D6.3 – Results of Comparative Evaluation after 3rd Cycle	
Project Number:	690705
Classification **Deliverable No.:**	Public D6.3
Work Package(s): **Document Version:** **Issue Date:** **Document Timescale:** Start of the Document: Final version due:	WP6 Vs. 1.0 Month 35 Project Start Date: September 1, 2016 Month 31 Month 35
Deliverable Overview:	**Main document:** D6.3
Compiled by:	Jürgen Pichen, ULM
Authors:	Jürgen Pichen, ULM Fei Yan, ULM Martin Baumann, ULM Andrea Castellano, REL Elisa Landini, REL Fabio Tango, CRF Massimo Fossanetti, CRF Gurucharan Srinivas, DLR David Käthner, DLR Mark Eilers, HMT Leonardo Espindola, HMT Stefan Suck, OFF Elsa Yousfi, VED Lynda Halit, VED Alain Giralt, CAF Ádám Knapp, BIT Maximilian Graf, ULM
Technical Approval:	Fabio Tango, CRF
Issue Authorisation:	Andreas Lüdtke, OFF

DISTRIBUTION LIST		
Copy type[1]	Company and Location	Recipient
T	AutoMate Consortium	all AutoMate Partners

RECORD OF REVISION		
Date	Status Description	Author
13.03.2019	Document Structure	ULM
19.09.2019	Merge of all contributions	ULM+ALL
20.09.2019	First draft for review	ALL
30.09.2019	Final version	ALL

[1] Copy types: E=Email, C=Controlled copy (paper), D=electronic copy on Disk or other medium, T=Team site (Sharepoint)

Table of Contents

List of Figures

List of Tables

List of Acronyms

A2H Automation to human
H2A Human to automation
DMS Driver monitoring system
DIR Driver intention recognition
HMI Human-Machine Interface
AR Augmented reality
GUI Graphical user interface
KSS Karolinska sleepiness scale
IVS Intelligent Vehicle Symposium
V2X Vehicle to X
ETSI European Telecommunications Standards Institute
OBD on-board diagnostics
CAN controller area network
CAM cooperative awareness message
DENM Decentralized Environmental Message
DLL Dynamically Linked Library
OWL Web Ontology Language
SWRL Sematic Web Rule Language
API Application Programming Interface
PC personal computer
MLN *Markov Logical Network*
CYRA constant yaw-rate and acceleration
IDM Intelligent Driver Model
DBN Dynamic Bayesian Networks
ADAS Advanced Driver Assistance Systems
ACC Adaptive cruise control
RSS responsibility-sensitive safety
PC Probability Of Collision
H-Metaphor Horse metaphor
RPM revolution per minute
V&V Verification & Validation
HUD Head-Up display
AR-HMI Augmented Reality human-machine interface
TM TeamMate
AM Automated mode
SCR sharing control request
FCW Forward Collision Warning
LDW Lane Departure Warning
LAS Lateral Active Support
BSD Blind Spot Detection

1 Executive Summary

This document describes the results of comparative evaluation conducted in driving simulators and also with real vehicles on test tracks of the third project cycle, to demonstrate the added values of the integrated enablers in the TeamMate car. It consists of two parts (section3 and section 4). The first part (section 3) mainly introduces the updated of the individual enablers (E1.1 "Driver monitoring system", E1.2 "V2X communication", E2.1 "Driver intention recognition", E3.1 "Situation and vehicle model", E4.1 "Planning and execution of safe manoeuvre", E4.2 "Learning of intention from the driver", E5.1 "Online risk assessment", E6.1 "Interaction modality", E6.2 "TeamMate multimodal HMI", E6.3 "Augmented reality"). In section 3, there is a subsection for each mentioned enabler above that addresses the development within AutoMate and the improvements in comparison to state of the art, and the final status. The second part documents the results of comparative evaluations conducted in driving simulators and also with real vehicles in the section 4. For each demonstrator, a TeamMate system setup with several integrated enablers was compared against a simulated baseline system for the AutoMate scenarios (PETER, EVA, MARTHA).

Section 4.1 and section 4.5 describe the evaluation study of TeamMate concept in the PETER scenario on rural roads conducted in the driving simulator and with a real vehicle on test tracks. The baseline car was a state-of-the-art automated car. For the TeamMate car, all enablers mentioned above were integrated in the ULM driving simulator, whereas the enablers of Planning and execution of safe manoeuvre, Interaction modality, TeamMate multimodal HMI (Cluster + audio) were integrated in the ULM vehicle. The evaluation results in the ULM simulator show a benefit of the TeamMate car compared to the baseline car regarding efficiency, usage of automation, usability, workload and willingness to buy and pay. Besides, the performance of the integrated enablers in the TeamMate car was rated relatively high. The evaluation results in the ULM vehicle show that TeamMate car doesn't show the added value regarding trust, acceptance and safety compared to the baseline where a human driver carried out the overtaking maneuverer. The lateral control was neither pleasant nor accustomed nor predictable in the TeamMate condition and the test person's skin conductance level increased over the time. However, the TeamMate car was rated higher than the baseline condition regarding usability and willingness to buy.

Section 4.2 and section 4.6 describe the evaluation study of TeamMate concept in the EVA roundabout scenario conducted in the driving simulator and with a real vehicle on test tracks. The baseline, an autonomous vehicle which follows the driverless approach, was compared against a TeamMate car. In the REL simulator, the Team Mate system was integrated TeamMate HMI, interaction modality, Driver Monitoring System and learning of intention from the driver, whereas situation and vehicle model, planning and execution of safe manoeuvre, TeamMate HMI (Cluster + audio, Central stack display, HUD) were integrated in the CRF vehicle. The evaluation results in the REL simulator show a benefit of the TeamMate car compared to the baseline car regarding trust, acceptance, workload and willingness to buy and pay. Besides, it also demonstrates the added value of TeamMate system in terms of efficiency and safety.

Section 4.3 and section 4.4 describe the evaluation study of TeamMate concept in the MATHA roundabout conducted in the driving simulator and with a real vehicle on test tracks. The baseline, an autonomous vehicle which follows the driverless approach, was compared against a TeamMate car. For the VED simulator and VED vehicle, the evaluation results show no benefit of the TeamMate system regarding acceptance, trust and, usability compared to the baseline. However, participants prefer the TeamMate system and their willingness to buy is higher for the TeamMate system than the baseline system.

For CRF vehicle, the TeamMate system show its benefit with regard to acceptance, willingness to buy and willingness to pay compared to the baseline car. However, the workload with the TeamMate system is higher than the baseline car.

2 Introduction

This document describes the results of the evaluation studies of the integrated TeamMate system run in cycle 3 of the project. Based on the results of the evaluation studies performed in the second cycle in the different demonstrators and the previously defined scenarios both the different enablers and the integrated TeamMate systems have been improved and further developed. Based on these developments it was possible to integrate the enabling technologies of the TeamMate car not only in the driving simulators of the AutoMate project but also in three demonstrator vehicles to demonstrate and evaluate the TeamMate car functionality in the three defined scenarios on real road in test-track studies.

The basic principle of the TeamMate car concept is that driver and TeamMate car functionality work together as team players. This means that both the driver and the automation support each other if necessary when performing driving manoeuvres. This creates basically two different cooperation situations that were coined in D6.2 as **A2H support**, when the automation supports the human driver and **H2A support** when the human driver supports the automation. In D1.3 and D1.5 different use cases and scenarios were defined that serve as critical test cases for the evaluation of this interplay between human driver and automation and that demonstrate the limits of currently available traditional vehicle automation approaches. These scenarios have been used to evaluate the TeamMate car concept in the evaluation studies reported in D6.2 and they were used again in the evaluation studies reported in this deliverable D6.3. The PETER scenario exemplifies a scenario where the human driver can support the automation to solve a situation more efficiently than the automation could do as the automation’s environment perception is impaired. The EVA scenario represents a scenario where the situation is too complex for the automation and the driver needs to be brought back into the loop to monitor the automation in handling the complex situation. The MARTHA scenario stands for those class of situations where the human driver has to be efficiently brought back into the loop to take back the control of the vehicle from the automation.

The cooperative interaction of human driver and vehicle automation in these different scenarios was possible by integrating the identified required enabling technologies, such as Driver Monitoring System to check whether the driver is available in case the driver should take some or full control of the driving task, the Driver Intention Recognition to understand the human driver’s plans in given traffic situations and to best support these planes, the Online Risk Assessment to be able to suggest and perform only safe manoeuvres and sophisticated interaction strategies (including HMI, AR and a concept of interaction modalities) that facilitate the driver's understanding of the automation behaviour and its plans and to easily change the automation's plans according to changing priorities and changing environmental conditions without losing the maximum possible support by the automation.

This TeamMate car system was tested in various instantiations adapted to the requirements of the different scenarios to optimize the project efficiency and to be able to address these many classes of situations under different conditions. In cycle 3 we carried out in total six evaluation studies, three in real vehicles demonstrating the systems principal feasibility and positive effects under realistic conditions and three evaluation experiments in high-end state-of-the art driving simulators that allowed the evaluation of the TeamMate car system under more complex and critical conditions.

To adequately evaluate the TeamMate car system in the different scenarios with their different requirements the methodology described in D6.1 was applied. Specific baselines and KPIs have been used as described D6.1 for each demonstrator in the different scenarios. This allowed us to evaluate the specific gain of the cooperative driver-vehicle interaction realized in the TeamMate care systems in terms of safety, efficiency, trust in automation and acceptance of the new technology in the different scenarios.

3 Update of Enablers in Cycle 3

This section describes the updates of the individual enablers since the last cycle and also the final status of the enablers in the TeamMate car.

3.1 Description of Enabler Updates

3.1.1 E1.1 Driver monitoring system with driver state model for distraction and drowsiness

This section presents a synthesis of the Driver Monitoring System (DMS) overall related work performed in Automate. It includes 3 main parts:

- Work around the DMS integration in the demonstrators
- Work for the improvement for the Drowsiness model
- Work for the improvement of the driver attention model including the identification of the areas the driver is looking at.

The Driver Monitoring System (DMS) is a monocular vision-based system observing the driver's face which estimates the driver physiological and behavioural states including drowsiness and visual distraction (see Figure 1). The system detects, tracks the driver's face and computes features as eye closure, eye/head gaze, head pose required to model the different driver states. DMS is fully automatic, works in real time by night and day conditions.

The Automate Human Machine Interaction (HMI) module makes use of the state estimation to adapt the takeover strategies and warnings.

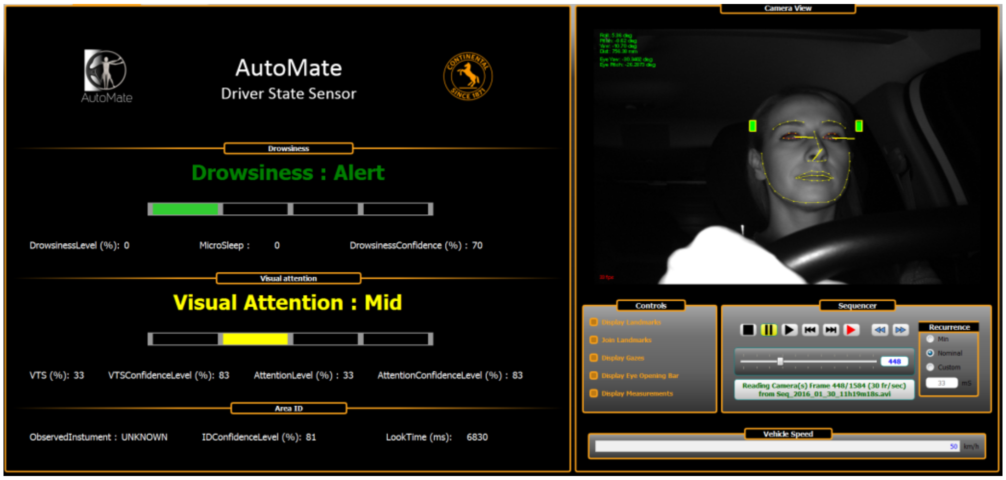


Figure 1: DMS graphical user interface

DMS integration in vehicles and simulators

Within the Automate project the DMS has been integrated in the following demonstrators:

- VED real vehicle demonstrator (see Figure 2)
- ULM simulator demonstrator
- REL simulator demonstrator
- CRF real vehicle demonstrator

The integration in the Automate demonstrators brings issues which requested some specific improvement/adaptation of the tooling, process, communication interfaces. For each demonstrator the following integration tasks have been done:

Physical integration objective is to determine the best camera pose (position and orientation) in compliance with the vehicle integration constraints (camera occultation, intrusiveness, etc.) It includes an analysis of the DMS performances for different selected camera pose for each demonstrator

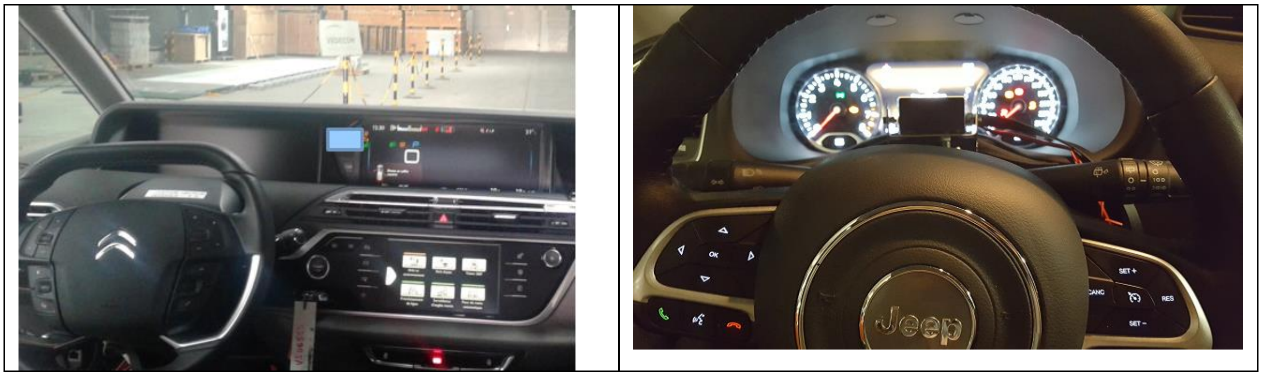


Figure 2: DMS Camera integrated in the Vedecom car (left; blue overlay) and CRF car (right)

DMS calibration aims to determine the camera pose in the vehicle coordinate system. This is done using a set of targets and Continental tools. Within the Automate project the tools were improved to ease the calibration process and improve the camera calibration accuracy. The calibration process used within the integration in the Vedecom car is described in detail in the deliverable 6.2.

The DMS parameters/configuration are determined to optimize the DMS functionalities according to the camera pose and cockpit configuration. This task consists first in collecting recordings of a set of drivers performing a specific protocol. During this protocol the drivers must look at different areas of the vehicle (Instrument cluster, mirrors, ahead, central display, etc.), move and incline their head, and perform some facial related actions (blinking, talking, etc.). The comparison of the DMS output on these videos are compared to the protocol ground truth in order to determine the best set of parameters. It must be noted that the protocol was defined according to the Automate requirements.

The communication interface and protocol have been adapted to the software platform of the demonstrators. Validation tests have been performed jointly with the demonstrator technical team ensuring a high reliability.

The graphical user interfaces (GUI) have been adapted to the partners requirement providing understanding and visibility on the DMS functionalities.

The drowsiness model

The Continental's algorithm makes a direct estimation of the drowsiness mainly based on driver blinking behavior. This algorithm has good performances, however, there exists some limit cases, typically when the driver wears Infrared-blocking glasses, in this case the algorithm is unusable because the camera cannot see the driver's eyes. Within Automate Continental has been focusing the development on improving the eyelid/eye opening based model by a drowsiness model based on non-eye features. In deliverable 2.4 we present the first concept based on head movements only. The work has been pursued by extending the model to all non-eye signals the DMS tracker provides; such as head pose/activity related signals and mouth related signals

The algorithm principle of the non-eye drowsiness model makes use of a learning base approach based on Random Forest classifiers (RF) applied to a set of features (Mean, Variance, Energy, etc.) computed over a defined time-period (150s, 180s, 240s and 300 seconds) for each selected signal.

Evaluations have been done on labelled drowsiness recordings collected on 30 subjects in a simulator.

The results obtained without a preliminary phase of normalization have highlighted the necessity of a feature normalization.

The Figure 3 below shows the recall results obtained on the 30 drivers after a phase of feature normalization on the first 10 minutes of highway driving where the driver is considered perfectly awake.

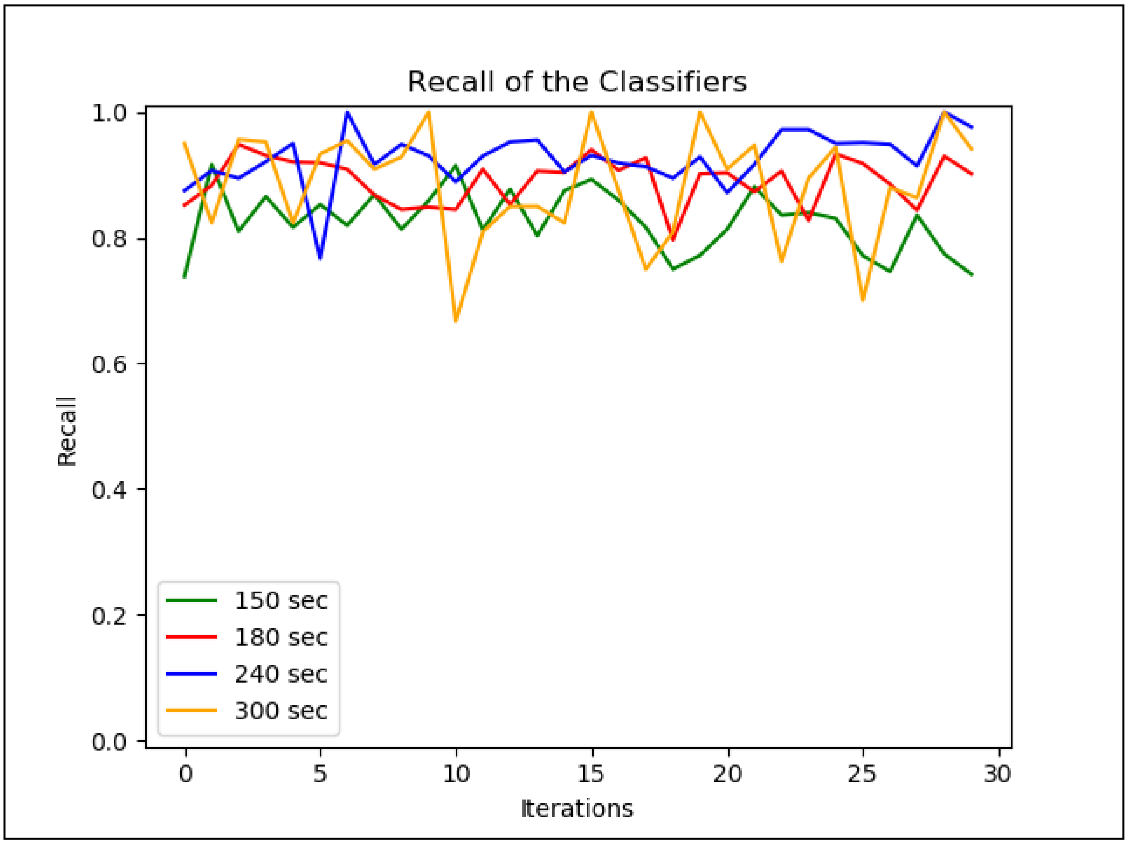


Figure 3: Drowsiness recall of 30 drivers

It must be noted that only highly drowsy states and clearly non-drowsy states have been considered. States ranging from 4 to 7 in the KSS scale have been excluded.

At the end we can say that this algorithm works well to detect highly drowsy or clearly non-drowsy drivers, but it will be much more difficult for it when it comes to evaluate sequences where the drivers is between these 2 classes.

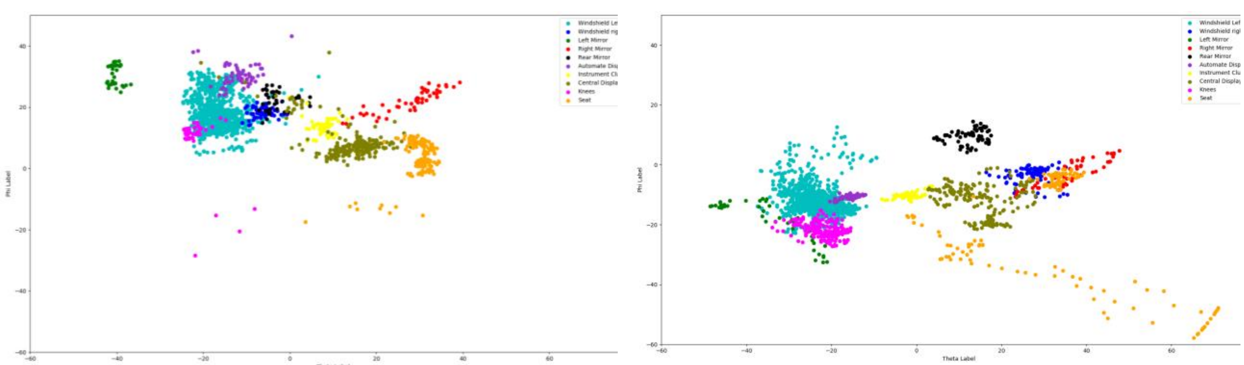
Visual attention model and identification of the area the driver is looking at.

The objectives of the work were to optimize the “Off-road/On-road” detection, adapt the visual attention model to the Human Machine Interface design and finally improve the identification of the area the driver is looking at.

The works have been done mostly on the video database collected in static and driving conditions at the Satory test track with the Vedecom demonstrator car.

The Figure 4 below shows for 2 drivers the values in degree of the pitch (vertical axis) and yaw (horizontal axis) angles computed by the DMS for the different areas the driver is looking at during the test protocol.

As one can see the angles values can be significantly different for the same instrument which of course degrades the identification of the vehicle areas the driver is looking at. This issue calls for an eye gaze calibration which needs to be done automatically while driving without interfering with the driver.



Driver 1 Driver 2

Figure 4: AOI (Areas of Interest) for two drivers during the test protocol

Within Automate we have developed a concept based on the 3 hypotheses:

- Statistically the driver looks much more in front than in other direction. This hypothesis allows to calibrate the front (to the road) eye gaze.
- The major number of extreme head yaws are because the driver looks at the left or right mirror. This hypothesis allows to calibrate the lateral mirrors eye gaze.
- The offset angles applied to a calibrated area (front, left mirror, right mirror) can also be applied to the areas nearby the calibrated one.

We have developed this concept on simulation. We achieved better results for the calibrated areas: the detection of the road, left and right mirror are above 85% for all tested subjects. Still the performances for the other areas are much lower. This is mainly due to the eye gaze estimation noise and the non-optimal position of the camera.

Intelligent Vehicles demonstrations

The DMS integrated in the CRF demo car (Eva scenario) and Vedecom demo car (Martha scenario) has been successfully demonstrated during the track tests of the Intelligent Vehicles demonstration event at Satory.

In both Eva and Martha scenario DMS is used to inform the HMI if the driver is distracted or not.

These demonstrations have shown the very good performances of the DMS for the different scenarios even in adverse light conditions (bright sunny day with direct sun light). The implemented strategy to trigger the distracted flag worked well activating the “distraction” flag with the appropriate timing and according to the driver distraction state. It must also be noted that during these tests no detection lack and no false detection have been observed. These demonstrations have also proven the reliability of the integration for different cockpit position and the reliability of the communication interfaces developed specifically for the CRF and Vedecom systems.

3.1.2 E1.2 V2X communication

In this section, the V2X communication system related developments and their final statuses are summarized.

At the beginning of the project, off-the-shelf Cohda Wireless MK5[2] V2X communication devices were brought. These devices provide state of the art V2X communication features including the ETSI G5 protocol stack. During the project, several development and tests were carried out to utilize the capabilities of the equipment.

First, a robust and flexible application were developed that is able to transmit custom messages between cars (i.e. on board units) or infrastructure (i.e. road side units). The concept of such application was born during the AutoNet2030[3] project. The benefit of this feature is the possibility of rapid implementation of new kind of messages or the newer version of the existing ones. Furthermore, the application is able to transform data streams between different transport layer protocols: IP/TCP, IP/UDP, GeoNetworking/BTP. Using the vehicles' OBD connector, it can also capture the data stream from CAN bus.

Besides that, tests were performed to understand how the standardized Cooperative Awareness Message (CAM) and Decentralized Environmental

[2] https://cohdawireless.com/solutions/hardware/mk5-obu/

[3] http://www.autonet2030.eu/

Message (DENM) can be produced. During laboratory tests, the compliance of the standards was investigated. This is very important for the interoperability of different devices that implement the same standard.

Secondly, the previously mentioned application was improved to be able log and record the V2X communication (and also any other local communication if necessary). The recorded data streams can be replayed in real-time, therefore the number of field tests can be reduced.

Furthermore, several field tests were carried out to record real data for relevant AutoMate scenarios, as well as to measure the capabilities of the MK5 devices in real environment. It was found that the performance of these devices meets the expectations, i.e. they similarly perform as other state of the art equipment.

Finally, based on the recording and replaying features, a visualization framework was developed to be able to show what the V2X communication is capable of. The framework has web-based frontend that runs in any modern browser. The prototype version is deployed on a Raspberry PI 3. Its Wi-Fi works in AP mode, thus the communicated information can be followed by the users easily using a smartphone or tablet. Of course, it is able to work with live data as well, which makes field testing more convenient.

A V2X communication device has been deployed as road side unit in Satory test track at Vedecom that broadcasts road works warning message for field testing.

3.1.3 E2.1 Driver intention recognition

This section summarizes the development and final status of E2.1, “Driver intention recognition” as previously described in the deliverables of WP2 [1, 2, 3, 4, 5, 6] and to be published in [7].

E2.1 provides the TeamMate car with knowledge about probable and desired current and future manoeuvre intentions of the driver. Such knowledge is required to develop a shared understanding between the driver and the automation. When the driver is in control, such knowledge can be used to assess the safety of an intended maneuver prior to its execution and provide adequate information and warnings. If the automation is in control, it can be used to select intention-compliant behavior of the automation or to detect and communicate mismatches between the driver’s intention and the TeamMate car’s behavior.

To realize E2.1, we developed a conceptional model for intention and maneuver recognition based on (conditional) Dynamic Bayesian Networks, whose structure and parameters can be estimated from annotated time-series of human driving behavior. The model represents the statistical and causal relations between the driver’s intentions, the performed driving maneuvers, and available sensor information about the traffic situation and vehicle state. The model then addresses the problem of intention and maneuver recognition from the available situational context, where the situational context is given by a set of observable features, comprised and derived from the state of the TeamMate vehicle, including its position in the road, and the traffic situation, i.e., the state of other traffic participants. We treat the state of traffic participants as a set of observable *inputs* or *causes* for the formation of intentions and the state of the TeamMate vehicle as a set of observable *outputs* or *effects* resulting from the driving behavior. The model assumes that the intentions of the driver evolve based on the situational input encountered. The intentions then manifest themselves in the execution of driving maneuvers whose effects can be observed. The model formalizes these assumptions in a conditional Dynamic Bayesian Network that is composed of a variable set of sub-models, e.g., to model the probability distribution over intentions given the observable inputs. The detailed structure of these sub-models and the parameters of their probability distributions and density functions involved are estimated from annotated experimental data. Parameter estimation is achieved via Bayesian parameter estimation, structure learning is achieved via a greedy hill-climbing search in a search space of model structure using a discriminative variant of the Bayesian Information Criterion [7].

During runtime, the model can be used in two different settings, akin to *intention and maneuver recognition* and *intention prediction*. If the driver is in control of the vehicle, both observable inputs and outputs can be used to simultaneously perform intention and maneuver recognition by continuously inferring the joint belief state over the current intentions and maneuvers given all available inputs and outputs observed thus far. If the automation is in control, the model can be used for intention prediction by ignoring observable effects resulting from the automation, and continuously inferring a belief state over the intentions given the available situational input instead.

Throughout AutoMate, we adapted the conceptional model to three different scenarios using corresponding datasets: real-world motorway, simulated rural road, and simulated roundabout scenarios. In [6], we reported on the development of models for rural road and roundabout scenarios, as used for the Peter and Eva scenarios. For [7], we further refined these models and developed an additional model for highway scenarios, as used for the Martha scenario. The model for rural road scenarios has been integrated in the ULM simulator and VED real vehicle demonstrator, the model for roundabout scenarios has been integrated into the REL simulator demonstrator.

The final version of D2.1 has been evaluated on unseen test data. Summarizing the latest results [7], the model for intention recognition on two-lane motorways achieves an accuracy of 0.888, precision of 0.617, recall of 0.831, F-score of 0.708, and a false positive rate (FPR) of 0.101. The model for intention recognition on rural roads achieves an accuracy of 0.952, precision of 0.838, recall of 0.844, F-score of 0.841, and a false positive rate of 0.029. Lastly, the model for predicting the intention of a driver to enter roundabouts achieves comparative results with an accuracy of 0.850, precision of 0.886, recall of 0.808, F-score of 0.845 and false positive rate of 0.107. To allow for a numerical comparison with other approaches for driver intention recognition on motorways reported in the literature [8, 9, 10], we analysed the time span between the model for driver intention recognition on motorways consistently predicting a lane-change intention and the TeamMate car's centre crossing the lane boundary. Evaluated on unseen test data, the model reaches an average prediction horizon of 6.08s. Discarding individual prediction times greater than 10s (the overall execution time of a lane change manoeuvre is usually assumed as approx. 10s [11]) results in a more conservative prediction time of 5.57s. A similar analysis for intention recognition on rural roads shows that the model is able to predict a lane change intention 4.60s prior to the TeamMate car crossing the lane boundary (or 4.44s when discarding values greater than 10s).

E2.1 has been successfully integrated in the VED real vehicle, the ULM simulator, and the REL simulator demonstrator demonstrator to help enabling our vision of the TeamMate concept. For this, E2.1 has been implemented together with the functionality for the prediction of the spatial and temporal evolution of the traffic scene (E3.1), online risk assessment for dynamic objects (E5.1), and online learning (E4.2) into a single C++ Dynamically Linked Library. Within the second and third cycle, this DLL was embedded into functional plug-in modules for the simulation environment SILAB, used by the ULM simulator demonstrator, and the third-party software RTMaps, used by the VED real vehicle demonstrator, enabling the utilization of these functionality in corresponding demonstrators. For the REL simulator, we used the TeamMate Extension SDK [12] to compile E2.1 to an executable that connects to the REL simulator. The resulting VED real vehicle demonstrator has been demonstrated during the final event. First and final versions of the ULM simulator demonstrator have been evaluated at the end of the second [13] and third cycle (Section 4.1), the final version of the REL simulator demonstrator has been evaluated at the end of the third cycle (Section 4.2).

3.1.3.1 Comparison with the state of the art

This section primarily summarizes [4] and the discussion and results to be published in [7], to which we refer for more information. Driver intention recognition addresses the problem of anticipating driving manoeuvres, a driver is likely to perform in the near future. As early knowledge about potentially dangerous manoeuvre intentions may serve as a potential enabler to generate adaptive warnings and early interventions, driver intention recognition is an increasingly important topic for the development of advanced driver assistance systems and has become a popular research topic. Approaches reported in the literature (some comparative reviews are provided e.g. in [14] and [15]) mainly differ in respect to the selected scenarios and addressed manoeuvres, modelling techniques used, and the sensor input considered.

Concerning the sensor input, we distinguish between different kind of information, causes and effects. Here, causes should be understood as information perceived by the driver that results in the formation of an intention, e.g., a slow lead vehicle in the case of overtaking intentions. In contrast, effects should be understood as the observable effects on the overall behaviour of the driver and vehicle, resulting from the existence of an intention, e.g., head movements to check the blind spot or the initiation of an overtaking manoeuvre.

Traditional driver intention recognition commonly focusses on modelling the relations between manoeuvre intentions and their effect on the behaviour of vehicle and driver. Existing approaches commonly focus on information about the vehicle state, e.g. provided via the Controller Area Network (CAN) bus, and the location of the vehicle in the lane to recognize driving manoeuvres as early as possible [16, 17, 18, 19, 20, 21, 22]. An obvious limitation of such approaches is the necessity for a manoeuvre to be initialized before it can be recognized. In order to overcome these limitations and extend the predictive capabilities, more sophisticated approaches consider the inclusion of driver-based input obtained from camera systems, e.g., by tracking head and eye movements of the driver, to recognize characteristic preparatory measures preceding the execution of a manoeuvre, e.g. shoulder checks [14, 8, 23, 18, 24]. Driver-based input provides valuable information, but their inclusion only shifts the recognition of manoeuvre intentions to earlier stages of execution and with the increasing introduction of automation to the vehicle, driver-based input for driver intention recognition may become misleading and, in the extreme case of fully autonomous driving, obsolete.

For the development of driver intention recognition in AutoMate, we primarily focussed on causes for intentions, given by the situational context, esp. the traffic situation, i.e., information about vehicles in the vicinity of the driver. Up to now, potentially due to limited sensor capabilities, such information has not been used thoroughly for intention recognition, but is either neglected entirely [25, 16, 21, 26, 27, 17, 18, 24], or restricted to the immediate surrounding of the driver, namely the lead vehicle [28, 23, 22, 29, 19] and vehicles in the blind spots [8, 9]. This is surprising, as where the inclusion of driver-based input only shifts the recognition of manoeuvres to earlier stages of the execution, information about the current traffic situation should be able to provide information suitable to actually *predict* the intentions of the driver, e.g., a slow driving lead vehicle may be the reason why the driver may form the intention to overtake, while an acceptable gap may provide the reason why a driver intends to return to the original lane.

Within AutoMate, we developed a model for driver intention recognition that refrains from driver-based input but instead explores the utilization of information about the traffic situation to extend the predictive capabilities of the model and enable the use in highly automated or autonomous driving.

Models for driver intention recognition have been widely studied in context of different scenarios and modelling techniques [14, 15]. Many studies address lane change manoeuvre on motorways and rural roads [28, 25, 8, 9, 16, 23, 21, 10] or turning and stopping manoeuvres at intersections [26, 30, 27, 22, 29]. In contrast, roundabout scenarios are relatively uncharted. Muffert [31] developed a method for the safe entrance to roundabouts using stereo cameras, however, [32] proposed a model for recognizing driver’s intentions to exit or remain in a roundabout. In AutoMate, we developed a conceptional model that was adapted to three different scenarios: real-world highway, simulated rural road and simulated roundabout scenarios.

Modelling techniques primarily include probabilistic *generative* approaches like Dynamic Bayesian Network (including Hidden Markov Models and their variants) [26, 22, 17, 18, 19, 20], supposed to be better suited for modelling temporal aspects [24], or probabilistic and non-probabilistic *discriminative* approaches, including Support Vector Machines (SVMs) [28, 16, 30, 32], Multi-Layer Perceptrons [27], or logistic regressions [15], which are better suited to include complex feature set, like e.g. head and eye tracking data, for which the definition of a generative model may be complicated to define correctly [33]. In the context of decision-making, the problem of predicting driver manoeuvres based on the traffic situations is also addressed by gap acceptance models [34, 35]. Gap acceptance models assume the existence of a latent critical gap at which a driver is indifferent between accepting and rejecting a gap in traffic [34]. This gives rise to a gap acceptance function describing the probability that a driver accepts an offered gap, usually realized as a logistic regression [34, 35]. Unfortunately, the limitation to logistic regressions can be overly restricted in more complex scenarios. In AutoMate, we tested both generative and discriminative approaches, and settled on conditional Dynamic Bayesian Networks composed of sub-networks, which can be interpreted as a combination of both generative and discriminative approaches.

One of the most sophisticated of such approaches for intention recognition on *motorways* implemented in real vehicles up to date is the discriminative model described by [8] and evaluated in [9]. They used Relevance Vector Machines as a probabilistic alternative to SVMs for learning a model for online recognition of lane change intentions based on information about the vehicle state, head-tracking, the lead vehicle and vehicles in the blind spot. The resulting model can recognize lane change intentions of human drivers up to approx. three seconds [8, 9] prior to the actual crossing of the lane (in the following denoted as *prediction horizon*). Building on these results, [23] proposed the use of discriminative Latent-Dynamic Conditional Random Fields and extended the driver-based input by hand and foot motion cues. They the state improved prediction horizons, but do not report actual numbers. Just recently, [10] presented a model for recognizing and predicting lane changes, realized as a (non-dynamic) Bayesian Network that incorporates both driver-based input and the traffic situation, with the traffic situation being condensed into three discrete levels of occupancy for each lane. They report a vastly improved average prediction horizon of 7.8s at a recall of 0.7. Our model for intention recognition on motorways achieves an average prediction horizon of 6.08s (or 5.57s when discarding individual prediction times greater than 10s) at a recall of approx. 0.8. This exceeds the performance of [8, 9] but falls short of the results presented by [10], showing the potential benefit of driver-based input, if available. For rural roads our model is able to predict a lane change intention 4.60s prior to the TeamMate car crossing the lane boundary (or 4.44s when discarding values greater than 10s).

3.1.3.2 Pre-existing developments

As previously described in [6], for the development of the models for driver intention recognition in AutoMate, we started with a pre-existing framework, consisting of libraries and algorithms for the creation and utilization of (Dynamic) Bayesian Networks, originally developed during the former EU project HoliDes. Within AutoMate, this framework was significantly updated and extended, e.g., to allow for the learning and utilization of more complex model structures and parametric distributions, enabling the update of model parameters during runtime (as required by E4.2 "Learning of intention from the driver"), and enabling the use in rural road and roundabout scenarios.

To start the model development during the first cycle of AutoMate, prior to the conduction of any data collection experiments, we made use of experimental data obtained during the former EU project HoliDes. This data represented real-world recorded in a CRF prototype vehicle with human drivers manually performing overtaking vehicles. The data has been used for the development of libraries and tools for the development and evaluation of models for driver intention recognition until explicit experimental data for the development of model for driver intention recognition on rural roads became available.

3.1.3.3 Facing the cold start problem

In regard to driver intention recognition in AutoMate, the cold start problem can be understood as the problem of recognizing and predicting the intentions of an individual driver during the introduction period of the system, when insufficient information about the specific driver is available. The term "cold start problem" originated in the context of recommender systems, where it refers to the problem of performing inferences for a user or item before the necessary information for such inferences have been gathered [36].

As a mitigation strategy to face the cold start problem for driver intention recognition in AutoMate, we rely on the utilization of a prior or default model, representing the average or a group of drivers that can then be adapted to the individual driver, once such data is available. In AutoMate, this default model is given by E2.1 "Driver intention recognition". Utilizing E4.2 "Learning of intention from the driver", this default model can then be adapted to the individual driver in an online fashion (c.f. Section 3.1.6).

As described in the deliverables of WP2 [1, 2, 3, 4, 5, 6], enabler E2.1 “Driver Intention Recognition” is realized in terms of probabilistic models whose parameters are structures have been learnt offline, using datasets obtained in simulator studies, as conducted by OFF, ULM, and HMT during the second and third cycle. To start the model development during the first cycle of AutoMate, prior to the conduction of any data collection experiments, we made use of experimental data obtained during the former EU project HoliDes (c.f. Section 3.1.3.2). The experiments and datasets to train the models have been described in deliverables D2.4, Section 4.3 [4], D2.5, Section 3.1.2.3 [5], and D2.6, Section 4.3.1.2 [6]. The experiments have been designed to favor multiple participants with comparable few iterations over the contrary. The resulting models represent groups of drivers or the average driver, hopefully being reasonable applicable, although not perfectly adapted, to a broad spectrum of potential individuals.

3.1.3.4 Driver profiles

To some extent, E2.1, after being adapted by E4.2, can be interpreted as a *user model* or *driver profile* to infer useful information about an associated driver, in this specific case limited to the potential driving intentions given the current traffic situation. This raises the question of whether and how it would be possible to extend the capabilities of the user model to additional information of interest, e.g. lateral/longitudinal driver preferences or whether the driver prefers risk averse or friendly driving behavior. We note that driver profiles were not planned to be addressed in AutoMate and would require substantial effort in profile, privacy, and security management beyond the scope of AutoMate. That said. we belief that the models for driver intention recognition, adapted to the individual driver using E4.2, could be extended to or embedded in driver profiles in future research. As described in the deliverables of WP2 [1, 2, 3, 4, 5, 6] and [7], the models for driver intention recognition basically encode a conditional probability / density distribution over the temporary evolution of a set of latent states of the driver, e.g. intentions and currently performed maneuvers, and observable effects of the driving behavior, e.g. speed and control signals, given the observable vehicle state and traffic situation. A natural first step for the extension into driver profiles would be the addition of the distribution over the observable vehicle state and traffic situation to obtain the joint distribution over all variables. If modelled correctly, the resulting joint distribution could then readily be used for driver intention recognition and to infer additional queries of interest, e.g., velocity preferences in different situational contexts, while relying on the same algorithmic foundation already in place. Due to the use of embedded Bayesian classifiers [6, 7], information necessary for the realization of such driver profiles is already partially encoded in the models for driver intention recognition. However, due to the use of discriminative machine-learning techniques for feature selection that focused on maximizing the performance of intention recognition, potentially valuable information for the realization of driver profiles may not be included.

We note however, that the Dynamic Bayesian Networks used for driver intention recognition exploit knowledge about temporal dependencies that may not be necessary for the realization of driver profiles. Simpler non-dynamic Bayesian Networks may be sufficient for modelling driver profiles, resulting int easier to learn, more efficient, and potentially more robust models. Nonetheless, such simpler models could easily be utilized in conjunction with the models for driver intention recognition, utilizing the same algorithms for parameter and structure learning, performing inferences, and online adaptation. The AutoMate partner OFF currently investigates the modelling of driver profiles in the BMWI project "AutoAkzept"[4] [37].

3.1.4 E3.1 Situation and vehicle model

The final status of Enabler 3.1 "Vehicle and Situation Model" within WP2 [1] [2] [3] [4] [5] [6] is summarized in this section. Representing the state and semantic information about a scene is the desired capability in autonomous driving systems. Therefore the situation model is an intermediate layer between the sensor and communication platform updating the state information for the driver models and vehicle model. Based on the sensor information, TeamMate vehicle's current belief about the world is represented, and at constant interval the situation model is update via the sensor and communication platform.

Within the scope of AutoMate project the development of the enabler 3.1 is focused by considering two features; *Semantic enrichment of the situation model*, which extends inputs from perception layer with semantic information and the *prediction of the future evolution of the traffic scene* based on the enriched and current state of the situation - and driver model.

3.1.4.1 Semantic Enrichment of the situation model

The semantic enrichment model extends the inputs of the perception layer with semantic information from the scene model. Semantic information such as legal drivable maneuvers is inferred, on the basis of the modeled

[4] AutoAkzept – Erhöhung der Akzeptanz automatisierten und vernetzen Fahrens. https://www.dlr.de/ts/desktopdefault.aspx/tabid-10704/20365_read-54052/, last visited 24.09.2019.

relationship between the agents and the scene elements and the set of traffic rules. In [2] we proposed extended ontology with logical rules. An ontology is a semantic model allowing to express the domain knowledge, the modeled ontology is used to reason about the complex relations and facts. For this work to model the relations between scene elements a *Web Ontology Language (OWL)* was used, and a *Sematic Web Rule Language (SWRL)* was used to extend the modeled ontology with traffic rules. With amalgamating the two, OWL and SWRL we successfully model the complete domain knowledge to infer possible maneuvers for vehicles in the given scene. Figure 5 provides the overview of sample ontology taxonomy describing the spatial, temporal and semantic relations between scene objects. As a matter of visualization a sample taxonomy tree is presented here. In [2] we define the classes for modeling the relations between the scene elements

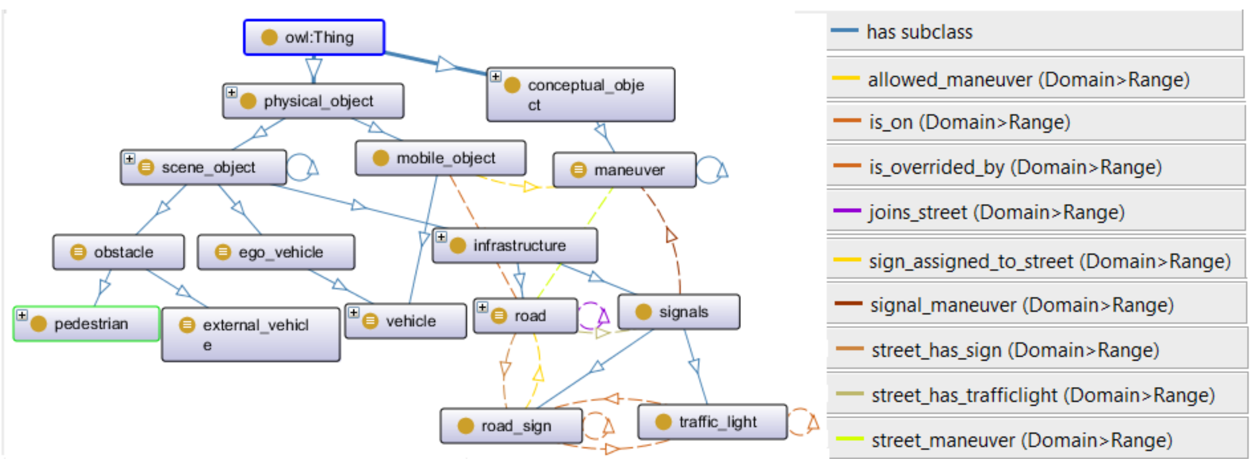


Figure 5: Illustration of ontology taxonomy (left) and the relations legend (right). For more detail please view it in colour version.

illustrates the sample set of traffic rule that are based on the relations and concepts of the ontology taxonomy to build the complete domain knowledge. With the help of the reasoner the allowed manoeuvre for each of the vehicle in the scene can be inferred, the inferred manoeuvre could be used as a prior for predicting the traffic evolution.

Name	Rule	Meaning
R1	$stop\_sign(?s) \wedge maneuver(?m) \wedge signal\_maneuver(?s, ?m) \rightarrow stop(?m)$	Stop sign allows stop maneuver
R2	$give\_way\_sign(?s) \wedge maneuver(?m) \wedge signal\_maneuver(?s, ?m) \rightarrow slow\_down(?m)$	Give way sign allows slow maneuver
R3	$traffic\_light(?l) \wedge has\_tl\_state(?l, ?s) \wedge red\_light(?s) \wedge signal\_maneuver(?l, ?m) \rightarrow stop(?m)$	Red traffic light allows stop maneuver
R4	$traffic\_light(?l) \wedge road(?r) \wedge road\_sign(?s) \wedge trafficlight\_assigned\_to\_street(?l, ?r) \wedge sign\_assigned\_to\_street(?s, ?r) \wedge signal\_maneuver(?s, ?m) \wedge signal\_maneuver(?l, ?m2) \rightarrow street\_maneuver(?r, ?m2)$	Traffic lights has high priority comparing to traffic signs, if both are assigned to the same road
R5	$road(?r) \wedge road\_sign(?s) \wedge sign\_assigned\_to\_street(?s, ?r) \wedge no\_trafficlight\_assigned\_to\_street(?r, true) \wedge signal\_maneuver(?s, ?m) \rightarrow street\_maneuver(?r, ?m)$	maneuver allowed on that road depend on the assigned traffic sign where there is no traffic light
R6	$mobile\_object(?o) \wedge road(?r) \wedge street\_maneuver(?r, ?m) \wedge is\_on(?o, ?r) \rightarrow allowed\_maneuver(?o, ?m)$	Traffic participants allowed maneuvers depend on the road they're are on
R7	$road(?r) \wedge road\_vehicle(?o) \wedge is\_on(?o, ?r) \wedge has\_max\_speed\_value(?r, ?v) \rightarrow has\_max\_speed\_value(?o, ?v)$	Vehicles allowed maximal velocity depend on the road they are on

Table 1 Basic traffic rules expressed with the help of SWRL

The semantic enrichment module was successfully integrated in the vedecom demo vehicle and was demonstrated at Satory during the end demo event. For packaging the model to facilitate seamless integration, we had developed *JNIOWLBridge in C++ as presented in [4]. The JNIOWLBridge* allows to access the OWL ontology and the reasoner as a C++ function within our semantic enrichment module. The *JNIOWLBridge* is the bridge between the Java OWL/reasoner API and C++ module. Furthermore, in the 3rd cycle we shipped our semantic enrichment module as a C++ Dynamic Linking Library (Dll) to wrap it as a plugin within RTmaps environment of vedecom vehicle, allowing us to have a seamless integration. A communication module to communicate between the perception layer and our semantic enrichment module was also integrated within the vedecom vehicle environment as a Dll [6]. Figure 6 shows the semantically enriched information about permissible driving manoeuvres inferred using the reasoner from the semantic enrichment module. The module was developed within the scope of the AutoMate, and no feature of the component is being inherited from external projects.

