Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.

HEIDI

Report on Evaluation Tests

Deliverable Number	D7.1
Deliverable Type	R – Document, Report
Dissemination Level	PU (Public)
Author(s)	A. Sjörs Dahlman (VTI), P. Hoffmann (BMW), R. Leute (MAR), J. Edelbrunner (NISYS)
Document Version & Status	V1.0 | Final

Project Acronym	HEIDI
Project Title	Holistic and adaptivE Interface Design for human-technology Interactions
Grant Agreement Number	101069538
Project Coordinator	Virtual Vehicle Research GmbH
Project Website	https://heidi-project.eu/

HEIDI has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101069538.

Author(s)

Name	Organisation
Robert Leute	MAR
Anna Sjörs Dahlman	VTI
Johann Edelbrunner	NISYS
Philipp Hoffmann	BMW

Reviewers

Name	Organisation	Date
Paolo Pretto	VIF	2025-04-29
Miguel Sotelo	UAH	2025-04-29

Change History

Version	Date	Name/Organisation	Description
0.1	2025-03-05	Leute / MAR	creation
0.2	2025-04-22	Hoffmann / BMW	edited
0.3	2025-04-29	Edelbrunner / NIS	edited
1.0	2025-04-30	Hoffmann / BMW	Edited due to reviewers input / final version

Table of Contents

List of Figures

List of Tables

1. Executive Summary

This delivery presents the results of the evaluation tests of the HMIs developed over the course of the HEIDI project. This includes real world studies 8 and 10, which evaluate the efficiency and performance of internal and external HMIs respectively. The latter was performed by MAR in collaboration with RUAS, the former by BMW in collaboration with NISYS.

Both studies used the prototypes as described in D6.1, fully integrated in test vehicles.

Study 12 which was also done as part of WP7 evaluates the cooperative HMI and was done as a simulator study by VTI in collaboration with HRI-EU.

All evaluation tests strongly profited from results obtained during WP5 and presented in deliverable D5.3, which presented small-scale simulator studies and provided insight into validation and evaluation methods for HMIs.

This deliverable includes full descriptions of the methods used in the respective studies as well as results and conclusion from each study. It is therefore in continuation of WP2, WP3 and WP5 and gives a well-grounded evaluation of these systems.

Study 8 was designed to assess the impact of iHMIs on pedestrian safety and traffic flow during pedestrian-vehicle interaction from the driver's perspective. The study took place within a cordoned-off test area. The main use case of an pedestrian using an indicated crosswalk (zebra crossing) was observed from the perspective of the driver while heading towards it in a prepared test vehicle. The study had 16 participants and several hundred vehicle pedestrian interactions.

The objective of Study 10 was to evaluate the impact of eHMIs on pedestrian safety and traffic flow during the interaction between pedestrian and vehicle. The study took place inside the Motion Capture System of the AIDA Hall of partner RUAS at Innoport in Reutlingen, allowing measurements with both spatial and temporal high resolution. We investigated the use case of a pedestrian intending to cross a road without having priority, i.e. no zebra crossing, as described in deliverable D1.2. The study had 27 participants in two age groups, one younger and one older vulnerable road users (VRUs) and we investigated well over 1000 vehicle-pedestrian interactions.

The sophisticated and realistic setup allowed concise conclusions regarding traffic flow and pedestrian safety. Measurements of the MoCap System were supported with questionnaires, where feedback from the participants was collected.

The objective of study 12 was to evaluate the cooperative HMI (cHMI) with regular drivers and regular and older pedestrians. The evaluation was done using distributed simulation with two agents at the same time, one driver in a driving simulator with XR setup and one pedestrian wearing a VR headset and motion trackers around their waist and ankles. Two test blocks were performed, one with cHMI and one without cHMI. Each test block had 8 interaction scenarios centered around an unsignalized crossing. Evaluation of efficiency and usability was done using objective measures collected via the simulators and trackers, and subjective measures collected via questionnaires and semi-structured interviews.

For all three studies the combination of both objective, i.e. time-to-resolve (TTR), time-to-collision (TTC), etc., and subjective, i.e. questionnaires, validation tools gives a broad and close to complete picture for validating the effectiveness of HMIs. TTC (Time to Contact) refers to the amount of time that elapses between a driving test vehicle and a pedestrian walking (or standing). It is a measured value, which could be calculated from the (constant) speed of the test vehicle (km/h) and the remaining distance from test vehicle to pedestrian in meter (m). The TTR (time to resolve) describes the duration in seconds (s) within a given test situation from beginning of a warning (t=0) to bringing the test vehicle to a complete stop (v=0km/h).

It is noted however, that the three studies are all evaluating distinct and different aspects of the wide thematic field of HMIs and their roles in traffic situations. They have different roles of participants in each study, i.e. pedestrian vs. driver, different environments, i.e. simulator vs. controlled real world indoor setting vs real traffic and different types of HMIs, iHMI vs eHMI vs cHMI. While they all show similar conclusions, they are certainly to be interpreted as complementing each other and NOT being directly compared to each other.

Keywords: iHMI, eHMI, cHMI, simulator studies, real world studies, validation methos, older vulnerable road users

2. Objectives

The work described in this deliverable is divided into three distinct studies which evaluate internal HMIs, external HMIs and cooperative HMIs respectively.

The study on internal HMIs (iHMI), study 8, was done under the lead of BMW, the study on external HMIs (eHMIs), study 10, was done under the lead of MAR and the study on cHMIs, study 12, was done under the lead of VTI.

2.1 Study 8

The objective of this study was to evaluate the internal fluid human-machine interface (iHMI), previously examined in simulation under Work Package D.2, by implementing it within an actual vehicle prototype. The iHMI continuously monitors the driver's focus in relation to the surrounding traffic and intervenes only by issuing notifications or warnings when momentary distractions are detected. This targeted intervention strategy is intended to distil the information provided to the driver down to its most essential elements, while simultaneously enhancing pedestrian safety by alerting the driver precisely when their inattention poses a risk to an appropriate traffic response.

2.2 Study 10

Study 10 aims to assess the impact of enhanced communication between vehicles and pedestrians, focusing on Vulnerable Road Users (VRUs), to improve urban traffic efficiency and safety using an external Human-Machine Interface (eHMI). By analysing interactions between vehicles and pedestrians, the study seeks to gain a deeper understanding of the information flow in traffic and identify potential measures for advancing future urban mobility. It focuses on evaluating the effectiveness of an eHMI in displaying specific symbols on the front of a test vehicle to assist pedestrians in making safer and quicker decisions about crossing a street, thereby potentially enhancing overall traffic flow. Additionally, the study will determine whether clear messages communicated via the eHMI, such as "I am slowing down," "I am stopping," or "I am continuing," influence pedestrian behaviour to facilitate more rapid and secure decision-making.

2.3 Study 12

The purpose of the study was to evaluate the usability and effectiveness of the cHMI with regular drivers and regular and older pedestrians. It focuses on the interaction between a freely acting human driver and pedestrian at an unsignalized crossing. The cooperative HMI synchronizes driver information and information from other road users to facilitate an optimal joint action between the actors. The use of the cHMI system was hypothesized to lead to a quicker and safer resolution of ambiguous traffic situations.

3. Validation Studies

The methods used in the three small-scale validation studies are described below. The studies differed in their design depending on the objectives and research questions of the respective study and the context in which it was conducted.

3.1 Study 8 –Evaluation test of iHMI

3.1.1 Experimental setup

Study 8 was conducted on a purpose-built professional test track. At a designated pedestrian crossing, a test pedestrian - guided by a controlled traffic light signal - was instructed to either step onto the crosswalk or take cover behind a parked vehicle, contingent on the specific test scenario. Simultaneously, drivers operating the prototype vehicles were directed to come to a complete stop when the pedestrian was visible, thereby ensuring a safe opportunity for the pedestrian to cross. For the execution of the study, real-time processing of GPS data, vehicle CAN information, and driver monitoring signals was employed to generate seamless iHMI outputs, to present targeted queries regarding driver distraction, and to dispatch trigger signals to the pedestrian traffic light. All of these data streams were meticulously synchronized and recorded for subsequent validation.

Figure 3-1: Test route displayed in OpenStreetMap, showing the position of the pedestrian, the position of the vehicle, the vehicle's trajectory (red line), and the test area (orange rectangle). All measurements were conducted within the test area. The total length of the test track is approximately 100 metres, with the measuring area spanning about 70 metres.

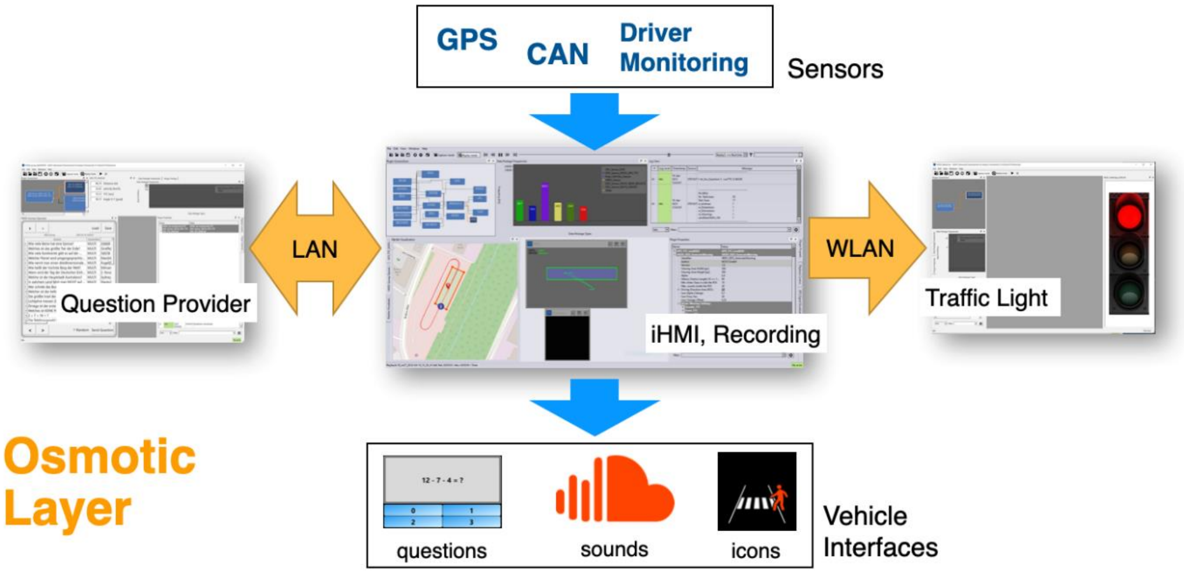


Figure 3-2: Schematic representation of the Osmotic Layer used in the test vehicle.

Figure 3-2 illustrates a schematic representation of the osmotic layer within the test vehicle. Sensor data (encompassing GPS, CAN, and driver monitoring inputs) are conveyed to the (in vehicle) iHMI computer via either USB or Ethernet. Subsequently, the vehicle iHMI computer channels this information to the output devices (namely, the TFT display and loudspeaker) through the iHMI interface or an additional USB connection. Furthermore, the osmotic layer underpins communications with the pedestrian display device (traffic lights), manages the transmission of driver distraction inquiries and their corresponding responses, and facilitates the data exchange between the iHMI modules and the recording module.

Figure 3-3: Test setup in the area of the pedestrian crossing and the interior of the test vehicle with the TFT for visualising the questions on distraction to the driver as well as the control screen for the test supervisor.

Figure 3-3 illustrates the test setup. In the left-hand image, the scene at the pedestrian crossing is depicted: a test pedestrian, concealed behind a grey vehicle, is prompted by a green traffic light signal to cross at the crosswalk. The test subjects operating the vehicle were initially instructed to bring the vehicle to a stop at the yellow line marking the crosswalk whenever the pedestrian became visible, thereby ensuring the pedestrian’s safe passage. On the right-hand side of Figure 3-3, the interior of the test vehicle is shown, featuring both the display used for presenting questions on driver distraction and the interface for the test supervisor.

3.1.2 Use Case

The test cases were deliberately designed to obscure any direct correlation between a distraction alert and the appearance of a pedestrian. Figure 3-4 displays the potential scenarios (Cases 1 through 4), while Cases 5 to 8 were omitted either because they were not pertinent to the experiment or would have indicated a system error within the iHMI.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8
Pedestrian	+	+	+	-	+	-	-	-
(Distraction) Questions	+	+	-	+	-	-	-	+
Warning	+	-	-	-	+	-	+	+
Case Description	iHMI - driver distraction - warnings	iHMI - driver distraction - NO warnings	iHMI - NO driver distraction - NO warnings	normal driving situation - driver distraction	HMI without driver monitoring	normal drive situation	System Error	System Error
System description	iHMI System	current systems	current systems	current systems	current systems	current systems		
Driver	driver does NOT observe the situation	driver does NOT observe the situation	driver observes the situation	driver does NOT observe the situation	gaze direction is not considered			
TEST CASES covered	+	+	+	+				

Figure 3-4: Combinatorial creation of test cases

3.1.3 iHMI solution

The objective of the study was to evaluate the effectiveness of the Fluid-iHMI. Unlike a conventional HMI that simply warns of dangerous situations, the Fluid-iHMI takes the driver's state into account and issues warnings only when the driver is distracted. Driver state monitoring was accomplished using eye-tracking, and the alert system operates in three progressive stages, each distinguished by unique auditory signals and icons (see Figure 3-5). Specifically, the first warning is triggered when the time-to-collision (TTC) drops below 10 seconds, the second when TTC falls under 7 seconds, and the final warning, or an instruction to stop the vehicle, when TTC is less than 3 seconds.

