to be related to experiencing Mediator's behaviour as more confirm with the participants' expectations (i.e. less unexpected behaviour).

Table 3.11 Correlations between acceptance-related items and the participants' ATI score and general opinion about driving automation.

Statement	ATI	General opinion about driving automation
Mediator worked as it should work.	*r* = - .05 *p* = .672	*r* = .12 *p* = .325
Mediator acted appropriately in all situations.	*r* = - .07 *p* = .561	*r* = .03 *p* = .810
Sometimes Mediator behaved unexpectedly.	*r* = .00 *p* = .999	*r* = - .27 *p* = .017
I would use the time Mediator was active to do other activities.	*r* = .25 *p* = .033	*r* = .20 *p* = .094

L3 Pilot questionnaire (intention to buy / use). The following driver characteristics were analysed descriptively: Age, gender and driving experience. Similar as to the acceptance-related items, the influence of these driver characteristics on the different questionnaire items were relatively small. For instance, both female and male participants gave quite similar answers to the eight items. Comparing the four age groups, a few, more considerable differences were found. For instance, a trend was found that the older the participants are, the more they seemed to want to use Mediator during their everyday trips. In addition, the younger age groups seemed to disagree more with regard to making more trips with Mediator compared to the older age groups, who were more neutral. Regarding the driving experience (i.e. annual mileage), the different groups also showed relatively similar answers. See Appendix 5.2.3.1 for more detailed descriptions of the influences of age, gender and driving experience (i.e. annual mileage) and corresponding figures.

In addition, the participants' answers regarding the usage- and purchase-related items were also correlated with their ATI scores as well as their general opinion (see Table 3.12). Regarding the ATI scores, three significant, positive correlations were found. These indicated that the higher the ATI scores, the more likely participants would buy Mediator (medium effect size), would recommend it to others (small effect size), and would use Mediator during their everyday trips (medium effect size).

Moreover, the participants' general opinion was significantly and positively correlated to seven out of eight items. The more positive the general opinion about driving automation, the more likely participants would buy Mediator (medium effect size), recommend it (medium effect size), use if it was in the car (large effect size), use it during everyday trips (large effect size), make more trips if it was in the car (large effect size) and select further destinations (medium effect size). Moreover, the more positive their general opinion, the more important were the benefits of Mediator for the participants.

Table 3.12 Correlations between items related to use/buy adapted from L3 Pilot questionnaire and the participants' ATI score and general opinion about driving automation.

Statement	ATI	General opinion about driving automation
I would buy Mediator.	r = .34 p = .003	r = .44 p < .001
I would recommend Mediator to others.	r = .23 p = .050	r = .37 p = .001
I would use Mediator if it was in my car.	r = .21 p = .076	r = .54 p < .001
I would use Mediator during my everyday trips.	r = .32 p = .006	r = .54 p < .001
I would make MORE trips if I had Mediator in my car.	r = .08 p = .49	r = .54 p < .001
I would select destinations further away if I had Mediator in my car.	r = - .15 p = .206	r = .29 p = .014
The cost of Mediator would be the most important thing I would consider before purchasing one.	r = - .034 p = .773	r = .10 p = .388
The benefits of Mediator would be the most important thing I would consider before purchasing one.	r = .20 p = .086	r = .31 p = .007

Comparison Mediator vs. no Mediator. The driver characteristics age, gender and driving experience were also analysed descriptively for this questionnaire. The differences between female and male participants were relatively small. The comparison of the four age groups revealed somewhat greater differences. For instance, with increasing age, a greater increase in the “possibility to drive longer (more years)” was expected. The oldest age group expected the highest increase compared to the youngest age group. Lastly, the four driving experience groups were compared as well. The four groups had no noticeable, systematic influence on participants’ answers.

See Appendix 5.2.3.1 for more detailed descriptions of the influences of age, gender and driving experience (i.e. annual mileage) and corresponding figures.

In addition, the participants’ answers were also correlated with their ATI scores as well as their general opinion (see Table 3.13). The negative correlations between the ATI scores and the four items were not significant and all had a small effect size. The participants’ general opinion about driving automation revealed only one significant, positive correlation with the “possibility to drive longer (more years)” (medium effect size). Here, the more positive the general opinion about driving automation was, the higher the expected increase in the participants’ possibility to be able to drive longer (i.e. with greater age) was.

Table 3.13 Correlations between items related to use/buy from Comparison Mediator vs. no Mediator questionnaire and the participants’ ATI score and general opinion about driving automation.

Statement	ATI	General opinion about driving automation
Possibility to drive longer (longer trips)	r = - .11 p = .351	r = - .08 p = .476
Possibility to drive longer (more years)	r = - .1 p = .404	r = .38 p = .001
Possibility to drive when driver is not fully capacitated.	r = - .03 p = .823	r = .12 p = .313
Possibility to do other tasks	r = - .05 p = .652	r = .07 p = .558

3.5.2. RQ2: Comfort

To answer the second research question (RQ2) “*How is Mediator’s potential to increase drivers’ comfort?*”, again, questionnaires as well as interview questions were used. Firstly, a comfort questionnaire on automated driving style was administered, which originated from the L3 Pilot project (Metz et al., 2019). This questionnaire asked for participants’ satisfaction with the driving style of the automated vehicle. It can be expected that participants will not differentiate between Mediator and the automated driving system implemented in the simulator. Hence, results can also give indications regarding participants’ comfort rating with respect to Mediator. The questionnaire was applied once after the first drive with Mediator, hence after drive 2 with CM. In addition, a few comfort-related items from the questionnaire comparison Mediator vs. no Mediator adapted from the L3 Pilot and ADAPTIVE projects (Metz et al., 2019; Rodarius et al., 2015) were used. Moreover, three questionnaire items were designed especially for the current driving simulator study that assessed the participants’ comfort during the drives with Mediator. Two of these items were administered after each of the three drives with Mediator (i.e. drive 2-4). However, the third item, which focused on NDRT execution, was only administered after the two drives with SB (i.e., drive 3 and 4), during which the participants were asked to execute a NDRT (i.e., reading and answering an email). Lastly, the questionnaire adapted from the L3 Pilot project, already mentioned in the previous sections, also included a few items focusing on comfort. Finally, participants' comfort was also evaluated through various questions within the interviews executed throughout the simulator study.

Summary of results. On average, participants evaluated the automated driving style positively (*M* = 4.22 on a 5-point Likert scale ranging from 1 to 5). More than 90% of the participants agreed that driving with Mediator active was comfortable. Participants valued Mediator's support for the traffic jam scenarios (*M* > 5 on a 6-point Likert scale ranging from 1 to 6) with slightly lower values after experiencing the close approach. The possibility to perform NDRTs while driving was rated a bit less positively (*M* ~ 4 on a 6-point Likert scale ranging from 1 to 6). Participants expected higher comfort and possibility to relax/rest, as well as reduced stress and irritation when using Mediator, but they also expect less enjoyment of driving, more boredom and fatigue.
No considerable effects of driver characteristics on driving comfort were revealed.

3.5.2.1. Questionnaire and interview results on comfort

Comfort questionnaire on automated driving style. Participants could indicate how (un-) comfortable the behaviour of the automated vehicle made them feel ranging from "*1 – very uncomfortable*" via "3 – *neutral*" to "*5 – very comfortable*". The mean values of all six items ranged from *M* = 3.76 to *M* = 4.66 and where thus in the positive answer range (Figure 3.10). This indicates that the behaviour of the automated vehicle made the participants feel (almost) *somewhat comfortable* to (almost) *very comfortable*. The participants evaluated the automated vehicle's "smoothness of driving" the most positively (*M* = 4.66, *SD* = .58), indicating an (almost) *very comfortable* experience. This was followed by the "distance kept to the road markings" (*M* = 4.43, *SD* = 0.8) and the "distance kept to the vehicle in front" (*M* = 4.27, *SD* = 1.04), indicating a more than *somewhat comfortable* experience. Only with regard to the "braking behaviour of the vehicle", the mean value was lower (*M* = 3.76, *SD* = 1.21), indicating a slightly less than *somewhat comfortable* experience. Since the questionnaire was administered after drive 2 with CM, the evaluation of the braking behaviour was not influenced by the close approach of drive 4.

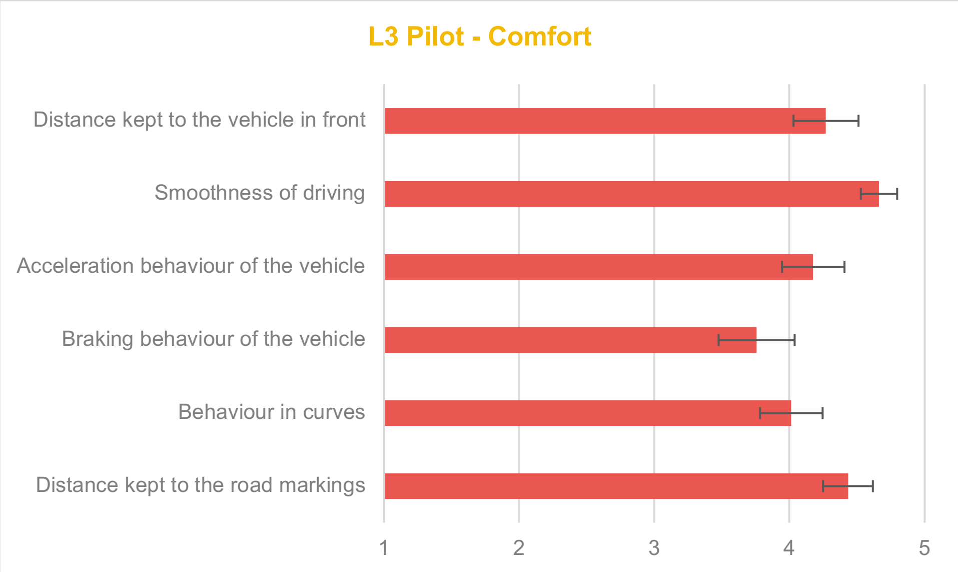


Figure 3.10 Comfort questionnaire on automated driving style. Scale range: 1 – very uncomfortable, 5 – very comfortable; Error Bars: 95th-CI.

Comfort-related items (self-designed). Figure 3.11 shows the results with regard to the three self-designed comfort items. Participants could indicate their agreement with the items from "*1 – completely disagree*" to "*6 – completely agree*". Mean values for the statement "It was pleasant to be able to use Mediator during the traffic jam" ranged from *M* = 4.88 to *M* = 5.39, indicating *somewhat* to *mostly agreement*. The statistical comparison revealed a significant effect ($F(1.68, 122.75) = 8.75$, $p < .001$, $\eta_p^2 = .11$). As can be seen in Figure 3.11, participants' agreement with the statement was significantly higher after drive 2 with CM (*M* = 5.39, *SD* = .81) and drive 3 with SB and the NDRT (*M* = 5.35, *SD* = .1) compared to drive 4 with SB, the NDRT and the close approach (*M* = 4.88, *SD* = 1.41). Hence, the usage of Mediator felt less pleasant within the traffic jam of drive 4 compared to the other two drives. It seems likely that participants felt less safe and comfortable when Mediator drove them through the traffic jam during drive 4, since they just experienced the rather unpleasant close approach of the end of the traffic jam. Moreover, mean values for the statement "Thanks to Mediator, I didn't mind driving through the traffic jam" ranged from *M* = 4.72 to *M* = 5.23, indicating *somewhat* to *mostly agreement*. The statistical comparison revealed a significant effect as well ($F(1.63, 119.19) = 8.02$, $p < .001$, $\eta_p^2 = .10$). As can be seen in Figure 3.11, participants' agreement with the statement was significantly higher after drive 3 (*M* = 5.23, *SD* = 1.0) compared to drive 4 (*M* = 4.72, *SD* = 1.33). Similar to the previous item, participants seemed to mind it more to drive through the traffic jam with Mediator active during drive 4 than during drive 3. Again, the close approach could have negatively influenced their experience. Agreement with the statement after drive 2 (*M* = 5.03, *SD* = 1.0) was in the middle and did not differ significantly from the other two drives. Lastly, the item "I liked having the option to perform a NDRT (such as answering messages) while driving" was only administered after drive 3 and 4, where the participants were asked to execute the NDRT. For both drives, participants *somewhat agreed* with the statement. The statistical comparison of the two drives, did not reveal significant differences ($t(73) = .731$, $p = .233$, $r = .085$) between drive 3 (*M* = 4.15, *SD* = 1.45) and drive 4 (*M* = 4.05, *SD* = 1.51). Hence, similarly for both drives participants seemed to like that they could perform a NDRT while automated driving was active.

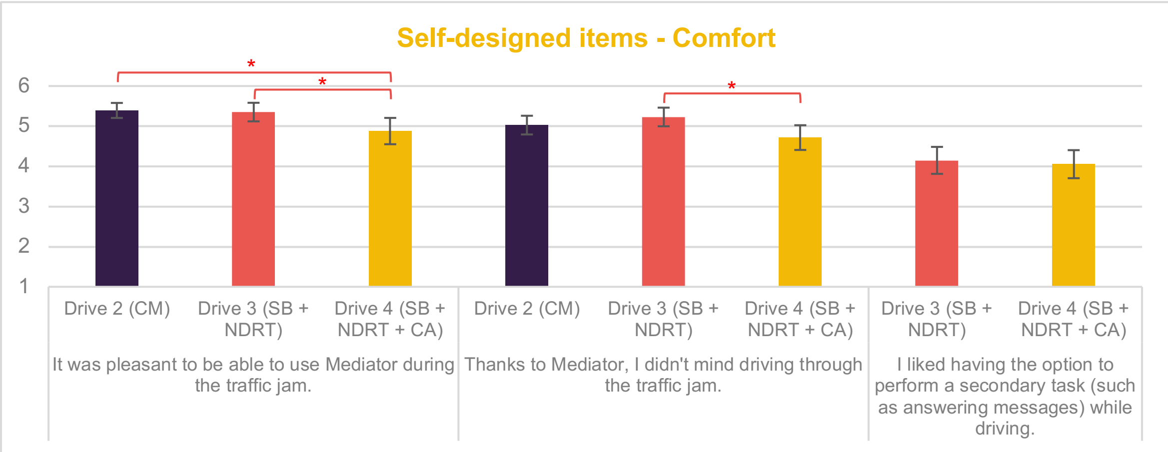


Figure 3.11 Comfort-related items (self-designed). Scale range: "1 – completely disagree", "6 – completely agree"; Error Bars: 95th-CI.

L3 Pilot questionnaire (comfort). Table 3.14 provides the mean and standard deviation (range from "*1 - totally disagree*" to "*5 – totally agree*"), as well as the share of (dis-)agreement with regard to the comfort-related items. Overall, participants evaluated the comfort during the drives with Mediator quite positively, as they indicated *agreement* (*M* = 4.22, *SD* = .75) with the statement "Driving with Mediator active was comfortable*"*. Actually, more than 90% of the participants agreed with this statement, whereas less than 3% disagreed. Correspondingly, participants also indicated *disagreement* with the statements that "Driving with Mediator was difficult" (*M* = 1.49, *SD* = .73), "(…) demanding" (*M* = 1.99, *SD* = 1.1) and "(…) stressful" (*M* = 1.95, *SD* = .93). The vast majority of the participants (almost 95%) experienced driving with Mediator as not being difficult. With respect to whether driving with Mediator was demanding and stressful, almost 73% of the participants each disagreed with the statement, while less than 15% of the participants indicated that driving with Mediator was actually demanding, and a little more than 8% indicated that it was stressful. Lastly, participants' mean values for the item "Driving with Mediator on long journeys would make me tired" indicated opinions that were more *neutral* (*M* = 3.16, *SD* = 1.14). Almost 42% agreed with the statement, while a little more than 31% disagreed. Overall, the results indicate that participants experienced driving with Mediator as quite comfortable, easy and not particularly demanding or stressful.

Table 3.14 Comfort-related items adapted from L3 Pilot questionnaire: Average (M; range from 1 – totally disagree to 5 – totally agree) and standard deviation (SD), share of (dis-) agreement with each statement.

Statement	*M (SD)*	Disagreement (%)	Neutral (%)	Agreement (%)
Driving with Mediator active was comfortable	4.22 (.75)	2.70	6.76	90.54
Driving with Mediator was difficult.	1.49 (.73)	94.59	1.35	4.05
Driving with Mediator was demanding.	1.99 (1.1)	72.97	12.16	14.86
Driving with Mediator was stressful.	1.95 (.96)	72.97	18.92	8.11
Driving with Mediator on long journeys would make me tired.	3.16 (1.14)	31.08	27.03	41.89

Comparison Mediator vs. no Mediator. Table 3.15 shows the results for the comfort-related items. The answers could range from "*1 – strong decrease*" to "*5 – strong increase*". Firstly, participants indicated a *slight increase* with a tendency towards a *great increase* (*M* = 4.55, *SD* = 1.58) regarding their “comfort” when driving with Mediator. This increase was expected almost unanimously by more than 98% of the participants. In line with this, participants also expected a *slight increase* regarding their “possibility to relax/rest” (*M* = 3.88, *SD* = 0.76). Almost 73% of the participants expected an increase, while only a little more than 4% expected a decrease and the remaining ca. 23% expecting no difference at all. However, while the participants expect their comfort and possibility to relax or rest to increase, they also expected that their “boredom” might increase (*M* = 3.41, *SD* = 1.01). Almost 53% of the participants expected an increase, while a third expected no differences and almost 15% expected a decrease. In line with this, a little more than 40% of the participants expected a decrease in “enjoyment of driving”, while only a little more than 36% expected an increase. The mean values indicated *no difference* (*M* = 2.95, *SD* = 1.28). Regarding their “stress” participants expected on average a *slight decrease* (*M* = 2.2, *SD* = 0.92). More than two thirds of the participants expected a stress decrease, while a little more than 25% expected no difference and less than 7% expected an increase in stress. Participants expected on average more or less *no difference* regarding their “irritation” (*M* = 2.72, *SD* = 0.94), with almost 42% expecting a decrease, less than 22% expecting an increase and the rest of the participants (ca. 37%) expecting no difference. Lastly, participants tended towards a *slight increase* in “fatigue” when driving with Mediator (*M* = 3.46, *SD* = 1.02). Almost 60% of the participants expected an increase, while the rest of the participants were equally distributed between a decrease and no differences.

Hence, while participants expected some positive outcomes of using Mediator, such as higher comfort and possibility to relax/rest, as well as reduced stress and irritation, they also seemed to expect less enjoyment of driving, more boredom and fatigue. The expected reduction of enjoyment of driving might be related to handing over vehicle control to the automated system, which might reduce the joy some drivers experience when actively driving themselves. Moreover, fatigue and boredom might also be related to handing over vehicle control. Especially during a drive with CM, where the drivers are responsible for monitoring the system and driving environment and are not allowed to do other activities, this monotone monitoring task might induce boredom and fatigue as well as reduce the enjoyment of driving.

Table 3.15 Comfort-related items of the questionnaire comparing Mediator vs. no Mediator [adapted from L3 Pilot / ADAPTIVE]. Average (M; range from “1 – great decrease” to “5 – great increase”) and standard deviation (SD).

What differences do you expect when using the Mediator compared to driving without Mediator?	*M (SD)*	Decrease (%)	No difference (%)	Increase (%)
Comfort	4.55 (0.58)	1.35	0	98.65
Possibility to relax/rest	3.88 (0.76)	4.05	22.97	72.97
Boredom	3.41 (1.01)	14.86	32.43	52.70
Enjoyment of driving	2.95 (1.28)	40.54	22.97	36.79
Stress	2.20 (0.92)	67.57	25.68	6.76
Irritation	2.72 (0.94)	41.89	36.49	21.62
Fatigue	3.46 (1.02)	20.27	20.27	59.46

Results from the interviews. During the interviews, participants were asked how they experienced the opportunity to be driven through the traffic jam and how they evaluated the possibility to read and answer an email while in SB mode. Ninety-two percent of the participants evaluated the opportunity to be driven through the traffic jam positively. Forty-nine percent mentioned the possibility to read and answer an email positively, especially the increased efficiency in using available time was mentioned by 17% of these drivers. Nevertheless, six percent of these drivers added that they only feel comfortable with the NDRT execution in a driving simulator. This indicates that perceived safety is an important aspect in this regard. Fifty percent of the participants rated the NDRT execution as less comfortable. Mentioned reasons by these participants were a reduced familiarity with this task while driving (22%), the close approach in drive 4 (19%) and perceived loss of control and missing trust (20%).

3.5.2.2. Influences of driver characteristics

L3 Pilot questionnaire (comfort). The influences of age, gender and driving experience were analysed descriptively. Neither of the driver characteristics seemed to have a particularly strong influence on the participants' comfort when driving with Mediator.

See Appendix 5.2.3.2 for more detailed descriptions of the influences of age, gender and driving experience (i.e. annual mileage).

In addition, the participants' answers regarding the comfort-related items were also correlated with their ATI scores as well as their general opinion (see Table 3.16). Only two significant correlations were found. Firstly, a positive correlation with a small effect size was found between the participant's general opinion about driving automation and the item "Driving with Mediator active was comfortable". Hence, participants who had a more positive opinion about driving automation seemed to have experienced driving with Mediator as more comfortable. In addition, a negative correlation with a small effect size was found between the participants' general opinion about driving automation and the item "Driving with Mediator on long journeys would make me tired". Hence, lower general opinion about driving automation seemed to be related to higher expectations to get tired on long journeys with Mediator.

Table 3.16 Correlations between comfort-related items and the participants' ATI score and general opinion about driving automation.

Statement	ATI	General opinion about driving automation
Driving with Mediator active was comfortable.	r = .22 p = .058	r = .24 p = .043
Driving with Mediator was difficult.	r = - .06 p = .627	r = - .01 p = .941
Driving with Mediator was demanding.	r = - .11 p = .357	r = - .15 p = .203
Driving with Mediator was stressful.	r = - .18 p = .116	r = - 1.34 p = .255
Driving with Mediator was on long journeys would make me tired.	r = - .12 p = .299	r = - .23 p = .047

Comparison Mediator vs. no Mediator. The driver characteristics age, gender and driving experience were also analysed descriptively for this questionnaire. No systematic influence could be found in terms of participants' gender, age or driving experience (i.e. annual mileage).

See Appendix 5.2.3.2 for more detailed descriptions of the influences of age, gender and driving experience (i.e. annual mileage).

In addition, the participants' answers regarding the comfort-related items were also correlated with their ATI scores as well as their general opinion (see Table 3.17). Only one significant correlation was found between the participants' ATI scores and the comfort-related items. A significant, negative correlation (small effect size) was found between the ATI scores and the expected difference in "stress". Hence, the higher the participants' ATI score, the more the participants expected a decrease in stress. With regard to the participants' general opinion about driving automation, six significant correlations were found. Firstly, a positive correlation (small effect size) was found between the general opinion and "comfort". Hence, the more positive the general opinion of the participants was, the higher the expected increase in comfort was as well. Another positive correlation (small effect size) was found between the general opinion and the "enjoyment of driving". The more positive the participants were about vehicle automation, the higher they expected the enjoyment of driving to be. Another significant, but negative correlation (medium effect size) was found between "boredom" and the general opinion. Here, the more positive the general opinion, the higher the expected decrease in boredom was. Moreover, the participants' general opinion also correlated negatively (small effect sizes) with the expected changes in "stress", "irritation" and "fatigue". Hence, the more positive the general opinion was, the higher the expected decrease in stress, irritation and fatigue was.

Table 3.17 Correlations between comfort-related items of the comparison Mediator vs. no Mediator and the participants' ATI score and general opinion about driving automation.

Statement	ATI	General opinion about driving automation
Comfort	r = .23 p = .053	r = .26 p = .024
Possibility to relax/rest	r = .06 p = .604	r = - .05 p = .693
Boredom	r = - .10 p = .398	r = - .31 p = .007
Enjoyment of driving	r = .06 p = .630	r = .26 p = .027
Stress	r = - .23 p = .047	r = - .24 p = .039
Irritation	r = - .18 p = .120	r = - .24 p = .036
Fatigue	r = - .16 p = .172	r = - .29 p = .011

3.5.3. RQ3: Usability

To answer the third research question (RQ3) "*How is the developed HMI experienced by the drivers (e.g., usability, transparency, comprehensibility, predictability)?*" questionnaires as well as interview questions were used. Regarding the questionnaires, several questionnaire items were designed for the current driving simulator study to assess the usability of Mediator. These items were administered after each of the three drives with Mediator (i.e. drive 2-4). In addition, the SUS (Brooke et al., 1996) was administered in the post evaluation. Finally, participants were asked to evaluate the Mediator system's usability through various questions within the interviews executed throughout the experiment.

Summary of results. On average, participants rated the indications in the display as clear and very easy to understand (M = 5.26 on a 6-point Likert scale ranging from 1 to 6). The potential of

Mediator to distract the driver was rated as low (*M* = 1.93 on a 6-point Likert scale ranging from 1 to 6). Regarding the displayed time budget, it seemed that drivers' understanding increased over extended system experience. Additionally, the colour coding was rated as helpful to increase mode awareness (*M* = 5.36 on a 6-point Likert scale ranging from 1 to 6). Mediator's usability was evaluated as good up to excellent (SUS score 85 out of 100). No considerable effects of driver characteristics on driving comfort were revealed. A tendency was found that higher affinity for technology is related to a more positive evaluation of Mediator's usability.

3.5.3.1. Questionnaire and interview results on usability

Usability-related items (self-designed). Figure 3.12 provides the results for the self-designed, usability-related items. Participants could indicate their agreement with the item-statements from "*1 – completely disagree*" to "*6 – completely agree*". Mean values for the statement "The indications in the display were clear and very easy to understand" were quite similar for all three drives with Mediator. The mean values ranged from *M* = 5.24 to *M* = 5.27, indicating *mostly agreement* for all three drives. In line with the similar mean values, statistical analysis did not reveal a significant effect, $F(2, 146) = 0.05$, $p = .956$, $\eta_p^2 = .00$. With respect to the statement "The information in the display was distracting", mean values were again quite similar ranging from *M* = 1.89 to *M* = 1.97, which indicated *mostly disagreement*. Again, no significant effect was found ($F(2, 146) = 0.38$, $p = .687$, $\eta_p^2 = .01$). Mean values for the statement "The display of the time budget was very easy for me to understand" ranged from *M* = 3.87 to *M* = 4.30, which shows (close to) *mostly agreement*. Interestingly, the agreement with the statement increased over the course of the three drives, with lowest agreement after drive 2 with CM (*M* = 3.87, *SD* = 1.53), somewhat higher agreement after drive 3 with SB and NDRT (*M* = 4.15, *SD* = 1.47) and the highest agreement after drive 4 with SB, the NDRT and close approach (*M* = 4.30, *SD* = 1.73). In line with these changes, statistical analysis revealed a significant effect ($F(2, 146) = 5.30$, $p = .006$, $\eta_p^2 = .07$), with a significantly higher mean value (i.e. higher agreement) after drive 4 than drive 2. It seems that participants' understanding of the time budget increased over extended system experience. With respect to the statement "The different colour concepts helped me to identify the current driving mode with certainty", mean values were again quite similar across the three drives and ranged from *M* = 5.23 to *M* = 5.42, indicating *mostly agreement*. Again, no significant effect was found ($F(1.76, 124.94) = 0.90$, $p = .408$, $\eta_p^2 = .01$). Regarding the last usability-related item, "I was sometimes unsure which driving mode I was in", mean values ranged from *M* = 1.54 to *M* = 1.77, indicating *mostly disagreement*. Again, no significant differences were found between the three drives ($F(2, 146) = 0.38$, $p = .687$, $\eta_p^2 = .01$).

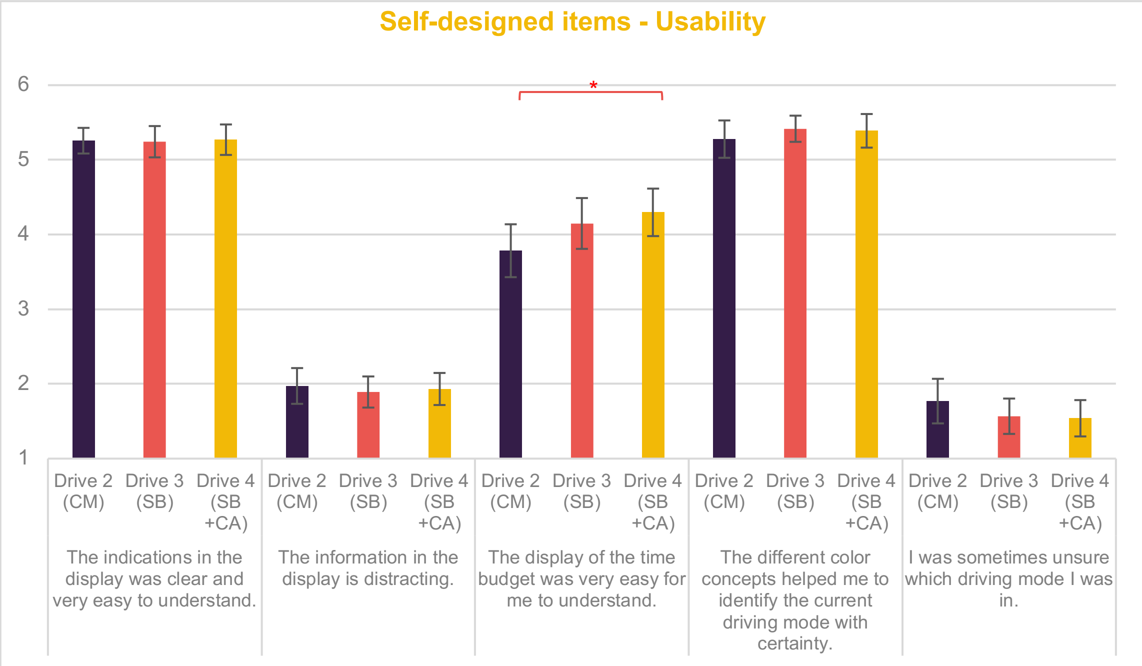


Figure 3.12 Usability-related items (self-designed). Scale range: "1 – completely disagree", "6 – completely agree"; Error Bars: 95th-CI.

SUS. Within the post-questionnaire, participants were also asked to evaluate the usability of the Mediator system considering all three drives with Mediator (i.e. drive 2-4). As can be seen in Figure 3.13, participants evaluated the Mediator's usability as good up to excellent and within the range of high acceptability (*M* = 84.86, *SD* = 11.96).

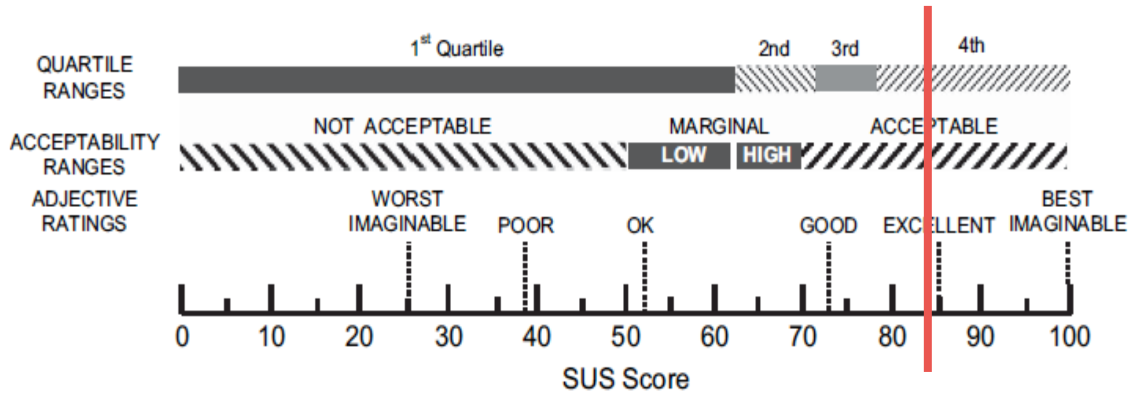


Figure 3.13 SUS-score (orange) with interpretation in quartiles, acceptability and adjective ratings. Graphic extracted from Bangor et al. (2008).

Results from the interviews. When asked about the different elements of Mediator, 16% of the drivers spontaneously evaluated the HMI concept positively, mentioning that it is understandable and easy to use. Drivers talked freely about their experiences with the Mediator HMI, hence, not all drivers mentioned all elements.

The messages in general were evaluated positively by the drivers with 66% of all drivers agreeing that the messages were understandable. Twenty-four percent of all drivers stated that sometimes they did not understand messages because, for instance, the reason for takeover or handover actions were not stated clearly enough. Fourteen percent explicitly expressed the wish for more transparency and more information. Seven percent stated that a warning when approaching the rear-end of the traffic jam would have been preferable. Regarding the audio messages, nearly 11% of all drivers mentioned the messages as helpful and understandable. One participant stated that he/she would prefer sound signals over voice messages especially for warnings. One participated suggested to have more verbal interaction between system and driver as a more social component of Mediator.

With respect to the amount of provided information, most of all participants (82%) stated that the amount of information was sufficient and not overloading. One participant stated that, although the amount of information was just right, he/she wants more information in other use cases like the messages shared via radio (e.g., persons on the road). Only two participants (3%) mentioned that they preferred less information (e.g., only the audio messages).