Evaluation of this module was performed on synthesized data and tests data; to quantify the measure of accuracy we use the F1 score as a metric. We generated synthetic scenes to account for a large variance in scene appearance and to perform robust evaluation. Due to the fact that our model is a formal model, we obtained 100% accuracy on synthesized data. As we have a core dependency of the perception output. On real data, the accuracy of the model could drop with respect to uncertainties within the outputs of the perception system [6]. We noticed the module performs at 82.95 ms and requires 600 MB for inference on a system with *Intel-i7*-CPU, 8 cores@2,8 GHz and 8 GB RAM. The car-pc in vedecom's real vehicle has the close specification, therefore we see the close performance in the test vehicle. The runtime is proportional to the complexity of the scene.

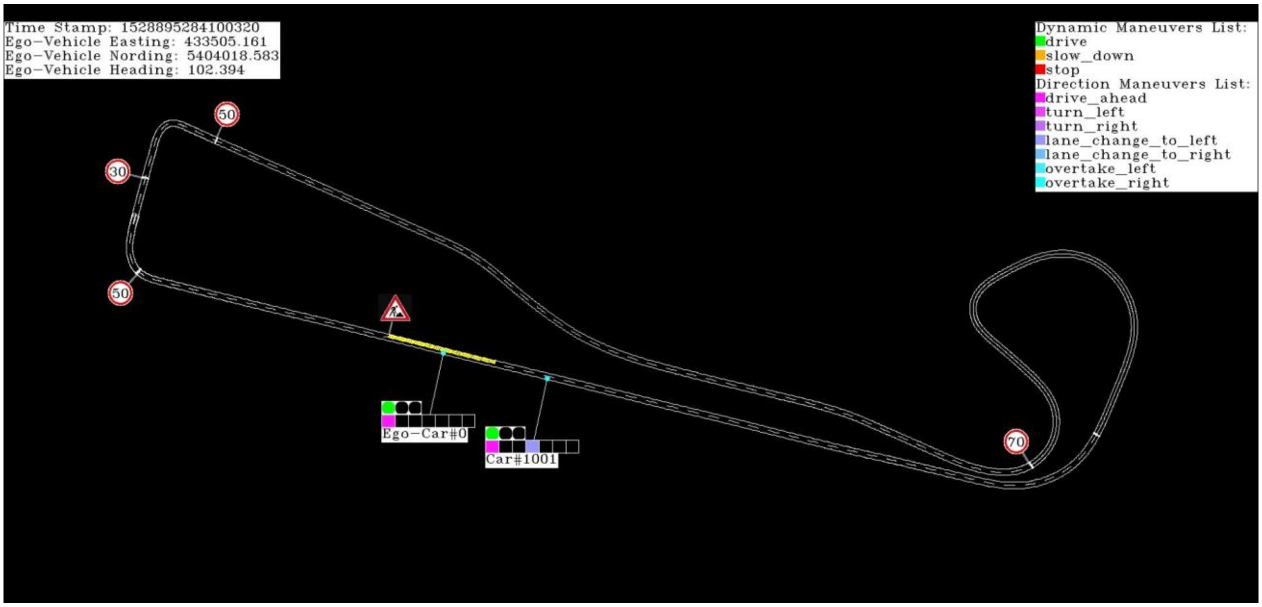


Figure 6: Inferred semantic information about permissible manoeuvres for vehicles based on the modelled domain knowledge from semantic enrichment module. The Ego-car (TeamMate) inferred dynamic maneuver is to "drive" and direction maneuver is to "drive-ahead" based on the semantic information of the scene. The coded colours represent the semantic meaning about inferred maneuvers (Legend). View it in colour to have better intuition.

3.1.4.1.1 Comparison of the state of the art

In [38] the author proposes a *Markov Logical Network* (MLN) framework for situation interpretation and rules mining to infer real-world events under uncertainty and ambiguous sensor information. Similar approaches as proposed in [38] are greatly used for visual surveillance applications, where the aggregation of complex scene information does not necessarily require updates as in autonomous driving task.

As the authors in [39] [40] [41] [42] present, *ontologies* are greatly used for formal representation of the domain knowledge. At times modelling of these ontologies could get complex when considering large space of domain information, and also remain laborious as they are designed manually from experts. In [40] the author reports an inference time of up to 1.17s for a single frame at a complex intersection for reasoning about the scene. It is far from real time inference requirements, although the author in [40] argues, for less-complex scenes 500ms would be enough in real time. Nevertheless, we require at least 100ms for real time driving scenarios.

To overcome the complexities and the high inference time issues, we restrict our domain knowledge to traffic rules. Moreover, limiting to traffic rules remain enough to predicted manoeuvres for TeamMate vehicle and its surrounding vehicles, which provides as prior to predict evolution of traffic. The author in [41] has a similar ontology for an automated vehicle's context model. Following the similar proposed method to design our knowledge base, we greatly reduce the inference time to be within the bounds of 100ms, which is more suitable for automated driving task. Nevertheless, our approach has a limitation towards the consistency check of the inputs from the perception model, these limitations are going to be considered in our future works.

3.1.4.2 Prediction of evolution traffic scene

Long-term prediction of traffic participants is crucial for the development of advanced driver assistance systems and advancement of autonomous driving on public roads [43], e.g., existing trajectory planning components already require prediction horizons of up to ten seconds [44]. To achieve a long-term prediction, the TeamMate car makes use of the *traffic prediction* component, a probabilistic model, comprising situation and vehicle models, to predict the likely temporal and spatial evolution of the TeamMate car and all traffic participants observed in its vicinity. The traffic prediction has been developed for two-lane rural road and motorway scenarios, where it provides a long-term prediction, with user-defined horizon and step size, with integrated recognition of *lane-keeping* and *lane-changing* behaviors for better prediction accuracy [6].

The traffic prediction is characterized by the following key aspects: Traffic prediction is performed at constant intervals for the TeamMate car and each dynamic object detected in its vicinity. Once started, the state of each object considered is predicted for equidistant points in the future, up to a user-defined maximum prediction horizon. The state of each object is represented by a six-dimensional Gaussian belief state over its location (in a two-dimensional global coordinate system), its yaw angle, velocity, acceleration, and yaw-rate. For predicting the state of an object into the future, the constant yaw-rate and acceleration (CYRA) motion model is used. The CYRA motion model is a physical motion model that describes the non-linear dynamics of location, yaw angle, and velocity (of a point mass) under the eponymous assumption of constant yaw-rate and acceleration in a set of motion equations [45]. As the assumption of constant yaw-rate and acceleration is insufficient for long-term prediction, a set of simple driver models, selecting appropriate yaw-rates and accelerations to perform specific manoeuvres based on a map of the environment, is used to enable context-specific alterations of the yaw-rate and acceleration during the prediction. Inference-wise, the prediction is achieved by unscented transformation, as used in unscented Kalman filters [46]. For this, the six-dimensional belief state is condensed into a set of characteristic sigma points, which are then passed through the equations of the CYRA motion model. The resulting transformed sigma points are then used to derive a new six-dimensional belief state, representing an approximation of the result when passing the original belief state through a non-linear function.

Given these components, the traffic prediction operates as follows (c.f. Figure 7Figure 7): The traffic prediction maintains a Gaussian belief state for the TeamMate car and each dynamic object in its vicinity, which is constantly updated whenever new sensor information is provided. At constant intervals and for each object considered, the CYRA motion model and the set of driver models is used to perform a short-term prediction for each considered maneuver hypothesis. At the next time step, the new sensor observations are incorporated into the short-term predictions to obtain the likelihoods for each maneuver. The most probable maneuver is then used for a long-term prediction. The result is a long-term prediction for the most probable maneuver for each dynamic object in the traffic scene, represented as a set of multivariate Gaussian distributions over the location, yaw angle, velocity, acceleration, and yaw-rate for the desired user-defined discrete time steps in the future. The long-term prediction is then used by E5.1 to perform online risk assessment for the TeamMate car (c.f. Section 3.1.7).

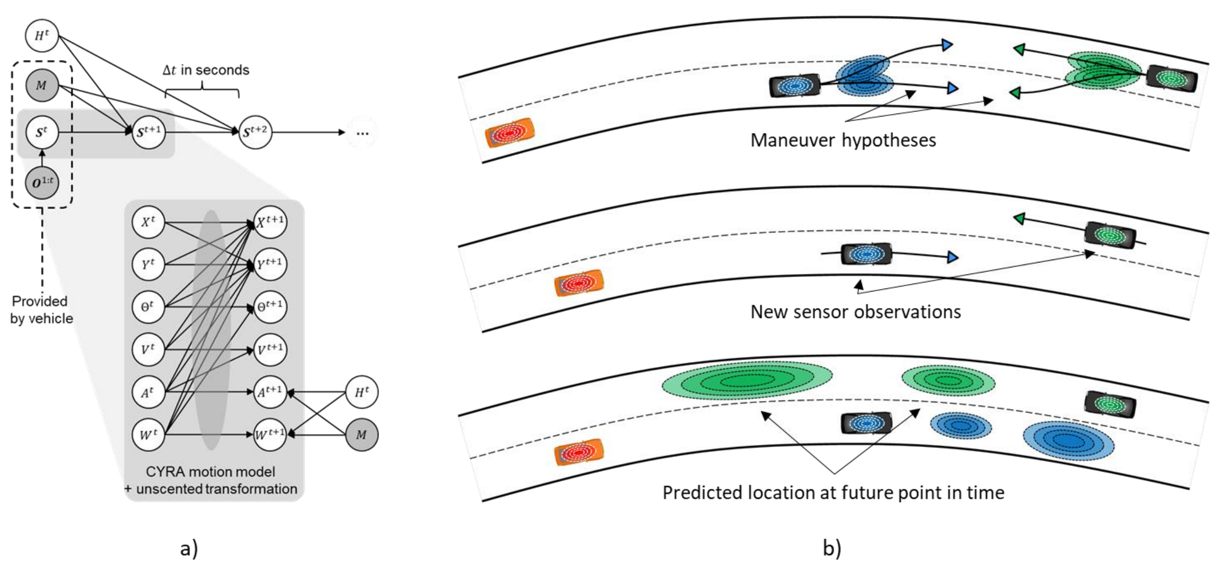


Figure 7: Schematic probabilistic model for traffic prediction (a). Overview of functionality (b).

Evaluated on test data obtained in simulator studies throughout AutoMate, the final version of the traffic prediction achieves a correct rate of prediction above 90% for prediction horizons up to five seconds for vehicles controlled by human drivers while requiring an average execution time of 0.139ms per predicted object and second on an i7-6700 CPU @ 3.40GHz, 16GB Ram desktop computer running a Microsoft Windows 10 64-Bit operation system [6].

Throughout the second and third cycle of AutoMate, the traffic prediction has been successfully integrated in the VED real vehicle and the ULM simulator demonstrator to help enabling our vision of the TeamMate concept. For this, the traffic prediction has been implemented together with the functionality for online risk assessment for dynamic objects (E5.1), the driver intention recognition (E2.1), and online learning (E4.2) into a single C++ Dynamically Linked Library. The DLL was then embedded into functional plug-in modules for the simulation environment SILAB, used by the ULM simulator demonstrator, and the third-party software RTMaps, used by the VED real vehicle demonstrator, enabling the utilization of all functionalities in the corresponding demonstrators. The resulting VED real vehicle demonstrator has been demonstrated during the final event, first and final versions of the ULM simulator demonstrator have been evaluated at the end of the second [13] and third cycle (Section 4.1)

3.1.4.2.1 Comparison with the state of the art

Traffic prediction must deal with uncertainties, arising e.g. from the inability to perfectly observe the current traffic situation, the hidden intentions of the traffic participants, and variability in how these intentions may be executed [47]. Approaches for traffic prediction can be broadly categorized as short-term prediction, long-term prediction, and abstract forms of situation prediction [47]. Short- and long-term prediction attempt to directly predict the evolution of vehicle states on different time scales, while abstract forms summarize the evolution in terms of manoeuvre or intention recognition (c.f. Section 3.1.3) or risk assessments (c.f. Section 3.1.7). As a general enabler for other technologies like risk assessment or trajectory planning, traffic prediction in AutoMate belongs to the former categories.

Short-term prediction relies on motion or vehicle models to predict the short-term motion of a vehicle in which the influence of driver and environment are minor and the prediction depends only on the driving physics and system dynamics [47]. The general topic of vehicle dynamics is well studied and understood (e.g. [48]) and numerous motion models with different degrees of complexity have been proposed for this task [45]: At the lower end of complexity lie linear motion models, assuming a vehicle to travel on a straight path with *constant velocity* or *constant acceleration*. Having linear state transition equations, they allow for a direct utilization in Kalman-filters. Curvilinear models such as the *constant yaw-rate and velocity* or the *constant yaw-rate and acceleration* (CYRA) motion models also incorporate rotation but deny any correlation between velocity and yaw-rate. This assumption is relaxed by the *constant steering angle and velocity* and the *constant curvature and acceleration* motion models. The latter motion models share many similarities with the *kinematic bicycle model* [49] [50], representing the lower end of complexity for the variety of *vehicle models*. Like motion models, vehicle models with many different degrees of complexity have been proposed. Unfortunately, the information necessary for their utilization (e.g., individual tire slip) are not observable by exteroceptive sensors, such that their use is limited to predicting the motion of the ego (TeamMate) vehicle. Schubert et al. [45] performed a comparison of many of these models for the task of vehicle tracking and found the CYRA motion model to be the most effective trade-off in terms of precision and efficiency. The CYRA motion model was also successfully used by [51] and [52]. Based on these findings, we decided to use the CYRA motion model within our traffic prediction.

Independent of the vehicle model utilized, the assumption of constant inputs is only reasonable for prediction horizons of less than a second [15]. For long-term prediction with prediction horizon above a second, the constraints on the possible trajectory of vehicles imposed by the road network and likely maneuvers and maneuver intentions have to be considered [53].

Constraints imposed by the road network are usually either incorporated implicitly using lane-based coordinate systems or explicitly by the use driver models. Many approaches operate on a lane-based reference of Frenet frame, where the x-axis is given by a mathematical function like the course of the road or a planned reference trajectory [20] [47] [43] [54] [55]. Working in such a transformed system greatly simplifies the problem of incorporating road network constraints in that a simple linear model in the transformed space will perfectly follow the road in the Cartesian space. Unfortunately, lane-based representations may not be possible or require complicated treatment for more complicated road networks [54] (e.g., parking lots, intersections, roundabouts). Furthermore, a transformation into Cartesian space, if required from other components along the processing chain, may be complicated and computational expensive. In contrast, our approach directly works in Cartesian space, using more complicated drive models.

[30] and [56] used a combination of Support Vector Machines and Bayesian Filtering for intention recognition and Rapidly exploring Random Trees for

trajectory prediction. [57] and esp. [20] used Hidden Markov Models for intention recognition and Gaussian Processes for trajectory prediction. They used the recent positions of traffic participants to estimate a Gaussian Process over trajectory, which could then be used to predict the (non-linear) trajectory without the need of dedicated driver models. The use of Gaussian Processes seems very promising but is (for now) limited by the prohibiting computational complexity.

[43] proposed a prediction based on particle filters, Monte-Carlo simulations, and a microscopic driver model called the Intelligent Driver Model (IDM) to predict the future longitudinal car-following behavior. More specifically, each traffic participant is modelled by an IDM model, whose parameters are maintained in a particle filter guided by the observable driving behavior. A Monte-Carlo simulation is then used to predict the future motion using the parameters of the IDM model provided by the particle filter. Equivalent to our approach, the prediction is performed for discrete steps in the future. Unlike our approach, but characterizing for sampling approaches, the resulting prediction is only implicitly represented by the different particles. For actual use, the particles must be approximated by some probability distributions, e.g. a (multivariate) Gaussian distribution. Our method directly works in a Gaussian space and requires less computational power.

Our approach was inspired by [58], proposing the use of four-dimensional Gaussian state space within a Kalman filter and control signals provided by path-following driver models. We extended this approach by the use of the CYRA motion model and unscented transformation (i.e. unscented Kalman filters) and additional lane changing driver models. Within the timeframe of AutoMate, [59] developed a very similar system, using a similar combination of multivariate Gaussian belief states, unscented transformation, CYRA motion, and driver models for simultaneous manoeuvre recognition and trajectory prediction. They however focused on turning behaviour at intersections instead of overtaking in rural road scenarios. For future development, it should be possible to combine both approaches to extend the set of scenarios and manoeuvres considered.

3.1.4.2.2 Pre-existing developments

Conceptualization, development, and implementation of the algorithm pipeline for the traffic prediction has been realized exclusively within the context of AutoMate. No part of the enabler has been inherited from previous projects nor addressed in any other European projects.

3.1.5 E4.1 Planning and execution of safe manoeuvre

In Automate it is intended to drive in structured environments such as rural roads or highways (see the Automate Demonstration scenarios). Therefore, in the following there are two popular state-of-the art trajectory planning algorithms presented.

The first one is an approach based on polynomial sampling [1]. Therein the center line of each lane which can e.g. be stored in a digital map has to be known. In the first step the ego vehicle is getting transformed from Cartesian coordinates (e.g. UTM) into the Frenét coordinates of the one center line dedicated to the lane the vehicle is supposed to drive on. In Frenét coordinates the vehicles position is described by the longitudinal distance from the beginning of the line and the lateral deviation from it. Each polynomial now describes the vehicles reference point’s (e.g. the gravity center) position along and lateral to the center line over time. Each longitudinal and lateral trajectory is described by a quintic polynomial. In order to be able to specify values for each coefficient, 6 conditions for each polynomial are required. The initial vehicle state already contains 3 of them (position, velocity acceleration). To be able to also obtain the remaining coefficients, terminal states for a specific terminal time are sampled as well (for further details have a look at [1]). Each longitudinal trajectory can then be combined with each lateral trajectory and the "best" one in terms of predefined costs that is also collisions free and kinematic feasible is selected to be forwarded to the vehicle controller.

Another approach is the one presented in [2]. Therein a driving corridor consisting of two-lane boundaries is used to mark the area in which the vehicle is supposed to stay in. Then a continuous optimization problem is stated to obtain an optimal solution that guides the vehicle central to the corridor by approaching the target speed. To make the drive more comfortable to the vehicle passengers, accelerations as well as the uncomfortable jerk (derivative of the acceleration) are getting penalized by using appropriate cost terms. To avoid collisions with other vehicles, the ego vehicle is approximated by circles and foreign vehicles by trapezoids. Subsequently according hard constraints are introduced to make sure the circles do not collide with these trapezoids. Furthermore, only trajectories that fulfill the kinematic constraints are considered as valid. The mightiness of this approach can e.g. be seen in [5]. Therein the equipment of the autonomous driving S-class "Bertha" is described. The applied trajectory planning concept is the one in [2]. Bertha completed the historic route of 103km from Mannheim to Pforzheim completely autonomously.

Within the Automate project a new trajectory planner based on the concept in [2] was developed. One major difference is that the cost functional was modified in a way to be able to explicitly consider information in concerns to social compliant behavior [3]. Therefore, reference trajectories which are calculated using appropriate driver models e.g. the "intelligent driver model" [4] are integrated in the first step. The nature of this reference trajectory allows it to incorporate this reference into the already mentioned cost functional of the optimization problem. Since the reference itself may not directly be forwarded to the vehicle controller, smoothing terms to penalize the acceleration and jerk are added as in [2]. The resulting solution aims to guide the vehicle along the road while approaching the target speed. The resulting behavior is social compliant in a way that e.g. necessary safety distances to other vehicles are held.

3.1.6 E4.2 Learning of intention from the driver

This section shall give a summary of the development and the final status of E4.2 "Learning of intention from driver" as previously described in the deliverables of WP3 [60, 61, 62, 63, 64, 65].

The Learning of intention from driver relies on the Driver Intention Recognition (DIR) model from WP2. The initial DIR model is trained offline with data from multiple different drivers and therefore represents the average driver. Enabler E4.2 personalizes the initial DIR model by adapting the model parameters during driving. For a warning-based system which tries do recognize driver intentions during manual driving this might reduce the number of false alarms for the individual driver. During automated driving, where the model could be utilized as a basis for manoeuvre decisions, it could lead to a more pleasant driving behaviour. In both cases a personalized model could increase the acceptance of and the trust in the system. Therefore, it is desirable that cooperative automated vehicles are able to adapt their automation strategies to the driver's preferences to guarantee a human expert-like driving behaviour.

The DIR model is based on (conditional) Dynamic Bayesian Networks (DBN). The nodes of DBNs can represent different types of probability distributions. The enabler E4.2 is able to personalize the initial DIR model for the individual driver by updating the parameters of probability distributions of the model while driving. The currently implemented online learning algorithms provide update methods for the parameters of discrete, multivariate Gaussian and Mixture of Gaussian distributions while the structure of the DBN stays unchanged. The update methods rely on Bayesian parameter learning and the usage of hyper-parameters which describe probability distributions over the model parameters. The hyper-parameters are updated as new evidence becomes available through observations, details were provided in [65], [63], and [61]. Since the update methods work in a supervised manner they require complete data samples. Thus, in order to be able to apply the algorithms during driving, an automated sample generation and labelling methods are necessary. The automated sample generation which was implemented for AutoMate relies on forward-backward inference, also known as smoothing, and employs the DIR to create labels for variable sequences of observed data points. Details on this process can be found in [65].

As described in [65] the smoothing based sample generation can be quite computational expensive depending on the complexity of the used model and the amount of data points, which have to be processed during the backward inference. In the worst case this leads to delays in the simulation environment or the dropping of data points. To avoid this, the enabler was extended by the option to perform the backward inference in a separate thread.

Additionally, a specific interpreter class was introduced that can be applied, for example, during the cooperative parts of the Peter scenario. In this case the driver has the opportunity to directly communicate the lane change intention via the HMI to the automation. Furthermore, the exact duration of the manoeuvre is known, since it is executed by the automation. Thus, by defining specific labelling rules the smoothing is not absolutely necessary for this case and the computational effort for can be reduced. However, implementing the specific interpreter requires some knowledge about the used DIR, e.g., names and values of the variables that shall be affected by the rules, while the smoothing based labelling requires usually no further knowledge of the internals of the DIR.

Since non-lane change data samples are predominant in the training data, as reported in [6], it can be expected that also during driving the amount of lane change sample is much lower than the amount of non-lane change samples. To reduce the imbalance in the samples and to somewhat increase the influence of the few samples that can be gathered for the individual driver during the experiments compared to the amount of data that was used to train the initial DIR model an oversampling functionality was implemented. The oversampling creates additional virtual samples for lane changes close to the actual samples during the experiments. This is achieved by multiplying the actual observed samples of the lane change manoeuvres with samples from a narrow Gaussian.

As mentioned in [65] this enabler was compiled into a C++ Dynamically Linked Library. For integration into the ULM simulator, this library is wrapped in a SiLab DPU. For the integration into the VED demonstrator the library is wrapped into a RTmaps package.

For the possibility to visualize the change from the initial DIR model to the current updated one during driving an additional stand-alone application was implemented. The application can receive the current model parameters from the aforementioned Silab DPU or RTmaps package via a socket connection and visualize every supported distribution available in the DIR model.

The distribution that shall be visualized can be selected via the “Distribution” cascading dropdown menu. The Application will draw the initial distribution together with the current distribution as shown in Figure 8. The current distribution graph is updated whenever a new message with updated model parameters is received via the aforementioned socket connection. The connection parameters can be configured via the “Options” menu. The communication is described in more detail in deliverable D5.7.

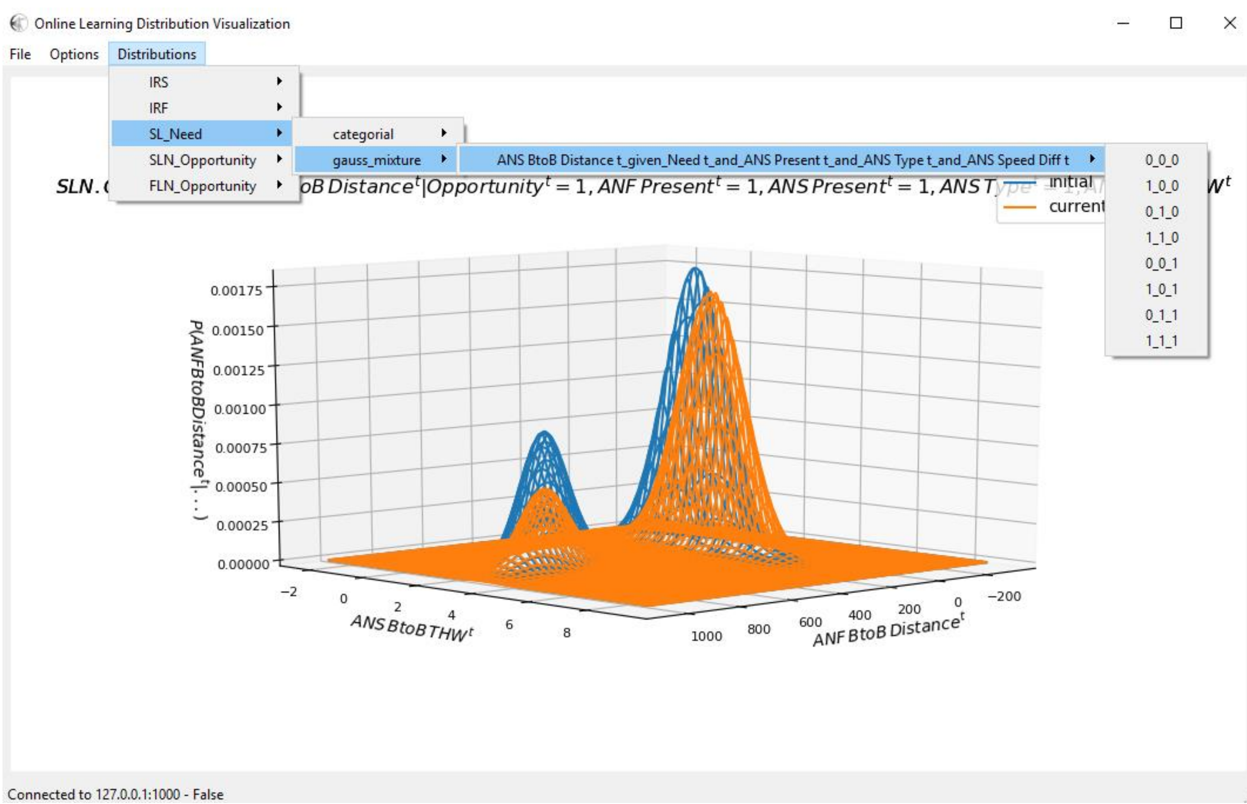


Figure 8: Distribution Visualization application for enabler E4.2

3.1.6.1 Comparison with the state of the art

Personalization of driver model with the application of online learning in the automotive domain is still some recent development. However, none of the approaches so far utilizes DBNs. In [66] the authors give an overview about some state-of-the-art approaches to the personalization of ADAS or driving style for automated vehicles. The approaches cover the following fields of personalization:

- ACC systems
- forward collision warning and brake assistance
- lane keeping
- cooperative assistance
- automated driving
- lane change

The personalization for ACC systems covers approaches where the driver is assigned to a certain driving style group and the ACC provides the appropriate control strategy, as well as approaches where the ACC attempts to mimic the driving style of the individual driver. The ACC approaches concentrate on gap preferences, acceleration profiles, and car following models.

The approaches for collision warnings as well as those for lane keeping provide warning thresholds for individual drivers.

Personalization for cooperative assistance mainly covers selective assistance functions or modalities dependent on direct requests or situations.

For the case of automated driving, the presented approaches either aim at learning individual driving styles for highway driving or general trajectory planning by imitating the driver, or their intention is to determine the individual driver prefers a defensive or a rather assertive driving style.

More related to the applications of AutoMate is the personalization for lane changes. In [66] only the work of [67] is presented. In this approach GMMs trained via EM are used to model lane-change and car following behaviour. In order to make the model responsive to individual drivers and behaviour changes the EM training is started again whenever a sufficient amount of new samples is available. Since the retraining consumes many resources the GMMs are retrained on a certain batch of recent data. In contrast to our approach the model only represents the recent driving behaviour and ignores older experiences.

A fuzzy Case-Based Reasoning and Situation-Operator modelling based approach to individualize and learn situation recognition for lane-changes is shown in [68]. The initially offline learned models are already individualized for a single driver and are then trained further online during a simulator experiment. However, the case base might grow over time leading to an increased time to check for known cases.

Another system for personalized lane change assistance is presented in [69]. In a highway scenario lane changes to the left and the right as well as lane keeping are modelled and predicted with HMMs. Starting form a general model, incremental batch learning for HMMs including several EM iterations on each new data batch is employed to implement a personalization for individual drivers. The approach should work while driving but the learning and evaluation is so far only performed with offline data. The automatic data labelling of this approach relies on the detection of an actual lane change and driver data to detect certain head movements of the driver. The author shows that the personalized models outperform the initial general model. In contrast to that, our approach does not require driver data.

A further topic for personalization and online learning in the automotive domain is the manoeuvre prediction at intersections. In [70] a manoeuvre forecast for other road users at intersections based on a Bernoulli-Gaussian Mixture Model is described. An update of the model is realized by means of sequential EM. In contrast to our approach, updating of the model while driving and an online sample generation are not covered.

Additionally in [71] the authors present an approach to individualize the prediction of stop, turn or straight manoeuvres at intersections for the current driver. Online Random Forest is used to learn from automatically labelled real driving data. This approach employs also an automatic data labelling but only for a fixed number of samples.

3.1.6.2 Pre-existing developments

As mentioned in [65], the development of this enabler for AutoMate could start with a pre-existing framework, consisting of libraries and algorithms for the creation and utilization of (Dynamic) Bayesian Networks. This framework was originally developed by OFF during the former EU project HoliDes[5]. For AutoMate many updates and extensions were implemented. With respect to E4.2 these are:

- the general ability to store model parameters in a way that they can be updated during runtime, e.g., as sufficient statistics

[5] www.holides.eu

- update methods for different distribution types used by the DIR (discrete, Gaussian, and Mixture of Gaussian)
- methods for online sample generation

3.1.7 E5.1 Online risk assessment

This section summarizes the development and final status of E5.1, "Online risk assessment" as previously described in the deliverables of WP3 [60, 61, 62, 63, 64, 65].

In the context of intelligent driving systems, the purpose of risk assessment is commonly associated with an early detection of situations that "may be dangerous for the driver, i.e. may result in harm or injury" [15]. This requires a concept to quantify and formalize the safety of the current and near-future traffic situation according to a metric of risk. The spatial and temporal region surrounding the TeamMate car in which there is no risk or acceptable levels of risk can intuitively be understood as *safety corridors*. The TeamMate car may occupy any point in the safety corridor without endangering the passenger or other vehicles. Once formalized in an appropriate form, safety corridors can be used by the TeamMate car to assess and plan safe and feasible trajectories, leading to a set of algorithms that allow identifying safe and reasonable arrangements of the driving process.

For AutoMate, the enabler E5.1 "Online risk assessment" has been developed to provide the TeamMate car with such safety corridors. Online risk assessment was divided into two independent parts that have been realized by different partners and shall be described in separate subsections: *online risk assessment with respect to dynamic objects,* like other traffic participants in the vicinity of the TeamMate car, and *online risk assessment with respect to static objects*, like obstacles and road boundaries.

3.1.7.1 Dynamic Objects

Online risk assessment for dynamic objects has been developed to formalize and quantify the safety of the current and near-future traffic situation according to a metric of risk into safety corridors.

As a metric of risk, we decided upon the *probability of collision*, i.e. the probability that the TeamMate car collides with another dynamic object. Following this idea, we developed a concept of safety corridors as geometric interpretations of the area in which the probability of the TeamMate car colliding with another object for a specific temporal interval is bounded by a user-defined threshold as a set of polygons.

Online risk assessment for dynamic objects requires knowledge about the probable current and future states of all dynamic objects observed in the vicinity of the TeamMate car, which we refer to as the prediction of the spatial and temporal evolution of the traffic scene (Figure 9a). In AutoMate, this prediction is provided in terms of probability density functions over the state of each dynamic object for future points in time by the traffic prediction (E3.1, c.f. Section 1.1.1.1). Given such a prediction, the predicted location and pose of vehicles at consecutive points in time are combined into polygons enclosing probable locations of vehicles for resulting temporal interval. Together the polygons implicitly define a safety corridor in which the TeamMate car may maneuver with a bounded risk of collision (Figure 9b). Once constructed, safety corridors can be used by the TeamMate car to plan safe trajectories, assess the safety of a trajectory planned by the automation, or assess the safety of a trajectory predicted for the human driver prior to its execution. The geometric interpretation of safety corridors allows for a quick assessment of potential trajectories as safe or critical, by checking whether the trajectory would force the TeamMate car to leave the safety corridor in a specific temporal interval (Figure 9c).

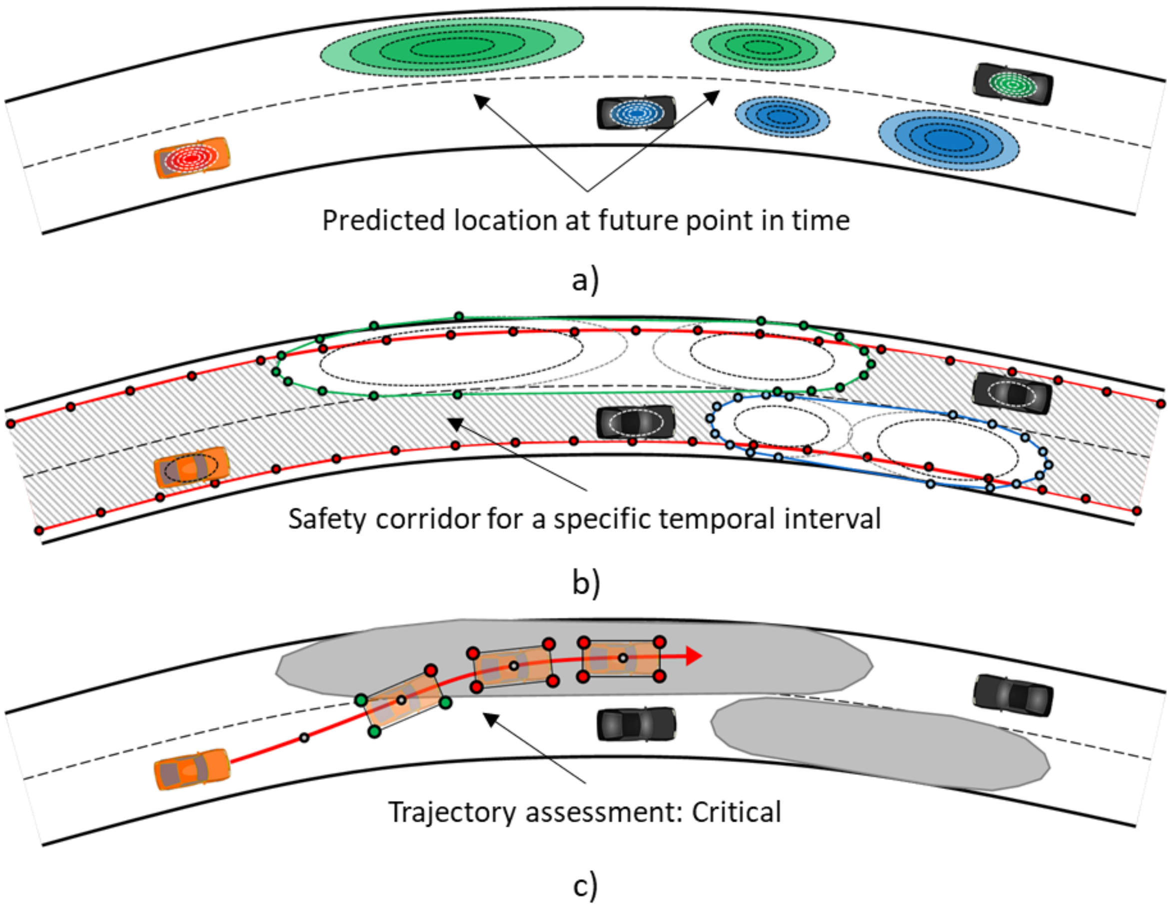


Figure 9: Visualization of safety corridors, a geometric interpretation of the area in which the probability of the TeamMate car colliding with another object for a specific temporal interval is bounded as and the use of safety corridors for trajectory assessment.

Evaluated on test data obtained in simulator studies throughout AutoMate, the final version of online risk assessment for dynamic objects achieves a correct rate of classification above 90% for prediction horizons up to 6 seconds [65].

The geometric interpretation of safety corridors as set of polygons allows the utilization of high-performance standard algorithms for trajectory assessment. Combined with the fact that the computational complexity for online risk assessment grows approx. linear with the number of objects considered, the achieved enabler promises a good scalability to more complex traffic scenarios [65].

Throughout the second and third cycle of AutoMate, online risk assessment for dynamic objects has been successfully integrated in the VED real vehicle and the ULM simulator demonstrator to help enabling our vision of the TeamMate concept. For this, online risk assessment in respect to other traffic participants has been implemented together with the functionality for the prediction of the spatial and temporal evolution of the traffic scene (E3.1), the driver intention recognition (E2.1), and online learning (E4.2) into a single C++ Dynamically Linked Library. The DLL was then embedded into functional plug-in modules for the simulation environment SILAB, used by the ULM simulator demonstrator, and the third-party software RTMaps, used by the VED real vehicle demonstrator, enabling the utilization of all functionalities in the corresponding demonstrators. The resulting VED real vehicle demonstrator has been demonstrated during the final event, first and final versions of the ULM simulator demonstrator have been evaluated at the end of the second [13] and third cycle (Section 4.1).

3.1.7.1.1 Comparison with the state of the art

As previously described in [65], approaches for risk assessment have been broadly classified into two categories [15]. The first category comprises approaches that relate risk to unexpected behaviour of traffic participants, i.e. the risk of a situation is assumed to be proportional to its unusualness. [72]

represented the nominal behaviour of driver by Gaussian mixture models which could then be used to detect "unusual" situations by assessing the likelihood of a driver's behaviour. [73] proposed to compare expectations about a driver's behaviour with estimated intentions, which allows the computation of the probability of a mismatch between expectation and intentions to indicate risk. Unfortunately, such approaches only allow to assess whether a situation is critical, but provide no additional information concerning the exact circumstances.

The second and more common category comprises approaches that relate risk with potential physical collisions between entities (e.g., vehicles) in the traffic scene [15]. Usually, these approaches combine both a prediction of future trajectories for all entities in the traffic scene and the assessment of these trajectories to detect potential collisions [15].

Many of such approaches in the literature focus on "Time-To-X" measures, e.g., the "Time-To-Collision", representing the remaining time to a collision under the assumption of constant velocities, or "Time-To-React" measures, representing e.g., the remaining time to initiate a braking or steering manoeuvre, which can be used as an indication of what action should be taken or to identify the least dangerous intervention manoeuvre [15].

For assessing whether a future trajectory of the driver or the automation is safe, the most popular metric of risk is based on the notion of the *probability of collision* [47, 74, 52, 75], based on the predicted trajectories of the driver and other traffic participants, which we adopted for online risk assessment in AutoMate. Assessing the probability of collision under uncertainty requires the integration over all possible trajectories and dimensions of all traffic participants [47, 74, 52, 75]. Due to the unsolvable nature of this integration in closed form, one must usually resort to Monte Carlo methods, limiting the real-time capacity of such approaches. Unfortunately, actual computation times are seldom reported. [47] limited the prediction horizon to a maximum of three seconds to achieve real-time capacity. [52] considered predictions up to four seconds by limiting the number of samples to a very low number of just 100. [75] proposed a novel approach, suitable if all vehicles are perfectly aligned with the road, that enables online risk assessment via testing for collisions of pairs of trajectories with average computation times of approx. just 0.007ms for each tested pair. Unfortunately, when dealing with uncertainties, one must once again resort to Monte Carlo methods, cancelling the computational advantages.

Although based on the same metric of risk, the probability of collision, our approach for online risk assessment for dynamic objects in AutoMate differs from these approaches by transforming the prediction of the temporal and spatial evolution of the traffic scene into polygonal safety corridors over time spans, removing the need for integration. The transformation comes with the caveat of an inability to provide an "exact" probability of collision for a given trajectory. Our safety corridors only provide the upper bound on the probability of collision, allowing the assessment whether a trajectory is safe in relation to a desired probability of collision. We argue however that an "exact" assessment is unnecessary, if it finally used to test it against a threshold, in which case this inability is of no effect. When testing against a threshold is sufficient, our approach allows the assessment of trajectories without the need for Monte Carlo methods, allowing for a much greater prediction horizon (e.g., 10s) and higher traffic densities than the state of the art.

Recently, the Intel company Mobileye[6] began to promote a new concept called responsibility-sensitive safety (RSS) in the hope to establish a new standard in online risk assessment. RSS is a mathematical model for the formalization of safe driving for autonomous vehicles based on a limited set of five "common sense" rules [54]. E.g., the rule "do not hit someone from behind" is translated into a simple mathematical formula to calculate the required distance to a lead vehicle necessary to avoid a crash, should said lead vehicle initiate an emergency braking, given user-defined assumption concerning the lead vehicle behaviour, like the maximal deceleration. Combining similar formulas derived for the other rules, RSS implicitly defines a similar kind of safety corridor to constraint the possible control actions of the autonomous vehicle and ensure a safe and collision-free behaviour. RSS promises a great coverage of potential scenarios while being simple enough for real-time utilization. However, as of now, RSS relies on very naïve models for traffic prediction and abstracts from sensor uncertainty by requiring that the vehicle's sensors operate with errors small enough to treat measurements as ground truth. Given these assumptions, RSS may be inappropriate when dealing with manoeuvres that require longer prediction horizons, like e.g., overtaking manoeuvres. We belief that our approach for online risk assessment is for the most part consistent with RSS and could be used as a first proposal to solve these limitations of RSS:

3.1.7.1.2 Pre-existing developments

Conceptualization, development, and implementation of the algorithm pipeline for online risk assessment with regard to dynamic objects has been realized

[6] https://www.mobileye.com/

exclusively within the context of AutoMate. No part of the enabler has been inherited from previous projects nor addressed in any other European projects.

3.1.7.2 Static Objects

A novel approach is presented within the scope of the project to assess the safeness of the extracted corridor from road boundaries and ego-pose information [60, 62, 64]. The same risk metric "Probability Of Collision (POC)", as in online risk assessment for dynamic object is used here to determine and quantify the bounds of the safety corridor in the uncertain environment. Lane boundary information in the digital map and the uncertainty of the ego-pose are used to define the corridor of the required spatial horizon. The POC, is used to quantify the influence of the uncertainty coefficient in the localization measurement in assessing safety corridor information for required bounds. The combination of the online risk assessment for dynamic objects and static objects provides assessed corridor information within certain bounds of the risk metric over the required spatial and temporal horizon. Furthermore, the risk of planned trajectories from the motion planning are quantified using a traffic safety surrogate measure like "Time-To-X" [64, 65].

In Figure 9b, the red lane illustrates the spatial geometry of the assed corridor using the boundary information from the digital maps. The spatial geometry is extracted as the set of points which defines the polygon. During the 3rd cycle, the component was evaluated and verified on the test data obtained from Vedecom's real test car [65]. Evaluation results show the algorithms sensitivity towards uncertainty in the localization measurements. Nevertheless, on Vedecom's test data we achieved 100% accuracy as the localization error was negligible. Furthermore, for the complexity test and to emphasize the importance of the localization measurement, a similar evaluation during the early stage of 3rd cycle was performed on one of internal test data that involved complex road networks and synthesized uncertainty in measurements. The results confirm that the module has an almost constant run-time for extracting and assessing the corridor for risk from digital maps with an increase in the complexity of the scene [65]. Our component was seamlessly integrated in Vedecom's real demo vehicle, and the capability of the component was demonstrated during the final demo event at Satory. The component was shipped as C++ Dynamic Linking Library (DLL). Callback functions from our DLLs were exported to the RTmaps middleware used by the Vedecom's system environment [65].

3.1.7.2.1 Comparison with the state of the art

In general as in [15], risk assessment are broadly categorised with methods that take into account the risk as a collision between physical scene entities, and approaches defining the risk associated to unexpected behaviour of traffic agents. In our approach, we define safety corridor by considering risk as the collision between the TeamMate vehicle and road boundaries in the scene [60].

An illustrative and simple form of safe corridor is obtained using the concept of occupancy grid map [76] to represent the planar environment around ego car. Similarly [77] [78] [79] propose a different representation of maps which differ in memory and computational requirements and their potential use for path planning. In the literature we come across many metrics for risk assessment, our proposed approach considers probability of collision [52] as a metric for defining upper bound to the represented safety corridor.

In many literatures [76] [52] [79] [78] complex representation of the road map are proposed. However, our approach uses a simple modelling technique, which yet at the same time is robust enough to represent the required safety corridor. Many of those proposed approaches in the literature have a proportional complexity with the scene, our validation and verification results show that our component is easily scalable for complex situations and has an almost constant run time. Furthermore, the represented safety corridor is used to assess the planned trajectory using the Time-To-Collision metric [65]; it could also be easily extended to obtain the severity of collision based on the velocity of the TeamMate vehicle.

3.1.8 E6.1 Interaction modality

Based on the literature defining the requirements of a successful cooperation and our empirical studies within the project, the following design implications are proposed to help design interaction concepts for highly automated cooperative driving.

3.1.8.1 Choice of Interaction Modality

Already existing control units should be used for the interaction [29]. That means that the steering wheel or the gas and breaking pedals can be used in situations where they are also used while driving manually. As an example, the steering wheel is dominantly used in overtaking maneuvers. Therefore, drivers know how to interact with these control units automatically without the need for focusing on the interaction.

3.1.8.2 Definition of Interaction Intention

Using control units as input interaction device in automated mode as a mean to provide support to the automation requires that the automation system is able to distinguish between three possible types of inputs. Three different types of input intentions need to be distinguished:

1. Unintentional – the driver's intention was not to initiate the maneuver it was only an accidental interaction with the interaction device.
2. Intention to provide input to the automation – the driver wants to fulfill the assigned task as a team partner and interacts with the system.
3. Intention to take back control – the driver wants to take back full control and deactivates the system.

The driver intention recognition can additionally be linked to rigorous preconditions like the visual focus [45] to assure drivers intention.

3.1.8.3 Feedback

The debriefing interviews of the user study implicated that users want to get haptic feedback while turning the steering wheel. The interaction should be designed considering the ideas of the H-Metaphor [13]. The driver should experience haptic feedback – other feedback modalities like auditory or visual feedback can additionally be included. An excellent example of haptic feedback in today´s cars is the kickdown [26] function, where an automatic transmission car is downshifting for better acceleration. The driver feels a force feedback while pushing the car’s gas paddle to the floor.

3.1.8.4 Provided Information

Unnecessary information should not be provided. The driver should only see the task related necessary information. An example is that the RPM Counter is redundant in automatic transmission – this only distracts the drivers and is not useful to operate the car. Therefore, the driving task must be modeled, and only information necessary to accomplish the goal must be provided to lower the complexity of the drivers´visual search task. As Musk [1] pointed out “The more automated a car is, the less dash info you need. How often do you look at the instrument panel when you drive in a taxi?”.

3.1.8.5 Driver Monitoring

In a cooperative interaction mode, the system should also be aware of the drivers' state and their future actions [56]. Therefore, a continuous driver monitoring is crucial to adapt the interaction appropriately. The driver monitoring should assess the involvement of the driver in the driving task based on values like fatigue [42] or distraction.

3.1.8.6 Adaptivity

The definition of the thresholds (see Implication 2) should adapt according to the driver's attention (see Implication 5). An automation input interfering with the automation's planned action should be associated with a higher effort if the driver is distracted than if the driver is attentive. This is necessary to rule out an unintentional interaction or an uninformed input and to enhance the immersion into the driving task.

3.1.8.7 Time to interact

Designing an interaction, the right time when to notify the user of a potential cooperation is crucial [18]. If it comes to time-critical situations the time to interact with the system should be reasonable and as fast as possible. The interface design should strive for the lowest interaction time possible. As an adaption to Fitt´s law [28], one can say that the distance and size of existing in-car control elements are appropriate, and in comparison to, for example, a touch display more suitable for the interaction.

3.1.8.8 Task Engagement

The design of the interface should achieve a high task engagement [31, 32]. Walch et al. [54] showed that even after a maneuver initiation drivers tend to be engaged in the task and show safety assurance behavior. A cooperative system should strive for a high task engagement during the cooperation sequence even after the cooperative part is finished but the maneuver is still being executed. This is necessary to give the driver time to react in an unforeseen situation (e.g., if there is unexpected oncoming traffic).

3.1.8.9 Design for calibrated trust

It is essential that the drivers do not “overtrust” the system [39, 53, 25] which would lead in a hazardous situation. Therefore, an accurate impression of the vehicles current abilities and responsibilities must be instilled by the driver [19]. The system should clearly state its limitations and confidence. The interface must be designed to distribute the responsibilities between the two cooperative partners in a transparent way.

3.1.8.10 Habituated Interaction Concept

As shown above, the cooperative action should be designed using task-related control units. These control units should be used during the manual operation of the vehicle. The physical movement performed during the manual operation can activate the according habituated action schemes and, therefore, trigger specific behavioral patterns [20]. If possible the physical movement while using an interaction concept should imitate iconic physical movements which are performed during the according manually operated task.

3.1.9 E6.2 TeamMate multimodal HMI

3.1.9.1 Benchmark of Human Machine Interface for highly automated vehicles

The automation level reached by the cars currently on the market is between 2 and 3, meaning that they can drive autonomously in certain given conditions. To achieve that, they are made of the combination of different ADAS. Currently, due to technical, legal and behavioural constraints, vehicles are not able to fully replace the driver in performing each driving task.

