Integration of Evaluation Results from the MEDIATOR Studies

Deliverable D3.5 – WP3 – Public

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

Integration of Evaluation Results from the MEDIATOR Studies

Work package 3, Deliverable D3.5

Please refer to this report as follows:
Rauh, N., Springer-Teumer, S., Spaan, M., Borowsky, A., Fiorentino, A., Ahlström, C., Thalya, P. (2023). Integration of evaluation results from the MEDIATOR studies, Deliverable D3.5 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:	30/04/2023
Submission date:	30/04/2023

Lead contractor for this deliverable:
Nadine Rauh – Chemnitz University of Technology

Report Author(s):	Rauh, N., Springer-Teumer, S. - Chemnitz University of Technology (TUC), Germany Spaan, M. – Delft University of Technology (TUD), The Netherlands Borowsky. A. – Ben-Gurion University of the Negev (BGU), Israel Fiorentino, A. – Fiat Chrysler Automobiles (FCA), Italy Ahlström, C. – Swedish National Road and Transport Research Institute (VTI), Sweden Thalya, P. – Veoneer (VEO), Sweden

Revision history

Date	Version	Reviewer	Description
03.04.2023	First Draft	Haneen Farah – Delft University of Technology (TUD), The Netherlands	Internal QA review
03.04.2023	First Draft	David Shinar – Ben-Gurion University of Negev, Israel, Department of Industrial Engineering and Management	External QA review
03.04.2023	First Draft	Borowsky. A. – Ben-Gurion University of the Negev (BGU), Israel, Fiorentino, A. – Stellantis, Italy, Ahlström, C. – VTI, Sweden, Thalya, P. – Veoneer, Sweden	Review by co-authors
25.04.2023	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
30.04.2023	Final deliverable	Michiel Christoph (Technical project coordinator) – SWOV Institute for Road Safety Research, The Netherlands	Approved

Table of contents

List of abbreviations

BGU	Ben-Gurion University of the Negev
CM	Continuous Mediation (automation level)
DL	Decision Logic
ECG	Electrocardiogram
FCA	Fiat Chrysler Automobiles
GPS	Global Positioning System
HMI	Human Machine Interface
HRV	Heart Rate Variability
KPI	Key Performance Indicator
KSS	Karolinska Sleepiness Scale
M	Mean value
Md	Median
N	Total number of population
NDRT	Non-Driving Related Tasks
ODD	Operational Design Domain
PAV	Partially Automated Vehicle
SAE	Society of Automotive Engineers
SB	Driver Standby (automation level)
SUS	System Usability Scale
TOC	Transition Of Control
TUC	Technische Universität Chemnitz / Chemnitz University of Technology
TUD	Technische Universiteit Delft / Delft University of Technology
TTAF	Time to automation fitness
TTAU	Time to automation unfitness
TTDF	Time to driver fitness
TTDU	Time to driver unfitness
UC	Use Case
VEO	Veoneer
VTI	Swedish National Road and Transport Research Institute
WoOz	Wizard of Oz
WP	Work Package

List of tables

List of figures

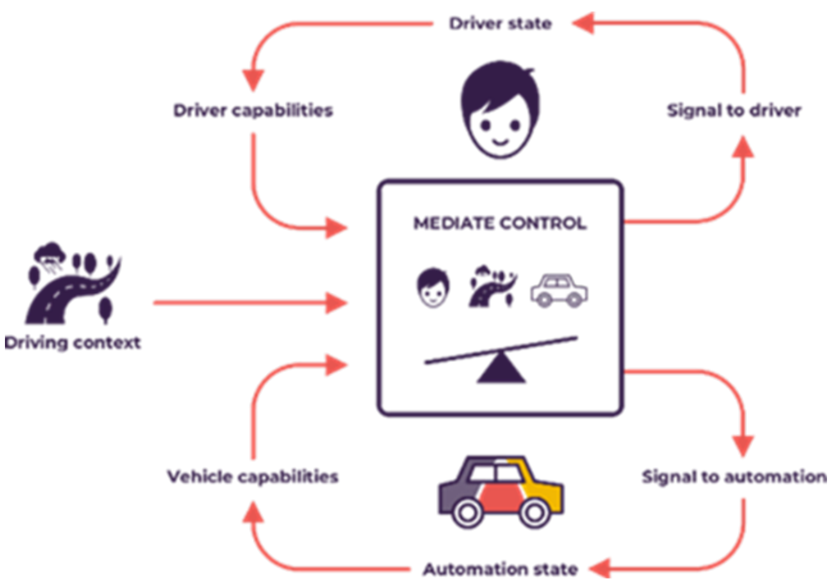
About MEDIATOR

MEDIATOR, a 4-year project coordinated by SWOV Institute for Road Safety Research, has come to an end after four years of hard work. The project has been 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, represented all transport modes, maximising input from, and transferring results to aviation, maritime and rail (with mode-specific adaptations).

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 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.

The MEDIATOR project aimed to develop an in-vehicle system, the Mediator system, that intelligently assesses the strengths and weaknesses of both the driver and the automation and mediates between them, while also taking into account the driving context. It assists the timely take-over between driver and automation and vice versa, based on who is fittest to drive. This Mediator system optimises the safety potential of vehicle automation during the transition to full (level 5) automation. It would reduce risks, such as those caused by driver fatigue or inattention, or on the automation side by imperfect automated driving technology. MEDIATOR has facilitated market exploitation by actively involving the automotive industry during the development process.

To accomplish the development of this support system MEDIATOR integrated and enhanced existing knowledge of human factors and HMI, taking advantage of the expertise in other transport modes (aviation, rail and maritime). It further developed and adapted available technologies for real-time data collection, storage and analysis and incorporated the latest artificial intelligence techniques. MEDIATOR has developed working prototypes, and validated the system in a number of studies, including computer simulation, virtual reality, driving simulator and on-road studies.

With MEDIATOR we further paved the way towards safe and reliable future vehicle automation that takes into account who is most fit to drive: the human or the system.

https://mediatorproject.eu/

Executive summary

The Mediator system is an intelligent driver assistant system for SAE levels 2-4 driving automation mediating in real-time between the human driver and the automated system. Based on which is fittest to drive, the Mediator system will propose or request a change of driving mode and will ensure a safe handover / takeover action. In Work Package 3 (WP3) “Testing & Evaluation”, seven studies were conducted to evaluate the functionality of the Mediator system including assessing the system’s performance, reliability, and effects on driving safety, as well as the acceptance, trustworthiness, perceived safety, and user-friendliness of different user groups.
This deliverable presents the aggregated and integrated results of all conducted studies (three driving simulator studies, one computer simulation study without real users, three on-road studies) with respect to different topics that are important for the evaluation of a system like Mediator. Public deliverables D3.2 (Athmer et al., 2022), D3.3 (Borowsky et al., 2023) and D3.4 (Fiorentino et al., 2023) provided the basis for this deliverable and will provide further details on examined research questions, used methods and revealed results.

The overall evaluation of the Mediator system was positive, and participants evaluated the Mediator system as beneficial for (partly) automated driving.

Participants indicated quite high acceptance of the Mediator system, quite high intention to buy, use and recommend the system, quite high trust in the Mediator system, quite high experienced comfort while driving with the Mediator system, and rated perceived safety as medium to good. Additionally, participants expect an improvement in driving safety when driving with Mediator. Especially, the time budget information was rated as reliable information. Further, the time budget information was evaluated as helpful providing added value for partly automated driving. Nevertheless, the time budget needs to be explained (because it is not intuitively understandable) and, maybe even, be represented in a better way (e.g., presentation of time to automation fitness but also of upcoming automation unfitness in parallel). The timing of Mediator’s suggestions and requests were experienced as good to very good. Drivers rated the provided messages as understandable and the amount of provided information as mostly sufficient and not overloading. The workload caused by Mediator was rated as low. Further, drivers evaluated the transitions from one driving mode to another positively with instructions that are easy to understand and to follow. Nevertheless, some drivers would like to have more information (e.g., why a mode change is possible or necessary) and better transparency. Over all studies, Mediator’s usability was perceived as good with a tendency to excellent. All HMI elements were evaluated positively, especially the light concept to increase mode awareness and acoustic and haptic feedback for potentially dangerous situations. Further, it could be shown that drivers would appreciate fatigue as well as distraction detection with respective warnings.

Regarding the decision logic, the computer simulation study revealed that a more conservative approach for the DL should be prioritized to avoid unsafe or critical driving situations.

Important prerequisites for drivers accepting, trusting and using the Mediator system are, for instance, the fulfilment of drivers’ expectations regarding the system’s functionalities and limitations, the possibility for personalization, the final price and high transparency as well as reliability of the system.

1. Introduction

1.1. Main aims and summary of use cases

Within the scope of the MEDIATOR project, an intelligent mediating system enabling safe switching between the human driver and the automated system was developed. Work Package 3 (WP3) “Testing & Evaluation” aimed to evaluate the functionality of the new developed Mediator system (in the following referred to as Mediator) conducting a computer simulation study (deliverable D3.2: Athmer et al., 2022), driving simulator studies (deliverable D3.3: Borowsky et al., 2023) and on-road studies (deliverable D3.4: Fiorentino et al., 2023). During the evaluations, Mediator’s performance, reliability, and effects on driving safety, as well as drivers’ acceptance, trust, perceived safety, and experienced user-friendliness were assessed taking different user groups into account. The research carried out on Mediator Human Machine Interface (HMI) is complementary. Hence, the studies focused on different elements and functionalities of Mediator, and examined different research questions and use cases. This deliverable is focused on the integration of results achieved within all studies that were conducted in WP3. Thereby, results will be summarized and compared by topics.

Since each study implemented different driving modes and focused on different use cases, the following section defines the applied driving modes and gives a short overview of all use cases.

In deliverable D.1.1 (Christoph et al., 2019) three different automation levels were defined for the MEDIATOR project that rely on a human-centred approach focusing on drivers’ needs and challenges as well as on the support required. The three automation levels are as follows:

1. **Continuous Mediation (CM) – Driver in the Loop** describes assisted driving (similar to SAE (Society of Automotive Engineers) level 2), where the driver is responsible but is supported by the automation. There is a division of tasks between driver and automation: The automation generally performs all active control tasks, while the human driver takes on the task of monitoring and maintaining sufficient situation awareness. A takeover by the driver need to be performed immediately in case the automation becomes unfit (even without a takeover request). Challenges in this level of automation are creating mode awareness and supporting the driver with their part of the driving task by creating an optimal task load.
2. **Driver Standby (SB) – Short Out of the Loop** describes conditional automation (similar to SAE level 3), where the driver can be out of the loop for some periods of time but can be requested to takeover anytime when the automation becomes unfit. Thereby, the takeover request will be triggered several seconds in advance. The driver has to remain ‘on standby’ to take back control. Challenges in this automation level are related to regaining driver fitness and balancing the time until the automation or driver becomes unfit, making sure always one is fit enough for the driving task. This challenge extends to the HMI challenges of communicating these time budgets and mediating comfortable and safe takeovers over a relatively short time span.
3. **Time to Sleep (TtS) – Long out of the Loop** describes driving with high automation (similar to SAE level 4 and higher), where the driver can be out of the loop for long periods of time. Drivers can then truly immerse themselves in non-driving related tasks and even fall asleep. Challenges in this level of automation are to bring the driver back in the loop after full

disengagement and to predict when this will be required long enough in advance. A takeover request is triggered several minutes in advance.