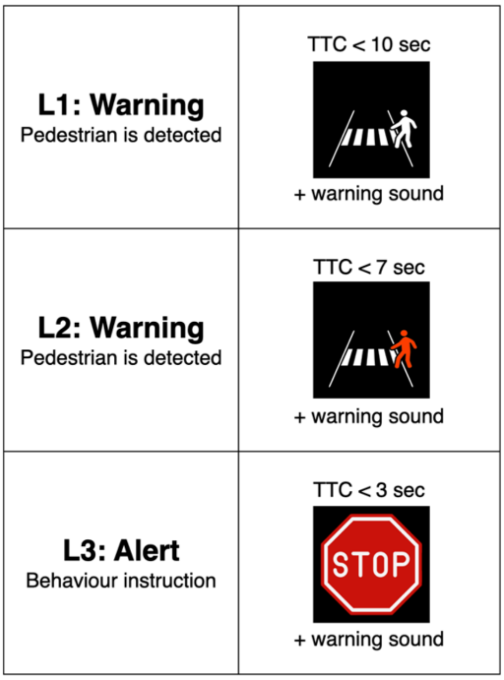


Figure 3-5: Overview of the iHMI design to be tested

In contrast to simulator tests reported in D5.3 “Results and conclusions from the validation tests”, there was no preliminary indication of the crosswalk’s location provided, as the known test setup rendered such information superfluous. Furthermore, the warning signals, both auditory and visual, ceased immediately once the driver’s gaze, as tracked by the eye-tracking system, returned to the traffic situation.

Figure 3-6: Questions about driver distraction and warning symbols

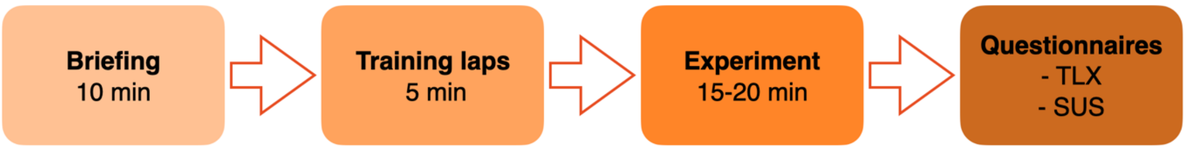
3.1.4 Procedure

All tests were conducted with strict safety measures to preclude any hazardous situations arising from driver error. Notably, the pedestrian was instructed not to cross the road but rather to position themselves safely at a distance from the crosswalk to ensure visibility, while the driver was required to allow the pedestrian a safe passage.

3.1.4.1 Experimental Procedure

In addition to the primary test subject, two personnel were involved—the test supervisor seated in the vehicle and the individual performing the role of the pedestrian. Before the tests commenced, the entire procedure was thoroughly explained to all participants. The test subjects then navigated two familiarization laps in the presence of the test supervisor to become accustomed to the vehicle. For the trials, the supervisor initiated a predefined protocol that randomly generated various use cases over the course of 20 laps, with the traffic light signals for the pedestrian transmitted via WLAN.

3.1.4.2 Participants Procedures



Upon arrival, participants were warmly greeted by the test administrators and informed that their participation was entirely voluntary, with the option to withdraw at any time. They were also assured that their identities would remain confidential. After verifying their driving licenses and familiarizing them with the testing procedures, participants were instructed to adjust their seating to their optimal position prior to commencing the training laps. Upon completion of these laps, they were asked whether additional practice was desired or if they wished to proceed to the experimental phase. At the conclusion of the experiment, the participants were requested to complete two questionnaires (TLX and SUS) to provide evaluative feedback on their experience [3][4].

3.1.5 Participants

A total of 16 participants were recruited. The average age of the male participants (11) was 29.7 ± 6.7 years and the average age of the female participants (5) was 32.2 ± 5.9 years.

3.1.6 Data collection

During the tests, various data streams were recorded in a synchronized manner. The positions of the test vehicle and the pedestrian were determined using GPS data. The gaze direction of the driver was tracked using an eye-tracking system, while vehicle state data (accelerator position, braking status, speed) was captured via the CAN bus.

This data was also utilized in real time to trigger questions for the test subjects at appropriate moments and to generate warnings depending on the driving situation. Throughout the test route, to induce driver‘s distraction were presented independently of any warnings generated by the system under test (iHMI) in order to minimize positive correlations between driver queries and system warnings.

Figure 3-7 and Figure 3-8 present time series plots of the recorded data. In these plots:

• The **red curve** represents the warning level (ranging from 1 to 3),

• The **green curve** indicates the driver’s attention state (distracted or attentive),

• The **blue curve** shows the status of driver questions (whether active or already answered).

When a question appeared, it was accompanied by an acoustic signal in addition to the visual display. The plots also include the Time-to-Collision (TTC) and the distance to the pedestrian.

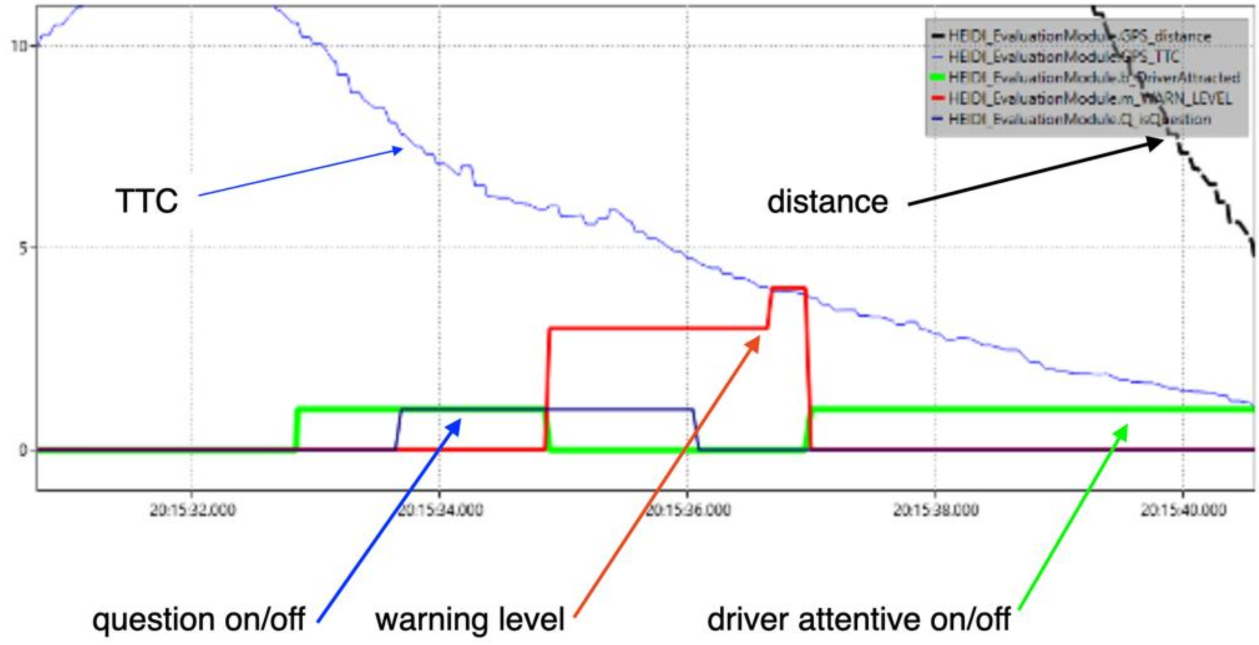


Figure 3-7: Time series plot of iHMI-Data

Figure 3-8 provides further detailed information beyond that shown in Figure 3-7, including:

- The vehicle’s position relative to the designated measurement area (“inside area on/off”),
- Accelerator pedal angle,
- Brake status (on/off),
- Vehicle speed.

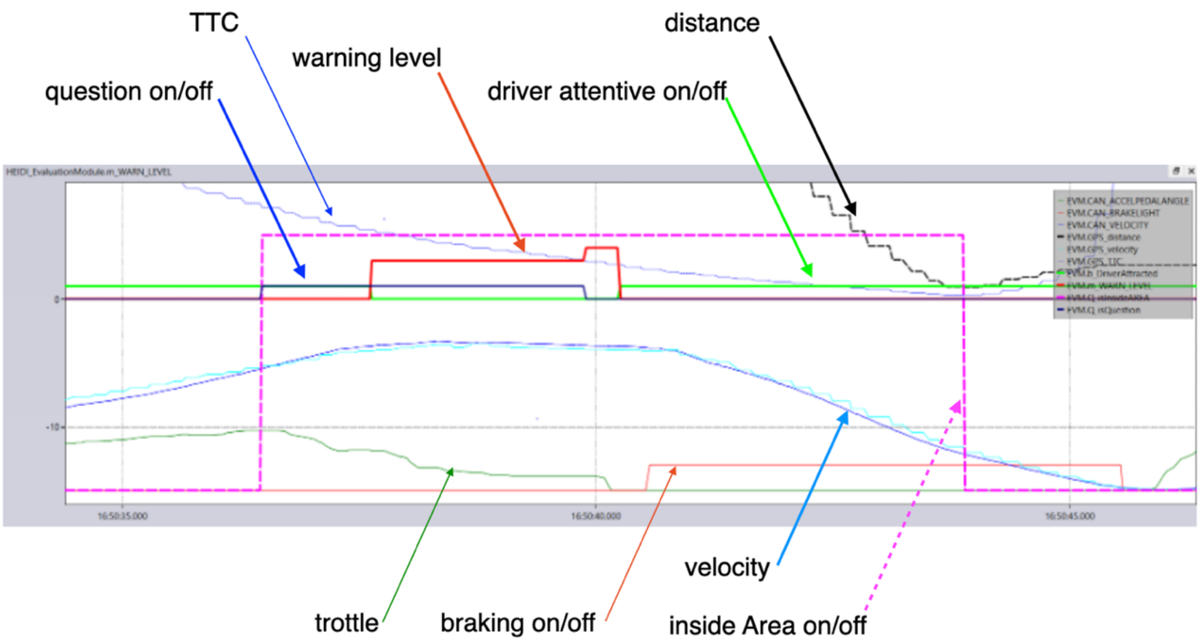


Figure 3-8: Time series plot of iHMI and vehicle state data

3.1.7 Results

The results are divided into two sections: subjective measurements based on questionnaire responses and objective results based on sensor data.

3.1.7.1 Subjective Measures (Data from questionnaires)

To enable a direct comparison between the results obtained with the real test vehicle (Figure 3-3) and those from the simulator study (linked to Document D2.3 'Initial Simulator Studies Results' and further results in D5.3), the same set of evaluation questions for the iHMI system was used.

Notable differences emerged in certain responses. For Question 6 ("I think it is dangerous to use the system"), participants in the real-vehicle test rated the system approximately 1.5 points lower in perceived danger compared to the simulator test. A possible explanation for this deviation lies in the differences in test design: in the real vehicle, the system featured fewer warning levels, and simple arithmetic tasks with predefined answer options were used to induce distraction. In contrast, the simulator test employed the input of words as the distraction task, which may have had a greater cognitive load and perceived risk. Similarly, differences were observed for Question 24 ("The system made it easier for me to recognise pedestrians around me in time"). Here, the rating in the simulator study was approximately one point higher than in the real-vehicle test. This may be attributed to the simpler and more controlled test environment of the real-world test, where the locations of potential pedestrians were known in advance, potentially reducing the perceived value of the system's assistance.

The NASA-TLX results show a high degree of satisfaction with task completion, but also considerable mental strain and effort. This increased cognitive load may partly be due to answering the questions, which, in this test, were only intended as a secondary task to induce distraction.

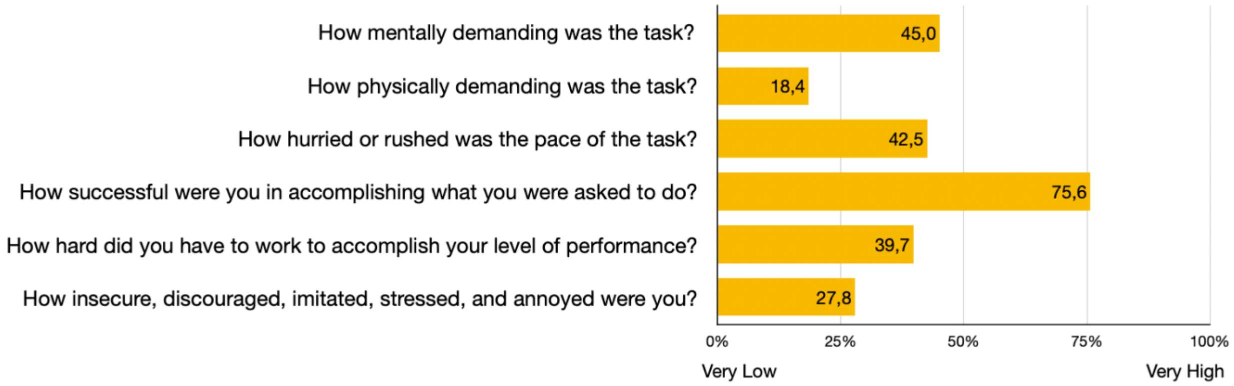


Figure 3-9: Results of the NASA-TLX Questionnaire on Task Load and Effort

Figure 3-10: Average ratings of the various statements about the iHMIs

3.1.7.2 Objective Measures

During the test drives, all relevant sensor data (GPS, CAN, driver monitoring), the states of the iHMI, as well as the start and end times of questions and warnings, were synchronously recorded in Hierarchical Data Format Five (.h5) [1] using the NISYS framework. The recorded data was subsequently analysed through an automated process.

For the evaluation, the warning levels, defined by Time-to-Collision (TTC) thresholds (Warn Level 1: TTC < 10 seconds, Warn Level 2: TTC < 7 seconds, and Warn Level 3: TTC < 3 seconds), served as the primary reference points. The TTC values at the initiation of braking manoeuvres were analysed with respect to two scenarios:

Case 1: Pedestrian presence combined with a distraction task and a warning — examining braking initiation in response to the system warning under distraction conditions.