In most advanced cases, the car systems are able to replace the driver in controlling the vehicle under certain conditions to perform certain maneuvers or in some parts of the road, but in any case, the driver is always asked to supervise the driving operations. The autonomy currently available is a combination of longitudinal control systems such as Adaptive Cruise Control, lateral control systems such as lane Centering and automatic braking functionalities, deployed with different technical approaches. The HMI, in this context, should be able inform the driver about the level of support provided by the automation.

The main vehicles provided with automated driving features that will be considered in this analysis (also for the HMI design solutions adopted) ate the Tesla model S and the Audi Q8.

The Tesla features a 17-inch digital cluster, which allows access to all vehicle functions (see Figure 10).

Figure 10: Tesla Model S exteriors and instrument cluster

When the vehicle is in manual driving mode (see Figure 11), on the cluster are generally reported the same information that we find on all electric vehicles (i.e. the speed indicator, battery consumption, date and time and if any, the

Figure 11: Cluster Tesla model S in manual drive mode

navigator).

When the vehicle is in “Autopilot” mode (despite the name, this system cannot be considered as an autonomous driving feature), on the cluster are displayed different information regarding the different systems that are intervening, such as the adaptive cruise control, the Lane Change, Blind spot detection and the Lane keeping.

Figure 12: Cluster Tesla model S in autonomous drive mode

In Figure 12 it is showed that the relevant information is the position of the Ego-vehicle inside the lane and the detection of the other vehicles. The blue color is commonly used to induce autonomous driving. The presence of the blue steering wheel indicates that the vehicle is proceeding autonomously.

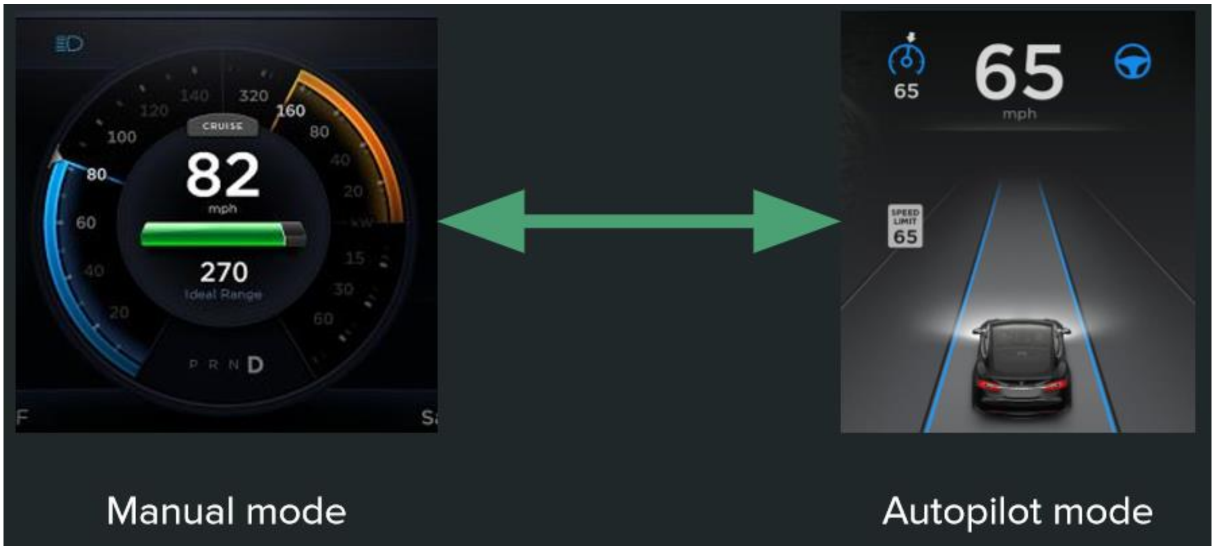


Figure 13: Cluster Tesla model S in manual drive mode vs autonomous drive mode

In Figure 13 the direct comparison of information in the two modes is reported: the focus is shifted to the set of ADAS systems that allow autopilot mode. In this case the systems that provide an indication by an icon are the ACC and the speed limiter, while lane centering signals the activation of the lane demarcation lines using the blue color. The vision of the automation designed in Tesla HMI, understands the automation not as a whole but as the combination of multiple systems that dialogue with each other: if one of the different systems stops working, the automation is no longer ensured.

In the take-over request a small pop-up appears at the bottom (see Figure 14), with a very long label written in small characters, with a very low legibility impact.

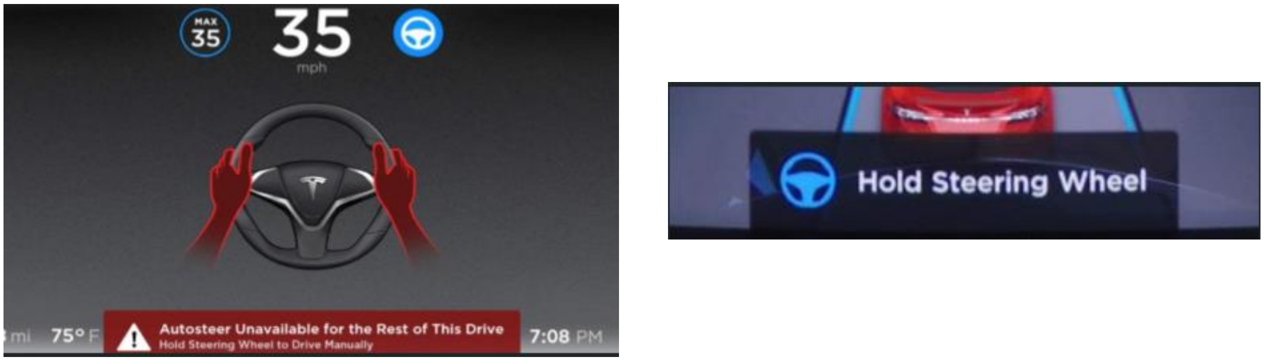


Figure 14: Pop-up take over request

Another example of autonomous driving is provided by the Audi Q8 (see figure 15). It has a digital cluster, a superior MMI touch response display and a secondary display (see Figure 16).

Figure 15: Audi Q8 exteriors and interiors

When the vehicle is in manual driving mode, on the instrument cluster are reported the same information that we can find on all electric vehicles (e.g. the speed indicator, date and time and if any, fuel level indicator).

Figure 16: Audi Q8 Cluster

Relevant information about the position of the Ego-veicle inside the lane and the detection of the other vehicles are reported in automated mode. In Tesla

Figure 17: Audi Q8 in automated drive mode (Traffic jam pilot)

this configuration is called autopilot, while Audi calls it the Traffic jam pilot (see Figure 17). The Traffic Jam Assist helps drivers get more relaxed at their destination, even in heavy traffic or traffic jams. As a partially automated comfort function, the system takes over the longitudinal and lateral guidance of the vehicle. This means that the car can drive off, accelerate and brake automatically, as well as steer the vehicle within certain constraints. The driver has permanently supervised the system and is ready to take control over the vehicle at any time. Green is the color chosen by Audi to indicate autonomous driving.

In Figure 18 is displayed the comparison among the different modes.

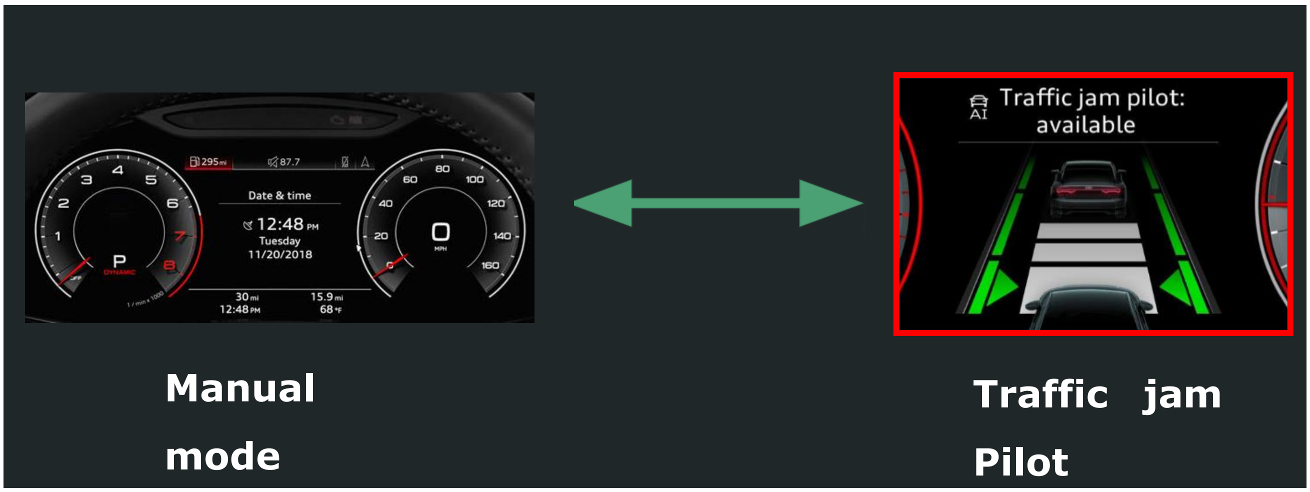


Figure 18: Audi Q8 in manual drive mode vs autonomous drive mode

Tesla and Audi, even if with different graphical layouts, share the same concept and high-level application. Both conceive automated features as the set of multiple systems allowing separated control functions. Figure 19 shows a comparison between the two interfaces in automated mode, both based on bird-eye view reconstruction.

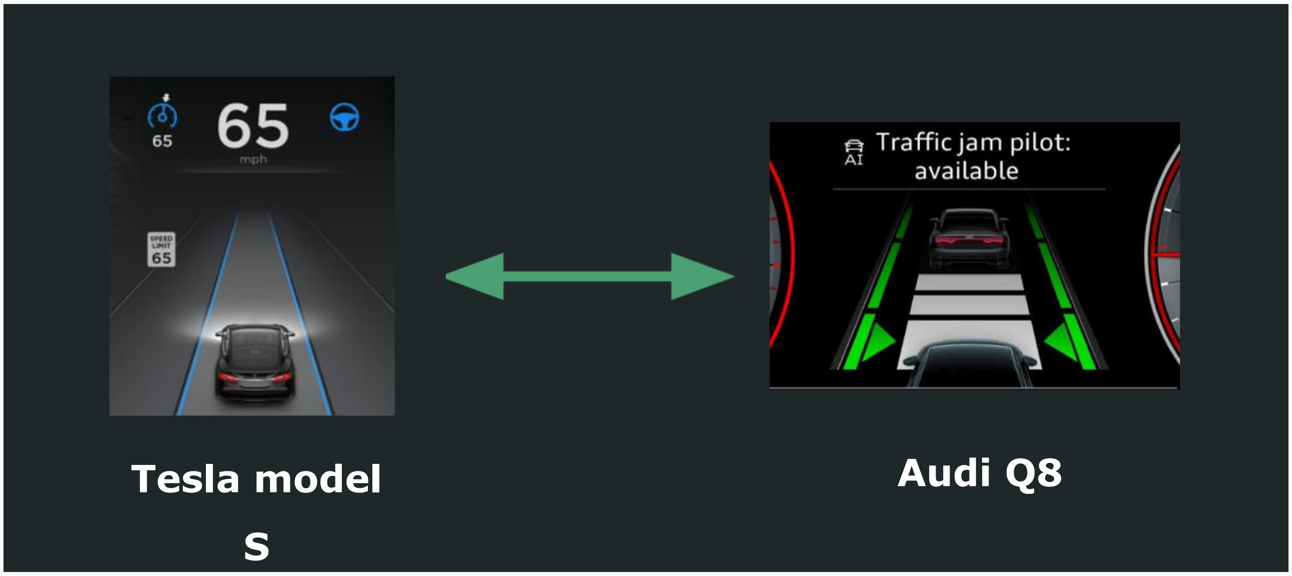


Figure 19: Tesla model S vs Audi Q8 in autonomous drive mode

Besides vehicles already on the market, also research projects are working on exploring the most effective interaction modalities with highly automated vehicles. One of these projects is AdaptIVe-IP: over 42 months (ended in June 2017), 28 partners from all over Europe collaborated in this largescale project to advance the performance of automated driving systems for cars and trucks. Taking automation to higher levels, AdaptIVe's results support the goals of making driving safer and more comfortable, and of reducing congestion and fuel consumption.

The results collected in the experiments showed that the driver manifested a preference for a two-mode interface (see Figure 20, with a reduced information load).

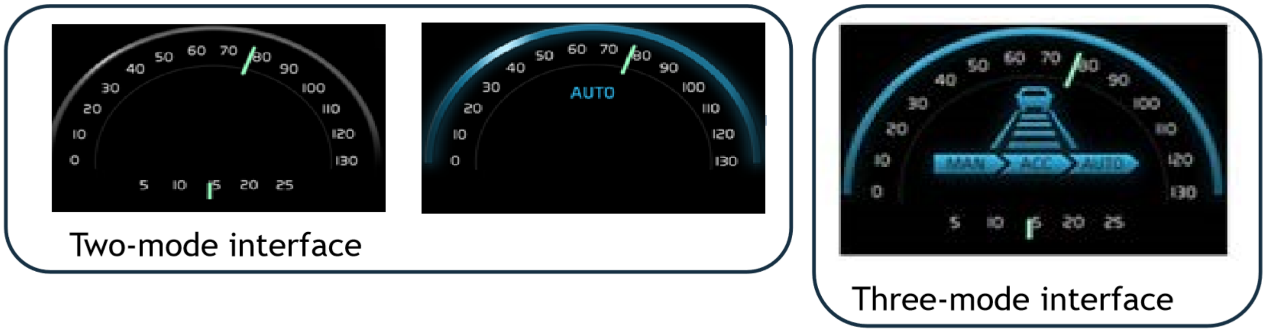


Figure 20: AdaptIVe Interface mode

The experiments also showed how the announcements about the situation of the system increases the awareness of approaching the limits of the system, this helps to avoid inconvenient transitions (see Figure 21). The automation should inform in advance about an upcoming automation or vehicle limit, so that the driver will be able to safely take-over the driving task.

Figure 21: AdaptIVe Pop-up alert

In general, it has been noted the relevance to promptly inform the driver before facing a situational change using visual and auditory feedback.

Peripheral vision is very effective in bringing the driver's attention back to supervision. The indication of the detected vehicles helps drivers to anticipate future automation maneuvers. The indication of the detected vehicles supports the drivers in anticipating the failure of the automation.

3.1.9.2 Final updates of E6.2

The TeamMate multimodal HMI developed in AutoMate and integrated in 3 simulators and 3 vehicles demonstrators, has been designed from scratch with a process tailored around users' tasks: the possible driver states, behaviors and interactions with the automation have been modelled into an HMI strategy and then deployed into a multimodal interaction tool, made of software (the visual, audio and distributed HMI) and hardware (the haptic seat).

The detailed description of the design process, the iterative implementation made with incremental prototypes, the final layouts and the V&V results are described in WP4 deliverables. This paragraph will focus on describing the final updates in the HMI development.

The activities performed in the last project cycle in order to update the enabler "TeamMate multimodal HMI" had the following aims:

- To focus the development on the last topic of interest, i.e. the distributed HMI on mobile device;
- To tailor the releases according to the final functional and non-functional requirements from the demonstrator owners (especially real vehicles), also checking the consistency of the elements in the HMI with all the elements needed for the use case;

- To perform the final integration and functional tests (data communication tests, deployment on target control unit and screens, data stream latency) on vehicle demonstrators, in order to use the enabler as part of the final evaluation.

In order to achieve these objectives, iterative tests have been performed in close collaboration with demonstrator owners. Several incremental releases, including translations and bug fixing, have been realized before the final evaluation phase. Moreover, several adjustments have been made on the existing software, thanks to the dry run tests performed on tracks and on the road by the demo owners. In particular, possible back-up screens (e.g. related to unexpected while possible car behaviors) and adjustments according to the vehicle signal map have been included in the last releases.

3.1.9.3 Summary of improvements in comparison with the SoA

The Human Machine Interaction concept developed in AutoMate represents a novel and innovative interaction design solution to address the cooperation between the human and the automation agents.

In particular, the concept of increasing the transparency of the automation, by explaining to the driver the reason that leaded to the request of support, is becoming in recent years a relevant topic both for applied research and automotive industrial domain.

In this context, it is important to highlight that the reference domain is experiencing an extremely rapid evolution, deriving from (i) the ever-greater investments done by car makers in on-board information technologies, (ii) the pervasive digitalization of vehicle interiors and (iii) the evolution of automotive industry, with the raise of new players (brand new companies - such as Tesla, Faraday Future and so on - and companies coming from other domains, such as Dyson).

In this framework, the “flash forward evolution” carried out in scientific community in the 3-years duration of the project influenced the design process in AutoMate. For example, the “distributed HMI” on mobile app (part of the multimodality of this enabler) was not expected at the beginning of the project, while it became a relevant topic in contemporary research. At the same time, as emerged in recent studies (e.g. experiments performed by ULM partner inside the project), some interaction modes such as speech interaction (even if very effective for other needs) have not been considered as the most appropriate interaction modality for automation-related features.

The TeamMate multimodal HMI developed in AutoMate showed a significant improvement against the state of the art, as emerged from the V&V cycles performed and described in WP4 deliverables (in particular D4.2, D4.4 and D4.6) and scientific publications in relevant international conferences (e.g. Automotive User Interfaces, ITS World Congress – the full list is reported in D7.7). In particular, the HMI designed in this project showed to provide a relevant contribution in increasing the trust in automation and to contribute in balancing the workload. It also showed, in comparison with the baseline, a significant role in improving the drivers’ awareness and to be in general perceived as useful, even if some users suggested that some training would be needed to improve the interaction. This was confirmed also, for example, by the evaluation of the overall TeamMate system performed on real vehicles, e.g. the CRF car, in which the users performed several roundabouts interacting with the HMI and showed constant improvements during the test session, with significant differences between the first roundabouts and the last ones (N.B.: CRF evaluation scenario includes more than 20 roundabouts).

Moreover, the concept implemented in project, strongly characterized by the cooperative approach, raised new possible research questions to be asked. Among then it is important to mention, as possible future steps:

- Further research on "Take Over Anxiety", a novel notion deriving from "Range Anxiety" experienced in electric vehicles; this concept, consisting in possible anxiety deriving from the end of a comfort area (i.e. the battery range for the electric vehicles, the area in which the car is confident to handle a high automation level for automated vehicles), has been addressed in AutoMate by providing the driver with an estimation, based on maps and CAN data, of the time to the next expected take-over request. Currently, this option is under investigation also in other research projects.
- The creation of an ontology of reasons that may lead to a take-over request, in order to map the main causes of transfer of control from the automation to the driver. In AutoMate, it has been done by including 3D videos in the HMI to explain the behavior of the artificial agents; in further research, this may include a structured ontological analysis able to cover several limits of the automation and to facilitate the creation of challenging use cases for automated vehicles.
- Further research on state-adaptive distributed HMI, in order to increase the effectiveness of the interaction through the combination of driver monitoring, connectivity technologies and interaction design.

3.1.10 E6.3 Augmented reality

This section summarizes the development and final status of E6.3, “Augmented Reality HMI” as previously described in the deliverables of WP4.

3.1.10.1 Comparation with the state of the art

Facilitated by technological advancement, Augmented reality (AR) for automotive is gaining impact on the driving experience by overlaying virtual 3D information with the real driving scene. It is favorably realized using Head-Up displays (HUD), which project a virtual image on the windshield. The decreased distance between the driver’s line of sight and the information displayed in a HUD exerts a positive effect on driving performance and mental workload, especially for time critical and dynamic information [80] [81]. In AutoMate we developed an Augmented Reality HMI (AR-HMI) to increase the level of awareness by displaying relevant information in the field of view of the driver.

Longer glances at the AR display are possible without negatively affecting control of the vehicle [82]. In AutoMate we minimized the amount of information displayed on the AR-HMI during manual driving to not divert the attention of the driver with unnecessary information.

The integration of Augmented Reality in a HUD offers a wide range of important information that, otherwise, is not readily available to the driver, such as the stopping path of the vehicle or the indication of visually concealed hazards [83]. Thus, situation awareness is further improved even in comparison to classical HUD [84] [85]. In AutoMate, the AR-HMI is connected to other enablers in order to provide the relevant information to the driver. These enablers are E2.1 “Driver intention recognition” by displaying a path (corridor) that the vehicle will follow according to the intention of the driver and E5.1 "Online risk assessment" by displaying the location of the risk and changing the color of the path depending of the risk of the maneuver (green for a safe maneuver red for an unsafe maneuver and blue for a non-assessed risk maneuver). Also, the AR-HMI displays the location of potential hazards, such us incoming and oncoming traffic, work construction, etc.

AR concepts are realized and evaluated in current simulator studies [86]; challenges for the implementation in real vehicles such as accounting for the driver's head position and movements are being investigated [87]. In AutoMate we made several studies in simulator and real vehicle, testing the level of awareness of the driver related with the vehicle's behavior during autonomous mode.

3.1.10.2 Implementation on real vehicle.

For the demonstration and evaluation on the 3rd cycle of the project, we decided to develop an augmented reality application with the purpose of simulating a head-up/windshield display. The main function of the application is to display an image that shows the behavior of the vehicle while driving in automated mode.

The development of this application was made for the Epson Moverio BT-200, which are semi-transparent smart glasses (see Figure 22). The Moverio BT-200 runs the OS Android Version 4.0.4 (API 15). Due to that, the development process was implemented with the Android Studio IDE and the programming language Java. The structure of the application is split mainly in two parts.

Figure 22: Epson Moveiro BT-200

On the one hand, there is the connection between the AR-glasses (client) and the car (server). This was implemented through a network socket. The client connects to the server and receives the data that are generated by the server. The data is sent in the protocol buffers format and represents the behaviour of the car. To configure the connection, the application provides a form for IP address and port input.

On the other hand, the application has the task to interpret the incoming data and transform it to display the proper image on the display. This is implemented through mapping the data to the image resources.

The following UML sequence diagram (see Figure 23) shows the interaction between components of the application. In the first step, the application triggers the connection between the client and the server. The IP address and port of the server can be configured in a form after launching the application. In step 1.1 the client connects to the server. If the connection has been successful, the server sends data to the client component (1.3). The client passes the data to the FullscreenActivity class of the application (1.3.1), which sets the image based on the data (1.3.1.1). Step 1.3 and its sub steps proceed in a loop, while the client and the server are connected.

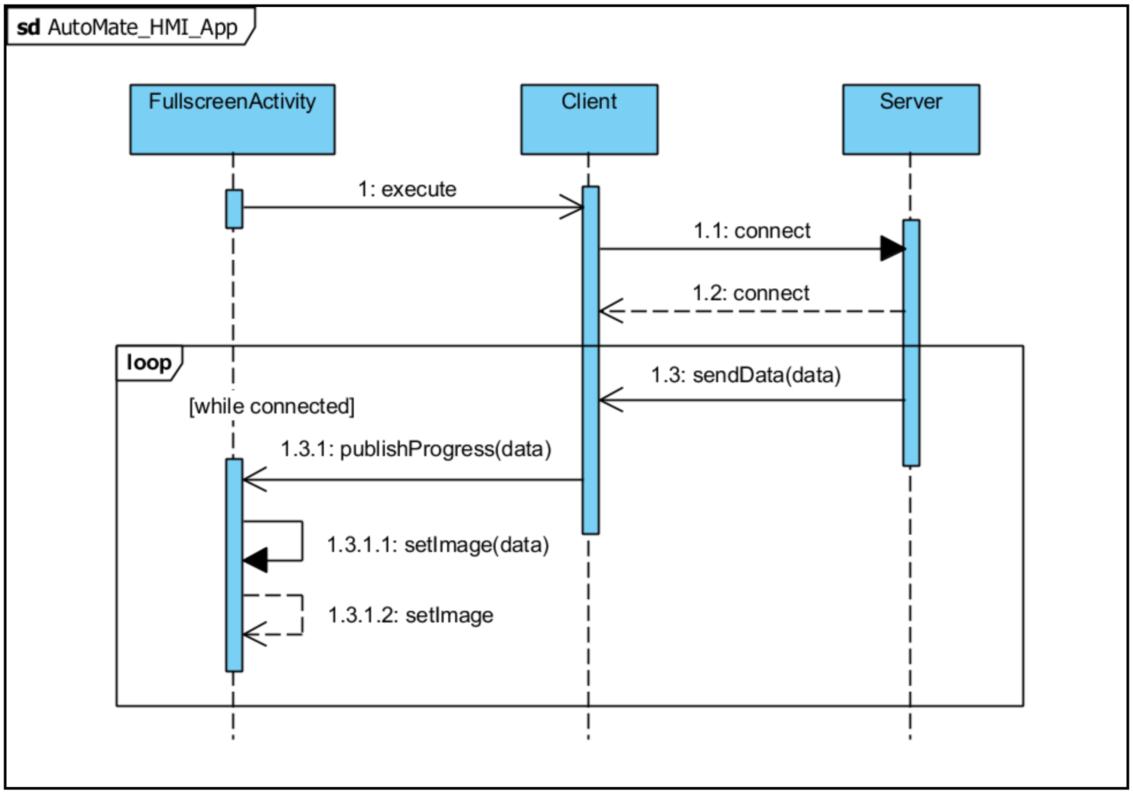


Figure 23: UML sequence diagram

Due to the limitation of the AR-glasses, the graphics in form of icons were designed to represent the different possible scenarios to ensure understanding of the vehicle's behavior (see Figure 24).

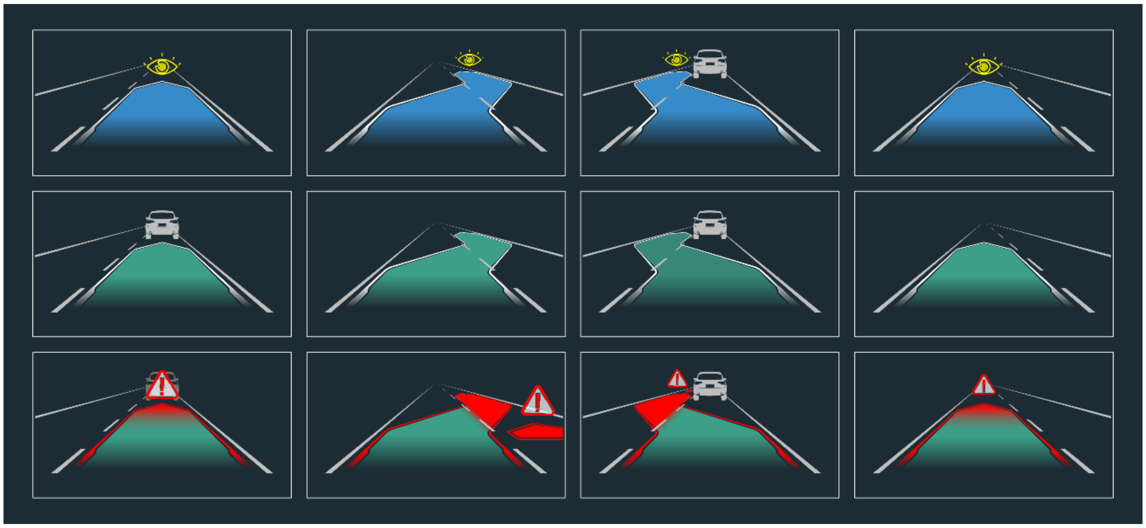


Figure 24: Graphics implemented in the Epson Glasses.

The graphics display the representation of the maneuver that the vehicle will performance (see Figure 25). This graphics don't match with reality as in the original concept, due to the limitation of the device and the movement of the head during driving, but it proves the concept of the information shown in the driver's field of view.

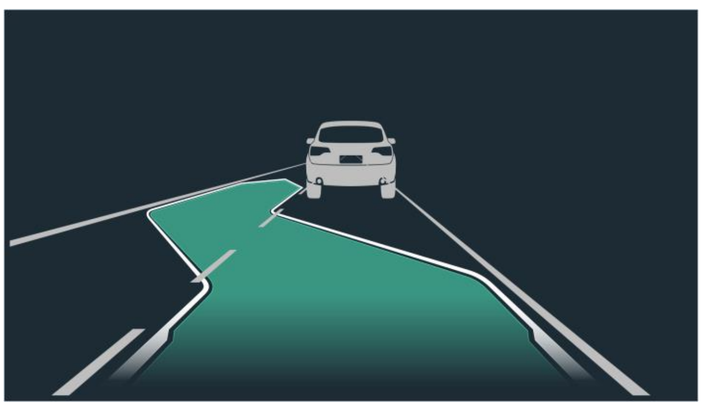


Figure 25: example of the graphics.

The image displayed on the AR-glasses only uses a small percentage of the driver's field of view, without obstructing completely the view on the road (see Figure 26).

Figure 26: image displayed through the AR device.

To create the graphics, 12 case scenarios were studied and implemented in the AR-glasses to cover several maneuvers during an overtaking situation. In the first 4 scenarios, the vehicle cannot assess any risk, from scenario 5 to 8, the vehicle can assess that it is a safe maneuver and from scenario 9 to 12, the vehicle predicts an unsafe maneuver (see Figure 27).

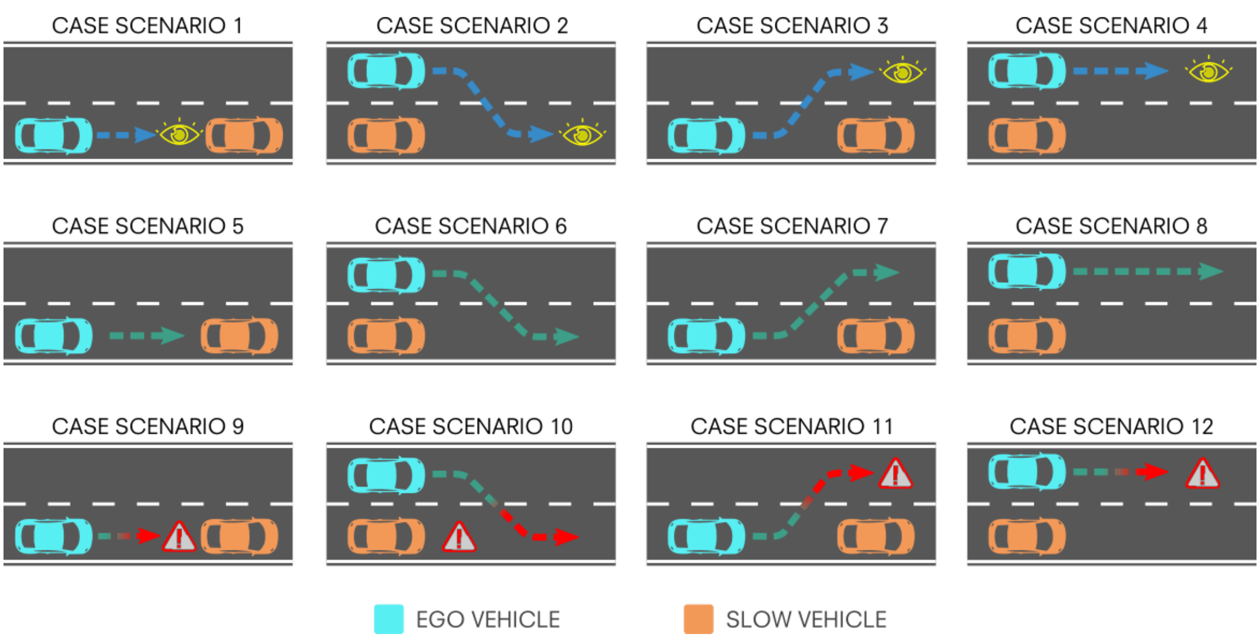


Figure 27: Overtaking case scenarios implemented in the AR-glasses

Evaluation studies were carried out during the demonstration of the AutoMate concept in Satory on June 2019. There, the AR-glasses were used to demonstrate the augmented reality's concept with good acceptance by the participants of the demonstration (see Figure 28 and Figure 29). The communication between the glasses and the vehicle worked very well, showing the graphics to the participants and improving the understanding of the vehicle's behavior. The vehicle where it was implemented was the VEDECOM'S real demonstrator.

Figure 28: use of the AR-glasses during the showcase in Satory.

Figure 29: graphical representation of the AR-glasses during the showcase in Satory.

The results of the studies are described in the chapter 4 of this document.

3.1.10.3 Implementation on Vedecom and ReLab simulators

In deliverable D4.6 (4.3.3.1), it was mentioned that the AR-HMI was planned to be implemented in VEDECOM and ReLab simulators. Many efforts were made in order to implement it, design concepts and experimentation in the Unity software were developed (see Figure 30), but due to the limitations of the SDK of the software used in both VEDECOM and ReLab (SCANeR™ studio), the implementation was not successful.

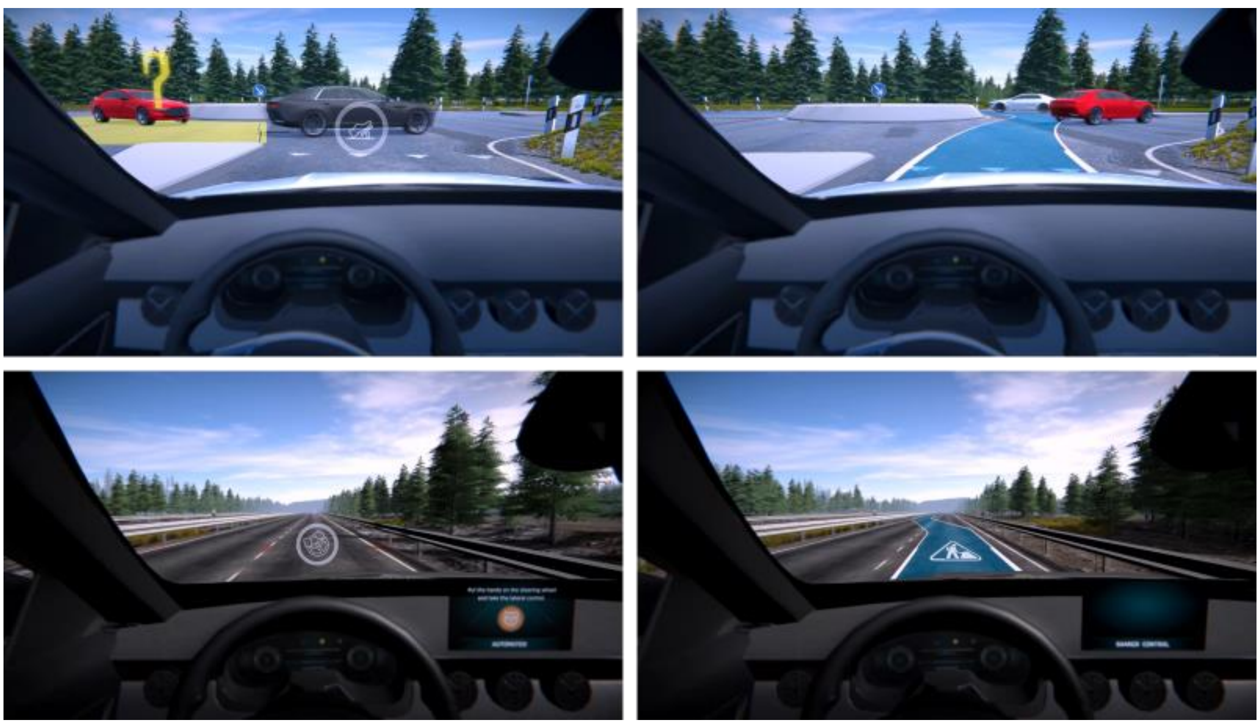


Figure 30: Eva and Martha's AR-HMI concept in Unity software.

3.2 Additional Empirical Studies for Enabler Updates

3.2.1 Input from Driver task modelling and the validation experiment to the PETER scenario evaluation study

At DLR, a study with 12 participants (6 females, 6 males) was conducted to obtain data modelling how users of the TeamMate car interact with the car while performing an occupation as it might occur during an automated drive (see Figure 31). Preliminary results from this experiment were reported in deliverable D2.6. The scenario was the Peter scenario, an automated drive along a winded, mountainous rural road. From time to time, slower lead vehicles appeared, and participants' task was to commandeer an overtaking manoeuvre of the TeamMate car. Subsequently, for each of the subjects at DLR, sequence diagrams were produced from the data gathered during the experiment which gave a very good impression of the motor, perceptual, and cognitive actions performed by the participants, as well as the goals pursued at each stage of the overtaking manoeuvre. The study served two purposes: 1) validation of the underlying method, DriveGOMS, and 2) delivering input to ULM's evaluation study of the Peter scenario reported in this document. Regarding 1), the data were compared with results obtained from a pre-study with 5 participants at ULM's laboratory. The resulting sequence diagrams and qualitative impression of participants' behaviour resembled those obtained at DLR's driving simulator closely, revealing the same underlying goal-plan-action structure. Because the situation (Peter-scenario), task (commandeering an overtaking scenario) and artefact (automated driving in a driving simulator) were fixed but the participants changed, the method can be considered to be of high face validity.

Figure 31: The Peter-scenario with a secondary task occupying participants implemented in DLR's MOSaiC driving simulator.

Regarding 2), a workshop was held with DLR and ULM at University of Ulm's laboratory, discussing the results obtained at DLR and the relevance of the insights for ULM's final evaluation study. Among other results, it was revealed that participants both in the pre-study and the main study had a great desire to be informed about their environment, even if they were not directly involved in the driving task. Further, a follow-up interview showed that every participant did not trust the automation entirely and thus tracked the vehicle during the overtaking manoeuvre. Finally, the one information every participant wished for to decide whether to initiate the overtaking manoeuvre or not was the distance to the oncoming vehicle. Of course, the lack of this information is the defining feature of the Peter-scenario, but the result shows that even in novel interaction situations such as with the TeamMate-car human drivers quickly recognize crucial elements of the situation and act accordingly. The probability of an oncoming vehicle reaching the ego-car during an overtaking manoeuvre was the critical variable participants attempted to gauge when deciding for or against initiating the overtaking manoeuvre.

4 Evaluation studies

4.1 Final Evaluation study of TeamMate concept in the PETER scenario (driving simulator)

4.1.1 Introduction

The experiment conducted in the ULM driving simulator aimed to evaluate the TeamMate Car with the integrated enablers in the last period. The PETER scenario has been used and adapted to measure the values of the integrated enablers in the ULM driving simulator.

In detail, with respect to the 1st evaluation performed at M24, the following updated features have been included in the TeamMate mode:

- The input modality has been extended, in order to simulate a more realistic behaviour of the driver-vehicle team. A haptic feedback was included because of the suggestion by the defined Guidelines (see Guidelines)
- The TeamMate multimodal HMI was updated and translated into German
- The DMS was integrated in the logic of the simulation; as a precondition to initiate the overtaking manoeuvre, participants had to be attentive
- All Enablers regarding the AR, DIR and RA were updated
- The online learning was integrated into the simulator
- A more realistic environment graphic, due to a new simulation version

Enabler	**Title**
E1.1	Driver monitoring system with driver state model for distraction and drowsiness
E2.1	*Driver intention recognition*
E3.1	*Situation and vehicle model*

E4.1	*Planning and execution of safe manoeuvre*
E4.2	*Learning of intention from the driver*
E5.1	*Online risk assessment*
E6.1	*Interaction modality*
E6.2	*HMI*
E6.3	*Augmented reality*

Table 2: List of implemented enablers in the ULM driving simulator

4.1.1.1 Description of the TeamMate Car and Baseline Car in Peter Scenario

TeamMate Car

Peter is driving in a narrow rural road in automated mode. There is a slowly driving car driving in front of the TeamMate car. The TeamMate car can learn Peter's normal behaviour in overtaking situations. Based on this, Teammate car can recognize Peter's intention on overtaking and decide for Peter if he wants to overtake. The TeamMate car supports Peter by showing him the crucial information, like the trajectory for the following manoeuvre as well as the risks of the oncoming traffic or, in case Peter would overtake in a situation in which is not possible, by informing him that it is too dangerous to overtake via augmented reality.

When the TeamMate car detects that the leading vehicle obstructs the view and is not confident of the oncoming traffic due to the limitation in perception. As a consequence, it would follow the slowly driving car (car-following mode) until the sensors provide enough information to safely overtake it or until the slowly driving car changes the lane. In such situations, the TeamMate car will ask Peter to check by himself in the form of augmented reality. When Peter confirms the possibility of overtaking, the TeamMate car will perform the overtaking manoeuvre in Automated Mode.

Baseline Car

The baseline car is representative for the state-of-the-art automated car that can execute longitudinal and lateral control. It means that the baseline car can drive through the rural road fully autonomously. However, different to the TeamMate car, the baseline car cannot suggest or be able to conduct an overtaking manoeuvre. It means that if there is a slowly driving car in front, the baseline car would follow it. Therefore, in order to execute an overtaking with the baseline car, Peter has to first turn off the automation and overtake the slowly driving car in front manually.

4.1.2 Method

4.1.2.1 Participants

Totally 18 participants with valid a German driving license for at least one year (Mean = 6.5; SD =2.7) have been recruited for the experiment (8 males, 10 females). They had an average age of 25.7 years (SD = 7.4). In average, participants travelled about 5000 km/year. After the experiment, they were paid 12 euros for 1.5-hour participation.

4.1.2.2 Material

4.1.2.2.1 Questionnaires

System Usability Scale

Systems usability was measured using a German translation of Brooke's (1996) System Usability Scale (SUS) consisting of 10 items, which provides a usability score ranging from 1 to 100. Participants were asked to give their ratings of agreement on a Likert scale ranging from 1 to 5.

Trust Questionnaire

Trust was measured after each experimental condition (Baseline car/TeamMate car) using Körber's "Trust in Automation" questionnaire in German Version, which consists of six scales (Reliability/Competence, Understandability/Predictability, Propensity to Trust, Intention of Developers, Familiarity, and Trust in Automation) containing a total of 19 items (Körber,2018). The participants were asked to rate their trust in the tested systems from 1 (Strongly disagree) to 5 (Strongly agree).

Acceptance Questionnaire

To measure driver acceptance of new technology, a German translation of the acceptance questionnaire from Van der Laan et al (1996) was used. This questionnaire consisted of 9 items. The participants were asked to indicate their judgements of the tested system regarding these 9 items and give scores for individual items from -2 to +2 with a pair of opposed adjectives (ie. "*useful*" versus "*useless*", or "*assisting*" versus "*worthless*").

NASA-TLX

Driver workload after each scenario was measured with a German translation of the NASA-TLX (Hart & Staveland, 1988). This questionnaire consists of 6 dimensions (mental demand, physical demand, temporal demand, performance, effort and frustration linked to the completion of a specific task). Participants were asked to rate their perceived workload in this six Likert scales ranging from 0 to 100.

Willingness to Buy and Willingness to Pay

Participants' willingness to buy a vehicle equipped with the baseline system and the TeamMate system was measured by scales ranging from 1 to 5. Participants were asked if they would buy the vehicle equipped with the

TeamMate system and with the baseline system. Responses were collected by mean of two scales ranging from 1 to 5.

Besides, a scale ranging from 0 € to 50 000€ was used to measure how much money participants were willing to pay to purchase the baseline system in addition to the price of the vehicle.

Questions regarding Enablers

In addition to the mentioned the questionnaires above, there are seven questions aiming to ask participants' satisfaction with the functions of the integrated enablers on a 5-point Likert scale ranging from 1 to 5:

- Are you satisfied with the Driver Monitoring System (allows to detect the driver's distraction and gives feedback to the driver) which is integrated in the TeamMate car?
- Are you satisfied with the Human Machine interface (information displayed on the dashboard that allows the communication of system's status and events to the driver), which is integrated in the TeamMate car?
- Have you realized that the system is trying to predict your overtaking intention?
- Is the predicted intention by the TeamMate car displayed via the AR-HMI same to your actual intention?
- Have you realized that the intention recognition integrated in the TeamMate car adapts to your overtaking behaviour?
- Do you like the fact that the intention recognition integrated in the TeamMate car adapts to your overtaking behaviour?

- Do you consider the overtaking behaviour of the TeamMate car as safe?

4.1.2.2.2 Simulator

The Simulator (described in D5.6) was equipped with all relevant enablers for the Peter scenario (see Table 2). The driving related parameters (speed, time lateral position and distance driven) are logged within the simulation and analysed with R-Studio.

The scenario was built in the simulation software SILAB (Version 6.0). It consisted of 12 overtaking possibilities where a slower vehicle was on the rural road in front of the ego-vehicle. The speed of the slowly driving vehicle was 60 km/h, whereas the ego-vehicle was driving with the allowed speed of 100 km/h. The weather condition changed via a hedgehog, a static point on the road where the weather got foggier in the second half of the course. There was also oncoming traffic throughout the course to manipulate the risk of overtaking.

4.1.2.2.3 Enabler Interaction

The demonstrator integrates a variety of different enablers to realize our concept of a TeamMate car. This section will present on overview of the interaction of these enablers throughout the experiment.

In general, the experiment consisted of a series of situations in which the baseline or TeamMate car approached a slower lead vehicle in the potential presence of incoming traffic limiting the opportunities to overtake. The TeamMate phase of the experiment consisted of two phases, a learning and a utilization phase. The learning phase imitated the cold start phase, in which the system is not yet adapted to the individual driver. In contrast, the utilization phase imitated the phase in which the system is assumed to be reasonable adapted to the individual driver, such that no more learning is required, although the system will continue to adapt to the driver. In the experiment, these phases were clearly separated and controlled. In reality, a gradual shift between learning and utilization phase would be expected.

As the TeamMate car approached a slower lead vehicle during the learning phase, the enablers interacted as follows: Enabler E4.2 "Learning of intention from the driver" (c.f. Section **Fehler! Verweisquelle konnte nicht gefunden werden.**) detects the need for user input and triggers E6.3 "Augmented reality" to request information of the driver, of whether and when he'd like to trigger an automatic overtaking manoeuvre, which is then automatically realized by the TeamMate car. For the sake of the experiment the trigger to show the request interface was designed rather simple. At the beginning of the learning phase the system was set to a mode which basically corresponds to the H2A support in perception mode. This was realized with trigger points in the simulation environment placed on certain points at the scenario track. In this mode, whenever the time headway to the lead vehicle on the slow lane (right lane) fell below a certain threshold and no driver feedback was given yet, the E6.3 would display its feedback request overlay, as shown in Figure 32.

Figure 32: Schematic illustration of the overlay shown by E6.3 to request the driver feedback

E4.2 then uses the new information concerning the driver's choice and the traffic situation to adapt the parameters of E2.1 "Driver intention recognition" (c.f. Section **Fehler! Verweisquelle konnte nicht gefunden werden.**).

Once the training phase has been completed (in the experiment triggered by a set number of overtaking instances) the utilization phase begins. Suppose the TeamMate car is once again approaching a slower lead vehicle. While approaching, E.2.1 "Driver intention recognition" will constantly assess whether the TeamMate car should change to the left lane in an attempt to overtake the lead vehicle, based on the information about the driver's behaviour. Once an overtaking intention has been recognized, the TeamMate car will plan a potential overtaking trajectory. E3.1 "Vehicle and situation models" (c.f. Section **Fehler! Verweisquelle konnte nicht gefunden werden.**) will then be used to predict the possible future evolution of all vehicles detected in the traffic scene using a sufficient temporal horizon to cover the duration of the trajectory. Both the planned trajectory and the evolution of the traffic scene are then used by E5.1 "Online risk assessment" (c.f. Section **Fehler! Verweisquelle konnte nicht gefunden werden.**) to calculate safety corridors and assess the safety of the planned trajectory in respect to the predicted evolution of the traffic scene as either safe or critical.

The results of E2.1, in terms of recognized intentions, and E5.1, in terms of trajectory assessments, are passed to E6.3 "Augmented reality" (c.f. Section **Fehler! Verweisquelle konnte nicht gefunden werden.**) to provide intention- and safety-dependent information to the driver, as indicated in figure 33. If at some point E2.1 indicates an overtaking intention and E5.1 deems the trajectory proposed by the TeamMate car as safe, the realization of the trajectory is triggered and TeamMate car changes to the left lane.