The light concept was evaluated positively by 66% of all drivers. Especially the colour coding was mentioned by 24% of all drivers. They stated, for instance, that the colours were selected wisely (no confusion with traffic lights or blue light of emergency vehicles), that they were perfectly visible in the periphery and helped to identify the current driving mode easily. One participant even mentioned the colour amber as activating and, hence, fitting for continuous mediation mode. Purple was identified as calming and, therefore, fitting to driver stand-by mode. Eight percent of the drivers would prefer colours other than amber and purple. Thirty-five percent of all drivers stated that the lighting concept was a bit too much (e.g., the additional LEDs on the steering wheel are not necessarily needed in their opinion; ambient light in the lower part of the vehicle was nearly not visible). Nevertheless, a few participants (3%) suggested to have additional LED at the doors. Nine percent of all drivers directly expressed the wish to personalize the colours, and 3% mentioned the need to adapt brightness of the light to the environmental conditions. Two people mentioned that they want to have another colour or a fading of colours showing the transition between two driving modes.

Drivers were directly asked if they understood the colour coding of the displayed route (grey vs. amber or purple). Fifty-five percent of all drivers interpreted the grey part of the route correctly (20% not), and 74% interpreted the purple part of the route correctly (13% not).
Eighty percent of all drivers stated that it was always clear to them, who is responsible for the driving task, for instance, because of the colour coding (mentioned by 23% of all drivers), Mediator's messages and displayed icons (mentioned by 15% of all drivers). Two drivers expressed their wish to have the choice which automated driving mode they want to activate (CM vs. SB) if both are available (not only get one suggestion they can choose or not chose).
Nearly 20% of all drivers mentioned positive aspects regarding the display like good visibility, dynamics of the displayed route, time budget information, display of the current driving mode, as well as layout and design. The most positive aspect was the announcement of the upcoming traffic jam mentioned by 10% of all drivers. Nineteen percent of all drivers did also mention aspects that need to be improved in their opinion like the displayed route or time budget information. The latter was mentioned by 16% of all drivers. The drivers stated that it was confusing and not perfectly understandable for them.

During the interviews, several screenshots from the display including the time budget information were shown to the participants. They were asked to explain the time budget concept. For driving in manual mode (time budget showed how long assisted/autopilot driving will be available as soon as the respective part is reached), 27% of all drivers explained the shown time budget correctly. Thirty-nine percent were not able to explain the time budget. For driving in autopilot mode (time budget showed how long autopilot driving is still available), 53% of all drivers explained the shown time budget correctly. Eleven percent were not able to explain the time budget. When asked about the time budget for the upcoming part of the route which stayed stable until the driver reached the section, 43% of the participants were able to explain the time budget correctly. Twenty-six percent were not able to explain the time budget correctly. When asked why the countdown of the time budget started when reaching the respective part even if not changing driving mode, 34% were able to explain and 54% were not.

With respect to the animation of the displayed route, 41% of participants recognized the animation when driving in manual mode (i.e., the amber or purple section on the upper part of the displayed route was getting bigger when approaching the section) and 15% did not. In automated driving mode, 26% of participants recognized the animation (i.e., the amber or purple section became smaller when reaching the end of ODD) and 30% did not.

When asked about the steering wheel icon, 32% of all drivers interpreted the icon correctly for manual driving and autopilot driving. Fifty-two percent provided a correct interpretation for manual driving but were unsure about autopilot driving (e.g., mixed-up assisted and autopilot driving, misinterpreting the icon for autopilot driving as request to have on hand on the steering wheel or to take back control). Fifteen percent of the participants stated that they didn't pay attention to the icon at all and, hence, cannot provide any interpretation or provided a wrong interpretation (e.g., icon is displaying what actually happens in the cockpit or reflects the ego-vehicle).
Ten percent of all drivers expressed the wish to have navigational information included in the HMI, 11% would have preferred an additional Head-up Display showing most important information. Another 7% of all drivers suggested to have the possibility for voice commands.

3.5.3.2. Influences of driver characteristics

As can be seen in Table 3.18, only one of the five analysed driver characteristics correlated significantly with the SUS score. A positive correlation of (almost) moderate effect size was found between the ATI score and the SUS score. Hence, a higher affinity for technology seemed to be related to a more positive evaluation of Mediator's usability. Moreover, another positive correlation with a small effect size was found between the participants' general opinion about driving automation and the SUS score, which was, however, not significantly. Nevertheless, it indicated that a more positive opinion towards driving automation seemed to be related to a more positive evaluation of Mediator's usability as well.

Table 3.18 Correlations between System Usability Scale (SUS) and various driver characteristics.

	Gender	Age	Driving experience (i.e. annual mileage)	ATI	General opinion about driving automation
SUS	$r = .00$ $p = .998$	$r = .04$ $p = .753$	$r = .04$ $p = .724$	$r = .29$ $p = .012$	$r = .21$ $p = .080$

3.5.4. RQ4: Trust

To answer the fourth research question (RQ4) "*How much do the drivers trust Mediator and its functionalities?*" the following questionnaires were used, which were complemented by interview questions as well. Firstly, the trust in automation questionnaire (Jian et al., 2000) was used.

Similarly to the Van der Laan Acceptance scale, the trust in automation scale was administered five times during the simulator study:

- (1) after participants received a detailed explanation of Mediator and its functionalities but before the first actual drive with the system (i.e., Drive 2) to assess participants' expected trust in Mediator,
- (2) once after each drive with Mediator (i.e., Drive 2-4) to assess participants' trust in Mediator in different use cases (e.g., CM vs. SB, close approach), and
- (3) in the post evaluation to assess participants' overall trust of Mediator after all drives.

Moreover, participants' trust was also evaluated by means of the L3 Pilot questionnaire (Metz et al., 2019), which was administered in the post evaluation. Again, only the results regarding the trust-related items will be discussed in the following subsections.

Summary of results. On average, participants trusted Mediator (*M* = 5.5 on a 7-point Likert scale from 1 to 7). After actual system experience, participants' trust in automation increased compared to solely reading about its functionalities and capabilities. Similar to the participants' acceptance, their trust in Mediator decreased after drive 4, where they experienced the uncomfortable close approach. Additionally, 73% of the participants want to monitor Mediator's performance constantly. No considerable effects of driver characteristics on trust were revealed. Affinity for technology as well as general opinion about vehicle automation are positively correlated to drivers' trust in Mediator.

3.5.4.1. Questionnaire and interview results on trust

Trust in automation. Figure 3.14 shows the results regarding participants' trust in automation for the three drives with Mediator, the expected trust (pre drive 2) and the post evaluation (overall evaluation after experiencing all drives with Mediator). Generally, trust in automation ratings were quite high throughout all five data evaluation points, ranging from *M* = 4.73 (*SD* = 1.70) for the expected trust to *M* = 5.63 (*SD* = 1.74) for drive 3 with SB. Statistical analysis also revealed a significant effect of the data collection points, $F(2.48, 180.94) = 30.32$, $p < .001$, $\eta_p^2 = .29$. The expected trust (pre drive 2; *M* = 4.73, *SD* = .70) was significantly lower than trust ratings after the three drives with Mediator (drive 2 with CM: *M* = 5.57, *SD* = 1.70; drive 3 with SB and NDRT: *M* = 5.63, *SD* = 7.40; drive 4 with SB, NDRT and close approach: *M* = 5.34, *SD* = 1.11) and the post evaluation (*M* = 5.50, *SD* = 1.89). Hence, participants' trust in automation significantly increased after actual system experience compared to solely reading about its functionalities and capabilities. Moreover, similar to the participants' acceptance, their trust in Mediator also decreased after drive 4, where they experienced the uncomfortable close approach at the end of the traffic jam, compared to drive 3. However, for their trust ratings the different was not significant. In the post evaluation, considering all three drives, an increase in trust in Mediator was revealed. Similar to acceptance ratings, trust ratings did not rise back up to the levels of drive 2 and drive 3. Hence, the close approach slightly reduced the participants' overall trust evaluation.

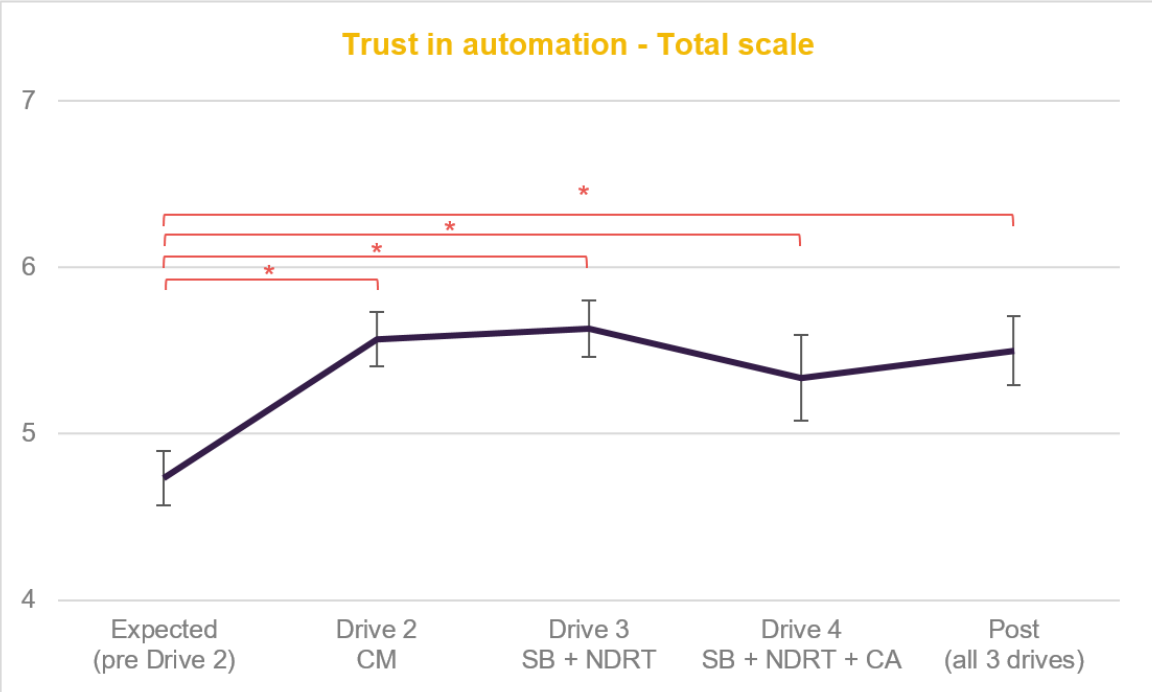


Figure 3.14 Trust in automation. Scale range: "1 – totally disagree", "7 – totally agree"; Error Bars: 95th-CI.

L3 Pilot questionnaire (trust). Table 3.19 provides the mean and standard deviation (range from "*1 - totally disagree*" to "*5 – totally agree*"), as well as the share of (dis-)agreement with regard to the trust-related items. Regarding the item "I trust Mediator to drive", the mean value trended towards *agreement* (*M* = 3.78, *SD* = 1.94), indicating that participants seemed to trust Mediator. In line with this, almost 79% of the participants agreed with the statement, while less than 14% disagreed and online slightly more than 8% were undecided. Mean value for the statement "I would want to monitor Mediator's performance" indicated a similar trend towards *agreement* (*M* = 3.78, *SD* = .93). Hence, even though participants seemed to trust Mediator they still would want to monitor its' performance. Here, almost 73% of the participants agreed, while almost 15% were undecided and a little more than 12% disagreed. With respect to the last statement "During driving with Mediator active, I monitored the surrounding environment more than in manual driving", the mean value indicated a slight trend towards *disagreement* with the statement (*M* = 2.64, *SD* = 1.17). However, only a little more than 50% of the participants disagreed with the statement, while more than 31% agreed with it and almost 18% were undecided. Hence, even though participants indicated to trust Mediator, they still wanted to monitor its' performance and seemed to monitor the surrounding environment similarly as during manual driving.

Table 3.19 Trust-related items adapted from L3 Pilot questionnaire: Average (M; range from 1 – totally disagree to 5 – totally agree) and standard deviation (SD), share of (dis-) agreement with each statement).

Statement	*M* *(SD)*	Disagreement (%)	Neutral (%)	Agreement (%)
I trust Mediator to drive	3.78 (1.94)	13.51	8.11	78.38
I would want to monitor Mediator's performance.	3.78 (1.93)	12.16	14.86	72.97
During driving with Mediator active, I monitored the surrounding environment more than in manual driving.	2.64 (1.17)	51.35	17.57	31.08

Results from the interviews. Eighty-seven percent of all participants stated that they trusted Mediator fully, 8% trusted the system partly and another 8% expressed low trust in Mediator. Participants were asked in which functionalities or components of Mediator they trust. Almost one-third (31%) said that they fully trust in Mediator performing the driving task sufficiently including keeping a safe distance to other vehicles, maintain the allowed speed, lane keeping, and recognizing and reacting to the driving situation appropriately. A quarter of all participants (26%) expressed confidence in the technology, such as the sensors in the car.
Further, participants were asked why the trust Mediator. Sixty percent stated positive experience with Mediator during the first drives as important reason for building up trust. Experience with similar (assistance) systems was mentioned by 8% of the participants. Also, 8% of the participants stated that they only trust Mediator because of the simulated and, hence, safe environment. Additionally, participants were asked to think about possibilities to increase their trust in Mediator. More than half of all participants (55%) mentioned that more experience with the system would help them gain more trust in Mediator, for instance, in real traffic and over longer distances. Experiencing the system in dangerous or more complex situations was also considered to be helpful in increasing confidence in Mediator. Fifteen percent of the participants stated that positive experiences and recommendations from others (e.g., scientific surveys and tests as well as experiences from friends in real traffic) will increase their trust in Mediator further. Some participants (11%) would also have liked to receive more information about the status of the system or about the driving environment from Mediator while driving (e.g., via displays in the cockpit and announcements inside the vehicle). Another 10% would have more confidence in Mediator if they had been able to acquire more information in advance, specifically, about the technical details, functionalities and limitations.

Finally, participants were also asked about instances that might lead to a reduction of their trust in Mediator. Sixty-five percent assumed that any failure of the system will lead to a loss of confidence (e.g., accidents, software errors, no or wrong reaction to certain traffic situations, reacting too late). Another 15% indicated that disagreements between Mediator and the driver would cause a decrease in trust (i.e., Mediator reacts differently than the driver would react in a certain situation, hence, unpredictable or unexpected behaviour from Mediator). Reports of bad experiences with Mediator (e.g. by other drivers or in the media) were cited by 10% of the participants as a reason for declining trust. Five percent of the participants expressed concerns that the vehicle can be hacked.

3.5.4.2. Influences of driver characteristics

Trust in automation. In addition to comparing whether the five different data collection points influenced the participants' trust in Mediator differently, a rmANOVA was conducted including the participants' age, gender and driving experience (i.e. annual mileage) and affinity for technology (ATI score). In addition to the earlier discussed effect of the data collection points, there were no significant main effects of the driver characteristics or interaction effects of the driver characteristics with each other or the five data collection points (see Appendix 5.2.3.3). Hence, the participants' age, gender, driving experience and their affinity for technology did not seem to influence their trust in automation significantly.

Additionally, the relationship between the data collection points of the trust in automation scale and the participants' ATI scores as well as their general opinion about vehicle automation was analyzed (Table 3.20). The participants' ATI scores correlated positively with the participants' expected trust rating (medium effect size), trust ratings after drive 2 (medium effect size) and after drive 3 (medium effect size). Hence, for these three data collection points the higher the participants' ATI scores, the higher the participants' trust in Mediator. The ATI scores also correlated positively with the trust ratings after drive 4 (small effect size) and the post evaluation (small effect size). However, these correlations were not significant. The participants' general opinion about vehicle automation correlated significantly with all data collection points and all correlations had medium effect sizes. Hence, the more positive the participants' opinion about automating parts or the entire driving task was, the higher was their trust in Mediator throughout the entire experiment. In contrast to the relationship with the ATI scores, the correlation between the general opinion and trust after drive 4 and the post evaluation was strong and statistically significant.

Table 3.20 Correlations between the Trust in automation scale and the affinity for technology (ATI) as well as the general opinion about vehicle automation.

	Expected (pre drive 2)	Drive 2 CM	Drive 3 SB + NDRT	Drive 4 SB + NDRT + close approach	Post (all three drives)
ATI	$r = .32$ $p = .005$	$r = .37$ $p = .001$	$r = .36$ $p = .002$	$r = .18$ $p = .322$	$r = .12$ $p = .341$
General opinion about vehicle automation	$r = .31$ $p = .008$	$r = .34$ $p = .003$	$r = .43$ $p < .001$	$r = .40$ $p < .001$	$r = .39$ $p = .001$

L3 Pilot questionnaire (trust). For this questionnaire, the influence of age, gender and driving experience were analysed descriptively as well. Both male and female participants gave quite similar responses regarding the three trust-related items. The comparison of the four age groups revealed a few interesting differences. Firstly, with increasing age the participants seem to agree more with the statement "I trust Mediator to drive" as well as disagree with the statement "I would want to monitor Mediator's performance". In both cases the oldest age group (55 years and older) agreeing the most with the former and disagreeing the most with the latter statement. This might be related to more driving experience as well as more experience with ADAS systems. Moreover, the higher trust might be related to a reduced need to monitor Mediator's performance, while the other groups, with somewhat lower trust, wanted to monitor it more. Lastly, the oldest age group also agreed the most with the statement "During driving with Mediator active, I monitored the surrounding environment more than in manual driving". This result is also in line with the previous finding. The oldest age group had enough trust in Mediator to not monitor Mediator's performance constantly and, hence, had time to concentrate on the surrounding environment. Lastly, the comparison of the driving experience (i.e. annual mileage) groups revealed no considerable and systematic differences .

See Appendix 5.2.3.3 for more detailed descriptions of the influences of age, gender and driving experience (i.e. annual mileage).

In addition, the answers of the three items were also correlated with the participants' ATI scores and their general opinion about driving automation (Table 3.21). No significant correlation could be revealed for the participants' ATI scores. With respect to the general opinion about driving automation, a significant positive correlation with the item "I trust in Mediator to drive" was found (medium effect size). Hence, the more positive their general opinion was, the more the participants tended towards trusting Mediator.

Table 3.21 Correlations between the trust-related items of the questionnaire adapted from L3 Pilot and the affinity for technology (ATI) as well as the general opinion about vehicle automation.

Statement	ATI	General opinion about driving automation
I trust Mediator to drive.	r = .17 p = .160	r = .36 p = .001
I would want to monitor Mediator's performance.	r = - .21 p = .077	r = - .22 p = .057
During driving with Mediator active, I monitored the surrounding environment more than in manual driving.	r = .10 p = .384	r = .16 p = .166

3.5.5. RQ5: Perceived Safety

The fifth research question (RQ5) „*How is the perceived safety rated by the drivers?*" was addressed by means of the following questionnaires. Firstly, the safety-related items of the questionnaire comparison Mediator vs. no Mediator adapted from the L3 Pilot and ADAPTIVE (Metz et al., 2019; Rodarius et al., 2015) projects were used. In addition, safety-related items from the L3 Pilot questionnaire (Metz et al., 2019) were used as well. Lastly, three questionnaire items were designed for the purpose of this simulator study. Two of these items were administered after each of the three drives with Mediator (i.e. drive 2-4). The third item was only administered after the two drives with SB (i.e., drive 3 and drive 4), where the participants were asked to execute the NDRT. In addition, perceived safety was also addressed within the interviews.

Summary of results. On average, more than 86% of the participants expected an increase in safety when driving with Mediator, for instance, in terms of decreased number of accidents (85% agreed) or accidents' severity (82% agreed). 78% of the participants felt safe when driving with Mediator active. Perceived safety while approaching the end of the traffic jam was rated quite good (M = 4.32 on a 6-point Likert scale from 1 to 6), although the rating was lower after drive 4 with the close approach (M = 2.88 on a 6-point Likert scale from 1 to 6). No considerable effects of driver characteristics on perceived safety were revealed. Drivers' general opinion about vehicle automation seem to be positively correlated to drivers' perceived safety of Mediator.

3.5.5.1. Questionnaire and interview results on perceived safety

Comparison Mediator vs. no Mediator. Table 3.22 provides the results with regard to the safety-related items comparing using Mediator while driving versus not using Mediator. The answers could range from "*1 – strong decrease*" to "*5 – strong increase*". With regard to "safety", participants expected on average a *slight increase* (M = 4.22, SD = 0.86). In total, more than 86% of the participants expected the increase in safety when driving with Mediator, while only slightly more than 5% expected a decrease. In line with this, participants expected on average a *slight decrease* for the “number of accidents” (*M* = 1.82, *SD* = 0.93). Here, a little more than 85% of the participants expected a decrease, while the rest of the participants were almost evenly split between no differences and an increase in the number of accidents. Similarly, participants expected on average a *slight decrease* (*M* = 1.72, *SD* = 1.05) regarding the “accidents’ severity”. A little more than 82% expected a decrease, while the rest of the participants expected either no differences or an increase in accidents’ severity. Lastly, participants expected on average a *slight decrease* (*M* = 2.09, *SD* = 1.22) regarding the “attention towards the road”. Here, ca. 77% expected a decrease, while less than 18% expected an increase. Overall, the participants seemed to be quite positive about the potential of using Mediator to improve their safety compared to not using Mediator. Although, expecting a decrease regarding the attention of the road, this does not necessarily implies a reduction in safety. While driving in SM mode, drivers do not need to pay attention to the road and are allowed to, for instance, engage in NDRTs. Nevertheless, it should be avoided that participants pay less attention towards the road while driving manually or in CM because they are distracted by Mediator (e.g., by the information displayed on the screens).

Table 3.22 Safety-related items of the questionnaire comparing Mediator vs. no Mediator [adapted from L3 Pilot / ADAPTIVE]. Average (M; range from “1 – great decrease” to “5 – great increase”) and standard deviation (SD).

What differences do you expect when using the Mediator compared to driving without Mediator?	*M (SD)*	Decrease (%)	No difference (%)	Increase (%)
Safety	4.22 (.86)	5.41	8.11	86.49
Number of accidents	1.82 (.93)	85.14	6.76	8.11
Accidents’ severity	1.72 (1.05)	82.43	8.11	9.46
Attention towards the road	2.09 (1.22)	77.03	5.41	17.57

Safety-related items (self-designed). Figure 3.15 shows the results regarding the self-designed, safety-related items. Participants could indicate their agreement with the items from “*1 – completely disagree*” to “*6 – completely agree*”. Mean values for the statement “I felt very safe while Mediator was driving through the traffic jam for me” ranged from *M* = 4.54 to *M* = 5.18, indicating *mostly agreement* with a trend towards *complete agreement* for all three drives. The statistical analysis did reveal a significant effect as well ($F(1.77, 128.92) = 12.76$, $p < .001$, $\eta_p^2 = .15$). Participants indicated significantly lower agreement with the statement after drive 4 with SB, the NDRT and the close approach (*M* = 4.54, *SD* = 1.5), compared to drive 2 with CM (*M* = 5.18, *SD* = .9) and drive 3 with SB and the NDRT (*M* = 5.12, *SD* = 1.05). Overall, participants seemed to feel quite safe when Mediator drove them through the traffic jam during the three drives. However, they felt significantly less safe during the traffic jam after they experienced the close approach in drive 4 compared to the other two incident-free drives (i.e. drive 2 and 3). The mean values for the second statement “I felt very safe while approaching the end of the traffic jam” ranged from *M* = 2.88 to *M* = 4.51, indicating an answers range between *rather do not agree* (drive 4) to *rather agree* to even *largely agree* (drive 2 and 3). In line with this relative broad range, the statistical analysis revealed a significant effect ($F(2, 146) = 41.41$, $p < .001$, $\eta_p^2 = .36$). Participants indicated significantly lower agreement after drive 4 (*M* = 2.88, *SD* = 1.79), compared to drive 2 (*M* = 4.12, *SD* = 1.39) and drive 3 (*M* = 4.51, *SD* = 1.26). Again, participants seemed to feel considerably less safe during the approach of the traffic jam in drive 4, where they experienced the close approach, compared to the other two drives where the approach of the traffic jam was done slowly and smoothly. Moreover, participants indicated significantly higher agreement with the statement after drive 3 compared to drive 2. Participants might have felt safer during drive 3 than drive 2, due to the different weather conditions, i.e. clear weather during drive 3 versus foggy weather during drive 2, as well as increased system experience in drive 3 compared to drive 2. Lastly, the mean values for the item "I felt safe while answering the email", which was only administered after the two drives with SB (i.e. drives 3 and 4), differed significantly between the two drives ($t(73) = 2.11$, $p = .019$, $r = .234$). Participants indicated significantly higher agreement with the statement after drive 3 ($M = 4.24$, $SD = 1.42$) compared to drive 4 with the close approach ($M = 3.95$, $SD = 1.59$). Hence, participants seemed to feel safer while answering the email during drive 3 than drive 4, which could be due to the close approach. Overall, while the participants seemed to feel quite safe when driving through the traffic jam with Mediator, the close approach clearly resulted in a decrease in their perceived safety. Especially during the approach of the traffic jam participants felt considerably less safe during drive 4 than drive 2 and 3. However, also after the close approach, when Mediator drove the participants through the traffic jam, participants still felt less safe than after the incident-free approaches in drive 2 and 3.

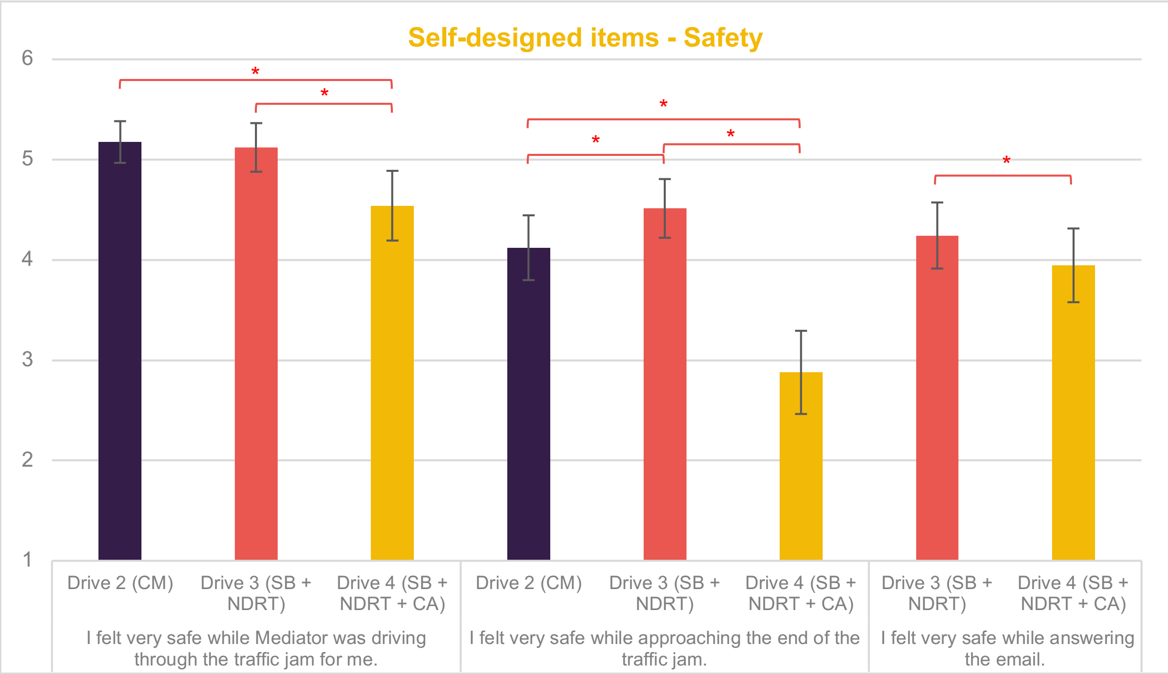


Figure 3.15 Safety-related items (self-designed). Scale range: "1 – completely disagree", "6 – completely agree"; Error Bars: 95th-CI.

L3 Pilot questionnaire (safety). Table 3.23 provides the mean and standard deviation (range from "*1 - totally disagree*" to "*5 – totally agree*"), as well as the percentage of (dis-)agreement with regard to the safety-related items. Regarding the first item "I felt safe when driving with Mediator active", the mean value indicates a trend towards *agreement* with the statement ($M = 3.93$, $SD = .87$). In line with this, a little more than 78% of the participants agreed with the statement, while less than 10% disagreed and online slightly more than 12% were undecided. The mean values for the second item "During driving with Mediator active, I was more aware of hazards in the surrounding environment than in manual driving" indicated *neutral* answers trending towards *disagreement* ($M = 2.41$, $SD = 1.0$). A little more than 60% disagreed with the statement, while a little more than 25% indicated neutral opinions and less than 14% agreed with the statement. Results reveal that Mediator didn't help the drivers to increase their situational awareness.

Table 3.23 Safety-related items adapted from L3 Pilot questionnaire: Average (M; range from 1 – totally disagree to 5 – totally agree) and standard deviation (SD), percentage of (dis-) agreement per statement.

Statement	M (SD)	Disagreement (%)	Neutral (%)	Agreement (%)
I felt safe when driving with Mediator active.	3.93 (.87)	9.46	12.16	78.38
During driving with Mediator active, I was more aware of hazards in the surrounding environment than in manual driving.	2.41 (1.0)	60.81	25.68	13.51

Results from the interviews. During the interviews, participants were not directly asked about perceived safety while driving with Mediator, but were asked about potential benefits if distraction detection and fatigue detection would have been implemented in Mediator. Sixty-two percent of the participants anticipated an increase in driving safety with added distraction warning. Fifty-eight percent of the participants were in favour of fatigue detection.

3.5.5.2. Influences of driver characteristics

Comparison Mediator vs. no Mediator. The differences regarding the participants' age, gender and driving experience were analysed descriptively as well. With respect to the participants' gender, no considerable differences were found between male and female participants. The comparison of the four age groups revealed somewhat greater differences. Firstly, the oldest age group (55 years and older) tended towards a *great increase* regarding the "safety", while the other three groups only expected a *slight increase.* Moreover, the oldest age group also expected the biggest decrease (i.e. tendency towards *great decrease*) in the "number of accidents" and "accident severity", while most of the other three groups tended more towards a *slight decrease*. Lastly, with increasing age, the participants seemed to expect smaller differences in the "attention towards the road". The youngest age group (18-24 years) expected the greatest decrease in attention towards the road, while the oldest age group tended the most towards no differences. Regarding the four driving experience (i.e. annual mileage) groups, no substantial, systematic influence on perceived safety could be identified.

See Appendix 5.2.3.4 for more detailed descriptions of the influences of age, gender and driving experience (i.e. annual mileage).

In addition, the participants' answers regarding the safety-related items were also correlated with their ATI scores as well as their general opinion (see Table 3.24). The ATI scores of the participants did not correlate significantly with their answers regarding the four safety-related items. However, the four items all correlated with the participants' general opinion about driving automation. Firstly, a positive correlation (medium effect size) was found between the general opinion and "safety". Hence, a more positive general opinion was related with an expected increase in safety. In line with this, the general opinion correlated negatively with the "number of accidents" (small effect size) and the "accidents' severity" (small effect size). Hence, a more positive general opinion was also related to an expected decrease in the number and severity of accidents. Lastly, the general opinion was positively correlated (small effect size) with the "attention towards the road". Hence, the more positive the general opinion was, the higher the expected increase in attention towards the road was as well.

Table 3.24 Correlations between safety-related items from Comparison Mediator vs. no Mediator questionnaire and the participants' ATI score and general opinion about driving automation.

Items	ATI	General opinion about driving automation
Safety	r = .02 p = .892	r = .38 p = .001
Number of accidents	r = - .02 p = .129	r = - .24 p = .038
Accidents' severity	r = - .1 p = .377	r = - .29 p = .013
Attention towards the road	r = .15 p = .202	r = .24 p = .042

L3 Pilot questionnaire (safety). For this questionnaire, the influences of age, gender and driving experience were analysed descriptively as well. Both male and female participants gave very similar responses regarding the three safety-related items. Comparing the four age groups revealed differences that were somewhat more considerable. Firstly, while all four groups seemed to feel relatively safe, the oldest age group (55 years and older) seemed to feel even a little safer than the other three age groups. Moreover, all four groups tended to *slightly disagree* with the second statement that they were more aware of hazards in the surrounding environment during Mediator active than in manual driving. However, the oldest and the youngest (18-24 years) age group disagreed a little less with the statement than the other two age groups. Hence, while all four groups indicated that they were not more aware of hazards when driving with Mediator active, the youngest and oldest group might have been a little more aware than the other groups. Lastly, with respect to the four driving experience (i.e. annual mileage) groups no considerable, systematic influcne could be found.

See Appendix 5.2.3.4 for more detailed descriptions of the influences of age, gender and driving experience (i.e. annual mileage).

Additionally, the participants' answers regarding the two safety-related items were also correlated with their ATI scores and their general opinion about driving automation (Table 3.25). The ATI scores did not reveal any significant correlations with the two items. With respect to the general opinion, a positive correlation was found for the first item "I felt safe when driving with Mediator active" (medium effect size). The more positive the general opinion was, the more agreement with the statement was indicated, hence the safer the participants felt with Mediator active.