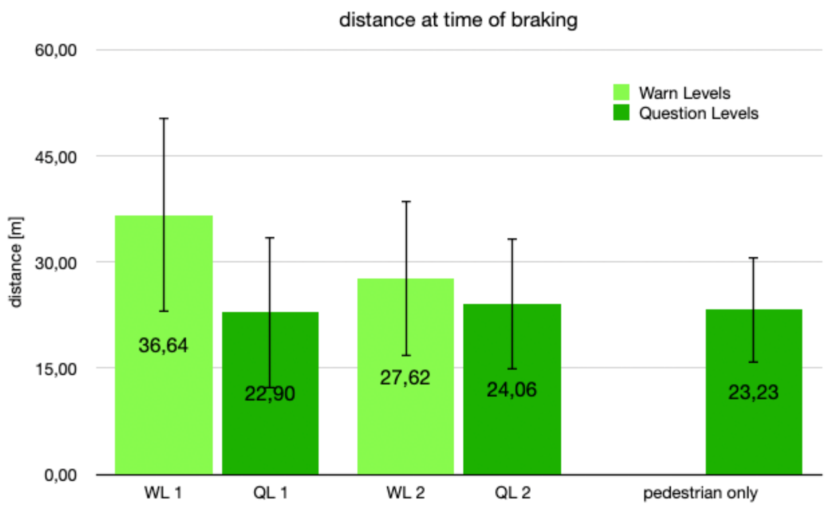
Case 2: Pedestrian presence combined with a distraction task (question) — examining braking initiation in response to the presence of a pedestrian while the driver was distracted.

Baseline values were derived from **Case 3** (pedestrian presence only), which represented 'normal behaviour' where the driver, undistracted, recognised the pedestrian and initiated a standard braking manoeuvre.

Due to system parameterisation, test cases from Case 2 could not be directly compared with those from Case 1. Table 3-1 presents the number of recorded instances for each test case.

Table 3-1: Average ratings of different statements about the iHMIs

Test Case	Number Events
Case 1	102
Case 2	70
Case 3	48
Sum	220



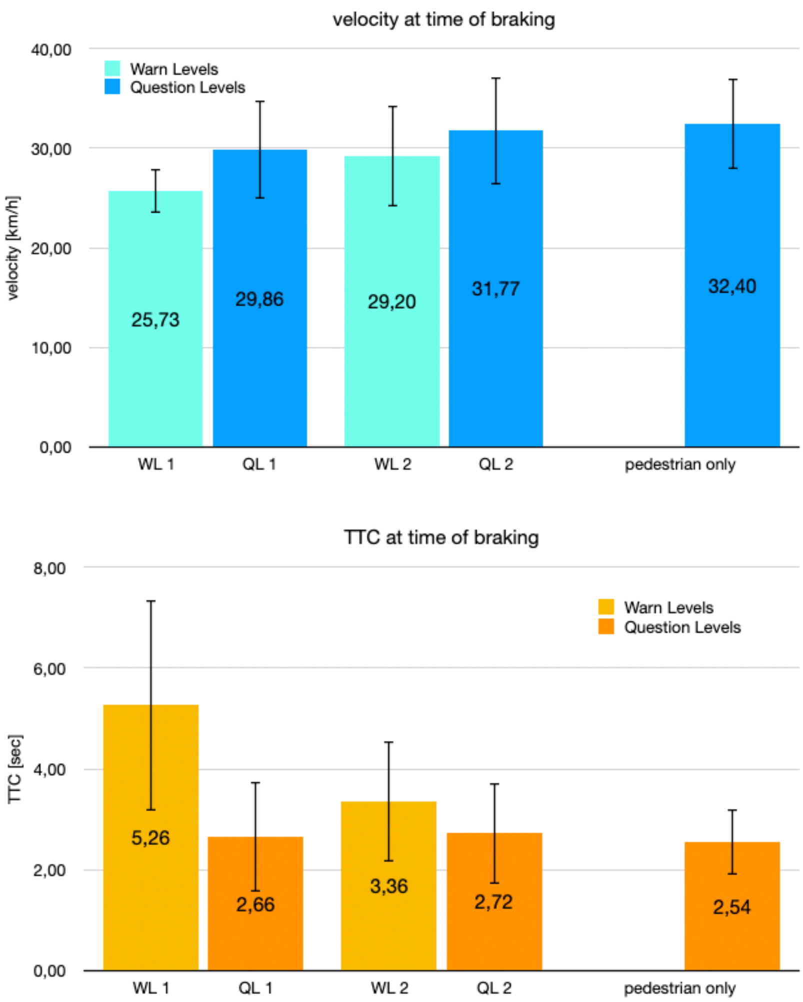


Figure 3-11: Comparison of measurement results for Case 1, 2, 3; "Normal Behaviour", Questions and Warnings

Table 3-2: Correlations between Case 1, Case2 and Case 3 for Warning Levels 1 and 2with focus on TTC

	p-value	t-value	df
pedestrian only / QL1	0,52	0,64	88
pedestrian only / QL2	0,34	0,95	74
pedestrian only / WL2	< 0,0001	4,35	116
pedestrian only / WL1	< 0,0001	5,98	49

Table 3-2 presents the results of the statistical analysis. As the results indicate, there is no significant difference in the Time-to-Collision (TTC) between 'Normal Behaviour' (Case 3: pedestrian only) and Case 2 (pedestrian with distraction through a question) at both warning levels (QL1: TTC < 10 seconds and QL2: TTC < 7 seconds). However, when a distraction was detected based on the question (via eye-tracking) and a warning was subsequently triggered, the braking manoeuvre was initiated significantly earlier.

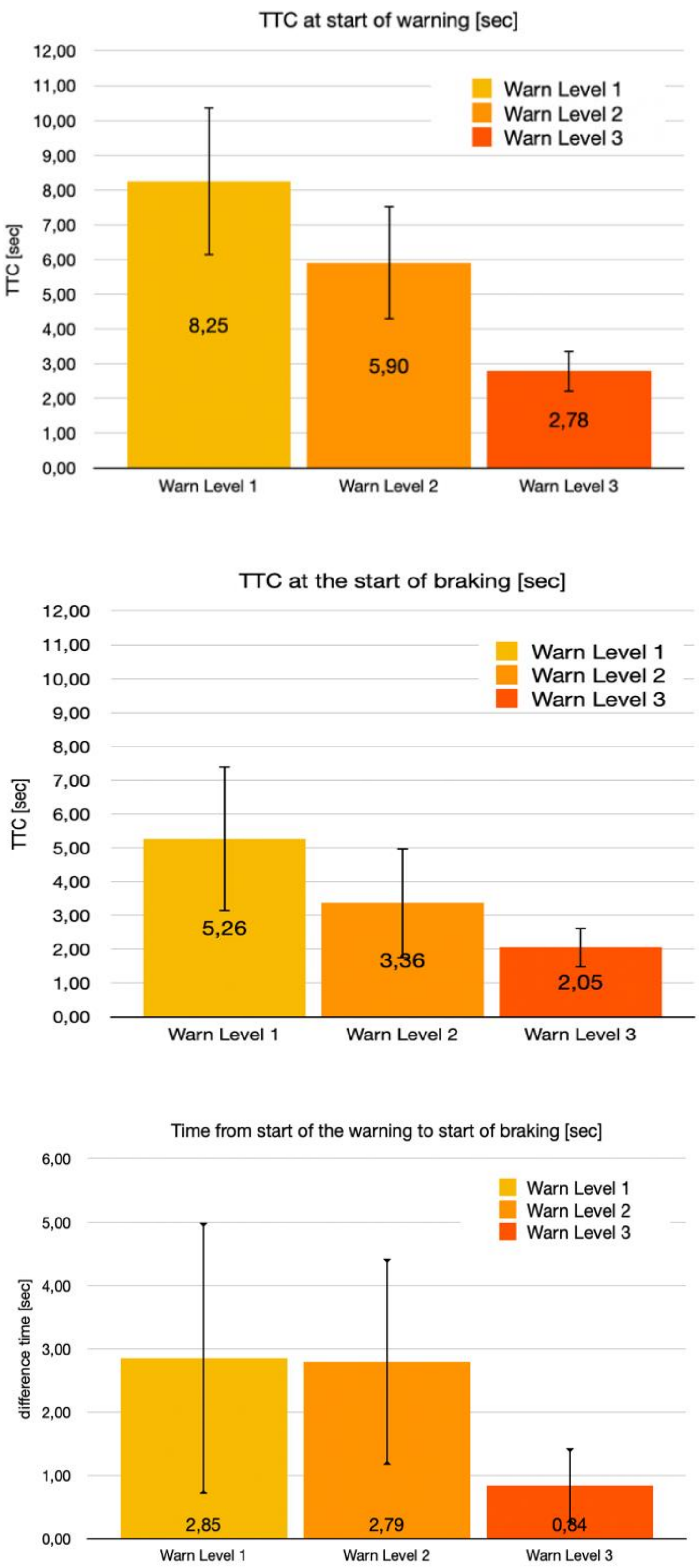


Figure 3-12: Comparison of the warning levels

Table 3-3: Correlation of Warning Levels with focus on reactions time (from warning to breaking Figure 3-12)

	p-value	t-value	df
Warn Level 1 / Warn Level 2	0,95	0,06	71
Warn Level 1 / Warn Level 3	0,0002	4,17	30
Warn Level 2 / Warn Level 3	<0,0001	6,26	97

When comparing the warning levels (Case 1), a significant difference was observed between warning levels 3 and 2 compared to warning level 1. The reaction time — defined as the time between the warning and the initiation of braking — was drastically reduced at warning level 3 compared to warning levels 1 and 2. This effect is most likely due to the more concise warning tone and the use of an icon signalling a hazard at warning level 3. These differences were intentionally incorporated into the design of the warning levels and successfully achieved the desired outcome: an immediate driver reaction to warning level 3.

3.2 Study 10 – Evaluation test of eHMI

3.2.1 Experimental setup

Study 10 is a real life study in controlled environments. Participants interacted as pedestrians with a real test vehicle, equipped with a real eHMI. The eHMI and its capabilities are described in D6.1. For data collection, Study 10 used a large-scale motion capture system provided by the AIDA project (https://aida.reutlingen-university.de), which facilitates advanced motion tracking and analysis within a motion capture volume of 55m x 30m x 5m (Figure 3-13). This system comprises 26 high-performance Vicon Valkyrie VK16 cameras and 12 VK8 cameras. The VK16 cameras feature a 16.1 MP resolution at 240 fps, with a maximum frame rate of 2000 fps and a latency of 5.6 ms. The VK8 cameras provide an 8 MP resolution at 500 fps and can also reach up to 2000 fps with a latency of 2.4 ms. The cameras are equipped with varifocal lenses offering a horizontal field of view ranging from 54° to 72°, ensuring optimal volume coverage and extended tracking distances.

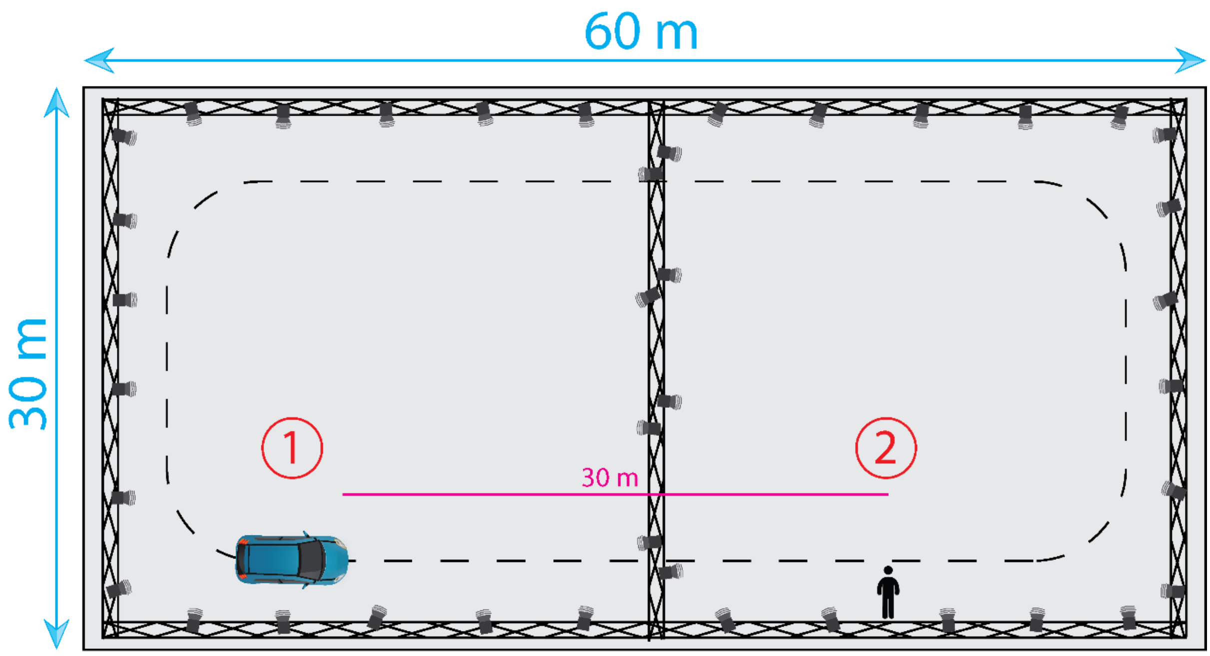


Figure 3-13: Schematic Layout of AIDA hall with MoCap System.

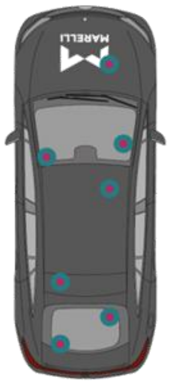


Figure 3-14: Test vehicle with IR reflectors. Schematic positions (left), vehicle inside test venue with reflectors attached (right).

3.2.2 Use Case

This evaluation study investigated use case 2, experiment 1, pedestrian intending to cross the street, as defined in D1.2 “Use case definition” with the objective to evaluate the effect of an eHMI on the interaction between vehicle and VRU. Both variants were investigated with the participants crossing the street during each interaction, thus changing between near side sidewalk (variant 1) and far side sidewalk (variant 2).