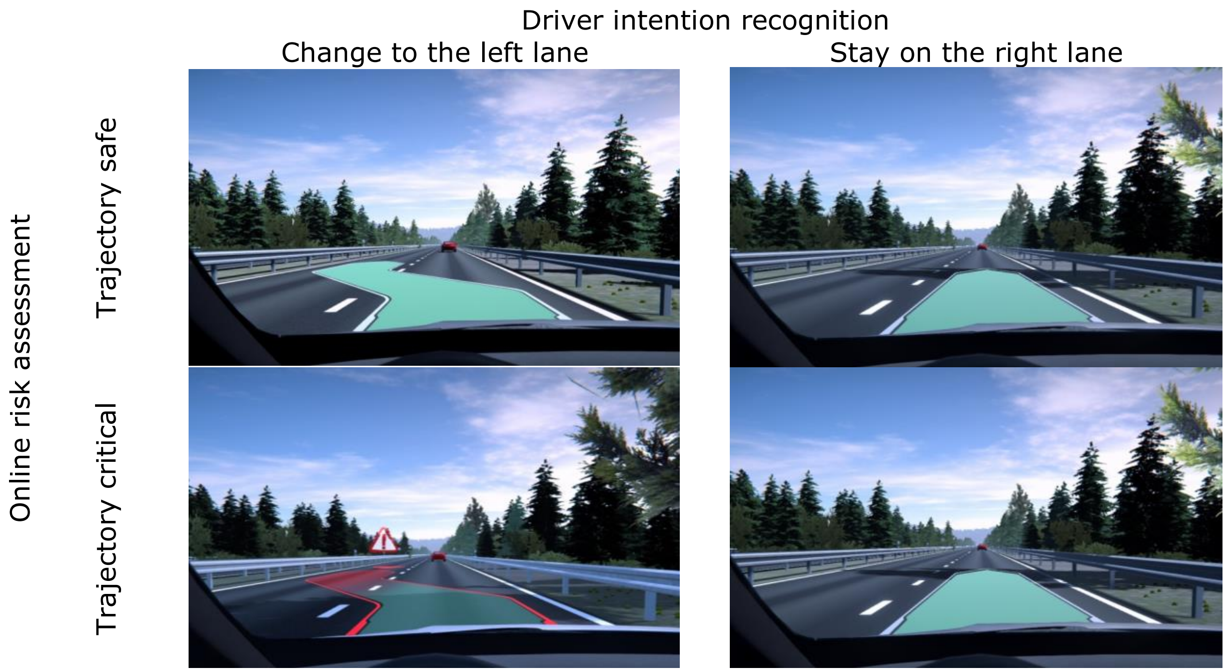


Figure 33: Schematic overview of information shown by E6.3 in respect to the output of E2.1 and E5.1 when driving on the right lane (conceptual graphics).

Analogously, while the TeamMate car is driving on the left lane to overtake the slower lead vehicle, E2.1 "Driver intention recognition" tries to recognize when the driver would have the intention to change back the right lane, triggering the generation of a potential return trajectory. Once again, the future evolution of all vehicles detected in the traffic scene predicted by E3.1 "Situation and vehicle model" and the potential trajectory is passed to E5.1. "Online Risk Assessment", which calculates safety corridors and assesses the safety of the TeamMate trajectory. The results of E2.1 and E5.1 are passed to E6.3 to provide intention- and safety-dependent information to the driver as indicated in Figure 34. The overall process is repeated until the trajectory is considered safe and the TeamMate vehicle returns to the right lane.

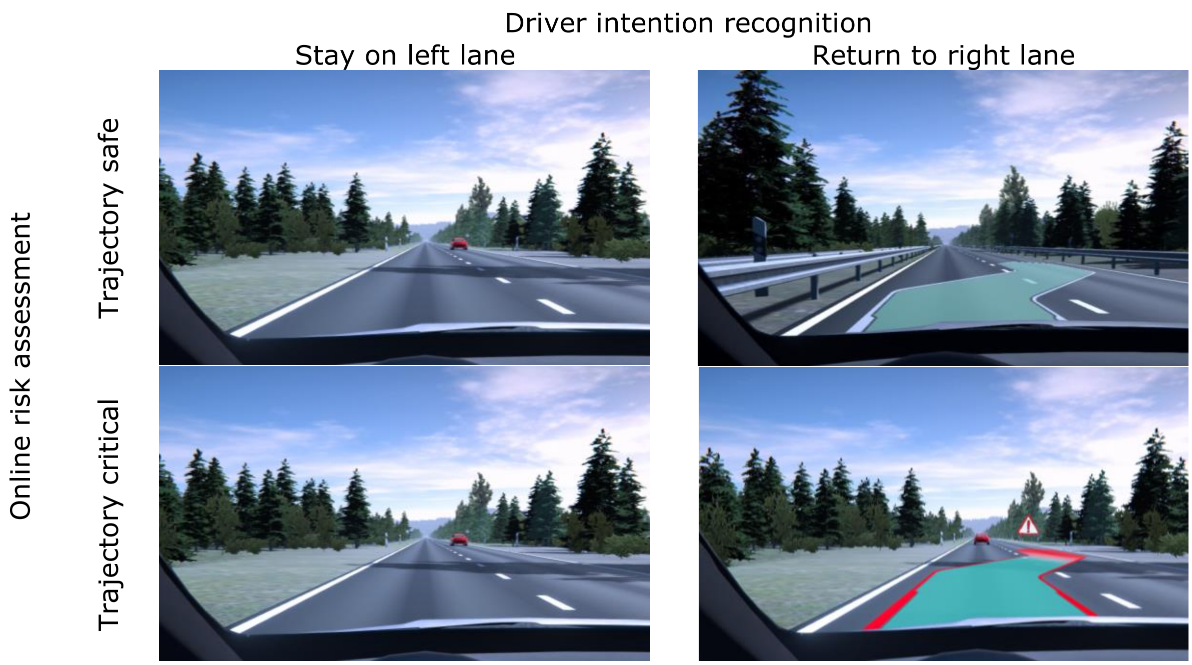


Figure 34: Schematic overview of information shown by E6.3 in respect to the output of E2.1 and E5.1 when driving on the left lane (conceptual graphics).

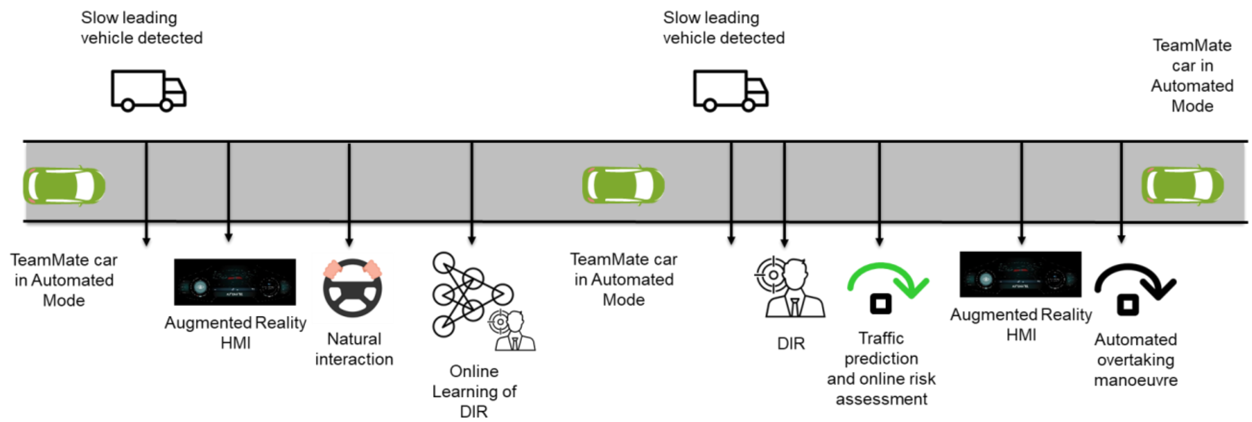


Figure 35: TeamMate enabler used in the final evaluation of the PETER scenario

All integrated enablers were needed at some point of the evaluation scenario (represented in Figure 35). The slowly driving vehicle was, different from the previous evaluation experiment, a normal vehicle and not a tractor anymore. This was due to a higher realism, because participants criticised the high density of tractors on a rural road before. Figure 35 shows two overtaking scenarios: one where the driver has to cooperate with the TeamMate car and the other one where the vehicle can overtake the slower vehicle on its own.

1.1.1.1 Experiment Design

A within subject design was used for this experiment. The manipulated factor was the function of the highly automated vehicle: Baseline car and TeamMate car. The dependent variables were measured objectively and subjectively. For the objective part, several parameters like speed, time, number of overtakes, accidents and steering wheel angle were logged in the simulation.

For the subjective part, usability, trust in automation, acceptance, workload, and willingness to buy and pay as well as the satisfaction with each integrated enabler was measured through questionnaires (see 4.1.3.1).

4.1.2.3 Experiment Design

A within subject design was used for this experiment. The manipulated factor was the function of the highly automated vehicle: Baseline car and TeamMate car. The dependent variables were measured objectively and subjectively. For the objective part, several parameters like speed, time, number of overtakes, accidents and steering wheel angle were logged in the simulation.

For the subjective part, usability, trust in automation, acceptance, workload, and willingness to buy and pay as well as the satisfaction with each integrated enabler was measured through questionnaires (see 4.1.3.1).

4.1.2.4 Procedure

First, participants were welcomed and asked to sign an informed consent form and data protection agreement. Then they were brought to the driving simulator and an explanation regarding the control of the driving simulator was given. Afterwards, SMI eye tracker was calibrated to record their eye movements and two electrodes were placed on the inside of the participants' left foot to measure the skin conductance.

After the training, the aim of the experiment and the functions and limitations of the Baseline car was first introduced to participants. All participants drove first with the Baseline car, followed by the TeamMate car. After about 20 minutes driving with the Baseline car, participants were asked to fill in the questionnaires regarding usability, trust, acceptance, workload, willingness to buy (see 4.1.3.1) on an Apple iPad. When they finished the questionnaire for the Baseline car, participants were given the instructions about the functions of the TeamMate car and asked to fill in the questionnaires as well as some questions concerning the performance of the integrated enablers.

After driving through both the Baseline car and TeamMate car, participants were asked to fill out the demographic questionnaire and give comments regarding the performance of two systems. In the end, they were paid with 12 Euros.

4.1.3 Results

4.1.3.1 Objective Results

One person had to be excluded from the analysis of the objective data due to recording problems. Therefore, the number of fully recorded participants was 17.

The evaluation of the objective driving data from the simulator showed that participants took longer to drive through the course with the baseline car (M = 1262.33s, SD = 145.61s; M = 21:02min, SD = 2:26min) than with the TeamMate car (M = 1145.82s, SD = 110.28s; M = 19:06min, SD = 1:50min). This difference was significantly lower, $t(16) = -4.185$, $p = .001$.

Participants used the automation 70.69% (SD = 19.92) of the time while driving with the baseline car. While driving with the TeamMate car, participants used the automation 93.55% (SD = 11.14) of the time in average. This was statistically significant, $t(16) = 15.64$, $p<.001$.

During the experiment, whenever the driver was cooperating with the automation during the learning phase, or the driver would deactivate the automation, enabler E4.2 gathered data to update the DIR model of enabler E2.1 for the respective current driver starting from the initial model provided by E2.1. As mentioned in [6] the DIR model consists of sub-models which are related to the current ego lane. On average 91825 new samples were added to the sub-model which is responsible when the ego vehicle is driving on the slow/right lane. This is the more important model if we are interested in parameters that influence the lane change intention to the fast lane, thus the initiation of an overtaking manoeuvre. As an example for the updates of the DIR model, the Figure 36 shows the initial and the current/final distribution $\mathrm{p}(\mathrm{ANF\ Speed\ Diff}^{\mathrm{t}}|\mathrm{Opportunity}^{\mathrm{t}} = 1, \mathrm{ANF\ Present}^{\mathrm{t}} = 1)$for several subjects. This distribution is a Gaussian Mixture Model and describes which speed differences to the oncoming traffic can be observed if the driver sees a subjective opportunity to change the lane and therefore to overtake.

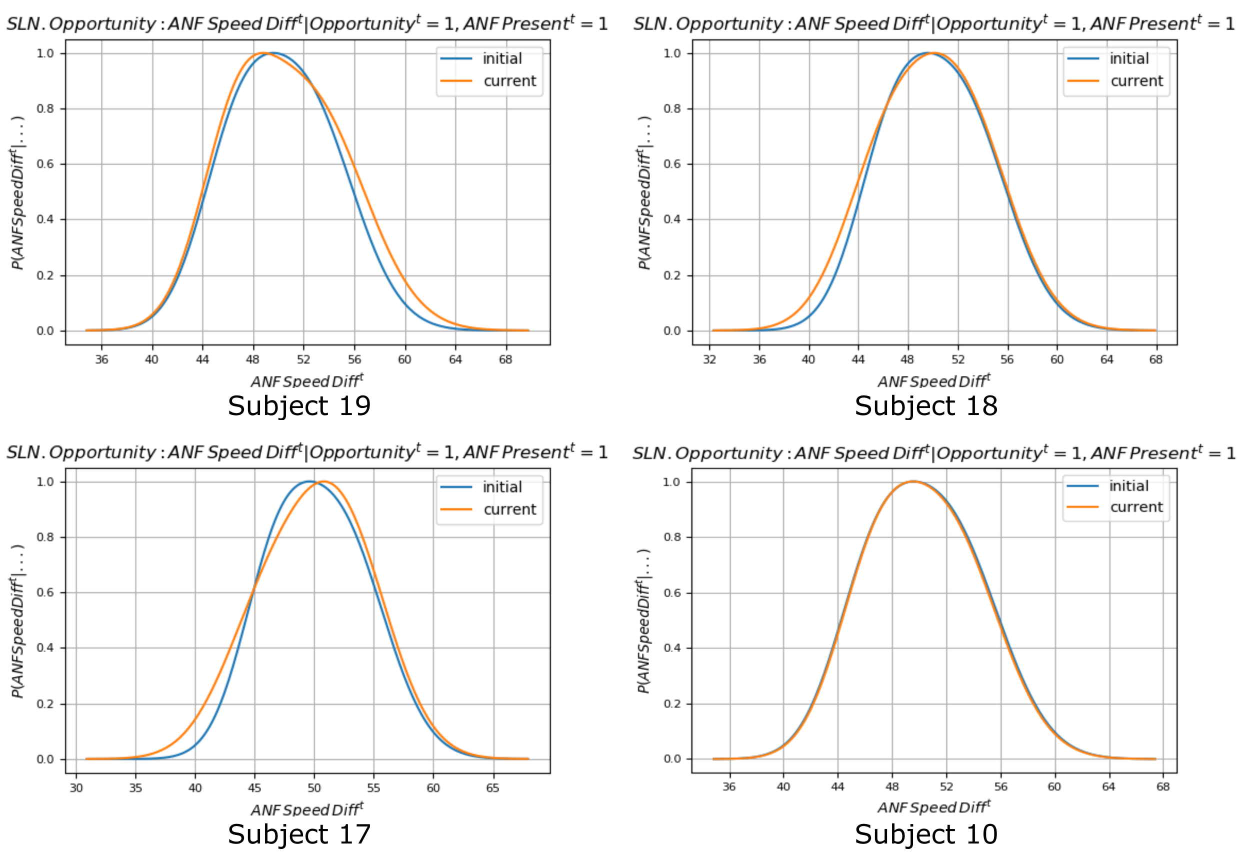


Figure 36: Exemplary initial and updated distributions for several subjects

For the subjects 17, 18, 19 the distribution shows clear changes due to the new samples, subject 10 is at least for this distribution still close to the initial model.

4.1.3.2 Questionnaires

System Usability Scale

The System Usabilty Scale (Brooke, 1996) has been used to evaluate the usability of both the Baseline car and the TeamMate car. The results of the analysis (see Figure 37) showed that there was a significant difference ($t(17) = -2.321$, $p = 0.03$) regarding the rating of the usability between the Baseline car ($M = 72.5$, $SD = 14.31$) and the TeamMate car ($M = 80.78$, $SD = 6.82$), where the usability of the TeamMate car was higher than the Baseline car.

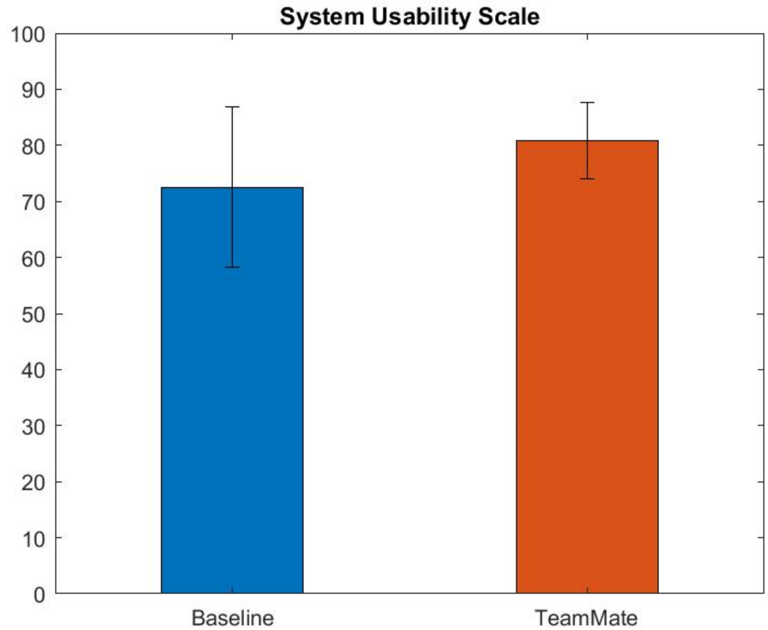


Figure 37: SUS score of the baseline and the TeamMate car (The error bars depict standard deviation)

Trust Questionnaire

To evaluate the trust in automation for both the Baseline car and the TeamMate car, the questionnaire Trust in Automation (Körber, 2018) has been used. The results of the analysis (see Figure 38) showed that there was no significant effect ($t(17) = -2.034$, $p = 0.058$) regarding the trust in automation between the Baseline car (M = [3.23, SD = 0.47) and the TeamMate car (M = 3.43, SD = 0.31).

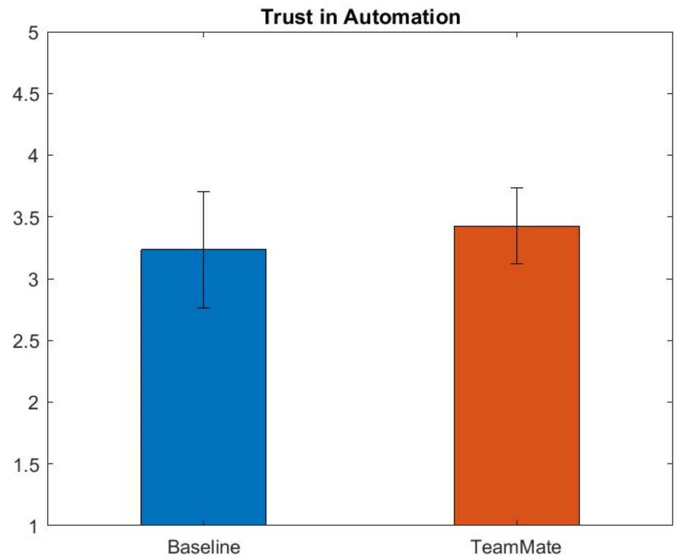


Figure 38: **The rating of trust of the baseline and the TeamMate car (The error bars depict standard deviation)**

Acceptance Questionnaire

The Acceptance Questionnaire by Van der Laan et al. (1997) has been used to evaluate the acceptance of both the Baseline car and the TeamMate car. This questionnaire consists of 2 dimensions: usefulness and satisfying.

The results of the analysis (see Figure 39) showed that there was no significant difference ($t(17) = 2.093$, $p = 0.052$) regarding the perceived usefulness of the new technology between the Baseline car ($M = -0.41$, $SD = 0.49$) and the TeamMate car ($M = -0.58$, $SD = 0.45$).

The results of the analysis (see Figure 39) showed that there was no significant difference ($t(17) = 0.45$, $p = 0.66$) regarding the satisfaction with the new technology between the Baseline car ($M = -0.07$, $SD = 0.33$) and the TeamMate car ($M = -0.125$, $SD = 0.44$).

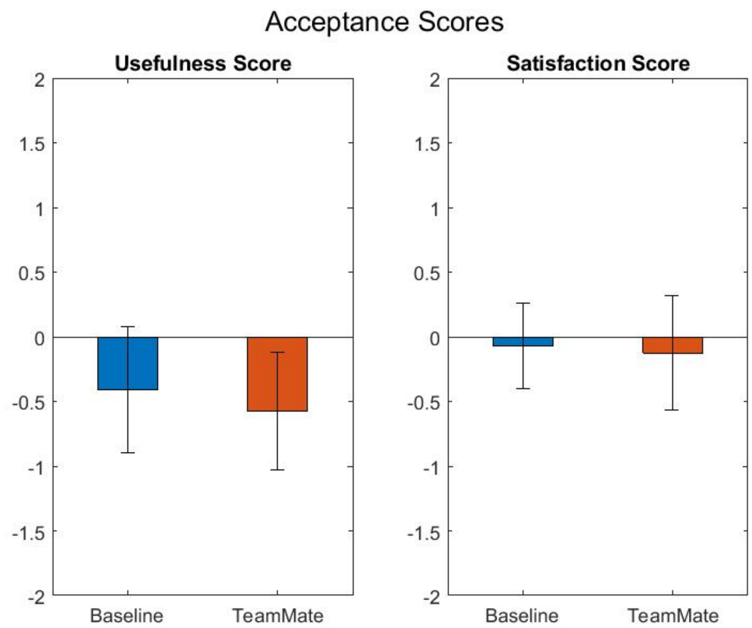


Figure 39: The mean Usefulness – Score (left) and Satisfaction – Score (right) of the baseline and the TeamMate car (The error bars depict standard deviation)

NASA-TLX

To evaluate the driver workload for both the Baseline car and the TeamMate car the NASA-TLX (Hart & Staveland, 1988) has been used. The results of the analysis (see Figure 40) showed that there was a significant effect ($t(17) = 5.21$, $p < 0.001$) regarding the workload between the Baseline car ($M = 8.10$, $SD = 2.37$) and the TeamMate car ($M = 5.49$, $SD = 1.54$).

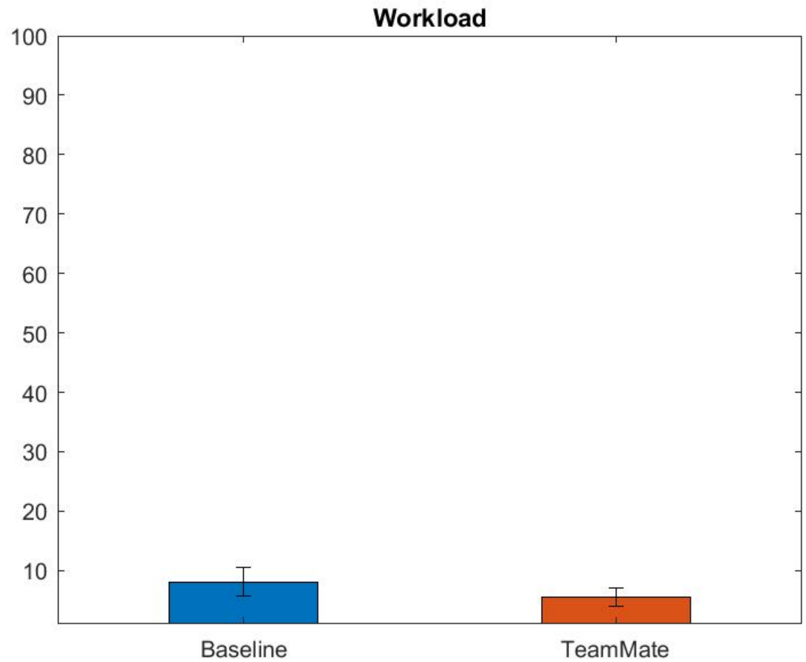
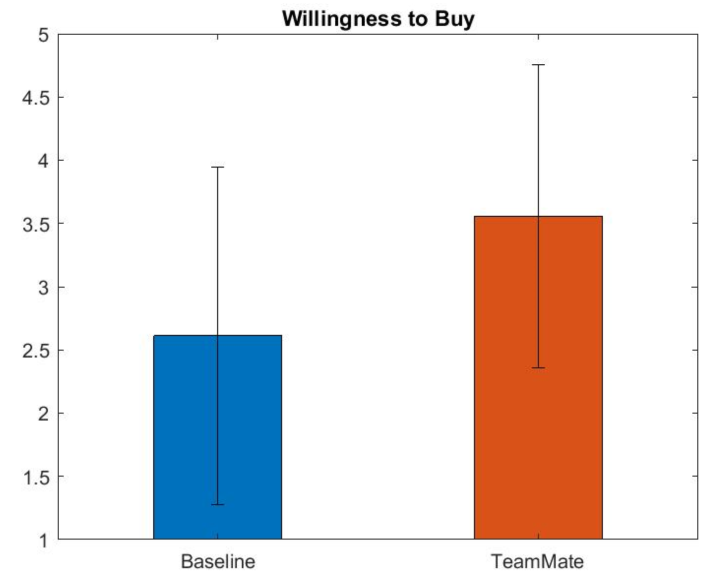


Figure 40: The workload score of the baseline and the TeamMate car (The error bars depict standard deviation)

Willingness to Buy and Willingness to Pay

Willingness to buy:

The results of the analysis (see Figure 41) showed that there was a significant difference ($t(17) = -2.46$, $p = 0.025$) regarding the willingness to buy between the Baseline car (M = 2.61 , SD = 1.33) and the TeamMate car



(M = 3.56, SD = 1.20).

Figure 41: The willingness to buy of the baseline and the TeamMate car (The error bars depict standard deviation)

Willingness to pay:

In general, participants would like to pay the function of the Baseline car with around 19444 euros (SD = 8024 €), while 23889 euros (SD = 9785 €) for the TeamMate car (see Figure 42).

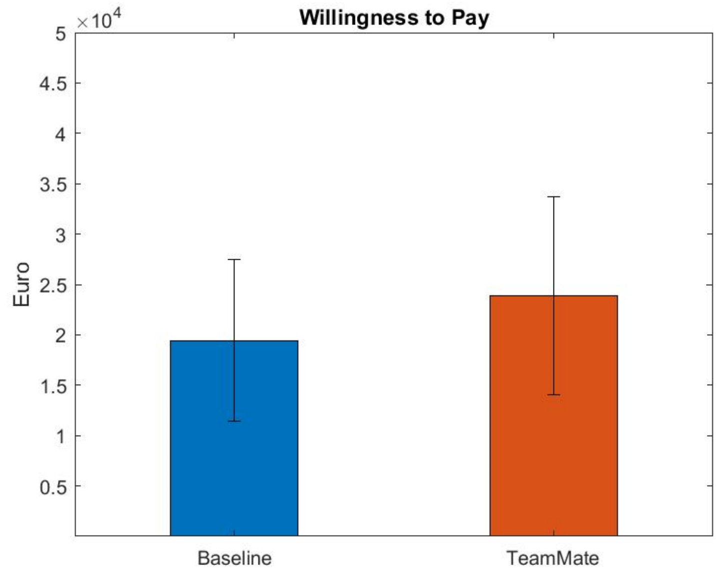


Figure 42: **The willingness to pay of the baseline and the TeamMate car (The error bars depict standard deviation)**

Questions regarding Enablers

With regard to the Driver Monitoring System integrated in the TeamMate car, the participants gave an average score of 3.56.

Concerning Human Machine interface integrated in the TeamMate car, the participants gave an average score of 4.11.

Regarding the prediction of driver's overtaking intention, participants gave an average score of 3.78.

Regarding the match between the predicted intention by the TeamMate car displayed via the AR-HMI same and the driver's actual intention, participants gave an average score of 3.44.

Concerning the question if the intention recognition integrated in the TeamMate car adapts to drivers' overtaking behaviour, participants gave an average score of 3.5.

Concerning the question if drivers like the fact that the intention recognition integrated in the TeamMate car adapts to their overtaking behaviour, participants gave an average score of 4.06.

Regarding the question that drivers consider the overtaking behaviour of the TeamMate car as safe, participants gave an average score of 3.56.

4.1.4 Discussion

The objective results show a significant improvement in the time efficiency of the cooperative TeamMate approach. That means that with the TeamMate system the participants could reach the point of destination faster without reducing the safety. There were no accidents – neither in the baseline nor the TeamMate condition. Therefore, it is highly recommended to use the cooperative approach.

The usage of the automation was also significantly higher in the TeamMate condition. The aim of increasing the usage of the automation for e.g. safety reasons could be met by using the TeamMate cooperative approach. Two participants never deactivated the automation which means that they were completely satisfied with the functionality of the system and had a appropriate level of trust towards it.

The online learning could only be shown objectively. The enabler was working properly during the experiment but was not “felt” by the participants due to the duration of the experiment and the number of overtaking manoeuvres. This could be further investigated in future long-term studies or in demonstrator vehicles where it is possible to conduct longer studies without risking the simulator sickness effect.

Summarizing the results of questionnaires, the TeamMate car had a higher usability rating, higher willingness to buy and pay, and lower workload than the Baseline car. However, there was no difference between both systems with regard to the trust and acceptance in automation. The reasons for this can be found in the questionnaires which was reported by participants: The TeamMate car steers too early to the left once the oncoming vehicle passes by, or the TeamMate car often executes the overtaking manoeuvre in some situations, in which participants actually don't prefer to overtake. Besides, another criticism regarding the TeamMate car is the overtaking in the fog situation: Usually participants prefer not to overtake in the fog, but TeamMate can execute overtaking also in the fog. However, participants have reported that it is unclear for them to know how the TeamMate car estimate the criticality of the overtaking manoeuvre, as the TeamMate car doesn't explain it to drivers.

Regarding the rating of the performance of the integrated enablers, the TeamMate car got a relatively higher average score (around 3.5) in the range of 1 to 5. The Human Machine interface got the highest rating (4.11). Besides, participants also gave a higher rating (4.06) concerning the intention recognition enabler that was integrated in the TeamMate car that adapts to their overtaking behaviour. On the other hand, the match between the predicted intention by the TeamMate car displayed via the AR-HMI same and the driver's actual intention got a lowest rating (3.44), which was consistent with the reported comments by the participants. In addition, the Driver Monitoring System integrated in the TeamMate car got a relatively lower rating of 3.56. Regarding this, participants reported in the questionnaire that the TeamMate car was reliable, but they were quickly distracted by the TeamMate car.

4.2 Final evaluation of the EVA scenario (driving simulator)

4.2.1 Introduction

The aim of the experiment conducted for the final evaluation cycle at REL premises was to evaluate the added value of the ecosystem of enablers integrated in the last period in REL simulator demonstrator. The EVA use case, described in D1.3 and D1.5, has been selected and adapted in order to answer new research questions, and to measure the value of the enablers integrated in the last cycle.

In this experiment new interaction modalities (in terms of input modalities and information provided to the driver) and new capabilities of the automation have been included.

In detail, with respect to the 1st evaluation performed at M24, the following updated features have been included in the TeamMate mode:

- The input modality has been extended, in order to simulate a more realistic behaviour of the driver-vehicle team.
- The TeamMate multimodal HMI integration included, in this phase, also the distributed HMI on mobile application.
- The automation has been customized in order to consider a human-like behaviour, through the integration of the Driver Intention Recognition module. This module allowed, when activated, to enter the roundabout according to the driver preferences.

Table 3 summarizes the enablers integrated in REL simulator.

Enabler	Title
E1.1	Driver monitoring system with driver state model for distraction and drowsiness
E4.2	Learning of intention from the driver
E6.1	Interaction Modality
E6.2	TeamMate multimodal HMI

Table 3: Enablers integrated in REL simulator

4.2.2 Scenario

The most relevant part of EVA scenario is the roundabout. Two types pf roundabout have been designed in order to test different situations: a small roundabout, repeated twice to measure the added value of the Driver Intention Recognition, and a big roundabout in which the vehicle, since is no longer able to recognize the lanes, performs a take-over request.

The scenarios include also other road elements and vehicle in order to increase the realism of the simulation. It was designed as a “ring” in order to simplify the data collection. The same simulation “Terrain” used in first evaluation cycle has been used in this cycle, with minor changes (especially in terms of landscape and traffic situations).

In Figure 43 the simulation scenario is reported.

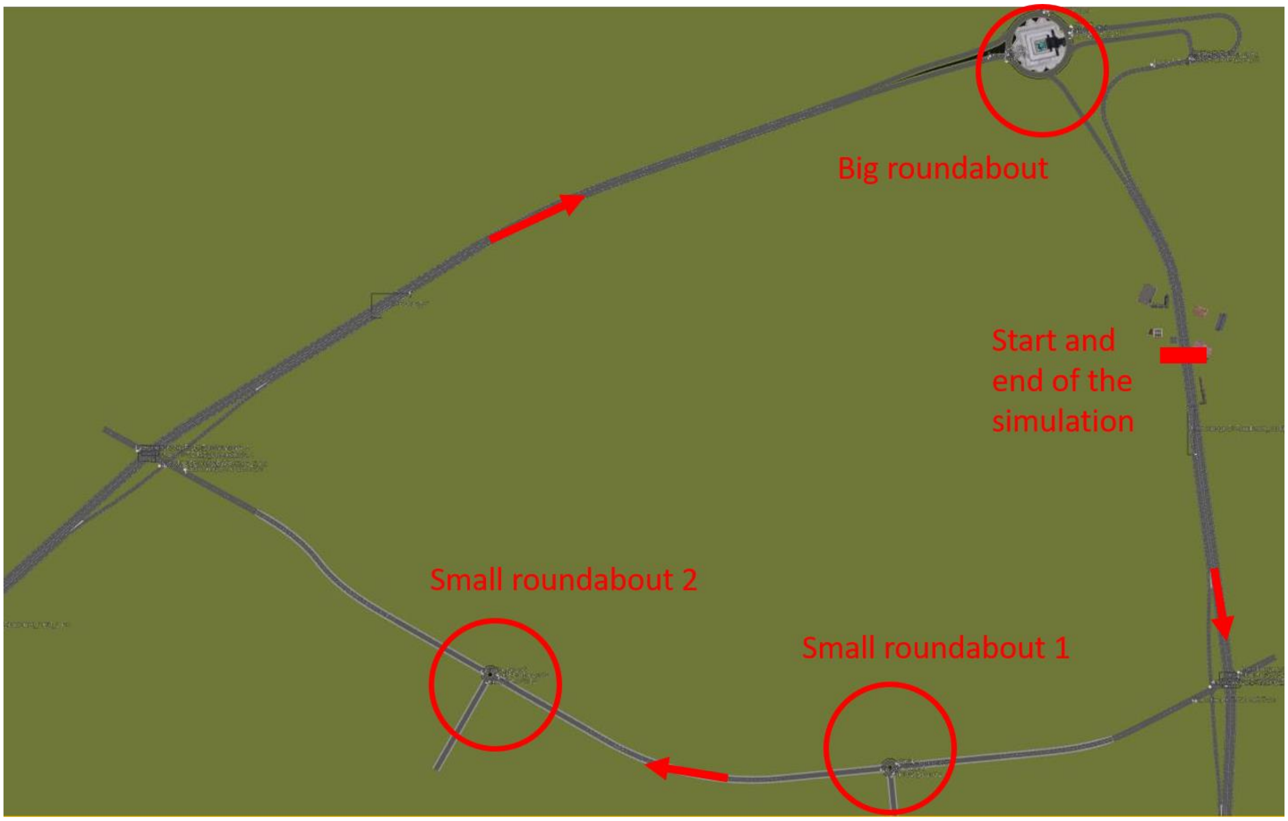


Figure 43: Baseline and TeamMate itinerary in EVA scenario

In TeamMate mode, the scenario starts in Automated Mode. When the car approaches the first roundabout the car would normally perform an over-safe manouver, waiting until the roundabout is completely empty. In this case, the Driver Intention Recognition (DIR) is activated in order to allow a more effective and human-like behavior. The entering in the roundabout is triggered by the DIR (that acts in background through all the scenario): when the probability to enter the roundabout exceeds the 80% (according to the traffic conditions and the preferences collected to train the model), the car automatically enters the roundabout. This part of the scenario is repeated twice, with different traffic conditions.

On the way between the second and the third roundabout, the driver is asked by the experimenter to read aloud a text (in Italian language) on a smartphone placed in the central tunnel. When approaching the third roundabout, since the car needs the driver's intervention and the DMS detects that the driver is looking in that specific Area of Interest, the Take Over Request is given directly on the smartphone. If the driver reacts properly (i.e. with natural interaction, by pressing a pedal) he/she takes the control and performs the roundabout manually; if he/she does not react, the car performs a Minimum Risk Manouver, by stopping before entering it. After the roundabout, the car, when detecting again the lanes and is able to regain the control, performs an automatic switch to Automated Mode, informing the driver of the transition.

In this case the main research questions of the first part of the scenario concerned the effectiveness at the roundabout (a comfort-related parameter) and its effect in terms on impact on the driver, while the second part of the scenario concerned safety related parameters, such as the time needed to the driver to take back the control when supported by an adaptive distributed HMI.

In particular, the following crucial indicators have been taken into account:

- The time to take over (and in particular the delta time between the baseline and the TeamMate)
- The number of safe maneuvers, since this can be considered a safety critical at the roundabout, with a significant impact on traffic situation.

Figure 44 reports the flowchart with the TeamMate scenario.

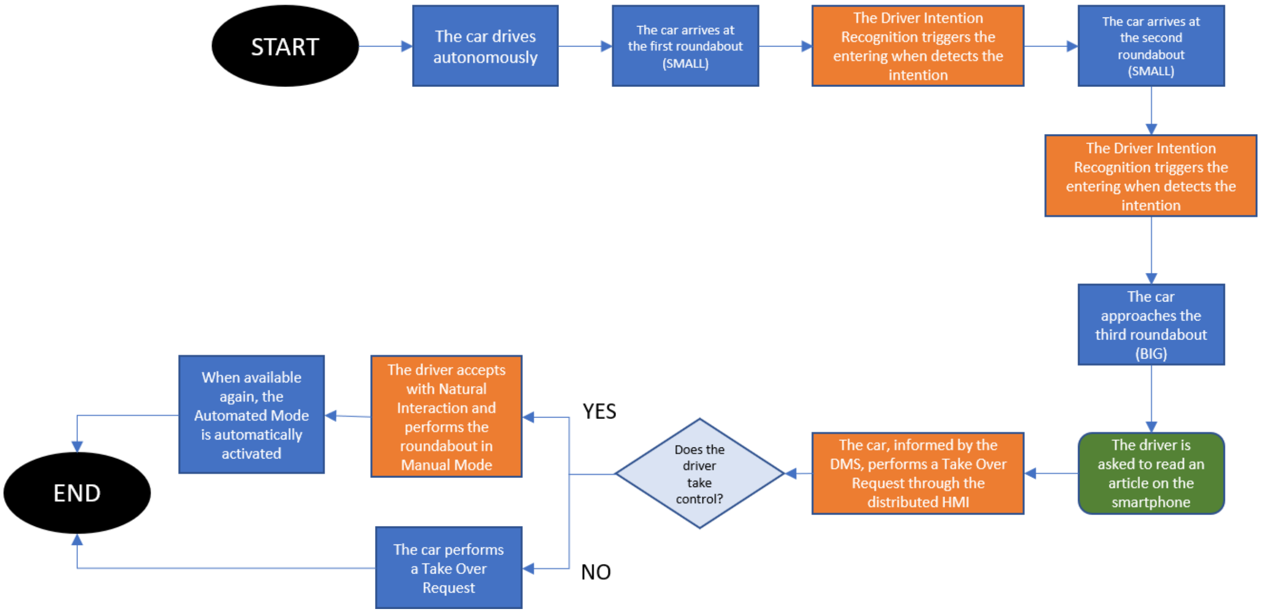


Figure 44: TeamMate scenario in REL simulator

4.2.3 Baseline

According to the approach described in the common evaluation framework in D6.1, the Baseline scenario consists in performing the same driving scenario without the enablers, in order to evaluate the added value of the technologies developed in AutoMate.

Figure 45 describes with a flowchart this scenario. As in the TeamMate scenario, the simulation starts in Automated mode. However, when vehicle approaches the roundabout 1 and 2, the manoeuvre is performed without the DIR (since the DIR is not included in the baseline): it enters the roundabout only when it is empty. At the second roundabout, even if the driver is distracted (he/she has been asked to read aloud a text) the take over request is given only in the instrument cluster (since there are no DMS and no distributed HMI). If the driver does not react, the car performs a minimum risk manouver; if the driver takes control, after the roundabout he/she is asked to manually activate the automated driving.

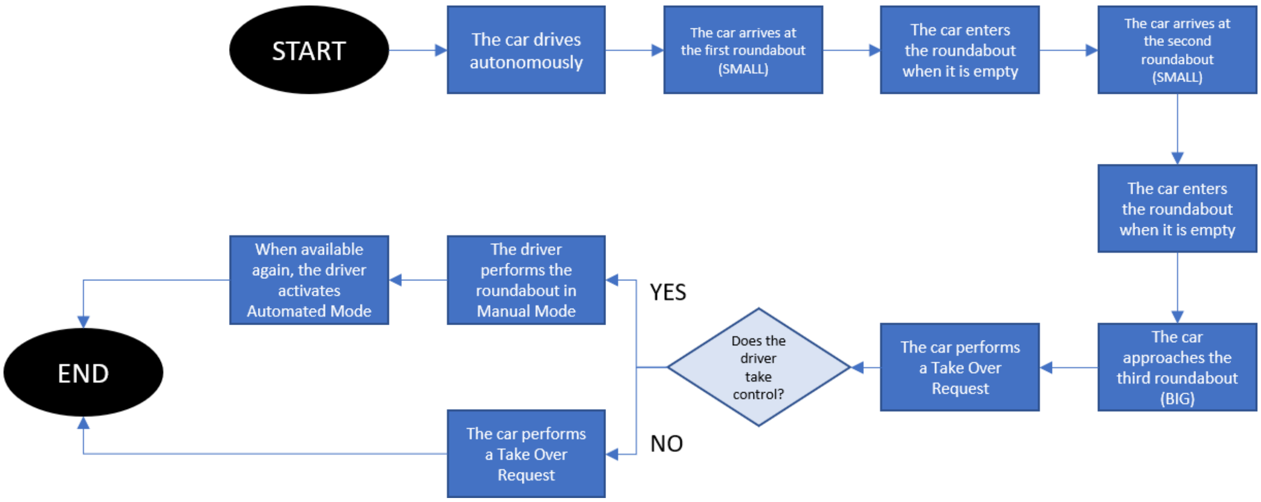


Figure 45: Baseline scenario in REL simulator

4.2.4 Method

The test has been designed as a between-subjects experimental design, i.e. each participant performed the same scenario. The participants had to drive the scenarios twice, one in Baseline and one in TeamMate mode (see Figure 46). The order in which the users performed each scenario was alternated in order to avoid biases. Twenty subjects have been recruited, balanced for gender in order to reduces biases (11 males and 9 females).

The users were welcomed and asked to sign a consent form module for the data protection. Then, basics demographics data (e.g. gender, age, driving experience, driving habits) have been collected in order to allow the creation of data clusters. The users were asked to have a 5-minutes trial, alternating

Manual and Automated driving, in order to become familiar with the driving simulator and the automation logics. They were introduced to AutoMate concept and watched a video describing the main pillars of the project.

Figure 46: Experimental setup at driving simulator

The evaluation focused on measuring mostly comfort- and acceptability-related parameters. After each scenario, the users were asked to answer a questionnaire aimed at assessing the user satisfaction in using the TeamMate system compared to a baseline.

The following items were considered the most relevant for this cycle (according to the use case tested in this demonstrator):

1. The user acceptance

2. The trust in the automated system
3. The workload in using the system
4. The willingness to buy (and to pay) the system
5. The efficiency

As stated in the Common Evaluation framework, the following tools have been used to respectively measure them:

1. The Van der Laan questionnaire
2. The Koerber questionnaire
3. The NASA-TLX
4. A custom questionnaire, created ad-hoc to evaluate these propensity
5. The execution time in the roundabout and the reaction times at the Take Over Request

All the questionnaires (i.e. items 1, 2, 3 and 4) are reported as annex.

Moreover, the implications in terms of safety, related to the Safe Manoeuvre eventually performed at the roundabout, have been considered a Key Performance Indicator and will be discussed in the next paragraphs.

Table 4 summarizes the KPIs considered in REL simulator demonstrator.

KPI ID	KPI	KPI Type	Recording Tool
KPI1	Time to enter the roundabout	Objective	Simulator's logs
KPI2	Acceptance	Subjective	Van der Laan questionnaire
KPI3	Trust	Subjective	Koerber questionnaire
KPI4	Workload	Subjective	NASA-TLX

KPI5	Willingness to buy	Subjective	Custom questionnaire
KPI6	Willingness to pay	Subjective	Custom questionnaire
KPI7	Time to take over	Objective	Simulator's logs
KPI8	Number of safe manouver	Objective	Simulator's logs

Table 4: List of KPIs for REL demonstrator

Qualitative data have been collected too, i.e. through users' comments and observations.

4.2.5 Results

As reported in Table 5, the TeamMate system showed relevant improvements compared to the Baseline. These improvements can be seen both for objective and subjective measures.

KPI ID	KPI	Baseline results	TeamMate result	Delta
KPI1	Time to enter the roundabout	12,75 seconds	6,07 seconds	6, 68 seconds (-52,39%)
KPI2	Acceptance	-0,02	+0,66	+0,64
KPI3	Trust	-0,02	+0,35	+0,37
KPI4	Workload	- 2,67 (weighted score of NASA-TLX)		
KPI5	Willingness to buy	-0,05	+0,85	+0,9
KPI6	Willingness to pay	4880 €	7000 €	2120 €
KPI7	Time to take over	7,65 seconds	4,49 seconds	3,16 seconds (-41,34 %)

KPI8	Number of safe manouver	4	1	3

Table 5: KPI results for REL demonstrator

The TeamMate system showed significant improvements in user acceptance (see Figure 47), according to the Van der Laan Scale. The total average score of the TeamMate system, in fact, was +0,66, against the -0,02 obtained by the Baseline. In particular, the users found the TeamMate system as "Effective" (+1,00 against the +0,25 of the Baseline) and "Useful" (+0,8 against the +0,15 of the Baseline). From qualitative data, collected through the thinking aloud method, it has been noted that the users appreciated the behaviour of the car in the small roundabout, in which the Driver Intention Recognition allowed a faster and more human-like entering manouver. The users considered this behaviour more realistic and closer to the actual behaviour they would have eventually performed in Manual Mode.

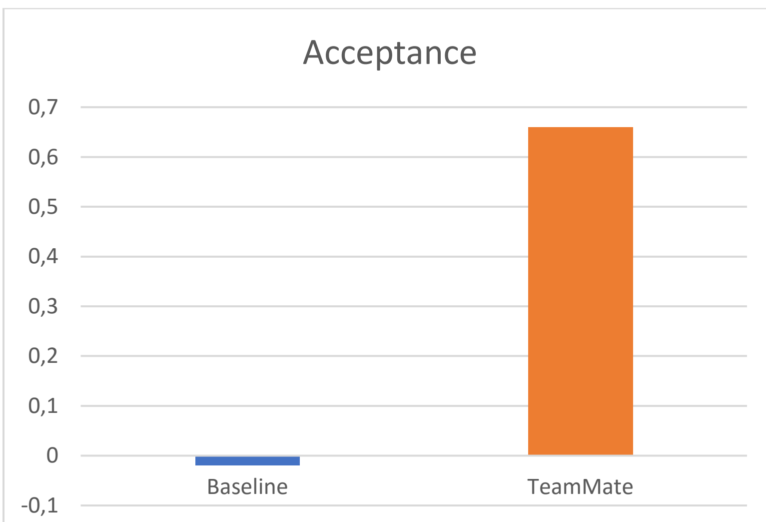


Figure 47: User Acceptance in REL simulator

The results of the experiment show improvements also in terms of trust in the system (see Figure 48). The average score of the Baseline was -0,02, while the average score of the TeamMate system was +0,35. In particular, the users found the TeamMate "System state" more clear than the Baseline (+0,15 against -0,35 on average) even if, according to the comments, the state was not always evident. A significant improvement has been found in understanding the reasons leading to the system behaviours: the item "I was able to understand why things happened" obtained an average score of +0,45 in TeamMate Mode against the -0,3 obtained by the Baseline. The users appreciated the explanation provided through the HMI, considering it as useful to increase the transparency of the automation's behaviour. This is particularly relevant since it strengthens the concept designed in AutoMate, based on cooperation: the users found the negotiation-based approach as a promising way to allow an effective and smooth interaction between the human and the automation.

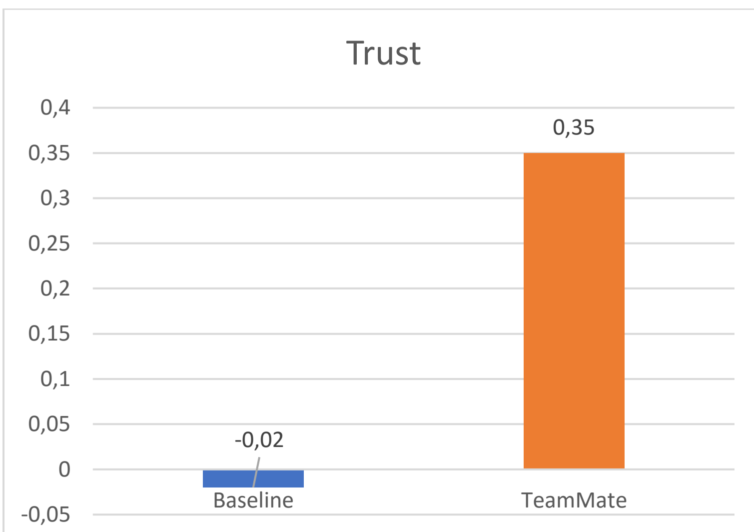


Figure 48: Trust in automation in REL simulator

The users interviewed after the test considered the idea of buying a car with the features they experimented in the evaluation; the TeamMate system obtained an average score (on a 5-points Likert Scale, measured with a -2/+2 interval) of +0,85 against -0,05 of the Baseline (see Figure 49). The users would be willing to pay 7.000 € on average for the features tested in the TeamMate System against the 4.880 € for the Baseline (see Figure 50). However, as indicated by the high Standard Deviation (SD=5770,62 for the TeamMate System), the judgement on this topic is not particularly indicative, and some users did not consider the idea of buying (and paying) the system.