Table 3.25 Safety-related items adapted from L3 Pilot questionnaire: Average (M; range from 1 – totally disagree to 5 – totally agree) and standard deviation (SD), percentage of (dis-) agreement per statement.

Statement	ATI	General opinion about driving automation
I felt safe when driving with Mediator active.	r = .20 p = .085	r = .4 p < .001
During driving with Mediator active, I was more aware of hazards in the surrounding environment than in manual driving.	r = .14 p = .250	r = .21 p = .079

3.5.6. RQ6: Comfort-related transfers human to machine

To address the sixth research question (RQ6) „*How are the comfort-related TOC from manual to automated driving experienced by the drivers?*" the following questionnaires or items were used. Firstly, three items from the takeover questionnaire adapted from L3 Pilot (Metz et al., 2019) that focused on the transfer from human to machine were used. In addition, three items were designed especially for the purpose of the study to gain more information. All of these six items were administered after each of the three drives with Mediator.

Summary of results. On average, participants understand why Mediator offered to take over the driving task for them very well (M = 4.16 on a 5-point Likert scale from 1 to 5). Further, participants understand the instructions to hand over the driving task similarly well (M = 4.59 on a 5-point Likert scale from 1 to 5). The timing of the offer to have Mediator take over the driving task was evaluated very good by the participants (M = 4.14 on a 5-point Likert scale from 1 to 5). Participants felt quite safe while the handover was performed (M = 3.84 on a 5-point Likert scale from 1 to 5).

3.5.6.1. Questionnaire and interview results on comfort-related transfers

Takeover questionnaire adapted from L3 Pilot (TOC human to machine). Figure 3.16 provides the results for the three items from the L3 Pilot questionnaire. Participants could indicate their agreement with the items from "*1 – strongly disagree*" to "*5 – strongly agree*", with "*3 – neutral*". Mean values for the statement "It was obvious to me why Mediator's takeover was offered" were quite similar for all three drives with Mediator (Figure 3.16). The mean values ranged from M = 4.12 to M = 4.27 indicating *agreement* for all three drives. Hence, overall the participants seemed to understand why Mediator offered them to take over the driving task for them throughout all three drives. The statistical analysis did not reveal a significant effect ($F(1.64, 119.47) = 2.09$, $p = .127$, $\eta_p^2 = .03$). With respect to the statement "I would have liked more information about why a takeover by Mediator was offered", the mean values were again quite similar (Figure 3.16), ranging from M = 2.57 to M = 2.64, which indicated *disagreement* with a strong tendency towards a *neutral* opinion. Again, no significant effect was found ($F(2, 146) = .16$, $p = .855$, $\eta_p^2 = .00$). Hence, participants did not seem to be very keen to have more information about the reasons for the takeover by Mediator for all three drives similarly. Lastly, the mean values for the statement "During the takeover by Mediator, I always felt safe" differed somewhat more between the three drives than for the other two items. Here, the mean values ranged from M = 3.51 to M = 4.01 indicating *agreement* as well as the middle between *agreement* and a *neutral* opinion. In line with this greater range, the statistical analysis revealed significant differences ($F(2, 146) = 8.81$, $p < .001$, $\eta_p^2 = .11$). After drive 4 with SB, the NDRT and the close approach (M = 3.51, SD = 1.24) participants agreed significantly less with the statement compared to drive 2 with CM (M = 3.99, SD = .99) and drive 3 with SB and the NDRT (M = 4.01, SD = .93). Hence, participants seemed to feel significantly less safe during the takeover by Mediator in drive 4 than in the other two drives. It seems likely that the experience of the close approach, which happened after Mediator had already taken over the driving task, negatively influenced the participants' evaluation of their perceived safety during the takeover.

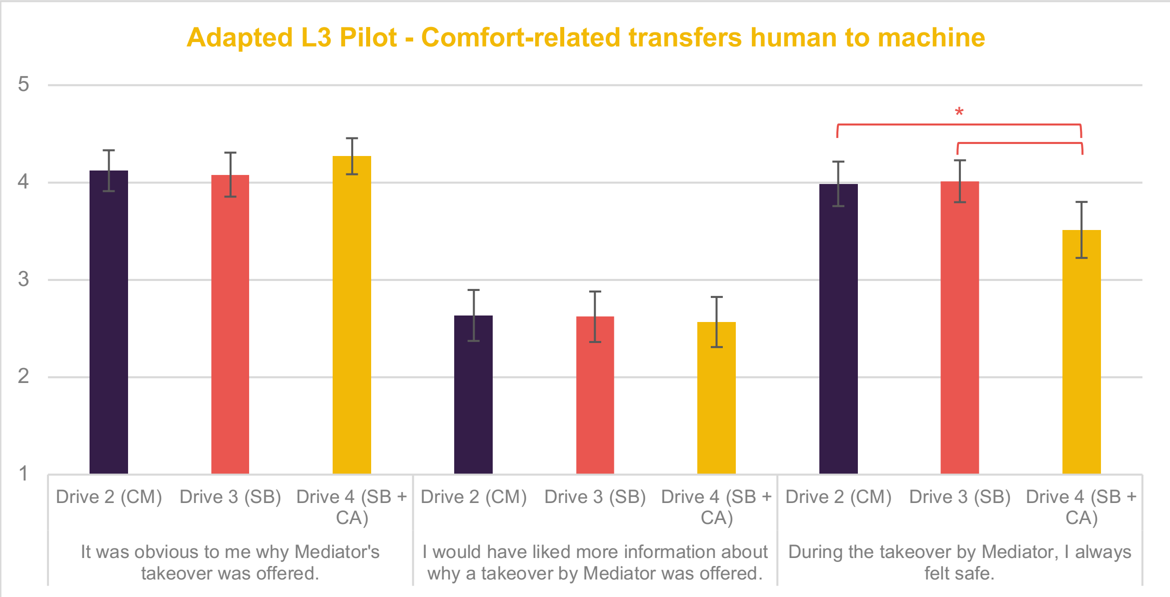


Figure 3.16 Comfort-related transfers human to machine (Adapted L3 Pilot). Scale range: 1 – strongly disagree, 5 – strongly agree; Error Bars: 95th-CI.

Comfort-related transfers items (self-designed). Figure 3.17 provides the results for the three self-designed items. Participants could indicate their agreement with the items from "*1 – strongly disagree*" to "*5 – strongly agree*", with "*3 – neutral*". Mean values for the statement "I only agreed to Mediator taking over because I was instructed that way at the beginning of the study" ranged from *M* = 2.01 to *M* = 2.36 indicating dis*agreement* with a slight tendency towards *neutral* for the three drives. The statistical analysis revealed a significant effect ($F(2, 146) = 6.104$, $p = .003$, $\eta_p^2 = .08$). Participants disagreed significantly more after drive 3 with SB and the NDRT (*M* = 2.08, *SD* = 1.07) and drive 4 with SB, the NDRT and close approach (*M* = 2.01, *SD* = 1.03) than after drive 2 with CM (*M* = 2.36, *SD* = 1.17), where participants tended slightly more towards *neutral.* Hence, overall participants seemed to be handing over the driving task to Mediator mostly because they wanted to and not only because of the study's instructions. However, during drive 3 and 4, hence the second and third drives with Mediator offering to take over the driving tasks, the participants seemed to hand over the driving task even a little more out of their own will. It seems likely that the advantages of Mediator experienced in drive 2 influenced the drivers' willingness to hand over the driving task in the following two drives. Moreover, the participants might also been slightly more reluctant to hand over the driving task during drive 2 due to the foggy weather conditions. With regard to the second item „It was very easy for me to understand the instructions to hand over the driving task to Mediator" the mean values ranged from *M* = 4.54 to *M* = 4.65 indicating *agreement* with a tendency towards *strongly agree* for all three drives. In line with these similar mean values, the statistical analysis did not reveal any significant effects ($F(2, 146) = 1.20$, $p = .303$, $\eta_p^2 = .02$). Hence, for all three drives the participants seem to understand the instructions to hand over the driving task similarly well. Lastly, the mean values for the item "The timing of the offer to have Mediator take over the driving task was very good" were very similar and ranged from *M* = 4.11 to *M* = 4.20 indicating *agreement*. The statistical analysis did not reveal a significant effect either ($F(1.84, 133.99) = .39$, $p = .679$, $\eta_p^2 = .01$). Hence, for all three drives the timing seemed to be very good according to the participants.

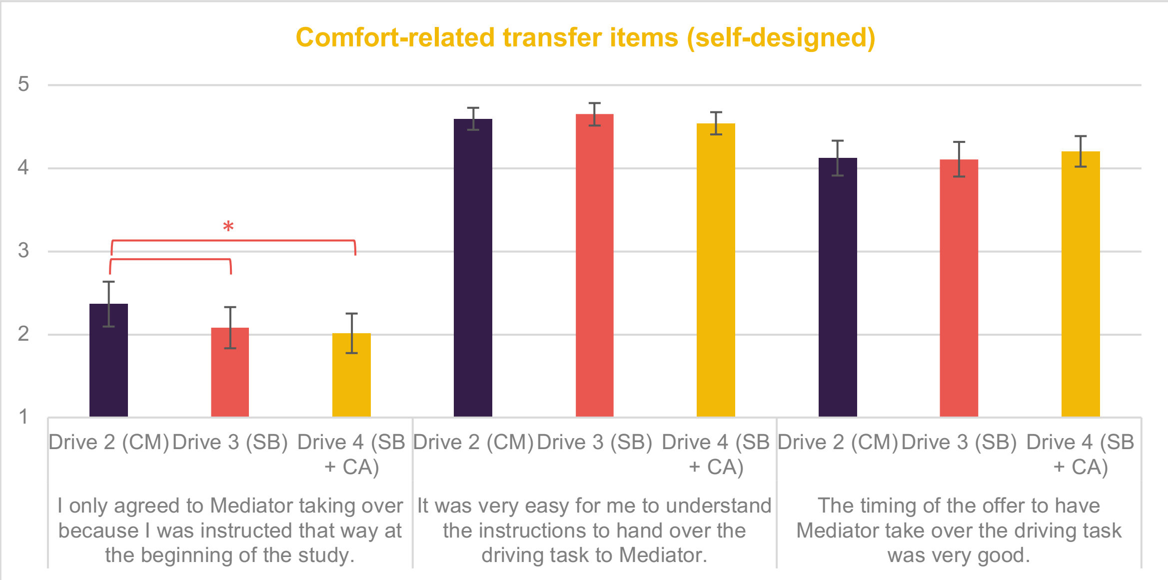


Figure 3.17 Comfort-related transfers human to machine (self-designed). Scale range: 1 – strongly disagree, 5 – strongly agree; Error Bars: 95th-CI.

Results from the interviews. Only three participants explicitly mention the transitions from manual to automated driving. Two of them (3% of all drivers) stated that the process was understandable.

3.5.7. RQ7: (Un-)planned transfers from machine to human

To address the seventh research question (RQ7) „How are the planned and unplanned transfers of control from automated to manual driving experienced by the drivers?" the following questionnaires or items were used. Firstly, five items from the takeover questionnaire adapted from L3 Pilot (Metz et al., 2019) that focused on the transfer from human to machine were used. In addition, one item was designed especially for the purpose of the study to gain more information. All of these six items were administered after each of the three drives with the Mediator system.

Summary of results. On average, participants understand why TOR occurred quite well (M = 3.66 on a 5-point Likert scale from 1 to 5). Further, participants understand the instructions to take over the driving task very well (M = 4.44 on a 5-point Likert scale from 1 to 5).
Further, participants agreed that they were warned in an appropriate way (M = 4.05 on a 5-point Likert scale from 1 to 5) and with sufficient time to retake control safely (M = 4.01 on a 5-point Likert scale from 1 to 5). Participants felt quite safe while the takeover from machine to human was performed (M = 3.87 on a 5-point Likert scale from 1 to 5).

3.5.7.1. Questionnaire and interview results on (un-) planned transfers

Takeover questionnaire adapted from L3 Pilot (TOC machine to human). Figure 3.18 provides the results for the five items from the L3 Pilot questionnaire. Participants could indicate their agreement with the items from "*1 – strongly disagree*" to "*5 – strongly agree*", with "*3 – neutral*". Mean values for the statement "It was obvious to me why all takeover requests occurred" ranged from M = 3.55 to M = 3.86 indicating a tendency towards *agreement* for all three drives. The statistical analysis revealed no significant effect (F(1.45, 101.63) = 2.68, p = .072, η_p^2 = .04), while the post-hoc tests revealed a significant difference between drive 3 with SB and the NDRT (M = 3.55, SD = 1.09) and drive 4 with SB, the NDRT and the close approach (3.86, SD = 1.00). Hence, after drive 4 it seemed to be more obvious to the participants why the TOR occurred compared to drive 3. The participants' agreement after drive 2 with CM (*M* = 3.68, *SD* = 1.19) was in the middle of the two other drives and did not differ significantly from the other two drives. With respect to the item "I would have liked more information about why a takeover request was triggered" the mean values were quite similar and ranged from *M* = 2.70 to *M* = 2.94 indicating *disagreeing* answers with a clear tendency towards a *neutral* opinion for all three drives. In line with this, the statistical analysis did not reveal a significant effect ($F(2, 146) = .16$, $p = .855$, $\eta_p^2 = .00$). Hence, similarly for all three drives, the participants did not seem to want more information about the reasons why a TOR was triggered. Hence, the provided information seemed sufficient for the participants. The mean values for the item "During the takeover I always felt safe" ranged from *M* = 3.76 to *M* = 4.01 indicating a strong tendency towards *agreement*. No significant effect was found ($F(2, 140) = 2.33$, $p = .101$, $\eta_p^2 = .03$). While participants seemed to feel slightly less safe during the takeover in drive 4 (*M* = 3.76, *SD* = 1.02) compared to drive 2 (*M* = 4.00, *SD* = .85) and drive 3 (*M* = 4.01, *SD* = .89), which was likely related to the close approach in drive 4, the difference was not significant. Hence, the participants seemed to feel relatively safe during the takeover from machine to human during all three drives. With respect to the item "When Mediator asked me to retake control, I was warned in an appropriate way" the mean values ranged from *M* = 3.87 to *M* = 4.17 indicating *agreement*. Here, the statistical analysis revealed a significant effect, $F(1.84, 129.03) = 3.52$, $p = .032$, $\eta_p^2 = .05$. Post-hoc tests revealed a significant difference between drive 3 (*M* = 3.87, *SD* = 1.03) and drive 4 (*M* = 4.17, *SD* = .79), which indicates that Mediator's warning to retake control seemed to feel more appropriate for the participants during drive 3 than drive 4. Even though a similar difference can be seen between drive 3 and drive 2 (*M* = 4.13, *SD* = .82), this was not significant. It seems likely that the execution of the NDRT distracted the participants during drive 3, when they first started executing it. Through focusing on the NDRT, the participants might have missed some of the information Mediator provided during the retake. During drive 4 it seems likely that participants were monitoring Mediator more closely after they experienced the close approach or were already done with the NDRT, wherefore they might have missed less of the information and felt more appropriately informed. Lastly, the mean values for the item "When Mediator asked me to retake control, I was warned and given enough time to do so safely" ranged from *M* = 3.90 to *M* = 4.14 indicating *agreement* for all three drives. The statistical analysis revealed no significant effect ($F(2, 140) = 2.08$, $p = .129$, $\eta_p^2 = .03$). Hence, for all three drives similarly participants seemed to feel that they were warned and given enough time to safely retake control.

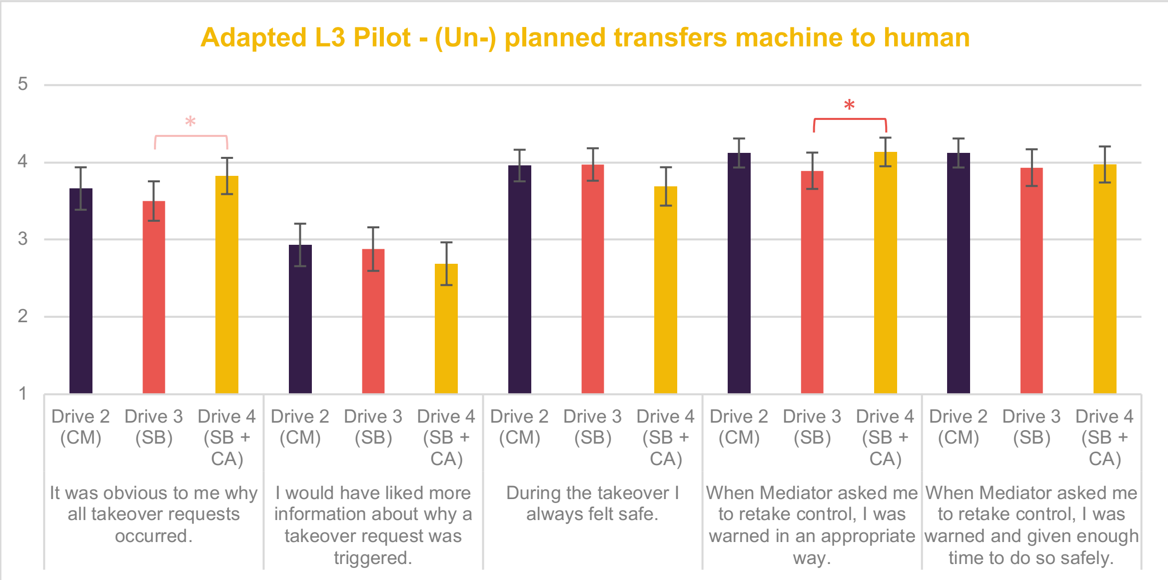


Figure 3.18 (Un-)planned transfers machine to human (Adapted L3 Pilot). Scale range: 1 – strongly disagree, 5 – strongly agree; Error Bars: 95th-CI.

(Un-) planned transfer-related items (self-designed). Figure 3.19 provides the results for the three self-designed items. Participants could indicate their agreement with the items from "*1 – strongly disagree*" to "*5 – strongly agree*", with "*3 – neutral*". Mean values for the statement "It was very easy for me to understand Mediator's instructions to take over the driving task" ranged from M = 4.37 to M = 4.58 indicating *agreement* with a tendency towards *strongly agree* for all three drives. The statistical analysis was not significant, $F(2, 140) = 2.99$, $p = .054$, $\eta_p^2 = .04$. Overall participants indicated that they understood the instructions quite well.

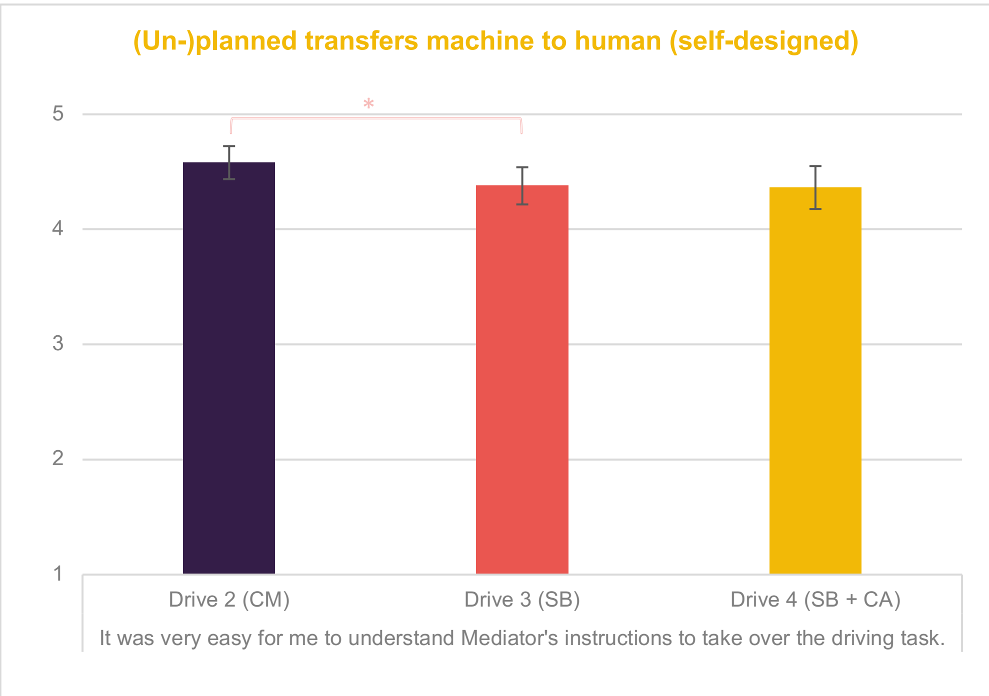


Figure 3.19 (Un-)planned transfers machine to human (self-designed). Scale range: 1 – strongly disagree, 5 – strongly agree; Error Bars: 95th-CI.

Results from the interviews. Nine participants directly referred to the transitions from automated to manual driving. Four participants (5% of all drivers) stated that the time to take-over the driving task as a bit too short. Three (4% of all drivers) complained that the take-over message was not clear enough. Two drivers stated that the message was clear and the transition was smooth.

3.5.8. RQ8: Observable driver features as indicators for uncomfortable driving situations

Research question 8 "*Which observable driver features are indicators for uncomfortable driving situations?"* was assessed by an exploratory approach. To identify observable driver features, the close approach situation in condition 4 (Figure 3.20) was analysed from meter 3,400 (sign "reduction to one lane in 800m" visible) until meter 7,300 (end of construction zone). In condition 4 all participants were instructed beforehand to work on the laptop (reading and answering the email) in the automated driving mode.

Figure 3.20 Close approach situation to the end of the traffic jam in automated mode (SB) with laptop work.

Basis of the exploratory analyses was a visual inspection and classification of behaviour using synchronized and playable data containing driving parameters (speed, distance, pedals, steering wheel angle, time-to-collision etc.), automation status (automation active, manual takeover), HMI (messages) as well as data on driver's behaviour such as head tracking and hands-on detection (example screenshot in Figure 3.21).

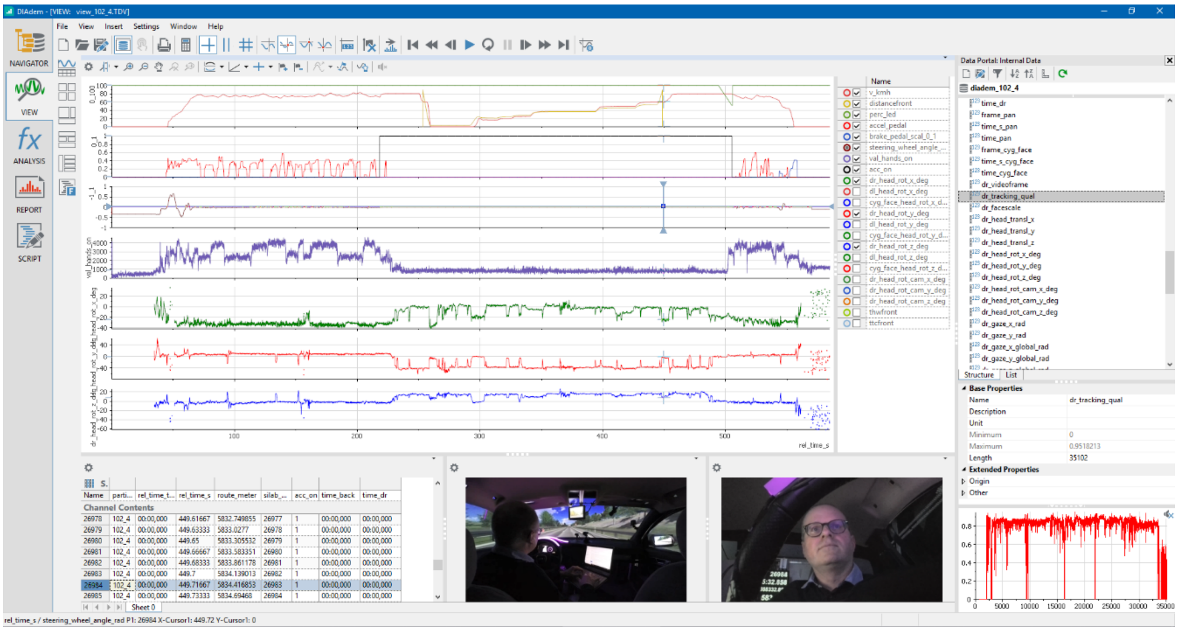


Figure 3.21 Example screenshot of playable synchronized video and measurement data for one person (Software NI DIADEM).

For each of the 74 participants it was analysed

1. if there were TOC from automated to manual driving mode,
2. if, when and how the participants may have interrupted and restarted laptop work until reaching the end of the construction zone,
3. if there was an observable emotional reaction (surprised / scared) in facial expressions and
4. how the glance behaviour towards the laptop and the approach situation evolved during the approach.

Based on these classifications, an explorative typology of behavioural reactions was created (Table 3.26). The type of reactions showed a wider variety than initially expected.

Behavioural types 1 to 4 (40 of 74 participants = 54%) did not take over manual control and left the automation always on:

- Type 1 (9%) did not show any sign for noticing the close approach. There were no glances towards the road, the work on the laptop was never interrupted and no other behavioural reactions could be observed.
- Type 2 (13%) showed short glances ahead to the approach situation. However, the work on the laptop was not interrupted and no signs for being surprised or scared could be noted.
- Type 3 (19%) showed slightly surprised and/or scared facial expressions and observed the situation longer by interrupting the work on the laptop.
- Type 4 (12%) was markedly surprised and/or scared showing corresponding facial reactions, sometimes combined with gestures such as throwing the hands in the air or covering the eyes (however, without manual intervention at pedals or steering wheel).

Types 5 and 6 (34 of 74 participants = 46%) took over manual control by pressing the brake pedal and placing the hands on the steering wheel:

- Type 5 (31%) showed clear signs for being surprised and/or scared in facial expressions. Laptop work was interrupted to take over manual driving control by pressing the brake pedal and placing the hands immediately on the steering wheel.
- Type 6 (15%) observed the approach situation already longer beforehand with an early interruption of laptop work. There were no signs for being surprised or scared and the takeover of manual control was planned and controlled.

To sum up, the behavioural reactions in this uncomfortable close approach situation differed greatly between the participants. Reactions ranged from not noticing at all up to almost panic reactions with sometimes rather uncontrolled manual interventions. One participant crashed into the truck due to simultaneously pressing the brake pedal (deactivation of automation) and the accelerator pedal. Having the experience of a smooth approach in condition 3 (with not even one manual takeover), some participants mentioned in the interviews that they relied on the automation also in condition 4. Despite the clear result, that automation should always avoid such situations, some early behavioural indicators for beginning discomfort could be the stop of NDRT engagement and concentrated observation of situations on the road. However, this behaviour per se cannot be generalized as general sign for uncomfortable automated driving situations, but needs to be interpreted in the context of the current driving situation.

Table 3.26 Typology of behaviour in uncomfortable automated approach situation

Behaviour type No.	Number of participant	Manual takeover	Description of behaviour type
1	7	No	Situation not noticed, no glances on the road, no observable reaction, uninterrupted continuous laptop work
2	10	No	Short glances to the approach situation on the road but no interruption of laptop work
3	14	No	Slightly surprised / scared, interruption of laptop work, observation of the situation
4	9	No	Markedly surprised / scared, interruption of laptop work, observation of the situation
5	23	Yes	Markedly surprised / scared, interruption of laptop work, manual takeover by pressing the brake pedal
6	11	Yes	Not surprised / scared, longer observation of the situation with early interruption of laptop work, planned and controlled manual takeover by pressing the brake pedal
Sum:	**74**		

3.5.9. RQ9: Eyes-off road time

Eyes-off road time can be an indicator on how much the driver is disengaged from the driving task, for instance, while performing a NDRT revealing driver's distraction. With regard to that, the research question 9 "*How good can eyes-off road time be detected by a camera system?"* was analysed quantitatively by a machine learning approach, i.e. a Bayes classifier based on head tracking.

The approach estimates glance directions based on head movements, and thus, only requires head tracking without the need of eye tracking to get more robust tracking data. Therefore, the drivers' glances during the execution of a NDRT, i.e. the instructed work on the laptop in condition 3 and 4, were annotated. These annotated glances were then used as output labels to train a

Bayes classifier for the eyes-off road detection with head position and rotation of the driver as input features. Head position and rotation were extracted by integrating the head tracking software FaceAPI from SeeingMachines and applied on replays of the recorded camera data in real-time at 30 fps in VGA resolution.

Figure 3.22 illustrates the labelled glances. Hence, *ahead* refers to eyes-on road whereas *laptop* refers to eyes-off road. Figure 3.23 shows the installed face camera position capable for head tracking whether the driver is looking ahead or to the laptop.

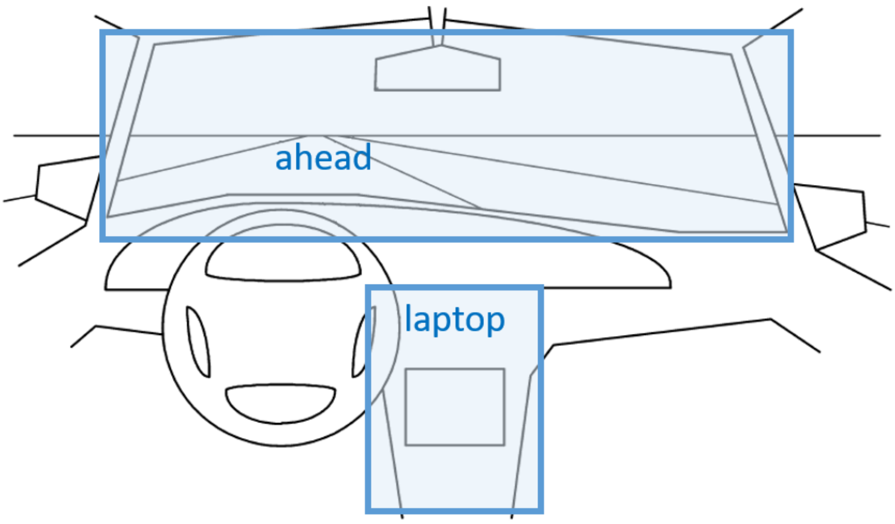


Figure 3.22 Labelled glances for eyes-off road detection.

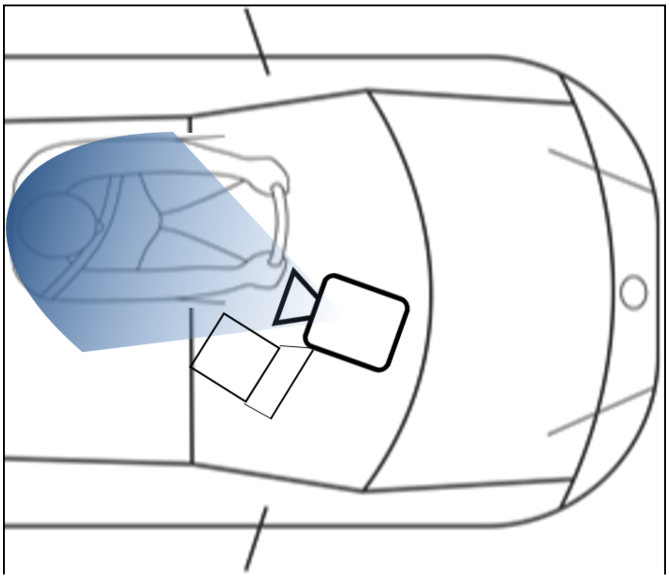


Figure 3.23 Driver face camera position for head tracking.

The input features for the Bayes classifier to determine the driver's glance direction are described in Figure 3.24. That are for each video frame the head position and rotation relative to the driver face camera on the x-, y-, and z-axis respectively.

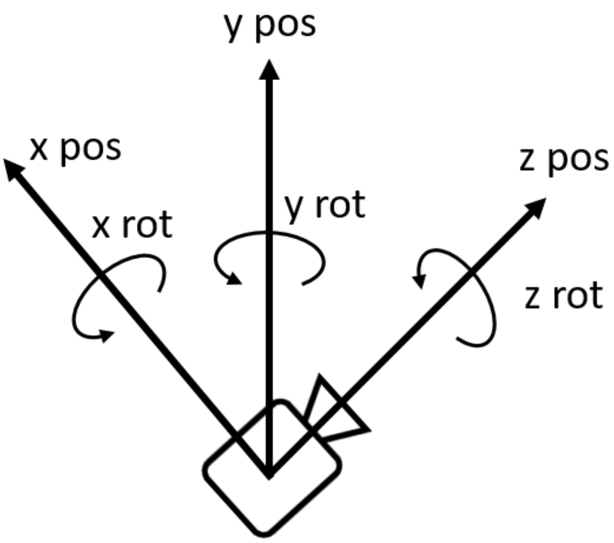


Figure 3.24 Input features of Bayes classifier from head tracking.

The performance of the Bayes classifier was evaluated in measures defined in signal detection theory, i.e. the hit and false alarm rates were calculated according to Table 3.27. These rates (see Figure 3.25) of a *hit rate* = .92 and a *false alarm rate* = .12 calculated from a leave-one-out cross-validation separated by participant result in a decent sensitivity index *d'* = 2.53 for the camera-based eyes-off road detection.