In total, ten uses cases (UCs) were defined (see D1.4, Cleij, et al., 2019) for the MEDIATOR project:

1. Mediator system initiates takeover (human to automation): Degraded human fitness, caused by either drowsiness (a) or distraction (b), is detected by the Mediator system. The system reacts by initiating a takeover to automation.
2. Driver takes back control: The driver uses the HMI to indicate a desire to take back. The Mediator system reacts by confirming that the driver is fit enough to drive and guiding the takeover.
3. Comfort takeover (human to automation): Either the driver (a) or the Mediator system (b) initiates a takeover from human to automation.
   a) The driver indicates via the HMI that he/she is not motivated to drive. The Mediator system reacts by confirming the automation fitness and guiding the takeover.
   b) 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.
4. Corrective Action (SB): While driving in SB the human driver becomes drowsy. The Mediator system reacts by initiating an action to improve the driver fitness and monitors the effect.
5. Mediator initiated takeover (automation to human): A planned (a) or an unplanned (b) takeover from automation to human is initiated by the Mediator system.
   a) 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.
   b) The automation indicates that its fitness is rapidly degrading and can soon no longer perform the driving task. The Mediator system reacts by informing the human driver and guiding the urgent takeover.
6. Comfort CM switch on: Either the driver (a) or the Mediator system (b) switches on driving in CM.
   a) The driver indicates via the HMI that he/she is not motivated to drive. The Mediator system reacts by confirming the automation fitness and switches on CM.
   b) 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 switching on CM.
7. Preventive Action (CM): While driving in CM, the driver is supported by the Mediator system in performing the monitoring task. The system does this by trying to prevent underload and keeping the driver in the loop.
8. Corrective Action (CM): While driving in CM, degraded driver fitness is detected by the Mediator system. The system reacts by initiating a corrective action to improve driver fitness.
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.
10. Smooth transition from TtS to SB: while driving in TtS the driver is fully disengaged from the driving task when the automation indicates that the current route will leave the operational design domain. The Mediator system detects sufficient automation fitness for driving in SB and reacts by informing the driver that SB will be switched on and subsequently monitors the required driver fitness.

1.2. Overview of WP3 studies

An overview of the different studies (e.g., investigated use cases and research focus, recruited participants, used materials and data, performed procedure and achieved results) can be obtained from Table 1 and Table 2. Thereby, Table 1 contains the overview of three driving simulator studies and one computer simulation study, and Table 2 contains the overview of three on-road studies.

Table 1. Overview of driving simulator and computer simulation studies conducted within WP3 including study focus, use cases, participants, material, procedure, data, and topics of provided main findings.

	Driving Simulator Study on driver fatigue (BGU)	**Driving Simulator Study on driver distraction (BGU)**	**Driving Simulator Study on driver comfort (TUC)**	**Computer Simulation Study on decision logic (TUD)**
Focus	Effects of preventive & corrective actions on fatigue + hazard perception, longer term effects over 2 distinct sessions (1 week apart)	Effects of preventive actions on distraction + hazard perception, longer term effects over 2 distinct sessions (1 week apart)	Comfort, Transition of control, HMI, automation degradation, driver characteristics, user evaluation	Technical evaluation, simulation + evaluation High-Level Decision Logic (DL), based on data from the on-road study on automation state (VEO)
Use cases	1-8 (4 each participant)	1-8 (4 each participant)	2, 3, 5, 6, 9	All, reduced to 5 base UC for simulation
Participants	*N* = 24 (in each session; 50% male: *M* = 27 years, *SD* = 7; 50% female: *M* = 25, *SD* = 3 years)	*N* = 26 (in each session; 50% male: *M* = 26 years, *SD* = 1; 50% female: *M* = 24 years, *SD* = 1)	*N* = 74 (*M* = 40 years, *SD* = 17; 4 age groups: 18-24 y/o, 25-39 y/o, 40-54 y/o, ≥55y/o), balanced age + gender	none
Material	Hazard perception scenarios, Trivia game as preventive + corrective action against fatigue (based on KSS scores)	Hazard perception scenarios, WhatsApp-like task as distractor	HMI implemented by TUC, simulation of defined events	On-road data from VEO vehicle
Procedure	4 standardized driving scenarios each session, Trivia game as intervention vs. no intervention	4 standardized driving scenarios each session, WhatsApp-like task as distractor for experimental group in hazard perception scenarios vs. no dis-tractor in control group	Same driving scenario in 4 conditions (manual = baseline, CM, SB + non-driving related task (NDRT), SB + uncomfortable approach + NDRT)	Computer simulation (Reinforcement Learning, decision trees), DL implement-tation, evaluation key performance indicator (KPI), sensitivity analysis
Data	KSS scores (fatigue), ECG, eye tracking, Trivia game data, driving performance data, questionnaires	ECG, eye tracking (distraction), WhatsApp-like task data, driving performance data, questionnaires	Driver videos, driving data including HMI, questionnaires, interviews	Decision accuracy, driver + driving + automation context data, real world data
Main findings provided for the following topics	Acceptance & intention to use, trust & perceived reliability, perceived safety & expected impact on driving safety, usability, HMI evaluation, degraded driver states, workload & distraction	Acceptance & intention to use, perceived comfort, usability, HMI evaluation, workload & distraction	Acceptance & intention to use, trust & perceived reliability, perceived comfort, perceived safety & expected impact on driving safety, usability, take-over and hand-over actions, timings, distraction detection	DL, sensitivity analysis to test DL's robustness, reinforcement learning as alternative to decisions trees used by DL

Table 2. Overview of on-road Studies conducted within WP3 including study focus, use cases, participants, material, procedure, data, and topics of provided main findings.

	On-road study on HMI (FCA)	**On-road study on driver impairments and degraded automation (VTI)**	**On-road study on automation state (VEO)**
Focus	HMI (takeover requests, automation levels, manual driving), user evaluation	Functionality + user evaluation, fatigue + distraction, degraded automation, HMI (Mediator + baseline)	Technical evaluation of Automation module and DL module, interviews, questionnaire, evaluation of Driver State component + HMI, several drives over a longer period of time
Use cases	1a, 1b, 2, 3a, 3b, 4, 5a	1, 4, 5a, 8, 9, 10	5, 6, 9
Participants	*N* = 16 (*M* = 40 years, *SD* = 14; 56% male, 44% female; 2 age groups: 23-30 y/o, 50-60 y/o)	*N* = 50 (*M* = 40 years, *SD* = 14; 64% male, 36% female; 37 alert, 13 sleep-deprived 2 drives	*N* = 7 (*M* = 42 years, *n* = 1 female) professional drivers
Material	FCA Wizard of Oz (WoOz), Mediator HMI	FCA WoOz, 2 drives (baseline + Mediator HMI)	TI in-vehicle prototype, Mediator system (with different HMI configurations)
Procedure	46 km public road, 1 trip: Mediator HMI, NDRT, 7 scenarios	82 km public road, 10 scenarios, 2 drives (HMI: Mediator vs. baseline), 2 NDRTs	Standardized route public road (incl. stretches that affect automation performance)
Data	Thinking aloud, driving performance data, eye tracking (distraction), interview, questionnaires	Driver videos, driver state AttenD (distraction detection algorithm), driving performance data, physiological data, eye tracking, KSS scores, reaction time, questionnaires, interview	Vehicle data, HMI interactions, interviews, questionnaires
Main findings provided for the following topics	Acceptance & intention to use, trust & perceived reliability, usability, HMI evaluation, take-over and hand-over actions, degraded driver states	Acceptance & intention to use, trust & perceived reliability, perceived comfort, perceived safety & expected impact on driving safety, usability, HMI evaluation, take-over and hand-over actions, timings, degraded driver states, workload & distraction	Acceptance & intention to use, trust & perceived reliability, perceived comfort, perceived safety & expected impact on driving safety, HMI evaluation, take-over and hand-over actions, timings, degraded driver states, workload & distraction, time budget calculation

Further details on the methods and questionnaires applied in each study as well as an indication towards the according chapter where this is covered can be obtained from Note that the numbering of the studies is based on the following scheme:

- Study 1: Driving Simulator Study on driver fatigue – BGU (Israel)
- Study 2: Driving Simulator Study on driver distraction – BGU (Israel)
- Study 3: Driving Simulator Study on driver comfort – TUC (Germany)
- Study 4: Computer Simulation Study on decision logic– TUD (Netherlands)
- Study 5: On-road study on HMI – FCA (Italy)
- Study 6: On-road study on driver impairments and degraded automation – VTI (Sweden)
- Study 7: On-road study on automation state – VEO (Sweden)

Table 3 (see Appendix A). Note that the numbering of the studies is based on the following scheme:

- Study 1: Driving Simulator Study on driver fatigue – BGU (Israel)

- Study 2: Driving Simulator Study on driver distraction – BGU (Israel)
- Study 3: Driving Simulator Study on driver comfort – TUC (Germany)
- Study 4: Computer Simulation Study on decision logic – TUD (Netherlands)
- Study 5: On-road study on HMI – FCA (Italy)
- Study 6: On-road study on driver impairments and degraded automation – VTI (Sweden)
- Study 7: On-road study on automation state – VEO (Sweden)

1.2.1. Driving Simulator Studies on driver fatigue and driver distraction

The aims of the driving simulator studies on driver fatigue and driver distraction (in the following referred to as “study 1” and “study 2”) conducted under the lead of Ben-Gurion University of the Negev (BGU, Israel) were to evaluate

1. whether the Mediator HMI concept of preventive and corrective mediation assists in mitigating the adverse effects of distraction and fatigue,
2. long term effects of preventive and corrective mediation, and
3. participants’ acceptance, trust, and perceived usefulness of Mediator.

For this purpose, both studies used a fixed-base driving simulator consisting of an engineless vehicle and a 165° projection of the virtual world on a curved screen. In both studies, the Mediator HMI includes two interfaces (s. Figure 1), i.e. the instrument panel (displaying icons and speedometer) and the central stack (touchscreen, driver information, presentation of non-driving related tasks).

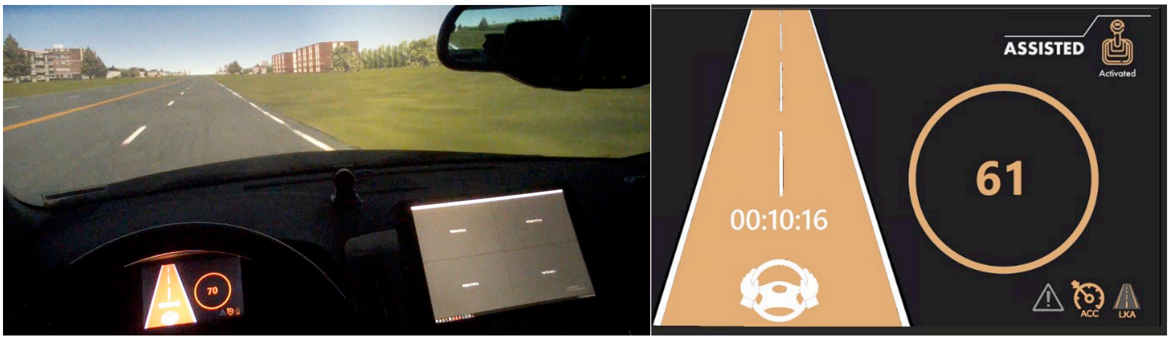


Figure 1. Dashboard and infotainment displays (adopted from Borowsky et al., 2023).

The simulated vehicle was designed in two modes: Manual driving mode and CM driving mode. All aspects of the driving task in the manual driving mode were controlled by the participants. The partially automated vehicle (PAV) controlled the car’s speed, acceleration, time headways from lead vehicles, and lane position during CM driving mode. The route in the fatigue and distraction studies included highways and urban environments in daylight. During the drive, participants navigated eight latent hazard scenarios, which were distributed throughout the drives. However, the main difference between the studies is the role of the secondary task. In the *fatigue* study, the secondary task is used as a countermeasure for fatigue (based on outcomes of KSS), and in the *distraction* study the secondary task is the distractor itself.

The focus of the *fatigue* study was on prolonged driving to induce passive fatigue. A secondary task consisting of an auditory Trivia game was used as countermeasure for fatigue. Each driving session lasted about 40 minutes and included a pre-defined set of four un-materialized hazardous scenarios, 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). Underload was induced by embedding road sections of eight minutes that included sparse traffic before each hazardous event, meanwhile drivers did not have to intervene. Participants were randomly assigned to one out of two experimental conditions: 1) CM driving with Trivia and notifications of upcoming hazards or (2) CM driving with notifications of upcoming hazards only (no Trivia). Participants were invited to two distinct sessions (1 week apart).