Figure 49: Willingness to buy in REL simulator

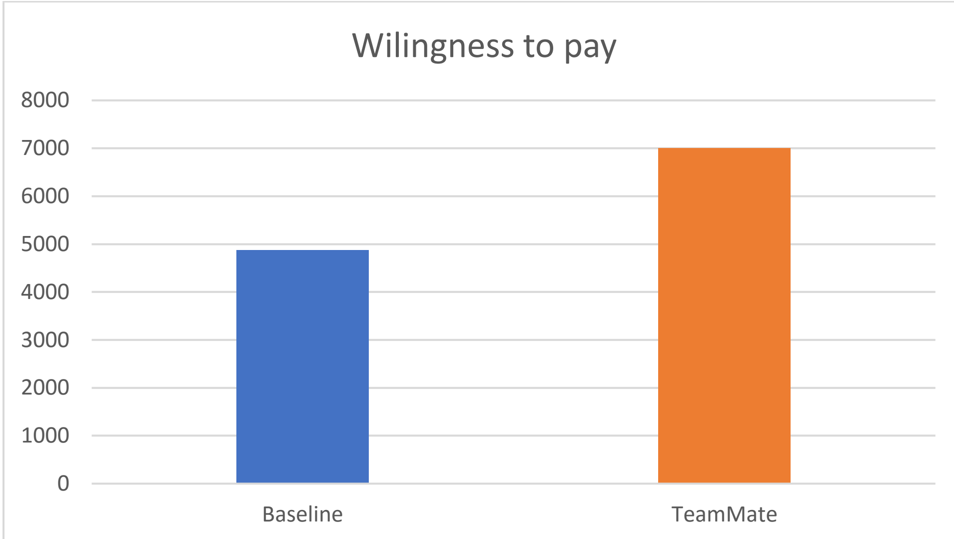


Figure 50: Willingness to pay in REL simulator

Figure 51 shows the comparison of the main items of the NASA-TLX, used to measure the overall workload experienced by the driver in the simulation. The TeamMate system showed a great potential in reducing the average Mental Demand (4,95 against the 6,9 of the Baseline), the Frustration (4,7 against the 10 of the Baseline) and in increasing the self-perceived Performance (9,8 on average against the 5,9 of the Baseline). The workload reduction is a relevant result for the overall system: it was also justified by users, who clarified that the possibility of having a tool (the Driver Intention Recognition) able to reduce the time to enter the roundabout, could significantly improve the performance in specific situations (e.g. when in a hurry); moreover, they confirmed that the possibility of having a state-adaptive distributed HMI (on the mobile phone) is able to increase the sense of safety and the Temporal Demand (5,6 against the 8,15 of the Baseline) needed to effectively interact with the system.

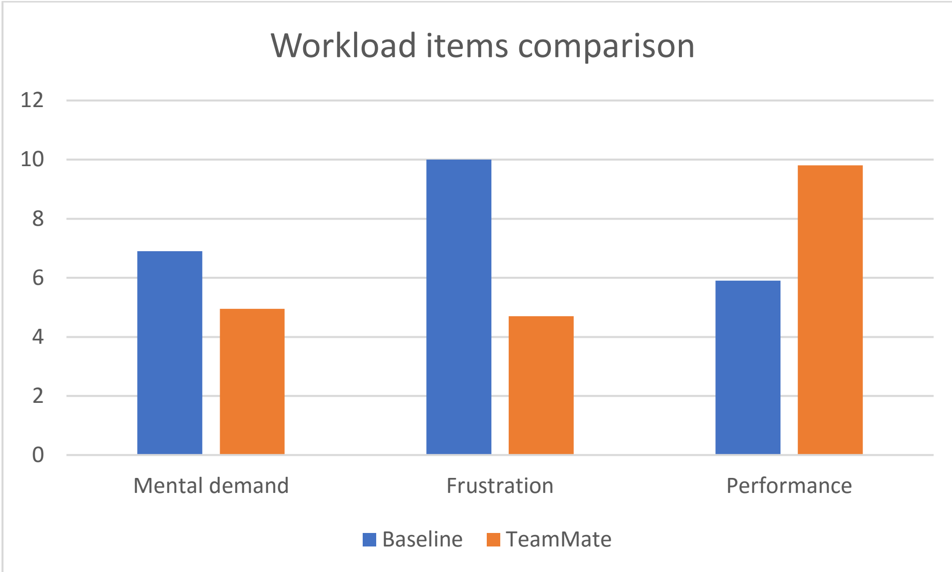


Figure 51: Workload items comparison in REL simulator

Also objective data collected from the driving simulator demonstrates the impact of the TeamMate solutions, in terms of efficiency and safety.

At the small roundabouts, when the car hesitated in entering due to traffic flows, the Driver Intention Recognition (DIR) allowed to significantly speed up the entering manouver. In Baseline (i.e. without the DIR), the car took 16,35 seconds on average to enter the first roundabout and 9,14 seconds on average to enter the second roundabout (on a sample of 20 roundabouts per each scenario), In TeamMate mode, the car took 8,70 and 3,44 seconds respectively. On average (considering both the roundabouts) the Baseline system took 12,75 seconds to enter the roundabout, while the TeamMate took 6,07 seconds. This data confirms that the TeamMate system is able to increase the efficiency, since it halves the execution time of the specific manouvers considered in this evaluation.

Another significant insight is represented by the number of Safe Manouvers (KPI8): in Baseline mode, the vehicle performed 4 automatic stops before entering the roundabout (since the driver did not react in time to a Take Over Request), while in TeamMate mode it performed just 1 stop. Since the experiments were randomized (i.e. alternating the Baseline and the TeamMate as first scenario administered to each user), no learning effects are considered: this can be explained with the combined role of the distributed HMI and the DMS; by offering information to the driver directly in the device he/she is looking at, the vehicle is able to ensure a safer, smoother and effective interaction.

This finding is also confirmed by the Time to Take Over measurement (KPI7): in Baseline mode, this indicator was 7,65 seconds (SD=2,76), while in TeamMate mode it was 4,49 seconds (SD=1,12), with a reduction of 3,16 seconds (41,34%), see Figure 52.

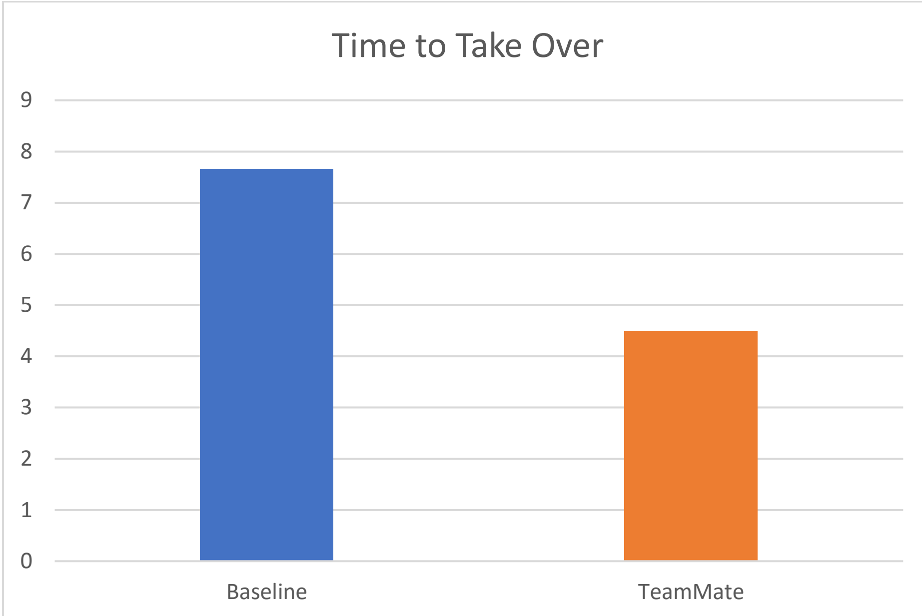


Figure 52: Time to Take Over in EVA scenario

This result has relevant implications on safety, since it witnesses the relevance if the TeamMate approach in ensuring an effective communication and interaction between the human and the technological agent.

Finally, no accidents were found in the experiment (neither in Baseline nor in TeamMate modes).

The detailed results of this experiment, including the simulator's logs and subjective surveys, are reported in D6.4 and published as Open Data.

4.2.6 Discussion

The results, collected in this last project cycle to evaluate the TeamMate system against the Baseline, show significant improvements related to both objective and subjective data.

In particular, the TeamMate system showed significant improvements in acceptance and workload reduction for this specific use case. Also trust in automation was slightly increased. The TeamMate system also showed to significantly improve the efficiency, e.g. in terms of time to enter the roundabout. The users considered the system as supportive, showing a general appreciation for the cooperative approach explored in the project.

Even if the EVA use case was not focused on safety-critical factors (e.g. no risks of accidents were induced and some safety related parameters, such as time to collision, were not considered as relevant KPIs for this use case), some indications confirm that the approach followed is able to increase the safety; for example, the simulator's data showed a significant reduction of the time needed to take over. Moreover, the reduction of "Safe Manoeuvres" (i.e. when the driver doesn't' respond to a take-over request and the car performs a minimum risk manouver in order to minimize the risks) implies a relevant improvement in terms of safety, since this can be considered as a cause of danger, with a negative impact on traffic congestion and potential risks of accidents.

Some data, even if encouraging, show room for potential improvements: for example, the willingness to pay the system was not significantly high, and the level of trust increased enough to keep the level beyond an acceptability threshold, but without a remarkable increase. These data can be correlated with personal dispositions and limited interest in having automated vehicles. The user sample, in fact, has been recruited to be representative of a generic audience, without considering previous experiences with automated driving features; moreover, questions on dispositional trust (i.e. the innate and individual predisposition to trust in technological – or automated - system) show that the user sample had a low willing to trust automated systems (e.g.

the question "Automated systems generally work well" obtained a neutral score of +0,27, as aggregated result of the item in TeamMate and Baseline modes).

In general, the results of the evaluation performed at REL driving simulator are in line with the project objectives. In particular, the results demonstrate the effectiveness of the TeamMate approach in terms on Interaction strategies (including HMI and input strategies), thanks to the increased acceptance and the improvements in the overall workload (i.e. Objective 1 "Develop solutions for flexible gradual and smooth distribution of tasks between driver and automation to better handle critical driving situations" and Objective 5 "Develop solutions for optimized human-machine interaction"). Moreover, thanks to the enabler "Driver Intention Recognition", the TeamMate system was able to meet the Objective 3 ("Develop solutions allowing the TeamMate car to plan and execute driving manouvers in a human expert-like way"). Finally, thanks to the enabler "Driver Monitoring System", the system was able to constantly monitor the driver state, in order to adapt the level of automation as well as the type and the modality of Human Machine Interface, in order to ensure the adaptiveness to the driver fitness to drive (Objective 2 "Develop solutions to monitor, understand, assess and anticipate the driver, the vehicle and the traffic situations").

4.3 Evaluation of the MARTHA scenario in the driving simulator and on the test-tracks

For the third cycle of evaluation, experiments performed on VED demonstrators aimed at evaluating the TeamMate concept with the MATRHA scenario. This cycle of evaluation included two comparative experiments carried out in parallel in VED driving simulator, and in VED vehicle platform on a test-track. The procedure was replicated in both virtual and physical demonstrators.

Another evaluation of the TeamMate concept was also performed during the final event of the project. This evaluation was done on a test-track with expert drivers with an extended version of the MARTHA scenario involving all the enablers developed in the project.

4.3.1 Participants

Participants that took part in the experiment had to hold a valid driving licence for at least 3 years, be younger than 60 years old and drive at least once a week. Each participant received a gift card of 50 € in exchange for their participation to the experiment.

Driving simulator: 23 drivers agreed to complete the experiment in the driving simulator. Among them, 3 did not complete the experiment due to simulation sickness, and 4 experienced technical issues during experiment so their data were not included in the analysis. The 17 remaining drivers (10 males and 7 females) were in average 37.65 years old (SD = 9.6), obtained their driving license in average 16.88 years ago (SD = 9.55), and drove in average 12941 km per year (SD = 7258).

Vehicle platform: 17 participants (10 males and 7 females) completed the experimentation in the vehicle platform, with an average age of 32.58 years old (SD=9.2). They hold their driving license for 13.76 years in average (SD=9.41), and they drove in average 14117 km per year (SD=7668).

4.3.2 Material

4.3.2.1 Driving simulator

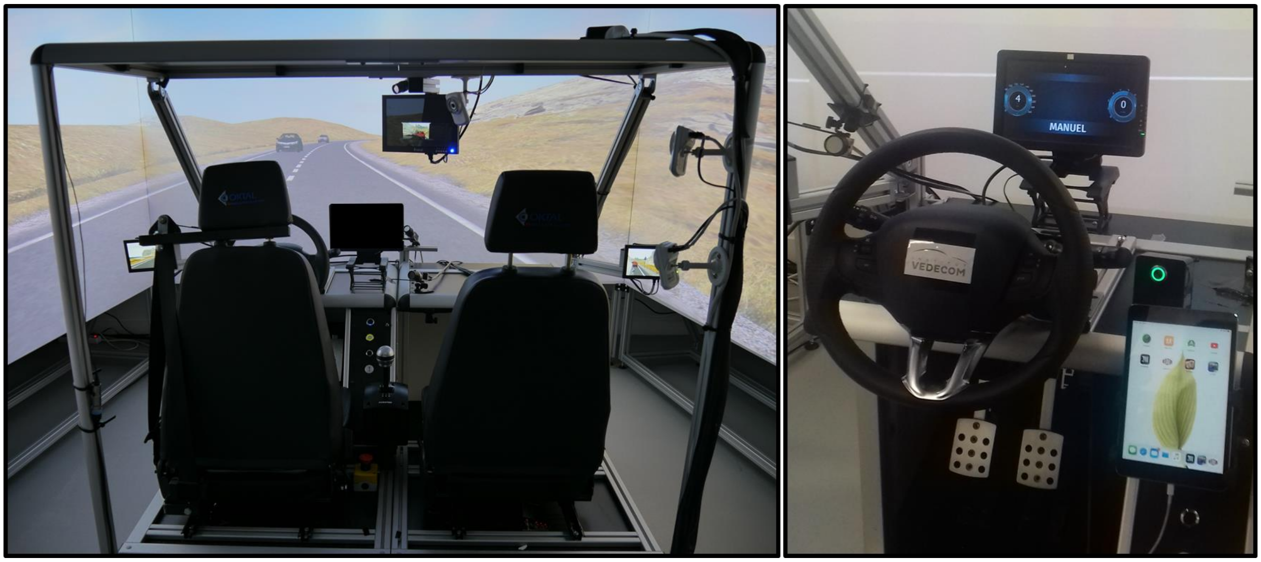


Figure 53: VED driving simulator set-up.

The cabin of the static driving simulator depicted in Figure 53 is composed of a driver seat and a front passenger seat, a force feedback steering wheel and, three pedals. A box equipped with a button is installed on the dashboard right to the steering wheel was used to activate automated mode. The driving environment was displayed on 3 140*240 cm panels offering a 230° horizontal field of view. Rear view was displayed on three separate screens. Lateral mirrors were two screens of 8*5 cm, and the central mirror was a 10*7 cm-screen. A 10'' screen set behind the steering wheel was used as a dashboard. A 7.9 iPad mini, mounted in the central console, was used to perform non-driving-related-task during automated driving. Driving scenarios were generated by SCANER® studio software 1.8 developed by Oktal.

4.3.2.2 Vehicle platform

The vehicle platform was a prototype derived from a series-produced car "C4 – Picasso" manufactured by Peugeot-Citroën (cf. Figure 54). VED has acquired this robotised prototype and equipped it with all necessary sensors and technologies in order to make the autonomous vehicle platform. Among the integrated material: 1 lidar 360° (Velodyne 16 layers, V FoV 30°, 100m range), 6 lidars (H FoV:145°, V FoV:3.2°, 4 layers of 0.8° each, 150m range), 2 radars, 3 cameras (2 in the front and 1 in the back), odometer, central unit, connectivity platform, GPS-RTK and different antennas (GNSS, GPS, 3G/4G). Specific power distribution with all necessary cabling have been realised ingeniously with an integrated relay system.

Figure 54: VED vehicle platform.

As depicted in the Figure 55, the visual HMI was displayed in the dashboard visible in the blue box. The automated mode was activated the same way in both the simulator and vehicle, namely, by using a dedicated button on the middle of the dashboard (cf. Figure 55). An iPad was also installed in the right-hand side of the driver, with an articulated holder in order to adapt the orientation of the screen according to driver's comfort.

Figure 55. Dashboard and HMI of the VED vehicle platform

The vehicle has also been equipped with dual pedal apparatus, similar to a vehicle used by driving schools. This is done not only for security reasons but also to make sure that legislative rules were respected (during the automated mode, an expert driver was aware and able to takeover control in case of unexpected event). Additional rear view mirrors, one on each side, and one central rear mirror, have also been installed to ensure that the expert driver installed in the front passenger car can properly observe and supervise the car's environment.

4.3.2.3 Driving environment

Evaluation with the car demonstrator was carried out on a closed test track in Satory, Versailles. The part of the test track used for the experiment was 3.4 kilometres long and was composed of two lines in the same direction.

The road geometry of the test track was modelized in the driving environment of SCANER studio software so that the road was same between the vehicle and the driving simulator. The itinerary on the test track and on the simulator encompassed zones with different speed limits set at 30, 50 and, 90 km/h. A roadworks zone (cf. Figure 56) was set at the same position in both driving environments.

Figure 56: overview of Satory's test track. The Itinerary used for both experiments is highlighted in green. The zone circled in red is where the roadworks zone was set. The zone circled in blue is the departure and arrival zone.

4.3.2.4 Automated driving system and Human Machine Interface

The automated mode could be activated only if the vehicle was travelling on the right lane in a low speed. For the driving simulator, the speed had to be below the speed limit. In the vehicle demonstrator, participants were asked to lower their speed under 30km/h before activating autonomous mode for security reasons. When those conditions were not met, the vehicle was in manual mode. As depicted in Figure 57, in manual mode the driving mode ("manuel" – "manual"), speed and RPM were displayed on the dashboard.

Figure 57: Screenshot of the dashboard in manual mode.

Once the conditions of activation were met, the vocal message “automated mode available” (“mode Autonome disponible”) was issued and displayed in the centre of the dashboard along with a specific pictogram as depicted in Figure 58. In the vehicle platform, a LED circling around the activation button was flashing green light.

Figure 58: Screenshot of the dashboard in manual mode when conditions of automated mode activation were met.

Automated mode activation was possible by pressing a dedicated button positioned on the right of the steering wheel as described in sections 4.3.2.1

and 4.3.2.2. A vocal message (“mode autonome activé” – “automated mode activated”) was issued to validate the automated mode activation.

In the vehicle platform, the LED circling the activation button becomes full green once automated mode is activated. The automated system maintained the vehicle in the centre of the lane and set speed according to speed limit. The system could be deactivated at any moment by pressing the brake or the accelerator. In automated mode, the system status on the dashboard was set to "automated" ("autonome") (cf. Figure 59).

Figure 59: Screenshot of the dashboard when automated mode was activated.

In case of a takeover request, the vocal message “Takeover Manuel driving” (“reprenez la main”) was issued. A specific pictogram was displayed on the dashboard, and the system status was set to “Takeover Manuel driving” (“reprenez la main”) as depicted in Figure 60. In case the driver could not takeover manual driving on time, the vehicle performed a safe stop.

Figure 60: Screenshot of the dashboard when manual takeover is requested.

4.3.3 Driving scenarios

4.3.3.1 BaseLine condition

After starting the scenario in manual mode in the departure zone (cf. Figure 56), the automated mode became available. After travelling about 1.5 kilometer, an eMail was received in the Mailbox of the tablet, and an acoustic signal was issued. The eMail was a text written in French composed of about 100 words.

When the distance to the roadwork zone was below 40 meters a takeover request was issued. The image displayed in Figure 61 was displayed in the centre of the dashboard when system was requesting manual takeover. After the manual takeover, the automated mode activation was not possible anymore and the course ended in the arrival zone (cf. Figure 56).

Figure 61: Image displayed in the center of the dashboard when the automated system was requesting a manual takeover because of the roadwork zone.

4.3.3.2 TeamMate condition

After starting the scenario in manual mode in the departure zone (cf. Figure 56), the automated mode became available. After travelling about 1.5 kilometer, an eMail was received in the Mailbox of the tablet, and an acoustic signal was issued. The eMail was a text written in French composed of about 100 words.

From the moment the distance to the roadwork was below 500 meters, the vocal message "We are approaching a roadwork zone, get ready to takeover manual driving" ("nous approchons d'une zone de travaux, préparez-vous à reprendre la main") was issued, and a pictogram depicting a roadwork zone was displayed with the distance to the roadwork zone (cf. Figure 62). When the distance to the roadwork zone was below 120 meters, and if the vehicle was still in automated mode a takeover request was issued. The image displayed in Figure 61 was displayed in the centre of the dashboard when system was requesting manual takeover. After the manual takeover, the automated mode activation was not possible anymore and the course ended in the finish zone cf. Figure 56).

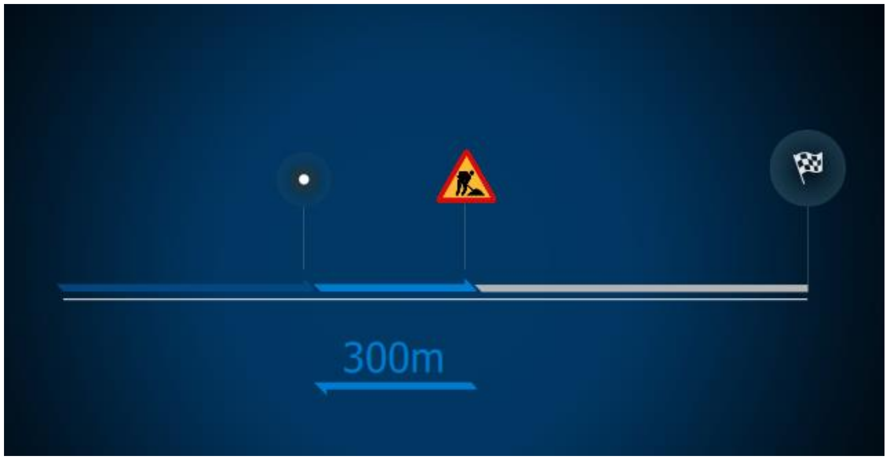


Figure 62: Center of the dashboard when displaying the pictogram depicting the distance to the roadwork zone.

4.3.4 Questionnaires

4.3.4.1.1 Trust questionnaire

Trust in the baseline system and in the TeamMate system was assessed with a French translation of a questionnaire composed of 19 items described in Körber (2018). Level of agreement with each item is assessed on a scale ranging from 1 to 5. Six dimensions of trust are assessed with this questionnaire: reliability of the system, predictability of the system, familiarity with the system, intentions of developers, propensity to trust and trust in automation (toward the system).

4.3.4.1.2 Acceptance questionnaire

A French translation of the questionnaire described by Van Der Laan, Heino and De Waard (1997) was used to compare baseline to TeamMate acceptance. This questionnaire aims at assessing two dimensions of automotive technologies acceptance: perceived usefulness and satisfaction, with nine items. Each item is composed of one scale ranging from -2 to 2 with a pair of opposed adjectives (ie. "*useful*" versus "*useless*", or "*assisting*" versus "*worthless*"). According to the calculation procedure described in Van Der Laan, Heino and De Waard (1997), this questionnaire provides a perceived usefulness score and a satisfaction score ranging from -2 to 2. Higher score indicates higher perceived usefulness and higher satisfaction.

4.3.4.1.3 NASA-TLX

Driver workload during manual takeover was assessed using a French translation of the NASA-TLX (Hart & Staveland, 1988). This questionnaire is composed of six scales ranging from 0 to 100. Each scale aims at evaluating a dimension of workload: mental demand, physical demand, temporal demand, performance, effort and frustration linked to the completion of a specific task.

4.3.4.1.4 System Usability Scale

Systems usability was assessed using a French translation of Brooke's (1996) System Usability Scale (SUS) composed of 10 items. Level of agreement with each item is evaluated on a scale ranging from 1 to 5. According to the calculation procedure provided in Brooke (1996), this questionnaire provides a usability score ranging from 1 to 100. Higher score means higher rates of usability.

4.3.4.1.5 Willingness to buy

Participants' willingness to buy a vehicle equipped with the baseline system and the TeamMate system was assessed by mean of a scale ranging from 1 to 5. Participants were asked if they would buy the vehicle equipped with the TeamMate system and with the baseline system. Responses were collected by mean of two scales ranging from 1 to 5, with 1 corresponding to a lower willingness to buy and 5 corresponding to a higher willingness to buy. Additionally, a scale ranging from 0 € to 50 000€ was used to evaluate how much money participants were willing to spend to purchase the system in addition to the price of the vehicle.

4.3.4.2 Procedure

After welcoming the participants, the aim of the experiment and the different steps of the procedure were explained. Risks and constrains were explicitly enlightened, indicating that they could stop the experiment at any moment. They were asked to read and sign an informed consent, and to answer a questionnaire to collect sociodemographic data.

After that, a description of vehicle functionalities was exposed along with a description of HMI states. Drivers were told that they will drive two automated cars which offer two driving modes: manual and automated.

It was said that in manual mode, the driver was responsible for the entire vehicle control. Automated mode was described as working thanks to sensors that allowed the vehicle to perceive the environment, localise itself and other objects, as well as to control vehicle’s trajectory. Activations conditions have then been explained, they were also told that they would be informed if activations conditions were validated by means of a vocal message and the display of a message and a pictogram on the dashboard. Then participants were shown the automated mode activation button, and explained that once activated, another vocal message would announce the validation of automated mode activation, and that they would have to release pedals and steering wheel.

Participants were explained that in automated mode, the vehicle would maintain itself on the center of the lane and would adapt speed according to speed limit and to other vehicles. The procedure to deactivate automated mode was explained to the participant. Experimenter said that participants did not have to monitor the road environment during automated mode activation, and that they were free to engage in any task on the tablet.

Participants were informed that the automated system was not able to deal with all the driving situations and that if, for example, the infrastructure was damaged or, if the driving situation was too complex, then the system would issue a takeover request to give back vehicle control to the driver. Participants were told that if the system could not be deactivated on time, it would perform a safe stop.

The participant carried on with the training scenario which allowed them to drive the car in manual mode. Participants were instructed to perform lane changes and to use the brake pedal as many times as they needed to get used to simulator's or car’s command. Then, automated mode became available and participants were instructed to activate it. During automated mode activation experimenter commented on the system functioning and instructed the participant to deactivate the automated system with each pedal. Driver could test the activation/deactivation process as many times as they needed.

Afterwards, the two driving scenarios (BaseLine and TeamMate) were completed in a counterbalanced order. Before each scenario, participants were instructed to comply with speed limit, to drive in the right lane as often as they could and, to activate the automated system as soon as it would be available. Participants were reminded that during the activation of the automated mode, the monitoring of the road environment was not required, and that they were free to use the tablet. Additionally, the participants were instructed to open and read aloud eMail in case they would receive one.

After each scenario, participants responded to questionnaires to evaluate trust, acceptance, mental workload, usability and, willingness to buy automated system.

The experiment ended with a semi-structured interview and participants were given the financial allowance. The whole protocol lasted about 1.5 hours.

The experimental design was a mixed design, with the factor "type of demonstrator" (simulator versus vehicle platform) as a between-subject variable and the factor "system" (BaseLine versus TeamMate) as a within-subject variable.

4.3.5 Results

4.3.6.1. Data analysis

Among the 6 KPIs analysed, 4 were related to the subjective evaluation of the system by the participants and were assessed by means of questionnaires described in the Material section:

- Acceptance: from -2 (lower acceptance) to 2 (higher acceptance).

- Trust: from 0 (lower trust) to 5 (higher trust)

- Usability: from 0 (lower) to 100 (higher)

- Willingness to buy and to pay: from 0 (lower) to 5 (higher) and from 0 to 50000 euros

Another KPI concerned the auto evaluation of the mental workload during manual takeover as measured by the NASA-TLX which gives a rating from 0 (lower) to 100 (higher) for each dimension of this questionnaire.

The last KPI was related to the safety during the manual takeover caused by the roadworks zone: the minimum time to collision (TTC) with the roadwork zone.

An ANOVA was conducted for each KPI; with the factor "type of demonstrator" (simulator versus vehicle platform) as a between-subject variable and the factor "system" (BaseLine versus TeamMate) as a within-subject variable. The minimum TTC in the BaseLine condition was compared to the minimum TTC in the TeamMate condition by means of two paired student tests.

4.3.5.1 Subjective evaluation of the systems

4.3.5.1.1 Acceptance

Responses to the acceptance questionnaires (cf. Figure 63) revealed a slightly higher acceptance for the TeamMate as compared to the Baseline for both the simulator (BaseLine: 1.32 ± 0.49; TeamMate: 1.37 ± 0.81) and the vehicle platform (BaseLine: 0.94 ± 0.56; TeamMate: 1.28 ± 0.31). However, statistical analysis did not point out any significant effect of the type of the demonstrator ($F(1,32) = 2.05$; p=.16), system ($F(1,32) = 3.37$; p=.75), nor any significant interaction between the two factors ($F(1,32) = 1.82$; p=.19).

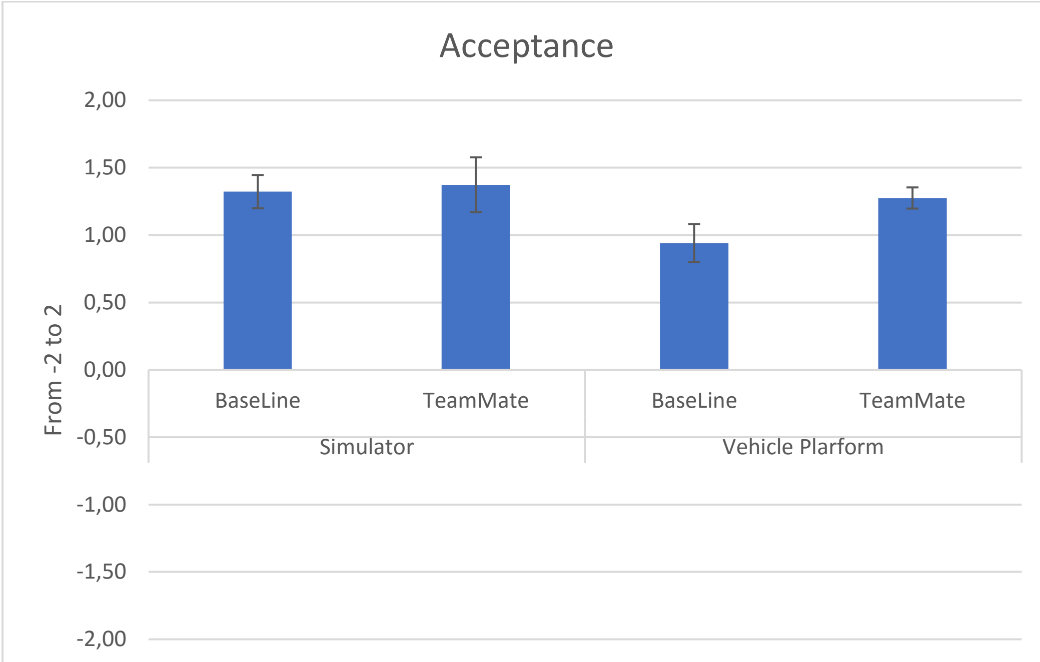


Figure 63: Mean acceptance as a function of type of demonstrator (simulator versus vehicle platform) and system (BaseLine versus TeamMate). Error bars stand for standard-error.

4.3.5.1.2 Trust

A slightly higher trust for the TeamMate system can be noted (cf. Figure 64), for both the simulator (BaseLine: 3.37± 0.66; TeamMate: 3.50 ± 0.59) and the vehicle platform (BaseLine: 3.06 ± 0.59; TeamMate: 3.23 ± 0.48). However, statistical analysis did not point out any significant effect of the type of demonstrator ($F(1,32) = 0.37$; $p=.55$), system ($F(1,32) = 0.37$; $p=.13$), nor any significant interaction between the two factors ($F(1,32) = 1$; $p=.32$).

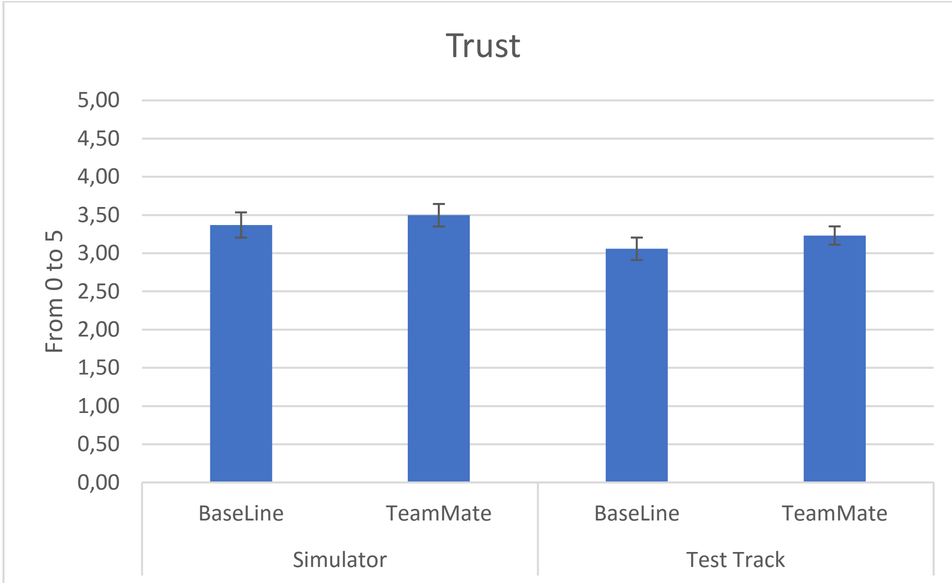


Figure 64: Mean trust as a function of type of demonstrator (simulator versus vehicle platform) and system (BaseLine versus TeamMate). Error bars stand for standard-error.

4.3.5.1.3 Usability

Results (see Figure 65) show a slightly higher usability for the TeamMate car than the baseline car for both the simulator (BaseLine: 84.26 ± 12.86; TeamMate: 87.79 ± 10.57) and the vehicle platform (BaseLine: 80.15 ± 11.97; TeamMate: 81.32 ± 13.52). However, statistical analysis did not point out any significant effect of type of demonstrator ($F(1,32) = 1.42$; $p=.24$), system ($F(1,32) = 0.61$; $p=.44$), nor any significant interaction between the two factors ($F(1,32) = 0.01$; $p=.94$).

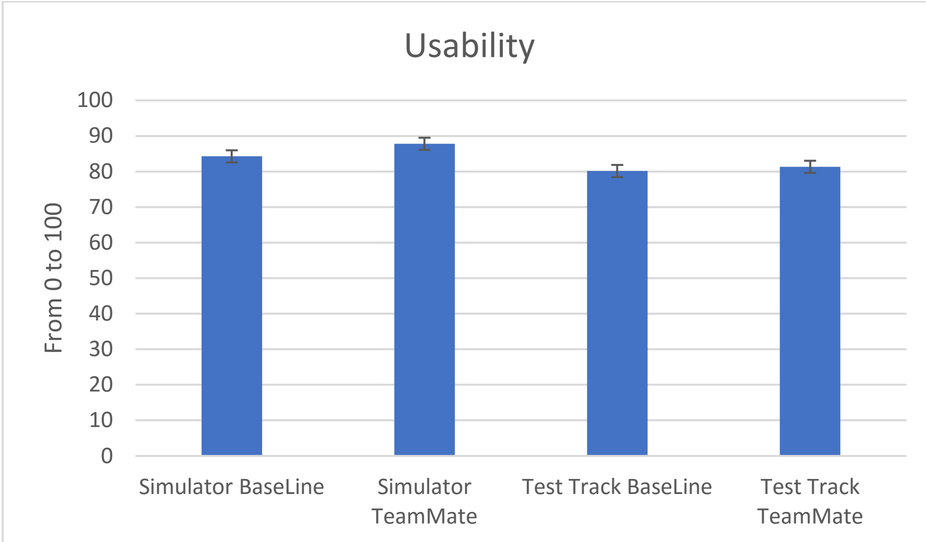


Figure 65: Mean Usability scores as a function of type of demonstrator (simulator versus vehicle platform) and system (BaseLine versus TeamMate). Error bars stand for standard-error.

4.3.5.1.4 Willingness to buy

Willingness to buy the TeamMate system is slightly higher compared to the willingness to buy the BaseLine system (cf. Figure 66). This effect is statistically significant ($F(1,32) = 9.17$; $p<.01$) and is observed in the simulator (BaseLine: 3.35 ± 1.17; TeamMate: 3.82 ± 1.19) and in the vehicle platform (BaseLine: 3.12 ± 1.36; TeamMate: 3.47 ± 1.28). No significant effect of the type of demonstrator is observed ($F(1,32) = 0.52$; $p=.48$), nor any significant interaction between the two factors ($F(1,32) = 0.19$; $p=.67$).

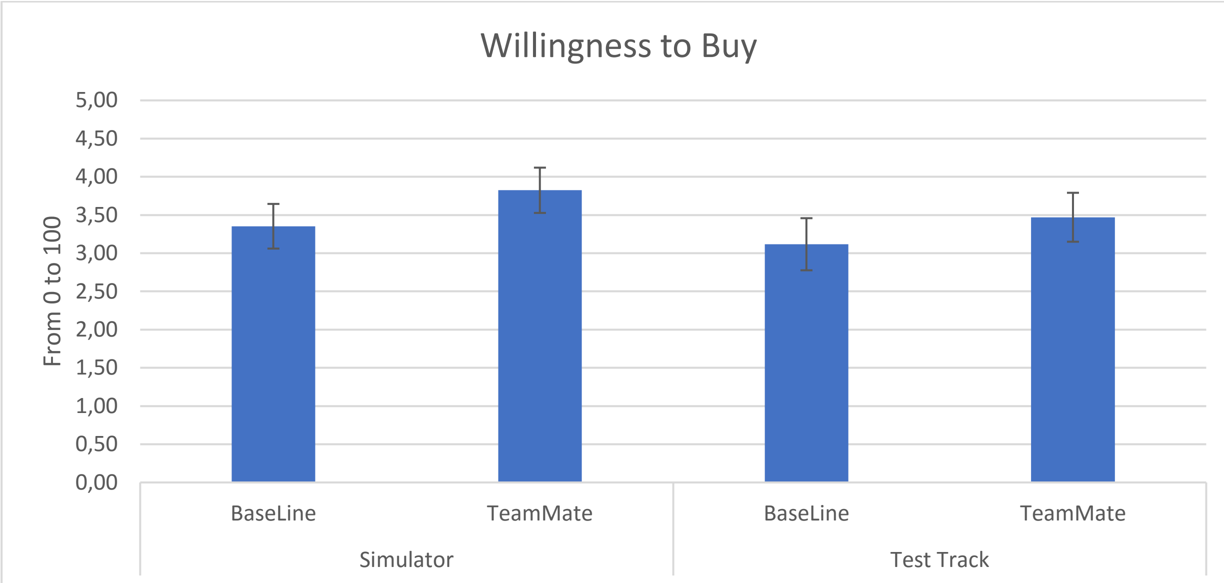


Figure 66: Mean willingness to buy as a function of type of demonstrator (simulator versus vehicle platform) and system (BaseLine versus TeamMate). Error bars stand for standard-error.

The amount that the participants were willing to spend to have a new personal vehicle equipped with the system was superior for the TeamMate system compared to the BaseLine system (cf. Figure 67). This was observed in the simulator (BaseLine: 9775 ± 9978; TeamMate: 10741 ± 9615) and in the vehicle platform (BaseLine: 7647 ± 6264; TeamMate: 8412 ± 5546). The effect of the system is significant ($F(1,32) = 8.57$; $p<.01$), while the effect of the type of demonstrator ($F(1,32) = 0.5$; $p=.48$) and the interaction effect are not significant ($F(1,32) = 0.97$; $p=.33$).

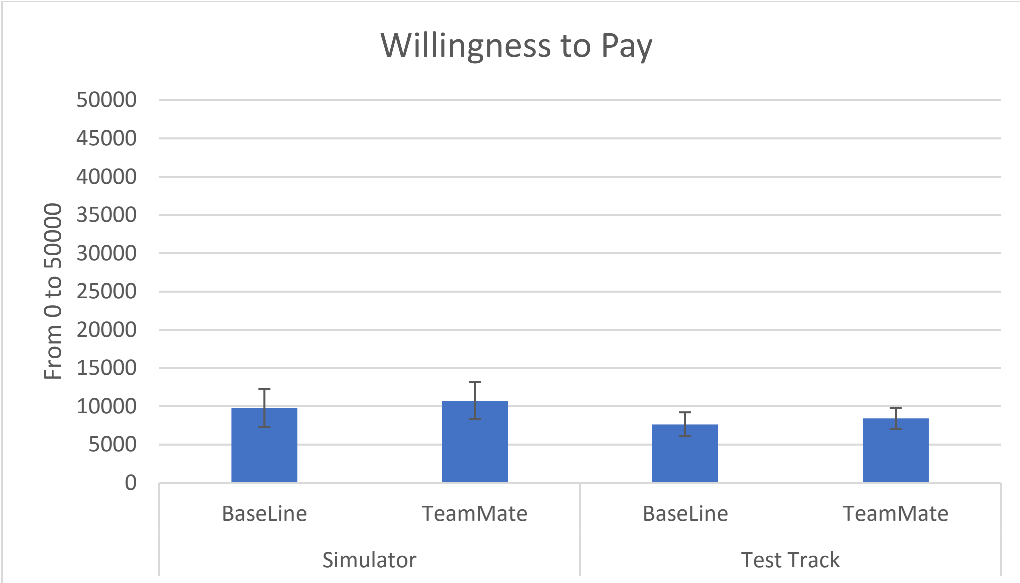


Figure 67: Mean willingness to pay as a function of type of demonstrator (simulator versus vehicle platform) and system (BaseLine versus TeamMate). Error bars stand for standard-error.

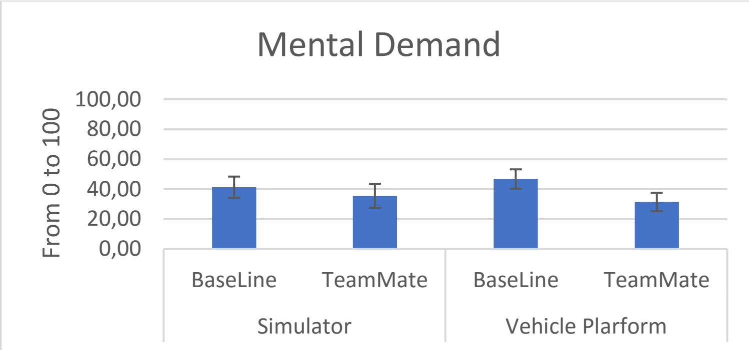
4.3.5.2 Mental workload

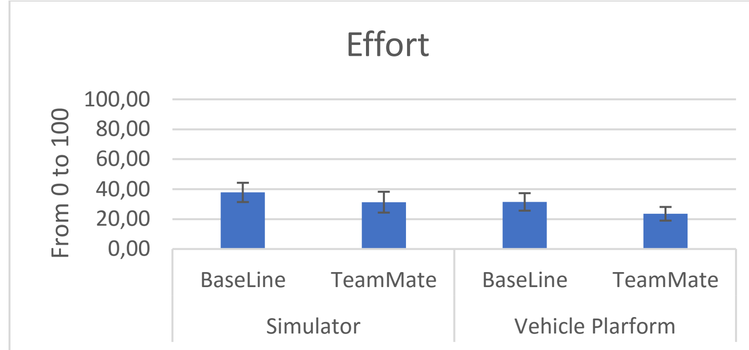
The six dimensions of the mental workload were analysed separately. Mental demand, physical demand, temporal demand, effort and frustration during manual takeover were rated as higher in the BaseLine condition compared to the TeamMate condition, in the simulator and in the real car. Besides, performance, measured in the simulator and in the real car, in the TeamMate condition is rated as higher compared to the BaseLine condition. Mean values and standard deviation for each dimension is presented in Table 6 and in Figure 68.

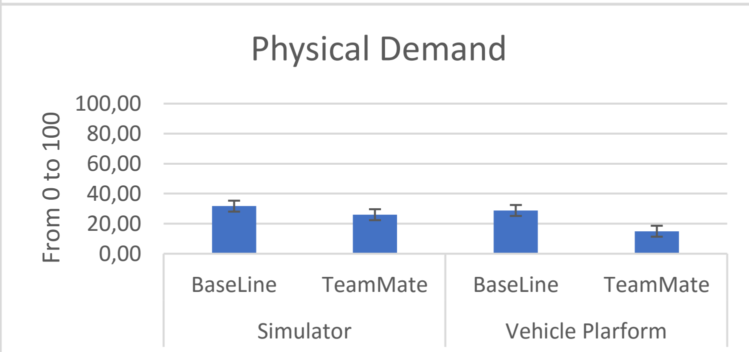
Statistical analysis revealed no significant effect of the type of demonstrator (mental demand: ($F(1,32) = 0.01$; $p=.94$); physical demand: ($F(1,32) = 1$; $p=.32$); temporal demand: ($F(1,32) = 3.15$; $p=.09$); performance: ($F(1,32) = 0.11$; $p=.74$); effort: ($F(1,32) = 1.03$; $p=.32$)), except for the factor frustration ($F(1,32) = 4.76$; $p<.05$), with a higher value for the vehicle platform. Besides, a significant effect of system is observed for the mental demand ($F(1,32) = 4.43$; $p<.05$), the physical demand ($F(1,32) = 9.4$; $p<.01$), the temporal demand ($F(1,32) = 11.07$; $p<.01$), and the performance ($F(1,32) = 5.83$; $p<.05$), but not for the effort ($F(1,32) = 2.63$; $p=.11$) nor the frustration ($F(1,32) = 0.62$; $p=.44$). No significant interaction between the two factors is observed for any dimension of the NASA-TLX(mental demand: ($F(1,32) = 0.91$; $p=.35$); physical demand: ($F(1,32) = 1.62$; $p=.21$); temporal demand: ($F(1,32) = 0.09$; $p=.77$); performance: ($F(1,32) = 1.55$; $p=.22$); effort: ($F(1,32) = 0.03$; $p=.88$); frustration: ($F(1,32) = 0.26$; $p=.61$)).

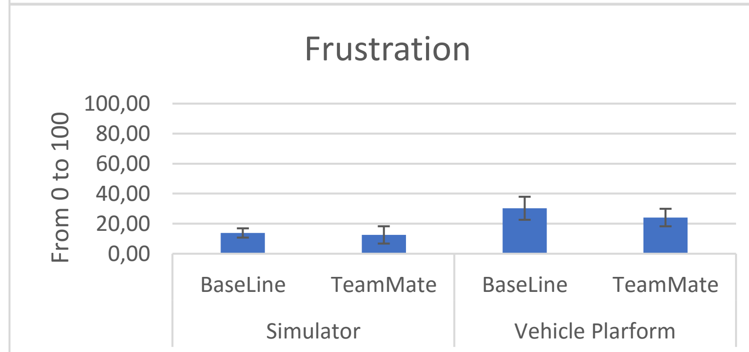
	Simulator		Real car	
	BaseLine	TeamMate	BaseLine	TeamMate
Mental Demand	41.35 ± 28.15	35.59 ± 31.87	46.76 ± 25.80	31.47 ± 24.67
Physical Demand	31.71 ± 25.28	26 ± 27.86	28.82 ± 19.33	15 ± 13.81
Temporal Demand	43.12 ± 32.42	23.94 ± 26.04	58.82 ± 28.59	35.88 ± 29.65
Effort	37.82 ± 25.71	31.29 ± 27.85	31.47 ± 23.37	23.53 ± 18.27
Frustration	13.82 ± 12.32	12.53 ± 23.08	30.29 ± 30.69	24.12 ± 23.33
Performance	82 ± 12.78	83 ± 10.53	81.18 ± 11.25	87.06 ±10.32

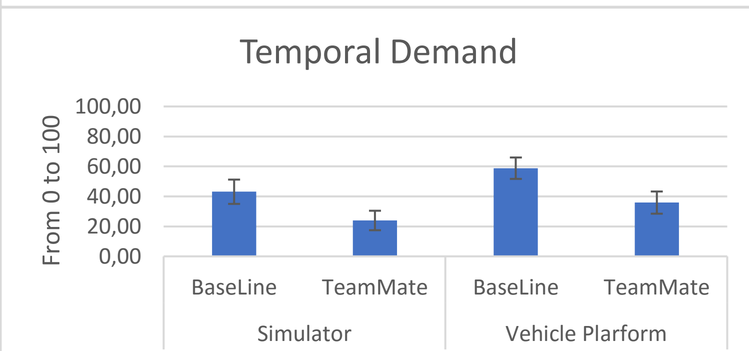
Table 6: Mean values and standard deviation for each dimension of the NASA-TLX according to experimental conditions.