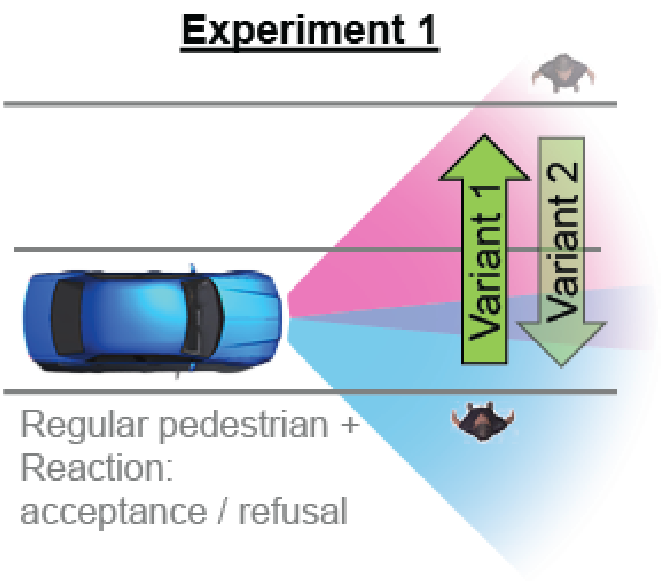


Figure 3-15: Use Case 2, experiment 1, as described in D1.2.

Specifically, 5 situations were evaluated.

1. Vehicle yielding priority to the pedestrian with active eHMI indicating vehicle intent. Message variant 1.
2. Vehicle yielding priority to the pedestrian with active eHMI indicating vehicle intent. Message variant 2.
3. Vehicle yielding priority to the pedestrian without active eHMI.
4. Vehicle proceeding without yielding with active eHMI indicating vehicle intent.
5. Vehicle proceeding without yielding without active eHMI.

All situations occurred with the pedestrian on either side of the road, so both variants of Use Case 2, Experiment 1 were investigated.

3.2.3 HMI solution

In WP6, an eHMI prototype was developed. It is fully integrated into the front of a battery electric vehicle (BEV). Using a BEV allows driving around in an enclosed space without the additional considerations of ventilation which would be necessary using an ICE test vehicle.

The eHMI consists of two displays with 768 RGB LEDs. Figure 3-16 shows the test vehicle in motion inside the test venue.

Figure 3-16: Test vehicle with integrated eHMI.

Figure 3-17 shows the scenarios of the study and the status of the eHMI. In the stop scenario, two messages were used. An animated chevron running downward, symbolizing a lowering front due to braking, indicates a decelerating vehicle. The second version shows an animated eye symbol before showing the animated chevron, with the eye indicating that the participant was "seen" and the vehicle will yield, supported by the chevron again showing deceleration.

For the drive on situation, animated chevron running upwards are shown, symbolizing a raising front due to acceleration, indicating that the test vehicle will not stop.

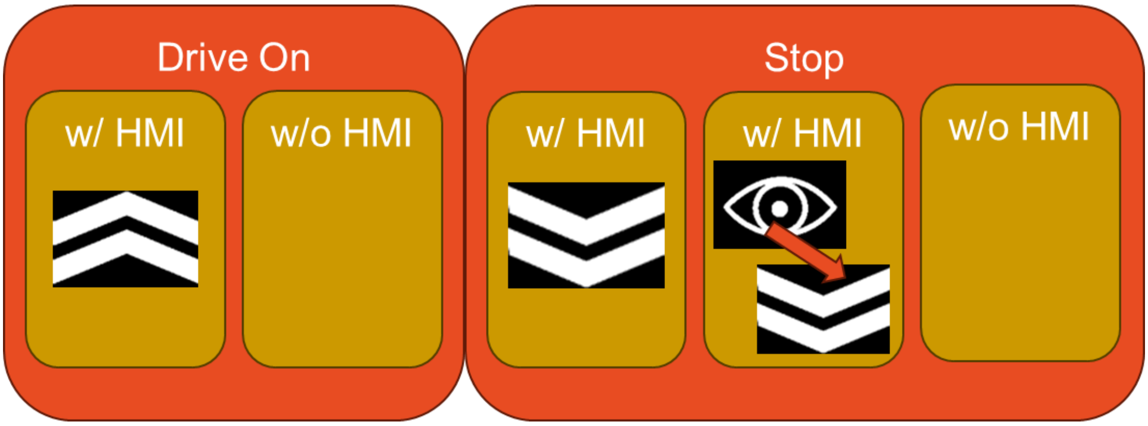


Figure 3-17: Five scenarios used in Study 10 and symbols shown on the eHMI respectively

3.2.4 Procedure

The procedure of Study 10 was carefully designed to ensure participant safety. It is the central element of the review process by the ethical review board of University Tübingen. After implementation of reviewer feedback, the ethical review board gave no reservations or concerns with conducting the study.

Official declaration was received on 21.11.2024 from „Ethik-Kommission an der Medizinischen Fakultät der Eberhard-Karls-Universität und am Universitätsklinikum Tübingen, Gartenstraße 47, 72074 Tübingen".

3.2.4.1 Experimental Procedure

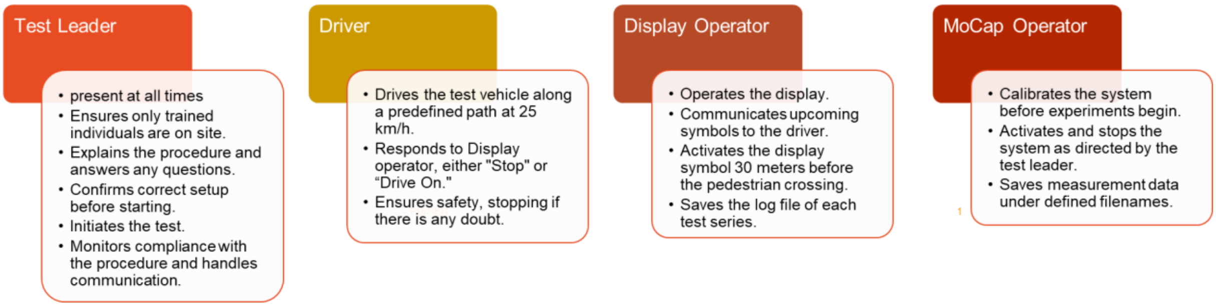


Figure 3-18: Staff roles within Study 10.

In the experimental setup, various roles are assigned to ensure a smooth and safe execution of the trials. The Test Leader is a central figure who must be present throughout the experiment. Their responsibilities include ensuring that only trained individuals are allowed in the test hall aside from the test person.

Before the start of each trial, the **Test Leader** explains the procedure to the test person and is available to answer any queries. They verify that the test vehicle is correctly positioned on the predefined path and initiate the experiment, remaining in a designated area to manage compliance with the procedural steps and communicate with the Motion-Capture System Operator when starting and stopping recordings.

The **Driver** plays a crucial role by driving the test vehicle along a predefined track at a controlled speed of 25 km/h within the hall. The driver's actions are guided by symbols displayed by the passenger, with instructions to "Stop" or "Drive On." Safety is paramount, and the driver is instructed to stop the vehicle if there is any uncertainty about the test person's actions.

Inside the vehicle, the **Display Operator**, located in the passenger seat, is responsible for operating a laptop to control the display, ensuring the correct symbol is shown at the right time. This role involves communicating upcoming symbols to the driver, activating the display symbol 30 meters before the pedestrian crossing, and saving the log file for each test series under a defined name.

The **Motion-Capture System Operator** is tasked with calibrating the system before starting the experiments. He activates the system when prompted by the Test Leader and ensure that the collected measurement data is saved under designated filenames.

The test takes place in a dedicated hall equipped with a Motion-Capture System, where the vehicle follows a specified track counterclockwise at a constant speed. This environment is carefully controlled to maintain safety and experiment integrity. Communication among the team members is facilitated via walkie-talkies to ensure seamless coordination during the trials.

Overall, the procedure emphasizes safety and compliance, with the Test Leader overseeing adherence to protocol and the Driver taking immediate action to adjust vehicle behaviour to ensure the test person's safety. All staff involved in the experiment have been instructed carefully and have confirmed their awareness of the respective guidelines.

3.2.4.2 Participants Procedures

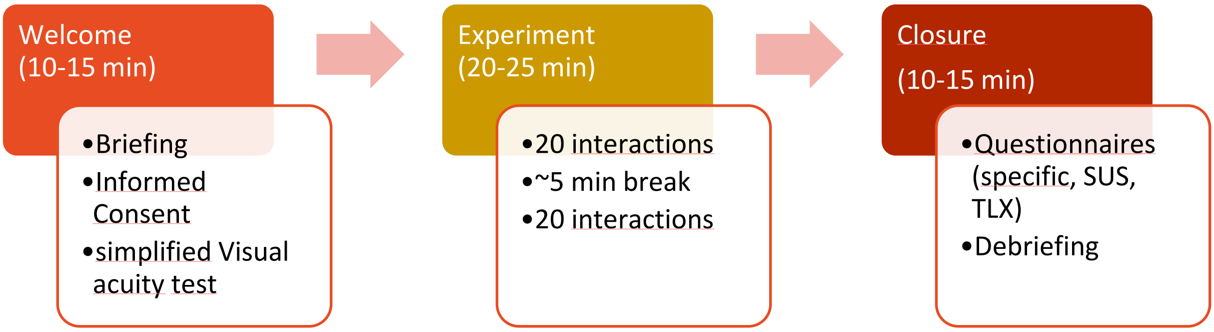


Figure 3-19: Procedure for each participant.

Upon arrival at the test facility, participants are welcomed by the Test Leader. They are informed about the content and procedure of the study, their tasks, the voluntary nature of their participation, and especially their rights, including the option to abort their participation at any time. After providing written consent, visual acuity is tested using a visual acuity chart to confirm subjects can recognize the optical signals. All visual aids that participants use while participating in traffic are allowed.

Next, subjects are taken to the simulated crossing situation where the actual experiments take place and are equipped with a helmet with IR reflectors for position tracking. Experiment 1, as shown in Figure 3-15 serves as an exemplary representation of a trial procedure: The test vehicle drives at 25 km/h on a circuit through the hall. At Key Point 1 (Figure 3-13), the speed should be 25 km/h, and the distance to Key Point 2 is 30 meters. At this point one interaction begins. Participants are instructed to cross the road during each interaction. The decision when to do it, is entirely their own, depending on their evaluation of vehicle speed and intention, i.e. the vehicle might slow down to yield or remain at speed. The vehicle intention is randomized, and the use of the eHMI to communicate vehicle intention is randomized as well. The interaction is complete, when the participants have crossed the road and the vehicle has fully passed, i.e. is along the curve.

The Test Leader remains close by but does not interact in the crossing. After crossing, the participants remain on that side of the road, so the experiment alternates between variant 1 and variant 2. After 20 interactions, there is a short pause. During this pause, successful data collection by the MoCap system is verified and the Test Leader checks on the participant's well-being as well as an update on the other experimenters, especially the test driver. If all parties involved give a positive feedback, the Test Leader initiates the second set of 20 interactions.

After completing the second set of 20 interactions, participants take off the reflector helmet and return to the welcome room. There, they are asked to fill out three questionnaires, one specific to the eHMI used in the study, the NASA TLX and the System Usability Score questionnaire [3], [4]. All questionnaires are attached in the Appendix. Then two open questions are asked and the participants are debriefed and thanked. This concludes the study for the participants and they are dismissed.

3.2.5 Participants

In total, 27 participants were recruited in two age groups. In order to allow conclusion about the effects of an eHMI with an elderly population, one age group included subjects below 50 years of age (39.4 ± 5.2 years, 4 female, 15 male) and the other group included subjects above 70 years (78.4 ± 6.2 years, 3 female, 5 male).

Age group 1 (younger than 50) had 19 participants.

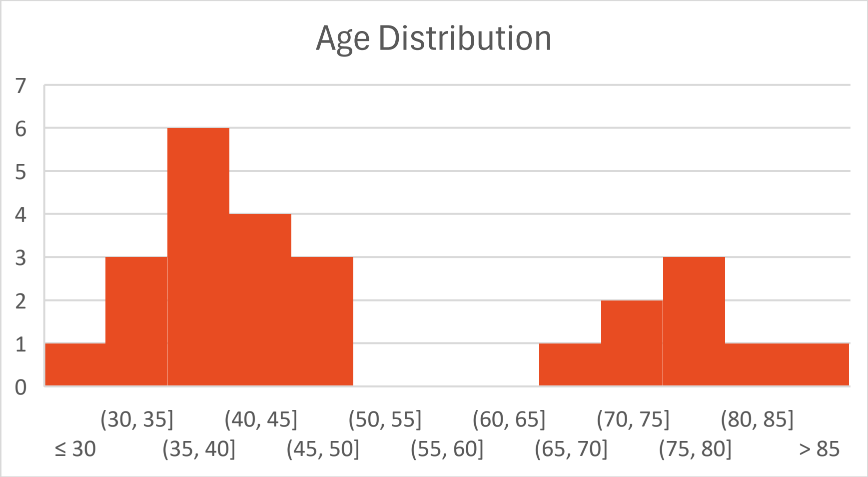


Figure 3-20: Age distribution of Study 10 participants.

3.2.6 Data collection

During the simulated traffic situation, the positions of both the vehicle and the participants are recorded using the MoCap system.

Since it solely records spatial positions without any information about the participants' body shapes, tracing back to individual participants is not possible. The only data linkage present is the assignment to the age groups.

Based on this data, several parameters are calculated: "Time to Resolve" (TTR), which is the time span between the start of the trial (when the test car passes Key Point 1 at 25 km/h and is visible to the participant) and the resolution of the situation (vehicle behind the participant's crossing point) , "Crossing Speed" (CS), the speed of the participant's movement, and covered distance while crossing. Further parameters which are inherent in the collected data are "Crossing Time" (CT), the time the participant takes to cross the road, "Reaction Time" (RT), the time between the display of a communication signal and a participant's reaction, such as deciding to start or stop moving, and "Time to Collision" (TTC), the time until the vehicle would reach the participant. The latter, however, bring no additional benefit in validating the effectiveness and are therefore not presented in 3.2.7. The former data, combined with feedback from the questionnaires, form the basis for evaluating the effectiveness of communication between vehicles and pedestrians.

Figure 3-21 shows the measurement points used in analyzing the 3D data collected by the MoCap System.