Table 3.27 Measures of signal detection theory, i.e. calculation of hit and false alarm rates

Label	Eyes-on road detected	Eyes-off road detected
Eyes-off road present	*Miss*	*Hit*
Eyes-on road present	*Correct rejection*	*False alarm*

Figure 3.25 Results of leave-one-out cross-validation for the camera-based eyes-off road detection; hit and false alarm rates.

Times without sufficient head tracking quality (e.g. due to the participant wearing a face mask because of the corona pandemic) were excluded. In summary, 1h 15min eyes-on road time (87%) and 5h 13min eyes-off road time (92%) out of a data set of 1h 26min of eyes-on road time and 5h 42min eyes-off road time over all 74 participants were successfully detected. Thus, the proposed approach to train a Bayes classifier on features of head tracking supports the feasibility of a camera-based system for eyes-off road detection.

3.5.10. RQ10: Safety impact of Mediator

Research question 10 “What is the safety impact of the Mediator system?” can partly be answered by the evaluation results of the driving simulator study and will be addressed in MEDIATOR WP4 “Impact and recommendations” as well. Results revealed that 86% of the participants expected an increase in safety, for instance, in terms of decreased number of accidents (85% agreed) or accidents’ severity (82% agreed) when driving with Mediator active.

In MEDIATOR deliverable 4.1 the safety impact will be addressed in more detail. General assumptions for the safety impact assessment are the following:

- Mediator is a system built on top of existing driving automation systems.
- Mediator is designed to:
  - Give recommendations for using automated driving based on:
    - Driver Fitness / Unfitness (e.g., driver states like distraction, fatigue, discomfort)
    - Automation Fitness / Unfitness (e.g., reaching end of operational design domain, reliability of sensors)
    - Environment (e.g., weather, traffic jams, highway exits)
  - Increase (for example):
    - Correct usage of automated driving systems
    - Predictability, quality and smoothness of takeover actions

        - Drivers‘ mode awareness
        - Safe driving by avoiding manual driving with impaired driver states
- Mediator is not mandatory, hence, drivers need to decide to:
    - Buy a vehicle equipped with Mediator
    - Activate Mediator on their trips
    - Follow the recommendations of Mediator
    - Not to override actions of the automated system initiated / recommended by Mediator

Based on these assumptions, TUC will concentrate on Mediator’s potential to reduce the burden of injury and identifying potential road-safety benefits of a large-scale implementation of Mediator. As a first step, a literature research will be conducted focusing on highly aggregated statistics on different types of accidents in Europe with different severity levels (slight, serious) that would cover different Mediator use cases, e.g. rear-end collisions in traffic jams). In a second step, specific assumptions regarding Mediator will be made, e.g., activated automation would prevent every rear-end collision in traffic jams (i.e. 100% efficiency) based on results from WP3 studies or derived from literature. Based on the percentage of users who would use Mediator (or the respective function) in their own vehicle (e.g. 88% in TUC simulator study; assumptions will be derived from the user studies in WP3) and assumptions of different penetration rates of the Mediator system (e.g. 90%), we could then calculate how many of the accidents would potentially be prevented. Further, based on existing publications, which estimate the total societal costs of minor to severe accidents in Euros (including for example treatment, pain compensation, etc.), the socio-economic savings can be calculated.

3.6. Summary and conclusions on comfort-related use cases

The driving simulator study at Chemnitz University of Technology focussed on comfort use cases with related TOC, simulated automation degradation and related TOC by the human driver, as well as driver characteristics. The study aimed primarily on the user’s evaluation of the Mediator system / HMI (i.e. acceptance, trust, usability, comfort, experience of TOC), including the analysis of driver characteristics effects such as age, gender and driving experience. It was expected that Mediator will be rated positively (e.g., high acceptance, trust, experienced comfort, perceived safety) and will be preferred compared to manual driving. Regarding the two automated driving modes, it was expected that driving in SB will be rated better than CM. Further, it was expected that the close approach will have a negative influence on drivers’ evaluation of Mediator (e.g., trust, acceptance, perceived safety, and perceived driving comfort).

Results revealed very high intention to use Mediator in future cars (88%) as well as very high driving comfort (90%), easiness to use (95%) and usability ratings (SUS score 85 out of 100) when driving with Mediator. Additionally, drivers’ expressed high acceptance (*M* = 1.21 on a scale ranging from -2 to +2) and trust (*M* = 5.5 on a scale ranging from 1 to 7). Further, 86% of the participants expected an increase in road safety when driving with Mediator.

On average, participants rated driving condition 3 (SB, NDRT execution, smooth approach to the rear end of the traffic jam) as the most favourable, followed by driving condition 2 (CM, smooth approach to the rear end of the traffic jam), driving condition 4 (SB, NDRT execution, close approach scenario) and driving condition 1 (manual driving). Hence, driving with Mediator and in automated driving mode was preferred compared to manual driving without Mediator. Further, driving in SB was preferred compared to driving in CM, although the close approach war rated less positively.

The condition with the close approach scenario led to a decreased rating for acceptance, trust, perceived comfort and perceived safety after the drive. Nevertheless, the overall evaluation of Mediator remained positive and was only slightly influenced by the close approach experience. The close approach scenario represents certain margins in which Mediator and the automated driving system might work in the future. As the results show, the situation can be rated as uncomfortable by drivers leading to decreased trust. Nevertheless, trust is important, otherwise participants cause dangerous situations by problematic interventions (i.e. evasive manoeuvres, pushing the acceleration instead of the brake pedal). To prevent such situations, transparency is very important (e.g., informing participants that the rear-end of the traffic jam was detected and a braking manoeuvre will be initiated).

In analysing the questionnaire data, great attention was paid to the influences of the driver characteristics on the participants' responses. Particularly the participants' age, gender, driving experience (i.e. annual mileage), as well as their affinity for technology and their general opinion regarding automating parts or the entire driving task were analysed. For this purpose, a balanced sample regarding age (19 to 75 years; balanced across the four age groups of the NHTSA) and gender (49% female) was recruited. Overall, the driver characteristics seemed to have relatively little or no systematic influences on the participants' responses. For instance, participants' gender seemed to have the least strong influence on their responses, since no noticeable, systematic differences were found between female and male participants. Similarly, no noticeable, systematic differences were found regarding the participants' driving experience. Only with regard to the participants' age, a few, more pronounced differences were found. For instance, with increasing age, participants seemed to rate Mediator's behaviour as less unexpected and more appropriate, and they seemed to want to use Mediator more during their everyday trips. Moreover, the oldest age group expected the highest increase in the possibility to drive longer (more years) and trusted Mediator to drive more compared to the younger age groups. Lastly, the oldest age groups expected greater increases in safety as well as greater decreases in number of accidents and their severity, as well as felt slightly safer than the other age groups. However, these differences were still relatively small and the statistical analyses for the acceptance scale and trust in automation scale did not reveal any significant effects of the driver characteristics.

In addition, the participants' affinity for technology and their general opinion were correlated with their responses. Overall, participants' with higher affinity for technology and a more positive general opinion towards automating parts or the entire driving task evaluated Mediator and its behaviour more positively. For instance, participants with a higher affinity for technology were more inclined to buy Mediator, to recommend it to others, to use it during everyday trips and to do other activities while Mediator was active. Moreover, participants with a higher affinity for technology also indicated higher trust in and acceptance of Mediator throughout the entire experiment. However, no statistically significant influences of the participants' affinity for technology were found regarding the acceptance scale or the trust in automation scale. Similarly, participants with a more positive opinion towards automating parts or the entire driving task, found Mediator's behaviour to be less unexpected, felt more comfortable, were more likely to buy Mediator, recommend it, use it if it was in their car, use it during everyday trips and select further destinations. Additionally, the participants with a more positive general opinion also showed higher trust in and acceptance of Mediator throughout the entire experiment.

Overall, the relatively small differences between the age groups as well as the overall lack of systematic differences between the participants' gender and their driving experience seem to indicate that Mediator follows the "Design for All" principle, since it does not systematically discriminate certain groups (i.e. female vs. male participants). Nonetheless, it needs to be taken into account that the sample had high values for ATI, high acceptance and a positive general opinion regarding vehicle automation, and high initial trust in Mediator.

During the interviews, participants mentioned several advantages of using Mediator, like support and relaxation for drivers (51%), possibility of performing secondary tasks (37%), providing a general positive experience (24%), increasing (perceived) safety (23%), and the opportunity to get familiar with a new ADAS and built-up trust (32%). Participants also mentioned challenges like trust in the system in terms of over-trust, not enough trust or decreased trust due to unexpected or even wrong decisions by Mediator (45%), reduced concentration / awareness, boredom or even fatigue (22%), the challenge to give up control of the driving task (12%), and too much distraction due to NDRTs (5%). One participant mentioned the danger of losing important driving skills and another one reported that he/she completely lost his/her sense of time.

Regarding the external validity of the results, it needs to be taken into account that that the sample had very high values for ATI, high acceptance and a positive general opinion regarding vehicle automation, and high initial trust in Mediator. In addition, the study was conducted in a driving simulator. Hence, results reflect drivers' evaluation of an "optimal working" Mediator system (e.g., no unplanned HMI reactions, no false positives or false negatives) and, therefore, represent the highest possible potential of the Mediator system. Further, the simulated environment might have influenced participants' evaluation of perceived safety and risk perception. Therefore, some results need to be interpreted carefully taking the simulated environment into account.

4. References

Arefnezhad, S., Eichberger, A., & Koglbauer, I. V. (2022). Effects of Automation and Fatigue on Drivers from Various Age Groups. Safety, 8(2), 30.

Ahlström, C., Zemblys, R., Jansson, H., Forsberg, C., Karlsson, J., & Anund, A. (2021). Effects of partially automated driving on the development of driver sleepiness. Accident Analysis & Prevention, 153, 106058.

Åkerstedt T, & Gillberg M. (1990). Subjective and objective sleepiness in the active individual. International Journal of Neuroscience, 52, 29–37.

Atchley, P., & Chan, M. (2011). Potential benefits and costs of concurrent task engagement to maintain vigilance: A driving simulator investigation. Human factors, 53(1), 3-12.

Bakker, B., Knauss, A., Rodrigues de Campos, G., Tanov, N., Andersson, S., Mano, D., Athmer, C., Enhuber, S.M., Beggiato, M., & Spaan, M. (2022). Development of integrated lab prototypes, Deliverable D2.9 of the H2020 project MEDIATOR.

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of usability studies, 4(3), 114-123.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. International Journal of Human-Computer Interaction, 24(6), 574–594. https://doi.org/10.1080/10447310802205776

Bartels, R., & Peçanha, T. (2020). HRV: A Pythonic package for heart rate variability analysis. Journal of Open Source Software, 5(51), 1867.

Bellem, H., Thiel, B., Schrauf, M., & Krems, J. F. (2018). Comfort in automated driving: An analysis of preferences for different automated driving styles and their dependence on personality traits. Transportation Research Part F: Traffic Psychology and Behaviour, 55, 90–100. doi:10.1016/j.trf.2018.02.036.

Borowsky A., & Oron-Gilad, T. (2013). Exploring the Effects of Driving Experience on Hazard Awareness and Risk Perception via Real-Time Hazard Identification, Hazard Classification, and Rating Tasks. Accident Analysis, & Prevention, 59, 548–565.

Borowsky, A., Oron-Gilad, T., Chasidim, H., Ahlström, C., Karlsson J.G., Bakker, B., … Christoph, M. (2020). Behavioural Markers for Degraded Human Performance. Deliverable D1.2 of the H2020 project MEDIATOR.

Brooke, J. (1996). SUS: a „quick and dirty" usability scale. In P.W. Jordan, B. Thomas, B.A. Weerdmeester, and I.L. McClelland (Eds.). Usability Evaluation in Industry, 189-194. London: Taylor and Francis.

Charness, N., Yoon, J. S., Souders, D., Stothart, C., & Yehnert, C. (2018). Predictors of attitudes toward autonomous vehicles: The roles of age, gender, prior knowledge, and personality. Frontiers in psychology, 9, 2589.

Choi, J. K., & Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. International Journal of Human-Computer Interaction, 31(10), 692-702.

Christoph, M., Cleij, D., Ahlström, C., Bakker, B., Beggiato, M., Borowsky, A., ..., Van Nes, C.N (2019). Mediating between human driver and automation: state-of-the art and knowledge gaps. Deliverable 1.1 of the H2020 project MEDIATOR.

Cleij, D., Bakker, B., Borowsky, A., Christoph, M., Fiorentino, A., van Grondelle, E., & van Nes, N. (2020). Mediator System and Functional Requirements, Deliverable D1.4 of the H2020 project MEDIATOR.

ERTRAC. (2019). Connected Automated Driving Roadmap. European Road Transport Research Advisory Council. Retrieved from https://www.ertrac.org/uploads/documentsearch/id57/ERTRAC-CAD-Roadmap-2019.pdf.

Fiorentino, A., Ahlström C., S. Andersson, Toffetti A., (2023). On-road evaluations of the vehicle-integrated Mediator system, Deliverable D3.4 of the H2020 project MEDIATOR.

Fisher, D. L., Lohrenz, M., Moore, D., Nadler, E. D., & Pollard, J. K. (2016). Humans and intelligent vehicles: The hope, the help, and the harm. IEEE Transactions on Intelligent Vehicles, 1(1), 56-67.

Franke, T., Attig, C. & Wessel, D. (2019). A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. International Journal of Human–Computer Interaction, 35(6), 456–467. doi: 10.1080/10447318.2018.1456150.

Gershon, P., Ronen, A., Oron-Gilad, T., & Shinar, D. (2009). The effects of an interactive cognitive task (ICT) in suppressing fatigue symptoms in driving. Transportation research part F: traffic psychology and behaviour, 12(1), 21-28.

Ginsburg, A. P. (2003). Contrast sensitivity and functional vision. International ophthalmology clinics, 43(2), 5-15.

Golbabaei, F., Yigitcanlar, T., Paz, A., & Bunker, J. (2020). Individual Predictors of Autonomous Vehicle Public Acceptance and Intention to Use: A Systematic Review of the Literature. Journal of Open Innovation: Technology, Market, and Complexity, 6(4), 106.

Greenwood, P. M., Lenneman, J. K., & Baldwin, C. L. (2022). Advanced driver assistance systems (ADAS): demographics, preferred sources of information, and accuracy of ADAS knowledge. Transportation research part F: traffic psychology and behaviour, 86, 131-150.

Grondelle, E.D. van, Zeumeren, I. van, Bjorneseth F., Borowsky, A., Chandran, T., Cleij, D.,…Christoph, M. (2021). HMI Functional Requirements, D1.5 of the H2020 project MEDIATOR.

Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. Human factors, 31(5), 519–537. https://doi.org/10.1177/001872088903100503.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Advances in Psychology. Human Mental Workload (p. 139–183).

Hecht, T., Zhou, W., & Bengler, K. (2022). How to keep drivers attentive during Level 2 automation? Development and evaluation of an HMI concept using affective elements and message framing. Safety, 8(3), 47.

Hergeth, S., Lorenz, L., & Krems, J. F. (2016). Prior Familiarization With Takeover Requests Affects Drivers' Takeover Performance and Automation Trust. Human Factors: The Journal of the Human Factors and Ergonomics Society, 59(3), 457–470. doi:10.1177/0018720816678714

Hockey G. R. (1997). Compensatory control in the regulation of human performance under stress and high workload; a cognitive-energetical framework. Biological psychology, 45(1-3), 73–93. https://doi.org/10.1016/s0301-0511(96)05223-4

Mediator

Hoff, K.A., & Bashir, M. (2015). Trust in Automation. Human Factors, 57(3), 407–434. doi:10.1177/0018720814547570.

Horswill, M.S., & McKenna, F.P. (2004). Drivers' hazard perception ability: Situation awareness on the road. A cognitive approach to situation awareness: Theory and application, 155-175.

Jian, J. Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. International Journal of Cognitive Ergonomics, 4(1), 53–71.

Jarosch, O., Bellem, H., & Bengler, K. (2019). Effects of task-induced fatigue in prolonged conditional automated driving. Human factors, 61(7), 1186-1199.

Kircher, K., & Ahlström, C. (2017). The driver distraction detection algorithm AttenD. In Driver distraction and inattention (pp. 327-348). CRC Press.

Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance decrement and passive fatigue caused by monotony in automated driving. Procedia Manufacturing, 3, 2403-2409.

Kraft, A. K., Naujoks, F., Wörle, J., & Neukum, A. (2018). The impact of an in-vehicle display on glance distribution in partially automated driving in an on-road experiment. Transportation research part F: traffic psychology and behaviour, 52, 40-50.

Krishnan, A., Samuel, S., Yamani, Y., Romoser, M. R. E., & Fisher, D. L. (2019). Effectiveness of a strategic hazard anticipation training intervention in high risk scenarios. Transportation Research Part F: Traffic Psychology and Behaviour, 67, 43–56. https://doi.org/10.1016/j.trf.2019.10.004

Kurpiers, C., Biebl, B., Mejia Hernandez, J., & Raisch, F. (2020). Mode Awareness and Automated Driving—What Is It and How Can It Be Measured? Information, 11(5), 277.

Lee, J., Hirano, T., & Itoh, M. (2020). Making Passenger Conversation in Partial Driving Automation: Effects of Relationship Between Driver and Passenger on Driver Fatigue and Driving Performance. IIEEE International Conference on Systems, Man, and Cybernetics, 1718- 1722

May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. Transportation research part F: traffic psychology and behaviour, 12(3), 218-224.

Mayring, P. (2004). Qualitative content analysis. A companion to qualitative research, 1(2), 159-176.

MEDIATOR (2023). MEDIATOR Driving Simulator Study Germany: Questionnaire Data User Evaluation HMI. https://doi.org/10.5281/zenodo.7638299.

Merat, N., Seppelt, B., Louw, T., Engström, J., Lee, J. D… & Keinath, A. (2019). The "Out-of-the-Loop" concept in automated driving: proposed definition, measures, and implications. Cognition, Technology and Work, 21(1). https://doi.org/10.1007/s10111-018-0525-8

Metz, B., Rösener, C., Louw, T., Aittoniemi, E., Bjorvatn, A., Wörle, J., … (2019). Evaluation methods. Deliverable 3.3 of the H2020 project L3Pilot. Retrieved from https://l3pilot.eu/download/

Naujoks, F., Purucker, C., & Neukum, A. (2016). Secondary task engagement and vehicle automation – comparing the effects of different automation levels in an on-road experiment. Transp. Res. Part F: Psychol. Behav. 38, 67–82.

NHTSA (2016). Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices. National Highway Traffic Safety Administration. Washington, DC, USA.

Oron-Gilad, T., Ronen, A., & Shinar, D. (2008). Alertness maintaining tasks (AMTs) while driving. Accident Analysis & Prevention, 40(3), 851-860.

Oron-Gilad, T., & Hancock, R. A. (2005). Road environment and driver fatigue. In Driving Assesment Conference (Vol. 3, No. 2005). University of Iowa.

Pereira Cocron, M., Vallejo, A., Delgado, M., Wilbrink, M., Anund, A., Krupenia, S., … (2019). Evaluation Framework. Deliverable 7.3 of the H2020 Project ADAS&ME. Retrieved from https://www.adasandme.com/dissemination/public-deliverables/

Petermeijer, S. M., Cieler, S., & De Winter, J. C. (2017). Comparing spatially static and dynamic vibrotactile take-over requests in the driver seat. Accident analysis & prevention, 99, 218-227.

Post, J.M.M., Ünal, A.B., & Veldstra, J.L. (2020). Deliverable 1.2. Model and guidelines depicting key psychological factors that explain and promote public acceptability of CAV among different user groups. SUaaVE. Retrieved from http://www.suaave.eu/wp-content/uploads/sites/17/2021/02/SUaaVE_WP1_D1.2_20200930_V100.pdf

Reagan, I. J., Teoh, E. R., Cicchino, J. B., Gershon, P., Reimer, B., Mehler, B., & Seppelt, B. (2021). Disengagement from driving when using automation during a 4-week field trial. Transportation research part F: traffic psychology and behaviour, 82, 400-411.

Robertson, R. D., Meister, S. R., Vanlaar, W. G., & Hing, M. M. (2017). Automated vehicles and behavioural adaptation in Canada. Transportation Research Part A: Policy and Practice, 104, 50-57.

Rodarius, C., Duflis, J., Fahrenkrog, F., Rösener, C., Várhelyi, A., Fernandez, R., … (2015). Test and Evaluation Plan. Deliverable 7.1 from the FP7 Project AdaptIVe. Retrieved from https://www.adaptive-ip.eu/index.php/deliverables_papers.html

Sarter, N.B., & Woods, D.D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. Human factors, 37(1), 5-19.

Schaefer, K.E., Chen, J.Y.C., Szalma, J.L., & Hancock, P.A. (2016). A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Understanding Autonomy in Future Systems. Human Factors, 58 (3), 377-400. doi:10.1177/0018720816634228.

Schömig, N., Hargutt, V., Neukum, A., Petermann-Stock, I., & Othersen, I. (2015). The interaction between highly automated driving and the development of drowsiness. Procedia Manufacturing, 3, 6652-6659.

Schömig, N., & Metz, B. (2013). Three levels of situation awareness in driving with secondary tasks. Safety Science, 56, 44-51.

Shahid, A., Wilkinson, K., Marcu, S., & Shapiro, C. M. (2011). Karolinska sleepiness scale (KSS). In STOP, THAT and one hundred other sleep scales (pp. 209-210). Springer, New York, NY.

Shupsky, T., Morales, K., Baldwin, C., Hancock, P., Greenlee, E. T., Horrey, W. J., & Klauer, C. (2020). Secondary Task Engagement During Automated Drives: Friend and Foe?. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 64, No. 1, pp. 1926-1930). Sage CA: Los Angeles, CA: SAGE Publications.

Smith, S., Koopmann, J., Rakoff, H., Peirce, S., Noel, G., Eilbert, A., & Yanagisawa, M. (2018). Benefits Estimation Model for Automated Vehicle Operations: Phase 2 Final Report. U.S. Department of Transportation. Retrieved from https://rosap.ntl.bts.gov/view/dot/34458/dot_34458_DS1.pdf?

Solís-Marcos, I., Ahlström, C., & Kircher, K. (2018). Performance of an additional task during level 2 automated driving: an on-road study comparing drivers with and without experience with partial automation. Hum. Factors 60 (6), 778–792.

Song, W., Woon, F. L., Doong, A., Persad, C., Tijerina, L., Pandit, P., ... & Giordani, B. (2017). Fatigue in younger and older drivers: effectiveness of an alertness-maintaining task. Human factors, 59(6), 995-1008.

Techer, F., Ojeda, L., Barat, D., Marteau, J.-Y., Rampillon, F., Feron, S., & Dogan, E. (2019). Anger and highly automated driving in urban areas: The role of time pressure. Transportation Research Part F: Traffic Psychology and Behaviour, 64, 353–360. doi:10.1016/j.trf.2019.05.016.

Van der Laan, J. D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. Transportation research. Part C, Emerging technologies, 5(1), 1-10. http://www.hfes-europe.org/accept/accept.htm.

Van Erp, J. B., Toet, A., & Janssen, J. B. (2015). Uni-, bi-and tri-modal warning signals: Effects of temporal parameters and sensory modality on perceived urgency. Safety science, 72, 1-8.

Vlakveld, W., van Nes, N., de Bruin, J., Vissers, L., & van der Kroft, M. (2018). Situation awareness increases when drivers have more time to take over the wheel in a Level 3 automated car: A simulator study. Transportation Research Part F: Traffic Psychology and Behaviour, 58, 917–929. https://doi.org/10.1016/j.trf.2018.07.025

Zangi, N., Srour-Zreik, R., Ridel, D., Chasidim, H., & Borowsky, A. (2022). Driver distraction and its effects on partially automated driving performance: A driving simulator study among young-experienced drivers. Accident Analysis & Prevention, 166, 106565. https://doi.org/10.1016/j.aap.2022.106565

5. Appendices

5.1. Appendices BGU driving simulator studies on fatigue and distraction

5.1.1. Questionnaire demographics

Age:	____________________ years

Gender:	☐ Female ☐ Male ☐ Other

What is your current employment status?	☐ Employed full-time ☐ Employed part-time ☐ Self-employed ☐ Homemaker ☐ Unemployed ☐ Retired ☐ Student ☐ Other

5.1.2. Questionnaire mobility / ADAS

Do you have a car available for your use?	☐ Yes, (nearly) always ☐ Yes, sometimes ☐ No or hardly ever

On average, how often do you drive a car?	☐ (Nearly) every day ☐ 3-5 days / week ☐ 1-2 days / week ☐ Less often or never

On average, how many kilometres do you drive in 1 year?	____________________ km/year

For how long do you hold a vehicle driving license?	____________________ years

Please state if your current vehicle is equipped with the following systems:

	I have it and I use it	I have it but I don't use it	I don`t have it	Don`t know if I have it
Parking Assist System (A system that provides a camera view and/or auditory beeps to indicate how close you are to an object, while you are parking).	☐	☐	☐	☐
Self-parking Assist System (A system that controls the vehicle for parallel parking or reverse parking. Some of these systems control both steering and the throttle; others only control the steering and the driver presses the brake and throttle).	☐	☐	☐	☐
Cruise Control (A system that maintains vehicle speed while driving).	☐	☐	☐	☐
Adaptive Cruise Control (A system that automatically slows down or speeds up to keep a safe distance from a vehicle ahead).	☐	☐	☐	☐
Blind spot monitoring (A system that monitors the driver's left and right blind spots for other vehicles. Often, drivers receive a visual or audio alert whenever a vehicle is present).	☐	☐	☐	☐
Lane departure warning (A system that provides assistance with lane-keeping, by sounding warnings when the vehicle travels outside of the lane markings/boundaries).	☐	☐	☐	☐
Lane keeping assistance (A system that helps motorists to avoid inadvertently moving out of the intended driving lane).	☐	☐	☐	☐
Forward Collision Warning (A system that provides warnings for potential collisions with the vehicle in front).	☐	☐	☐	☐
Driver alert - sleepiness warning (A system that provides warnings when falling asleep).	☐	☐	☐	☐

5.1.3. Adoption and trust

What is your general opinion about functions in the vehicle that can automate parts or the entire driving task?

☐ Very negative	☐ Somewhat negative	☐ Neutral	☐ Somewhat positive	☐ Positive

Please indicate the extent to which you agree or disagree with the statements on automated vehicles below.

	Completely disagree	disagree	Slightly disagree	Neither disagree nor agree	Slightly agree	agree	Completely agree
The use of automated vehicles is acceptable.	☐	☐	☐	☐	☐	☐	☐
It is acceptable that a part of the traffic will consist of automated vehicles.	☐	☐	☐	☐	☐	☐	☐
It is acceptable that people will use automated vehicles.	☐	☐	☐	☐	☐	☐	☐
I would trust an automated vehicle to behave as intended.	☐	☐	☐	☐	☐	☐	☐
I trust that automated vehicles would correctly detect other road users.	☐	☐	☐	☐	☐	☐	☐
I trust the computer systems of automated vehicles cannot get hacked.	☐	☐	☐	☐	☐	☐	☐
Automated vehicles would meet my driving needs.	☐	☐	☐	☐	☐	☐	☐
Driving an automated vehicle would be convenient, since it would make my journeys more efficient.	☐	☐	☐	☐	☐	☐	☐
An automated vehicle would be safe.	☐	☐	☐	☐	☐	☐	☐
Automated vehicles would pose minimal risk to its driver and passengers.	☐	☐	☐	☐	☐	☐	☐
Automated vehicles would pose minimal risk to other road users.	☐	☐	☐	☐	☐	☐	☐

5.1.4. Acknowledgment of CM automation functionality

בחר\י את התשובה הנכונה ביותר בהנחה שהרכב נמצא במצב נהיגה אוטונומית רמה 2.

מה קורה למיקום הרכב הכביש כאשר הנהג לא נוגע בהגה?	☐ הרכב יסטה בפתאומיות מהנתיב כמו בנהיגה ידנית. ☐ לאחר 30 שניות הרכב יסטה מעט מהנתיב. ☐ כל עוד שולי הנתיב מסומנים על הכביש הוא ישמור על מיקומו בנתיב. ☐ הרכב יסטה בהדרגה מהנתיב.
מה קורה למהירות הרכב כאשר הנהג לא לוחץ על הדוושות?	☐ הרכב מתחיל להאיץ עד שמגיע למהירות 60 קמ"ש. ☐ הרכב שומר על המהירות שהוגדרה לו. ☐ הרכב מאט עד שהנהג ילחץ על דוושת הגז. ☐ הרכב שומר על מהירות קבועה תוך התחשבות במרחק מהרכב שמלפנים.

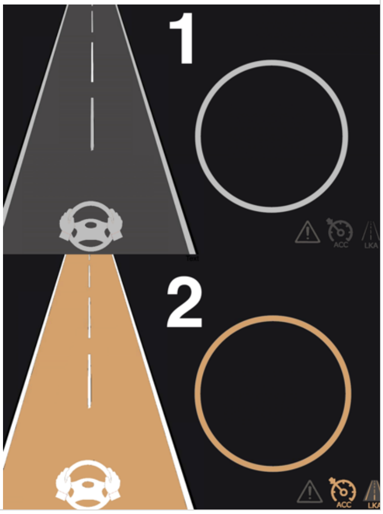
על מנת לעבור מנהיגה אוטומטית לידנית בסימולטור הנהיגה עליך (ניתן לבחור יותר מתשובה אחת):	☐ ללחוץ על דוושת הגז. ☐ ללחוץ על דוושת הברקס. ☐ לסובב את ההגה. ☐ ללחוץ על הכפתור באוטומציה מימן לנהג.
כיצד יתנהג הרכב שלך במידה והרכב מלפנים מאט?	☐ הרכב יגביר את מהירותו ויעקוף. ☐ הרכב ישמור על מהירות קבועה. ☐ הרכב יאט על מנת לשמור על מרחק קבוע מהרכב מלפנים. ☐ הרכב יעביר את הנהיגה למצב ידני.

5.1.5. Acknowledgment of Mediator system indication functionality

שאלון חווי מערכות MEDIATOR

לפניך מספר שאלות בנוגע לסמלים שונים המופיעים במערכת.

מה משמעות הסמל הבא? 	☐ אזהרה! אוטומציה פסקה! ☐ אזהרה! תרחיש מסוכן לפניך! ☐ אזהרה! תקלה ברכב, לא ניתן להמשיך נסיעה! ☐ אזהרה! אוטומציה החלה!
מה משמעות הסמל הבא? 	☐ התראת עייפות קלה. ☐ התראת עייפות כבדה. ☐ התראת עייפות קיצונית! יש לעצור את הרכב מיד! ☐ שימ\י לב לתנאי הדרך.
מה משמעות הסמל הבא?	☐ התראת עייפות קלה. ☐ התראת עייפות כבדה. ☐ התראת עייפות קיצונית! יש לעצור את הרכב מיד! ☐ תחנת ריענון בהמשך הדרך.

	
☐ מלבד הצבע אין כל הבדל. ☐ 1 - נהיגה אוטונומית פעילה. 2 - נהיגה אוטונומית כבויה. ☐ 1 - נהיגה אוטונומית כבויה. 2 - נהיגה אוטונומית פעילה. ☐ 1 - נהיגה אוטונומית פעילה. 2 - תקלה במערכת האוטונומית.	מה משמעות ההבדל בין המסכים הבאים? 

5.1.6. Workload

Please place an “X” along each scale at the point that best indicates your experience with the Mediator system.

Mental demand: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)?

Low		High

Physical demand: How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)?

Low		High

Temporal demand: How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred?

Lo w		Hig h

Performance: How successful do you think you were in accomplishing the goals of the task?

Lo w		Hig h

Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Lo w		Hig h

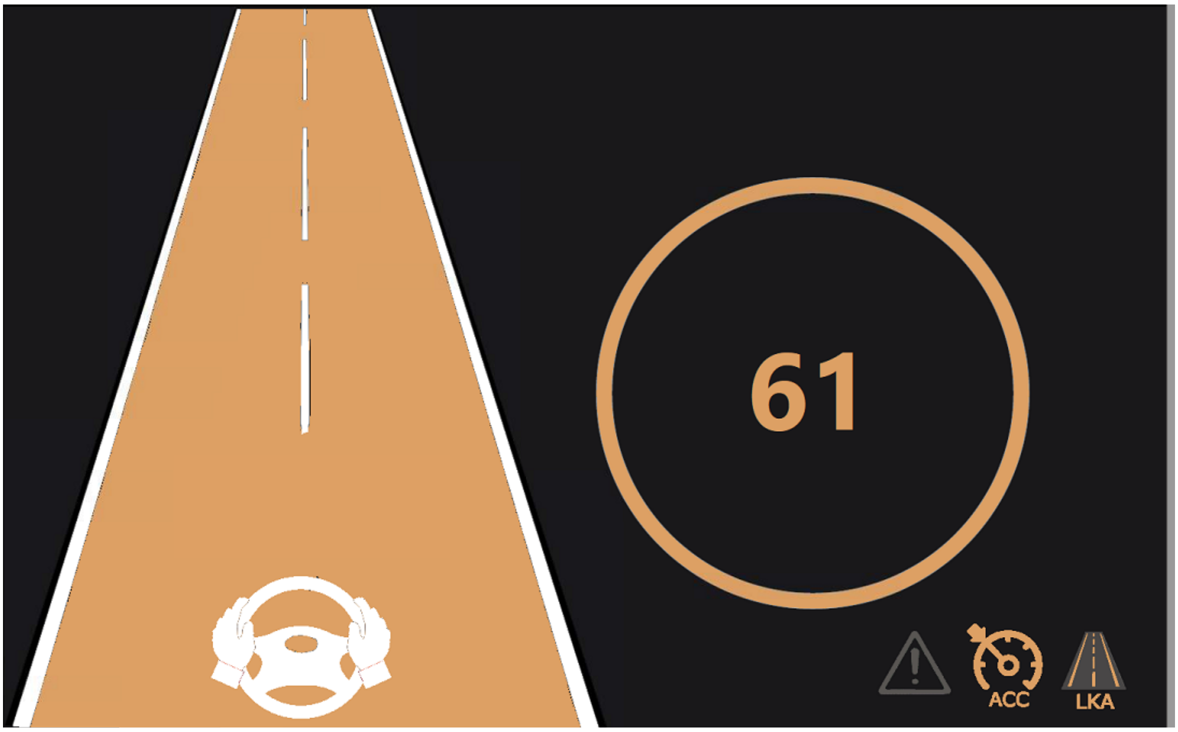
Frustration level: How insecure, discouraged, irritated, stressed and annoyed did you feel during the task?