Mental Demand
From 0 to 100
100,00
80,00
60,00
40,00
20,00
0,00
BaseLine
TeamMate
BaseLine
TeamMate
Simulator
Vehicle Plarform


Effort
From 0 to 100
100,00
80,00
60,00
40,00
20,00
0,00
BaseLine
TeamMate
BaseLine
TeamMate
Simulator
Vehicle Plarform


Physical Demand
From 0 to 100
100,00
80,00
60,00
40,00
20,00
0,00
BaseLine
TeamMate
BaseLine
TeamMate
Simulator
Vehicle Plarform


Frustration
From 0 to 100
100,00
80,00
60,00
40,00
20,00
0,00
BaseLine
TeamMate
BaseLine
TeamMate
Simulator
Vehicle Plarform


Temporal Demand
From 0 to 100
100,00
80,00
60,00
40,00
20,00
0,00
BaseLine
TeamMate
BaseLine
TeamMate
Simulator
Vehicle Plarform

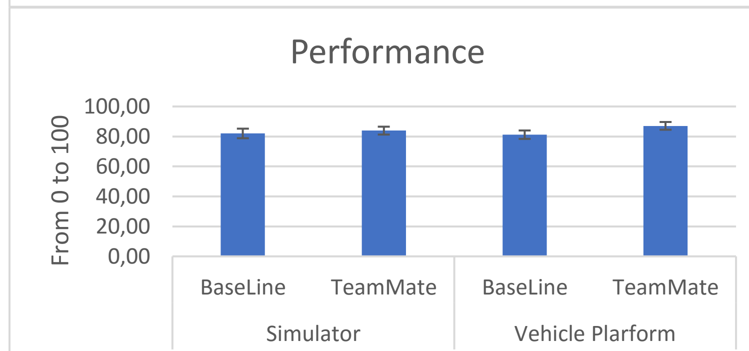

Performance
From 0 to 100
100,00
80,00
60,00
40,00
20,00
0,00
BaseLine
TeamMate
BaseLine
TeamMate
Simulator
Vehicle Plarform

Figure 68: Mean values and standard deviation for each dimension of the NASA-TLX as a function of type of demonstrator (simulator versus vehicle platform) and system (BaseLine versus TeamMate). Error bars stand for standard-error

4.3.5.3 Minimum Time to Collision with the roadwork zone

Minimum TTC with the roadwork zone is longer with the TeamMate system (Simulator: 3.72 ± 4.29; vehicle platform: 3.96 ± 2.05) compared to the BaseLine system (Simulator: 0.18 ± 0.31; vehicle platform: 3.45 ± 1.31) (cf. Figure 69 and Figure 70). This difference is significant in the simulator (t(16) = 3.21; $p < 0.01$) but not in the vehicle platform (t(16) = 1.65; $p = 0.12$).

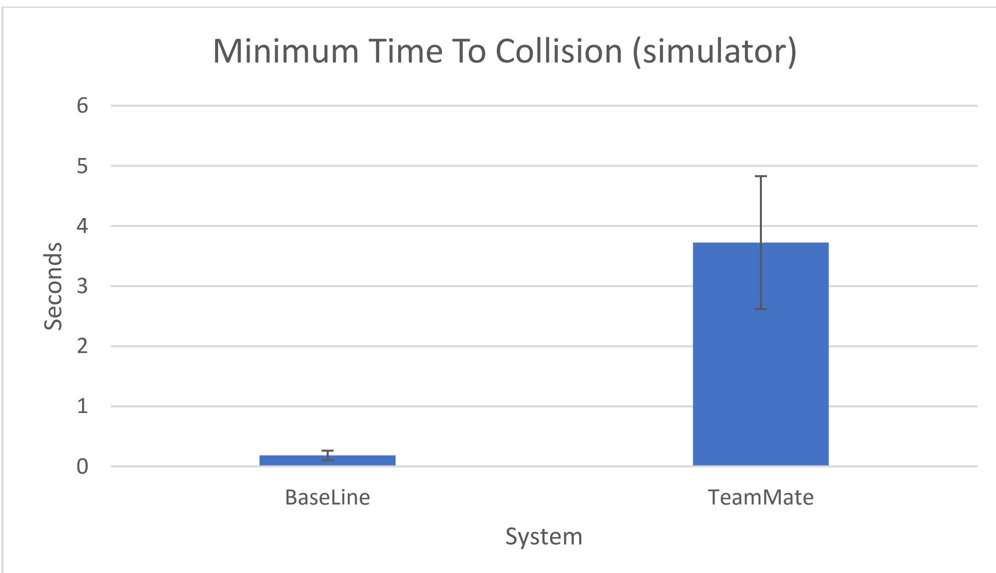


Figure 69: Minimum Time To Collision as a function of system (BaseLine versus TeamMate) for the simulator. Error bars stand for standard-error.

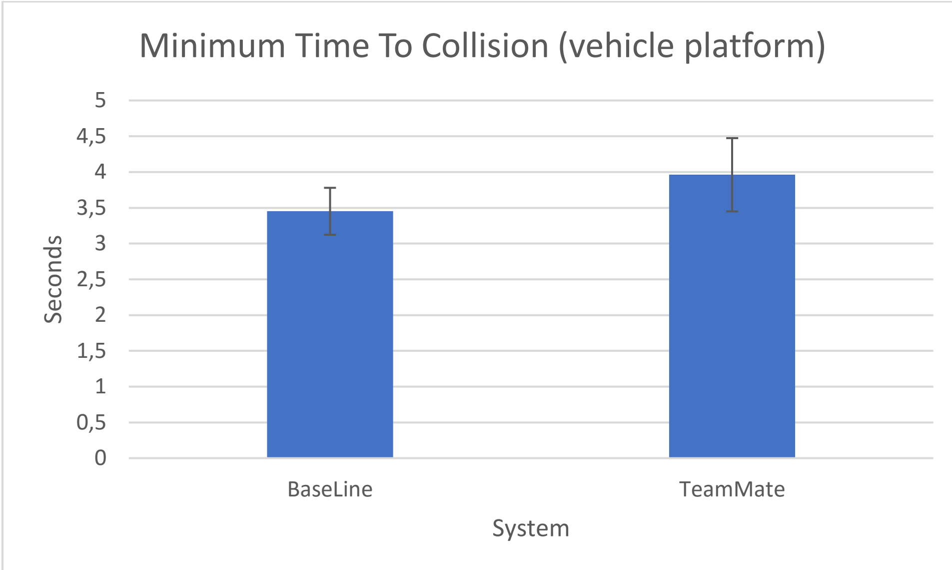


Figure 70: Minimum Time To Collision as a function of system (BaseLine versus TeamMate) for the vehicle platform. Error bars stand for standard-error.

4.3.6 Discussion

Subjective evaluation of the systems revealed that acceptance, trust and usability were not significantly higher for the TeamMate system. However, a preference for the TeamMate system was expressed by the participants during the interviews and the willingness to buy was significantly higher for the TeamMate system. Indeed, the takeover request of the BaseLine system was often described as an experience eliciting negative emotions by the participants, evoking "panic" (Participant 17 – simulator) and "surprise" (Participant 02 – simulator). This manual takeover request was even described as "brutal" and "stressful" (Participant 02 – simulator).

Besides, manual takeover with the BaseLine system was perceived as an event with a high temporal pressure and demanding a fast-cognitive processing of a complex situation "At a certain point, I stopped paying attention and, all of a sudden, the system requested me to take over. It was a little bit fast taking into account what was happening. We were already very close to the roadwork zone. I used the indicators and changed line to the left. I thought "Damn! I'm not sure that I checked the mirrors before changing lanes" (Participant 7 - vehicle platform). This effect is reflected in the responses participants gave to the NASA-TLX which revealed a significantly higher workload during the manual takeover with the BaseLine system. When asked what they disliked in the BaseLine system, some participants reported that the system « did not give a high enough feeling of safety » (Participant 3 - simulator)

On the contrary, the announcement of the incoming roadwork zone and the display of the distance to it by the TeamMate system was acknowledged by the drivers as a useful piece of information that helped them to reengage progressively in the driving activity "It warned me about the roadwork 500 meters in advance, which was really convenient. That way, I could really get ready for the roadwork zone." (Participant 23 - simulator). Moreover, TeamMate system allowed an increased safety during manual takeover with a longer minimum time to collision, even if this effect was not significant in the vehicle platform.

The TeamMate notification announcing the upcoming takeover request was appreciated but some participants criticized the fact that this event was too early taking into account the speed of the vehicle at this moment (30 km/h): "It warned me about a roadwork zone really in advance. So much in advance that it seemed to last forever because it was driving really slow to go there" (participant 8 – Simulator). Also, some participants seem to be confused by the first announcement of the upcoming roadwork zone as they did not understand if they could takeover manual driving immediately or if they had to wait for a formal takeover request « It [the manual takeover] was ok, but it was long. I did not know if I could take over right away » (Participant 2-Vehicle platform).

Besides, some participants pointed out limitations of the TeamMate system by noticing that late takeover requests will still be possible, for example if a non-planned critical event happens: "if, for example, there is an accident occurring immediately in front of me, how will the system react?" (Participant 2-simulator).

To summarize, the manual takeover request strategy rolled out by the TeamMate system allowed to reduce the driver's workload during manual takeover and to increase safety with a higher minimum time to collision. It allowed to increase the driver's willingness to buy the automated system. However, further research is needed in order to find the right timing for the requests as this parameter was criticized by participants.

4.4 Final evaluation Martha Scenario

The final evaluation has been performed on the VEDECOM vehicle platform where all enablers have been integrated. The goal of this evaluation, performed during the final event, is to show the benefits of each enabler and how they can all interact together in order to support drivers in their actions and improve the safety of the TeamMate vehicle.

The evaluation is performed with questionnaires on passive passengers using the extended version of Martha Scenario described in section 4.4.1.2. Indeed, VEDECOM expert is the only authorized driver during the IV conference for safety reasons. In fact, the site of test tracks was very crowded with vehicles and the short time for each demonstration did not allow enough time for a learning phase of each single participant.

4.4.1 Method

4.4.1.1 Participants

Participants were experts (engineers and researchers) from IEEE-IV conference (The Institute of Electrical and Electronics Engineers- Intelligent Vehicles Symposium, June, 12th, 2019 at Satory, Versailles), who have agreed to evaluate the “TeamMate” vehicle. They are from all around the world: France, China, Netherlands, Germany, Spain, Israel, Japan, Sweden, Italy, and the United States of America.

43 expert participants in total completed this final evaluation of the TeamMate vehicle (41 males, 2 females) with an average age of 37 years old (SD=11), and they hold their driving license on average for 15.5 years (SD=11). Participants were asked to evaluate their knowledge of Autonomous Vehicles by means of a five level Likert Scale (from 1=Very Low, to 5= Very High). They reported in average a level of knowledge of 4.3 (SD = 0.63).

4.4.1.2 Extended Martha scenario

Martha is driving the TeamMate car in manual mode when she encounters a slower vehicle. The intention recognition function learnt that Martha is willing to overtake. The online risk assessment evaluates the maneuver and communicates with Martha through the Augmented Reality HMI and informs her when it is safe to overtake. Afterwards, Martha looks for information on her iPad, therefore the DMS detects that she is distracted, and the automated mode activation is suggested. Martha activates the automated mode and can engage in non-driving related tasks. Thanks to V2I communication, TeamMate detects in advance an upcoming roadwork zone and asks Martha to overtake manually. The early takeover request allows a comfortable manual takeover and a safe avoidance of the roadwork zone.

This scenario was possible thanks to the integration of the following enablers:

- Driver Intention Recognition detects the driver’s intention to overtake a slower vehicle
- Online-Risk assessment to detect that the maneuver is not safe
- V2I communication system to communicate the roadworks
- HMI to warn Martha that the maneuver is not safe
- Augmented reality to support the driver in performing the maneuver
- Driver Monitoring System (DMS) to detect the distraction of the driver
- Interaction Modality to facilitate the safe and robust hand-over of vehicle control

4.4.1.3 Material

The vehicle platform described in (4.3.2.2) has been used for this final evaluation.

4.4.1.3.1 Driving Environment

This evaluation has been carried out in Satory test tracks (Versailles in France), on the so-called speed test track, which consists of straight road design for 2km long (Figure 71). This track has only one important curve of a very large radius that ends with a roundabout. Roadworks were installed on the way back (after negotiating the roundabout), almost at the end of the circuit (cf. Figure 72). The exact location of the roadwork zone is highlighted in orange in Figure 71. The itinerary on the test track is composed of zones with speed limits set at: 80, 50, 70 and 30 in the roadwork zone.

Figure 71. Overview of "Speed Test Track"- Satory, Versailles

Figure 72. Roadworks zone

4.4.1.4 Automated Driving System and Human Machine Interface

Automated driving system and the human machine interface were the same as described in (4.3.2.4). The human machine interface was translated in English for international participants as shown in Figure 73, Figure 74 and Figure 75.

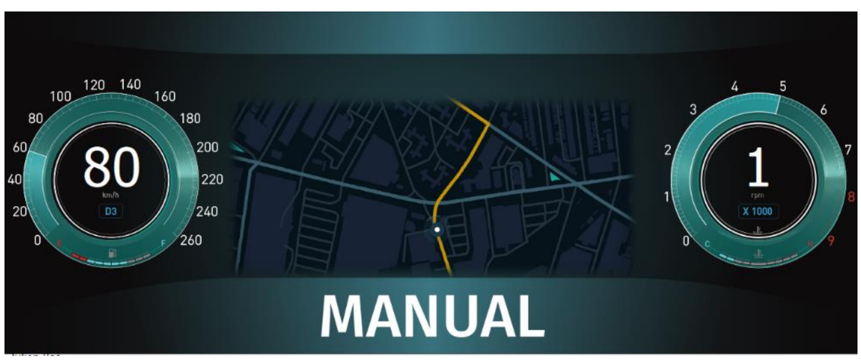


Figure 73: Screenshot of the dashboard in manual mode

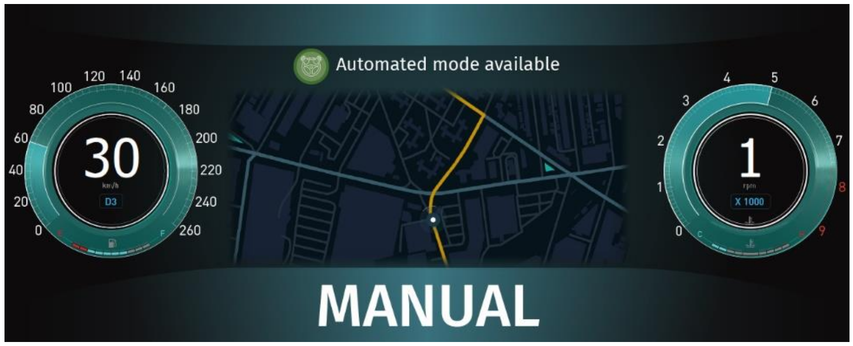


Figure 74: Screenshot of the dashboard in manual mode when conditions of automated mode activation were met

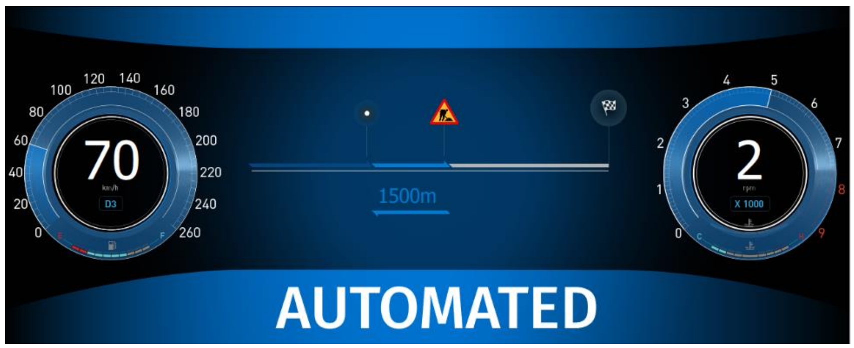


Figure 75: Screenshot of the dashboard when automated mode is active and when approaching Roadwork zone.

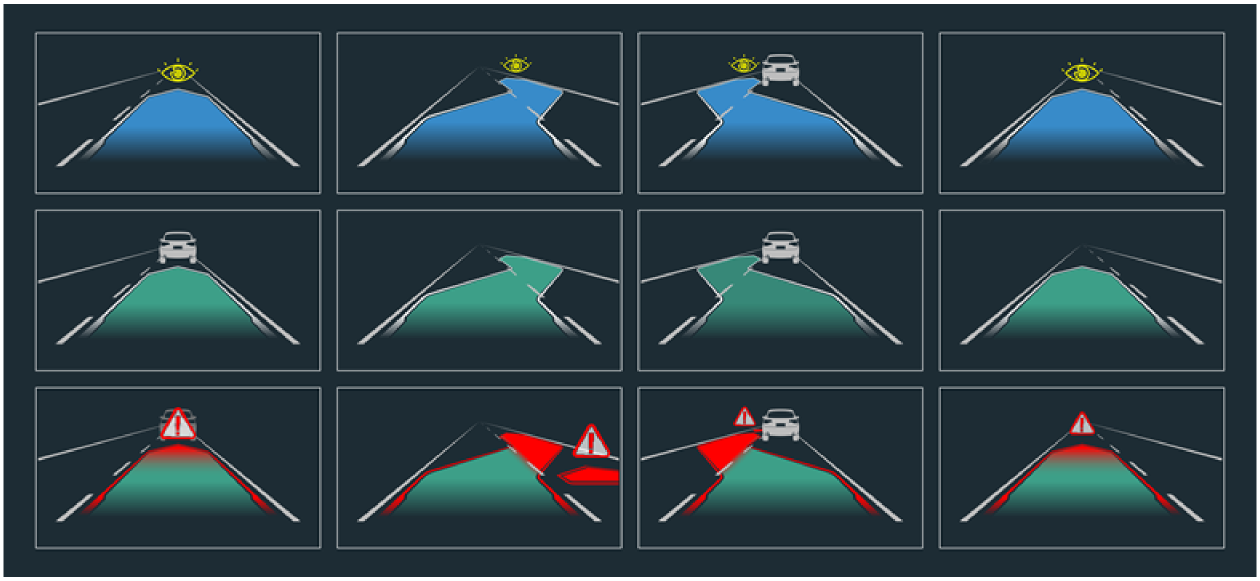


Figure 76: Information displayed on Augmented Reality Glasses.

Augmented Reality Glasses are used to assist the driver for lane change to overtake another vehicle and to come back to the right lane. When the driver intention detection knows that the driver is willing to overtake another vehicle a safety corridor is displayed, using the Risk assessment module, to inform the driver about the safety of the maneuver (cf. Figure 76).

4.4.1.4.1 **Questionnaires**

Subjective evaluation of the TeamMate car has been performed with questionnaires. Trust, acceptance and usability questionnaire (described in 4.3.4) have been used to evaluate the whole system (integrating all enablers of the AutoMate project). Afterwards, participants were asked to evaluate their satisfaction for each enabler on a five level Likert Scales (from 1=Not Satisfied, to 5= Very Satisfied).

4.4.1.5 Procedure

For each driving session, two participants were invited in the TeamMate Vehicle (one in the front passenger seat and one in the back-passenger seat).

Augmented Reality Glasses were only evaluated by drivers seated on the front passenger seat, this is why the evaluation study has been performed by less participants than all the other enablers.

First, a description of the AutoMate Project was presented to participants, and they were informed that they will be asked to evaluate the vehicle at the end of the session. The whole Martha scenario was then performed as described in section 4.4.1.2. The first part of the scenario is in manual mode. During this part of the scenario the goal is to show how the Driver Intention Recognition detects the driver's intention to overtake a slower vehicle, and at the same time give the feedback to the driver from Online-Risk assessment, which shows if the maneuver is safe or not. The person who is seated in the front passenger seat wore Augmented Reality Glasses which showed pictograms that indicated whether it was safe to change the lane or not (for overtaking and then to get back on the right lane). It has been explained to participants that those glasses were a proof of concept, and the goal of the Augmented Reality information shown was to support the driver in performing the overtaking maneuver.

The automated mode is activated just before the roundabout when entering the "delegation zone", the HMI informs the driver that "Automated Mode is available". After pushing the dedicated button, the automated mode was activated. An embedded map allowed to retrieve all needed traffic information. During the driving scenario, the expert driver took over control of the vehicle just after the roundabout. After some time, the expert driver became intentionally distracted, thus the Driver Monitoring System (DMS) detected the distraction of the driver and the system requested the driver through a vocal message to activate "Automated Mode" (Vocal Message: "You are distracted, please activate Automated Mode"). The VEDECOM expert driver activated the Automated Mode, the system stops giving feedback to the driver about the distraction state. When approaching the roadwork zone, the communication platform (communication between vehicle and infrastructure V2I) allowed to know the exact position and shares this information with the TeamMate connected vehicle. The TeamMate requested the driver to take over control of the vehicle. The expert driver did not take overcontrol, therefor the TeamMate vehicle performs a safety maneuver to stop the car in security, waiting for the driver to take overcontrol.

After the scenario, participants were asked to fill the questionnaires described in section 4.4.1.4.1.

4.4.2 Results

4.4.2.1 Acceptance

The average usefulness level (evaluated from -2 to 2) of the TeamMate system is 0.96 (SD = 0.58), the satisfaction level is 0.89 (SD = 0.69) and, the overall acceptance is 0.93 (SD = 0.60) (cf. Figure 77).

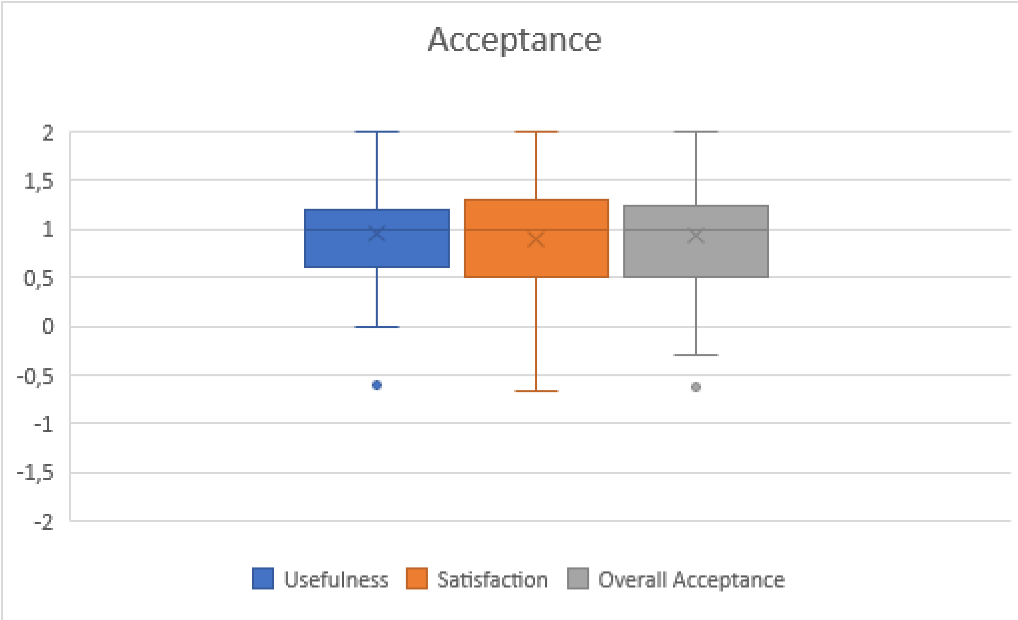


Figure 77. Boxplot of the acceptance

4.4.2.2 Usability

Usability evaluation of the TeamMate vehicle is performed in a scale from 0 (lower) to 100 (higher). Results show an average value of 61.96 with a SD of 17.31 (cf. Figure 78).

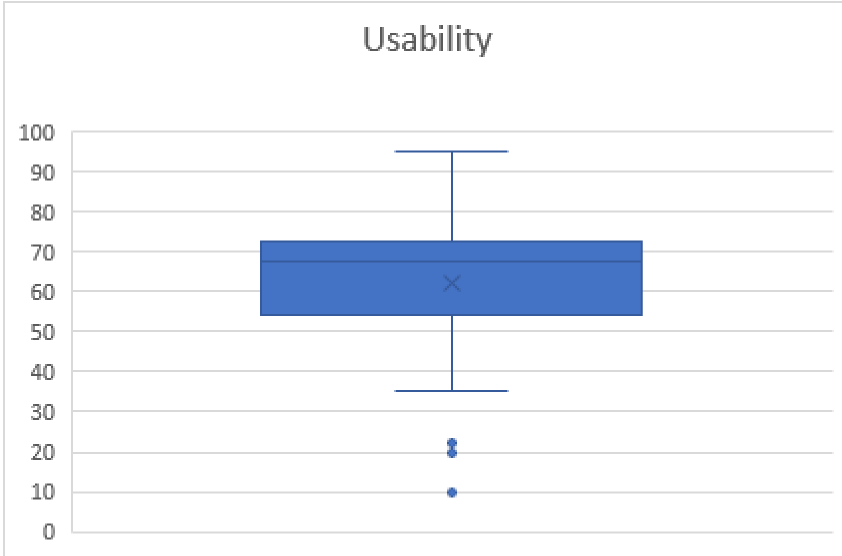


Figure 78: Boxplot of the Usability.

4.4.2.3 Trust

Mean score and standard deviations for each subscale of the trust questionnaire as well as the overall trust (mean of each subscale) are presented in Table 7. A graphical representation of those scores are presented in Figure 79.

Trust subscale	Mean	Standard Deviation
Reliability	3.28	0.59
Understanding	3.62	0.42
Familiarity	3.38	1.10
Intention of developers	3.60	0.67
Propensity to trust	2.9	0.70
Trust in automation	3.34	0.81
Overall trust	3.36	0.43

Table 7: : Mean value and standard deviation for each subscale of the trust questionnaire. Overall trust stands for the mean of the others subscale.

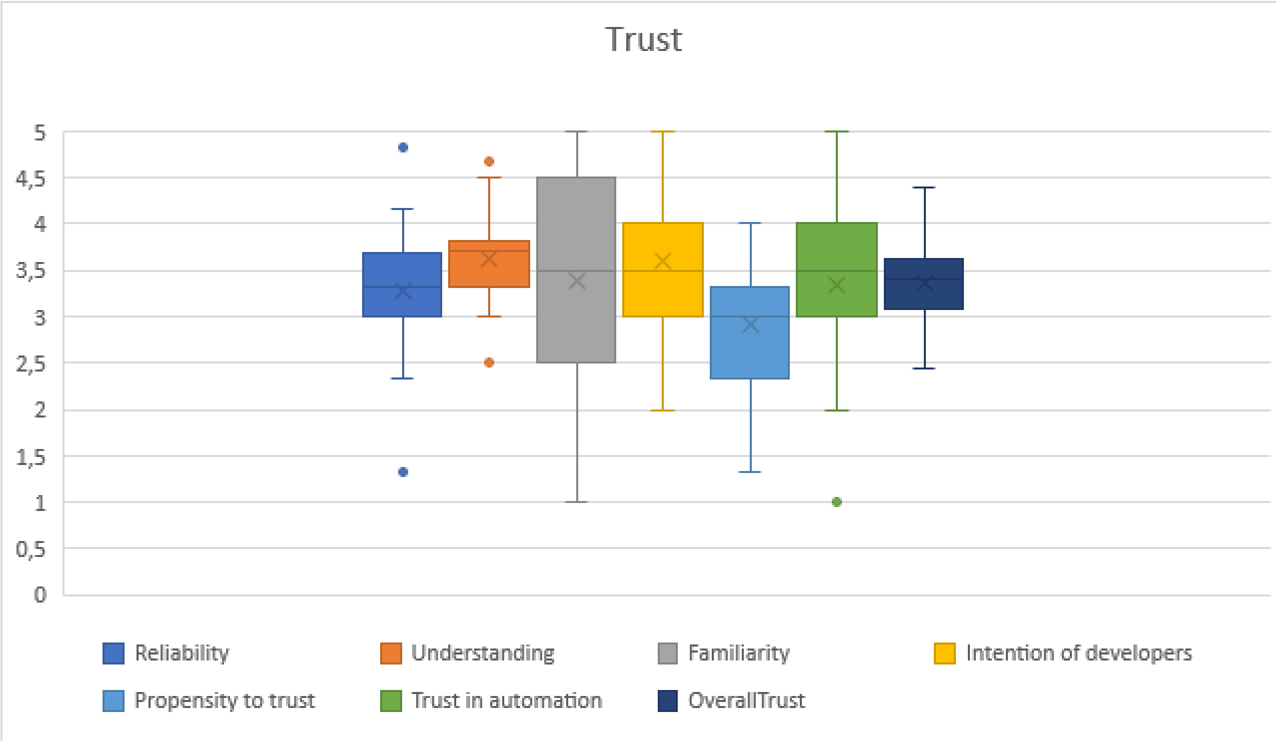


Figure 79: Boxplots for each subscale of the trust questionnaire. Overall trust stands for the mean of the others subscale

4.4.2.4 Enablers evaluation

Mean satisfaction and standard deviations for each enabler evaluated are presented in Table 8. A graphical representation of those scores are presented in Figure 80.

TeamMate Enabler	Mean	Standard Deviation
DMS- Driver Monitoring System	3.97	0.64
Visual HMI- Human Machine Interface	3.66	0.75
Audio HMI- Human Machine Interface	3.71	0.66
DID & RAE – Driver intention detection & Risk Assessment Evaluation	3.58	0.88
V2I- Connectivity platform- vehicle to infrastructure communication	4.02	0.83
ARG- Augmented Reality Glasses	2.81	1.23

Table 8: Mean satisfaction score and standard deviation for each enabler.

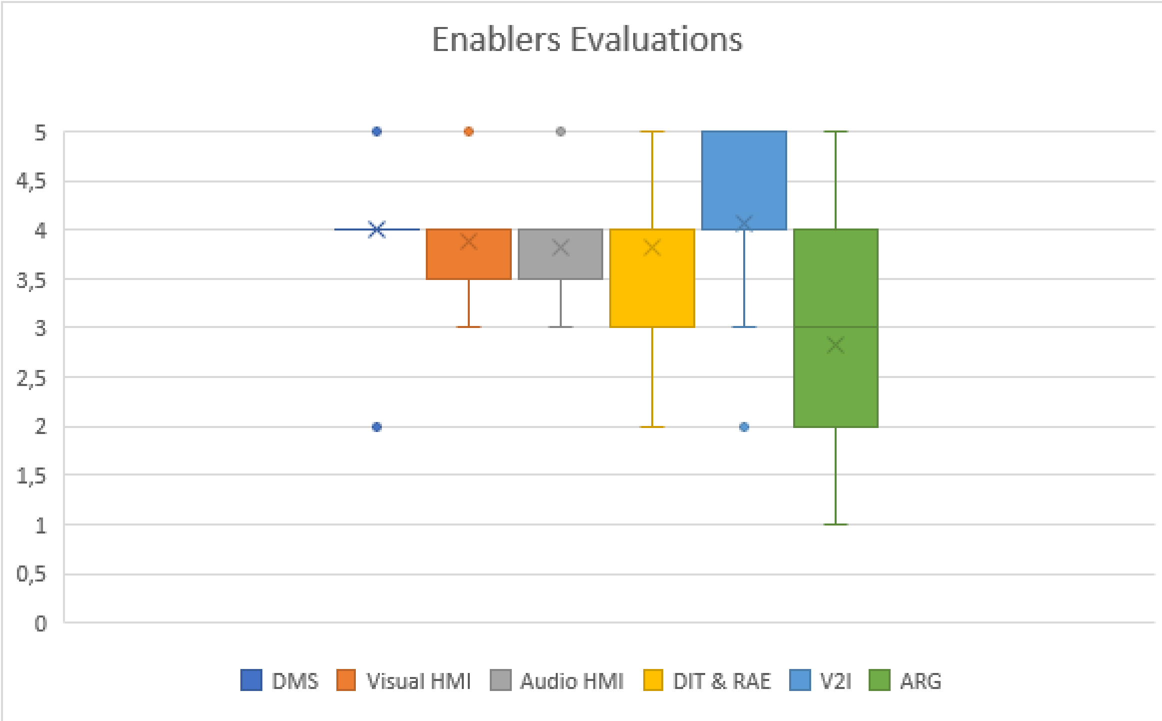


Figure 80: Boxplots of satisfaction score for each enabler.

4.4.3 Discussion

Acceptance, trust and usability analysis revealed results of the subjective evaluation above the average. TeamMate Enablers analysis are also all above the average except for the augmented reality glasses where the result is above the average (but very close).

The Driver Monitoring System is satisfying for participants (**3.97 ± 0.09**). This enabler is easily visible to participants, and the feedback was enhanced with an audio message, which marked even more participants. Visual Human Machine Interface was displayed on the dashboard just behind the steering wheel. It was visible during all the driving session and subjects rated satisfaction (**3.66± 0.12**).

Audio messages were displayed to inform the driver that the autonomous mode was available, when autonomous mode is activated, when autonomous mode is disactivated, inform the driver if he is distracted during manual mode and suggest to activate autonomous mode if available, to warn the driver about the presence of roadworks ahead with an estimation of the distance. The overall audio human machine interface is rated (**3.71 ± 0.10**).

Driver intention detection & Risk Assessment Evaluation are explained to participants when the expert driver from VEDECOM is driving manually the TeamMate Vehicle. They were informed that the goal of these enablers is to assess safety of intended maneuvers of the driver and calculation of safe and feasible trajectories which are shown on a computer screen in real time (for drivers seated in the back). Figures with safety corridors within the boundaries of the road and taking in consideration information from sensors are also shown. These two enablers combined are rated (**3.58 ± 0.14**).

The connectivity platform allowed to receive information from the infrastructure, in the tested use case of the scenario it sends the exact position of the roadwork, and in the visual HMI the left distance to roadwork is updated in Realtime. The road side unit used for this purpose was visible on the speed test tracks. The audio message enhanced the presence of these roadworks and distance to roadworks is also announced when asking the driver to take overcontrol. The V2I is rated (**4.02± 0.13**).

Augmented Reality glasses were worn only by front seat passengers. A total of 17 passengers rated this technology (associated with the driver intention detection and risk assessment evaluation). They were rated (**2.81 ± 0.29**). information displayed on ARG are optimized when using lidar information, although for technical difficulties, radar information was used for the final evaluation to show this proof of concept. The reliability of the system was lower with the radar than lidar, which might explain the notation below the average for this system.

4.5 Final evaluation of the PETER scenario (demonstrator vehicle)

This chapter describes the final evaluation of the "Peter scenario" in the demonstrator vehicle. The goal of this evaluation was to show the added value of the integrated enablers (see Table 9) in an overtaking manoeuvre. The main feature in Peter scenario was the overtaking manoeuvre. Therefore, an adequate scenario (see Scenario) was chosen to show the improvements of the integrated enablers.

The following table summarizes the enablers integrated in the ULM vehicle:

ID	Enabler
E4.1	Planning and execution of safe maneuver
E6.1	Interaction modality
E6.2	TeamMate HMI (Cluster + audio)

Table 9: Enablers integrated in ULM vehicle demonstrator.

These enablers have been implemented to exploit and show the "TeamMate" (TM) car concept: the aim is to prove the benefits of TM use with reference to the baseline car (described in D5.3).

4.5.1 Scenario

The most relevant part of the Peter scenario is the moment the TeamMate car takes over and the baseline car is not able to take over on its own. Therefore, a straight test-track has been chosen where the car could overtake a slowly driving vehicle without any safety issues (see Figure 81).

Figure 81: Evaluation test-track for the Peter scenario with two around 500m long straights.

The test-track has a total length of around 950m in total. For the evaluation a track, consisting of two 500m long straight sections and turning points. A second vehicle was driving in front of the baseline and TeamMate car.

4.5.2 Method

4.5.2.1 Participants

In total 9 participants with a valid German driving license for at least one year have been recruited for the experiment. For participating in the experiment, participants were not compensated monetarily because the experiment took place during their working hours. All participants were research assistance of the University of Ulm.

4.5.2.2 Material

4.5.2.3 Questionnaires

4.5.2.3.1 Trust Questionnaire

Trust was measured after each experimental condition (Baseline car/TeamMate car) using a custom scale, where participants had to indicate how much they trust the system on a scale ranging from 0 to 100, and Körber's "Trust in Automation" questionnaire in German Version, which consists of six scales (Reliability/Competence, Understandability/Predictability, Propensity to Trust, Intention of Developers, Familiarity, and Trust in Automation) containing a total of 19 items (Körber,2018). Participants were asked to rate each item on a 5-point Likert scale from 1 "strongly disagree" to 5 "strongly agree".

4.5.2.3.2 Acceptance Questionnaire

To measure the driver's acceptance of the new technology, a German translation of the acceptance questionnaire from Van der Laan et al (1996) was used. This questionnaire consisted of 9 items. Participants were asked to rate each item consisting of a pair of opposed adjectives (e.g. "*useful*" versus "*useless*", or "*assisting*" versus "*worthless*") from -2 to +2.

4.5.2.3.3 Questions regarding safety [custom]

Participants feeling of safety was measured using a custom questionnaire consisting of 5 items. One item was for example: "I felt safe during the drive". Participants were asked to rate each item on a 5-point Likert scale ranging from 1 "strongly disagree" to 5 "strongly agree".

After the last drive and following the questions above, participants were given the option to give feedback by answering the following three open questions: 1) What increased your feeling of safety?, 2) What would increase your feeling of safety that is currently missing? and 3) space for additional comments.

4.5.2.3.4 Questions regarding the driver's workload [adapted from NASA-TLX]

Since participants were only passengers who observed a driver either conducting the overtaking maneuver manually or using the TeamMate system, 5 items of the German translation of the NASA-TLX (Hart & Staveland, 1988) were selected and adjusted to this different perspective. One item was for example: "The driving task was mentally demanding for the driver".

Participants were asked to rate each item on a 5-point Likert scale ranging from 1 “strongly disagree” to 5 “strongly agree”.

4.5.2.3.5 Questions regarding the interaction between driver and HMI [custom]

A custom-built questionnaire was used to measure the interaction between the driver and the human machine interface evaluated by the participant who only observed the interaction as a passenger. Among five selected items was for example: “the interaction was pleasant”. Participants were asked to rate each item on a 5-point Likert scale ranging from 1 “strongly disagree” to 5 “strongly agree”.

4.5.2.3.6 Questions regarding the HMI [adapted from VisAWI]

To evaluate the human machine interface, 5 items from the VisAWI [88] were selected and adjusted to the context. Participants were asked to rate each item on a 5-point Likert scale ranging from 1 “strongly disagree” to 5 “strongly agree”.

4.5.2.3.7 Questions regarding the vehicle motion behavior [custom]

A custom-built questionnaire was used to evaluate the vehicle motion behavior during the overtaking maneuver. Participants had to indicate how pleasant, unusual and predictable they perceived the longitudinal and lateral motion behavior of the vehicle by rating each item on a 5-point Likert scale ranging from 1 “strongly disagree” to 5 “strongly agree”. The questionnaire consisted of five items.

4.5.2.3.8 Willingness to Pay [custom]

Participants' willingness to pay for a vehicle equipped with the TeamMate system was assessed via three questions. Participants were asked 1) how much money they would be willing to pay more for a car with the TeamMate system in relation to a car without the system, 2) how much money they would be willing to pay less for it (in case they did not like the system), and 3) how much money they would be willing to pay for a vehicle with such a system.

4.5.2.3.9 System Usability Scale

System usability was measured using a German translation of Brooke's (1996) System Usability Scale (SUS) consisting of 10 items, which provides a usability score ranging from 1 to 100. Participants were asked to rate each item on a 5-point-Likert scale ranging from 1 “strongly disagree” to 5 “strongly agree”.

4.5.2.3.10 Questions regarding Enablers [custom]

In addition to the abovementioned questionnaires, participants were asked whether the 10 enablers should be implemented in the car or not. More specifically, participants could choose for every enabler one of the following answers: “no”: the system is not necessary, “possible”: the system could be implemented or “necessary”: the system should be implemented, creating a

3-point Likert scale ranging from 1 “not necessary at all” to 3 “extremely necessary”.

4.5.2.4 Experiment Design

A within subject design was used for this experiment. The manipulated factor was the function of the highly automated vehicle: Baseline car and TeamMate car. The dependent variables were measured objectively and subjectively. For the objective part, electrodermal activity was recorded and the X, Y and Z-coordinated were tracked during all drives.

For the subjective part, trust in automation, acceptance, feeling of safety, workload of driver, interaction, HMI, and willingness to pay, usability, as well as the necessity of enablers was measured through questionnaires (see questionnaires).

The following Table 10 summarizes the KPIs considered in the demonstrator vehicle:

KPI ID	KPI	KPI type	Recording Tool
	Electrodermal activity (skin conductance)	Objective	Brain Vision Recorder
	Coordinates of the vehicle	Objective	GPS
	Trust	Subjective	Körber’s questionnaire
	Acceptance	Subjective	Van der Laan’s questionnaire
	Safety	Subjective	Custom questionnaire

	Driver workload	Subjective	Adapted from NASA-TLX
	Interaction	Subjective	Custom questionnaire
	HMI	Subjective	Adapted from VisAWI
	Vehicle Motion Behavior	Subjective	Custom questionnaire
	Willingness to Pay	Subjective	Custom questionnaire
	System Usability Scale	Subjective	Brooke’s questionnaire
	Necessity of Enablers	Subjective	Custom questionnaire

Table 10: KPIs considered in the ULM demonstrator vehicle

4.5.2.5 Procedure

First, participants were welcomed and asked to sign an informed consent form and data protection agreement. Afterwards they were seated in the middle of the back seat and the EDA sensors were attached to the left hand. A short introduction was read to them, explaining the experiment and the tasks of the participants.

The experiment consisted of two blocs which are described in detail below. Throughout both blocs, electrodermal activity was recorded using the Brain Vision Recorder, and the vehicle coordinates were tracked using the vehicle´s GPS module. In total, every participant experienced 8 drives: 4 drives in the Baseline condition and 4 drives in the TeamMate condition. During every drive, an overtaking maneuver was performed, either manually by the driver or by the TeamMate system.

The first bloc was composed of two drives. After each one of them, participants filled out questionnaires assessing trust, acceptance and safety.

The second bloc was composed of six drives. One of the following three aspects was assessed for every two drives: 1) driver & interaction, 2) HMI and 3) vehicle motion behavior. After every drive, participants filled out the respective questionnaires. This setup allowed for direct comparison between the Baseline car and the TeamMate car for every assessed aspect. In the end, participants filled out the questionnaires concerning trust, acceptance and safety once again as well as the questionnaires measuring the willingness to pay, SUS and necessity of enablers.

The following table summarizes the course of the experiment:

Bloc	Drive	Condition	Assessed aspects
Bloc 1	1 2	Baseline TeamMate	Trust, Acceptance & Safety
Bloc 2	3 4	Baseline TeamMate	Driver & Interaction
	5 6	Baseline TeamMate	HMI
	7 8	Baseline TeamMate	Vehicle Motion Behavior
			Trust, Acceptance & Safety Willingness to Pay, System Usability Scale, Necessity of Enablers

Table 11: The course of the ULM vehicle experiment

4.5.3 Results

4.5.3.1 Quantitative Analysis of Electrodermal Activity and Vehicle Coordinates

There was no significant difference between Baseline and TeamMate condition regarding the electrodermal activity (EDA) with $t(60.945) = -0.144$, $p = .886$ indicating that the test person's skin conductance level was not influenced by who was conducting the overtaking maneuver, a human or the automation. This result is not in line with subjective measurements that clearly show perceived differences between the conditions in regard to e.g. trust and safety (see results below). A difference in EDA would be expected.

There was a significant difference between the first drive and the last drive regarding EDA with $t(28.054) = -2.530$, $p < .05$, indicating that the test person's skin conductance level increased slowly over the time. In the beginning, test persons seem to be more relaxed than at the end of the test drives.

Preceding the experiment two hypotheses were formulated concerning the impact of the vehicle movement behavior on the skin conductance level. H1: The lateral movement of the vehicle (represented by X coordinates) influences the skin conductance level, and H2: The longitudinal movement (represented by Y coordinates) influences the skin conductance level.

The analysis of EDA revealed that the maximum turning points of EDA correlated significantly with the maximum turning points of the X coordinates with $r(70)=.384$, $p < 0.01$, therefore confirming the first hypothesis (see Table 12).

Korrelationen

					Y-Wert Hochpunkt EDA	Y-Wert Hochpunkt X_Dir
Y-Wert Hochpunkt EDA	Korrelation nach Pearson				1	,384**
	Signifikanz (2-seitig)					,001
	N				72	72
	Bootstrap[c]	Verzerrung			0	-,004
		Standard Fehler			0	,103
		95% Konfidenzintervall	Unterer Wert		1	,156
			Oberer Wert		1	,565
Y-Wert Hochpunkt X_Dir	Korrelation nach Pearson				,384**	1
	Signifikanz (2-seitig)				,001	
	N				72	72
	Bootstrap[c]	Verzerrung			-,004	0
		Standard Fehler			,103	0
		95% Konfidenzintervall	Unterer Wert		,156	1
			Oberer Wert		,565	1

**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

c. Sofern nicht anders angegeben, beruhen die Bootstrap-Ergebnisse auf 1000 Bootstrap-Stichproben

Table 12: Pearson correlation of maximum turning points of EDA with maximum turning points of X-coordinates.

However, the second hypothesis was not confirmed since the maximum turning points of EDA did not correlate significantly with the maximum turning points of the Y coordinates with $r(70)=.074$, $p = 0.535$ (see Table 13).

Korrelationen

				Y-Wert Hochpunkt EDA	Y-Wert Hochpunkt Y_Dir
Y-Wert Hochpunkt EDA	Korrelation nach Pearson			1	,074
	Signifikanz (2-seitig)				,535
	N			72	72
	Bootstrap[c]	Verzerrung		0	,039
		Standard Fehler		0	,168
		95% Konfidenzintervall	Unterer Wert	1	-,182
			Oberer Wert	1	,436
Y-Wert Hochpunkt Y_Dir	Korrelation nach Pearson			,074	1
	Signifikanz (2-seitig)			,535	
	N			72	72
	Bootstrap[c]	Verzerrung		,039	0
		Standard Fehler		,168	0
		95% Konfidenzintervall	Unterer Wert	-,182	1
			Oberer Wert	,436	1

c. Sofern nicht anders angegeben, beruhen die Bootstrap-Ergebnisse auf 1000 Bootstrap-Stichproben

Table 13: Pearson correlation of maximum turning points of EDA data with maximum turning points of Y-coordinates.

This indicates that the lateral movement, i.e. the left-right movement of the car had an impact on the level of emotional arousal of participants, while the longitudinal movement, i.e. acceleration and deceleration of the car did not have any impact on participants. This result is in line with the subjective measurements revealing that participants disliked the lateral vehicle motion behavior during overtaking (see results below).

However, some limitations of this analysis have to be taken into consideration. Only maximum turning points of EDA were considered. When correlating all data points of EDA with vehicle coordinates (instead of focusing on minimal and maximal turning points), no consistent pattern could be found: EDA data does not correlate with X coordinates with an average correlation of r(X)=.069, a minimal correlation of r(X)=-.429 and a maximal correlation of r(X)=.440.

EDA does not correlate with Y coordinates with an average correlation of *r*(X)=-.141, a minimal correlation of *r*(X)=-.646 and a maximal correlation of *r*(X)=.414 (since the time it took a participant to experience the overtaking maneuver was different for every drive, the degree of freedom is different for every single correlation). Moreover, every correlation reached the level of significance as there were a multitude of data points for every drive (~20.000). Therefore, only effect sizes should be taken into account. Effect sizes were small and showed a lot of diversity with ranges from -.4 to +.4, hence, we cannot draw reliable conclusions from this analysis.

Another interpretation idea led to the consideration of latency (reaction) and anticipation times related to the overtaking maneuver (see exemplary Table 14 and Table 15). Nevertheless, when correlating the data considering these two aspects, no conclusive pattern could be found. Correlations of all data points between EDA and X as well as EDA and Y coordinates for varying time periods between -+0.5 to + 4 seconds (latency) and -0.5 to -4 (anticipation) were calculated (to do so, data points were cut in the beginning or end).

Latency times		**Reaction**		**Anticipation**	
|in s|	in t	$r_{EDA,X}$	$r_{EDA,Y}$	$r_{EDA,X}$	$r_{EDA,Y}$
0	1	.295	.264	.295	.264
0,1	2	.290	.262	.303	.265

0,2	3	.285	.261	.311	.266
0,3	4	.280	.260	.320	.267
0,4	5	.275	.258	.327	.269
0,5	6	.271	.257	.336	.272
1,5	8	.235	.240	.412	.306
2,5	10	.203	.217	.460	.302
3,5	12	.181	.196	.510	.215
4	13	.190	.174	.533	.165

Table 14: Exemplary calculation of correlations considering latency (reaction) and anticipation times of 0 -4 seconds for the first drive (Baseline) of test person 1

Latency times		**Reaction**		**Anticipation**	
|in s|	in t	$r_{EDA,X}$	$r_{EDA,Y}$	$r_{EDA,X}$	$r_{EDA,Y}$
0	1	.050	- .295	.050	- .295
0,1	2	.047	- .301	.053	- .294
0,2	3	.045	- .308	.057	- .293

0,3	4	.042	- .313	.061	- .292
0,4	5	.041	- .319	.065	- .291
0,5	6	.040	- .324	.070	- .291
1,5	8	.018	- .346	.117	- .300
2,5	10	.010	- .335	.135	- .332
3,5	12	.072	- .324	.122	- .397
4	13	.121	- .319	.106	- .427

Table 15: Exemplary calculation of correlations considering latency (reaction) and anticipation times of 0 -4 seconds for the first drive (TeamMate) of test person 2

Summarizing the objective results, it can be said that no reliable conclusions can be drawn from quantitative data analysis with a sample as small as 9 participants. Therefore, the abovementioned results should be considered with caution. It is suggested to repeat the experiment with a bigger sample size.