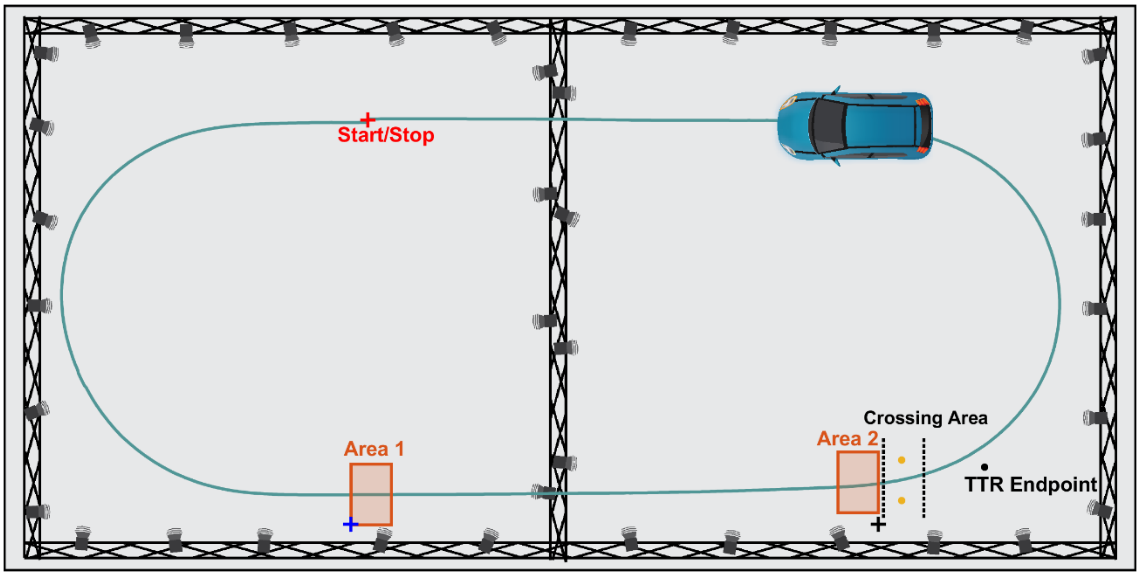


Figure 3-21: Measurement points and areas.

Table 3-4 shows a summary of the external coordinates in the Hall and their use for calculations.

Table 3-4: Measurements and their corresponding coordinates

Coordinates	Related measure-ment	Description
Start/Stop, red cross		Vehicle passing red cross marks separation of measurement data for new interaction, i.e. new round.
Blue Cross	TTR	Start of Interaction, point when vehicle is at speed along the straight and eHMI is switched on if in use.
Black Cross	CS, CT	IR marker indicating position of crossing area.
Crossing Area	CS, CT	Area, where participants are expected to cross the street.
TTR Endpoint	TTR	Point when vehicle has completely passed the crossing area and the situation is considered resolved.
Area 1	Vehicle speed	Area to measure vehicle speed at start of interaction
Area 2	Vehicle speed	Area to measure vehicle speed directly in front of the crossing area

3.2.7 Results

In the following chapter we present several aspects of the measured data and the high number of interactions (>1050) facilitated statistically significant results. Given error bars in graphs indicate standard deviations calculated for each quantity.

3.2.7.1 Effectiveness

The main goal of Study 10 is to investigate the effectiveness of an eHMI, i.e. to prove a measurable and positive difference between pedestrian behaviour with and without an eHMI. Figure 3-22 shows the time to resolve for all interactions.

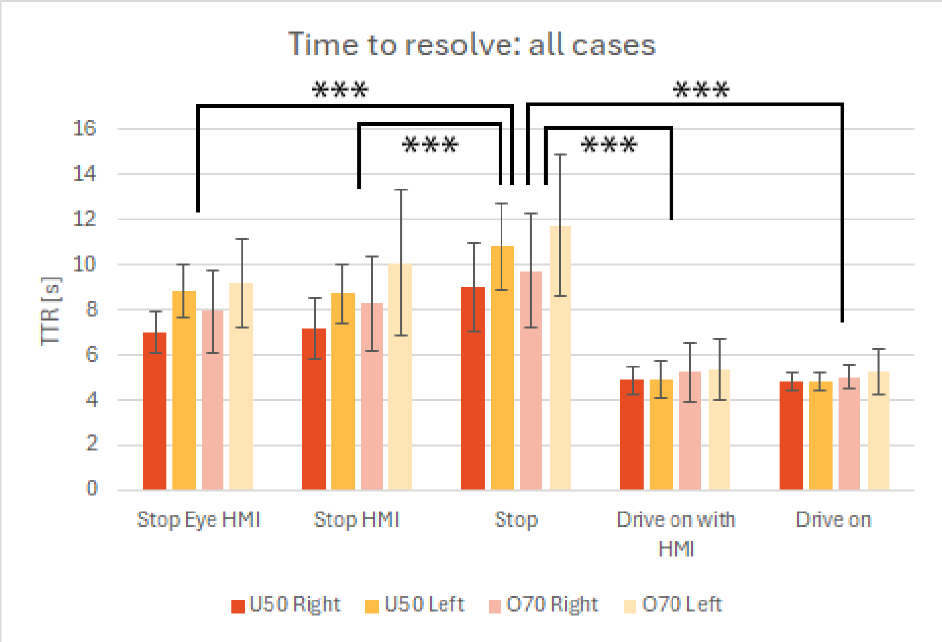


Figure 3-22: TTR for all interactions.

We have clear results. For the vehicle driving on, there is no significant difference between the scenarios with and without HMI, no significant difference between pedestrian being on left side or right side and an insignificant difference between age groups. TTR of the “drive on” situation can therefore be used as baseline to compare the effect of HMI on the “stop” situation.

Comparing the stop situation, i.e. the vehicle intends to let the pedestrian pass in front of it, we get three significant results.

First and foremost, with HMI active, the TTR is reduced statistically significant for both age groups, for both sides and for both symbol variants. On average over all interactions, TTR with HMI is reduced by 1.9 seconds. Using the Drive on situation as benchmark, the additional time spent to allow the pedestrian to cross in front is reduced by 40%.

This holds true with two secondary effects. Overall, TTR is higher with participants on the left side, which is expected for right hand traffic, and age group 70 years and older has a higher TTR which is expected due to an on average lower mobility.

Analyzing the effectiveness over the number of participants, Figure 3-23 shows a histogram over all participants showing the time difference in TTR of the stop situation for each individual.

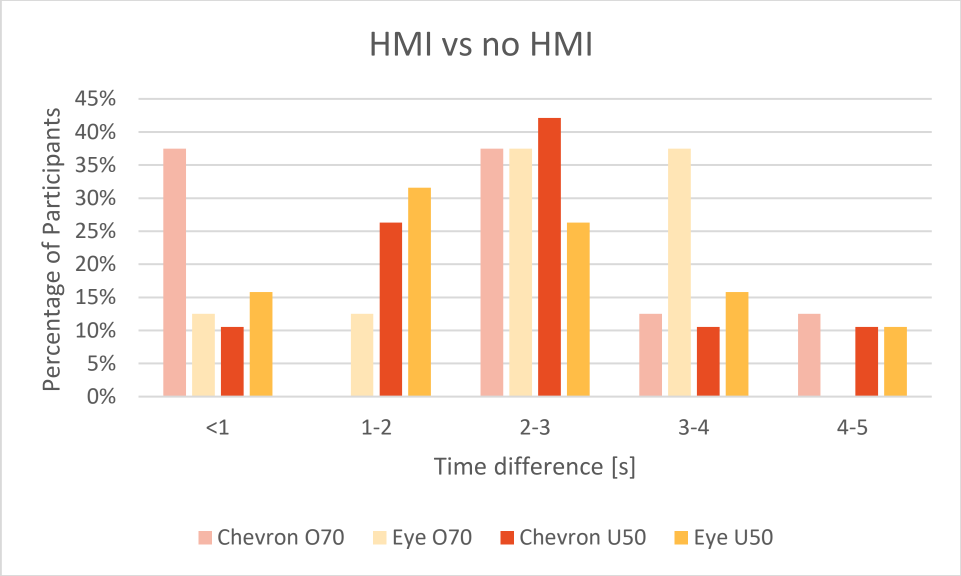


Figure 3-23: Histogram of the effect of eHMI on TTR.

When the eye symbol was used, 84 % of participants under 50 years old and 87 % of participants over 70 years have an average TTR reduced by over one second. That means, the vast majority of participants have a significant effect, for more than half of participants it is well over two seconds.

This effect is illustrated in the side-by-side comparison in Figure 3-24. The top row shows video frames of a “HMI off” situation, whereas the bottom row depicts the “HMI active” situation. At frame 2 the test car is in Area 1 approaching with 25km/h. Already at frame 3 the difference starts manifesting, because the pedestrian started crossing the street in the bottom row. With the HMI being off, the pedestrian waits until frame 4 and the vehicle slowing down and being closer to start crossing the street. At frame 5 the pedestrian has already crossed with HMI active, while without HMI an additional frame and almost three additional seconds are necessary.

Figure 3-24: Side by side comparison between Interactions w/ and w/o eHMI active. On resolution of the situation w/ eHMI the timer starts. W/o eHMI the situation takes three seconds longer to resolve.

3.2.7.2 Acceptance and Learning Curve

All participants were informed beforehand about the existence of the HMI and about the geometric shapes and patterns used as messages. However, they did not encounter the actual HMI in action before the testing. Each participant undertook 40 interactions with the 5 scenarios in random order. It was observed that the HMI improved the understanding of the vehicle intent very quickly. Thus, we observe in the data a learning curve, where the TTR of the "Stop"-Scenario decreases over the course of the experiment. Figure 3-25 shows the TTR of "Stop with HMI". Already after the first 10 interactions, the TTR decreased statistically significantly. This effect continues for the following 30 interactions, however not in a significant manner.

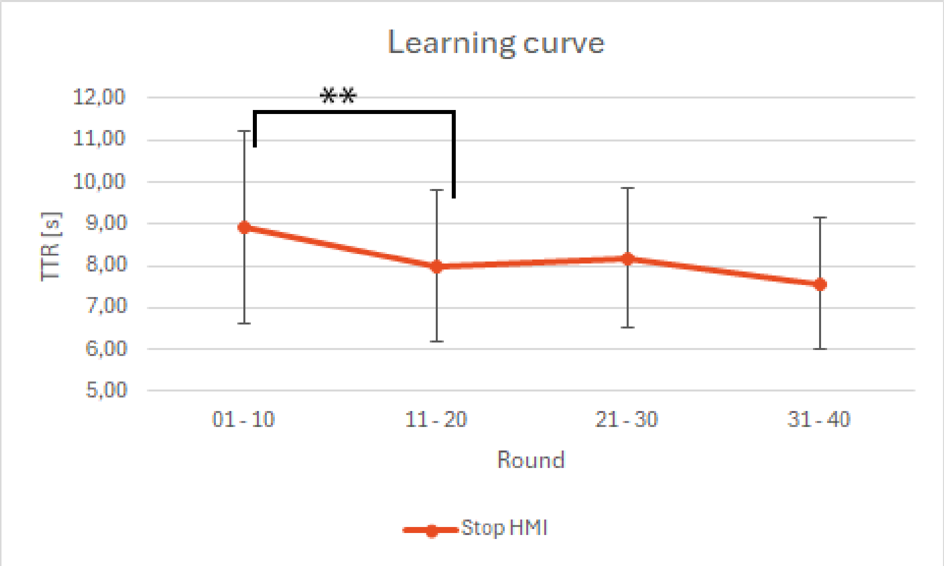


Figure 3-25: Evolution of TTR for Stop Scenario w/ eHMI over the course of the experiment for each participant averaged over all participants.

This allows us to presume, that following an information campaign, the HMI can be quickly accepted by the general road user population.

3.2.7.3 Distance & Traffic Flow

It is very challenging to correctly interpret safety of a participant from measurements of distance and speed. However, it is a clear result from our study, that due to an earlier understanding of the vehicle's intent, the pedestrians on average initiate their crossing earlier. While crossing, the distance between car and pedestrian is statistically significantly larger, see Figure 3-26, meaning there is more room for the vehicle to manoeuvre in case of an unforeseen event, e.g. pedestrian stumbling or similar. It bears noting that the HMI is only active when the vehicle has detected the pedestrian and is therefore monitoring the pedestrians' behaviour.

Figure 3-26: Distance between VRU and vehicle while crossing the street in front of the vehicle.

This increased distance allows the vehicle to retain its speed which results in an overall smoother and improved traffic flow. Figure 3-27 shows the average vehicle speed at areas A1 and A2, see Figure 3-21, at the beginning of the interaction and just in front of the crossing area. When the HMI is active, with little difference between the two HMI versions, the vehicle is statistically significantly faster than without HMI. This reduces the cost of collaboration, i.e. yielding, for the vehicle.

Figure 3-27: Average speed of vehicle at start of interaction (A1) and at the crossing area (A2).

3.2.7.4 Ease of Understanding and Usability

After the experimental part was finished, all participants were asked to complete several questionnaires (see also Annex). Figure 3-28 shows the results from the eHMI specific questions. They were positive throughout. Specifically, the technical HMI parameters were rated to be appropriate, the time and distance of activation was rated sufficient and therefore understandability was rated very positive. This increased the impression of safety while being deemed of low risk when used in traffic.