The focus of the *distraction* study was on distracted driving and the benefits of applying preventive mediation to mitigate the effects of distraction on driving performance and hazard perception. Participants were randomly assigned to two CM driving sessions each lasting about 20 minutes and including non-driving related tasks (NDRTs). To induce distraction in the experimental group, the content of information visualized via the secondary task (WhatsApp-like visual-manual task) was presented much longer and was more demanding than in the fatigue study. The control group drove the same scenarios without performing the secondary task. For the preventive mediation, only the notifications of upcoming hazards were used. The road segments connecting between hazards were shorter than in the fatigue study. Again, participants were invited to two distinct sessions (1 week apart).

It needs to be taken into account, that the study results reflect an optimal working Mediator and, thus, represent its highest possible potential. Despite driving simulator studies do not reach the same external validity as on-road studies, their decisive advantage is the controllability of the experimental environment and the guaranteed safety of the participants.

For further information, see deliverable D3.3, chapter 2 (Borowsky et al., 2023).

1.2.2. Driving Simulator Study on driver comfort

The driving simulator study conducted under the lead of Chemnitz University of Technology (TUC, Germany; in the following referred to as “study 3”) focused mainly on

1. comfort transitions of control (TOC) from manual to automated driving,
2. simulated automation degradation and related TOCs by the human driver,
3. comfort critical situations (i.e., close approach to the rear-end of a traffic jam), and
4. the influence of driver characteristics.

Thereby, an “optimal working” Mediator was presented to the participants avoiding unplanned HMI reactions and false alarms. All drivers experienced four different conditions of the same driving scenario in a fixed order: Baseline, continuous mediation (CM), driver standby (SB), and driver SB with close approach. In all conditions of the driving scenario, drivers started on a highway in manual mode and, corresponding to the predefined condition, either continued driving in that mode or switched to an automated driving mode (i.e., CM or SB). After several kilometres, participants approached a white truck at the rear-end of a traffic jam, came to a full stop, drove some kilometres in the traffic jam, and afterwards, continued the trip on the highway. After each condition, participants evaluated, for instance, the acceptance of and trust in Mediator, its functionalities and usability as well as the experienced TOCs.

For the study, a fixed-base driving simulator with a mock-up front vehicle section was used including a 180° field of view projection of the simulation (s. Figure 2).

Figure 2. Driving simulator at TUC (adopted from Borowsky et al., 2023).

The HMI concept included a screen behind the steering wheel displaying driving relevant data (e.g., speedometer, driving mode, time to automation unfitness), instructions for the driver (e.g., keep hands on steering wheel), and information about the context (e.g., detection of traffic jam). Further, the HMI concept included 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 (e.g., by different colours, brightness and pulsation frequencies; s. Figure 3).

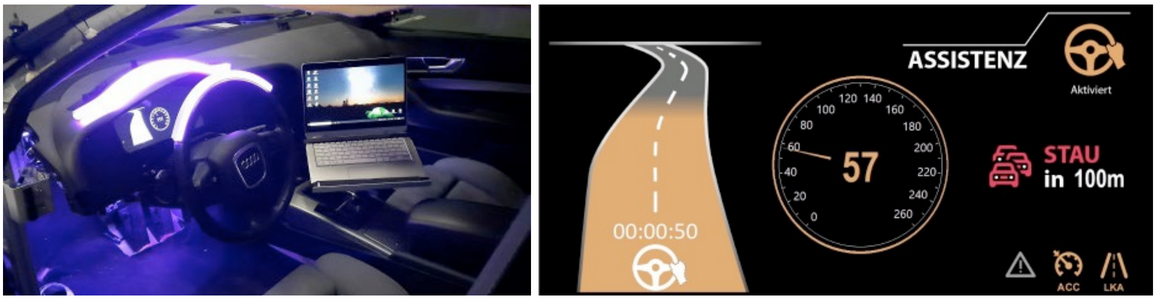


Figure 3. Left: HMI concept for the TUC lab prototype in Mediator SB mode. Right: Example of the TUC lab prototype system in Mediator CM mode (adopted from Borowsky et al., 2023).

A sound system for sound alerts and spoken messages as well as a laptop mounted at the centre console to run the secondary task were also part of the HMI concept.

Note that the results of study 3 also reflect the participants' perception and evaluation of an optimal working Mediator, which in turn demonstrate the systems' highest possible potential. However, as with driving simulation studies, the simulated environment might have biased the participants' feeling of safety and risk perception. Thus, when interpreting results, this fact needs to be taken into account.

For further information, consider deliverable D3.3, chapter 3 (Borowsky et al., 2023).

1.2.3. Computer Simulation Study on decision logic

The computer simulation study (in the following referred to as "study 4") conducted under the lead of Delft University of Technology (TUD, Netherlands) aimed to develop the high-level decision logic (DL) module of Mediator and a simulation environment. The latter was used to simulate all aspects of (semi-) autonomous driving on the road and to evaluate and improve simulation accuracy as well as the DL. Besides different Mediator components developed by various partners in the project, real-world driving data from the TI in-vehicle prototype vehicle (study 7) including driving context, automation context, driver context, and actions served as input for the DL model and the simulation. The goal was to generate proposals and requests for a specific driving mode based on who is fittest to drive (system vs. driver) in order to improve safety and comfort using logical trees with a modular structure. Tests also included a sensitivity analysis to determine, for instance, the systems' capability of dealing with noisy (i.e., wrong) observations or the effect of optimistic and pessimistic estimates for available automation levels.

For more information, see deliverable D3.2 (Athmer et al., 2022).

1.2.4. On-road Study on HMI

The Italian on-road study (in the following referred to as “study 5”) conducted under the lead of Fiat Chrysler Automobiles (FCA) aimed at evaluating usability, acceptance, and perceived trustworthiness of the Mediator HMI (for level CM and SB) on a public road implementing a Wizard of Oz (WoOz) setup to simulate vehicle automation. Even though participants do not have any real control of the vehicle, which may bias their evaluation of Mediator, this solution enables running experiments on automated driving on public roads by simulation higher levels of automations that are not yet available. Hence, in study 5, a driving wizard (experimenter) controlled the vehicle and sat on the right-hand side while participants sat on the left-hand side without any control of the vehicle. They were instructed to pretend to drive in a real automated vehicle. Participants experienced Mediator during an on-road session while they experienced the defined use cases. Meanwhile, eye-tracking data were used to evaluate the simulated driver distraction that experimenters asked for from the participants. Participants were asked to perform a NDRT and use the Think Aloud technique by answering specific questions regarding the HMI. Further, before and after the on-road sessions, participants completed standard questionnaires. The Mediator HMI consisted of several elements (s. Figure 4), namely:

1. a cluster behind the steering wheel presenting icons and messages depending on use case and driving mode,
2. a head unit interface in the centre of the dashboard showing the navigation map and different messages depending on use case,
3. LEDs on the non-functional steering wheel that changes colour depending on the driving mode,
4. vocal messages giving advice depending on use case and driving mode,
5. short and long acoustic feedback, and
6. haptic feedback via seat belt depending on use case.

Figure 4. Left: HMI stimuli on cluster, head unit and LEDs on steering wheel. Right: Automated driving mode icon (adopted from Fiorentino et al., 2023).

When interpreting participants’ answers it need to be taken into account that the distraction detection algorithm operates in real-time and thus triggers immediate alerts. However, the decision logic runs in 1Hz, which causes long delays of up to several seconds between the detection and the warning. This delay has implications on the transparency of the system, making it difficult for the participants to understand the warnings. Further, participants’ experience of Mediator and automated driving might be influenced by the WoOz-setup (e.g., presence of a driving wizard, no manual driving after respective transitions).

See deliverable D2.10 (Fiorentino et al., 2022) for a detailed technical description of the HMI and deliverable D3.4, chapter 2 for further information regarding this on-road study (Fiorentino et al., 2023).

1.2.5. On-road Study on driver impairments and degraded automation

In the on-road study conducted under the lead of Swedish National Road and Transport Research Institute (VTI, Sweden; in the following referred to as “study 6”), Mediator was evaluated on Swedish public roads in terms of its functionality, safety effects, and user acceptance under degraded driver performance conditions (distraction and fatigue) and degraded automation. For this purpose, the same HF in-vehicle prototype was used as in study 5 (see 1.2.4; cf. deliverable D.3.4, chapter 1.4 and deliverable D2.10). The road segment was chosen so that the driving modes CM, SB and Time to Sleep were covered. Each participant took part in the study on two different days in order to experience both the Mediator and baseline condition. Questionnaires were applied before and after the on-road session. The total test session for one condition (requiring about 3 hours including 1 hour of driving) ended with an interview to gather information on why participants appreciated or disliked the system. The Mediator HMI aimed to establish mode awareness by communicating the current as well as the upcoming automation mode to the driver. For this purpose, a colour scheme entailing yellow/amber for CM and purple/magenta for SB and Time to Sleep was implemented on all screens via ambient light as well as via LED strips on the dashboard and on the steering wheel (s. Figure 5).

Figure 5. Left: LED strips and ambient lights visualizing Driver Standby. Right: Visual warning messages shown to distracted participants at levels 1 (top) and 3 (bottom).

Furthermore, the corresponding automation mode was also displayed via icons. Participants could shift between the different automation modes by using a gear shifter that changed the colour similar to all HMI elements. In addition, a stylized road widget including colours, timers (time budget), and icons were used to convey information regarding the current and upcoming automation mode to the participant. For the manual driving condition, no time budget was presented. Whenever Mediator detected distraction, a degraded fitness message and distraction icon occurred on the displays accompanied by an audio alert (s. Figure 5). If necessary, a warning escalation started by a second notification and stronger audio alert. An emergency take-over was triggered on the highest escalation level accompanied by a start of LED-pulsation in red colour. A similar procedure was run for detected fatigue. The baseline HMI was a down-scaled version, whereby the touch screen was used instead of the gear shifter, a simple icon was presented for the current automation mode, take-over routines made use of simple sound and icon, distraction detection was turned off, and a coffee cup icon with a simple sound was presented as fatigue alerts.

For more details, see deliverable D3.4, chapter 3 (Fiorentino et al., 2023).

1.2.6. On-road Study on automation state

In the on-road study conducted under the lead of Veoneer (VEO, Sweden; in the following referred to as “study 7”) the fully realised technical Mediator Proof-of-Concept system was evaluated in terms of functionality and user acceptance. In detail, this is the evaluation of the automation state component under different driving contexts (e.g., how well can the system predict bad automation performance), the driver state component, the DL, and the HMI. The test vehicle was equipped with a CM system (adaptive cruise control + lane keeping assistance). As exemplarily depicted in Figure 6, a display at the centre stack displayed driving relevant data (e.g., driving mode, time to automation unfitness), instructions for the driver (e.g., keep hands on steering wheel), and information about the context (e.g., construction zone). Additional LED strips on the dashboard and the steering wheel as well as ambient lighting indicated the current driving mode, the intended or necessary changes in the driving mode as well as remaining time (e.g., by different colours, brightness and pulsation frequencies). Further, a sound system for sound alerts was implemented.

During the test session, professional test drivers drove on a pre-defined route twice in a row (overall ~2 h driving time). The route included road segments affecting automation performance and where the driver is assumed to deactivate assisted driving, or where automation would reach its limitations and deactivate itself. Each driver participated on average five times distributed over different days and performed overall ten drives. The drives were divided into four blocks that reflected the following different configurations of Mediator: (1) “complete Mediator system”, (2) very basic system without proactive automation and time budget information”, (3) “shorter time budget and more frequent TOC suggestions”, (4) “complete Mediator system with distraction detection”. After drive two, six, and 10, interviews were conducted with the participants regarding, for instance, the drivers’ experience of different configurations, their perception of the system and its functionalities.

Figure 6. Top: LED strips used inside the cabin of the test vehicle to show the driving mode of the in-vehicle prototype. Bottom: HMI example shown in the centre stack of the test vehicle (adopted from Fiorentiono et al., 2023).

Since this study entailed a small sample of seven participants, there is a larger variance between the drivers' responses. Further, results have a limited validity regarding the representativeness on a population level. However, this sample size suits the purpose of the exploratory study with a specific and rather rare sample, i.e., professional test drivers.