4.5.3.2 Questionnaires

Considering the small number of participants (N=9), all questionnaire analyses were done descriptively & qualitatively.

4.5.3.3 Trust

Participants trusted the system more when a human driver carried out the overtaking maneuver (79%) compared to the TeamMate driving concept (69%) (see Figure 82). Trust didn't change after participants got more familiar with the system: even after the last drive they trusted the system more when a human driver was involved (79%) vs the TeamMate driving concept (64%). For the latter, it descriptively looks like people trust the system even less than before.

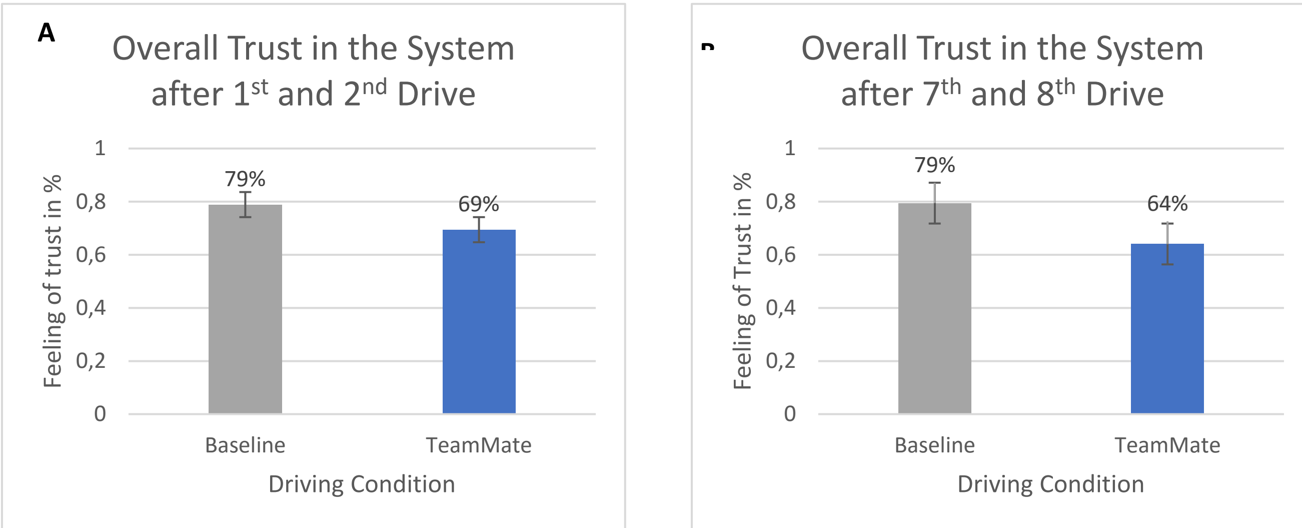


Figure 82: Trust score of the baseline and the TeamMate car (the error bars depict the standard deviation) after the first drives (A) and after the last drives (B).

The item "I am confident about the system's capabilities" received the lowest score (m=2.77) on the 5-point Likert scale which could indicate that after experiencing the system for the first time, participants were not yet convinced about its capabilities.

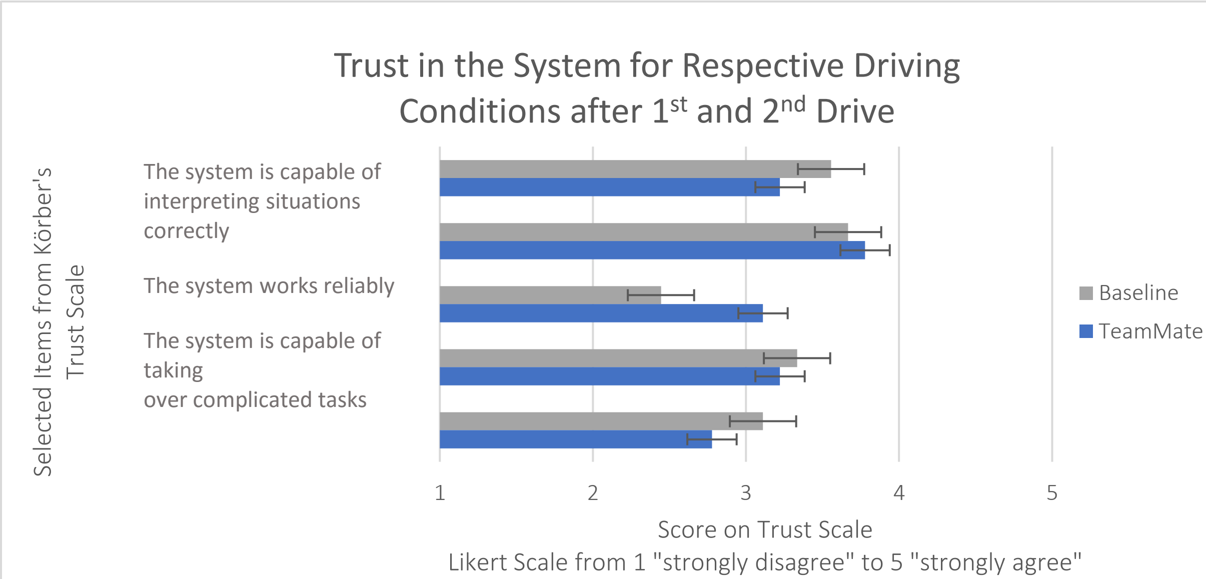


Figure 83: Trust score on a 5-point Likert scale for the Baseline and TeamMate car after the first drives (the error bars depict the standard deviation).

After participants got familiar with the system, most items were evaluated in a similar fashion. However, the item "the system is capable of taking over complicated tasks" was then better evaluated for the Baseline condition.

Participants trust the manual driver more to take over complicated tasks, while trust stays the same (or descriptively even decreases) in the TeamMate condition.

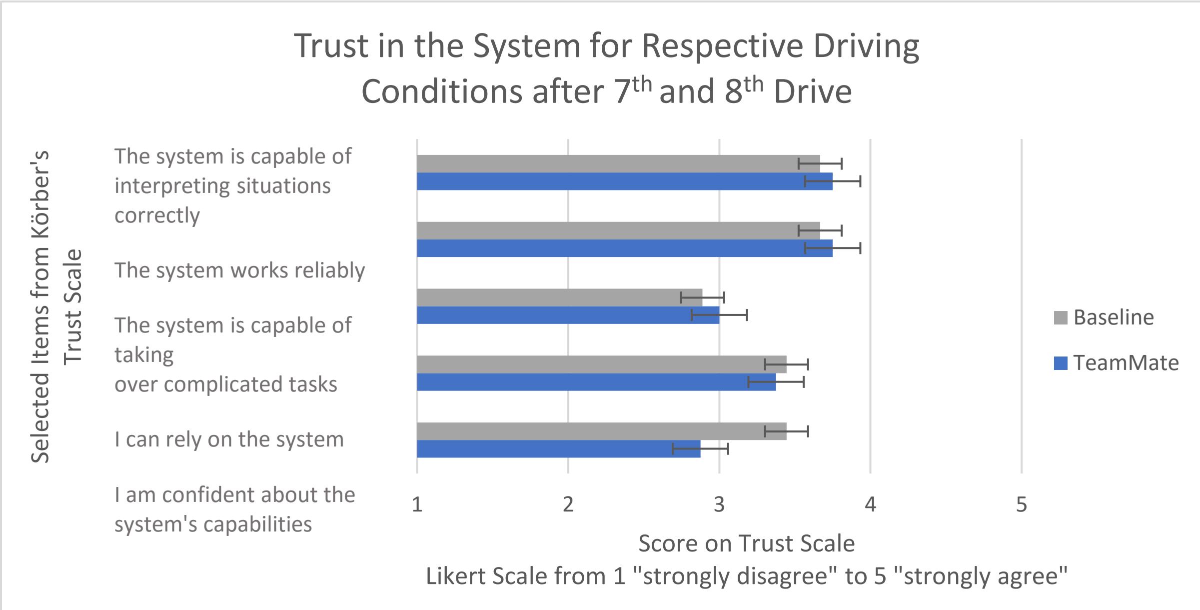


Figure 84: Trust score on a 5-point Likert scale for the Baseline and TeamMate car after the last drives (the error bars depict the standard deviation).

4.5.3.4 Acceptance

The results for the first contact clearly show that usefulness is considered high for the TeamMate car, however, satisfaction is very low. For the Baseline car satisfaction is a lot higher with a lower usefulness.

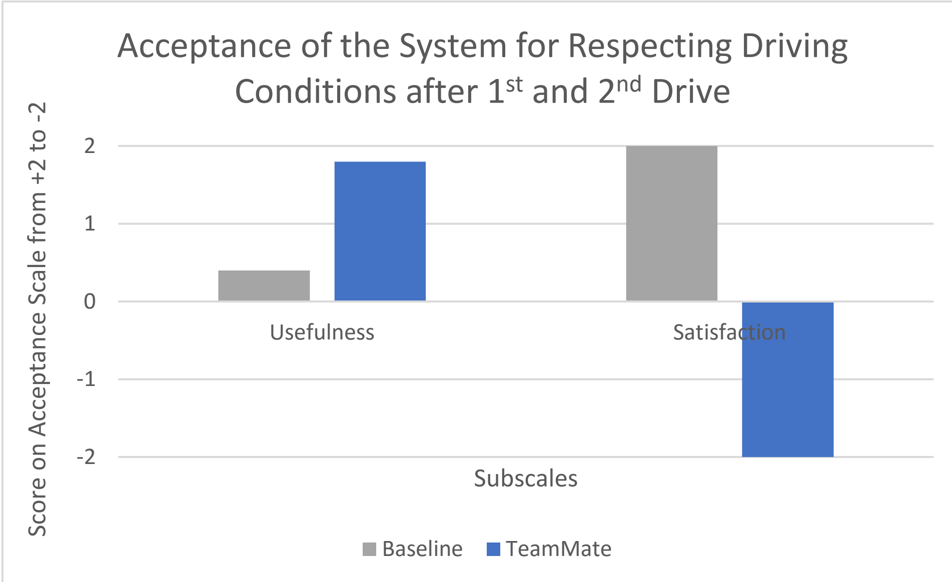


Figure 85: Acceptance score for Baseline and TeamMate car for the subscales usefulness and satisfaction after the first drives.

After participants got more familiar with the system, it was not considered useful anymore, and satisfaction was even lower than for the first drives. For the Baseline condition, the usefulness increased while satisfaction with decreased by a lot.

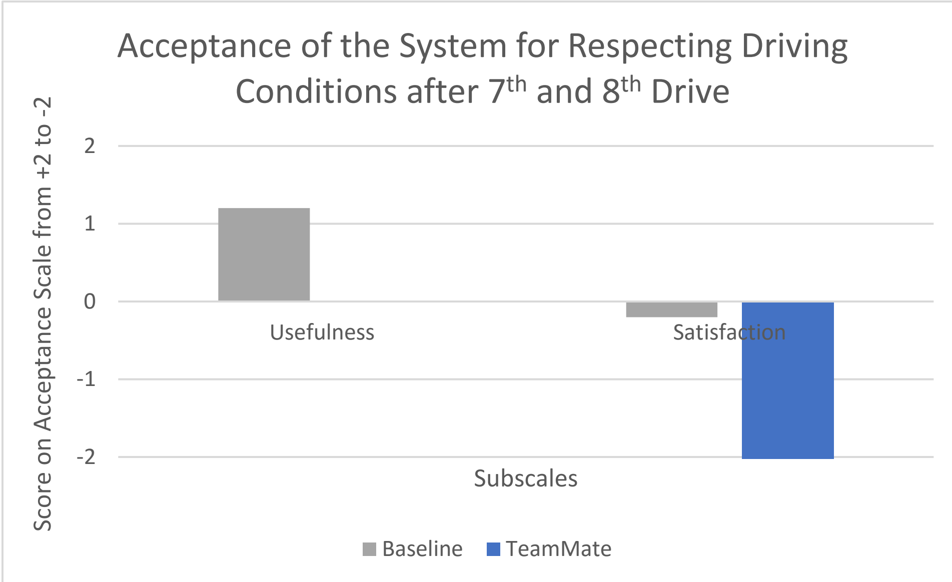


Figure 86: Acceptance score for Baseline and TeamMate car for the subscales usefulness and satisfaction after the last drives.

The teammate driving concept is considered unpleasant (m=3.78). This unpleasant feeling did not go away when participants got more familiar with the system (m = 3.75).

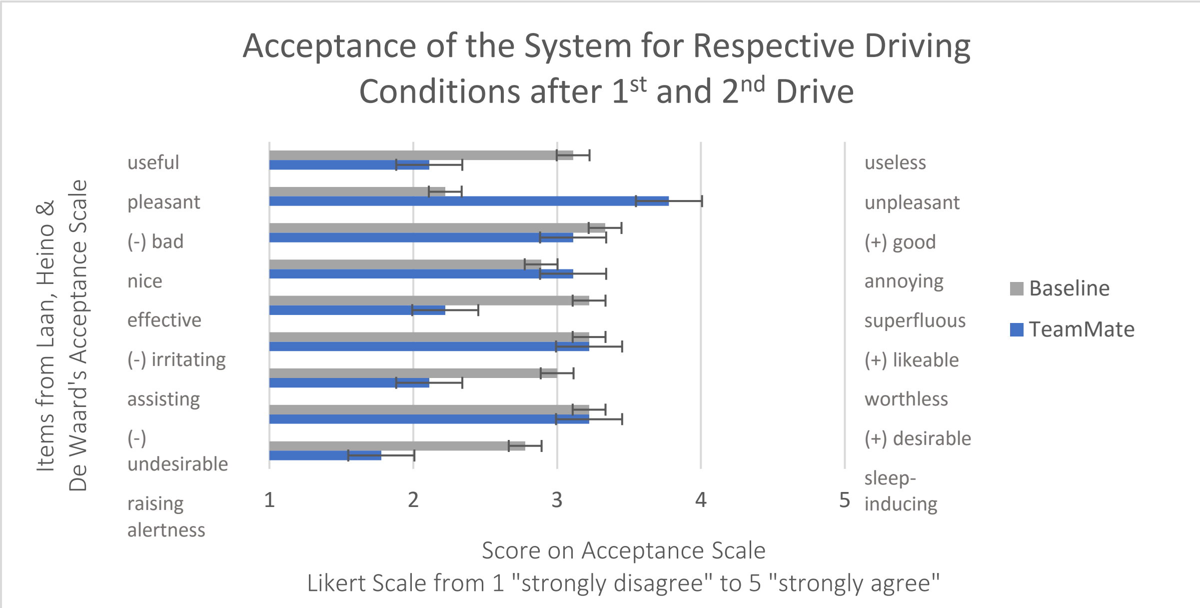


Figure 87: Acceptance score on a 5-point Likert scale for the Baseline and TeamMate car after the first drives (the error bars depict the standard deviation).

However, it was also considered useful, effective, assisting, and raising the alertness for the first contact, as well as after getting to know it more in detail.

The acceptance of the Baseline condition improved with familiarity: it was considered less worthless and less sleep-inducing for the last drives.

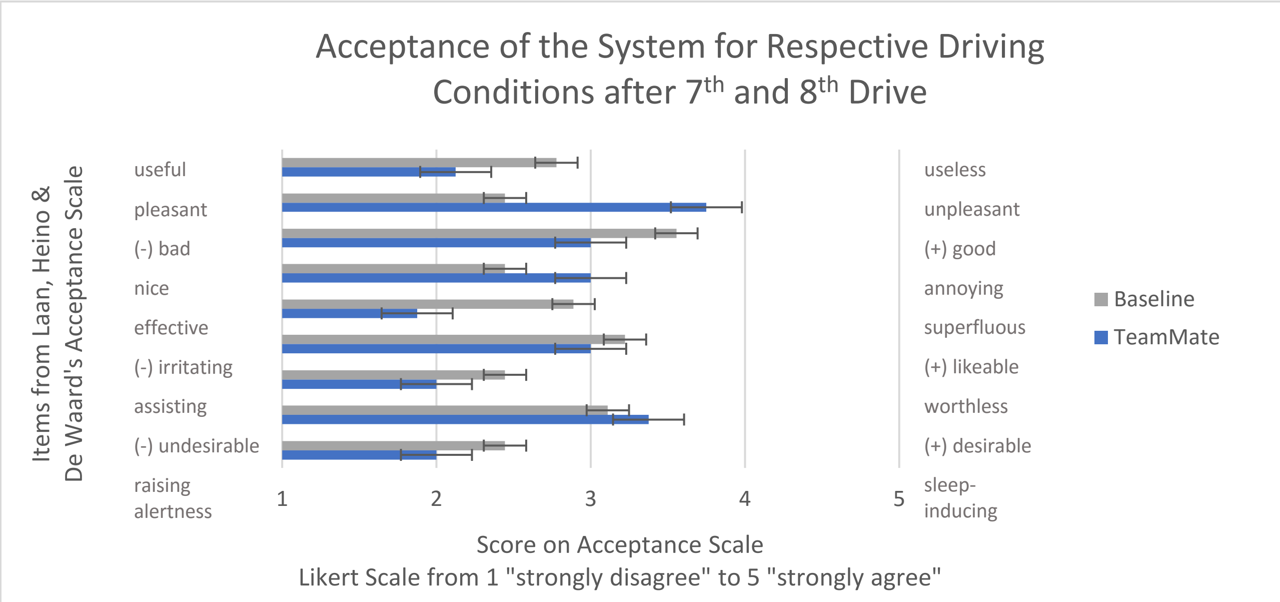


Figure 88: Acceptance score on a 5-point Likert scale for the Baseline and TeamMate car after the first drives (the error bars depict the standard deviation).

4.5.3.5 Safety

Participants felt safer and more relaxed in the Baseline condition while they were somewhat worried and concerned during the TeamMate condition (see Figure 89).

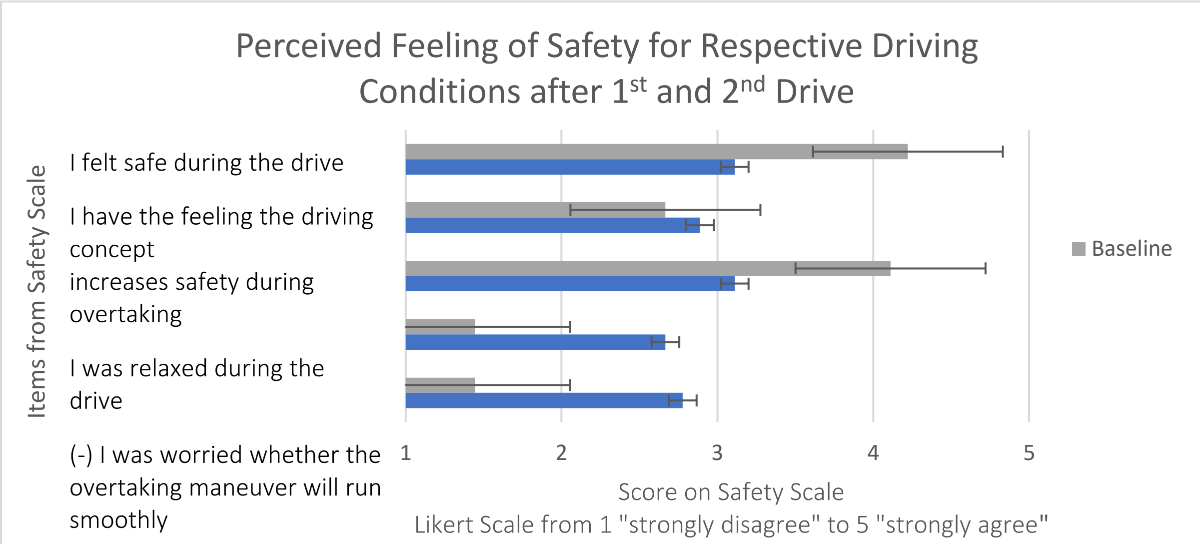


Figure 89: Safety score on a 5-point Likert scale for the Baseline and TeamMate car after the first drives (the error bars depict the standard deviation).

After a few drives, participants felt even safer in the Baseline condition and less safe in the TeamMate condition compared to the first contact with the system (see Figure 90).

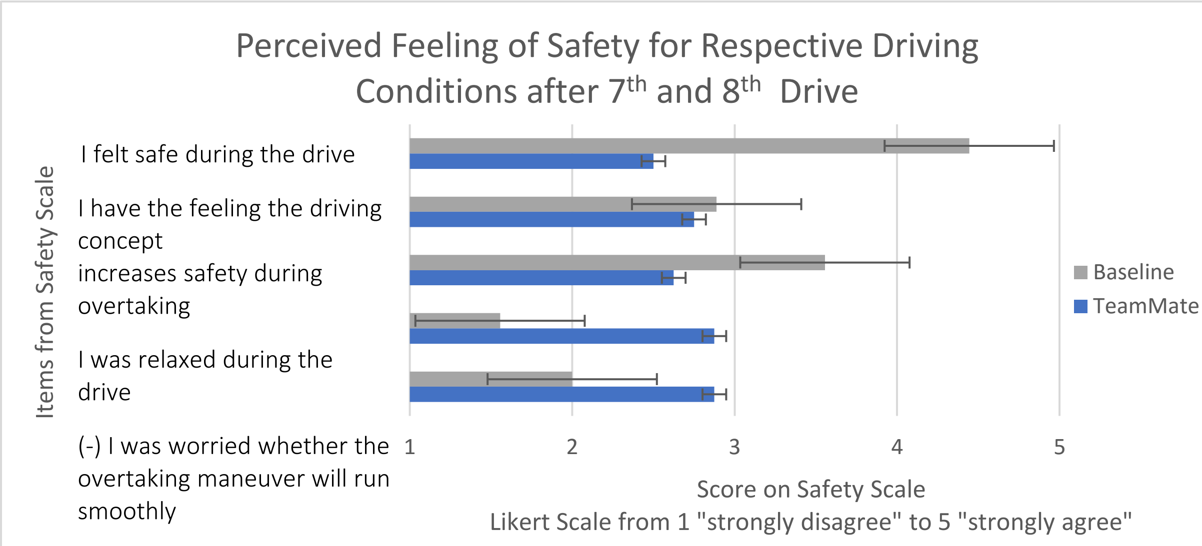


Figure 90: Safety score on a 5-point Likert scale for the Baseline and TeamMate car after the first drives (the error bars depict the standard deviation).

Insights from the open questions concerning safety showed that most people negatively evaluated the lack of communication and information in the Baseline condition in order to feel safer. Moreover, they preferred a human and calm driving behavior over the driving style of the TeamMate driving concept: a major problem was the lateral control of the vehicle during the overtaking process.

4.5.3.6 Driver & Interaction

With equally good performance, the TeamMate condition was rated better in

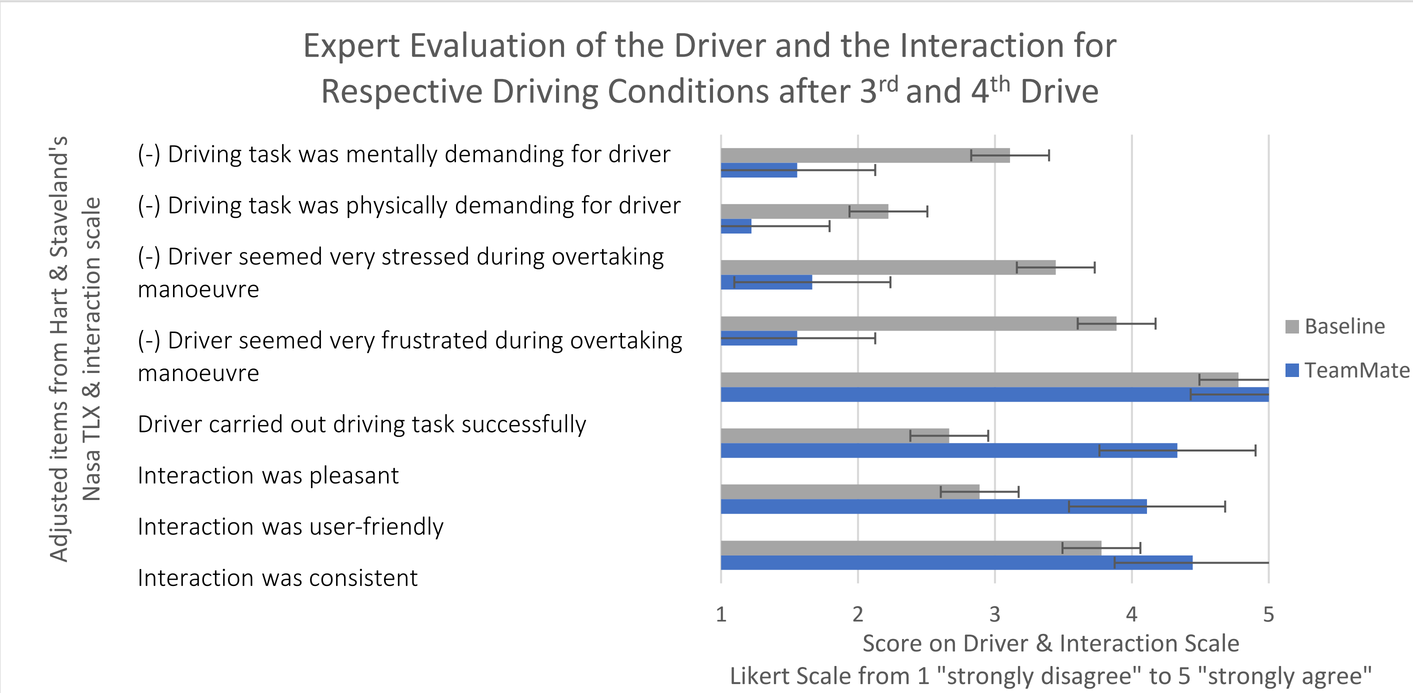


Figure 91: Evaluation of the driver's workload and the interaction between driver and human machine interface by the expert (participant) (the error bars depict the standard deviation).

terms of interaction (+) and driver demands (-): the interaction was perceived as pleasant, consistent and user-friendly while mental and physical demand as well as the level of stress and frustration of the driver was considered low (see Figure 91).

4.5.3.7 HMI

The human machine interface was not considered intuitive in the TeamMate condition and the Baseline condition did not display enough information (see Figure 92). In general, scores were not very good for the HMI (none of the items reached high scores, all ≤ 4).

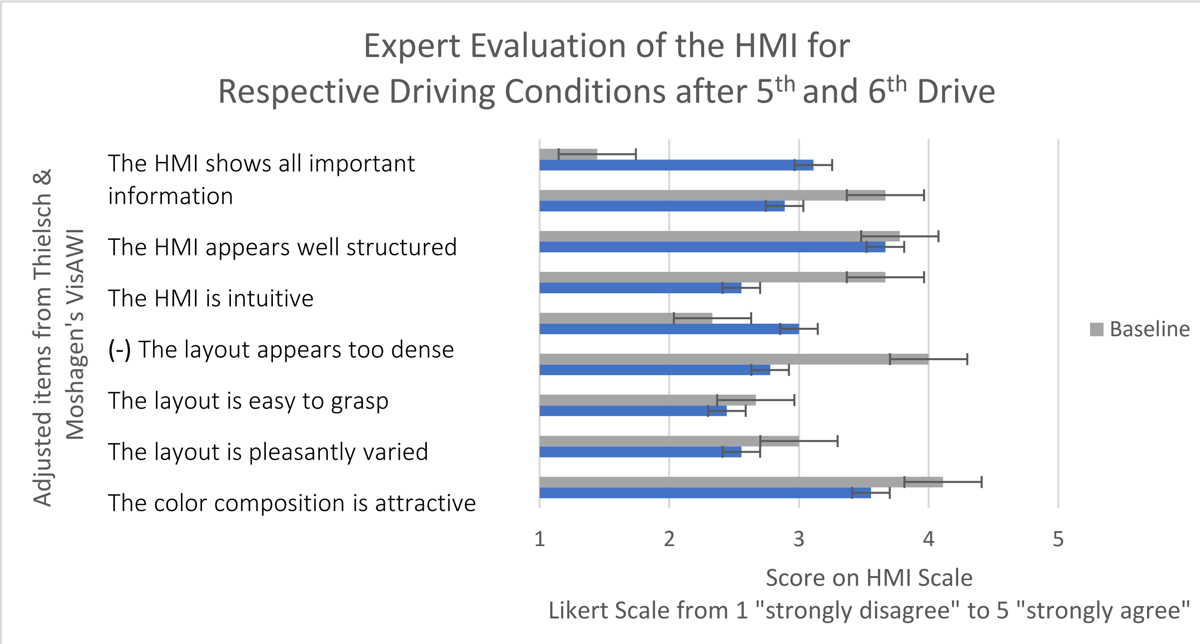


Figure 92: Evaluation of the human machine interface by the expert (participant) (the error bars depict the standard deviation).

4.5.3.8 Vehicle Motion Behavior

The lateral control was neither pleasant nor accustomed nor predictable in the TeamMate condition (see Figure 93).

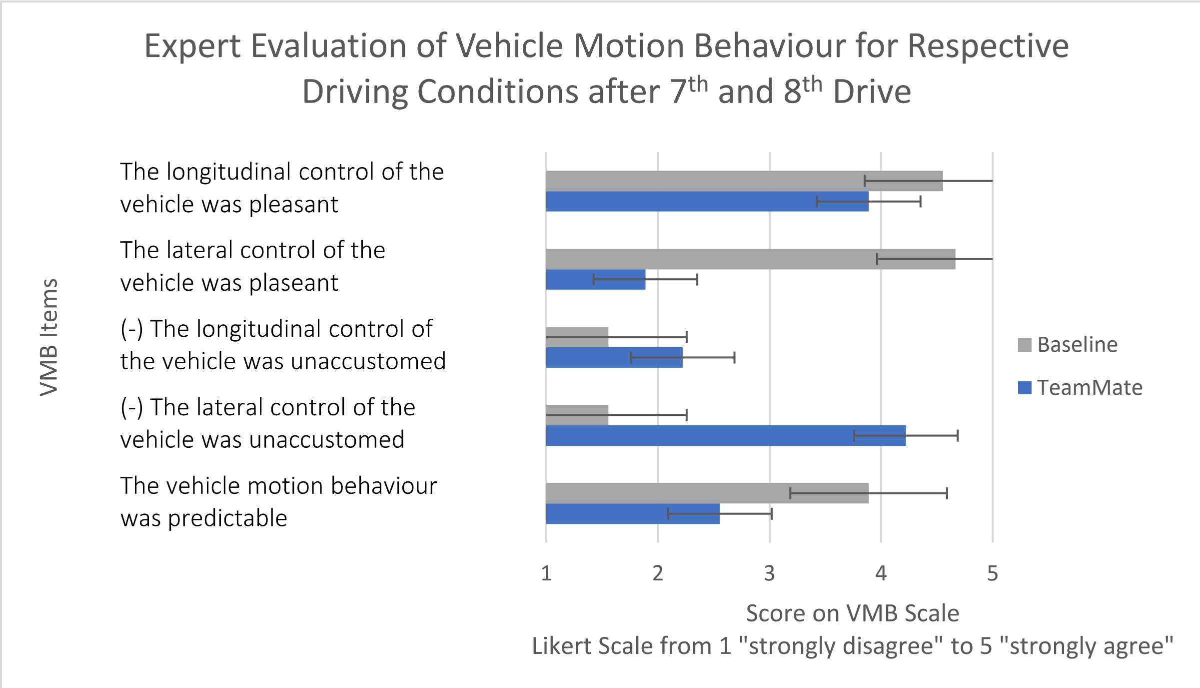


Figure 93: Evaluation of the human machine interface by the expert (participant) (the error bars depict the standard deviation).

4.5.3.9 Willingness to pay

78% of participants would be willing to pay more for a vehicle with TeamMate driving concept. Those 78% would be willing to pay in average 3.785,71€ more (see Figure 94).

The 22% of participants not willing to pay more, would even pay in average 500,5€ less.

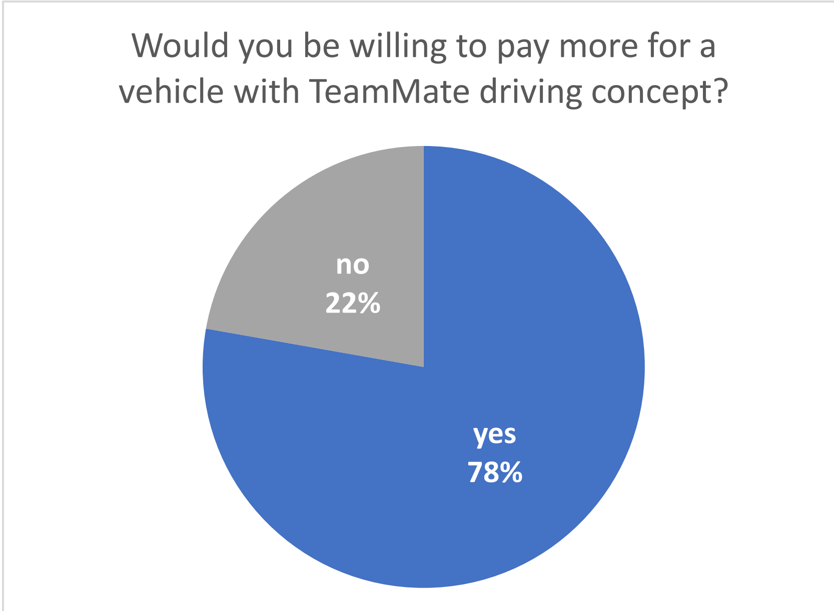


Figure 94: Participan's willingness to pay more money for a vehicle with TeamMate system.

4.5.3.10 System Usability Scale

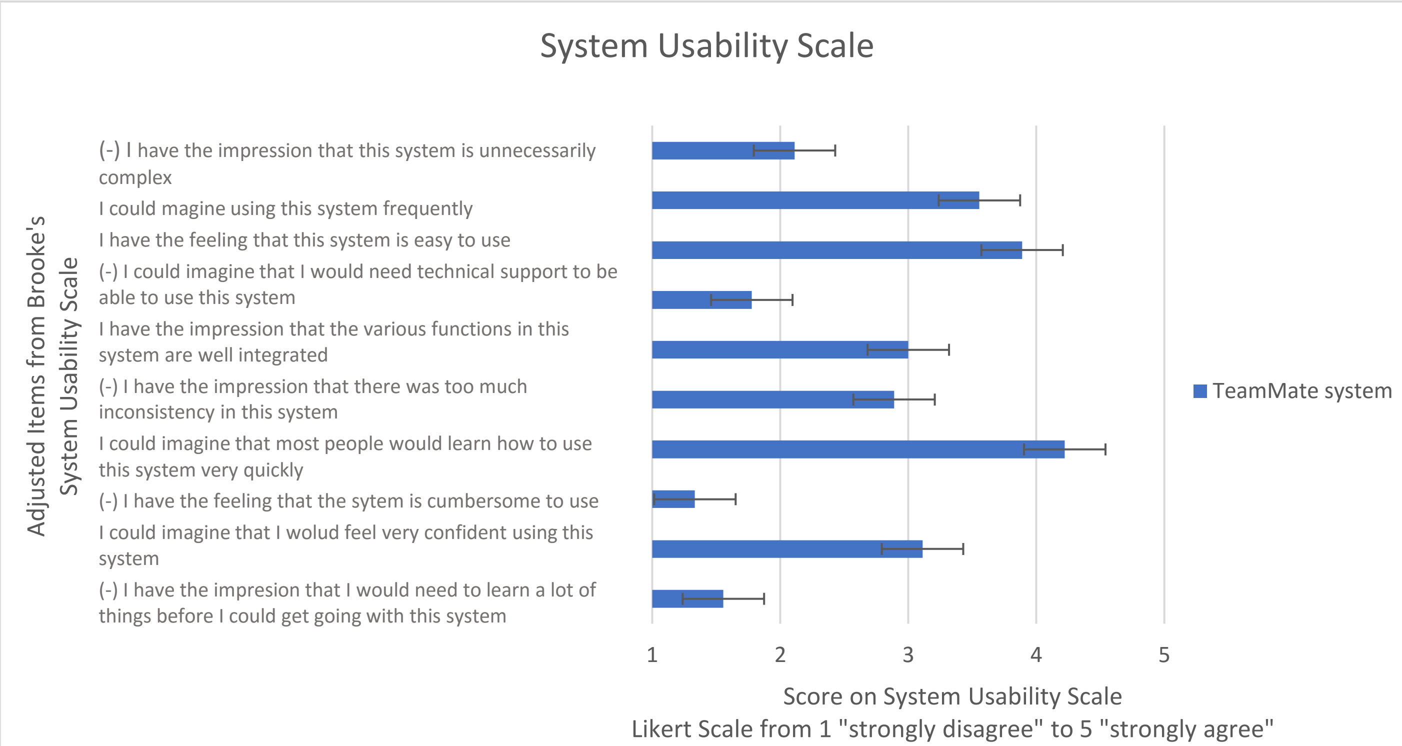


Figure 95: SUS score of the TeamMate system (the error bars depict standard deviation)

With a value of 57.78, system usability is below the lower acceptable limit, improvement is needed indisputably (see Figure 95 und Figure 96).

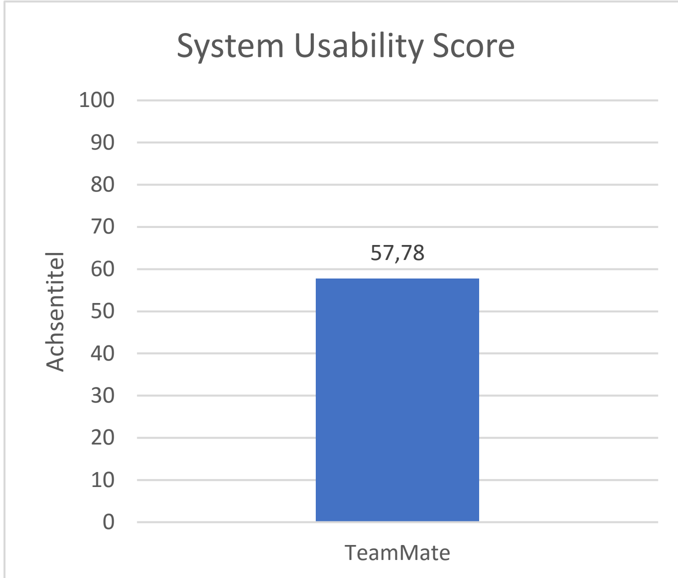


Figure 96: SUS score of the TeamMate system after the final drive (no distinction was made after the last drive, only the TeamMate system was evaluated).

4.5.3.11 Enabler

Finally, participants evaluated the necessity of integrating different enablers into the system. They considered the "planning and execution of safe manoeuvre" as the most necessary enabler and the "driver intention recognition" as the least necessary enabler (see Figure 97).

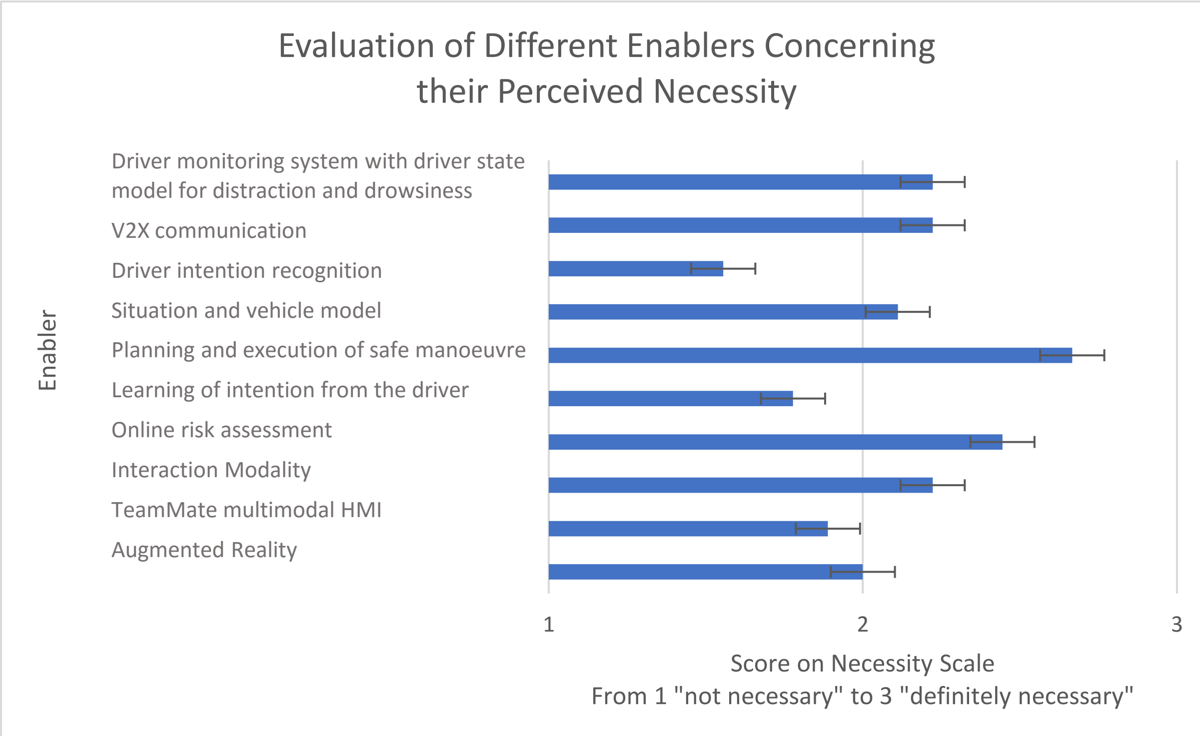


Figure 97: Evaluation of the necessity of enablers (error bars depict the standard deviation)

4.5.4 Discussion

The main motivation of this study was the comparison of the Baseline car with the TeamMate car to assess the impact of implemented enablers on participant's trust, acceptance and feeling of safety. Moreover, the mental workload of the driver, the interaction, the human machine interface as well as the vehicle motion behavior were examined.

4.5.4.1 Objective measurements

The quantitative results of the analysis of EDA data did not allow to draw any reliable conclusions since only 9 participants were assessed.

4.5.4.2 Subjective measurements

The qualitative results of questionnaires can be summarized as follows: Participants trust a human driver more than the automation when it comes to overtaking maneuver. This is true for the first contact and even more pronounced for the following contacts and could be explained with participant's comments about the lateral control of the TeamMate car. It was stated that the lateral control felt uncomfortable and unnatural since the vehicle steered too abruptly and too fast to the left lane while keeping an unexpected big distance to the other car. Even though the vehicle was programmed, due to safety reasons, two cross two lanes during the overtaking maneuver, the trajectory needs to be improved in future. The maneuver could be adjusted to a smoother and more human like overtaking maneuver.

The results for acceptance are in line with those for trust showing that people prefer the Baseline car over the TeamMate car. While people see the benefits of the driving concept of the TeamMate car, they do not like its current implementation.

Moreover, the qualitative questionnaires show that participants do not feel safe driving in the TeamMate car. This is in line with the trust results and could again be explained with the lateral control during overtaking. This finding is especially concerning, since a lack of safety could lead to people misusing or disusing the TeamMate system. A focus for improvement should therefore be on implementing a smoother and more natural lateral movement behavior, to increase people's feeling of safety.

From the open comments it becomes clear that people wish for more information, they prefer a more detailed communication between TeamMate car and passenger/driver. More information for the passenger/driver could also lead to better situation & intention awareness.

The human machine interface is considered as not intuitive by participants. In contrast, the evaluation of the interaction between the driver and the system as well as the perceived mental workload of the driver reached satisfying results.

Even though people do not trust the TeamMate car as much as the Baseline car, do not feel as safe in it and complain about the lack of information, a high number of people would be willing to pay more for the TeamMate car. This shows that people see the potential benefits of the TeamMate car and believe that the technical aspects would improve. This is underlined by the good system usability rating. A final questionnaire about the necessity of enablers revealed that among all options, it is most important to people that the vehicle can plan and execute safe maneuvers. This also shows that people are interested in being supported during driving. Since people clearly like the idea of a TeamMate car, it can be said that with better implementation of the technical aspects, especially related to the lateral control, trust, acceptance and safety would most probably increase.

4.6 Final evaluation of the EVA scenario (demonstrator vehicle)

This chapter describes the final evaluation of the "Eva scenario" in the demonstrator vehicle. As aforementioned in the previous section, for the "Eva scenario" in the REL driving simulator, the goal is to evaluate the added value of the ecosystem of enablers integrated in the last period on the CRF prototype. The EVA use case[7], described in D1.3 and D1.5, has been selected and adapted in order to answer new research questions, as well as to measure the value of the enablers integrated in the last cycle.

The following table summarizes the enablers integrated in CRF vehicle:

ID	Enabler
E1.1	Driver monitoring system with driver state model for distraction and drowsiness
E3.1	Situation and vehicle model
E4.1	Planning and execution of safe maneuver

[7] The Eva scenario is described as follows: "A TeamMate Car is driving through a complex roundabout with different traffic and driving status conditions". In particular, we have considered the type of support "Human To Automation" (H2A), with two different modes: cooperation in perception and in action. On CRF demonstrator, we took into account these two kinds of support: H2A support in perception and H2A support in action.

E6.2 - E6.3 - E6.4	TeamMate HMI (Cluster + audio, Central stack display, HUD)

Table 16: Enablers integrated in CRF demonstrator.

These enablers have been implemented to exploit and show the "TeamMate" (TM) car concept: the aim is to prove the benefits of TM use with reference to the baseline car (described in D5.3).

4.6.1 Scenario

The most relevant part of EVA scenario is the roundabout. Therefore, a dedicated test-site has been selected including real roads in the nearby of Orbassano town (the area where CRF is located). The following figure show this test-site:

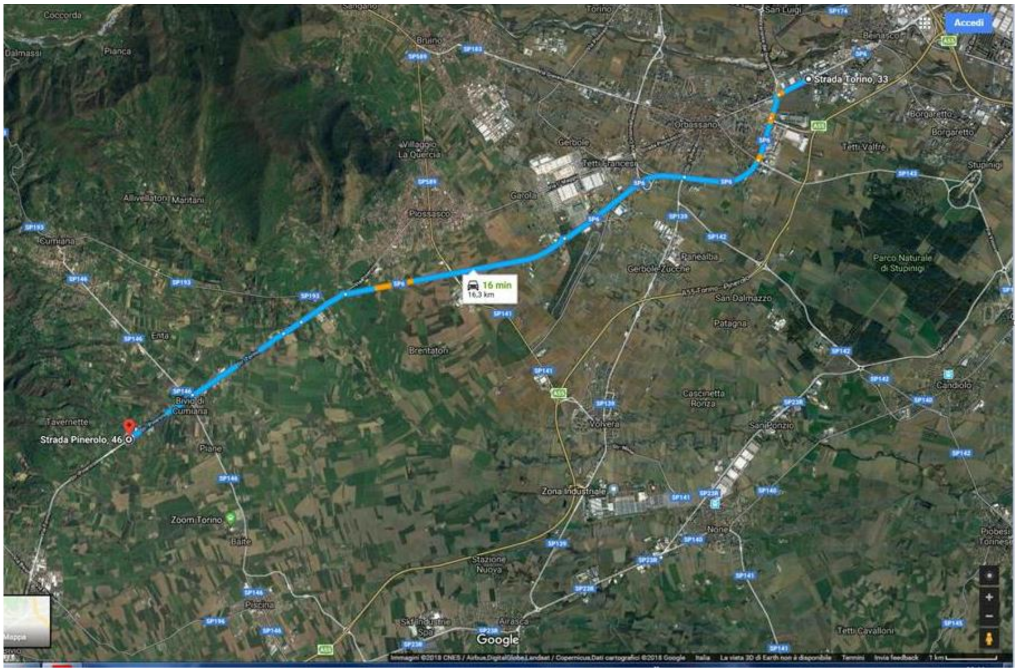


Figure 98: Baseline and TeamMate itinerary (test-site) in EVA scenario for CRF evaluation.

The test-site has is composed by extra-urban roads from "Orbassano" town towards "Pinerolo" village, with the following characteristics:

- Total length = around 40km
- Total number of roundabouts = around 23
- Road structure = Two-lanes or one-lane for each direction
- Speed limits = segments with 70km/h or 90km/h
- Volume of traffic = Medium

In TeamMate (TM) mode, the scenario always starts in Manual Mode (MM). Then, the user selects the Automated Mode (AM) and, if all conditions are met, the TM car starts moving autonomously. When the car approaches a roundabout, since it cannot deal with in AM (no lanes are present in these roundabout and thus the vehicle cannot manage the lateral control) three situations can happen:

1. The driver is attentive, the TM car asks for a sharing control and this is immediately accepted.
2. The driver is distracted by watching a video on a smartphone, thus the system takes back the user into the control-loop before asking for a sharing control; and then, it is accepted.
3. The driver is distracted by talking with a passenger, thus the system takes back the user into the control-loop before asking for a sharing control; and then, it is accepted.