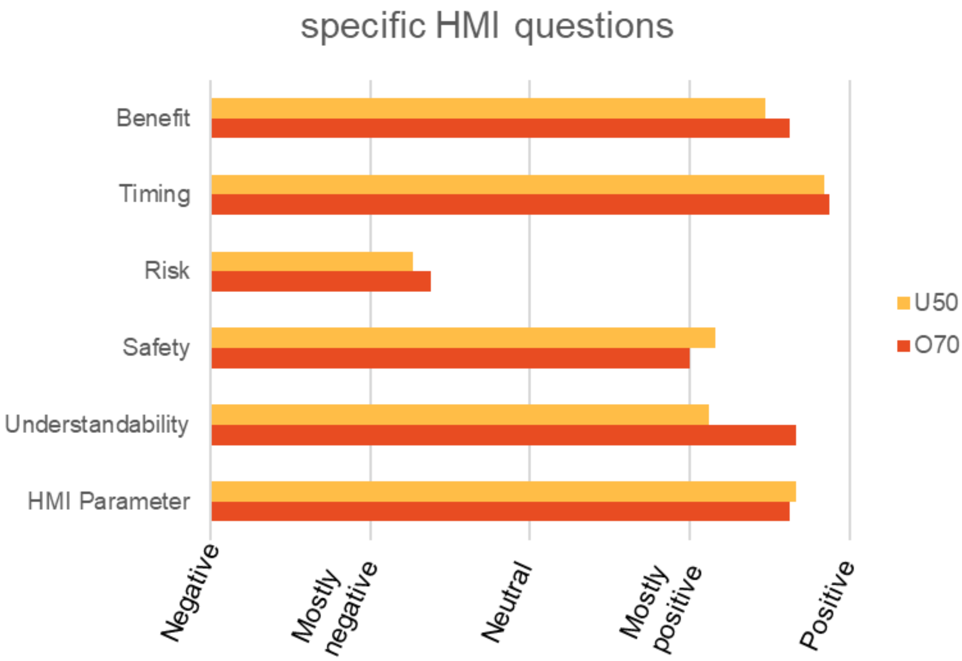


Figure 3-28: Results from questionnaires specific to eHMI solution.

Furthermore, the ease of understanding is reflected in an outstanding System Usability Score (SUS), see also Annex, where the system was rated 86.4, averaged over all participants. A score of 68 is deemed an average system, and values >80 are exceptional [4].

This result reflects well on the experiences made in previous work packages which influenced the prototype eHMI as well as the selection of patterns. Furthermore, it reflects on the experimental setup feeling natural to the participants, and them being able to use their everyday skill of navigating traffic. This is in contrast to the simulator studies, where the element of “unrealness” certainly had an effect on participants.

The naturalness of the situation and the ease of understanding the HMI is also reflected in the results from the Task Load Index (TLX) questionnaire. Shown in Figure 3-29, the participants experienced an overall little load of demand when interacting with the HMI, they were very content with their performance and experienced little frustration. This again reflects well on a future introduction of such an eHMI to public traffic situations.

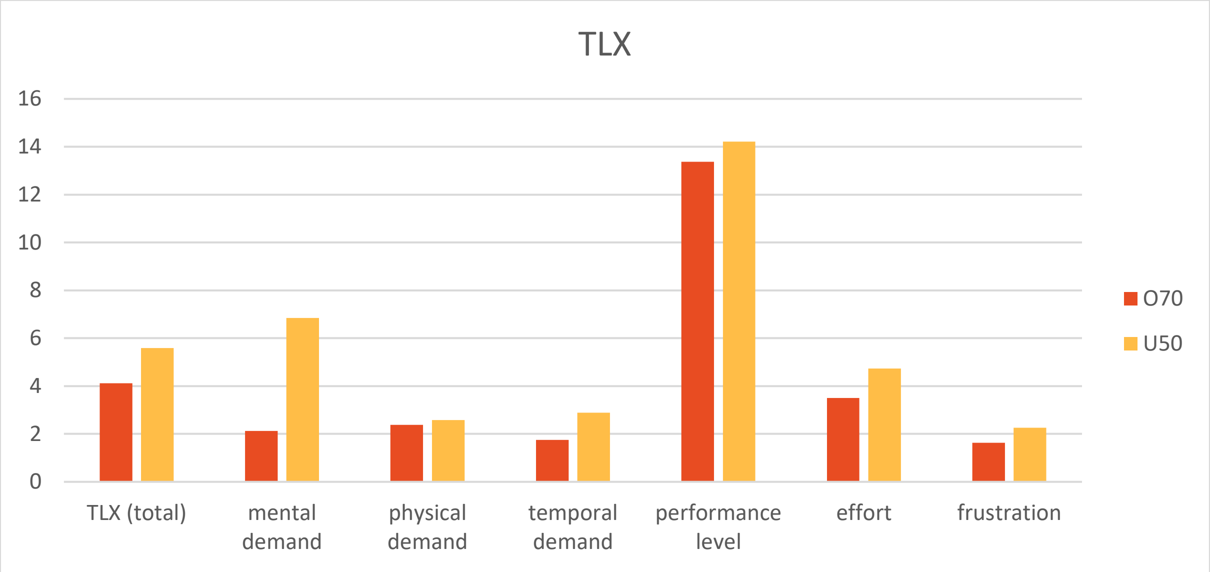


Figure 3-29: Result from TLX questionnaire.

3.3 Study 12 – Evaluation test of cooperative HEIDI HMI

3.3.1 Experimental setup

The study was conducted with one driver and one pedestrian simultaneously to enable interactions between them in the simulated world. All interactions were centred around an un-signalized, unmarked, pedestrian crossing. The study had a mixed within-between design with participant type (driver or pedestrian) as the between-subjects factor and cHMI condition (cHMI activated or not activated) as the within-subjects factor. In addition, distraction tasks for drivers and pedestrians were used in a subset of the pedestrian-driver interactions.

Two test blocks were performed, one with cHMI and one without cHMI. Each test block had 8 interaction scenarios. The study had a cross-over design where the order of test blocks with and without cHMI was counterbalanced between participants. The driver was engaged in a distraction task on a tablet in 25% of the interactions and the pedestrian had a distracting auditory task in 25% of the interactions. The two participants were able to either adhere or not adhere to the communicated cHMI message. The evaluation was done using objective data from the simulators and cHMI system and subjective data from questionnaires that participants answered before, in between, and after their test, and a debriefing interview at the end of the trial. The roles of driver and pedestrian were not changed between test blocks meaning that one participant acted as the driver in both test blocks and the other as pedestrian in both test blocks.

3.3.2 Use Case

A subset of the situations outlined for UC3 in D1.2 were tested. UC3 focuses on the interaction between two interaction partners, an ego vehicle driver and a pedestrian, that are both human agents. The experiments were designed to evaluate if the cHMI can facilitate cooperation in urban traffic. The evaluated experiments were:

1. Interaction between a regular driver and a regular pedestrian at an unsignalized crossing
2. Interaction between a distracted driver and a regular pedestrian at an unsignalized crossing

3. Interaction between a regular driver and a distracted pedestrian at an unsignalized crossing

Adherence or non-adherence to the communicated HMI message was not manipulated in the study design. Instead, adherence was regarded as an outcome of the interaction.

3.3.3 Co-simulation setup

Distributed simulation tests were performed using a networked virtual reality (VR) pedestrian simulator and a fixed-base driving simulator with augmented reality (AR) features. Detailed description of the co-simulation setup can be found in D6.2 “cHMI prototype”. The pedestrian was equipped with a Varjo VR3 headset and three HTC VIVE 3.0 trackers (two at the ankle joints, one around the waist). The position and orientation of the pedestrian was tracked using the headset together with the trackers and represented as an animated avatar. The driver was equipped with a Varjo XR-3 headset, which allowed the driver to see their own hands, the steering wheel, and the dashboard while the rest of the environment was simulated. The position and orientation of the car was tracked and visualized to the pedestrian. Hence, this simulation setup allowed realistic interactions between human drivers and pedestrians in a simulated VR environment.

The virtual environment had eight 700 m segments in an urban environment Figure 3-30. The crossing scenario was on a single lane road section in the last 170 m of each segment. The pedestrian area was 4 x 6.5 m with flat floor and no curbs. The pedestrians were free to walk around as they wished when the car was in other sections. A red grid appeared when the participant was close to the edges of the interaction area.

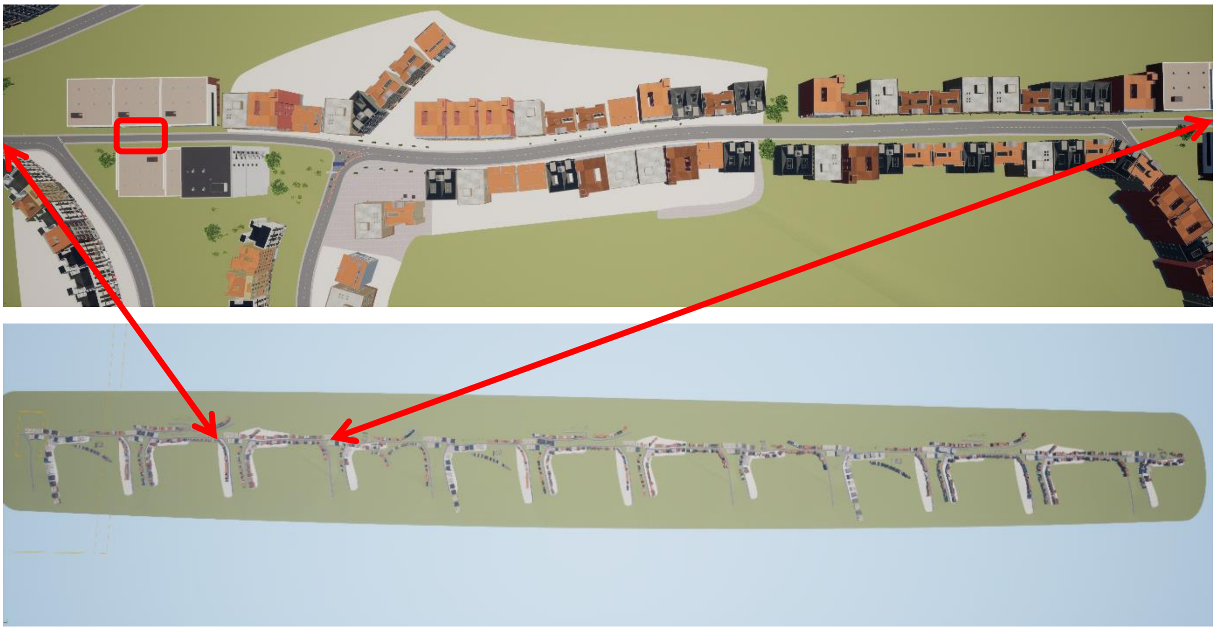


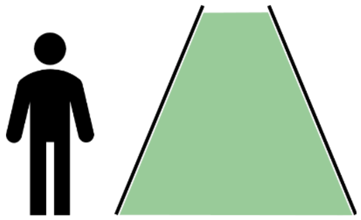
Figure 3-30: Map of the entire simulated environment (bottom) and an enlarged section (top) including one of the interaction areas. The red square indicates where the interaction took place.

3.3.4 HMI solution

The cooperative HMI system has three main components, the cooperative behavior planning system, the cooperative HMI logic, and internal and external message displays. Detailed descriptions of the cHMI solution can be found in deliverables D4.2 “Cooperative HMI Concept” and D4.3 “Cooperative HMI system”.

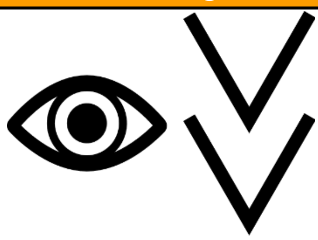
The cooperative behavior planning system uses state information from the ego vehicle, other traffic participants, and map information to calculate the cost of different potential combinations of vehicle and pedestrian trajectories, and outputs information about the optimal behavior to the cooperative HMI logic. The logic determines which internal (iHMI) and external (eHMI) messages should be shown based on the driver's and pedestrian's behavior. The iHMI logic compares driver behavior with the recommended behavior. If the driver behaves according to the recommended behavior, no iHMI message is shown. If the driver does not behave in accordance with the recommendation, different messages can be shown in a head-up display, depending on the situation. The iHMI symbols tested in study 12 were designed to communicate 1) that "vehicle first" is recommended via a green road and standing pedestrian and 2) that "yielding" is recommended using a symbol with an amber colored road and a walking pedestrian (Table 3-5).

Table 3-5: iHMI symbols

Vehicle first	Yielding
	

The eHMI messages are activated only if the driver's behavior is in accordance with the recommendation. Otherwise, an external message is considered unsafe, and no message is shown. The eHMI symbols are shown on displays on the front of the car (right and left). Three different eHMI symbols are be activated via the eHMI logic, 1) animated upward pointing arrows if the driver behavior is "vehicle first", 2) animated down pointing arrows and eye symbol if the driver behavior is "yielding", and 3) eye only if the vehicle speed is low and the driver behavior is "yielding" (Table 3-6).

Table 3-6: eHMI symbols

Vehicle first	Yielding	Yielding (low speed)
		

3.3.5 Procedures

Figure 3-31 presents an overview of the test procedures. The participants completed a background questionnaire, and the pedestrians underwent a short cognitive screening (MoCA) before the simulator test. Thereafter, the drivers were seated in the simulator's car seat and instructed to adjust the seat and put on seat belt before putting on the VR headset. The pedestrians put on sensors around the waist and ankles for motion capture and wore a VR headset. Both agents' headsets were calibrated.

The tests started with a short training session, where the driver and pedestrian got acquainted with the virtual reality milieu. Participants were informed that the aim of the study was to evaluate a new communication system in cars. Before the test block with cHMI, the participants were given short instructions about the HMI concept and what the symbols meant. The drivers were instructed to drive as usual. The pedestrians were informed they were going to be standing at the side of the road, and when the rain started pouring down it was their cue to cross the street. The drivers were free to decide whether they wanted to yield to the pedestrian and the pedestrians could decide for themselves if they wanted to cross the road before or after the car had passed. After crossing the road, the pedestrian answered two questions emerging in the VR-world. When the pedestrian was ready for the next crossing, the test leader teleported the pedestrian and the driver approached the new crossing point. The driver continued to drive straight after each crossing scenario.

Distraction was manipulated using secondary tasks. The driver performed visual search task on a tablet in the center panel of the driving simulator. The pedestrians performed an auditory task presented to them in their headphones.

After each test block, the participants removed the headsets and answered questionnaires about their experiences during the test. Lastly, a brief interview was performed with the driver and pedestrian together.