For more details, see deliverable D3.4, chapter 4 (Fiorentino et al., 2023).

1.3. General result integration strategy

In this deliverable, the integrated results for each examined topic will be summarized, compared and discussed. Public deliverables D3.2 (Athmer et al., 2022), D3.3 (Borowsky et al., 2023) and D3.4 (Fiorentino et al., 2023) provided the basis for this deliverable. Therefore, only results provided in the deliverables were considered. All results were collected and integrated to provide an overall evaluation of Mediator. Since different questionnaires or no questionnaires but interview questions were used to assess the same constructs in the underlying studies, the results are integrated, compared, and discussed at a descriptive level. When available, concrete measures of central tendency were integrated. However, in the underlying studies different measures of central tendency were provided. Therefore, both mean (*M*) and median (*Md*) values are reported in this deliverable. Each chapter provides (1) a short introduction to emphasize the importance of the examined topics, (2) an overview of the integrated results and (3) a short summary.

The first chapter presents results on acceptance and intention to use (2.1) followed by a chapter focusing on results regarding trust and perceived safety (2.2). The third chapter discusses the results on perceived comfort (2.3). Integrated results on perceived safety and drivers' expectations regarding Mediators impact on driving safety will be presented in the fourth chapter (2.4). The following chapter will summarize the results on usability and user friendliness, as well as users' evaluation of the Mediator HMI including the experienced impact on mode awareness (2.5). In the following chapters, results on Mediator's timings and time budget information (2.6), drivers' experiences during take-over and hand-over actions (2.7), and experienced workload and distraction caused by Mediator (2.8) will be presented. The following chapter discusses Mediator's influence on drivers' fatigue and distraction (2.9). The last chapter presents the results from the computer simulation study regarding the decision logic (2.10). A last chapter summarizes the results, provide conclusions and gives an outlook to deliverables providing concrete implications of Mediator results.

Further details on used methods and questionnaires for each topic in each study are shown in Note that the numbering of the studies is based on the following scheme:

- Study 1: Driving Simulator Study on driver fatigue – BGU (Israel)
- Study 2: Driving Simulator Study on driver distraction – BGU (Israel)
- Study 3: Driving Simulator Study on driver comfort – TUC (Germany)
- Study 4: Computer Simulation Study on decision logic– TUD (Netherlands)
- Study 5: On-road study on HMI – FCA (Italy)
- Study 6: On-road study on driver impairments and degraded automation – VTI (Sweden)
- Study 7: On-road study on automation state – VEO (Sweden)

Table 3 (see Appendix A).

2. Integration of MEDIATOR study results

2.1. Acceptance & Intention to use

Introduction: A technology designed to make driving safer, more efficient and more comfortable is only successful if it not only operates in a technically flawless manner but is also accepted and used by its target groups (Davis, 1989; Park et al., 2014). Mediator is a system that provides active proposals for a driving mode change to human drivers. To reveal the full potential of Mediator, it is crucial that drivers will follow Mediator's suggestions. Hence, assessing drivers' acceptance of and intention to use Mediator in their everyday life is an important prerequisite to reach the intended aims of the system (Kassner & Vollrath, 2006; Adell et al., 2014).

Results: In one driving simulator (*study 3*) and one on-road study (*study 5*), participants were asked about their general opinion on vehicle automation and acceptance of Mediator before actual usage. It was found that participants had quite high acceptance rates for vehicle automation (e.g., $M = 4.07$ on a 5 point Likert scale ranging from 1 to 5), and the expected acceptance of Mediator before actual usage was quite high (e.g., $M = 0.97$ on a 5 point Likert scale ranging from -2 to +2). After repeated exposure to Mediator, participants indicated significantly higher acceptance rates compared to the initial ratings. Overall acceptance was rated as quite high in all conducted studies ($4.6 <= M <= 5.9$ on a 7 point Likert scale ranging from 1 to 7, or $0.7 <= M <= 1.2$ on a 5 point Likert scale ranging from -2 to +2).

Acceptance was even higher for drivers who experienced Trivia and hazard notifications as preventive mediation (*study 1*). However, uncomfortable driving situations like a close approach to the rear-end of a traffic jam can cause a decrease in acceptance (*study 3*). Potential reasons are decreased perceived comfort and safety, maybe even trust, when experiencing the automated vehicle performing a quite late and sudden brake. Nevertheless, even after experiencing a close approach, the overall acceptance rating for Mediator was still high. Two driving simulator studies (*study 1* and *study 2*) revealed gender effects (male indicated higher acceptance compared to female drivers), and one on-road study (*study 6*) revealed age effects (younger drivers expressed higher preferences for Mediator). Further, one study (*study 3*) could show that higher affinity for technology and more positive general opinion about driving automation are positively related to acceptance of and intention to use Mediator. Nevertheless, the influence of driver characteristics like age, gender or attitudes could not be found in all studies. Possible reasons are either that the Mediator concept was successfully designed mostly independently of user characteristics, and / or that self-selection processes led to specific user samples with, for instance, positive opinions and expectations outweighing possible effects of driver characteristics.

Over all studies, participants indicated quite high intentions to buy and / or use Mediator in their everyday lives. One driving simulator study (*study 3*) revealed that 88% of the drivers would want to use Mediator in their car. One on-road study (*study 6*) revealed that 71% would like to buy Mediator given that it was working as expected and the final price is acceptable. Participants in another on-road study (*study 7*) mentioned further prerequisites like the chance for personalization (e.g., shut off distraction warning or beeping sounds) and improvement of the system (e.g., distraction warning). Especially, professional drivers discussed the added benefit of Mediator for CM driving critically and indicated that they would use Mediator in their future cars only for higher automation levels (*study 7*). Participants of another on-road study (*study 6*) indicated high willingness to recommend and to buy Mediator as well (*Md* = 4 on a 5-point Likert scale ranging from 1 to 5).

Summary: Over all studies, participants indicated quite high acceptance of Mediator as well as a quite high intention to buy, use and recommend Mediator. Important prerequisites are high reliability, fulfilment of drivers' expectations, personalization and the final price. Different study conditions (driving simulator vs. on-road study), samples (naïve vs. professional drivers) as well as driver characteristics had no systematic and replicable effects on drivers' acceptance ratings.

2.2. Trust & Perceived reliability

Introduction: Next to drivers' acceptance of Mediator, trust in Mediator is also an important aspect that needs to be taken into account since it does not only predict whether an automated system will be used, but also how it will be used (Parasuraman & Riley, 1997). If trust exceeds the capabilities of the system, over-trust can be the result, which leads to misuse (Parasuraman et al., 1996; Riley, 2018). Misuse, in turn, results in excessive use of the system, which becomes particularly problematic when the system is used in situations for which it was not intended. In contrast, a lack of trust may lead to disuse (Parasuraman et al., 1996; Riley, 2018), which manifests itself in the fact that a system that is actually reliably working is under utilized. A person, who experiences, for instance, that warnings are missing in critical situations or that a system reacts too sensitively, might not trust the system in an emergency and will reject the possibilities of automation (Manzey, 2012). In general, it can be assumed that the higher the trust in a technical system, the more intensively it will be used resulting in the desired positive effects of automation: Increased safety as well as comfort and reduced workload (Zhang et al., 2019). Hence, for the appropriateness of usage and usage frequency of a system like Mediator, which provides suggestions and support for the driver of (partly) automated vehicles, trust plays an important role (Hoff & Bashir, 2015).

Results: Trust in Mediator was rated quite high right from the beginning, and further increased with actual and repeated system usage in all conducted studies. Overall, trust was rated as high to very high ($3.9 <= M <= 5.9$ on a 7-point Likert scale ranging from 1 to 7). During interviews (e.g., *study 3* and *study 6*), participants reported that they trusted the system (87% trusted the system fully, 66% trusted it more than the baseline system). They found it reliable and safe, trusted the technology (e.g., sensors), felt confident that the system would detect and cope with the traffic situation and would perform the driving task sufficiently including keeping a safe distance to other vehicles, maintaining the allowed speed, and keeping lane. Participants reported that positive experience with Mediator during the first drives were important for building up trust, and that more experience with the system would help them gain more trust in Mediator, for instance, in real traffic, over longer distances and in dangerous or more complex situations. Additionally, recommendations to use Mediator by others (e.g., family, friends, peers, role models) as well as more information about the system's status, technical details, functionalities and limitations as well as about the driving environment were mentioned as important factors to build up trust.

Most studies revealed on average ratings of 5 to 6 on a 7-point Likert scale indicating high to very high trust. Two driving simulator studies (*study 1* and *study 2*) revealed gender effects with higher ratings provided by male participants compared to female participants, and with quite low trust ratings for female participants in one of the two studies ($M = 3.9$). No gender effects could be detected in the other studies. It could be shown that participants who engaged with Trivia rated trust higher than the non-Trivia group (*study 1*). Further, affinity for technology and drivers' general opinion about vehicle automation could be shown to be positively correlated to drivers' trust in Mediator in one study (*study 3*). Trust ratings decreased after experiencing an uncomfortable driving situation (i.e., close approach to the rear-end of a traffic jam, *study 3*) and reduced system reliability (i.e., distraction detection, e.g., *study 6* and *study 7*). However, drivers' overall trust rating remained quite high even after the close approach scenario and the experienced low reliability of the distraction detection indicating high specificity of drivers' trust (cf. Lee & See, 2004). Nevertheless, in one study (*study 3*), 73% of the participants expressed their wish to monitor Mediator's performance constantly revealing some scepticism or need for control despite the high trust ratings. In one of the studies (*study 6*), participants described some difficulties with fully leaving control to the vehicle.

Drivers' behaviour in terms of glance behaviour and interventions in uncomfortable driving situations (i.e., close approach) allows for some inferences regarding drivers' trust as well. Nearly half of all drivers in one driving simulator study (study 3) did not intervene in the close approach scenario. One possible explanation might be that they trusted Mediator enough to handle the situation. Some participants did not even monitor the traffic situation in SB driving being fully engaged in a secondary task execution. One on-road study (*study 7*) could show that drivers tend to look less frequently on the Mediator screen after repeated exposure to the system indicating that they trust the audio signals enough to feel sufficiently informed about the current driving situation.

Drivers' trust was shown to be closely related to Mediator's reliability (e.g., drivers' trust ratings decreased when reliability was rated lower). In one on-road study (*study 7*), drivers rated Mediator's reliability on average as good especially after the first drives. During the last drives, the reliability was partly rated as quite low. One explanation might be the distraction warning which gave false alarms and, hence, reduced the perceived reliability of the system (cf. Hoff & Bashir, 2015). Overall, reliability rating was not affected too strongly in this study, indicating that drivers can distinguish between the reliability of the distraction warning system and reliability of the whole Mediator system (including, e.g., the time budget). The interview results show that the displayed time budget was rated as reliable information in all studies. By comparing the calculated times to automation fitness and unfitness with the GPS locations of known static operational design domain (ODD) changes in one on-road study (*study 7*), it could be shown that there was generally a good match between the predicted and actual time budget. Detected differences between predicted and actual time budget had a maximum of 60s, and most of them were around 20s.

Summary: Over all studies, participants indicated quite high trust in Mediator. Different study conditions (driving simulator vs. on-road study), samples (naïve vs. professional drivers) as well as driver characteristics had no systematic and replicable effects on drivers' trust ratings. Especially, the time budget information was rated as reliable information.

2.3. Perceived comfort

Introduction: Comfort is considered as one of the main drivers of higher levels of automated driving next to safety, efficiency, social inclusion and accessibility (ERTRAC, 2019). In automated driving conditions, comfort aspects are not only relevant for a pleasant driving experience and acceptance of automated systems, but can also have safety impacts. Unnecessary interventions by the human driver due to uncomfortable situations (e.g., when apparent safety appears as compromised), for instance, could lead to safety-critical and unnecessary take-over situations (Hergeth, Lorenz, & Krems, 2016; Techer et al., 2019). As comfort aspects are primarily related to dynamic situations, constant comfort evaluation is required in order to prevent discomfort by, for instance, adapting automation features such as the driving style.