Lo w		Hig h

5.1.7. Mediator acceptance and usability

הערכת חווית השימוש במערכות

חוו את דעתכם על המערכת הבאה (לוח מחוונים):



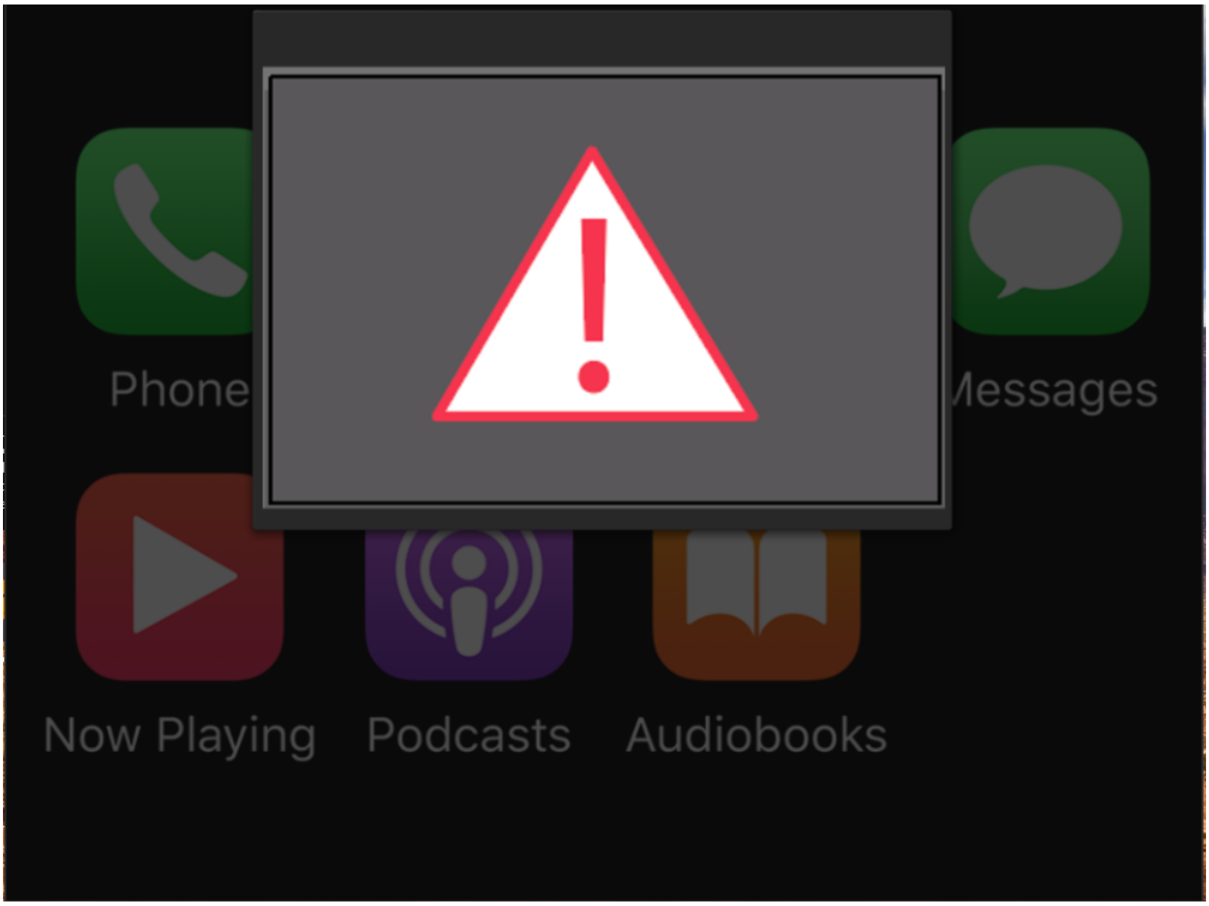
	1	2	3	4	5	
לא שימושי	☐	☐	☐	☐	☐	שימושי
נעים	☐	☐	☐	☐	☐	לא נעים
רע	☐	☐	☐	☐	☐	טוב
נחמד	☐	☐	☐	☐	☐	מעצבן
מועיל	☐	☐	☐	☐	☐	מיותר
מרגיז	☐	☐	☐	☐	☐	חביב
מסייע	☐	☐	☐	☐	☐	חסר ערך
לא רצוי	☐	☐	☐	☐	☐	רצוי
מעורר	☐	☐	☐	☐	☐	מרדים

לפניך רשימת היגדים להערכת הרושם לגבי המערכת. סמנ\י את המידה המתאימה ביותר המתארת את הרגשתך.

מאוד מסכימ\ה 5	4	3	2	מאוד לא מסכימ\ה 1	
☐	☐	☐	☐	☐	אני חושב\ת שהייתי רוצה להשתמש במערכת זו לעיתים תכופות.

מצאתי כי המערכת מסובכת ללא סיבה.	☐	☐	☐	☐	☐
חשבתי שהמערכת קלה לשימוש.	☐	☐	☐	☐	☐
אני חושב\ת שאזדקק לתמיכת איש טכני כדי שאוכל להשתמש במערכת זו.	☐	☐	☐	☐	☐
מצאתי כי הפונקציות השונות של המערכת היו מתואמות היטב.	☐	☐	☐	☐	☐
חשבתי כי היה יותר מידי חוסר עקביות במערכת זו.	☐	☐	☐	☐	☐
אני משער\ת כי רוב האנשים יהיו מסוגלים ללמוד להשתמש במערכת זו בקלות.	☐	☐	☐	☐	☐
מצאתי כי המערכת מאוד מסורבלת לשימוש.	☐	☐	☐	☐	☐
חשתי בטחון רב כאשר השתמשתי במערכת.	☐	☐	☐	☐	☐
עלי ללמוד הרבה דברים לפני שאוכל להשתמש במערכת זו.	☐	☐	☐	☐	☐

מה דעתך על המערכת הבאה (מערכת מולטימדיה):



	1	2	3	4	5	
לא שימושי	☐	☐	☐	☐	☐	שימושי
נעים	☐	☐	☐	☐	☐	לא נעים
רע	☐	☐	☐	☐	☐	טוב
נחמד	☐	☐	☐	☐	☐	מעצבן
מועיל	☐	☐	☐	☐	☐	מיותר
מרגיז	☐	☐	☐	☐	☐	חביב
מסייע	☐	☐	☐	☐	☐	חסר ערך
רצוי	☐	☐	☐	☐	☐	לא רצוי
מעורר	☐	☐	☐	☐	☐	מרדים

לפניך רשימת היגדים להערכת הרושם לגבי המערכת. סמנ\י את המידה המתאימה ביותר המתארת את הרגשתך.

	1 לא מאוד מסכים\ה	2	3	4	5 מאוד מסכים\ה
אני חושב\ת שהייתי רוצה להשתמש במערכת זו לעיתים תכופות.	☐	☐	☐	☐	☐
מצאתי כי המערכת מסובכת ללא סיבה.	☐	☐	☐	☐	☐
חשבתי שהמערכת קלה לשימוש.	☐	☐	☐	☐	☐
אני חושב\ת שאזדקק לתמיכת איש טכני כדי שאוכל להשתמש במערכת זו.	☐	☐	☐	☐	☐
מצאתי כי הפונקציות השונות של המערכת היו מתואמות היטב.	☐	☐	☐	☐	☐
חשבתי כי היה יותר מידי חוסר עקביות במערכת זו.	☐	☐	☐	☐	☐
אני משער\ת כי רוב האנשים יהיו מסוגלים ללמוד להשתמש במערכת זו בקלות.	☐	☐	☐	☐	☐
מצאתי כי המערכת מאוד מסורבלת לשימוש.	☐	☐	☐	☐	☐
חשתי בטחון רב כאשר השתמשתי במערכת.	☐	☐	☐	☐	☐
עלי ללמוד הרבה דברים לפני שאוכל להשתמש במערכת זו.	☐	☐	☐	☐	☐

(*) Second part of the questionnaire text translation to Hebrew from –

ד"ר נעמי שרויאר, פרופסור תמר וייס, ד"ר צבי קופליק - יעילות השימוש של מטפלים במערכת תומכת החלטה בעת התאמת טכנולוגיה מסייע, 2013

5.2. Appendices TUC driving simulator study on comfort-related use cases

5.2.1. Subscales Van der Laan acceptance scale

Acceptance of the Mediator system. Figure 5.1 and Figure 5.2 provide an overview of the participants' acceptance of the Mediator System for the five data collection points, as measured by the Van der Laan Acceptance Scale, for the Usefulness (Figure 5.1) and Satisfying Scale (Figure 5.2) subscales.

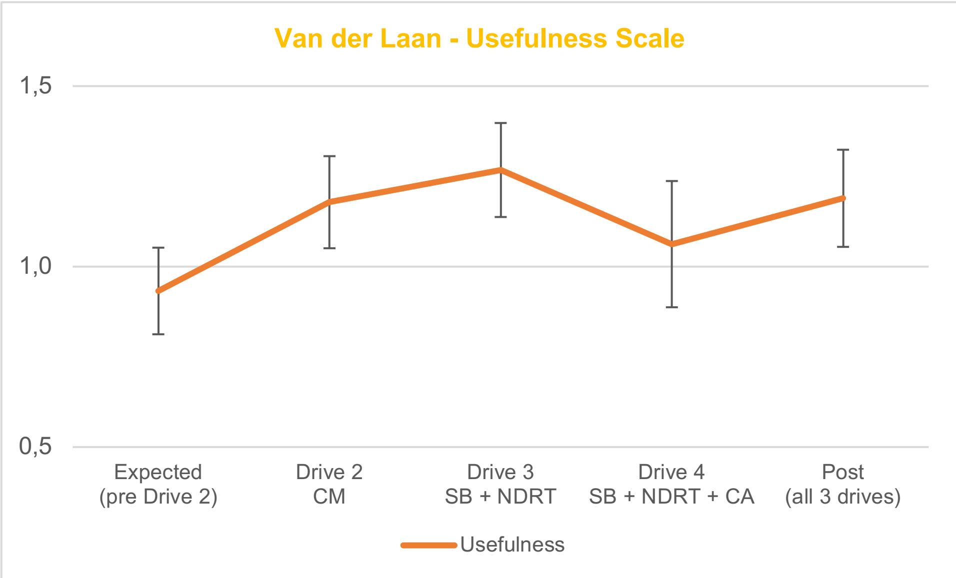


Figure 5.1 Acceptance of the Mediator System (Van der Laan – Usefulness scale) across all five evaluations. Scale range: -2 to 2; Error Bars: 95th-CI.

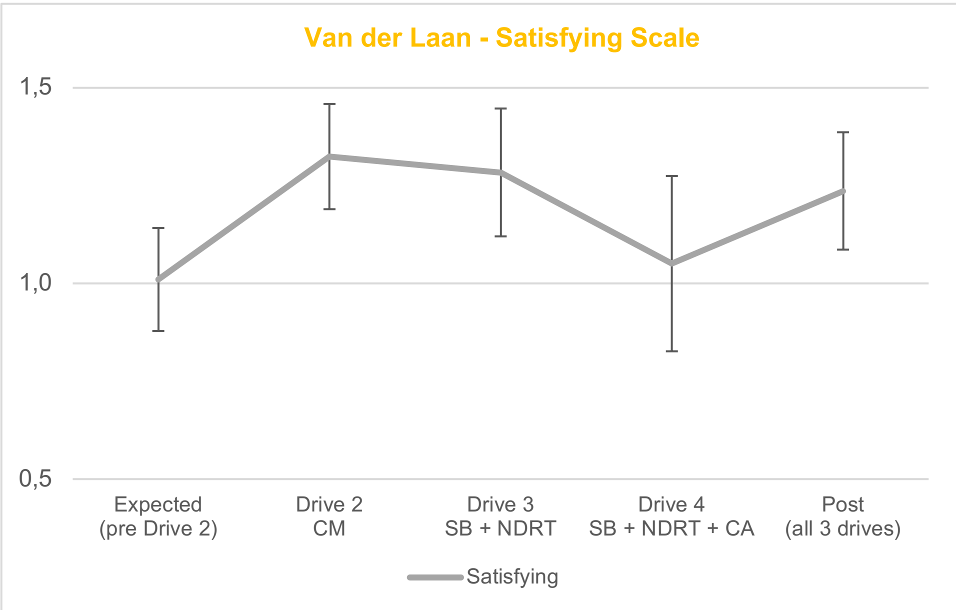


Figure 5.2 Acceptance of the Mediator System (Van der Laan – Satisfying scale) across all five evaluations. Scale range: -2 to 2; Error Bars: 95th-CI.

5.2.2. Additional participant characteristics

Educational degrees. With respect to the highest educational degrees, 32% of the participants had a manual, technical or occupational training, 31% had a university degree, 35% had other degrees (e.g., high school degrees, vocational university degrees, business management specialist etc.) and 2% did not answer the question.

Car availability and usage. 92% of the participants (almost) always had a car at their disposal. With respect to their average car usage, 56% of the participants reported that they used a car (almost) every day, while 28% used a car for three to five days per week. Sixteen percent of the participants used a car less than three days per week.

Mobility behaviour. Regarding the traffic environments (city, rural areas and highways) participants indicated driving most often within the city, followed by rural areas and highways, which they used similarly often.

In addition to the correlations discussed in section 3.4.1.2 (i.e. Table 3.2), the relationship between the participants' car availability as well as car usage with the participant characteristics (i.e., age, gender) and their driving experience were analysed as well (Table 5.1). The participants' age as well as drivers' license possession correlate positively with car availability (medium to large effect size) and car usage (medium effect size). One explanation might be that young drivers often lack the money to immediately buy a car after receiving their drivers' license or have no need to use a car during their studies or education.

Table 5.1 Correlations of driver characteristics, driving experience with car availability and usage.

	Car availability	Car usage
Age	$r = .31$ $p = .008$	$r = .47$ $p < .001$
Gender	$r = -.024$ $p = .842$	$r = .15$ $p = .199$
Annual mileage	$r = .49$ $p < .001$	$r = .26$ $p = .028$
Duration driver's license possession	$r = .5$ $p < .001$	$r = .29$ $p = .013$
Car availability	/	$r = .63$ $p < .001$
Car usage	$r = .63$ $p < .001$	/

5.2.3. Comparison of driver characteristics (age, gender, driving experience) – descriptive results

5.2.3.1. RQ1: Acceptance of and intention to buy / use Mediator

Van der Laan acceptance scale. Table 5.2 provides an overview of the main and interaction effects of the rmANOVA including age, gender, driving experience and ATI scores (grouped) regarding the Van der Laan acceptance scale.

Table 5.2 Van der Laan Acceptance Scale - Main and interaction effects of the rmANOVA including Age, Gender, Driving experience and ATI scores (grouped).

Main and interaction effects	Statistics
Age	$F(3, 27) = .78, p = .515, \eta_p^2 = .08$
Gender	$F(1, 27) = .20, p = .66, \eta_p^2 = .01$
Driving experience (i.e. annual mileage)	$F(3, 27) = 1.24, p = .315, \eta_p^2 = .12$
ATI	$F(1, 27) = 3.89, p = .059, \eta_p^2 = .13$

Main and interaction effects	Statistics
Age * Data collection points	F(5.94, 53.49) = .81, p = .564, η_p^2 = .08
Gender * Data collection points	F(1.98, 53.49) = .84, p = .438, η_p^2 = .03
Driving experience * Data collection points	F(5.94, 53.49) = .94, p = .477, η_p^2 = .09
ATI * Data collection points	F(1.98, 53.49) = .36, p = .697, η_p^2 = .01

L3 Pilot questionnaire (acceptance). For the L3 Pilot questionnaire, the influences of age, gender and driving experience (i.e. annual mileage) were analysed descriptively. In total, the differences were relatively small. Figure 5.3 shows that gender had no noticeable, systematic influence on participants' answers to the acceptance-related items.

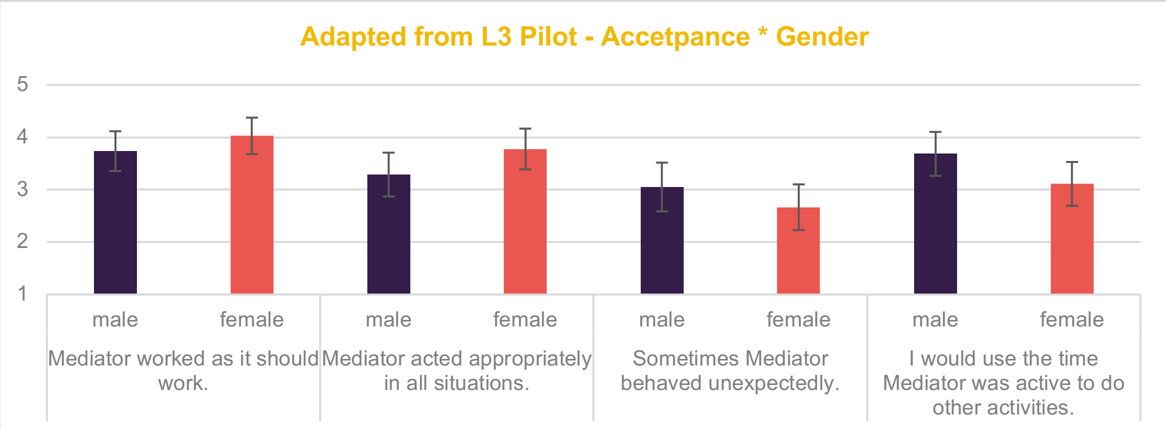


Figure 5.3 Acceptance-related items adapted from L3 Pilot questionnaire in light of participants' gender. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

Comparisons of the four age groups indicated differences that were more pronounced (see Figure 5.4). For instance, the oldest age group (55 years and older) seemed to agree the most with the statements "Mediator worked as it should work" and "Mediator acted appropriately in all situations". In line with this, the older age group also disagreed the most with the statement "Sometimes Mediator behaved unexpectedly". Interestingly, a tendency can be seen that with increasing age, the behaviour of Mediator seemed to be experienced less unexpected and more appropriately. Considering the relationship between participants' age, driving experience (i.e. license possession, annual mileage, car availability, car usage; see sections 3.4.1.2 and 5.2.2), it seems likely that older participants were more experienced with different driving situations and potentially with different ADAS, wherefore they might have been less surprised by the behaviour of Mediator than younger participants might have been.

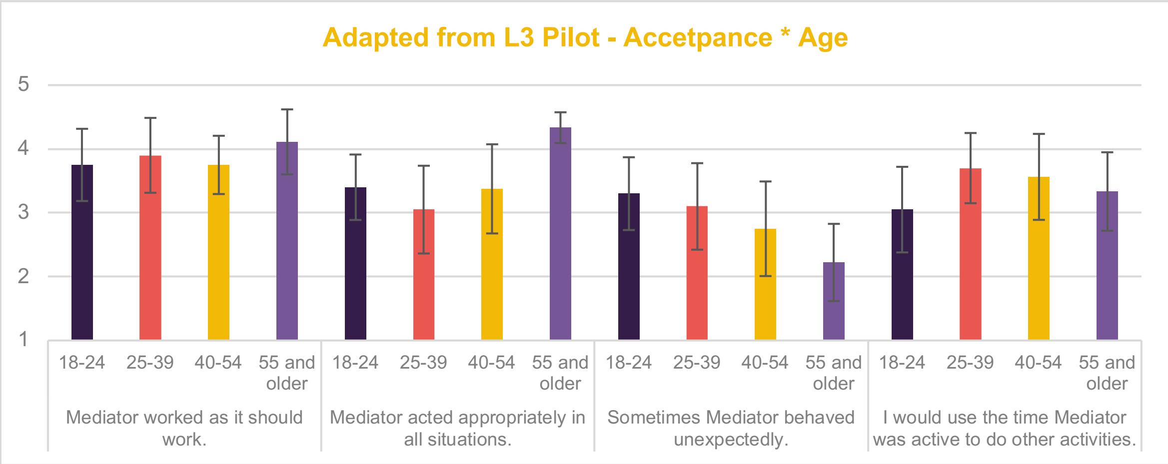


Figure 5.4 Acceptance-related items adapted from L3 Pilot questionnaire in light of participants' age. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

With respect to the participants' driving experience (i.e. annual mileage), there were only small differences (Figure 5.5). For instance, regarding the last statement, "I would use the time Mediator was active to do other activities", the group with very high annual mileage agreed the most with the statement, while the rest gave answers that were more neutral. Hence, those participants who drove the most seemed to be the most interested in using the time to do other activities. This group also included participants, who were driving more due to their occupation (i.e. truck or bus drivers).

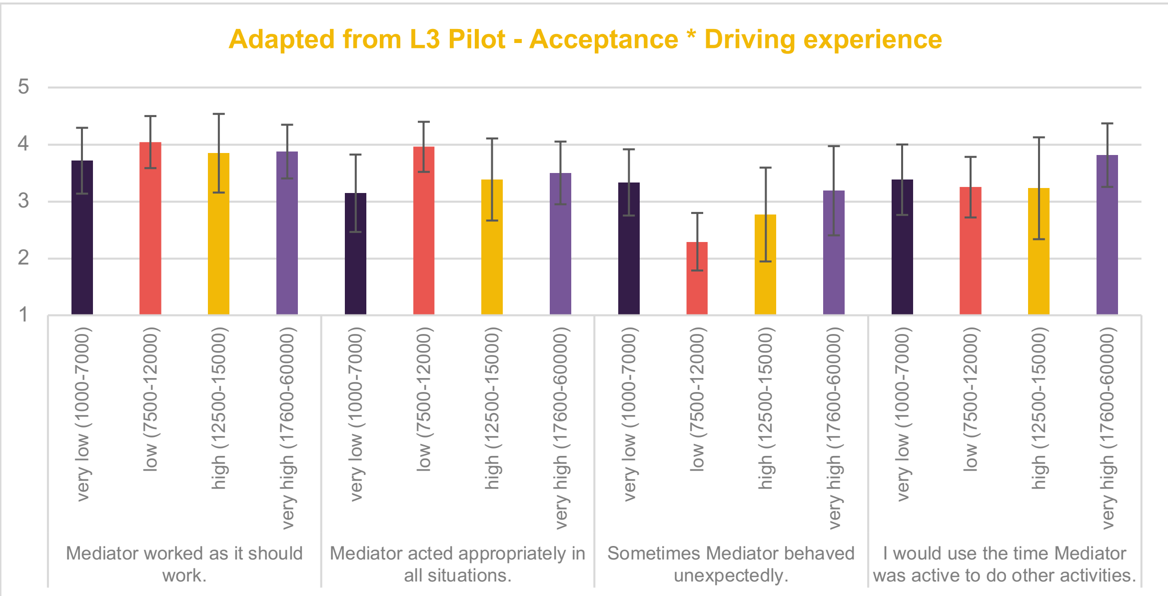


Figure 5.5 Acceptance-related items adapted from L3 Pilot questionnaire in light of participants' driving experience (i.e., annual mileage). Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

L3 Pilot questionnaire (intention to buy / use). The following driver characteristics were analysed descriptively: Age, gender and driving experience. As can be seen in Figure 5.6, Figure 5.7 and Figure 5.8, the influence of these driver characteristics on the different questionnaire items were relatively small. For instance, both female and male participants gave quite similar answers to the eight items as can be seen in Figure 5.6.

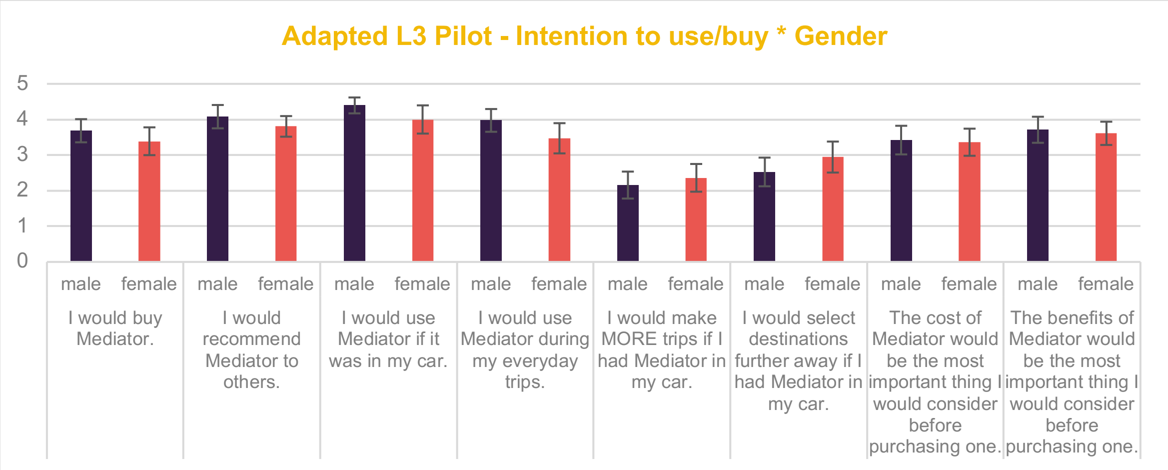


Figure 5.6 Items related to the intention to use/buy adapted from L3 Pilot questionnaire in light of participants' gender. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

Comparing the four age groups, a few, more considerable differences were found (Figure 5.7). Results regarding the item "I would use Mediator during my everyday trips" showed a trend that with increasing age, the agreement with this statement increased as well: The older the participants the more they seemed to want use Mediator during their everyday trips. A similar age related trend can be seen regarding the item "I would make MORE trips if I had Mediator in my car". The younger two age groups disagreed with the statement, while the older two age groups indicated opinions that were more neutral. Hence, the younger participants did not seem to be inclined to make more trips with Mediator.

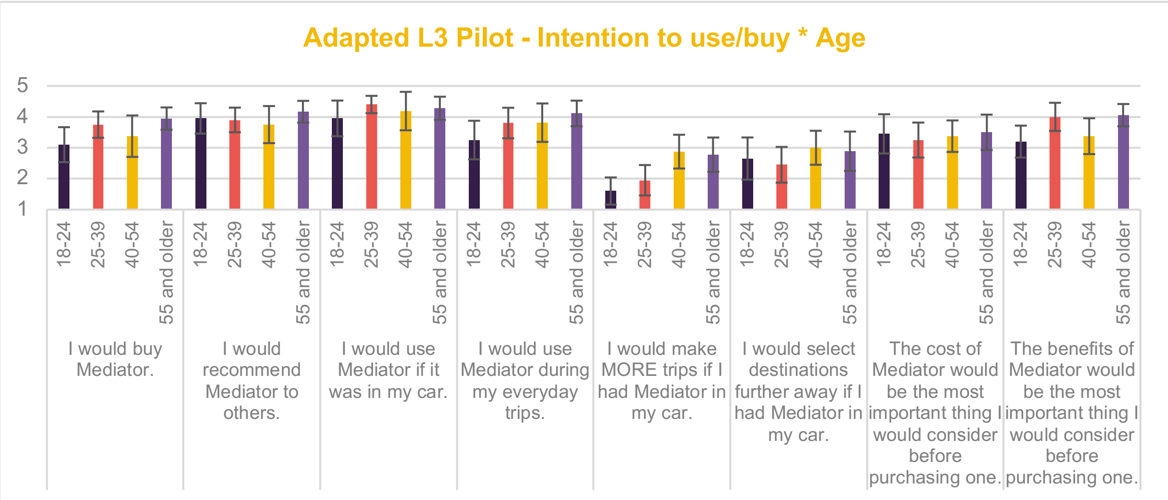


Figure 5.7 Items related to the intention to use/buy adapted from L3 Pilot questionnaire in light of participants' age. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

Regarding the driving experience (i.e. annual mileage), the different groups also showed relatively similar answers (Figure 5.8).

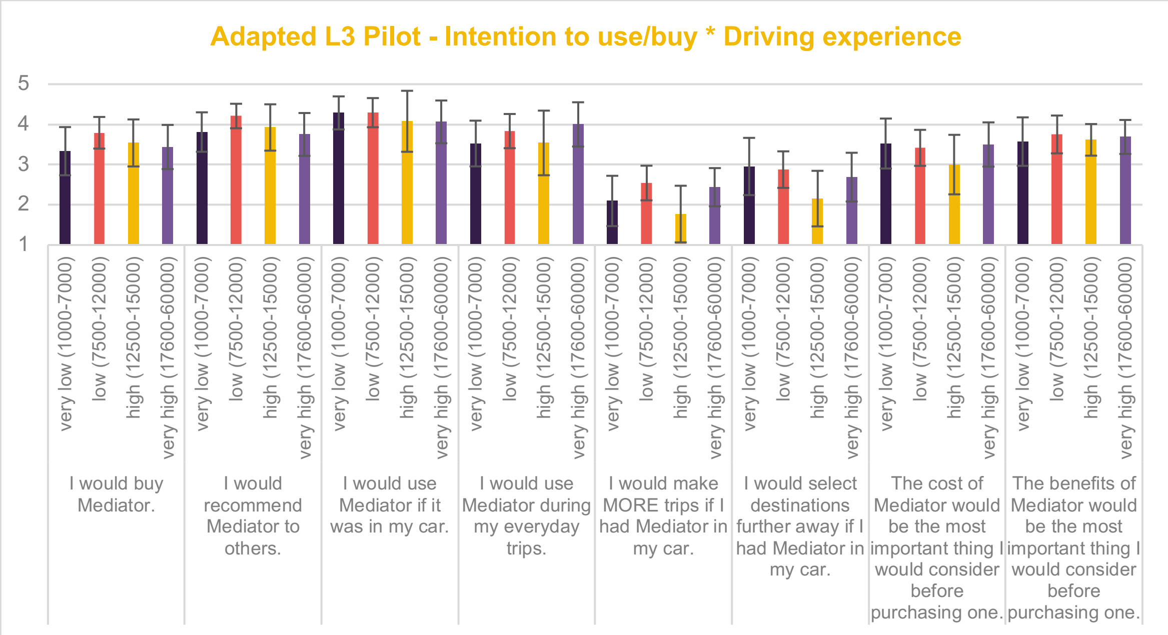


Figure 5.8 Items related to the intention to use/buy adapted from L3 Pilot questionnaire in light of participants' driving experience (i.e., annual mileage). Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

Comparison Mediator vs. no Mediator. The driver characteristics age, gender and driving experience were also analysed descriptively for this questionnaire. Figure 5.9, Figure 5.10 and Figure 5.11 provide the corresponding results. The differences between female and male participants were relatively small (Figure 5.9).

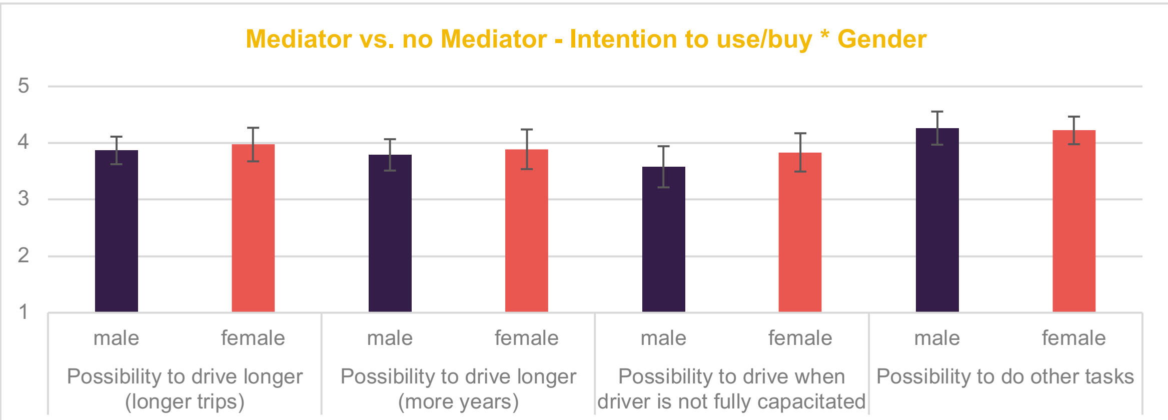


Figure 5.9 Items related to the intention to use/buy from Comparison Mediator vs. no Mediator questionnaire in light of participants' gender. Scale range: 1 – strong decrease; 5 – strong increase; Error Bars: 95th-CI.