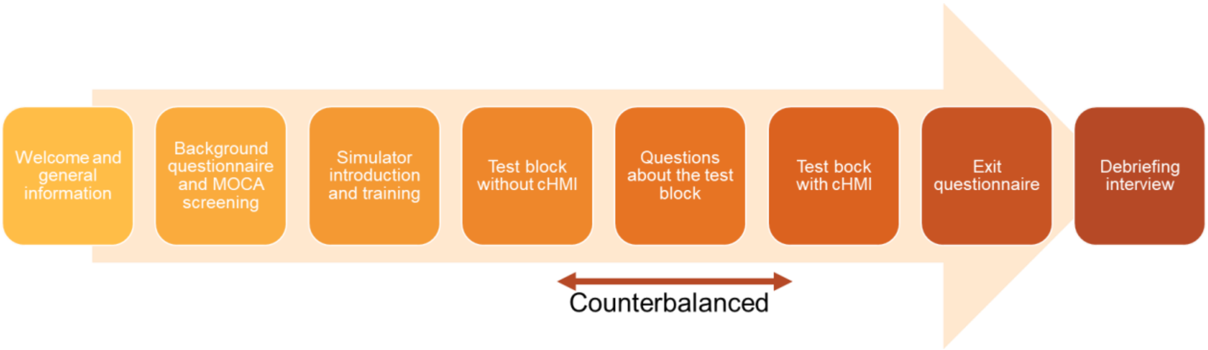


Figure 3-31: Test procedure

3.3.6 Participants

Seventy participants, 35 pedestrians and 35 drivers, took part in the experiment. The drivers were aged 25-55 years old, the pedestrians 25 to 55 years or 70 years and above (Table 3-7). The participants gave their written informed consent to participate in the study before the experiment started. The study was approved by the Swedish Ethical Review Authority (Dnr 2024-00861-01).

Table 3-7: Participant characteristics

	Drivers	Pedestrians
N	35	35
Gender, n		
Male	23	24
Female	12	11
Age, mean (SD)	40 (8)	55 (20)
Age group, n		
25-55 years	35	18
≥70 years	0	17

Years of education, mean (SD)	15 (2)	15 (3)

3.3.7 Data collection

Raw data including vehicle and pedestrian position, speed, etc. was logged from the simulator runs. Outcome measures were calculated as defined in (Table 3-8).

Table 3-8: Outcome measures from the simulators and their definitions

Measure	Unit	Definition
Start Speed	km/h	The speed of the vehicle when entering the crossing scenario (65 m from the position of the pedestrian)
Time to Brake	s	The time difference between when vehicle enters the crossing scenario and when the driver presses the brake with a force of at least 1N
Distance at Brake	m	The distance between the pedestrian and the vehicle when the driver presses the brake with a force of at least 1N
Standard Deviation of Lateral Position (SDlateral)	m	The standard deviation using an unbiased estimator of the lateral position of the vehicle measured from the left front tire to the left lane marker
Maximum vehicle deceleration	m/s2	The maximum deceleration of the vehicle in the crossing scenario
Crossing initiation delay (CID)	s	Time difference between vehicle entering the crossing scenario until Crossing initiation timestamp above
Crossing time	s	Time it takes for a pedestrian to walk from one side of the road to the other side in a crossing scenario.
Crossing speed	m/s	Average pedestrian crossing speed
Total pedestrian crossing time (TCT)	s	Time from the start of the crossing scenario until the pedestrian has reached the other side of the street (CID + crossing time).
Time to resolve (TTR)	s	Time from the start of the crossing scenario until the car has passed the pedestrian by 30 m
Crossing behavior	Categorical	Crossed in front of the vehicle (1) or crossed after the vehicle had passed (2)

The background questionnaire for drivers included the Driver Behavior Questionnaire (Mini DBQ) and for pedestrians it included the Pedestrian Behavior Scale (PBS), the Affinity for Technology Interaction (ATI) scale, and the 5-level EQ-5D version (EQ-5D-5L). After each test block, all participants completed the NASA Task Load Index (NASA-TLX) and one-item ratings of perceived stress (VSS) and sleepiness (KSS) both on a scale from 1 to 9. After the cHMI test block, both drivers and pedestrians answered questions about the cHMI system, including questions about which symbols they had seen, their interpretation of the symbols, and perceived safety of the cHMI system. The questionnaire included the User Experience Questionnaire (UEQ-Short), the SHAPE Automation Trust Index (SATI), and selected items from the Car Technology Acceptance Model (CTAM). Drivers also completed the System Usability Scale (SUS). After each interaction, the pedestrians were asked how safe they felt when crossing the road from a scale from 0 to 5, where 0 represented very unsafe and 5 very safe and how clearly they perceived who had priority in the crossing, where 0 represented very unclear and 5 very clear.

3.3.8 Results

The final dataset had 259 crossing scenarios without cHMI and 266 with cHMI.

There were significantly more crossings in front of the car in interactions with cHMI (77%) compared to interactions without cHMI (46%), which was regarded as a control condition (X^2=52.57, p<0.001). Moreover, there was a larger proportion of pedestrians crossing in front of the vehicle when the pedestrian distraction task was active (63%) as well as when the driver distraction task was active (71%), compared to interactions with no distraction (57%). This difference in behavior was statistically significant (X^2=7.46, p=0.024). Results from ANOVAs of all simulator measures with condition and distraction as fixed factors and test pair as a random factor are presented in Table 3-9. Crossing initiation delay (CID), crossing time, and total crossing time (TCT) were significantly shorter, whereas time to resolve (TTR) was significantly longer in interactions with cHMI. Max deceleration was significantly lower (indicating harder braking) when the cHMI was active. Distraction had a significant effect on start speed, SDlateral, CID, and TCT. The effect of test pair was significant in all tests, with p<0.001 for all parameters (data not shown). The longer TTR but shorter TCT indicates that the traffic efficiency decreased for drivers but increased for pedestrians when cHMI was active.

Table 3-9: Results from ANOVAs testing the main effects of condition (with or without cHMI) and distraction (no distraction, pedestrian distraction, or driver distraction), and condition*distraction interaction effects.

	Main effect of condition		**Main effect of distraction**		**Interaction effect condition*distraction**	
	F	p-value*	F	p-value*	F	p-value*
Start speed (km/h)	1.09	0.296	11.47	<0.001	0.56	0.573
Time to brake (s)	5.76	0.017	0.10	0.903	0.87	0.422
Distance at brake (m)	4.85	0.029	0.07	0.931	0.15	0.859
Max deceleration (m/s2)	52.21	<0.001	1.14	0.321	0.11	0.899
SDlateral (m)	6.24	0.013	25.41	<0.001	0.95	0.387
CID (s)	19.30	<0.001	6.04	0.003	2.40	0.092
Crossing time (s)	10.99	<0.001	0.24	0.787	0.89	0.411
Average crossing speed (m/s)	6.91	0.009	4.24	0.015	0.24	0.790
TCT (s)	23.44	<0.001	5.59	0.004	2.40	0.091
TTR (s)	33.42	<0.001	2.81	0.061	0.64	0.530

*Significance level α=0.005 after Bonferroni correction

Including crossing behavior as a covariate in the analyses changed the results regarding the main effect of HMI condition from significant to non-significant for the variables CID (p=0.865), crossing time (p=0.365), TCT (p=0.694), and TTR (p=0.813), but not for max deceleration (p=0.003). This indicates that yielding/crossing behavior had a mediating role. In interactions where the pedestrian crossed in front of the car, the cHMI did not affect pedestrian behavior or traffic efficiency (see Figure 3-32 for results regarding TTR).

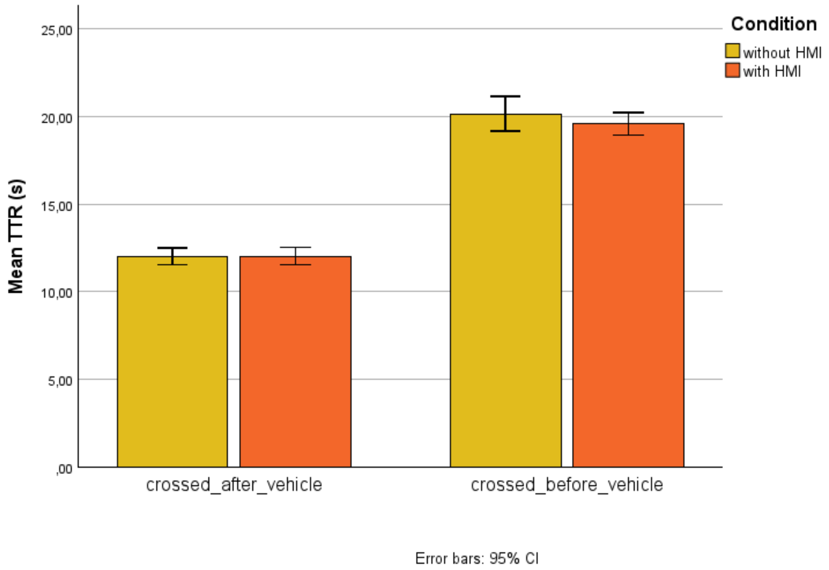


Figure 3-32: Mean TTR for interactions where the pedestrian crossed in front of the vehicle and after the vehicle separately.

Trust was evaluated using SATI and the mean score for pedestrians was 4.04 (SD 1.27) and for drivers 4.54 (SD0.93). SATI ranges from 0 to 5 and the mean scores indicated that both drivers and pedestrians had a high trust in the cHMI system [5]. User experience was evaluated with the short UEQ and mean overall score was 1.19 (SD 1.27) for pedestrians and 1.62 (SD 0.99) for drivers. The results from the short UEQ showed that the cHMI system was experienced as positive, with similar ratings for hedonic quality and pragmatic quality (Figure 3-33). Values between -0.8 and 0.8 represent a neutral evaluation of the corresponding scale, values > 0,8 represent a positive evaluation and values < -0,8 represent a negative evaluation [6]. Independent samples t-tests showed no significant differences between pedestrians and drivers in SATI or UEQ scores.

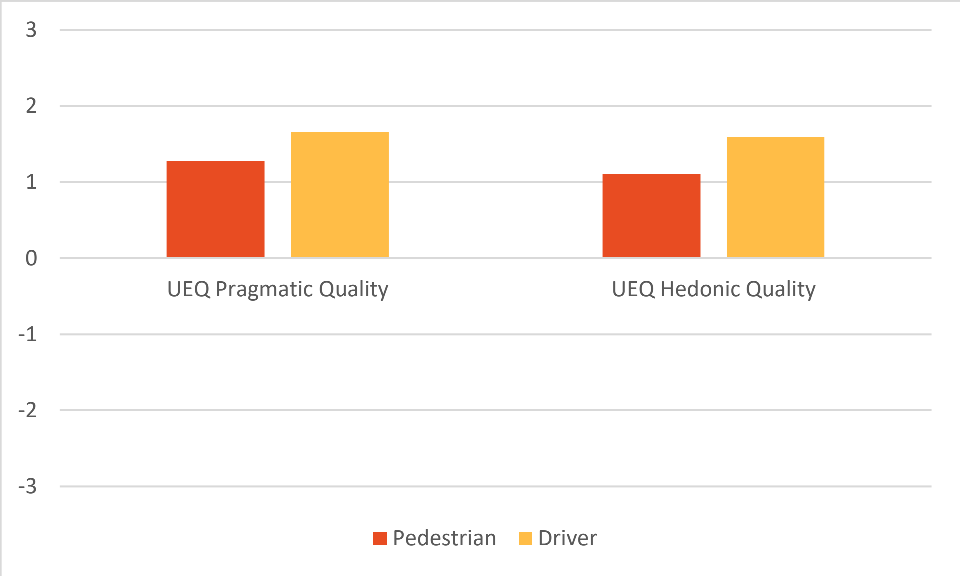


Figure 3-33: Short UEQ scores (range -3 to 3) for pedestrians and drivers.

The drivers also rated usability with the SUS. Mean score was 82.42 (SD 15.05) which is well above 68, which is a commonly used cut-off for good usability [4]. Trust and usability were improved compared with the results from the small-scale validation study of the first version of the cHMI reported in D5.3 (study 7).

The participants were asked to rate different statements about the cHMI, including items from CTAM as well as self-constructed items, on a scale from 1=do not agree at all to 6=totally agree [7]. Figure 3-34 provides an overview of the average ratings for each statement. In general, the cHMI was rated as relatively safe and understandable, indicated by mean scores around 4 for items about ease of understanding, safe behavior, and decreased accident risk and mean scores around 2 for items about accident risk, danger, and distraction. Mann-Whitney U-tests revealed no significant differences between pedestrians and drivers in their opinion about the cHMI. These results were also improved compared with study 7 (see D5.3 for results).

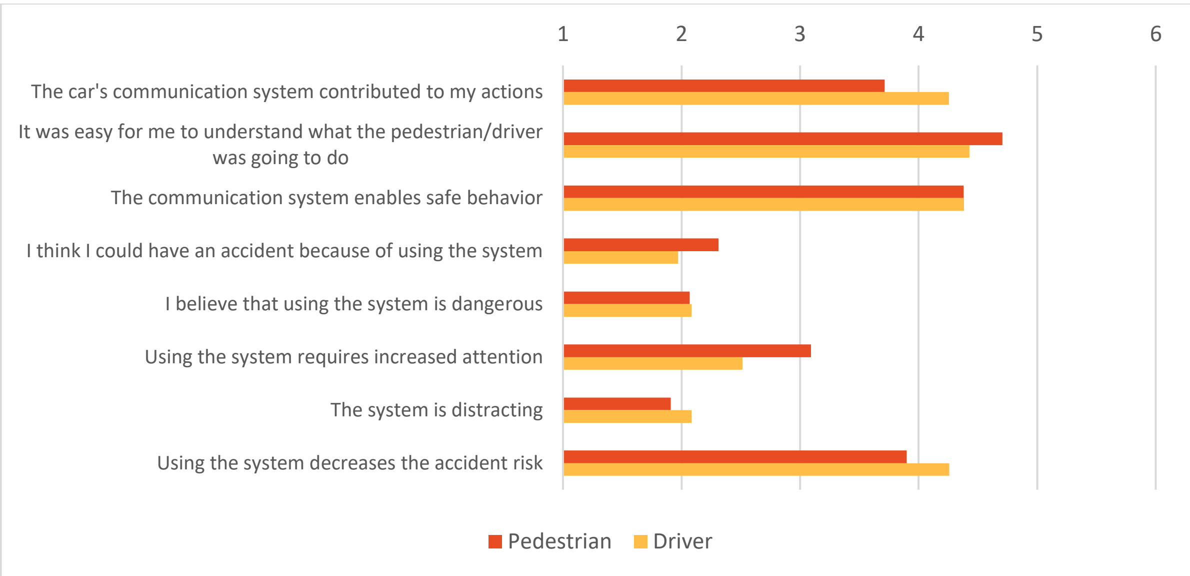


Figure 3-34: Average ratings of different statements about the cHMI on a scale from 1=do not agree at all to 6=totally agree.