Results: In one of the on-road studies (*study 7*), participants experienced medium comfort when driving with Mediator. Reduced reliability of Mediator (e.g., oversensitivity of distraction warning and repeated problems with the system) might be one reason for reduced experienced comfort. In another on-road study (*study 6*), less than 60% rated Mediator as more comfortable compared to a baseline system with reduced functionality. Drivers agreed that the take-over procedures felt comfortable. During the driving simulator studies, comfort was rated very high. In one of the driving simulator studies (*study 3*), more than 90% of the drivers agreed that driving with Mediator was comfortable. Especially, the possibility to relax/rest, as well as reduced stress and irritation when using Mediator were mentioned. Participants rated automated driving style positively as well (M = 4.2 on a 5-point Likert scale ranging from 1 to 5, *study 3*), and the possibility to perform secondary tasks while driving was rated quite positively ($M \sim 4$ on a 6-point Likert scale from 1 to 6, *study 3*). Additionally, participants valued Mediator's support for the traffic jam scenarios (*study 3*); although the ratings were slightly lower after experiencing the close approach.

It could be shown that drivers expressed their discomfort during the close approach scenario differently (*study 3*). Behavioural reactions ranged from not noticing up to almost panic reactions with sometimes rather uncontrolled manual interventions. It could be shown that early observable driver features like interruption of NDRT engagement or constant observation of the driving scenario can function as indicators for beginning discomfort. However, drivers' behaviour cannot be generalized as general sign for uncomfortable automated driving situations, but needs to be interpreted in the context of the current driving situation.

One driving simulator study (*study 1*) revealed that participants who experienced Trivia and hazard notifications as preventive mediation perceived Mediator as more convenient compared to the non-Trivia group.

Summary: Participants rated experienced comfort while driving with Mediator quite high. It seems that the study conditions might have an influence, as ratings in driving simulator studies were higher compared to on-road studies. One possible explanation might be that a driving simulator is a controlled and safe environment with all Mediator functionalities programmed in a way that they work in an optimal manner. During on-road studies, on the other hand, some technical problems with the prototype might have caused a decrease in driving comfort. Different samples (naïve vs. professional drivers) or driver characteristics had no systematic and replicable effects on drivers' comfort ratings.

2.4. Perceived safety and expected impact on driving safety

Introduction: Driver's assessment of a vehicle and its functionalities as well as the impression of safety are strongly interdependent. Thus, it is essential to provide an adequate feeling of safety (Heiderich et al., 2016). Moreover, for users' decision-making to buy or use a new car, the subjective evaluation of driving safety, well feeling and comfort are important criteria (Heißing et al., 2013) and were, therefore, investigated within the scope of the evaluation of Mediator.

Results: In one of the on-road studies (*study 6*), 57% of the participants rated Mediator as safer than the baseline system. In another on-road study (*study 7*), participants' perceived safety was on a medium level. In one of the driving simulator studies (*study 3*), 78% of the participants reported that they felt safe while driving with Mediator active. The perceived safety while approaching the end of the traffic jam was rated quite good, although the rating was lower after the close approach (*study 3*). Additionally, participants felt quite safe while the handover from manual to automated driving when comfort-related transfers were performed (*M* = 3.8 on a 5-point Likert scale ranging from 1 to 5, *study 3*). One on-road study (*study 6*) revealed that participants rated the safety of the handover rituals when Mediator actively proposed handovers to higher levels of automation as neutral. In one driving simulator study (*study 3*), participants felt quite safe while the takeover from automated to manual driving was performed (*M* = 3.9 on a 5-point Likert scale ranging from 1 to 5). During one on-road study (*study 6*), the transfer of control from automated to manual driving was rated as safe by the drivers as well (*Md* = 4 on a 5-point Likert scale ranging from 1 to 5).

No considerable effects of driver characteristics on perceived safety were revealed, but drivers' general opinion about vehicle automation seem to be positively correlated to drivers' perceived safety of Mediator (*study 3*).

Regarding the expected safety impact of Mediator, 86% of the participants in one of the driving simulator studies (*study 3*) expected an increase in safety, for instance, in terms of number and severity of accidents. Another driving simulator study (*study 1*) could reveal that participants who experienced Trivia as preventive mediation showed a higher probability of identifying important cues before an alert onset, of identifying hazard after the HMI alerted of an upcoming hazard, and had more glances at the hazard compared to the non-Trivia group.

Summary: Participants perceived safety while driving with Mediator was rated medium to good, with higher ratings in driving simulator studies compared to on-road studies. Additionally, drivers expect an improvement in driving safety when driving with Mediator. Different samples (naïve vs. professional drivers) or driver characteristics had no systematic and replicable effects on drivers' comfort ratings.

2.5. Usability & HMI evaluation

2.5.1. Usability and user friendliness

Introduction: Mediator is a system that is based on a cognitive design approach, whereby user and system are team partners aiming to use mutual dependencies in the sense of a joint activity effectively (Johnson et al., 2014). To ensure an effective interaction between the team partners, the HMI needs to be usable. According to DIN EN ISO 9241-11 (2018, p. 9), usability is defined as "the extent to which a product can be used by specific users in a specific context of use to achieve specific goals effectively, efficiently, and satisfactorily" (p. 9). An important prerequisite when developing interactive systems is to follow the human-centred design process that considers the users' needs at every developmental stage. Thereby, the aim is to create a product that is both understandable and useful, but also capable of performing the desired functions and providing a positive and pleasant user experience (Norman, 2016).

Results: In the on-road studies, perceived usability was rated as good (73 <= *M* <= 76 out of 100). During the interviews (*study 6*), 49% of the drivers reported that Mediator was easier to learn and to use than the baseline system. In the driving simulator studies, participants rated Mediator's usability as good to nearly excellent (79 >= *M* <= 85 out of 100).

In one study (*study 6*), age was identified as influencing factor indicating that older people gave slightly lower usability ratings. In another study (*study 1*), it could be shown that participants who experienced Trivia as preventive mediation rated Mediator's usability significantly higher than the non-Trivia group. Further, male participants evaluated the usability of the infotainment display higher than female participants did (*study 2*). Although, it could be revealed, that usability ratings of female participants increased with repeated usage indicating that higher familiarity with Mediator might have a positive influence on perceived usability. Neither age, nor gender effects could be found in the other studies. In one driving simulator study (*study 3*), it was found that higher affinity for technology might be related to a more positive evaluation of Mediator's usability.

Summary: Over all studies, Mediator's usability was percieved as good with a tendency to excellent. The ratings were a bit lower for on-road studies compared to driving simulator studies. Different samples (naïve vs. professional drivers) as well as driver characteristics had no systematic and replicable effects.

2.5.2. HMI Evaluation

Introduction: When developing HMIs particularly for such a crucial context as automated driving, several design principles need to be considered and evaluated in parallel (Lee et al., 2017; Wickens et al., 2013). An example of a model that can be used to derive concrete design guidelines for display and control concepts aiming to establish usability is Wickens' SEEV model (Wickens et al., 2013). The model can be used to predict the focus of users' selective attention on different areas of interest. These areas offer certain task-relevant information and are dependent on four perceptual factors consisting of salience of the elements, effort of attention shifting, expectation and value of the information. Beyond that, in their taxonomy of 13 design principles, Lee et al. (2017) define specific requirements regarding the design of display elements and the use of information of visual (e.g. colour schemes), acoustic (e.g. pitch level) and haptic (e.g. vibration strength) modalities and possibilities of combining them. According to Hoeger et al. (2011), the combination of visual, acoustic, and haptic information within an HMI in automated vehicles yields the best results in successful communication between the system and the user.

Results: In general, participants in all studies evaluated the HMI concept positively while driving, and during handover and takeover actions as well. One on-road study (*study 7*) could show that drivers tended to agree to the statement that Mediator is considered as beneficial for (partly) automated driving. One on-road study (*study 5*) revealed that all single HMI solutions (e.g., cluster interface, head unit interface, vocal message, acoustic and haptic feedback) were evaluated positively except LEDs on the steering wheel. Drivers stated that the LEDs on the steering wheel were hardly visible during daylight, and they expected the LEDs to be annoying during nighttime. In one driving simulator study (*study 3*), several participants mentioned that the lighting concept was a bit overloaded (e.g., the additional LEDs on the steering wheel and the dashboard; ambient light). In general, the light concept, especially the colour coding, was evaluated positively by most of the drivers during this driving simulator study (*study 3*). During one on-road study (*study 5*), drivers appreciated the change of the road colour because of its good visibility and comprehensibility. They rated the colour change as sufficient to inform the driver regarding the current driving mode. Participants of another on-road study (*study 6*) stated that the colour scheme (e.g., LED strip in the steering wheel, stylized road) contributed to the clarity of the system making it easier to identify and keep track of the current driving mode. Especially the colour scheme of the stylized route was mentioned as helpful as it also conveyed the time to the next shift in responsibility. Nevertheless, some participants perceived the colours as abstract and difficult to understand intuitively. In one driving simulator study (*study 3*), not all drivers understood the colour coding of the displayed route correctly. Less than half of the participants recognized the animation of the displayed route.

In one on-road study (*study 5*), participants appreciated acoustic and haptic feedback especially in more dangerous situations (e.g., distraction or fatigue) because it captured drivers' attention. In contrast, some of the participants experienced the haptic feedback as too strong and the acoustic feedback in general as less useful.

Additionally, drivers rated visibility and legibility of HMI messages as important for Mediator's usability (*study 5*). In one of the driving simulator studies (*study 3*), participants 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. Further, participants evaluated the messages in general positively and agreed that the messages were understandable. Nevertheless, some drivers stated that sometimes they did not understand the messages because, for instance, the reason for takeover or handover actions were not stated clearly enough. Hence, they expressed the wish for more transparency and more information. In one on-road study (*study 5*), icons were rated as very comprehensible, except the steering wheel icon with one dashed hand. This icon was not immediately comprehensible for all participants. Another on-road study (*study 6*) also revealed that the steering wheel icon was difficult to understand intuitively. In one driving simulator study (*study 3*), participants rated the indications on the display as clear and very easy to understand. It could be shown that the interpretation of the steering wheel icon was difficult for the drivers as well.

One driving simulator study (*study 2*) revealed a higher preference of male participants for the HMI compared to female participants. After repeated usage of the system, however, the ratings of female participants increased and were comparable to the evaluations of male participants.

Summary: Drivers evaluated the Mediator HMI concept positively and as beneficial for (partly) automated driving. Some elements of the HMI seem to have no added value (e.g., additional LEDs on the steering wheel when already having LEDs on the dashboard or vice versa). Overall, the light concept, especially the colour coding, was evaluated positively, for instance, to increase mode awareness. For potentially dangerous situations, acoustic and haptic feedback was experienced as helpful. The messages were rated as understandable, although some icons were not intuitively comprehensible. Hence, participants would like to have more information and better transparency. Different study conditions (driving simulator vs. on-road study), samples (naïve vs. professional drivers) as well as driver characteristics had no systematic and replicable effects on drivers' evaluation of the Mediator HMI concept.

2.5.3. Mode Awareness

Introduction: The displays in automated vehicles and the conveyed messages should be designed in such a way that users are at all times aware in which driving mode they currently are. Otherwise, serious consequences (e.g., accidents) may follow due to mode error (Bengler et al., 2017; Feldhütter et al., 2018). The implementation of mode awareness (Sarter & Woods, 1995) in the HMI concept should be realized using transparent messages regarding the systems' bases for decision-making, as it has a direct impact on acceptance and trust in the system (Yang et al., 2018).

Results: One on-road study (*study 5*) revealed that drivers would like to have more information why different driving modes are available. They stated that the sole information on automated driving availability is not enough to understand the situation. In another on-road study (*study 6*), participants expressed awareness of their responsibilities when using Mediator, although, drivers expressed higher uncertainty regarding their responsibilities in the SB condition and in the TtS condition, especially when it came to the difference between these two modes which were both mainly communicated as “piloted”. In the third on-road study (*study 7*), drivers tended to agree to the statement, that Mediator improved awareness of the automation system. During one of the driving simulator studies (*study 3*), drivers rated the colour coding in general as helpful to increase mode awareness. Most of the drivers stated that it was always clear to them, who is responsible for the driving task, for instance, because of the colour coding, Mediator’s messages and displayed icons.