The comparison of the four age groups revealed greater differences (Figure 5.10). For instance, with increasing age, a greater increase in the "possibility to drive longer (more years)" was expected. The oldest age group expected the highest increase compared to the youngest age group.

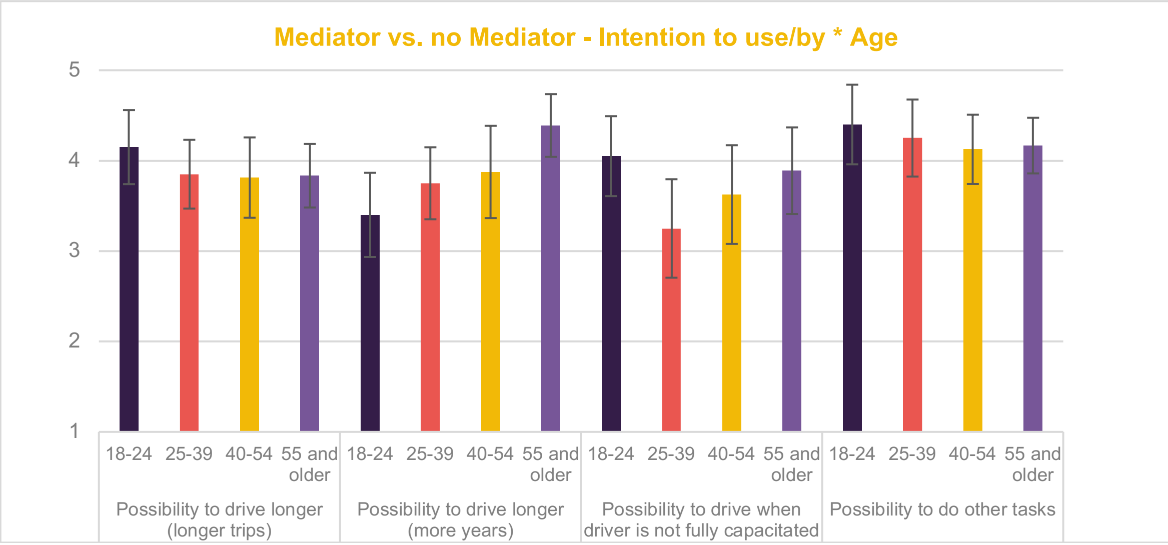


Figure 5.10 Items related to the intention to use/buy from Comparison Mediator vs. no Mediator questionnaire in light of participants' age. Scale range: 1 – strong decrease; 5 – strong increase; Error Bars: 95th-CI.

Lastly, the four driving experience groups were compared as well (Figure 5.11). The four groups had no noticeable, systematic influence on participants' answers.

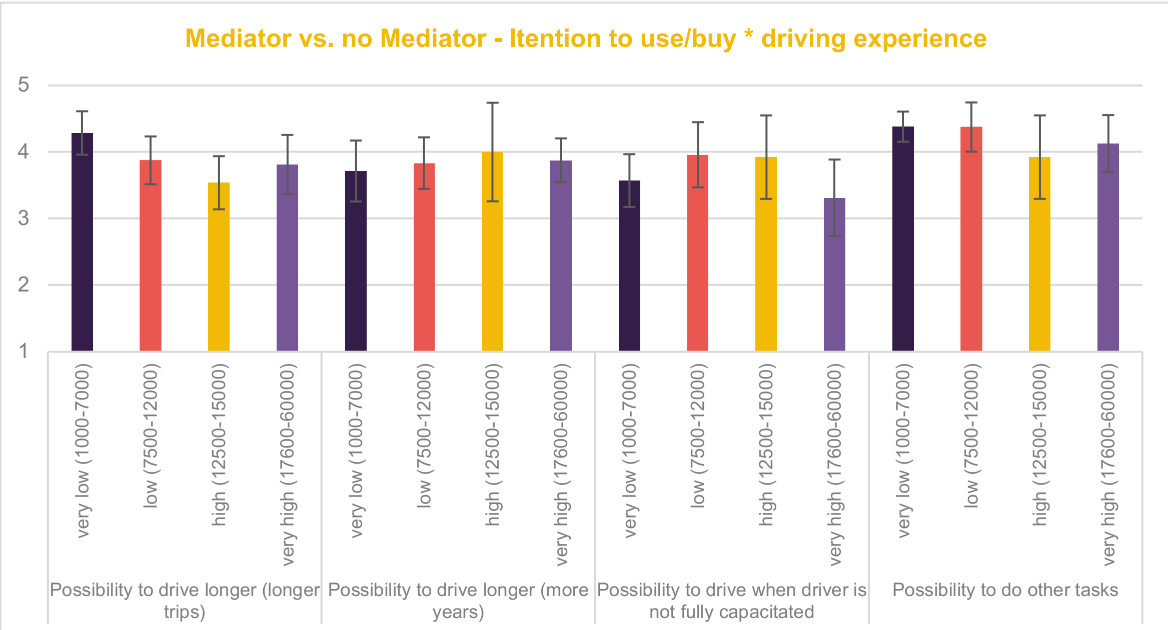


Figure 5.11 Items related to the intention to use/buy from Comparison Mediator vs. no Mediator questionnaire in light of participants' driving experience. Scale range: 1 – strong decrease; 5 – strong increase; Error Bars: 95th-CI.

5.2.3.2. RQ2: Comfort

L3 Pilot questionnaire (comfort). The influences of age, gender and driving experience were analysed descriptively. As can be seen in Figure 5.12, Figure 5.13 and Figure 5.14, neither of the driver characteristics seemed to have a particularly strong influence on the participants' comfort when driving with Mediator. Both, male and female participants gave quite similar answers regarding the five items (Figure 5.12).

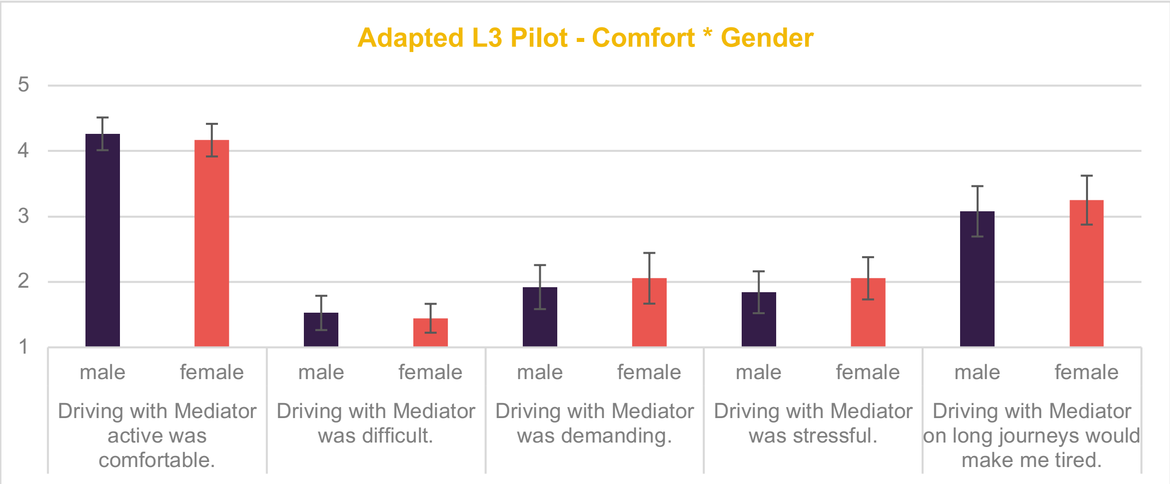


Figure 5.12 Comfort-related items adapted from L3 Pilot questionnaire in light of participants' gender. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

Comparisons between the four age groups (Figure 5.13) and the four driving experience (i.e. annual mileage) groups (Figure 5.14) revealed no noticeable, systematic influences.

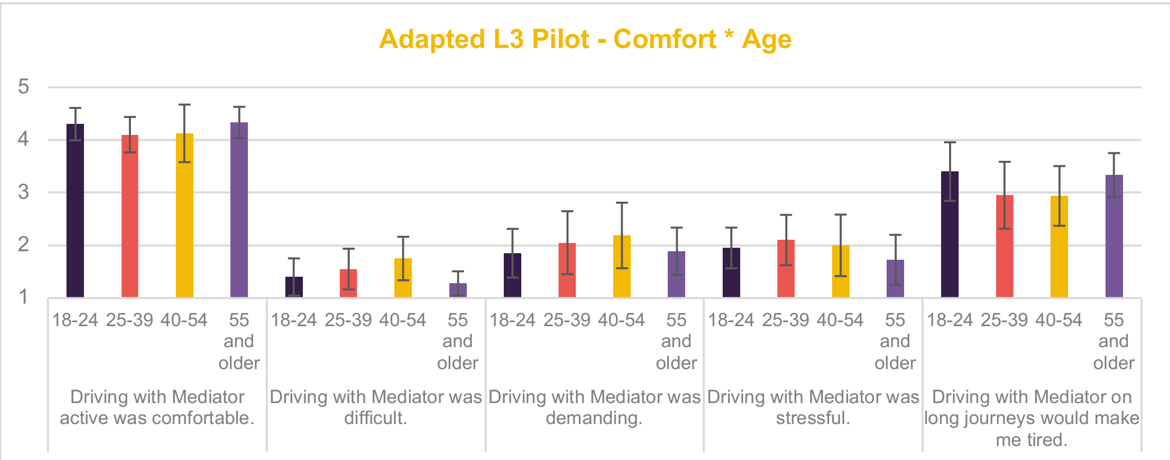


Figure 5.13 Comfort-related items adapted from L3 Pilot questionnaire in light of participants' age. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

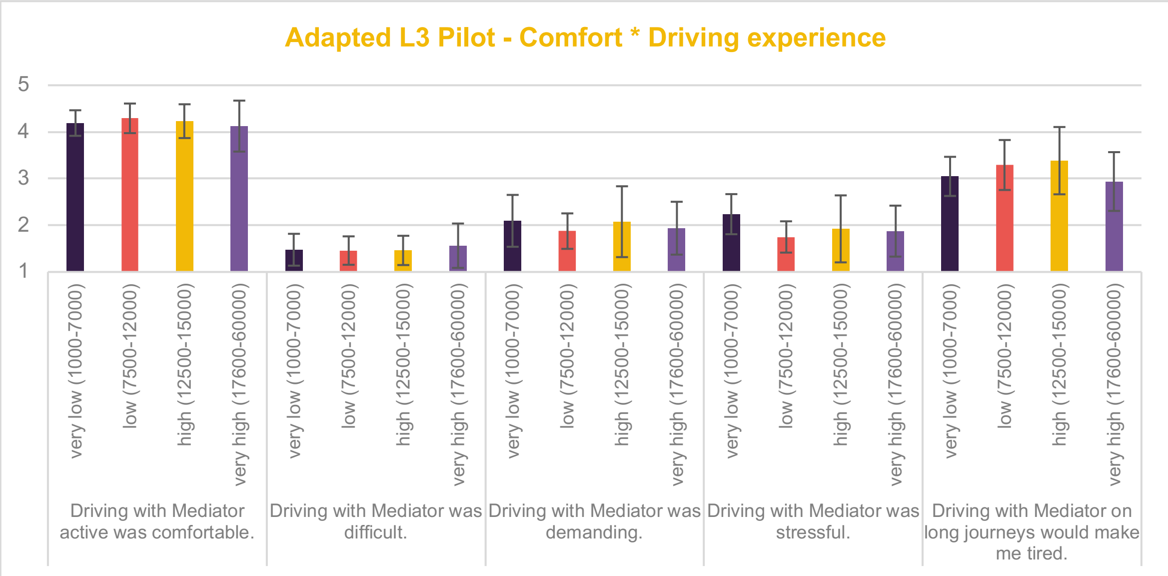


Figure 5.14 Comfort-related items adapted from L3 Pilot questionnaire in light of participants' driving experience (i.e., annual mileage). Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

Comparison Mediator vs. no Mediator. The driver characteristics age, gender and driving experience were also analysed descriptively for this questionnaire. Figure 5.15, Figure 5.16 and Figure 5.17 provide the corresponding results. No systematic influence could be found in terms of participants' gender (Figure 5.15).

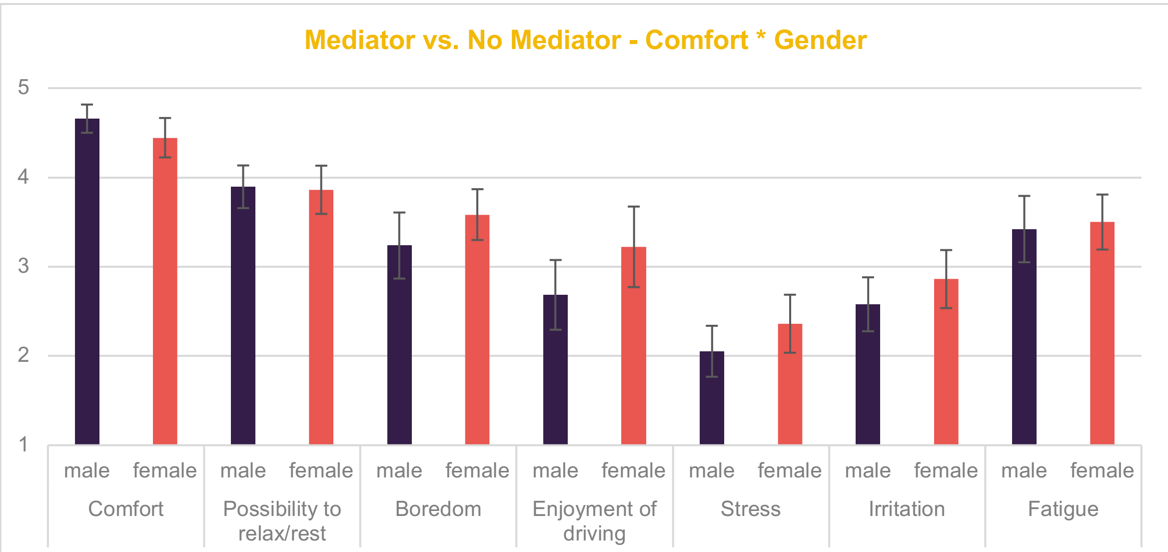


Figure 5.15 Comfort-related items from Comparison Mediator vs. no Mediator questionnaire in light of participants' gender. Scale range: 1 – strong decrease; 5 – strong increase; Error Bars: 95th-CI.

Also, comparisons between the four age groups (Figure 5.16) and the four driving experience (i.e. annual mileage) groups (Figure 5.17) revealed no noticeable, systematic influences.

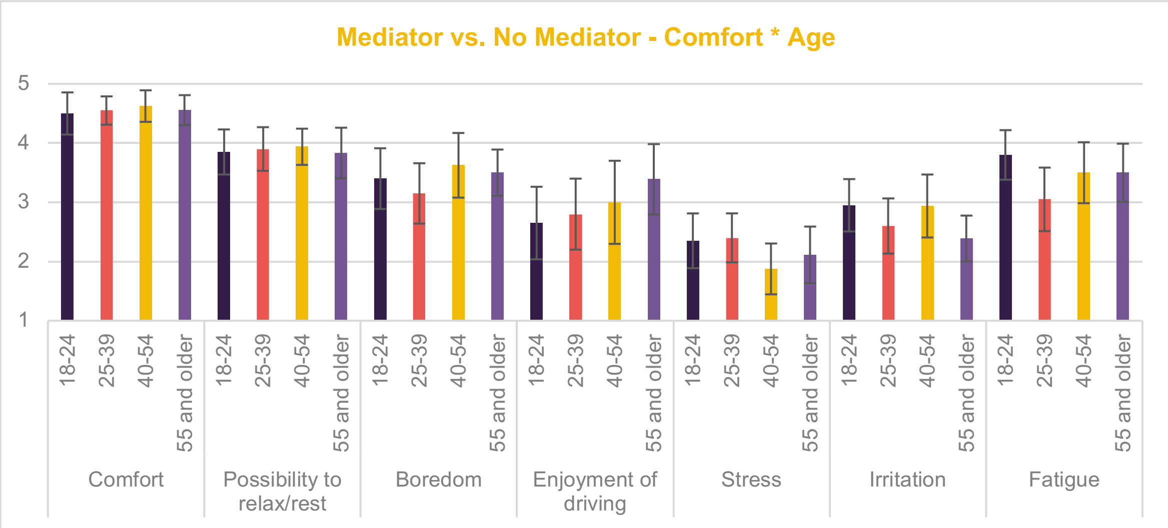


Figure 5.16 Comfort-related items from Comparison Mediator vs. no Mediator questionnaire in light of participants' age. Scale range: 1 – strong decrease; 5 – strong increase; Error Bars: 95th-CI.

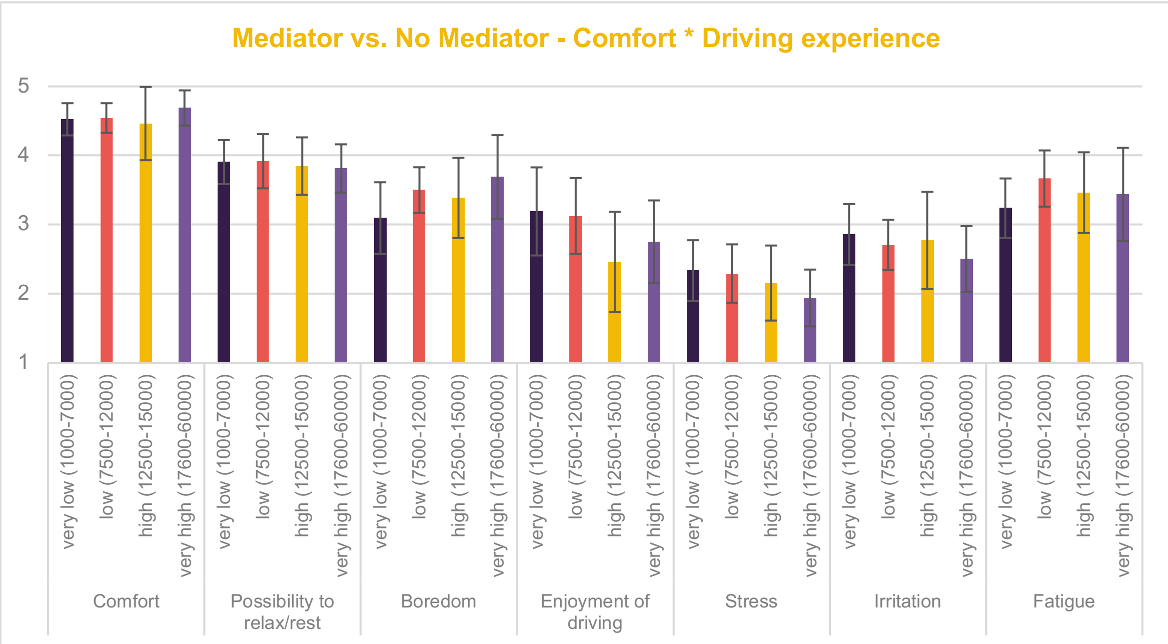


Figure 5.17 Comfort-related items from Comparison Mediator vs. no Mediator questionnaire in light of participants' driving experience. Scale range: 1 – strong decrease; 5 – strong increase; Error Bars: 95th-CI.

5.2.3.3. RQ4: Trust

Trust in automation. Table 5.3 provides an overview of the main and interaction effects of the rmANOVA including age, gender, driving experience and ATI scores (grouped) regarding the trust in automation questionnaire.

Table 5.3 Trust in automation - Main and interaction effects of the rmANOVA including Age, Gender, Driving experience and ATI scores (grouped).

Main and interaction effects	Statistics
Age	$F(3, 27) = .79, p = .511, \eta_p^2 = .08$
Gender	$F(1, 27) = .07, p = .799, \eta_p^2 = .00$
Driving experience (i.e. annual mileage)	$F(3, 27) = .70, p = .559, \eta_p^2 = .07$
ATI	$F(1, 27) = 2.67, p = .114, \eta_p^2 = .09$
Age * Data collection points	$F(7.11, 62.76) = .66, p = .710, \eta_p^2 = .07$
Gender * Data collection points	$F(2.37, 62.76) = .22, p = .841, \eta_p^2 = .01$
Driving experience * Data collection points	$F(7.11, 62.76) = .87, p = .537, \eta_p^2 = .09$
ATI * Data collection points	$F(2.37, 62.76) = 1.25, p = .297, \eta_p^2 = .04$

L3 Pilot questionnaire (trust). For this questionnaire, the influence of age, gender and driving experience were analysed descriptively as well (see Figure 5.18, Figure 5.19, Figure 5.20). Both male and female participants gave quite similar responses regarding the three trust-related items (Figure 5.18).

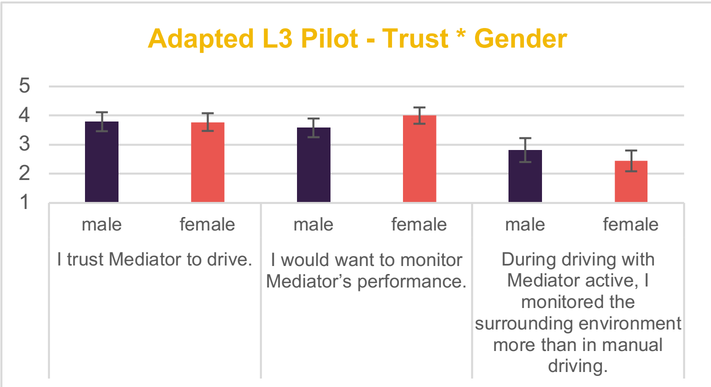


Figure 5.18 Trust-related items adapted from L3 Pilot questionnaire in light of participants’ gender. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

Comparison of the four age groups revealed a few interesting differences (Figure 5.19). Firstly, with increasing age the participants seem to agree more with the statement “I trust Mediator to drive” with the oldest age group (55 years and older) agreeing the most. This might also be related to more driving experience as well as more experience with ADAS systems, as discussed earlier. In line with this, with increasing age the agreement with the statement “I would want to monitor Mediator’s performance” seemed to decrease. Again, the oldest group agreed the least with the statement. Hence, their higher trust might be related to a reduced need to monitor Mediator’s performance, while the other groups, with somewhat lower trust, wanted to monitor it more. Lastly, the oldest age group also agreed the most with the statement “During driving with Mediator active, I monitored the surrounding environment more than in manual driving”. This result is also in line with the previous finding. The oldest age group had enough trust in Mediator to not monitor Mediator’s performance constantly and, hence, had time to concentrate on the surrounding environment.

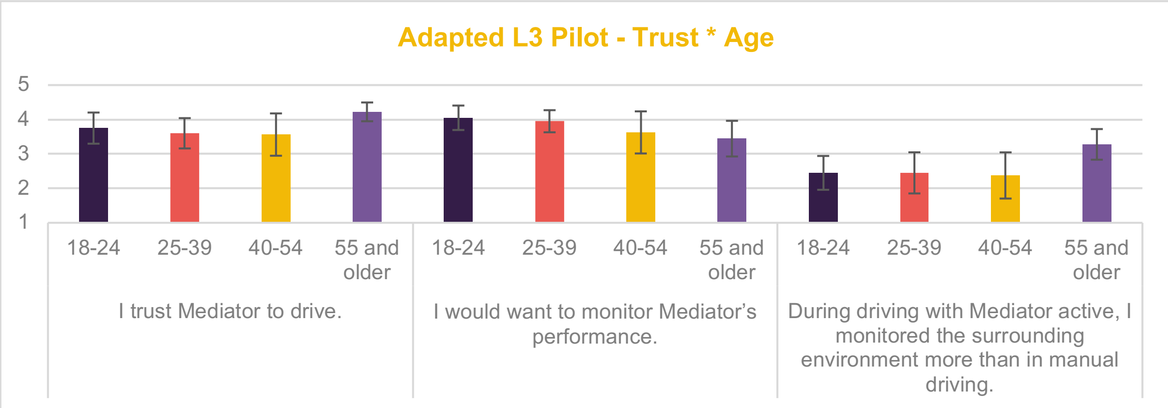


Figure 5.19 Trust-related items adapted from L3 Pilot questionnaire in light of participants' age. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

Lastly, the comparison of the driving experience (i.e. annual mileage) groups revealed no considerable and systematic differences (Figure 5.20).

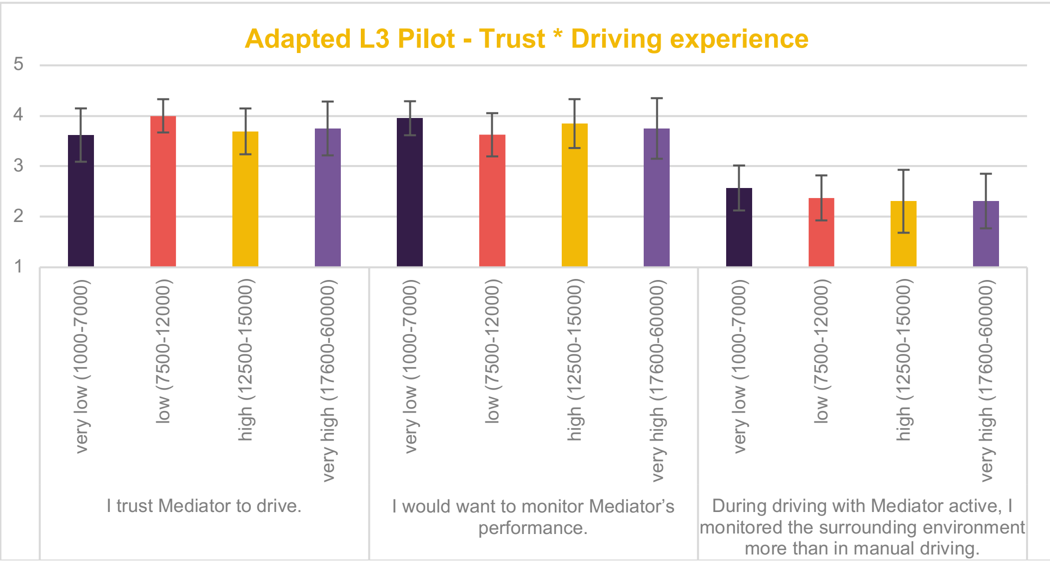


Figure 5.20 Trust-related items adapted from L3 Pilot questionnaire in light of participants' driving experience (i.e., annual mileage). Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

5.2.3.4. RQ5: Perceived safety

Comparison Mediator vs. no Mediator. The differences regarding the participants' age, gender and driving experience were analysed descriptively (Figure 5.21, Figure 5.22, Figure 5.23) as well. Figure 5.21 depicts the differences between male and female participants. As can be seen, female and male participants did not differ considerably from each other (Figure 5.21).

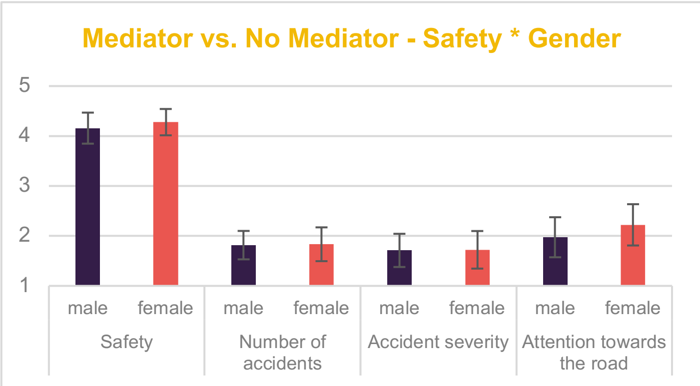


Figure 5.21 Safety-related items from Comparison Mediator vs. no Mediator questionnaire in light of participants' gender. Scale range: 1 – strong decrease; 5 – strong increase; Error Bars: 95th-CI.

The comparison of the four age groups revealed somewhat greater differences (Figure 5.22). Firstly, the oldest age group (55 years and older) tended towards a *great increase* regarding the "safety", while the other three groups only expected a *slight increase.* Moreover, the oldest age group also expected the biggest decrease (i.e. tendency towards *great decrease*) in the "number of accidents" and "accident severity", while most of the other three groups tended more towards a *slight decrease.* Lastly, with increasing age, the participants seemed to expect smaller differences in the "attention towards the road". The youngest age group (18-24 years) expected the greatest decrease in attention towards the road, while the oldest age group tended the most towards no differences.

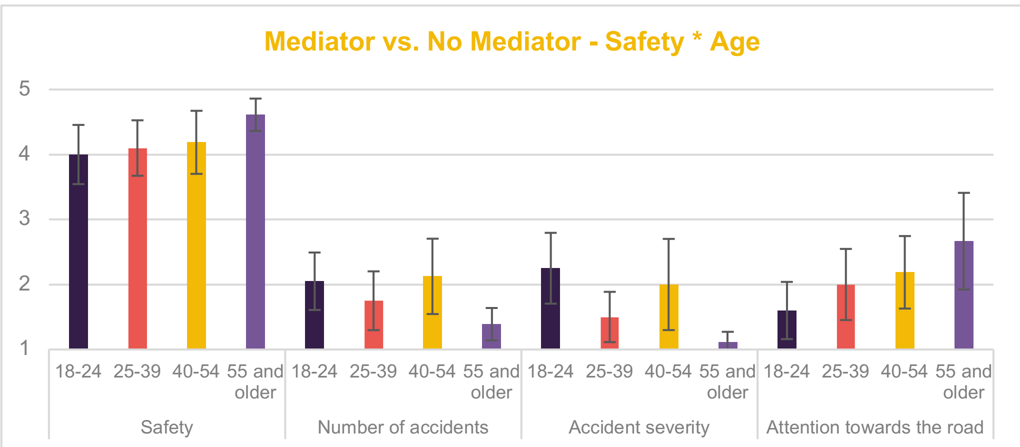


Figure 5.22 Safety-related items from Comparison Mediator vs. no Mediator questionnaire in light of participants' age. Scale range: 1 – strong decrease; 5 – strong increase; Error Bars: 95th-CI.

Regarding the four driving experience (i.e. annual mileage) groups, no substantial, systematic influence on perceived safety could be identified (Figure 5.23).

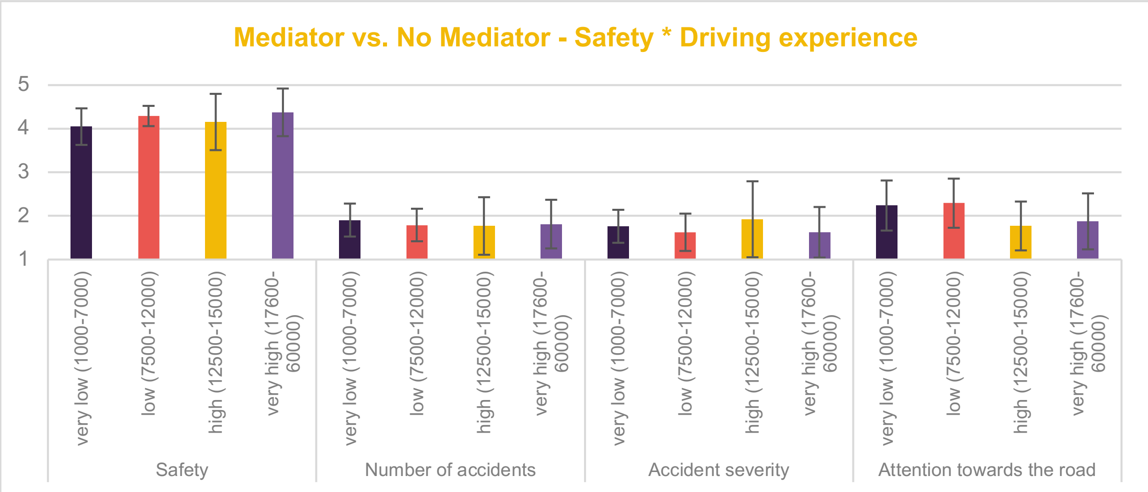


Figure 5.23 Safety-related items from Comparison Mediator vs. no Mediator questionnaire in light of participants' driving experience. Scale range: 1 – strong decrease; 5 – strong increase; Error Bars: 95th-CI.

L3 Pilot questionnaire (safety). For this questionnaire, the influences of age, gender and driving experience were analysed descriptively as well (see Figure 5.24, Figure 5.25, Figure 5.26). Both male and female participants gave very similar responses regarding the three safety-related items (Figure 5.24).