Most of the roundabouts is travelled considering the first point (point 1); two roundabouts in the test-site are used for the second situation (point 2) and other two for the third situation (point 3). This means that for every user, four distraction events are considered.

In situation 1), since the TM car is not able to drive autonomously through the roundabout, the system ask for human intervention, that is for a sharing control: the driver is in charge for the lateral control and the system is in charge for the longitudinal control. If he/she does not react, the car performs a Minimum Risk Maneuver (MRM), by stopping before entering it. After the roundabout, the car, when detecting again the lanes and is able to regain the control, performs an automatic switch to Automated Mode, informing the driver of the transition. The following figure reports the flowchart for this situation in the TeamMate scenario:

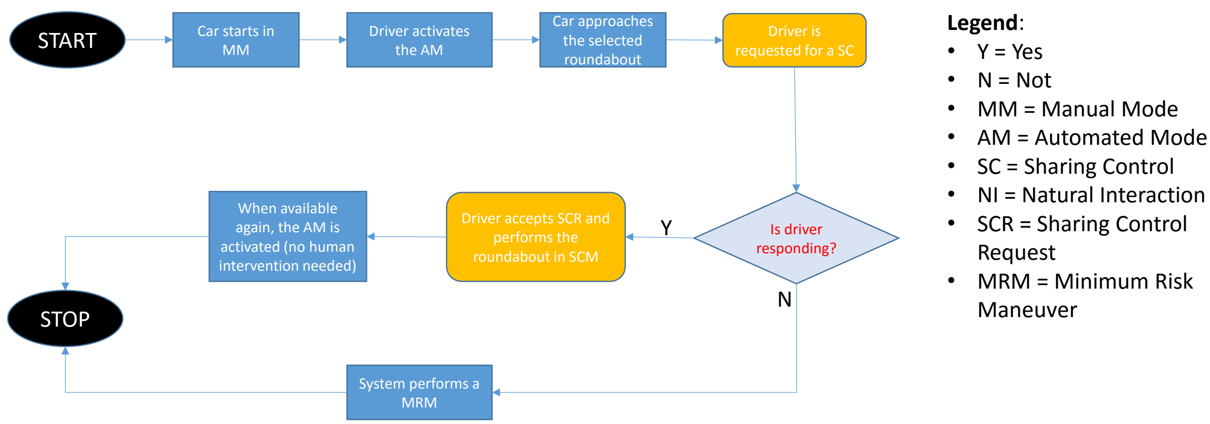
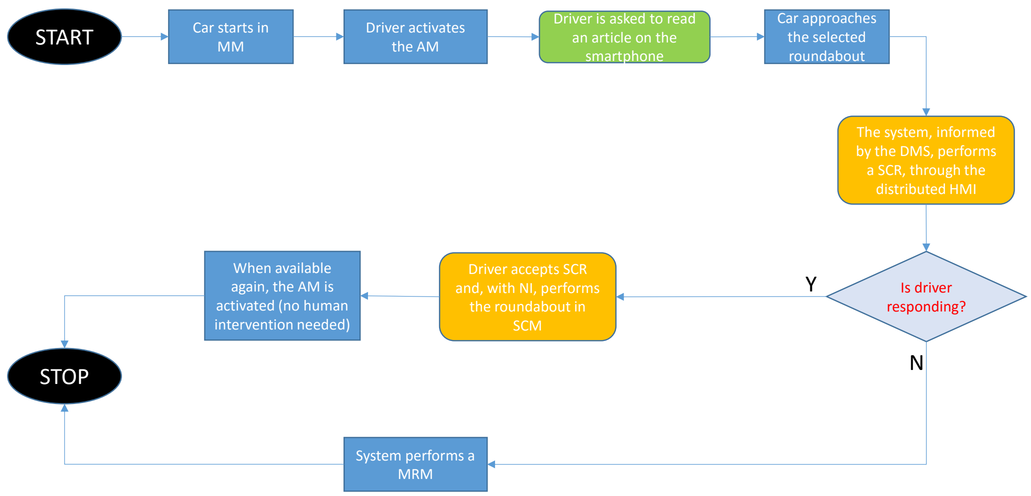


Figure 99: TeamMate scenario for situation 1) in CRF demonstrator car.

In situation 2), the driver is asked by the experimenter to read aloud a text (in Italian language) on a smartphone placed close to the central tunnel. When approaching the selected roundabout, since the car needs the driver's intervention and the DMS detects that the driver is looking in that specific Area of Interest, the sharing control request (SCR) is given directly on the smartphone. If the driver reacts properly (i.e. with natural interaction, by pressing a pedal) he/she takes the control and performs the roundabout in shared modality (longitudinal control to the vehicle, while lateral control to the driver); again, if he/she does not react, the car performs a Minimum Risk Maneuver (MRM), by stopping the vehicle before entering it. After the roundabout, also in this case, the car performs an automatic switch to Automated Mode (if and when possible), informing the driver of the transition.

The next figure reports the flowchart with the second situation for the TeamMate scenario:

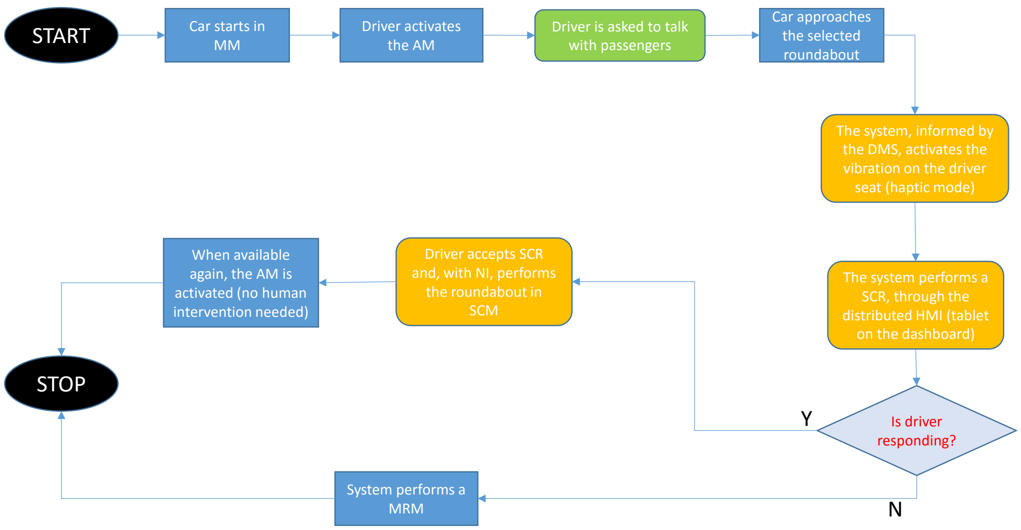


Legend:
- Y = Yes
- N = Not
- MM = Manual Mode
- AM = Automated Mode
- SC = Sharing Control
- NI = Natural Interaction
- DMS = Driver Monitoring System
- SCR = Sharing Control Request
- MRM = Minimum Risk Maneuver

Figure 100: TeamMate scenario for situation 2) in CRF demonstrator car.

Situation 3) is similar to the previous one, but now the driver is asked by the experimenter to talk with the passenger(s), looking at her/him. When approaching the selected roundabout, since the car needs the driver's intervention and the DMS detects that the driver is distracted (not looking ahead the road), first the HMI decides to activate the haptic device (tactile seat) to communicate to the driver that s/he has to come back to the control loop, then the sharing control request (SCR) is given in the main display (the tablet located on the dashboard) as distributed HMI. If the driver reacts properly (i.e. with natural interaction, by pressing a pedal) he/she takes the control and performs the roundabout in shared modality (longitudinal control to the vehicle, while lateral control to the driver); again, if he/she does not react, the car performs a Minimum Risk Maneuver (MRM), by stopping the vehicle before entering it. After the roundabout, also in this third case, the car performs an automatic switch to Automated Mode (if and when possible), informing the driver of the transition.

The next figure reports the flowchart with the second situation for the TeamMate scenario:



Legend:

- Y = Yes
- N = Not
- MM = Manual Mode
- AM = Automated Mode
- SC = Sharing Control
- NI = Natural Interaction
- DMS = Driver Monitoring System
- SCR = Sharing Control Request
- MRM = Minimum Risk Maneuver

Figure 101: TeamMate scenario for situation 3) in CRF demonstrator car.

In this case the main research questions of the first part of the scenario concerned the effectiveness at the roundabout (a comfort-related parameter) and its effect in terms on impact on the driver, while the second part of the scenario concerned safety related parameters.

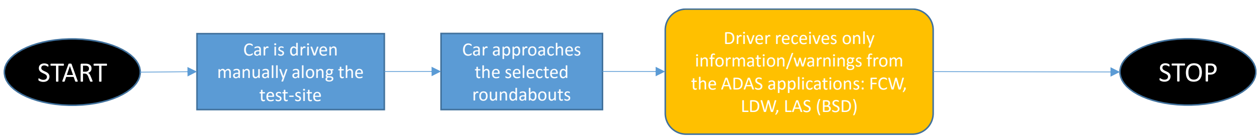
In particular, the following crucial indicators have been taken into account:

- Standard deviation of some indicators, such as position of the ego-vehicle (EV) in the lane, EV speed, Time-To-Collision (TTC) and Headway (HDW).
- The number of safe maneuvers, since this can be considered a safety critical at the roundabout, with a significant impact on traffic situation.

4.6.2 The values of these indicators and how they have been obtained are reported in the next sections, including the comparison with the baseline (BL).Baseline

According to the approach described in the common evaluation framework in D6.1, the Baseline scenario consists in performing the same driving scenario without the enablers, in order to evaluate the added value of the technologies developed in AutoMate. In other words, the CRF baseline car is a Jeep Renegade as currently available on the market, with "only" ADAS applications on-board.

The following figure describes this scenario with a specific flowchart:



Legend:
- ADAS = Advanced Driving Assistance System
- FCW = Forward Collision Warning
- LDW = Lane Departure Warning
- LAS = Lateral Active Support
- BSD = Blind Spot Detection

Figure 102: Baseline scenario in CRF demo-vehicle.

As in the TeamMate scenario, the vehicle driving starts in manual mode, but in this case the whole test is performed like that. Thus, when vehicle approaches the roundabouts, the manoeuvres are performed completely by the driver, without any support from the automation. The driver can only rely on the information and warnings from the ADAS applications, the ones available on the current Jeep Renegade available on the market. In particular, the applications present nowadays are listed below:

- FCW = Forward Collision Warning.
- LDW = Lane Departure Warning.
- LAS = Lateral Active Support (including BSD = Blind Spot Detection).

This means that the driver can receive support for the forward objects (e.g. approaching too close a vehicle ahead in a roundabout) and for the lateral manoeuvres (e.g. change lane for an overtaking, without using the proper indicators).

Also the HMI is the one available on the vehicle, without any distributed concept behind (thus, in case of distraction, there are no specific actions, unless the "normal" warnings from ADAS applications).

4.6.3 Method

Each user had travelled the test-site twice (randomly selected, in avoid any biases):

Baseline (namely, manual, with no assistance).

TeamMate

As for REL evaluation, these tests have been designed as a between-subjects experimental design, i.e. each participant performed the same scenario (of course, taking into consideration that – being a *real-roads* test – the traffic conditions cannot be under control). Twenty subjects (16 males and 4 females) have been recruited for the experiment.

The users were welcomed and asked to sign three consent form modules for the data protection (one for the project, one for CRF internal purposes and one use of images). Then, basics demographics data (e.g. gender, age, driving experience, driving habits) have been collected in order to allow the creation of data clusters. The users were asked to have a 5-minutes trial with the Jeep Renegade vehicle, in order to become familiar with it. Then, they were introduced to AutoMate concept, describing the main pillars of the project:

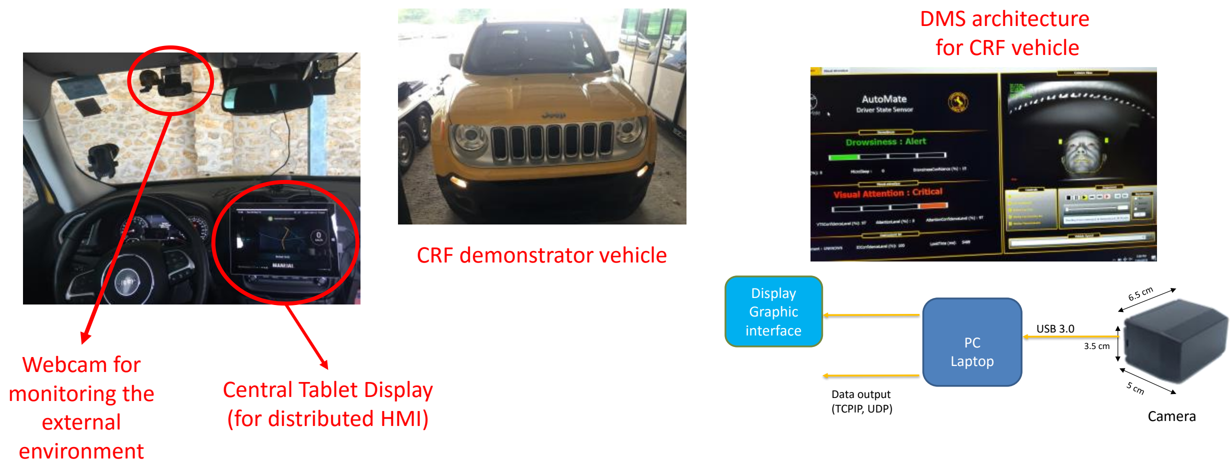


Figure 1034: experimental setup for CRF demonstrator car, including Driver Monitoring System (DMS).

The evaluation focused on measuring mostly safety-, comfort- and acceptability-related parameters. After each scenario, the users were asked to answer a questionnaire aimed at assessing the user-satisfaction in using the TeamMate system compared to a baseline.

The following items were considered the most relevant for this cycle (according to the use case tested in this demonstrator):

- The user acceptance

- The trust in the automated system
- The workload in using the system
- The willingness to buy (and to pay) the system
- The efficiency
- The system performances

As stated in the Common Evaluation framework, the following tools have been used for the first 4 items:

- The Van der Laan questionnaire
- The Koerber questionnaire
- The NASA-TLX
- A custom questionnaire, created ad-hoc to evaluate these propensity

The following table summarizes the KPIs considered in the AutoMate project:

KPI ID	KPI	KPI Type	Recording Tool
KPI1	Time to enter the roundabout	Objective	Vehicle logs
KPI2	Acceptance	Subjective	Van der Laan questionnaire
KPI3	Trust	Subjective	Koerber questionnaire
KPI4	Workload	Subjective	NASA-TLX
KPI5	Willingness to buy	Subjective	Custom questionnaire
KPI6	Willingness to pay	Subjective	Custom questionnaire
KPI7	Time to take over	Objective	Vehicle logs
KPI8	Number of safe manoeuvre	Objective	Vehicle logs

KPI9	Standard Deviation of Speed, TTC, HDW and Position in the Lane	Objective	Vehicle logs
KPI10	Max and mean of speed	Objective	Vehicle logs

Table 18: List of KPIs for REL demonstrator

Also qualitative data have been collected, i.e. users' comments and observations.

4.6.4 Results for subjective Data

In the following table are considered the KPIs used for the CRF demo vehicle:

KPI ID	KPI	Baseline results	TeamMate result	Delta
KPI2	Acceptance	0,76	1,08	+0,32
KPI3	Trust	NA (Baseline is in Manual)	0,72	NA
KPI4	Workload	6,75	7,45	+0,70
KPI5	Willingness to buy	0,55	0,80	+0,25
KPI6	Willingness to pay	1645 €	4300 €	+2655 €

Table 19: subjective KPI results for CRF demonstrator vehicle, in BL and TM modes.

This section illustrates the results for subjective data, next paragraph provide an overview for objective data.

Since the Baseline scenario was performed in Manual Mode, the Trust in automation was measured only in TeamMate modality. The results show satisfactory score in terms of trust: the TeamMate score was +0,72 (on a scale between -2/+2), beyond an acceptability threshold of "0". In particular, the system obtained excellent scores in terms of "Understandability" (+1,05), "Confidence" (+1,00) and judged the system as "Capable of interpreting complex situations" (+1,15).

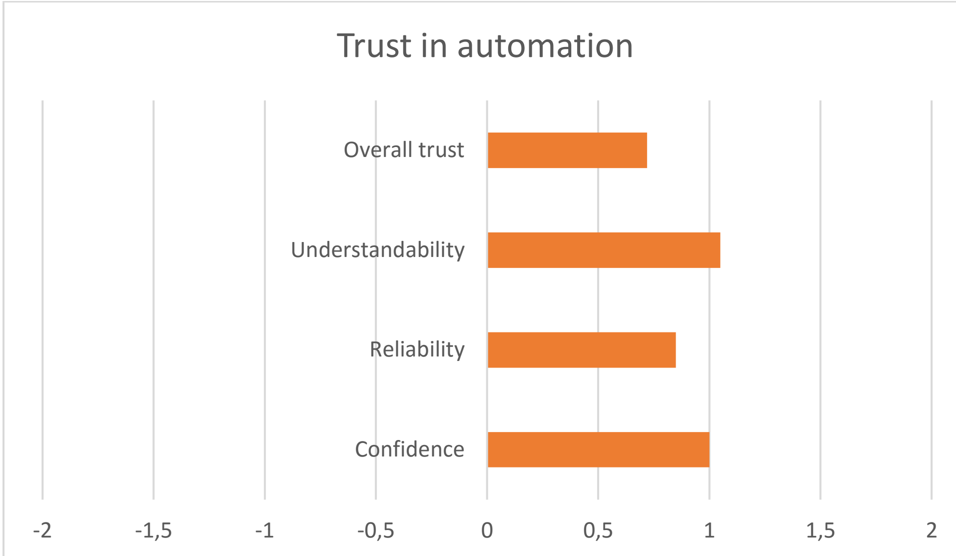


Figure 10405: Trust in automation in EVA scenario (vehicle)

The score related to the workload, collected through the NASA-TLX questionnaire, show that the TeamMate system affect the overall workload, by increasing it of 0,7 (7,45 against the 6,75 of the Baseline). In particular, the "Mental workload" (8,9 against the 7,4 in Baseline Mode) and the "Effort" (7,25 against the 6,4 in Baseline) increased. The "Physical demand" was the only item of the NASA-TLX improved by the TeamMate system.

From comments and observations collected during and after the experiment, this result can be explained by the fact that the TeamMate system requires a longer learning curve and a much less familiar interaction than a traditional manual vehicle. This is confirmed with the observations of the users' behaviour: since the scenario was designed as a ring (one section forward and one return), on the return way the users were significantly more confident in using the TeamMate system, showing that they were able to learn it properly.

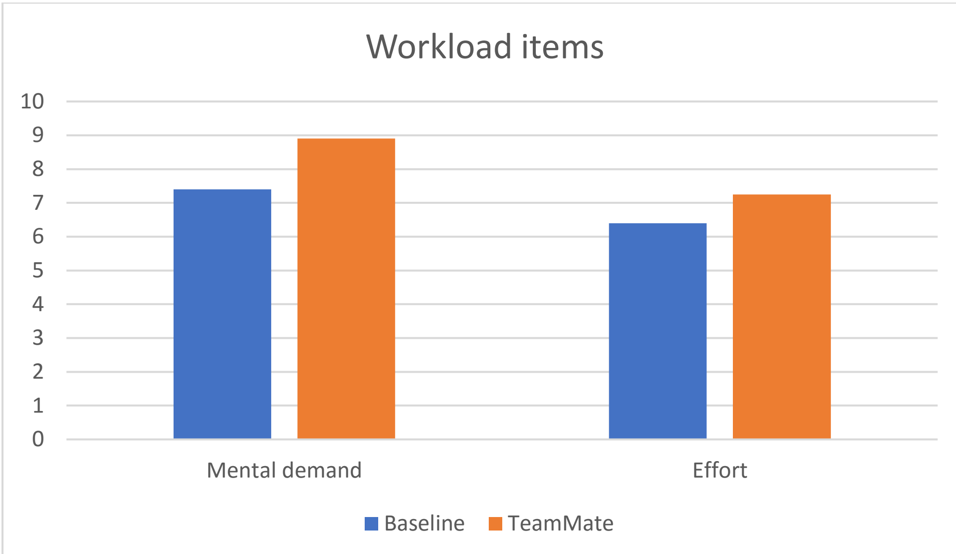


Figure 106: Workload results in EVA scenario (vehicle)

The Van der Laan items, measuring the User acceptance levels, shows a significant improvement of the TeamMate system against the Baseline: the overall score of the TeamMate system was +1,08 against the +0,76 reached by the Baseline. In particular, the users found the TeamMate system as "Desirable" (+1,2 against the +0,55 reached by the Baseline), "Supportive" (+1,40 against the +0,10 reached by the Baseline) "Useful" (+1,35 against the +0,90 reached by the Baseline) and "Effective" (+1,45 against the +1,15 reached by the Baseline). The results of the Acceptance test are of great significance, since they testify the appreciation of the proposed approach.

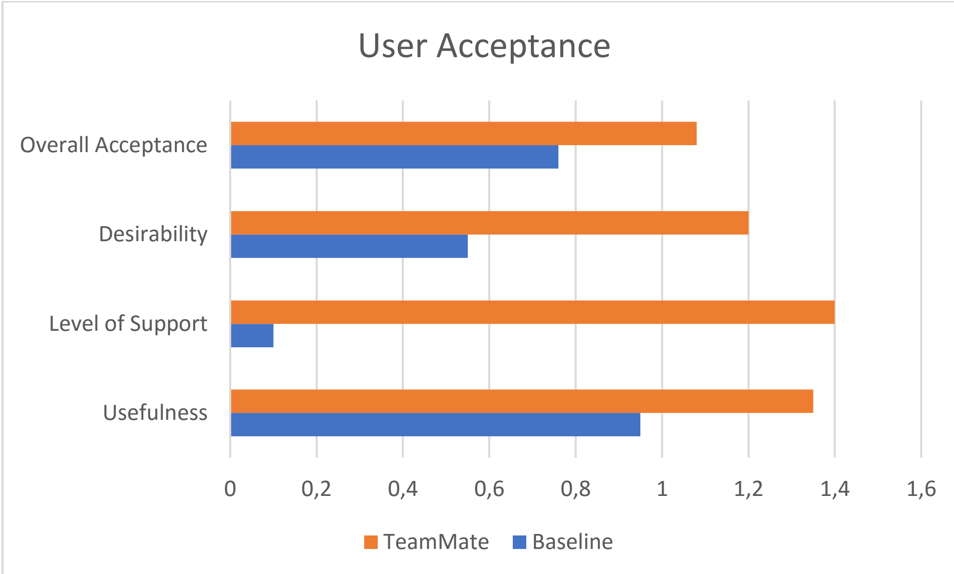


Figure 10507: Acceptance results in EVA scenario (vehicle)

Figure and Figure show, respectively, the "Willing to buy" and the "Willingness to pay" related to the TeamMate system compared to the Baseline.

Figure 108: Willingness to buy results in EVA scenario (vehicle)

The results of the "Willingness to buy" item for the TeamMate system was +0,80 against the +0,55 reached by the Baseline. The "Willingness to pay" results were 4.300 € for the TeamMate system and 1.645 € (on top of the price of the existing vehicle), with a significant increase of 261,4%.

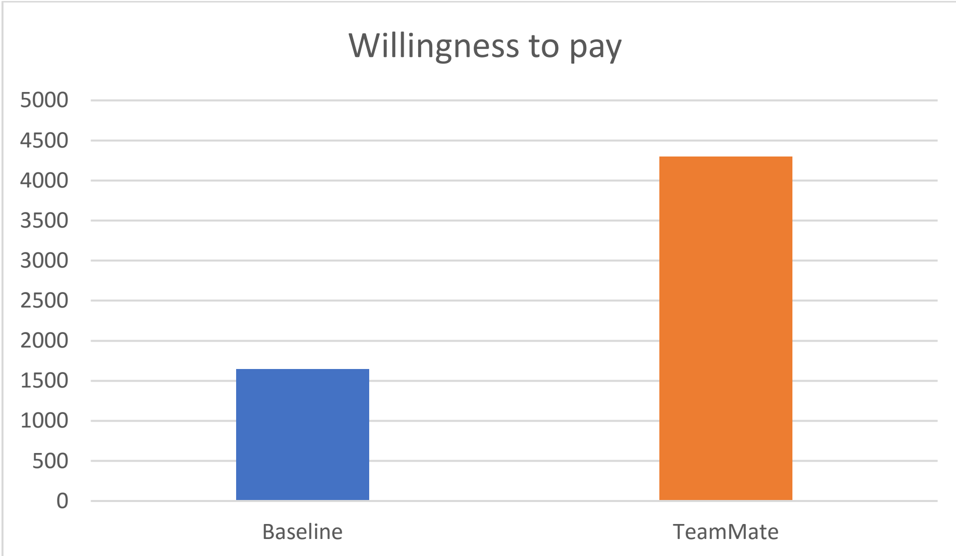


Figure 109: Willingness to pay results in EVA scenario (vehicle)

Since the user sample was not balanced per gender (due to the limitation of the experimental setup and the user recruitment, demanding CRF employees with a special license and not involved in AutoMate project), no insights can be collected about gender preferences.

However, some relevant findings related to the age clustering can be preliminary stated. The users have been divided into three groups (first group with age between 27 and 35, 2nd group between 36 and 44, third group over 45).

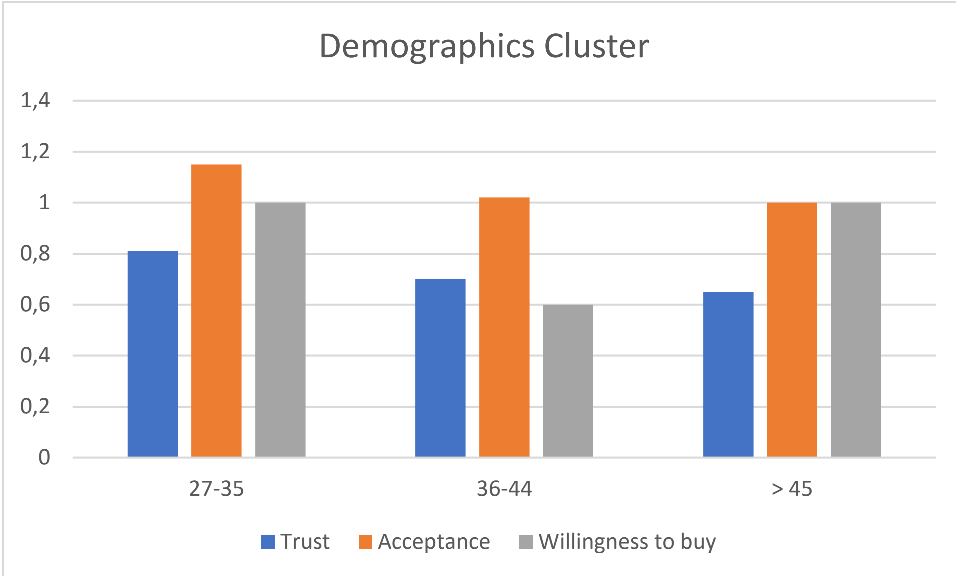


Figure 110: Demographics cluster results in EVA scenario (vehicle)

The results show that the youngest group (27-35 years old) is significantly more willing to trust and accept the TeamMate system, compared to the other groups (e.g. +1,15 of “User acceptance” for the first group against the 1,02 of the second group, and the +0,81 of “Trust” for the first group against the +0,65 of the third group). These results suggest that younger generations could be more willing to adopt highly automated vehicles than older generations, with the related impact on commercialization strategies and intention towards adoption.

4.6.5 Results for objective Data

With reference to the table in the previous paragraph, these are the KPIs considered:

KPI ID	KPI	Baseline results	TeamMate result	Delta
KPI8	Number of safe manoeuvre	NA	5/460	NA
KPI9a	Standard Deviation of Speed	12,5249	10,2919	2.233
KPI9b	Standard Deviation of Position in the Lane	0,6853	0,3422	0.3431
KPI10a	Max speed	94,0812	59,3256	34.7556
KPI10b	Mean of speed	79,4718	66,9556	12.5162

Table 20: objective KPI results for CRF demonstrator vehicle, in BL and TM modes.

As presented in the table, the first KPI of interest is the Number of safe manoeuvre that the system had to act when the driver is not responding in time to a take-over request (TOR). Of course this is applicable only to the TM mode, anyway it is interesting to note that on 460 roundabouts travelled in the evaluation, only 4 safe stops manoeuvre occurred, meaning that people were able to understand the communication from the system and thus act properly (coming back into the control loop, after a TOR).

For the standard deviation, both for position in the lane and for the speed, it is minor in TM mode than in BL mode, indicating that there is less dispersion respect to the mean of the set. Particularly interesting for the lane position, since the vehicle has minor fluctuations respect to the centre-line and thus the lateral driving is more "stable".

Also considering the KPI for the speed, the situation improves with TM. In fact, the mean and also the max are lower, meaning that people go faster when in manual mode that with the system support.

4.6.6 Discussion

In this section, we report a short summary of the main results both for the objective and subjective analysis.

In particular, the TM system has a positive effect: vehicle is more stable around the centre of the lane (less lateral displacement) and the ego-vehicle speed is minor (both as mean and max value), meaning that people respect more the traffic rules (e.g. go faster when driving manually).

The mental workload is worst in TM mode than in BL, but this is quite obvious: people are used to drive manually, thus they need some time to adapt to cooperate with a highly-autonomous driving system. On the other way around, the physical workload is minor, indicating that the interaction with the system requires more cognitive effort at the beginning, but less physical effort. This also suggests that users need a training period before being able to really understand and cooperate with the TM system, thus achieving a good level of trust and confidence. This is proved by the fact that subjects deemed "useful" the system, concerning the level of support.

5 Conclusions

The enablers developed within the AutoMate project are described in their final state and were proofed to improve the defined KPIs. Moreover, the integration and interaction of various enablers was successfully shown in this last evaluation in the different demonstrators.

The previous chapter described the final comparative evaluation of the TeamMate car against the baseline car in the AutoMate use cases of the Eva, Peter, and Martha scenario. The aim of the evaluation was to show the benefits of the different enablers in the bidirectional cooperation between the human and the automation in terms of safety, efficiency, comfort, trust and acceptance.

It was shown that the TeamMate concept improves different aspects of the driving task and could be a promising concept in future cars. Not all evaluation studies could show a significant improvement. This has various reasons such as the evaluation method. A long-term study could highlight the benefits of, for example, the online learning of the driver´s intention in the Peter scenario. Due to safety regulations the experiments for the manoeuvre planning were conducted on a test track and with a safety distance, which was criticized as too conservative and unnatural by the participants. Further research should investigate how participants experience the trajectory in real traffic scenarios and with an adapted safety distance.

Due to the state of technology there were limitations while evaluating enablers such as the augmented reality in the demonstrator vehicles. The developed concepts were proven useful in the simulator studies but were limited by the quality of the display´s hardware.

Overall the TeamMate cooperation seems to improve the safety, efficiency, comfort, trust and acceptance in the specific scenarios. Different scenarios or a higher complexity, such as the combination of the three scenarios, could verify the improvements and possibilities of the AutoMate enablers.

6 References

[1] AutoMate Consortium, „D2.1 - Metrics & Experiments for V&V of the models in the 1st cycle.," 2017.

[2] AutoMate Consortium, „D2.2 - Sensor Platform and Models incl. V&V results from 1st cycle," 2017.

[3] AutoMate Consortium, „D2.3 - Metrics & Experiments for V&V of the models in the 2nd cycle," 2017.

[4] AutoMate Consortium, „D2.4 - Sensor Platform and Models incl. V&V results from 2nd cycle," 2018.

[5] AutoMate Consortium, „D2.5 - Metrics & Experiments for V&V of the models in the 3rd cycle," 2018.

[6] AutoMate Consortium, „D2.6 - Sensor Platform and Models incl. V&V results from 3rd cycle," 2018.

[7] M. Eilers, E. Fathiazar, S. Suck und D. Twumasi, „Dynamic Bayesian networks for driver-intention recognition based on the traffic situation," in *Cooperative Intelligent Transport Systems: Towards High-Level Automated Driving*, 2019 [in press].

[8] A. Doshi, B. T. Morris und M. M. Trivedi, „On-Road Prediction of Driver's Intent with Multimodal Sensory Cues," *Automotive Pervasive Computing,* pp. 22-34, 2011.

[9] B. T. Morris, A. Doshi und M. M. Trivedi, „Lane-Change Intent Prediction for Driver Assistance: On-Road Design and Evaluation," *Proceedings of the 2011 IEEE Intelligent Vehicles Symposium,* pp. 895-901, 2011.

[10] V. Leonhardt, T. Pech und G. Wanielik, „Fusion of Driver Behaviour Analysis and Situation Assessment for Probabilistic Manoeuvre Prediction," in *UR:BAN Human Factors in Traffic*, Wiesbaded, Springer Fachmedien, 2018, pp. 223-244.

[11] D. Manstetten, “Behaviour Prediction and Intention Detection in UR:BAN VIE - Overview and Introduction,” in *UR:BAN Human Factors in Trafic - Approaches for Safe, Efficient and Stress-free Urban Traffic*, Wiesbaden, Springer Vieweg, 2018, pp. 153-162.

[12] AutoMate Consortium, „D5.7 - TeamMate Extension SDK,“ 2019.

[13] AutoMate Consortium, „D6.2 - Results of Evaluation 2nd Cycle,“ 2018.

[14] A. Doshi und M. M. Trivedi, „Tactical Driver Behavior Prediction and Intent Inference: A Review,“ *Proceedings of the 14th International IEEE Conference on Intelligent Transportation Systems,* pp. 1898-1897, 2011.

[15] S. Lefèvre, D. Vasquez und C. Laugier, „A survey on motion prediction and risk assessment for intelligent vehicles,“ *ROBOMECH Journal,* pp. 1-14, 2014.

[16] P. Kumar, M. Perrollaz, S. Lefèvre und C. Laugier, „Learning-based approach for online lane change intention prediction,“ *Proceedings of the 2013 IEEE Intelligent Vehicles Symposium,* pp. 797-802, 2013.

[17] A. Liu und A. P. Pentland, „owards real-time recognition of driver intentions,“ *Proceedings of the Intelligent Transportation System ITSC'97,* pp. 236-241, 1997.

[18] N. Oliver und A. P. Pentland, „Driver Behavior Recognition and Prediction in a SmartCar,“ *Proceedings of the SPIE,* pp. 280-290, 2000.

[19] K. Torkkola, S. Venkatesan und H. Liu, „Sensor Sequence Modeling for Driving,“ *Proceedings of the FLARS Conference,* pp. 721-727, 2005.

[20] C. Tay, Analysis of Dynamic Scenes: Application to Driving Assistance, France: Dissertation, Institut National Polytechnique de Grenoble INPG, 2009.

[21] L. Bi, C. Wang, X. Yang, M. Wang und Y. Lie, „Detecting Driver Normal and Emergency Lane-Changing Intentions With Queuing Network-Based Driver Models,“ *International Journal of Human-Computer Interaction,* pp. 139-145, 2015.

[22] M. Liebner, M. Baumann, F. Klanner und C. Stiller, „Driver Intent Inference at Urban Intersections using the Intelligent Driver Model,“ *Proceedings of the 2012 IEEE Intelligent Vehicles Symposium,* pp. 1162-1167, 2012.

[23] E. Ohn-Bar, A. Tawari, S. Martin und M. M. Trivedi, „Predicting Driver Maneuvers by Learning Holistic Features,“ *Proceedings of the IEEE Intelligent Vehicles Symposium,* pp. 719-724, 2014.

[24] A. Jain, H. S. Koppula, B. Raghavan, S. Soh und A. Saxena, „Car that Knows Before You Do: Anticipating Maneuvers via Learning Temporal Driving Models,“ *Proceedings of the International Conference on Computer Vision,* pp. 3182-3190, 2015.

[25] J. Berndt, J. Emmert und K. Dietmayer, „Continuous Driver Intention Recognition with Hidden Markov Models,“ *Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems,* pp. 1189-1194, 2008.

[26] T. Kumagai und M. Akamatsu, „Prediction of Human Driving Behavior Using Dynamic Bayesian Networks,“ *IEICE Transactions on Information and Systems,* pp. 857-860, 2006.

[27] M. Garcia-Ortiz, J. Fritsch, F. Kummert und A. Gepperth, „Behavior prediction at multiple time-scales in inner-city scenarios,“ *Proceedings of the IEEE Intelligent Vehicles Symposium,* pp. 1068-1073, 2011.

[28] H. M. Mandalia und D. D. Salvucci, „Using support vector machines for lane change detection,“ *Proceedings of the Human Factors and Ergonomics Society,* pp. 1968-1969, 2005.

[29] S. Klingelschmitt, M. Platho, H.-M. Gross, V. Willert und J. Eggert, „Combining behavior and situation information for reliably estimating multiple intentions,“ *Proceedings of the IEEE Intelligent Vehicles Symposium,* pp. 388-393, 2014.

[30] G. S. Aoude, B. D. Luders, K. K. H. Lee, D. S. Levine und J. P. How, „Treat assessment design for driver assistance at intersections,“ *Proceedings of the IEEE Intelligent Transportation Systems Conference,* pp. 25-30, 2010.

[31] M. Muffert, T. Milbich, D. Pfeiffer und U. Franke, „May I enter the roundabout? A time-to-contact computation based on stereo-vision,“ *Proceedings of the IEEE Intelligent Vehicles Symposium,* 2012.

[32] M. Zhao, D. Kathner, M. Jipp, D. Soffker und K. Lemmer, „Modeling driver behavior at roundabouts: Results from a field study," *Proceedings of the 2017 IEEE Intelligent Vehicles Symposium,* pp. 908-913, 2017.

[33] D. Koller und N. Friedman, Probabilistic Graphical Models: Principles and Technique, Cambridge, Massachusetts, London, England: The MIT Press, 2009.

[34] H. Farah, S. Bekhor, A. Polus und T. Toledo, „A Passing Gap Acceptance Model for Two-Lane Rural Highways," *Transportmetrica,* pp. 159-172, September 2009.

[35] U. Hassein, M. Diachuk und S. Easa, „Evaluation of dynamic passing gap acceptance on two-lane highways using field data," *Canadian Journal of Civil Engineering,* pp. 871-880, July 2017.

[36] J. Bobadilla, F. Ortega, A. Hernando und J. Bernal, „A collaborative filtering approach to mitigate the new user cold start problem," *Knowledge-Based Systems,* Bd. 26, pp. 225-238, 2012.

[37] A. Trende, D. Gräfing und L. Weber, „Personalized User Profiles for Autonomous Vehicles," *Proceedings of the AutomotiveUI 2019,* 2019 [in press].

[38] P. W. N. H. Atul Kanaujia, „Markov Logic Networks for Scene Interpretation and Complex Event Recognition in Videos," 2015.

[39] B. Hummel, „Description Logic for Scene Understanding at the Example of Urban Road Intersections," 2009.

[40] M. Huelsen, „Knowledge-based driver assistance systems. Traffic situation description and situation feature relevance," 2014.

[41] T. N. M. M. a. P. H. S. Ulbrich, „Graph-based context representation, environment modeling and information aggregation for automated driving," in *IEEE Intelligent Vehicles Symposium Proceedings*, Dearborn, MI, 2014.

[42] J. M. Z. a. C. W. M. Hülsen, „Traffic intersection situation description ontology for advanced driver assistance," in *IEEE Intelligent Vehicles Symposium (IV)*, Baden-Baden, 2011.

[43] S. Hoermann, D. Stumper und K. Dietmayer, „Probabilistic Long-Term Prediction for Autonomous Vehicles," *Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV),* pp. 237-243, 2017.

[44] J. Ziegler, P. Bender, T. Dang und C. Stiller, „Trajectory planning for Bertha - A local, continuous method," *Proceedings of the 2014 IEEE Intelligent Vehicles Symposium,* pp. 450-457, 2014.

[45] R. Schubert, E. Richter und G. Wanielik, „Comparison and Evaluation of Advanced Motion Models for Vehicle Tracking," *Proceedings of the 11th International Conference on Information Fusion,* pp. 730-735, 2008.

[46] E. A. Wan und R. van der Merwe, „The Unscented Kalman Filter for Nonlinear Estimation," *Proceedings of the IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium,* pp. 153-158, 2000.

[47] M. Schreier, Bayesian Environment Representation, Prediction, and Criticality Assessment for Driver Assistance Systems, Technische Universität Darmstadt: Fachbereich Elektrotechnik und Informationstechnik, 2015.

[48] T. D. Gillespie, Fundamentals of Vehicle Dynamics, Warrendale, PA: Society of Automotive Engineers, Inc., 1992.

[49] P. Polack, F. Altché, B. d'Andréa-Novel und A. de La Fortelle, „The Kinematic Bicycle Model: a Consistent Model for Planning Feasible Trajectories for Autonomous Vehicles?," *Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV),* pp. 812-818, 2017.

[50] J. Kong, M. Pfeiffer, G. Schildbach und F. Borrelli, „Kinematic and Dynamic Vehicle Models for Autonomous Driving Control Design," *Proceedings of the 2015 IEEE Intelligent Vehicles Symposium (IV),* pp. 1094-1099, 2015.

[51] A. Houenou, P. Bonnifait, V. Cherfaoui und Y. Wen, „Vehicle Trajectory Prediction based on Motion Model and Maneuver Recognition," *Proceedings of the 2013 IEEE/RSJ International Conference of Intelligent Robots and Systems,* pp. 4363-4369, 2013.

[52] A. Houénou, P. Bonnifait und V. Cherfaoui, „Risk assessment for collision avoidance systems," *Proceedings of the IEEE 17th International Conference on Intelligent Transportation Systems (ITSC 2014),* pp. 386-391, 2014.

[53] S. Lefèvre, C. Laugier und J. Ibanez-Guzmàn, „Intention-Aware Risk Estimation for General Traffic Situations, and Application to Intersection Safety," 2013.

[54] S. Shalev-Shwartz, S. Shammah und A. Shashua, „On a Formal Model of Safe and Scalable Self-driving Cars," 2017. [Online]. Available: https://arxiv.org/pdf/1708.06374.pdf.

[55] C. Hubmann, M. Becker, D. Althoff, D. Lenz und C. Stiller, „Decision Making for Autonomous Driving considering Interaction and Uncertain Prediction of Surrounding Vehicles," *Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV),* pp. 1678-1685, 2017.

[56] G. S. Aoude, B. D. Luders, D. S. Levine und J. P. How, „Threat-aware path planning in uncertain urban environments," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* pp. 6058-6063, 2010.

[57] C. Laugier, I. Paromtchik, M. Perrollaz, M. Yong, J. Yoder, C. Tay, K. Mekhnacha und A. Negre, „Probabilistic analysis of dynamic scenes and collision risks assessment to improve driving safety," *IEEE Intelligent Transportation Systems Magazine,* Bd. 3, Nr. 4, pp. 4-19, 2011.

[58] B. Kim, K. Yi, C. Jeong und J. Kim, „Probabilistic states prediction algorithm using multi-sensor fusion and application to Smart Cruise Control systems," *Proceedings of the 2013 IEEE Intelligent Vehicles Symposium (IV),* 2013.

[59] J. Schulz, C. Hubmann, J. Löchner und B. Darius, „Multiple Model Unscented Kalman Filtering in Dynamic Bayesian Networks for Intention Estimation and Trajectory Prediction," *10.1109/ITSC.2018.8569932,* 2018.

[60] AutoMate Consortium, „D3.1 - Metrics & plan for V&V of concepts and algorithms 1st cycle," 2017.

[61] AutoMate Consortium, „D3.3 - Concepts and algorithms incl. V&V results from 1 st cycle," 2017.

[62] AutoMate Consortium, „D3.4 - Metrics & plan for V&V of concepts and algorithms 2 nd cycle," 2017.

[63] AutoMate Consortium, „D3.5 - Concepts and algorithms incl. V&V results from 2 nd cycle," 2018.

[64] AutoMate Consortium, „D3.6 - Metrics & plan for V&V of concepts and algorithms 3 rd cycle," 2018.

[65] AutoMate Consortium, „D3.7 - Concepts and algorithms incl. V&V results from 3 rd cycle," 2019.

[66] M. Hasenjäger und H. Wersing, „Personalization in Advanced Driver Assistance Systems and Autonomous Vehicles: A Review," in *2017 IEEE20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017.

[67] V. Butakov, *Personalized driver assistance systems based on driver/vehicle models,* Los Angeles, California: University of Southern California, 2014.

[68] A. Sarkheyli-Hägele und D. Söffker, „Online learning for an individualized lane-change situation recognition system applied to driving assistance," in *IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, Savannah, GA, 2017.

[69] J. Geukes, „Personalized Driving Assistance to Predict Lane Change Manoeuvres," Darmstadt, 2016.

[70] J. Wiest, M. Karg, F. Kunz, S. Reuter, U. Kressel und K. C. J. Dietmayer, „A probabilistic maneuver prediction framework for self-learning vehicles with application to intersections".

[71] V. Losing, B. Hammer und H. Wersing, „Personalized maneuver prediction at intersections," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017.

[72] J. Salas, H. Jimanez, J. Gonzalez und J. Hurtado, „Detecting unusual activities at vehicular intersections," *Proceedings of the IEEE international conference on robotics and automation,* pp. 864-869, 2007.

[73] S. Lefèvre, C. Laugier und J. Ibanez-Guzmán, „Risk Assessment at Road Intersections: Comparing Intention and Expectation," *Proceedings of the IEEE Intelligent Vehicles Symposium,* pp. 165-171, 2012.

[74] G. Xie, X. Zhang, H. Gao, L. Qian, J. Wang und U. Ozguner, „Situational Assessments Based on Unertainty-Risk Awareness in Complex Traffic Scenarios," *Sustainability,* pp. 1-17, 2017.

[75] A. Lawitzky, D. Wollherr und M. Buss, „Maneuver-based Risk Assessment for High-Speed Automotive Scenarios," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),* pp. 1-6, 2012.

[76] C. L. e. al., „Probabilistic Analysis of Dynamic Scenes and Collision Risks Assessment to Improve Driving Safety," in *IEEE Intelligent Transportation Systems Magazine*, 2011.

[77] V. W. a. J. A. M. Schreier, „From grid maps to Parametric Free Space maps — A highly compact, generic environment representation for ADAS,“ in *2013 IEEE Intelligent Vehicles Symposium (IV)*, 2013.

[78] S. B. a. U. H. T. Weiherer, „An interval based representation of occupancy information for driver assistance systems,“ in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 2013.

[79] F. D. a. J. Eggert, „Predictive risk maps,“ in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Qingdao, 2014.

[80] M. Götze und K. Bengler, „Urban Driving: Where to Present What Types of Information – Comparison of Head-Down and Head-Up Displays,“ in *Yamamoto S. (eds) Human Interface and the Management of Information. Information and Knowledge in Context*, Bd. 9173, Springer International Publishing, 2015.

[81] M. Wittmann, M. Kiss, P. Gugg, A. Steffen, M. Fink, E. Pöppel und H. Kamiya, „Effects of display position of a visual in-vehicle task on simulated driving,“ *Applied Ergonomics,* Bd. 37, Nr. 2, pp. 187-199, 2006.

[82] M. Smith, J. L. Gabbard, G. Burnett und N. Doutcheva, „The Effects of Augmented Reality Head-Up Displays on Drivers' Eye Scan Patterns, Performance, and Perceptions,“ *International Journal of Mobile Human Computer Interaction (IJMHCI),* pp. 1-17, 2017.

[83] M. Plavšić, H. Bubb, M. Duschl, M. Tonnis und G. Klinker, „Ergonomic Design and Evaluation of Augmented Reality Based Cautionary Warnings for Driving Assistance in Urban Environments,“ 17th World Congress on Ergonomics, International Ergonomics Association, 2009.

[84] S.Langlois und B.Soualmi, „Augmented reality versus classical HUD to take over from automated driving: An aid to smooth reactions and to anticipate maneuvers,“ in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Rio de Janeiro, Brazil, 2016.

[85] P. Rane, H. Kim, J. L. Marcano und J. L. Gabbard, „Virtual Road Signs: Augmented Reality Driving Aid for Novice Drivers,“ *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* Bd. 60, Nr. 1, pp. 1750-1754, 2016.

[86] V. Charissis und S. Papanastasiou, „Artificial Intelligence Rationale for Autonomous Vehicle Agents Behaviour in Driving Simulation Environment,“ in *Advances in Robotics, Automation and Control*, Vienna, Austria, IntechOpen, 2008, p. 472.

[87] L. Halit, A. Kemeny, H. Mohellebi, S. Garbaya, F. Merienne, S. Michelin und V. Albaut, „Augmented Reality Head-Up-Display for Advanced Driver Assistance System: A Driving Simulation Study,“ in *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014*, At Kraków, Poland, 2014.

[88] M. &. T. M. T. Moshagen, „Facets of visual aesthetics,“ *International Journal of Human-Computer Studies,* Bd. 10, pp. 689-709, 2010.

[89] M. Eilers, E. Fathiazar, S. Suck und D. Twumasi, „Dynamic Bayesian networks for driver-intention recognition based on the traffic situation,“ in *Cooperative Intelligent Transport Systems: Towards High-Level Automated Driving*, In press, 2019.