Summary: Participants agreed that Mediator could help to increase mode awareness while using partly automated vehicles. Especially, the colour coding was evaluated as helpful to display the current as well as the upcoming driving mode. Used icons and messages were rated as helpful to increase awareness of the actual responsibilities of drivers. Nevertheless, drivers expressed the wish for more information why different automated driving modes are available or not longer available. Different study conditions (driving simulator vs. on-road study), samples (naïve vs. professional drivers) as well as driver characteristics had no systematic and replicable effects.

2.6. Timings and time budget information

Introduction: From SAE level 3 of automated driving, drivers are permitted to engage, for instance, in secondary tasks under certain conditions. In this context, the availability as well as the representation of time budgets for performing secondary activities plays a considerable role in the perceived usefulness of driver assistance systems (Danner et al., 2020). A study by Danner et al. (2020) revealed that when availability duration was displayed, acceptance and usability ratings towards the system increased, while perceived workload decreased. One explanation for this effect is that completing different tasks takes different amounts of time (e.g., reading a text vs. watching a video), and presenting the available time budgets helps in planning these secondary tasks before a take-over is required (Hecht et al., 2020). Thus, assessing the subjective evaluation of different timings is crucial.

Results: In one of the on-road studies (*study 7*), drivers referred to the time budget as a desirable and helpful functionality that can be a benefit for automated driving especially in higher automation modes and for longer drives. Additionally, it could be shown that the variability of drivers’ take-over times was reduced when driving with Mediator indicating that Mediator harmonizes the take-over time among drivers. Another on-road study (*study 6*) revealed that most participants rated the remaining time in the current mode as helpful and informative, and that they lacked the information during the baseline condition. Nevertheless, the time budget was partly not understood and some drivers suggested communicating the time budget information in a better way, for example with descriptive text. In one driving simulator study (*study 3*), some drivers stated that the time budget was confusing and not perfectly understandable for them.

During one driving simulator study (*study 3*), the timing for comfort-related transfers from manual to automated driving as well as for (un-)planned transfers from automated to manual driving were evaluated as very good (M = 4 on a 5-point Likert scale ranging from 1 to 5). Participants agreed that they were warned with sufficient time to retake control safely. One on-road study (*study 6*) could also show that timings for transfers of control from automated to manual driving were experienced as good, and timelines for warnings and when it was time to change automation mode were easy to understand. Some participants indicated that it was unclear to them how urgent a takeover request was, i.e., immediate need for a mode change right after receiving the message. Further, they suggested accompanying the countdown by a sound especially when the timer approached zero. In another on-road study (*study 7*), the timings of warnings when a take-over was necessary was on average rated as appropriate and sufficient. Interviews revealed that some drivers prefer earlier warnings, and that other drivers are more in favour of later warnings. In addition, drivers reported that their preference might change with more experience (e.g., repeated exposure to Mediator) or might vary according to the situation (e.g., fast approach to a highway exit vs. slower approach to an intersection in a city). After a drive with a higher frequency of take-over suggestions (*study 7*), participants' answers showed the highest variance indicating that some drivers preferred Mediator's recommendation to use assisted driving also for shorter sections meaning that the warnings to retake control appeared more often (and sometimes shortly after switching to assisted driving). Other drivers rated this behaviour of Mediator as less appropriate. Results indicate that, especially for retaking control over the driving task, the timing should be adaptable to drivers' preferences (that might change over time) and situations' requirements. During the third on-road study (*study 5*), participants stated that it is important for them to have enough time in advance to resume the control of the vehicle. Further aspects that were rated as crucial during one on-road study (*study 5*) were sufficient time to resume control of the vehicle, low latency times and quicker escalation times.

With respect to corrective actions for distraction, one on-road study (*study 6*) could show that participants generally had neutral opinions about understanding the distraction warnings and appreciating their timing. Some participants mentioned that the distraction warnings were badly timed. In another on-road study (*study 7*), participants reported an oversensitive distraction warning and suggested an improvement in sensitivity and timing.

Summary: The time budget information was evaluated as helpful information that provides added value for partly automated driving. Nevertheless, the representation of the time budget should be improved to avoid confusion (e.g., make it clear if the displayed time budget shows current time to automation fitness or expected time to automation unfitness on the next part of the route, i.e., time span of automation availability). Overall studies, the timing of Mediator's suggestions and requests for mode changes were experienced as good to very good. Drivers' expressed the wish for individualization of timings to meet personal preferences. Different study conditions (driving simulator vs. on-road study), samples (naïve vs. professional drivers) as well as driver characteristics had no systematic and replicable effects.

2.7. Take-over and hand-over actions

Introduction: For several levels of automation, the driver or passenger respectively is allowed to be out of the loop for short periods of time (e.g., SAE level 3). Yet, when the driver is asked to take over control again, he or she must be brought back into the loop, regain situation awareness, and take control of the vehicle. In case of distraction by secondary task or fatigue, the time to regain driver fitness might be too long potentially resulting in unfitness of both the human driver and automation. In order to avoid severe consequences, take-over and hand-over actions need to be evaluated.

Results: One driving simulator study (*study 3*) revealed that participants understood why Mediator offered to take over the driving task for them during comfort-related transfers from manual to automated driving very well (M = 4 on a 5-point Likert scale ranging from 1 to 5). They understood the instructions to hand over the driving task (M = 4.6 on a 5-point Likert scale ranging from 1 to 5). One on-road study (*study 6*) showed that participants found it easy to follow instructions for transfer when Mediator actively proposed hand-overs to higher levels of automation. Although, they could predict why the hand-over was proposed, they reported the preference for more information on why the hand-over took place. They also strongly agreed that they expected to make more use of automation if proactive proposals for activating automation were offered when either distracted or when a new automation level was available.

In one driving simulator study (*study 3*), participants indicated that they understood quite well why the take-over request occurred for (un)planned transfers from automated to manual driving (*M* = 3.7 on a 5-point Likert scale ranging from 1 to 5). They understood the instructions to take over the driving task very well (*M* = 4.4 on a 5-point Likert scale ranging from 1 to 5). Further, they agreed that they were warned in an appropriate way (*M* = 4.1 on a 5-point Likert scale ranging from 1 to 5).

One on-road study (*study 6*) could show that drivers rated the transfer of control from automated to manual driving on a 5-point Likert scale ranging from 1 to 5 as comfortable (*Md* = 4), appropriate (*Md* = 4), and experienced the instructions to hand over control as easy to follow (*Md* = 4). Although, the reason for occurrence of the request to take back control was rated as obvious (*Md* = 4), participants showed a tendency for a need of more information why the request of taking back control was triggered (*Md* = 2.5). Participants appreciated the way mode changes were communicated by the Mediator HMI more compared to the baseline HMI. Nevertheless, they would prefer to gather the information in one place instead of having it distributed across several screens; perceived the transfers of control to higher automation modes more as warnings rather than as invitations/requests; and perceived the text messages as too short (*study 6*).

Another on-road study (*study 5*) could show that drivers evaluated the HMI in general as positive and easy to understand during transfers of control. Some participants reported the need for more knowledge if and why the automated driving mode is available to better understand the situation. Further, they would prefer if the HMI suggested directly what to do (e.g., move shifter).

One on-road study (*study 7*) could show a general downward trend in number of transfers with continuous usage of Mediator. It seems that drivers use the automated system less with time. Two drivers had lower numbers of transfers to automated driving but drove longer durations in assisted driving mode in most of the drives compared to the other drivers. Interviews revealed that they tried to push Mediator's limits (driving longer in assisted driving mode than recommended).

Summary: Over all studies, drivers evaluated the transitions from one driving mode to another positively. The instructions were easy to understand and to follow. Most often, the reasons for take-over or hand-over actions were transparent. Nevertheless, some drivers would like to have more information on why the mode change is possible or necessary. The information should be presented in one place and should not be distributed over several screens or communication channels. Some drivers' reported that the difference between a request for a necessary mode change and an optional proposal to change driving mode was not clear enough. Different study conditions (driving simulator vs. on-road study), samples (naïve vs. professional drivers) as well as driver characteristics had no systematic and replicable effects.

2.8. Workload and distraction caused by Mediator

Introduction: Mental workload being defined as the amount of information-processing resources needed to perform a certain task (De Waard, 1996) is a crucial variable to be evaluated in regard of the development of Mediator. Hence, drivers of automated systems whose workload is too low (underload) or too high (overload) may have a decreased driving performance because of boredom or the inability of allocating sufficient resources to the intended driving task (De Waard et al., 1996). Moreover, previous studies revealed that mental workload increases with increased complexity of the driving context, e.g., distraction by secondary task (Törnros & Bolling, 2006; Cantin et al., 2009; Ma & Kaaber, 2005; Merat et al., 2012). Other studies found that at least in CM, mental workload seems to increase due to the drivers' task to monitor both the environment and the automated system, even without having an active part in driving (Stapel et al., 2019). In critical situations such as a take-over request, this is particularly problematic (Merat et al., 2012). Thus, balancing the amount of information provided to the driver via HMI with the total information load a driver is able to process is important and needs to be considered for Mediator.

Results: During one on-road study (*study 5*), participants mentioned information regarding the need for a takeover manoeuvre well in advance, information about the current driving mode and its duration but also, on the other hand, no information overload as crucial. In one driving simulator study (*study 3*) most of the participants stated that the amount of information was sufficient and not overloading.

Another on-road study (*study 6*) could show that driving with Mediator active was on average judged to cause low workload (*Md* = 20 on a scale from 0 to 100). During the interviews, only few drivers reflected on the workload. Some of them reported that there was too much information to take in and suggested to simplify the communication between the vehicle and the driver. Some drivers experienced the distraction warning as stressful.

Another on-road study (*study 7*) revealed a higher likelihood that drivers are not looking to the forward view when driving with Mediator available. One possible explanation might be that drivers inspect the Mediator HMI more frequently; hence, their eyes are away from the forward view. Nevertheless, the results do not impose higher risks compared to driving without Mediator, as the values are still close to each other. Another on-road study (*study 6*) could show that drivers were distracted for a longer duration when driving with the baseline HMI compared to the Mediator HMI (1,0 s <= *Md* <= 1,8 s). Although, the overall proportion of time drivers were detected to be distracted was quite low (0,7% <= *Md* <= 1,1%).

During one driving simulator study (*study 1*), it could be revealed that engagement with Trivia decreased the mental workload compared to the control condition. One on-road study (*study 6*) provided evidence that older drivers seem to experience the workload higher compared to younger drivers.

Summary: It could be revealed that drivers rated the amount of provided information as mostly sufficient and not overloading. The workload caused by Mediator was rated as low. Although the Mediator HMI with its different components providing information to the driver decreased the likelihood of forward views, the risk for drivers to be distracted by Mediator is interpreted as low. Different study conditions (driving simulator vs. on-road study), samples (naïve vs. professional drivers) as well as driver characteristics had no systematic and replicable effects.

2.9. Degraded driver states: Fatigue and distraction

Introduction: As the degree of automation increases, the role of humans as physically active decision-makers in the vehicle is replaced by automated systems. Behaviours that were previously essential (e.g., performing steering manoeuvres) are no longer needed. Simultaneously, users of automated vehicles have to acquire new skills and a new mental model of the system (Knecht et al., 2014; Maurer et al., 2016). This includes planning, monitoring, and decision-making regarding possible interventions in case of unforeseen events. However, tasks that are characterized by prolonged monitoring and require a high level of attention over a long period of time do not belong to the strengths of humans. Therefore, tasks linked to automation can quickly result in fatigue (e.g., task- or sleep-related fatigue), which in turn can massively impair performance (Knecht et al., 2014). A countermeasure for passive fatigue is monotony rupture. In case fatigue is caused by underload (e.g., long or tedious drives), activation, for instance, by performing NDRTs, is found to be an effective countermeasure to maintain awareness and vigilance (e.g., Oron-Gilad et al., 2008; Gershon et al., 2009; Lee et al., 2020). In contrast, engaging in NDRTs while the vehicle is in control of the automation may impair drivers' hazard perception performance (Zangi et al., 2022). So far, little attention has been paid to the trade-off between the benefits of NDRTs as a countermeasure of fatigue and the potential distraction it can cause at the same time and was therefore investigated within the project.