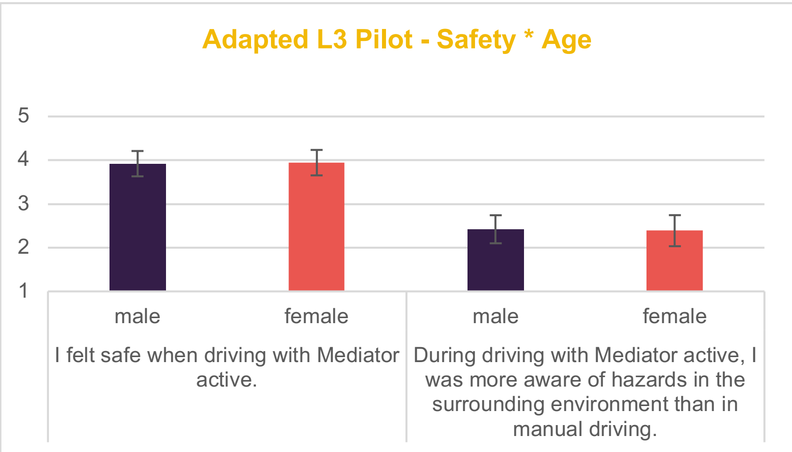


Figure 5.24 Safety-related items adapted from L3 Pilot questionnaire in light of participants' gender. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

Comparing the four age groups revealed differences that were somewhat more considerable (Figure 5.25). Firstly, while all four groups seemed to feel relatively safe, the oldest age group (55 years and older) seemed to feel even a little safer than the other three age groups. Moreover, all four groups tended to *slightly disagree* with the second statement. However, the oldest and the youngest (18-24 years) age group disagreed a little less with the statement than the other two age groups. Hence, while all four groups indicated that they were not more aware of hazards when driving with Mediator active, the youngest and oldest group might have been a little more aware than the other groups.

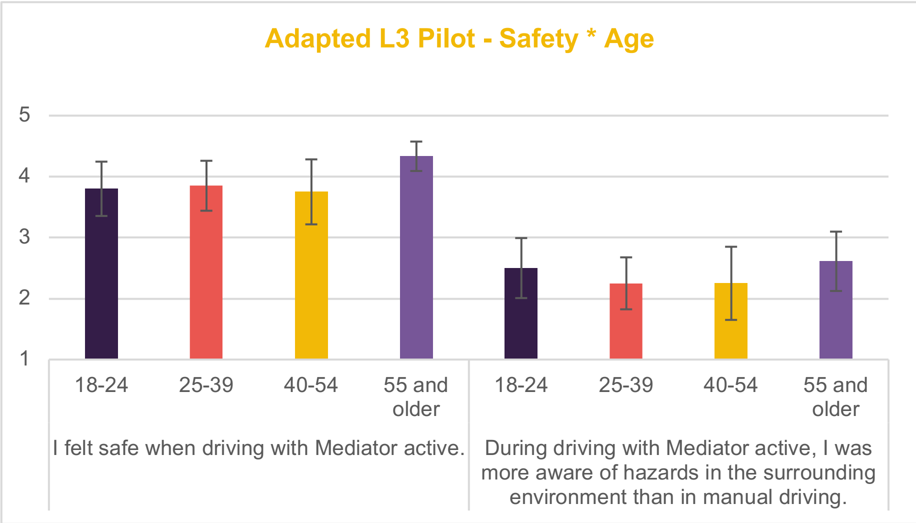


Figure 5.25 Safety-related items adapted from L3 Pilot questionnaire in light of participants' age. Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

Lastly, with respect to the four driving experience (i.e. annual mileage) groups no considerable, systematic influcne could be found (Figure 5.26).

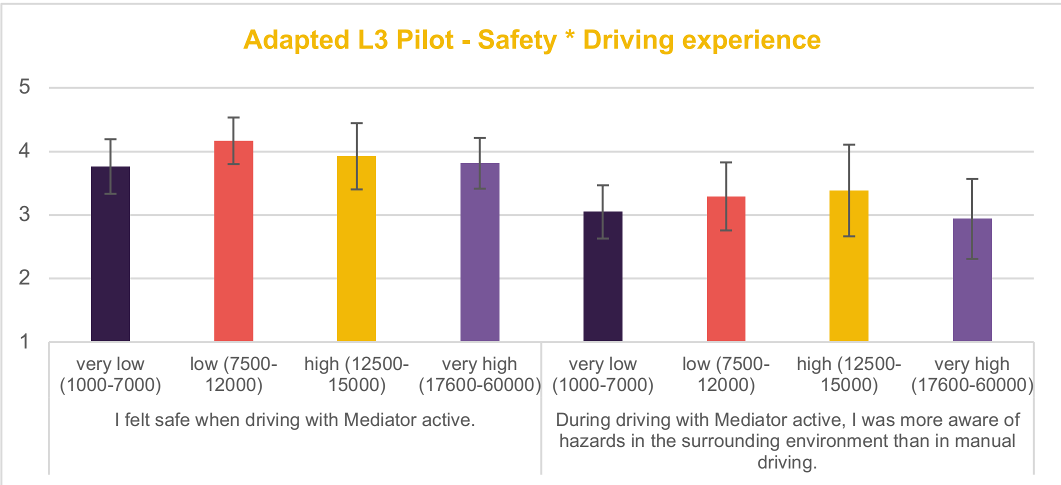


Figure 5.26 Safety-related items adapted from L3 Pilot questionnaire in light of participants' driving experience (i.e., annual mileage). Scale range: 1 – totally disagree, 5 – totally agree; Error Bars: 95th-CI.

5.2.4. Questionnaires

Pre-questionnaire.

Vorbefragung	
VP-Nummer	VP________

Fahrsimulatorstudie:
Beurteilung eines innovativen Anzeige- und Bedienkonzepts für das teil- und hochautomatisierte Fahren

Vielen Dank, dass Sie an der MEDIATOR-Studie teilnehmen. Mit dem folgenden Fragebogen möchten wir wichtige Informationen **über Sie** und Ihre **bisherigen Erfahrungen** sammeln.
Sollten Sie Fragen haben, zögern Sie bitte nicht, uns diese zu stellen.

Wie **alt** sind Sie?	___________ Jahre
Welchem **Geschlecht** fühlen Sie sich zugehörig?	☐ Weiblich ☐ Männlich ☐ Divers
Welches ist der **höchste** Bildungsabschluss, den Sie erworben haben?	☐ Handwerkliche/technische/berufliche Ausbildung ☐ Universitätsabschluss ☐ Sonstiges: ________________________________
Wie ist Ihr **derzeitiger** Beschäftigungsstatus?	☐ Vollzeitbeschäftigt ☐ Teilzeitbeschäftigt ☐ Selbstständig ☐ Hausfrau / Hausmann ☐ Arbeitslos ☐ Im Ruhestand / in Rente ☐ Student ☐ Sonstiges: ________________________________

Nun möchten wir gern wissen, wie sie sich **im Moment** fühlen.

1	2	3	4	5	6	7

Müde Wach

1	2	3	4	5	6	7

Ruhig Gestresst

1	2	3	4	5	6	7

Aufgeregt Gelangweilt

Wie viele Stunden haben Sie **letzte Nacht** schätzungsweise geschlafen?	_____Stunden

In diesem Abschnitt möchten wir mehr über Ihr **Mobilitätsverhalten** und Ihre Erfahrungen mit verschiedenen **Fahrerassistenzsystemen** erfahren.

Haben Sie ein **Auto** zur Verfügung, das Sie benutzen können?	☐ Ja, (fast) immer ☐ Ja, manchmal ☐ Kaum oder gar nicht
Wie oft fahren Sie **im Durchschnitt** mit dem Auto?	☐ (fast) jeden Tag ☐ 3-5 Tage / Woche ☐ 1-2 Tage / Woche ☐ Seltener oder nie
Wie viele Kilometer fahren Sie im Durchschnitt **pro Jahr**?	______________________ km / Jahr
Falls sich Ihr Mobilitätsverhalten in den letzten Jahren stark verändert hat (z.B. durch Homeoffice oder Renteneintritt) geben Sie bitte an, wie viele Kilometer Sie **vorher** durchschnittlich **pro Jahr** gefahren sind.	______________________ km / Jahr
Seit welchem Jahr sind Sie im Besitz eines **Führerscheins der Klasse B**?	______________________
Fahren Sie berufsbedingt (z.B. Fernkraftfahrer:in, Busfahrer:in, Pendler:in) **überdurchschnittlich viel**?	☐ Nein ☐ Ja, weil:___________________________
Bitte bringen Sie die folgenden Verkehrsumfelder in eine **Reihenfolge** hinsichtlich der Häufigkeit, in der sie in den jeweiligen Umgebungen fahren. Vergeben Sie eine **1** für die Umgebung, in der Sie **am häufigsten** fahren und eine **3** für die, in der Sie **am wenigsten** fahren.	Stadt: ______ Ländlicher Raum ______ Autobahn ______

Bitte geben Sie an, ob Ihr **derzeitig am häufigsten genutztes** Fahrzeug mit den folgenden Systemen **ausgestattet** ist:	Ich besitze und nutze das System	Ich besitze das System, aber nutze es nicht	Ich habe das System nicht	Ich weiß nicht, ob ich das System besitze
Einparkhilfe (Ein System, das Ihnen beim Einparken mit Hilfe einer Kamera und/oder akustischen Signaltönen anzeigt, wie nahe Sie sich einem Objekt befinden).	☐	☐	☐	☐
Einparkassistent (Ein System, das das Fahrzeug beim Parallel- oder Rückwärtseinparken steuert. Einige dieser Systeme steuern sowohl die Lenkung als auch das Gaspedal; andere steuern nur die Lenkung und der Fahrer betätigt die Bremse und das Gaspedal).	☐	☐	☐	☐
Tempomat (Ein System, das die Fahrzeuggeschwindigkeit während der Fahrt beibehält).	☐	☐	☐	☐
Adaptive Cruise Control (Ein System, das automatisch verlangsamt oder beschleunigt, um einen sicheren Abstand zu einem vorausfahrenden Fahrzeug zu halten).	☐	☐	☐	☐
Toter-Winkel-Assistent (Ein System, das den linken und rechten toten Winkel des Fahrers auf andere Fahrzeuge überwacht. Oft erhält der Fahrer ein optisches oder akustisches Warnsignal, wenn ein Fahrzeug in der Nähe ist).	☐	☐	☐	☐
Spurverlassenswarnung (Ein System, das beim Halten der Fahrspur hilft, indem es Warnsignale ausgibt, wenn das Fahrzeug die Fahrbahnmarkierungen/-begrenzungen verlässt).	☐	☐	☐	☐
Spurhalteassistent (Ein System, das Autofahrern hilft, das unbeabsichtigte Verlassen der vorgesehenen Fahrspur zu vermeiden).	☐	☐	☐	☐
Vorwärtskollisionswarnung (Ein System, das vor möglichen Kollisionen mit dem vorausfahrenden Fahrzeug warnt).	☐	☐	☐	☐
Müdigkeitswarnung (Ein System, das vor dem Einschlafen warnt).	☐	☐	☐	☐

Nun interessieren wir uns für Ihre **Erfahrungen** mit dem (teil-)automatisierten Fahren.

(Teil-)automatisiertes Fahren kurz erklärt:
Ein (teil-)automatisiertes Fahrzeug ist in der Lage, die Fahraufgabe **in bestimmten Umgebungen** (z.B., Staupilot, „Highway Assist“-Systeme) oder für die **gesamte Fahrt** (z.B., Autopilot) selbst zu übernehmen, d. h., das Fahrzeug übernimmt selbstständig die **Spurführung**, das Einhalten der erlaubten **Geschwindigkeit** und das Einhalten des **Abstandes** zu anderen Fahrzeugen. **Je nach Automatisierungsstufe** müssen Sie die **Fahrumgebung** und das **System** im Auge behalten und jederzeit in der Lage sein, die Fahraufgabe **spontan** wieder selbst zu übernehmen oder Sie können sich mit anderen, **nicht fahrrelevanten Aufgaben** beschäftigen und müssen nur nach einer **Aufforderung** die Fahraufgabe wieder übernehmen. Im Folgenden finden Sie ein paar Beispiele für Formen der Automatisierung, die bereits heute in Fahrzeugen zu finden sind: BMW Driving Assistant Plus, VW Travel Assist, Mercedes Drive Pilot oder Tesla-Autopilot.

Haben Sie schon einmal Formen des automatisierten Fahrens **getestet**? Falls ja, **beschreiben** Sie doch kurz, welche Erfahrungen Sie bereits gesammelt haben **(z.B., mit welchem System, in welchem Kontext, für wie lange)**.

Haben Sie schon einmal an **Studien** zum (teil-)automatisierten Fahren teilgenommen (z.B., Online-Befragungen, Fahrsimulatorstudien, Studien in einem Konzeptfahrzeug)? Falls ja, **beschreiben** Sie doch kurz, den **Inhalt** und die **Art** der Studien.

Der folgende Abschnitt enthält Fragen zu Ihrer **generellen Einstellung** zum Thema Fahrzeugautomatisierung.

Wie ist ihre **generelle Einstellung** zu Funktionen im Fahrzeug, die **Teile oder die gesamte Fahraufgabe** automatisieren können?				
☐ Sehr negativ	☐ Etwas negativ	☐ Neutral	☐ Etwas positiv	☐ Sehr positiv

Wir möchten es gern noch etwas konkreter wissen: Was halten Sie von Funktionen im Fahrzeug, die **Teile oder die gesamte Fahraufgabe** automatisieren können?

Bitte setzen Sie pro Zeile, d.h. zwischen den einzelnen Wortpaaren, ein Kreuz.

Nützlich	☐	☐	☐	☐	☐	Nutzlos
Angenehm	☐	☐	☐	☐	☐	Unangenehm
Schlecht	☐	☐	☐	☐	☐	Gut
Nett	☐	☐	☐	☐	☐	Nervig
Effizient	☐	☐	☐	☐	☐	Unnötig
Ärgerlich	☐	☐	☐	☐	☐	Erfreulich
Hilfreich	☐	☐	☐	☐	☐	Wertlos
Nicht wünschenswert	☐	☐	☐	☐	☐	Wünschenswert
Aktivierend	☐	☐	☐	☐	☐	Einschläfernd

Bitte geben Sie den Grad Ihrer **Zustimmung** zu folgenden Aussagen über **automatisierte Fahrzeuge** an.	Stimmt gar nicht	Stimmt weitgehend nicht	Stimmt eher nicht	Weder noch	Stimmt eher	Stimmt weitgehend	Stimmt völlig
Der **Einsatz** von automatisierten Fahrzeugen ist **akzeptabel**.	☐	☐	☐	☐	☐	☐	☐
Es ist akzeptabel, dass ein **Teil des Verkehrs** aus automatisierten Fahrzeugen bestehen wird.	☐	☐	☐	☐	☐	☐	☐
Es ist **akzeptabel**, dass die Menschen automatisierte Fahrzeuge **nutzen** werden.	☐	☐	☐	☐	☐	☐	☐
Ich würde darauf **vertraue**n, dass sich ein automatisiertes Fahrzeug **wie vorgesehen** verhält.	☐	☐	☐	☐	☐	☐	☐
Ich vertraue darauf, dass automatisierte Fahrzeuge andere Verkehrsteilnehmende **korrekt erkennen** würden.	☐	☐	☐	☐	☐	☐	☐
Ich vertraue darauf, dass die Computersysteme der automatisierten Fahrzeuge **nicht gehackt** werden können.	☐	☐	☐	☐	☐	☐	☐
Automatisierte Fahrzeuge würden meine **Fahrbedürfnisse erfüllen**.	☐	☐	☐	☐	☐	☐	☐
Das Fahren in einem automatisierten Fahrzeug wäre **bequem**, da ich meine Zeit mit anderen Dingen als dem Fahren verbringen könnte.	☐	☐	☐	☐	☐	☐	☐
Ein automatisiertes Fahrzeug zu fahren wäre **praktisch**, da es meine Fahrten effizienter machen würde.	☐	☐	☐	☐	☐	☐	☐
Ein automatisiertes Fahrzeug wäre **sicher**.	☐	☐	☐	☐	☐	☐	☐
Automatisierte Fahrzeuge würden nur ein **minimales Risiko** für **Fahrende und Passagiere** darstellen.	☐	☐	☐	☐	☐	☐	☐

Automatisierte Fahrzeuge würden nur ein **minimales Risiko** für andere **Verkehrsteilnehmende** darstellen.	☐	☐	☐	☐	☐	☐	☐

Die folgende Tabelle enthält Aussagen über Ihre **Interaktion mit technischen Systemen**. Der Begriff "technische Systeme" bezieht sich auf Apps und andere Softwareanwendungen sowie auf ganze digitale Geräte (z. B. Mobiltelefon, Computer, Fernseher, Autonavigation).

Bitte geben Sie den Grad Ihrer **Zustimmung** zu folgenden Aussagen an.	Stimmt gar nicht	Stimmt weitgehend nicht	Stimmt eher nicht	Stimmt eher	Stimmt weitgehend	Stimmt völlig
Ich beschäftige mich gern **genauer** mit technischen Systemen.	☐	☐	☐	☐	☐	☐
Ich **probiere gern** die Funktionen neuer technischer Systeme aus.	☐	☐	☐	☐	☐	☐
In erster Linie beschäftige ich mich mit technischen Systemen, weil ich **muss**.	☐	☐	☐	☐	☐	☐
Wenn ich ein neues technisches System vor mir habe, probiere ich es **intensiv** aus.	☐	☐	☐	☐	☐	☐
Ich verbringe sehr gern Zeit mit dem **Kennenlernen** eines neuen technischen Systems.	☐	☐	☐	☐	☐	☐
Es genügt mir, dass ein technisches System funktioniert, mir ist es **egal, wie oder warum**.	☐	☐	☐	☐	☐	☐
Ich versuche zu **verstehen**, wie ein technisches System genau funktioniert.	☐	☐	☐	☐	☐	☐
Es genügt mir, die **Grundfunktionen** eines technischen Systems zu kennen.	☐	☐	☐	☐	☐	☐
Ich versuche, die Möglichkeiten eines technischen Systems **vollständig** auszunutzen.	☐	☐	☐	☐	☐	☐

Vielen Dank für das Ausfüllen des Fragebogens!
Melden Sie sich bitte jetzt bei der Versuchsleitung.

Post-drive 1 questionnaire.

Nachbefragung Fahrt 1 (manuell)	
VP-Nummer	VP________

Fahrsimulatorstudie:
Beurteilung eines innovativen Anzeige- und Bedienkonzepts für das teil- und hochautomatisierte Fahren

Mit dem folgenden Fragebogen möchten wir herausfinden, wie es Ihnen nach **der ersten Fahrt** im Fahrsimulator geht und wie Ihnen das **manuelle Fahren** gefallen hat.
Sollten Sie Fragen haben, zögern Sie bitte nicht, uns diese zu stellen.

Nun möchten wir gern wissen, wie Sie sich **während der Fahrt** gefühlt haben.

1	2	3	4	5	6	7

Müde — Wach

1	2	3	4	5	6	7

Ruhig — Gestresst

1	2	3	4	5	6	7

Aufgeregt — Gelangweilt

Bitte geben Sie den Grad Ihrer **Zustimmung** zu folgenden Aussagen an.	Stimmt gar nicht	Stimmt weitgehend nicht	Stimmt eher nicht	Stimmt eher	Stimmt weitgehend	Stimmt völlig
Die Fahrt hat mir **großen Spaß** gemacht.	☐	☐	☐	☐	☐	☐
Während des Staus hätte ich mir **gewünscht**, (teil-) automatisierte Fahrfunktionen nutzen zu können.	☐	☐	☐	☐	☐	☐
Es hat mir **nichts ausgemacht**, durch den Stau zu fahren.	☐	☐	☐	☐	☐	☐
Ich hätte die Fahrt **angenehmer** empfunden, wenn es keinen Stau gegeben hätte.	☐	☐	☐	☐	☐	☐
Ich fand es sehr gut, dass mir der Stau **bereits im Vorfeld** im Display angezeigt worden ist.	☐	☐	☐	☐	☐	☐
Die **Stauwarnung** war sehr gut als solche zu erkennen.	☐	☐	☐	☐	☐	☐
Auf die Anzeige der **Entfernung** zum Stau könnte ich sehr gut verzichten.	☐	☐	☐	☐	☐	☐

Haben Sie während der Fahrt **Situationen** erlebt, in denen Sie die Fahraufgabe gern an ein (teil-) automatisiertes System **abgegeben** hätten? Wenn ja, **welche**?

Vielen Dank für das Ausfüllen des Fragebogens!
Melden Sie sich bitte jetzt bei der Versuchsleitung.

Pre-drive 2 questionnaire.

Vorbefragung Fahrt 2 (Mediator)	
VP-Nummer	VP________

Fahrsimulatorstudie:
Beurteilung eines innovativen Anzeige- und Bedienkonzepts für das teil- und hochautomatisierte Fahren

Mit dem folgenden Fragebogen möchten wir herausfinden, welche **Erwartungen** Sie **an Mediator** haben. Sollten Sie Fragen haben, zögern Sie bitte nicht, uns diese zu stellen.

Wenn Sie an **Mediator** denken, was sind Ihre **Erwartungen**? Markieren Sie bitte eine Position zwischen „stimme überhaupt nicht zu“ und „stimme vollkommen zu“.	Stimme überhaupt nicht zu			Weder noch			Stimme vollkommen zu
Ich kann Mediator **vertrauen**.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **irreführend**.	☐	☐	☐	☐	☐	☐	☐
Die Aktionen von Mediator sind **undurchsichtig**.	☐	☐	☐	☐	☐	☐	☐
Ich **misstraue** den Aktionen, Absichten oder Konsequenzen von Mediator.	☐	☐	☐	☐	☐	☐	☐
Ich bin Mediator gegenüber **wachsam**.	☐	☐	☐	☐	☐	☐	☐
Ich bin mit Mediator **vertraut**.	☐	☐	☐	☐	☐	☐	☐
Die Aktionen von Mediator führen zu **nachteiligen oder schädlichen** Konsequenzen.	☐	☐	☐	☐	☐	☐	☐
Ich **traue mir zu**, Mediator zu nutzen.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **glaubwürdig**.	☐	☐	☐	☐	☐	☐	☐
Ich kann mich auf Mediator **verlassen**.	☐	☐	☐	☐	☐	☐	☐
Mediator bietet **Sicherheit**.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **zuverlässig**.	☐	☐	☐	☐	☐	☐	☐

Woran machen Sie fest, ob Sie **erwarten** Mediator **vertrauen** zu können **oder nicht**?

Wie würden Sie Mediator einschätzen, basierend auf den **Informationen**, die Sie gerade erhalten haben? Bitte setzen Sie pro Zeile, d.h. zwischen den einzelnen Wortpaaren, ein Kreuz.						
Nützlich	☐	☐	☐	☐	☐	Nutzlos
Angenehm	☐	☐	☐	☐	☐	Unangenehm
Schlecht	☐	☐	☐	☐	☐	Gut
Nett	☐	☐	☐	☐	☐	Nervig
Effizient	☐	☐	☐	☐	☐	Unnötig
Ärgerlich	☐	☐	☐	☐	☐	Erfreulich
Hilfreich	☐	☐	☐	☐	☐	Wertlos
Nicht wünschenswert	☐	☐	☐	☐	☐	Wünschenswert
Aktivierend	☐	☐	☐	☐	☐	Einschläfernd

Vielen Dank für das Ausfüllen des Fragebogens!
Melden Sie sich bitte jetzt bei der Versuchsleitung.

Post-drive 2 questionnaire.

Nachbefragung Fahrt 2 (Mediator)	
VP-Nummer	VP________

Fahrsimulatorstudie:
Beurteilung eines innovativen Anzeige- und Bedienkonzepts für das teil- und hochautomatisierte Fahren

Mit dem folgenden Fragebogen möchten wir herausfinden, wie Ihnen die Fahrt mit **Mediator** gefallen hat. Sollten Sie Fragen haben, zögern Sie bitte nicht, uns diese zu stellen.

Nun möchten wir gern wissen, wie Sie sich **während der Fahrt** gefühlt haben.

1	2	3	4	5	6	7

Müde — Wach

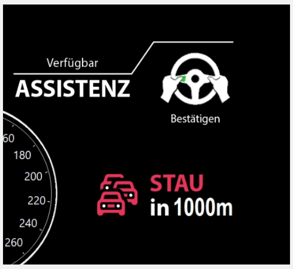
1	2	3	4	5	6	7

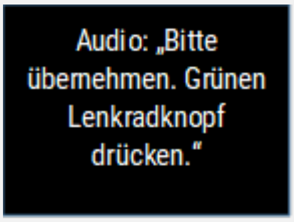
Ruhig — Gestresst

1	2	3	4	5	6	7

Aufgeregt — Gelangweilt

Wie wohl oder unwohl haben Sie sich bei dem **Verhalten des Fahrzeugs** gefühlt? Bitte antworten Sie zu den folgenden Themen:	Sehr unwohl	Etwas unwohl	Neutral	Etwas wohl	Sehr wohl
Eingehaltener **Abstand** zum vorausfahrenden **Fahrzeug**.	☐	☐	☐	☐	☐
Gleichmäßigkeit des Fahrens.	☐	☐	☐	☐	☐
Beschleunigungsverhalten des Fahrzeugs.	☐	☐	☐	☐	☐
Bremsverhalten des Fahrzeugs.	☐	☐	☐	☐	☐
Verhalten in **Kurven**.	☐	☐	☐	☐	☐
Eingehaltener **Abstand** zu **Fahrbahnmarkierungen**.	☐	☐	☐	☐	☐

Nun geht es um die erlebten **Vorschläge**, die Fahraufgabe **durch Mediator übernehmen** zu lassen.  Audio: „Stau voraus und Assistenz verfügbar. Grünen Lenkradknopf drücken, wenn Mediator für Sie fahren soll" Bitte geben Sie an, in wieweit Sie den folgenden Aussagen zustimmen oder diese ablehnen.	Starke Ablehnung	Ablehnung	Neutral	Zustimmung	Starke Zustimmung
Für mich war es **offensichtlich**, **warum** die Übernahme durch Mediator angeboten wurde.	☐	☐	☐	☐	☐
Ich hätte mir **mehr Informationen** darüber gewünscht, **warum** eine Übernahme der Fahraufgabe durch Mediator angeboten wurde.	☐	☐	☐	☐	☐
Während der Übernahme durch Mediator fühlte ich mich immer **sicher**.	☐	☐	☐	☐	☐
Ich habe der Übernahme durch Mediator **nur** zugestimmt, weil ich zu Beginn der Studie so **instruiert** wurde.	☐	☐	☐	☐	☐
Für mich war es sehr einfach, die **Anweisungen** zu **verstehen,** um die Fahraufgabe an Mediator abzugeben.	☐	☐	☐	☐	☐
Der **Zeitpunkt** für das Angebot, die Fahraufgabe durch Mediator übernehmen zu lassen, war **sehr gut** gewählt.	☐	☐	☐	☐	☐

Nun geht es um die erlebten **Aufforderungen**, die Fahraufgabe wieder **selbst zu übernehmen**.   Bitte geben Sie an, in wieweit Sie den folgenden Aussagen zustimmen oder diese ablehnen.	Starke Ablehnung	Ablehnung	Neutral	Zustimmung	Starke Zustimmung
Für mich war es für alle Übernahmeaufforderungen **offensichtlich**, **warum** sie notwendig waren.	☐	☐	☐	☐	☐
Ich hätte mir **mehr Informationen** darüber gewünscht, **warum** eine Übernahme der Fahraufgabe gefordert wurde.	☐	☐	☐	☐	☐
Während der Übernahme fühlte ich mich immer **sicher**.	☐	☐	☐	☐	☐
Als Mediator mich aufforderte, die Kontrolle wieder zu übernehmen, wurde ich **in angemessener Weise gewarnt**.	☐	☐	☐	☐	☐
Als Mediator mich aufforderte, die Kontrolle wieder zu übernehmen, wurde ich gewarnt und hatte **genügend Zeit**, dies sicher zu tun.	☐	☐	☐	☐	☐
Für mich war es sehr einfach, die **Anweisungen** von Mediator zu **verstehen**, um die Fahraufgabe zu übernehmen.	☐	☐	☐	☐	☐

Wie schätzen Sie **Mediator** ein? Markieren Sie bitte eine Position zwischen „stimme überhaupt nicht zu“ und „stimme vollkommen zu“.	Stimme überhaupt nicht zu			Weder noch			Stimme vollkommen zu
Ich kann Mediator **vertrauen**.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **irreführend**.	☐	☐	☐	☐	☐	☐	☐
Die Aktionen von Mediator sind **undurchsichtig**.	☐	☐	☐	☐	☐	☐	☐
Ich **misstraue** den Aktionen, Absichten oder Konsequenzen von Mediator.	☐	☐	☐	☐	☐	☐	☐
Ich bin Mediator gegenüber **wachsam**.	☐	☐	☐	☐	☐	☐	☐
Ich bin mit Mediator **vertraut**.	☐	☐	☐	☐	☐	☐	☐
Die Aktionen von Mediator führen zu **nachteiligen oder schädlichen** Konsequenzen.	☐	☐	☐	☐	☐	☐	☐
Ich **traue mir zu**, Mediator zu nutzen.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **glaubwürdig**.	☐	☐	☐	☐	☐	☐	☐
Ich kann mich auf Mediator **verlassen**.	☐	☐	☐	☐	☐	☐	☐
Mediator bietet **Sicherheit**.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **zuverlässig**.	☐	☐	☐	☐	☐	☐	☐

Woran machen Sie fest, ob Sie Mediator **vertrauen** oder nicht?

Bitte geben Sie den Grad Ihrer **Zustimmung** zu folgenden Aussagen an.	Stimmt gar nicht	Stimmt weitgehend nicht	Stimmt eher nicht	Stimmt eher	Stimmt weitgehend	Stimmt völlig
Die Fahrt hat mir **großen Spaß** gemacht.	☐	☐	☐	☐	☐	☐
Ich fand es **angenehm**, während des Staus Mediator nutzen zu können.	☐	☐	☐	☐	☐	☐
Dank Mediator hat es mir **nichts ausgemacht**, durch den Stau zufahren.	☐	☐	☐	☐	☐	☐
Ich habe mich sehr **sicher** gefühlt, während Mediator für mich **durch den Stau gefahren** ist.	☐	☐	☐	☐	☐	☐
Ich habe mich **während des Auffahrens** auf das Stauende sehr **sicher** gefühlt.	☐	☐	☐	☐	☐	☐
Ich hätte die Fahrt **angenehmer** empfunden, wenn es keinen Stau gegeben hätte.	☐	☐	☐	☐	☐	☐
Die Anzeige des Zeitbudgets war für mich sehr gut **nachvollziehbar**.	☐	☐	☐	☐	☐	☐
Ich war mir manchmal **unsicher**, in welchem Fahrmodus ich mich gerade befinde.	☐	☐	☐	☐	☐	☐
Die unterschiedlichen Farbkonzepte haben mir **geholfen**, den aktuellen Fahrmodus sicher zu identifizieren.	☐	☐	☐	☐	☐	☐
Die Anzeigen im Display waren **übersichtlich** und sehr gut **verständlich**.	☐	☐	☐	☐	☐	☐
Die Informationen im Display sind **ablenkend**.	☐	☐	☐	☐	☐	☐

Wie schätzen Sie Mediator, basierend auf Ihren während der **soeben beendeten** Fahrt gesammelten **Erfahrungen** ein? Bitte setzen Sie pro Zeile, d.h. zwischen den einzelnen Wortpaaren, ein Kreuz.						
Nützlich	☐	☐	☐	☐	☐	Nutzlos
Angenehm	☐	☐	☐	☐	☐	Unangenehm
Schlecht	☐	☐	☐	☐	☐	Gut
Nett	☐	☐	☐	☐	☐	Nervig
Effizient	☐	☐	☐	☐	☐	Unnötig
Ärgerlich	☐	☐	☐	☐	☐	Erfreulich
Hilfreich	☐	☐	☐	☐	☐	Wertlos
Nicht wünschenswert	☐	☐	☐	☐	☐	Wünschenswert
Aktivierend	☐	☐	☐	☐	☐	Einschläfernd

Vielen Dank für das Ausfüllen des Fragebogens!
Melden Sie sich bitte jetzt bei der Versuchsleitung.

Post-drive 3 questionnaire.

Nachbefragung Fahrt 3 (Mediator)	
VP-Nummer	VP________

Fahrsimulatorstudie:
Beurteilung eines innovativen Anzeige- und Bedienkonzepts für das teil- und hochautomatisierte Fahren

Mit dem folgenden Fragebogen möchten wir herausfinden, wie Ihnen die Fahrt mit **Mediator** gefallen hat. Sollten Sie Fragen haben, zögern Sie bitte nicht, uns diese zu stellen.

Nun möchten wir gern wissen, wie Sie sich **während der Fahrt** gefühlt haben.

1	2	3	4	5	6	7

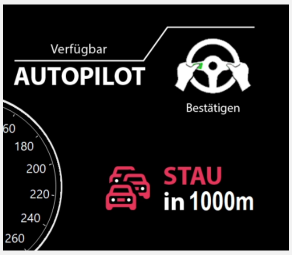
Müde — Wach

1	2	3	4	5	6	7

Ruhig — Gestresst

1	2	3	4	5	6	7

Aufgeregt — Gelangweilt

Nun geht es um die erlebten **Vorschläge**, die Fahraufgabe **durch Mediator übernehmen** zu lassen.  Audio: „Neue Nachricht und Autopilot verfügbar. Grünen Lenkradknopf drücken, wenn Mediator für Sie fahren soll" Bitte geben Sie an, in wieweit Sie den folgenden Aussagen zustimmen oder diese ablehnen.	Starke Ablehnung	Ablehnung	Neutral	Zustimmung	Starke Zustimmung
Für mich war es **offensichtlich**, **warum** die Übernahme durch Mediator angeboten wurde.	☐	☐	☐	☐	☐
Ich hätte mir **mehr Informationen** darüber gewünscht, **warum** eine Übernahme der Fahraufgabe durch Mediator angeboten wurde.	☐	☐	☐	☐	☐
Während der Übernahme durch Mediator fühlte ich mich immer **sicher**.	☐	☐	☐	☐	☐
Ich habe der Übernahme durch Mediator **nur** zugestimmt, weil ich zu Beginn der Studie so **instruiert** wurde.	☐	☐	☐	☐	☐
Für mich war es sehr einfach, die **Anweisungen** zu **verstehen,** um die Fahraufgabe an Mediator abzugeben.	☐	☐	☐	☐	☐
Der **Zeitpunkt** für das Angebot, die Fahraufgabe durch Mediator übernehmen zu lassen, **war sehr** gut gewählt.	☐	☐	☐	☐	☐

Nun geht es um die erlebten **Aufforderungen**, die Fahraufgabe wieder **selbst zu übernehmen**.