Results on fatigue: One driving simulator study (study 3) revealed that drivers would appreciate to have fatigue detection and respective warnings as part of Mediator. Another driving simulator study (study 1) compared participants who had the opportunity to play Trivia as one possible preventive action to avoid fatigue with participants who did not play the Trivia game. Results could show that Trivia had no influence on the development of drivers' fatigue. For all drivers, a comparable increase in fatigue could be revealed with continued system usage. Nevertheless, the fatigue level reached only a moderate level. The increased fatigue level was not evident in HRV measures indicating that drivers applied psychological adaptation but not physiological adaptation. Similar results could be revealed during one of the on-road studies (*study 6*). Drivers reported similar fatigue levels when driving with Mediator compared to a baseline. With increased driving time, a higher fatigue level was revealed for all drivers. Further, it could be shown that corrective alert-message used in the transfer of control ritual could reduce task-related fatigue a bit but not sleep-related fatigue. During the interviews, participants reported that they appreciated the fatigue warnings, and rated them as understandable, well timed and helpful. Especially, the haptic feedback was evaluated as effective. Some drivers expressed the wish for more interaction with Mediator, e.g., confirm that the warning was perceived.

Results on distraction: In one driving simulator study (*study 3*), participants indicated that they would appreciate distraction detection and respective warnings as part of Mediator. One on-road study (*study 7*) could also reveal that distraction detection and warning is a valuable functionality of Mediator but needs to work reliably. Another on-road study (*study 6*) revealed similar results. Participants generally appreciated the concept of distraction warning and Mediator's take-over suggestions in cases of detected distraction. Generally, drivers had neutral opinions regarding understandability of warnings and the respective timings. Some participants mentioned the need for more transparency why a warning was triggered and what action is suggested. Further, some drivers reported that the warnings were badly timed or were given in situations when the driver was not distracted but concentrated on oncoming traffic or checking the mirrors. In one on-road study (*study 7*), it could be shown that drivers were comparably often distracted when driving with distraction detection active vs. inactive. One driving simulator study (*study 3*) could show that, at least under controlled experimental conditions and stable light conditions in a driving simulator, camera-based detection of eyes-off road time by a Bayes classifier trained on features of head tracking is a fruitful approach.

Summary: Studies could reveal that drivers appreciate both fatigue detection and distraction detection with respective warnings. Especially, haptic feedback was evaluated as helpful. Important prerequisites are detection algorithms that will work reliably and transparently on why a warning was triggered, and which corrective actions are suggested. Trivia as one possible measure to prevent fatigue could not be shown to have an influence. Corrective alert-messages were shown to decrease task-induced fatigue slightly. Different study conditions (driving simulator vs. on-road study), samples (naïve vs. professional drivers) as well as driver characteristics did not show systematic and replicable effects.

2.10. Decision logic

Introduction: One core element of the Mediator system is the decision logic (DL) component. DL gathers all input data from the driver state component, the automation state component as well as input variables from the environment regarding the driving context. The different information will be analysed in real-time and based on the input variables, DL will decide who is fittest to drive (driver vs. automated system), and will calculate time to automation fitness (TTAF) / unfitness (TTAU) as well as time to driver fitness (TTDF) / unfitness (TTDU). Based on the results, Mediator will suggest or request driving mode changes.

Results: The quality of DL's output depends on the quality of input signals. Noise in the input signals can cause noise in the output, potentially leading to critical or unsafe situations. Examples for noise in the driver context are wrongly assessed driver states (e.g., distraction or fatigue), or the calculation of the resulting TTDF and TTDU. The values of TTDF and TTDU based on impaired driver states are a lot more subjective than the calculations of TTAF/ and TTAU, because predicting how long it will take before a driver really becomes unfit to drive is more difficult. Further, results show that the DL is currently not robust enough to deal with automation noise. One way to deal with this would be to add buffers within the decision trees. Implementing buffers will lead to a more conservative approach. With this approach, safety will be increased but driver comfort might be decreased because the percentage of time driving in manual driving is higher than in the default case. Driving in manual mode although (partly) automated driving is available is considered less comfortable for the driver. Using a more optimistic approach can lead to a big increase of events where the DL assumes the car can be driven in an automated mode that is realistically not available leading to unsafe situations. For a system like Mediator, safety should have the highest priority. Real world testing is necessary to estimate error margins and derive appropriate buffers. Together with the buffers, an optimal combination of the offsets for TTAF, TTAU, TTDF, and TTDU is important for safety. Another, more simple but possibly also more robust way, would be to counteract any negative noise by artificially adding time to TTAF and subtracting time from TTAU, respectively.

The default DL and DL with built-in buffers were tested on real-world data that was implemented into the computer simulation. In the default case, time driven in manual driving and CM is equally divided. With buffers, the time in manual driving is a little higher than the time in CM, which can be explained by the fact that the buffers make the DL more conservative. The average time between actions is almost 5 minutes, which seems like a reasonable time. Dynamic events causing quick takeovers cannot be predicted and, hence, recent switches cannot be prevented. Both methods are safe in over 99.9% of the examined cases. The few unsafe car-related situations had a duration of 0 seconds, meaning they are resolved instantly as they occur. The events related to driver unfitness (e.g., due to fatigue or distraction) last longer, but can be prevented by implementing an emergency stop function in the Mediator system concept. Overall, results revealed that the DL enables drivers to drive in a (semi-)autonomous driving mode for 50% of the time, while assuring safety.

In the scope of the computer simulation study, it was also tested if reinforcement learning (RL) instead of the currently used decision trees can be used as a basis for a decision-making system. RL will be more generalizable compared to logical reasoning. Therefore, a RL agent was trained on the distraction use case, and tested in an environment where the driver can get both fatigued and distracted. Results revealed that out of 100 runs, RL has 15 unsafe runs (15%), while the decision tree has 52 unsafe runs (52%), which can all be attributed to the tree not being able to handle driver fatigue. Hence, RL can handle slightly changing environments better because it learned to act based on TTDU. Nevertheless, RL is still unsafe in 15% of runs with a simplified scenario where only driver degradation is considered.

Summary: To avoid unsafe or critical driving situations, a more conservative approach for the DL should be prioritized. This might lead to reduced driver comfort, because the DL may not suggest a shift to a level up while realistically the higher level might be available for a long enough period. Ideally, the buffer is as small as possible to reduce this drawback. This requires the estimates to have a low error margin. More data and testing are needed to determine the current error margin, from which an appropriate buffer value can be deduced.

3. Overall summary, conclusions and implications

Mediator as an intelligent driver assistant system for SAE levels 2-4 driving automation mediates in real-time between the human driver and the automated system considering also environmental factors. Based on the decision who is most capable to perform the driving task (i.e., fittest to drive), Mediator system proposes or request a change of driving mode, and supports the driver to perform a safe handover / takeover action. In Work Package 3 (WP3) “Testing & Evaluation”, seven studies (three driving simulator studies, one computer simulation study without real users, three on-road studies) were conducted to evaluate the functionality of the Mediator system. Over all studies, Mediator was evaluated positively, and as beneficial for (partly) automated driving. Results revealed that over all studies and user groups:

- Acceptance of Mediator was quite high.
- Intention to buy, use and recommend Mediator was quite high.
- Trust in Mediator was quite high.
- Experienced comfort while driving with Mediator was quite high.
- Perceived safety was medium to good.
- Expectations regarding an improvement in driving safety when driving with Mediator were high.
- Workload caused by Mediator was rated as low.
- Comprehensibility and amount of provided messages were rated as quite good.
- Comprehensibility of instructions for transitions from one driving mode to another were rated positively, although some more information were suggested especially regarding the reasons for availability of different driving modes.
- Timing of Mediator’s suggestions and requests were experienced as good to very good.
- Time budget information was rated as reliable and helpful providing added value for partly automated driving, although better explanation and visualization is suggested.
- Usability was perceived as good with a tendency to excellent.
- All HMI elements were evaluated positively, especially the light concept to increase mode awareness and acoustic and haptic feedback for potentially dangerous situations.
- The functionalities of fatigue as well as distraction detection with respective warnings were rated as desirable and helpful.
- Prerequisites for drivers accepting, trusting and using Mediator are, for instance, the fulfilment of drivers’ expectations regarding the system’s functionalities and limitations, the possibility for personalization, the final price and high transparency as well as reliability of the system.
- A more conservative approach for the decision logic can avoid unsafe or critical driving situations.

Regarding results’ external validity, it needs to be taken into account that drivers in all studies participated voluntarily in the MEDIATOR studies. Hence, self-selection effects potentially leading to samples with, for instance, higher values for ATI, higher acceptance and a more positive general opinion regarding vehicle automation, and higher initial trust in Mediator might cause a more positive overall evaluation of the system. Further, for the driving simulation studies, results need to be interpreted carefully, because the study results reflect an optimal working Mediator and, thus, represent its highest possible potential. Additionally, the simulated environment might have biased the participants’ feeling of safety and risk perception.

More details regarding examined research questions, used methods, revealed results, limitations and implications can be found in the public deliverables D3.2 (Athmer et al., 2022), D3.3 (Borowsky et al., 2023) and D3.4 (Fiorentino et al., 2023).

The results have important implications for automated driving and are discussed in this regard in other public MEDIATOR deliverables, for instance, with respect to driving safety (D4.1), HMI design (D4.2) or assessment of degraded driver performance (D4.3).

4. References

Adell, E., Várhelyi, A., & Nilsson, L. (2018). The definition of acceptance and acceptability. In Driver acceptance of new technology, 11-22. CRC Press.

Athmer, C., Spaan, M., Li, Y., Liu, Y., Bakker, B. (2022). *Realistic Simulation of Decision Making in MEDIATOR System through use of Real-World Data*, Deliverable D3.2 of the H2020 project MEDIATOR.

Bengler, K., Winner, H., & Wachenfeld, W. (2017). No Human–No Cry?. at-Automatisierungstechnik, 65(7), 471-476.

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.

Cantin, V., Lavallière, M., Simoneau, M., and Teasdale, N. (2009). Mental workload when driving in a simulator: Effects of age and driving complexity. *Accident Analysis, & Prevention*, 41(4), 763-771.

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.* D1.1 of the H2020 project MEDIATOR.

Danner, S., Pfromm, M., & Bengler, K. (2020). Does information on automated driving functions and the way of presenting it before activation influence users' behavior and perception of the system? *Information*, 11(1), 54. https://doi.org/10.3390/info11010054

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quart*., 13(3), 319–340.

DIN EN ISO 9241-11. (2018). Ergonomie der Mensch-System-Interaktion - Teil 11: Gebrauchstauglichkeit: Begriffe und Konzepte (ISO 9241-11:2018). Beuth.

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

Feldhütter, A., Segler, C., & Bengler, K. (2018). Does Shifting Between Conditionally and Partially Automated Driving Lead to a Loss of Mode Awareness? In N. A. Stanton (Hrsg.), *Advances in Human Aspects of Transportation* (Bd. 597, S. 730–741). Springer International Publishing. https://doi.org/10.1007/978-3-319-60441-1_70

Fiorentino A., Bakker B., Brase D., Karlsson J., Knauss A., Santhià T., Thalya P., Toffetti A., van Grondelle E., Wang, D., (2022). *Final in-vehicle prototypes*, Deliverable D2.10 of the H2020 project MEDIATOR.

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

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology,* 52, 139-183. North-Holland.

Hecht, T., Feldhütter, A., Draeger, K., & Bengler, K. (2020). What do you do? An analysis of non-driving related activities during a 60 minutes conditionally automated highway drive. *International Conference on Human Interaction and Emerging Technologies*, 28-34. Springer, Cham. https://doi.org/10.1007/978-3-030-25629-6_5

Heiderich, M., Friedrich, T., & Nguyen, M. T. (2016). New approach for improvement of vehicle performance by using a simulation-based optimization and evaluation method. In *7th International Munich Chassis Symposium*, ATZ, Munich.

Heißing, B., Ersoy, M., & Gies, S. (2013). Zukunftsaspekte des Fahrwerks. Fahrwerkhandbuch: Grundlagen· Fahrdynamik· Komponenten· Systeme· Mechatronik· Perspektiven, 652-693.

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

Hoeger, R., Zeng, H., Hoess, A., Kranz, T., Boverie, S., Strauss, M., ... & Stratil, H. (2011). Final report, deliverable D61.1. Highly automated vehicles for intelligent transport (HAVEit). https://trimis.ec.europa.eu/sites/default/files/project/documents/20130628_174319_29918_HAVEit_FinalReport.pdf

Hoff, K. A., & Bashir, M. (2015). Trust In Automation: Integrating Empirical Evidence On Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 57(3), 407–434. https://doi.org/10.1177/0018720814547570

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. https://doi.org/10.1207/s15327566ijce0401_04

Johnson, M., Bradshaw, J. M., Feltovich, P. J., Jonker, C. M., van Riemsdijk, M. B., & Sierhuis, M. (2014). Coactive design: Designing support for interdependence in joint activity. *Journal of Human-Robot Interaction*, 3(1), 43. https://doi.org/10.5898/JHRI.3.1.Johnson

Kassner, A., & Vollrath, M. (2006). Akzeptanzmessung als Baustein für die Entwicklung von Fahrerassistenzsystemen. In *Integrierte Sicherheit und Fahrerassistenzsysteme*. 97–112. Düsseldorf: VDI-Verlag GmbH.

Knecht, C., Mühlethaler, C., Zimmermann, J., Binz, S., Painadath, R., & Wäfler, T. (2014). Vertrauen in die Automatisierung, fehlende Situation Awareness & Fertigkeitsverlust durch automatisierte Systeme–Eine subjektive Einschätzung aus der Linienpilotenperspektive. *56. Fachausschusssitzung Anthropotechnik der DGLR.*

Lee, J. D., Wickens, C. D., Liu, Y., & Boyle, L. N. (2017). Designing for people: An introduction to human factors engineering. *CreateSpace.*

Lee, J., & See, K. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

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.

Loh, S., Lamond, N., Dorrian, J., Roach, G., & Dawson, D. (2004). The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior research methods, instrument & computers*, 36(2), 339-346.

Ma, R., & Kaber, D.B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a mobile phone. *International Journal of Industrial Ergonomics*, 35, 939–953.

Manzey, D. (2012). Systemgestaltung und Automatisierung. *Human Factors: Psychologie sicheren Handelns in Risikobranchen*, 333-352.

Maurer, M., Gerdes, J. C., Lenz, B., & Winner, H. (Hrsg.). (2016). *Autonomous Driving*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-48847-8

Merat, N., Jamson, A.H., Lai, F.C., & Carsten, O.M. (2012). Highly Automated Driving, Secondary Task Performance, and Driver State. *Human Factors*, 54(5), 762-771.

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/

Norman, D. (2016). The design of everyday things (2., überarbeitete Auflage). Vahlen. https://doi.org/10.15358/9783800648108-114

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

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Parasuraman, R., Mouloua, M. & Molloy, R. (1996). Effects of adaptive task allocation on monitoring of automated systems. *Human factors*, 38(4), 665–679.

Park, E., Baek, S., Ohm, J., & Chang, H. J. (2014). Determinants of player acceptance of mobile social network games: An application of extended technology acceptance model. *Telematics and Informatics*, 31(1), 3-15.

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. http://www.suaave.eu/wp-content/uploads/sites/17/2021/02/SUaaVE_WP1_D1.2_20200930_V100.pdf

Riley, V. (2018). Operator reliance on automation: Theory and data. In *Automation and human performance: Theory and applications*, 19–35, CRC Press.

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: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 5–19. https://doi.org/10.1518/001872095779049516

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

Stapel, J., Mullakkal-Babu, F. A., & Happee, R. (2019). Automated driving reduces perceived workload, but monitoring causes higher cognitive load than manual driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*, 590-605.

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.

Törnros, J., & Bolling, A. (2006). Mobile phone use–effects of conversation on mental workload and driving speed in rural and urban environments. T*ransportation Research Part F: Traffic Psychology and Behaviour*, 9(4), 298-306.

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, 1-10. http://www.hfes-europe.org/accept/accept.htm.

Waard, D. de (1996). The measurement of drivers' mental workload. Netherlands: Groningen University, Traffic Research Centre.

Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., & Talleur, D. A. (2003). Attentional models of multitask pilot performance using advanced display technology. *Human factors*, 45(3), 360-380.

Yang, Y., Karakaya, B., Dominioni, G. C., Kawabe, K., & Bengler, K. (2018). An hmi concept to improve driver's visual behavior and situation awareness in automated vehicle. *21st International Conference on Intelligent Transportation Systems (ITSC)*, 650-655. IEEE.

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

Zhang, T., Tao, D., Qu, X., Zhang, X., Lin, R. & Zhang, W. (2019). The roles of initial trust and perceived risk in public’s acceptance of automated vehicles. *Transportation Research Part C: Emerging Technologies*, 98, 207–220.

5. Appendix A

Note that the numbering of the studies is based on the following scheme:

- Study 1: Driving Simulator Study on driver fatigue – BGU (Israel)
- Study 2: Driving Simulator Study on driver distraction – BGU (Israel)
- Study 3: Driving Simulator Study on driver comfort – TUC (Germany)
- Study 4: Computer Simulation Study on decision logic– TUD (Netherlands)
- Study 5: On-road study on HMI – FCA (Italy)
- Study 6: On-road study on driver impairments and degraded automation – VTI (Sweden)
- Study 7: On-road study on automation state – VEO (Sweden)

Table 3. Overview of topics covered in the Simulator and On-road studies with applied methods including indication of chapter of result discussion in this deliverable.

Topic	Study	Method	Questionnaires from literature / previous projects	Self-designed items	Results discussed in chapter
Acceptance & intention to use	1	Questionnaire	• Acceptance scale (van der Laan et al., 1997) • Choi and Ji (2015) • SUaaVE pre-trial questionnaire (Post et al., 2020)	/	2.1
	2	Questionnaire	• Acceptance scale (van der Laan et al., 1997) • SUaaVE pre-trial questionnaire (Post et al., 2020)	/	
	3	• Questionnaire • Interview	• Acceptance scale (van der Laan et al., 1997) • Questionnaire adapted from L3 Pilot (Metz et al., 2019) • Questionnaire adapted from ADAPTIVE (Rodarius et al., 2015)	/	
	5	Questionnaire	• SUaaVE Acceptance of automated vehicles (Post et al., 2020) • Acceptance scale (van der Laan et al., 1997)	/	
	6	Questionnaire	• SUaaVE Acceptance of automated vehicles (Post et al., 2020) • Acceptance scale (van der Laan et al., 1997)	/	

	7	• Questionnaire • Interview	/	Acceptance-related item (adapted from Jian et al., 2020)	
Trust & perceived reliability	1	Questionnaire	• Choi and Ji (2015) • SUaaVE pre-trial questionnaire (Post et al., 2020)	/	2.2
	3	• Questionnaire • Interview	• Trust in automation (Jian et al., 2000) • Questionnaire adapted from L3 Pilot (Metz et al., 2019)	/	
	5	Questionnaire	Trust in Mediator system (Jian et al., 2000)	/	
	6	• Questionnaire • Interview	Trust in Mediator system (Jian et al., 2000)	/	
	7	• Questionnaire • Interview • Eye tracking	/	Trust-related item (adapted from Jian et al., 2020)	
Perceived comfort	2	Questionnaire	SUaaVE pre-trial questionnaire (Post et al., 2020)	/	2.3
	3	• Questionnaire • Interview • Behavioural reactions	• Comfort questionnaire on automated driving style and takeover questionnaire (L3 Pilot; Metz et al., 2019) • Questionnaire adapted from ADAPTIVE (Rodarius et al., 2015)	Comfort-related items	
	6	Questionnaire	• Comfort questionnaire on automated driving style and takeover questionnaire (L3 Pilot; Metz et al., 2019) • Questionnaire adapted from ADAPTIVE (Rodarius et al., 2015)		
	7	• Questionnaire • Interview	/	Comfort-related item (adapted from Jian et al., 2020)	
Perceived safety & expected	1	• Questionnaire • Eye tracking	• Choi and Ji (2015) • SUaaVE pre-trial questionnaire (Post et al., 2020)	/	2.4

impact on driving safety	2	Questionnaire	SUaaVE pre-trial questionnaire (Post et al., 2020)	/	
	3	• Questionnaire • Interview	• Questionnaire adapted from L3 Pilot (Metz et al., 2019) • Questionnaire adapted from ADAPTIVE (Rodarius et al., 2015)	Safety-related items	
	6	Questionnaire	Questionnaire adapted from L3 Pilot (Metz et al., 2019)		
	7	• Questionnaire • Interview	/	safety-related item (adapted from Jian et al., 2020)	
Usability & HMI evaluation	1	Questionnaire	System Usability Scale (SUS; Brooke, 1996)	/	2.6
	2	Questionnaire	System Usability Scale (SUS; Brooke, 1996)	/	
	3	• Questionnaire • Interview	System Usability Scale (SUS; Brooke, 1996)	Usability-related items	
	5	• Questionnaire • Interview • Think Aloud	System Usability Scale (SUS; Brooke, 1996)	HMI-related items	
	6	• Questionnaire • Interview • Think Aloud	System Usability Scale (SUS; Brooke, 1996)	HMI-related items	
	7	• Questionnaire • Interview • Eye tracking		HMI-related item (adapted from Jian et al., 2020)	
Timings and time budget information	3	• Questionnaire • Interview	/	Usability-related items	2.6
	5	• Interview • Think Aloud	/	/	
	6	• Questionnaire • Interview	• Take-over questionnaire (adapted from L3 Pilot (Metz et al., 2019) • SUaaVE Acceptance of Automated Vehicles (Post et al., 2020) • Takeover questionnaire, adapted from L3Pilot questionnaire.		
	7	• Questionnaire • Interview • Time budget accuracy • Eye tracking	/	Timing-related single items	

Take-over and hand-over actions	3	• Questionnaire • Interview	Questionnaire adapted from L3 Pilot (Metz et al., 2019)	/	2.7
	4	Sensitivity analysis (automation and driver noise)	/	/	
	5	• Questionnaire • Vehicle data analysis Automation state analysis	Take-over questionnaire (adapted from L3 Pilot (Metz et al., 2019)	/	
	6	• Questionnaire • Interview • Vehicle data analysis • Automation state analysis	Take-over questionnaire (adapted from L3 Pilot (Metz et al., 2019)	/	
	7	• Interview • Vehicle data analysis • Automation state analysis • Eye tracking	/	/	
Workload & distraction caused by Mediator	1	Questionnaire	NASA-TLX (Hart & Staveland, 1988)	/	2.8
	2	Questionnaire	NASA-TLX (Hart & Staveland, 1988)	/	
	3	Eyes-off road time	/	/	
	6	Questionnaire	NASA-TLX (Hart & Staveland, 1988)	/	
	7	• Eye tracking • Eyes off-road time • Distraction Detection Algorithm (AttenD)	/	/	
Degraded driver states	1	• Questionnaire • HRV (Heart Rate Variability)	KSS (Shahid et al., 2011)	/	2.9
	5	Eye tracking	/	/	
	6	• Questionnaire • Interview • Eye tracking • ECG	• KSS (Shahid et al., 2011) • Vigilance test (Loh et al., 2004)	/	
	7	• Eye tracking • Eyes-off road time • Distraction Detection Algorithm (AttenD)	/	/	
Decision logic (DL)	4	• DL comparisons • Reinforcement Learning	/	/	2.10