Audio: „Bitte übernehmen. Grünen Lenkradknopf drücken."

Bitte geben Sie an, in wieweit Sie den folgenden Aussagen zustimmen oder diese ablehnen.	Starke Ablehnung	Ablehnung	Neutral	Zustimmung	Starke Zustimmung
Für mich war es für alle Übernahmeaufforderungen **offensichtlich**, **warum** sie notwendig waren.	☐	☐	☐	☐	☐
Ich hätte mir **mehr Informationen** darüber gewünscht, **warum** eine Übernahme der Fahraufgabe gefordert wurde.	☐	☐	☐	☐	☐
Während der Übernahme fühlte ich mich immer **sicher**.	☐	☐	☐	☐	☐
Als Mediator mich aufforderte, die Kontrolle wieder zu übernehmen, wurde ich **in angemessener Weise gewarnt**.	☐	☐	☐	☐	☐
Als Mediator mich aufforderte, die Kontrolle wieder zu übernehmen, wurde ich gewarnt und hatte **genügend Zeit**, dies sicher zu tun.	☐	☐	☐	☐	☐
Für mich war es sehr einfach, die **Anweisungen** von Mediator zu **verstehen**, um die Fahraufgabe zu übernehmen.	☐	☐	☐	☐	☐

Wie schätzen Sie **Mediator** ein? Markieren Sie bitte eine Position zwischen „stimme überhaupt nicht zu“ und „stimme vollkommen zu“.	Stimme überhaupt nicht zu			Weder noch			Stimme vollkommen zu
Ich kann Mediator **vertrauen**.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **irreführend**.	☐	☐	☐	☐	☐	☐	☐
Die Aktionen von Mediator sind **undurchsichtig**.	☐	☐	☐	☐	☐	☐	☐
Ich **misstraue** den Aktionen, Absichten oder Konsequenzen von Mediator.	☐	☐	☐	☐	☐	☐	☐
Ich bin Mediator gegenüber **wachsam**.	☐	☐	☐	☐	☐	☐	☐
Ich bin mit Mediator **vertraut**.	☐	☐	☐	☐	☐	☐	☐
Die Aktionen von Mediator führen zu **nachteiligen oder schädlichen** Konsequenzen.	☐	☐	☐	☐	☐	☐	☐
Ich **traue mir zu**, Mediator zu nutzen.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **glaubwürdig**.	☐	☐	☐	☐	☐	☐	☐
Ich kann mich auf Mediator **verlassen**.	☐	☐	☐	☐	☐	☐	☐
Mediator bietet **Sicherheit**.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **zuverlässig**.	☐	☐	☐	☐	☐	☐	☐

Woran machen Sie fest, ob Sie Mediator **vertrauen** oder nicht?

Bitte geben Sie den Grad Ihrer **Zustimmung** zu folgenden Aussagen an.	Stimmt gar nicht	Stimmt weitgehend nicht	Stimmt eher nicht	Stimmt eher	Stimmt weitgehend	Stimmt völlig
Die Fahrt hat mir **großen Spaß** gemacht.	☐	☐	☐	☐	☐	☐
Ich fand es **angenehm**, während des Staus Mediator nutzen zu können.	☐	☐	☐	☐	☐	☐
Dank Mediator hat es mir **nichts ausgemacht**, durch den Stau zufahren.	☐	☐	☐	☐	☐	☐
Ich habe mich sehr **sicher** gefühlt, während Mediator für mich **durch den Stau gefahren** ist.	☐	☐	☐	☐	☐	☐
Ich habe mich **während des Auffahrens** auf das Stauende sehr **sicher** gefühlt.	☐	☐	☐	☐	☐	☐
Ich hätte die Fahrt **angenehmer** empfunden, wenn es keinen Stau gegeben hätte.	☐	☐	☐	☐	☐	☐
Die Anzeige des Zeitbudgets war für mich sehr gut **nachvollziehbar**.	☐	☐	☐	☐	☐	☐
Ich war mir manchmal **unsicher**, in welchem Fahrmodus ich mich gerade befinde.	☐	☐	☐	☐	☐	☐
Die unterschiedlichen Farbkonzepte haben mir **geholfen**, den aktuellen Fahrmodus sicher zu identifizieren.	☐	☐	☐	☐	☐	☐
Die Anzeigen im Display waren **übersichtlich** und sehr gut **verständlich**.	☐	☐	☐	☐	☐	☐
Die Informationen im Display sind **ablenkend**.	☐	☐	☐	☐	☐	☐
Ich fand es **angenehm**, während der Fahrt die Möglichkeit zu haben, eine **Nebenaufgabe** (z.B. Nachrichten zu beantworten) durchführen zu können.	☐	☐	☐	☐	☐	☐
Ich habe mich während der **Bearbeitung der Email** sehr **sicher** gefühlt.	☐	☐	☐	☐	☐	☐

Wie schätzen Sie Mediator, basierend auf Ihren während der **soeben beendeten** Fahrt gesammelten **Erfahrungen** ein? Bitte setzen Sie pro Zeile, d.h. zwischen den einzelnen Wortpaaren, ein Kreuz.						
Nützlich	☐	☐	☐	☐	☐	Nutzlos
Angenehm	☐	☐	☐	☐	☐	Unangenehm
Schlecht	☐	☐	☐	☐	☐	Gut
Nett	☐	☐	☐	☐	☐	Nervig
Effizient	☐	☐	☐	☐	☐	Unnötig
Ärgerlich	☐	☐	☐	☐	☐	Erfreulich
Hilfreich	☐	☐	☐	☐	☐	Wertlos
Nicht wünschenswert	☐	☐	☐	☐	☐	Wünschenswert
Aktivierend	☐	☐	☐	☐	☐	Einschläfernd

Vielen Dank für das Ausfüllen des Fragebogens!
Melden Sie sich bitte jetzt bei der Versuchsleitung.

Post-drive 4 questionnaire.

Nachbefragung Fahrt 4 (Mediator)	
VP-Nummer	VP________

Fahrsimulatorstudie:
Beurteilung eines innovativen Anzeige- und Bedienkonzepts für das teil- und hochautomatisierte Fahren

Mit dem folgenden Fragebogen möchten wir herausfinden, wie Ihnen die Fahrt mit **Mediator** gefallen hat. Sollten Sie Fragen haben, zögern Sie bitte nicht, uns diese zu stellen.

Nun möchten wir gern wissen, wie Sie sich **während der Fahrt** gefühlt haben.

1	2	3	4	5	6	7

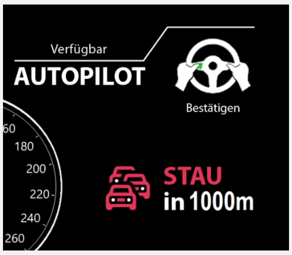
Müde Wach

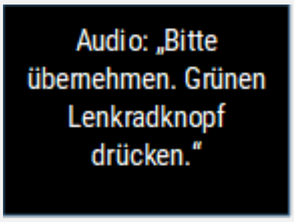
1	2	3	4	5	6	7

Ruhig Gestresst

1	2	3	4	5	6	7

Aufgeregt Gelangweilt

Nun geht es um die erlebten **Vorschläge**, die Fahraufgabe **durch Mediator übernehmen** zu lassen.  Audio: „Stau voraus und Autopilot verfügbar. Grünen Lenkradknopf drücken, wenn Mediator für Sie fahren soll" Bitte geben Sie an, in wieweit Sie den folgenden Aussagen zustimmen oder diese ablehnen.	Starke Ablehnung	Ablehnung	Neutral	Zustimmung	Starke Zustimmung
Für mich war es **offensichtlich**, **warum** die Übernahme durch Mediator angeboten wurde.	☐	☐	☐	☐	☐
Ich hätte mir **mehr Informationen** darüber gewünscht, **warum** eine Übernahme der Fahraufgabe durch Mediator angeboten wurde.	☐	☐	☐	☐	☐
Während der Übernahme durch Mediator fühlte ich mich immer **sicher**.	☐	☐	☐	☐	☐
Ich habe der Übernahme durch Mediator **nur** zugestimmt, weil ich zu Beginn der Studie so **instruiert** wurde.	☐	☐	☐	☐	☐
Für mich war es sehr einfach, die **Anweisungen** zu **verstehen,** um die Fahraufgabe an Mediator abzugeben.	☐	☐	☐	☐	☐
Der **Zeitpunkt** für das Angebot, die Fahraufgabe durch Mediator übernehmen zu lassen, war **sehr gut** gewählt.	☐	☐	☐	☐	☐

Nun geht es um die erlebten **Aufforderungen**, die Fahraufgabe wieder **selbst zu übernehmen**.   Bitte geben Sie an, in wieweit Sie den folgenden Aussagen zustimmen oder diese ablehnen.	Starke Ablehnung	Ablehnung	Neutral	Zustimmung	Starke Zustimmung
Für mich war es für alle Übernahmeaufforderungen **offensichtlich**, **warum** sie notwendig waren.	☐	☐	☐	☐	☐
Ich hätte mir **mehr Informationen** darüber gewünscht, **warum** eine Übernahme der Fahraufgabe gefordert wurde.	☐	☐	☐	☐	☐
Während der Übernahme fühlte ich mich immer **sicher**.	☐	☐	☐	☐	☐
Als Mediator mich aufforderte, die Kontrolle wieder zu übernehmen, wurde ich **in angemessener Weise gewarnt**.	☐	☐	☐	☐	☐
Als Mediator mich aufforderte, die Kontrolle wieder zu übernehmen, wurde ich gewarnt und hatte **genügend Zeit**, dies sicher zu tun.	☐	☐	☐	☐	☐
Für mich war es sehr einfach, die **Anweisungen** von Mediator zu **verstehen**, um die Fahraufgabe zu übernehmen.	☐	☐	☐	☐	☐

Wie schätzen Sie **Mediator** ein? Markieren Sie bitte eine Position zwischen „stimme überhaupt nicht zu“ und „stimme vollkommen zu“.	Stimme überhaupt nicht zu			Weder noch			Stimme vollkommen zu
Ich kann Mediator **vertrauen**.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **irreführend**.	☐	☐	☐	☐	☐	☐	☐
Die Aktionen von Mediator sind **undurchsichtig**.	☐	☐	☐	☐	☐	☐	☐
Ich **misstraue** den Aktionen, Absichten oder Konsequenzen von Mediator.	☐	☐	☐	☐	☐	☐	☐
Ich bin Mediator gegenüber **wachsam**.	☐	☐	☐	☐	☐	☐	☐
Ich bin mit Mediator **vertraut**.	☐	☐	☐	☐	☐	☐	☐
Die Aktionen von Mediator führen zu **nachteiligen oder schädlichen** Konsequenzen.	☐	☐	☐	☐	☐	☐	☐
Ich **traue mir zu**, Mediator zu nutzen.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **glaubwürdig**.	☐	☐	☐	☐	☐	☐	☐
Ich kann mich auf Mediator **verlassen**.	☐	☐	☐	☐	☐	☐	☐
Mediator bietet **Sicherheit**.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **zuverlässig**.	☐	☐	☐	☐	☐	☐	☐

Woran machen Sie fest, ob Sie Mediator **vertrauen** oder nicht?

Bitte geben Sie den Grad Ihrer **Zustimmung** zu folgenden Aussagen an.	Stimmt gar nicht	Stimmt weitgehend nicht	Stimmt eher nicht	Stimmt eher	Stimmt weitgehend	Stimmt völlig
Die Fahrt hat mir **großen Spaß** gemacht.	☐	☐	☐	☐	☐	☐
Ich fand es **angenehm**, während des Staus Mediator nutzen zu können.	☐	☐	☐	☐	☐	☐
Dank Mediator hat es mir **nichts ausgemacht**, durch den Stau zufahren.	☐	☐	☐	☐	☐	☐
Ich habe mich sehr **sicher** gefühlt, während Mediator für mich **durch den Stau gefahren** ist.	☐	☐	☐	☐	☐	☐
Ich habe mich **während des Auffahrens** auf das Stauende sehr **sicher** gefühlt.	☐	☐	☐	☐	☐	☐
Ich hätte die Fahrt **angenehmer** empfunden, wenn es keinen Stau gegeben hätte.	☐	☐	☐	☐	☐	☐
Die Anzeige des Zeitbudgets war für mich sehr gut **nachvollziehbar**.	☐	☐	☐	☐	☐	☐
Ich war mir manchmal **unsicher**, in welchem Fahrmodus ich mich gerade befinde.	☐	☐	☐	☐	☐	☐
Die unterschiedlichen Farbkonzepte haben mir **geholfen**, den aktuellen Fahrmodus sicher zu identifizieren.	☐	☐	☐	☐	☐	☐
Die Anzeigen im Display waren **übersichtlich** und sehr gut **verständlich**.	☐	☐	☐	☐	☐	☐
Die Informationen im Display sind **ablenkend**.	☐	☐	☐	☐	☐	☐
Ich fand es **angenehm**, während der Fahrt die Möglichkeit zu haben, eine **Nebenaufgabe** (z.B. Nachrichten zu beantworten) durchführen zu können.	☐	☐	☐	☐	☐	☐
Ich habe mich während der **Bearbeitung der Email** sehr **sicher** gefühlt.	☐	☐	☐	☐	☐	☐

Wie schätzen Sie Mediator, basierend auf Ihren während der **soeben beendeten** Fahrt gesammelten **Erfahrungen** ein? Bitte setzen Sie pro Zeile, d.h. zwischen den einzelnen Wortpaaren, ein Kreuz.						
Nützlich	☐	☐	☐	☐	☐	Nutzlos
Angenehm	☐	☐	☐	☐	☐	Unangenehm
Schlecht	☐	☐	☐	☐	☐	Gut
Nett	☐	☐	☐	☐	☐	Nervig
Effizient	☐	☐	☐	☐	☐	Unnötig
Ärgerlich	☐	☐	☐	☐	☐	Erfreulich
Hilfreich	☐	☐	☐	☐	☐	Wertlos
Nicht wünschenswert	☐	☐	☐	☐	☐	Wünschenswert
Aktivierend	☐	☐	☐	☐	☐	Einschläfernd

Vielen Dank für das Ausfüllen des Fragebogens!
Melden Sie sich bitte jetzt bei der Versuchsleitung.

Post-questionnaire.

Abschlussbefragung	
VP-Nummer	VP________

Fahrsimulatorstudie:
Beurteilung eines innovativen Anzeige- und Bedienkonzepts für das teil- und hochautomatisierte Fahren

Vielen Dank, dass Sie an der MEDIATOR-Studie teilgenommen haben. Sie haben Mediator **während der letzten Fahrten** in verschiedenen Situationen **kennen gelernt**. Bitte bewerten Sie im folgenden Abschnitt Mediator **im Allgemeinen**.
Sollten Sie Fragen haben, zögern Sie bitte nicht, uns diese zu stellen.

Nun möchten wir gern wissen, wie sie sich **im Moment** fühlen.

1	2	3	4	5	6	7

Müde — Wach

1	2	3	4	5	6	7

Ruhig — Gestresst

1	2	3	4	5	6	7

Aufgeregt — Gelangweilt

Bewerten Sie nun Mediator **über alle Fahrten**. Bitte setzen Sie pro Zeile, d.h. zwischen den einzelnen Wortpaaren, ein Kreuz.						
Nützlich	☐	☐	☐	☐	☐	Nutzlos
Angenehm	☐	☐	☐	☐	☐	Unangenehm
Schlecht	☐	☐	☐	☐	☐	Gut
Nett	☐	☐	☐	☐	☐	Nervig
Effizient	☐	☐	☐	☐	☐	Unnötig
Ärgerlich	☐	☐	☐	☐	☐	Erfreulich
Hilfreich	☐	☐	☐	☐	☐	Wertlos
Nicht wünschenswert	☐	☐	☐	☐	☐	Wünschenswert
Aktivierend	☐	☐	☐	☐	☐	Einschläfernd

Nachfolgend finden Sie eine Liste von Aussagen über **Mediator.** Bitte geben Sie an, in wieweit Sie den folgenden Aussagen zustimmen oder diese ablehnen.	Starke Ablehnung	Ablehnung	Neutral	Zustimmung	Starke Zustimmung
Ich würde Mediator **benutzen**, wenn es in meinem Auto wäre.	☐	☐	☐	☐	☐
Manchmal verhielt sich Mediator **unerwartet**.	☐	☐	☐	☐	☐
Ich habe mich **sicher** gefühlt, als ich mit Mediator gefahren bin.	☐	☐	☐	☐	☐
Ich würde Mediator **kaufen**.	☐	☐	☐	☐	☐
Die **Kosten** von Mediator wären das **Wichtigste**, was ich vor dem **Kauf** eines solchen Systems bedenken würde.	☐	☐	☐	☐	☐
Die **Vorteile** von Mediator wären das **Wichtigste**, was ich vor dem **Kauf** berücksichtigen würde.	☐	☐	☐	☐	☐
Ich würde Mediator **weiter empfehlen**.	☐	☐	☐	☐	☐
Das **Fahren** mit Mediator war **schwierig**.	☐	☐	☐	☐	☐
Das **Fahren** mit Mediator war **fordernd**.	☐	☐	☐	☐	☐
Das **Fahren** mit Mediator war **stressig**.	☐	☐	☐	☐	☐

Nachfolgend finden Sie eine Liste von Aussagen über **Mediator**. Bitte geben Sie an, in wieweit Sie den folgenden Aussagen zustimmen oder diese ablehnen.	Starke Ablehnung	Ablehnung	Neutral	Zustimmung	Starke Zustimmung
Mediator funktionierte so, wie es funktionieren **sollte**.	☐	☐	☐	☐	☐
Ich würde die Leistung von Mediator **überwachen** wollen.	☐	☐	☐	☐	☐
Mediator hat in allen Situationen **angemessen** gehandelt.	☐	☐	☐	☐	☐
Ich würde die Zeit, in der der Autopilot aktiv war, für **andere Aktivitäten** nutzen.	☐	☐	☐	☐	☐
Ich **vertraue** Mediator beim Fahren.	☐	☐	☐	☐	☐
Ich würde Mediator bei meinen **täglichen Fahrten** verwenden.	☐	☐	☐	☐	☐
Das Fahren mit Mediator war **angenehm**.	☐	☐	☐	☐	☐
Während der Fahrt mit Mediator habe ich die Umgebung **mehr beobachtet** als beim manuellen Fahren.	☐	☐	☐	☐	☐
Während der Fahrt mit Mediator war ich mir der **Gefahren** in der Umgebung **stärker bewusst** als beim manuellen Fahren.	☐	☐	☐	☐	☐
Das Fahren mit Mediator auf **langen Strecken** würde mich **müde** machen.	☐	☐	☐	☐	☐
Ich würde **mehr** Fahrten machen, wenn ich Mediator in meinem Auto hätte.	☐	☐	☐	☐	☐
Ich würde **weiter entfernte Ziele** auswählen, wenn ich Mediator in meinem Auto hätte.	☐	☐	☐	☐	☐

Wie schätzen Sie Mediator **ganz allgemein** über **alle 4 Fahrten** hinweg ein? Markieren Sie bitte eine Position zwischen „stimme überhaupt nicht zu“ und „stimme vollkommen zu“.	Stimme überhaupt nicht zu			Weder noch			Stimme vollkommen zu
Ich kann Mediator **vertrauen**.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **irreführend**.	☐	☐	☐	☐	☐	☐	☐
Die Aktionen von Mediator sind **undurchsichtig**.	☐	☐	☐	☐	☐	☐	☐
Ich **misstraue** den Aktionen, Absichten oder Konsequenzen von Mediator.	☐	☐	☐	☐	☐	☐	☐
Ich bin Mediator gegenüber **wachsam**.	☐	☐	☐	☐	☐	☐	☐
Ich bin mit Mediator **vertraut**.	☐	☐	☐	☐	☐	☐	☐
Die Aktionen von Mediator führen zu **nachteiligen oder schädlichen** Konsequenzen.	☐	☐	☐	☐	☐	☐	☐
Ich **traue mir zu**, Mediator zu nutzen.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **glaubwürdig**.	☐	☐	☐	☐	☐	☐	☐
Ich kann mich auf Mediator **verlassen**.	☐	☐	☐	☐	☐	☐	☐
Mediator bietet **Sicherheit**.	☐	☐	☐	☐	☐	☐	☐
Mediator ist **zuverlässig**.	☐	☐	☐	☐	☐	☐	☐

Nachfolgend finden Sie eine Liste von Aussagen zur Bewertung Ihres **allgemeinen Eindrucks** von Mediator. Bitte kreuzen Sie in jeder Zeile den Punkt an, der Ihr Gefühl am besten beschreibt.	Stimme überhaupt nicht zu				Stimme vollkommen zu
Ich denke, dass ich Mediator gerne **häufig benutzen** würde.	☐	☐	☐	☐	☐
Ich fand Mediator **unnötig komplex**.	☐	☐	☐	☐	☐
Ich fand Mediator **einfach** zu benutzen.	☐	☐	☐	☐	☐
Ich glaube, ich würde die **Hilfe** einer technisch versierten Person **benötigen**, um Mediator benutzen zu können.	☐	☐	☐	☐	☐
Ich fand, die verschiedenen Funktionen in Mediator waren **gut integriert**.	☐	☐	☐	☐	☐
Ich denke, Mediator enthielt zu **viele Inkonsistenzen**.	☐	☐	☐	☐	☐
Ich kann mir vorstellen, dass die meisten Menschen den Umgang mit Mediator **sehr schnell lernen**.	☐	☐	☐	☐	☐
Ich fand Mediator **sehr umständlich** zu nutzen.	☐	☐	☐	☐	☐
Ich fühlte mich bei der Benutzung von Mediator **sehr sicher**.	☐	☐	☐	☐	☐
Ich musste eine **Menge lernen**, bevor ich anfangen konnte Mediator zu verwenden.	☐	☐	☐	☐	☐

Welche **Unterschiede** erwarten Sie bei der **Nutzung** von Mediator im Vergleich zum Fahren ohne Mediator? Bitte kreuzen Sie Ihre Einschätzung auf der Skala an.	Starke Verringerung	Leichte Verringerung	Kein Unterschied	Leichte Erhöhung	Starke Erhöhung
Sicherheit	☐	☐	☐	☐	☐
Anzahl an Unfällen	☐	☐	☐	☐	☐
Schwere der Unfälle	☐	☐	☐	☐	☐
Komfort	☐	☐	☐	☐	☐
Aufmerksamkeit für die Straße	☐	☐	☐	☐	☐
Langeweile	☐	☐	☐	☐	☐
Freude am Fahren	☐	☐	☐	☐	☐
Stress	☐	☐	☐	☐	☐
Irritation	☐	☐	☐	☐	☐
Ermüdung	☐	☐	☐	☐	☐
Möglichkeit, länger zu fahren (längere Fahrten)	☐	☐	☐	☐	☐
Möglichkeit zur Entspannung/Ruhe	☐	☐	☐	☐	☐
Möglichkeit, andere Aufgaben zu übernehmen	☐	☐	☐	☐	☐
Möglichkeit zu fahren, wenn der Fahrer nicht voll einsatzfähig ist	☐	☐	☐	☐	☐
Möglichkeit, länger zu fahren (mehrere Jahre)	☐	☐	☐	☐	☐

Vielen Dank für das Ausfüllen des Fragebogens!
Melden Sie sich bitte jetzt bei der Versuchsleitung.

5.2.5. Interview guidelines

Short interview (after each drive).

1. **Wie haben Sie die Fahrt empfunden?**

   Evtl. nachfragen, falls die Probanden nicht von selbst etwas dazu gesagt haben:

   i. *Haben Sie während der Fahrt unangenehme oder kritische Situationen erlebt?*

   Falls ja – Frage je nach Fahrt stellen

   - Fahrt 1: *Wann und warum sind diese Situationen aufgetreten und wie haben Sie diese gelöst?*
   - Fahrt 2-4: *Wann und warum sind diese Situationen aufgetreten und wie konnte die Situation gelöst werden?"*

Slightly longer, more in-depth interview (after one of the automated drives 2, 3 or 4).

1. **Mediator hat Ihnen während der Fahrt angeboten, die Fahraufgabe für Sie zu übernehmen. Können Sie kurz erläutern, wann das war und was Sie glauben, warum der Vorschlag von Mediator gesendet wurde?**

2. **Während der Fahrt hat Mediator Sie aufgefordert, die Fahraufgabe wieder zu übernehmen. Können Sie kurz erläutern, wann das war und was Sie glauben, warum die Aufforderung von Mediator gesendet wurde?**

3. **Denken Sie bitte zurück an die erste Fahrt, als Sie die komplette Strecke selbst gefahren sind. Würden Sie in einer vergleichbaren Alltagssituation das Fahren mit oder ohne Mediator bevorzugen? Warum?**

Final, highly-detailed interview.

Hauptfragen, die **unbedingt** gestellt werden müssen	**Aspekte**, die **nachgefragt** werden sollen, falls die Teilnehmenden nicht von selbst etwas dazu gesagt haben	**Check**
	<<Mediator allgemein>>	
Welche Erfahrungen haben Sie mit Mediator gemacht?	Welche **Vorteile** haben Sie im Vergleich zum Fahren ohne Mediator festgestellt?	☐
	Welche **Nachteile**/**Herausforderungen** haben Sie im Vergleich zum Fahren ohne Mediator festgestellt?	☐
	War Ihnen immer **klar, warum** Sie die verschiedenen Mitteilungen von Mediator erhalten haben? (z.B. Übernahmeaufforderung; Angebot, dass Mediator für Sie fährt)	☐
	Wie fanden Sie es, dass das System Sie **durch den Stau** gefahren hat?	☐
	Wie fanden Sie es, dass Sie während des Fahrens eine **Email lesen und schreiben** konnten?	☐
Welche anderen Situationen fallen Ihnen ein, wo Sie gern eine solche Unterstützung hätten?	Stellen Sie sich vor, Mediator würde erkennen, dass Sie **müde** werden und Ihnen anbieten, eine Weile für Sie zu fahren, damit Sie sich **ausruhen** können. Hätten Sie gern eine solche Funktion?	☐
	Stellen Sie sich vor, Mediator würde erkennen, wenn Sie z.B. durch Gespräche mit Mitfahrenden sehr **abgelenkt** sind und Ihnen anbieten, eine Weile für Sie zu fahren, damit Sie die Gespräche in Ruhe **führen** können. Hätten Sie gern eine solche Funktion?	☐
	<<Vertrauen>>	
Haben Sie Mediator vertraut?	In **was** oder **worauf** genau vertrauen Sie?	☐
	Warum vertrauen Sie? **Wodurch entsteht** ihr Vertrauen in Mediator?	☐
	Was würde Ihnen helfen, um **mehr** Vertrauen in Mediator zu haben?	☐
	Können Sie sich Situationen vorstellen, die das Vertrauen in Mediator **schwächen** würden?	☐

<<Vergleich der 4 Fahrten>>		
Sie haben 4 Fahrten erlebt. Eine manuelle, eine mit Assistenz und 2 mit Autopilot. Bitte sortieren Sie diese von der angenehmsten zur unangenehmsten Fahrt und begründen Sie Ihre Entscheidung.	Warum ist diese Fahrt für Sie die **angenehmste**?	☐
	Wodurch unterscheiden sich die **beiden mittleren Fahrten** für Sie? Welche war angenehmer und wieso?	☐
	Warum ist diese Fahr für Sie die **unangenehmste**?	☐

<<Anzeige- und Bedienkonzept>>		
Wie würden Sie die verschiedenen Elemente des Anzeige- und Bedienkonzeptes bewerten?	Welche Meinung haben Sie zu dem **Lichtkonzept**? • LED-Band Lenkrad & Armaturenbrett • Indirekte Beleuchtung	☐
	Wie klar war es für Sie, **wer für das Fahren zuständig** ist und **was Sie tun dürfen**?	☐
	Waren die vom System gesendeten **Meldungen verständlich**?	☐
	Wie fanden Sie die **Menge an angezeigten Informationen**? Haben wichtige Informationen **gefehlt**?	☐
Schauen Sie sich diese Abbildungen an. Können Sie kurz erläutern, wann Sie diese während der Fahrt gesehen haben (in welchem Fahrmodus waren Sie unterwegs, was war gerade passiert, was wird demnächst passieren)?	Wie haben Sie die **Farbgebung der abgebildeten Straße** verstanden? (für Bild 1 und 3 erklären lassen) Was glauben Sie, was der **graue Bereich** bedeutet? Was glauben Sie, was der **violette Bereich** bedeutet?	☐
	Bild 1: Ist Ihnen aufgefallen, dass der violette Bereich während des Fahrens **größer wurde**? Was glauben Sie, was das bedeutet?	☐
	Bild 3: Ist Ihnen aufgefallen, dass der violette Bereich während des Fahrens **kleiner wurde**?	☐

	Wie haben Sie das **Zeitbudget** verstanden? Was zeigt es Ihnen an? (für Bild 1 und 3 erklären lassen)	☐
	Bild 1: Können Sie nachvollziehen, warum die angegebene Zeit zu diesem Zeitpunkt unveränderlich bleibt, es also **keinen Countdown** gibt?	☐
	Bild 2: Können Sie erklären, **wann und warum der Countdown eingesetzt** hat?	☐
	Wie würden Sie dieses **Lenkradsymbol** hier unten interpretieren? (für Bild 1 und 3 erklären lassen)	☐
<<Verbesserungspotenzial>>		
Würden Sie Mediator in Ihrem zukünftigen Auto verwenden? Bitte erläutern Sie, warum ja oder nein oder unter welchen Bedingungen?	Was würden Sie **ändern**, um Mediator zu verbessern?	☐

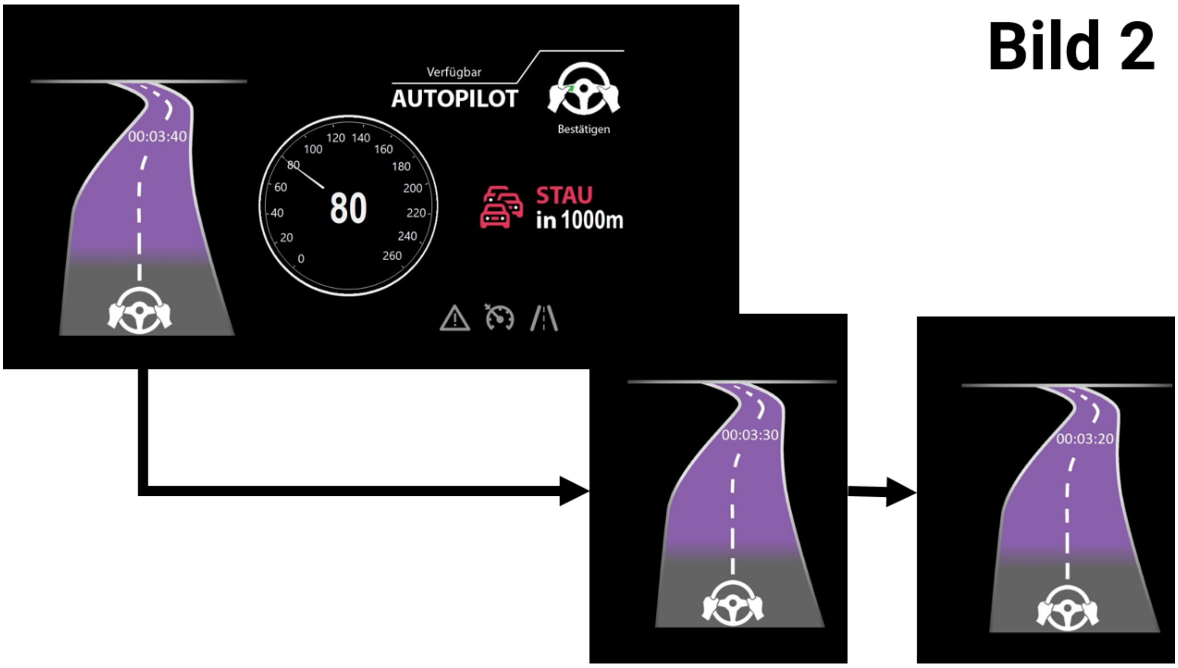

Verfügbar
AUTOPILOT
Bestätigen
00:03:40
80
STAU
in 1000m
00:03:30
00:03:20

Bild 2


00:02:50
5

Bild 3