Differences in workload (NASA TLX), stress and sleepiness scores were compared between cHMI conditions with paired samples t-tests. Performance scores were significantly lower ($t(69)=2.2$, $p=0.016$) and frustration scores were significantly higher ($t(69)=-2.7$, $p=0.013$) after test blocks with cHMI compared to test blocks without cHMI. The results were not statistically significant after correction for multiple tests. Workload scores were relatively low in both conditions (Figure 3-35).

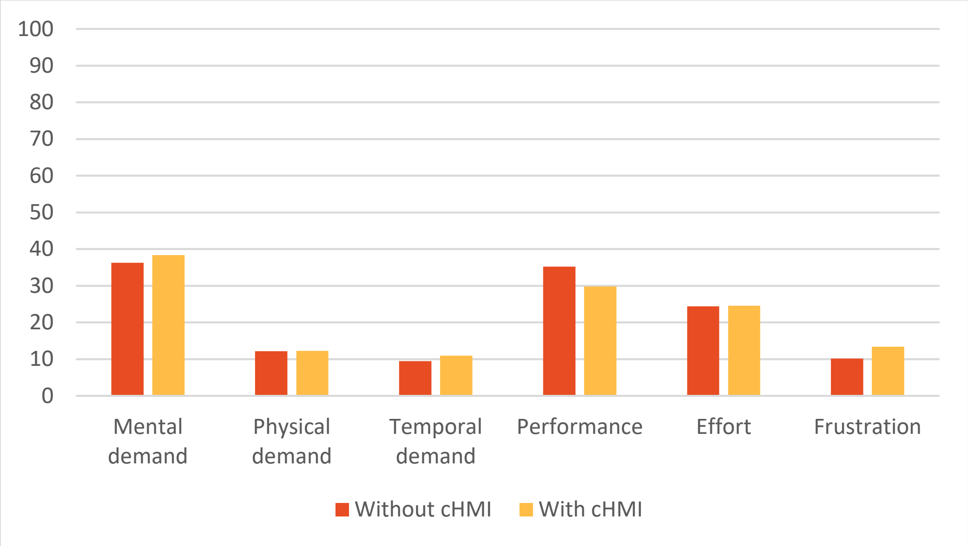


Figure 3-35: Mean NASA TLX scores (range 0-100) reported by pedestrians and drivers (combined) after test blocks with and without cHMI.

3.3.8.1 Adherence

When looking at the adherence of the participants to the joint behaviors recommended by the cHMI, we find that 78% of the drivers that received a yielding recommendation let the pedestrian pass first, while 22% decided to continue driving. Similarly, 76% of pedestrians followed an eHMI signal to cross first, whereas 24% still decided to yield to the vehicle. If the eHMI signalled that the vehicle will continue driving, all pedestrians adhered to the signal and yielded.

3.3.8.2 Age groups

Differences between age groups were analysed among pedestrians. A significantly larger proportion of the older pedestrians' crossings (69%) were in front of the car as compared to younger pedestrians' crossing behavior (56%) (chi-square=10.46, $p=0.001$). Older pedestrians walked slower and took longer to make the crossing decision (Table 3-10).

Table 3-10: Simulator measures related to pedestrian behavior in younger (25-55 years) and older (>70 years) pedestrians.

	Younger	Older
	Mean (SD)	Mean (SD)
CID (s)	8.0 (3.1)	9.6 (4.5)
Crossing time (s)	4.2 (0.8)	4.5 (1.1)
Avg crossing speed (m/s)	1.0 (0.2)	0.9 (0.2)
TCT (s)	12.3 (3.5)	14.1 (4.8)
TTR (s)	15.2 (4.5)	18.8 (6.3)

The subjective evaluation of the cHMI showed similar ratings of usability and trust in younger and older pedestrians. UEQ (t(33)=1.4, p=0.173) and SATI (t(33)=1.41, p=0.170) scores were not significantly different between age groups according to independent samples t-test. Older pedestrians generally rated workload on the NASA TLX scale lower than younger pedestrians (Figure 3-36) but the differences were not statistically significant according to ANOVAs with age group as a between subjects factor and HMI condition as a within subjects factor (data not shown).

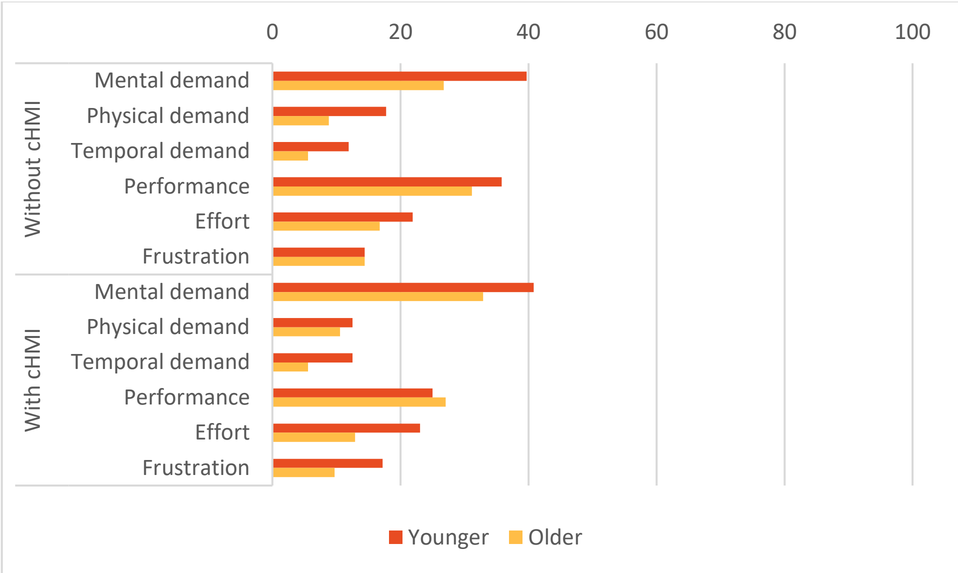


Figure 3-36: Mean NASA TLX scores (range 0-100) reported by older and younger pedestrians after test blocks with and without cHMI

4. Conclusion

The objective of this deliverable was to provide a comprehensive evaluation of the HMIs developed throughout the HEIDI project. It focused on assessing their efficacy and performance through various studies. Notably, real-world studies 8 and 10 examined the functionalities of internal and external HMIs, conducted by MAR in collaboration with RUAS and BMW alongside NISYS, respectively, utilizing fully integrated prototypes as outlined in D6.1. Furthermore, Study 12, conducted as part of WP7 through a simulator study by VTI in collaboration with HRI, explored the cooperative HMI.

These evaluations greatly benefited from prior research and methodologies established in WP5 and documented in deliverable D5.3, which offered foundational insights from small-scale simulator studies. This deliverable effectively extends the work of WP2, WP3, and WP5 by including detailed descriptions of the methods employed, as well as the results and conclusions drawn from each study, providing a thorough and evidence-based assessment of the systems.

All studies showed a significant effect of HMIs in improving situation recognition and resolution. The results of the studies, based on both simulator tests and real-world vehicle tests with iHMI, eHMI, and cHMI — even with a relatively small number of participants — demonstrate that these approaches, whether implemented individually (iHMI, eHMI) or as a combined system (cHMI), can significantly enhance both the comfort and safety of all road users.

Considering the ongoing development of driver assistance systems towards autonomous driving, it is clear that such functions will become increasingly important. On the one hand, systems like iHMI inform the driver selectively, only in situations where they are not actively engaged in the traffic environment, thereby avoiding unnecessary distractions and ensuring focused attention. On the other hand, systems like eHMI enable the vehicle — whether manually driven or operating autonomously — to actively communicate with its surroundings. This external communication helps vulnerable road users, such as pedestrians and cyclists, to better and more intuitively assess the traffic situation, ultimately improving overall road safety.

Based on the findings of this study, it is recommended that targeted methods be developed that pursue two main objectives: first, the standardisation of communication structures across all HMI systems (iHMI, eHMI, and cHMI); and second, the establishment of structured approaches for further investigations and validation of these systems in real-world operational environments.

Results from study 12 indicated that the traffic efficiency decreased for drivers but increased for pedestrians when cHMI was active, since TTR (time from start of the crossing scenario until the car had passed the pedestrian by 30 m) was longer but TCT (time from start of the scenario until the pedestrian reached the other side of the road) was shorter with cHMI. However, this difference was mediated by the change in yielding/crossing behavior when cHMI was active. Looking only at interactions where the pedestrian crossed in front of the car, there was no effect of cHMI on crossing initiation delay, TTR or TCT. Participants' evaluations regarding usability and trust were positive and the cHMI was perceived to have safety benefits.

Differences between studies can to some extent be attributed to different study designs. Most importantly, the behaviors of the pedestrian in study 8 and the driver in study 10 were controlled in the study design whereas both the driver and pedestrian were free to act in study 12. This introduced another level of uncertainty in study 12, which might have affected yielding and crossing behavior.

One aspect that could not be addressed within the scope of this study was the participants habituation to a new driver assistance system. Since the study was conducted over a relatively short period of time, it must be assumed that the drivers had not yet fully familiarised themselves with the system's behaviour, feedback, and intervention strategies. In practice, however, drivers require a certain amount of time and experience to build trust in and understanding of a new assistance system. Only after such a familiarisation phase can meaningful adjustments and further parameterisations of the system be carried out to optimise the level of assistance and ensure seamless integration into the driver's natural driving behaviour. Without this period of adaptation, drivers might either underutilise the system or react to its interventions with hesitation, both of which could impact the system's overall effectiveness and the validity of the measurements taken during the study. Future studies should therefore consider longer observation periods or longitudinal studies to systematically investigate how driver-system interaction evolves over time and to identify opportunities for system refinements based on real-world usage patterns.

5. Next steps

The current project documentation, D7.1, is one of four deliverables in WP7 and lays the foundation for the forthcoming project deliverables. In just one month, D7.2 - focused on the HMI rating - will be submitted. In addition, further in-depth discussions regarding D7.1 will follow, incorporating meta-analyses to fully extract all insights from the studies. These findings will, in turn, serve as the basis for D7.3, "Final Recommendations for Standardization," which will consolidate guidelines on how HMIs in road traffic should be designed—for instance, to reduce TTR—and potentially improve urban traffic flow in everyday European life. The project will conclude in four months with a real-world demonstration. At that time, the two test vehicles from the studies will be showcased in actual operational conditions, and the results will be summarized in D7.4, ensuring that the overall project ends with substantial, high-quality contributions.

6. Abbreviations

Term	Definition
AIDA	Human-centered Interactive **A**rtificial **I**ntelligence **DA**ta-Incubation Center
BEV	Battery electric vehicle
cHMI	Cooperative Human Machine Interface
CID	Crossin Initiation Delay
CS	Crossing speed
CT	Crossing Time
eHMI	External Human Machine Interface
HEIDI	Holistic and adaptivE Interface Design for human-technology Interactions
ICE	Internal combustion engine
iHMI	Internal Human Machine Interface
IR	Infrared
Km/h	Kilometers per hours
LED	Light emitting diode
MoCap	Motion Capturing
NASA TLX	NASA Task Load Index
PU	Public
R	Document, Report
RGB	Red, Green, Blue
RT	Reaction time
SW	Software
TTC	Time to collison
TTR	Time to resolve
VR	Virtual Reality
VRU	Vulnerable Road User
WP	Work Package

7. References

[1] Fortner, B. (1998). HDF: The hierarchical data format. Dr Dobb's J Software Tools Prof Program, 23(5), 42.

[2] R. Leute, F. Maier, P. Pretto, M. A. Sotelo Vázquez, „From Signaling to Human-Machine-Interface: new Roles for exterior Lighting", ISAL Proceedings (pp. 118-123), 2023.

[3] Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139-183). North-Holland. https://doi.org/https://doi.org/10.1016/S0166-4115(08)62386-9

[4] Lewis, J. R. (2018). The System Usability Scale: Past, Present, and Future. *International Journal of Human–Computer Interaction*, *34*(7), 577-590. https://doi.org/10.1080/10447318.2018.1455307

[5] Dehn, D. M. (2008). Assessing the Impact of Automation on the Air Traffic Controller: The SHAPE Questionnaires. Air Traffic Control Quarterly, 16(2), 127-146. https://doi.org/10.2514/atcq.16.2.127

[6] Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and evaluation of a short version of the user experience questionnaire (UEQ-S). International Journal of Interactive Multimedia and Artificial Intelligence, 4 (6), 103-108.

[7] Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012). *Predicting information technology usage in the car: towards a car technology acceptance model* Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Portsmouth, New Hampshire. https://doi.org/10.1145/2390256.2390264

Annex (Questionnaires used in Study 10)

ID:____________________________

NASA TLX Fragebogen (auf Deutsch)

Der NASA Task Load Index (NASA TLX) Fragebogen wird verwendet, um die subjektive Arbeitsbelastung bei der Durchführung einer Aufgabe zu bewerten. Bitte bewerten Sie jede Dimension der Arbeitsbelastung auf Basis Ihrer Erfahrung mit der durchgeführten Aufgabe.

Anweisungen:

Bewerten Sie jede der folgenden Aspekte auf einer Skala von 1 (sehr niedrig) bis 20 (sehr hoch).

1. **Mentale Anforderungen:**
   Wie viel geistige Anstrengung war bei der Informationsaufnahme und -verarbeitung erforderlich (z.B. Denken, Entscheiden, Rechnen, Erinnern, Hinsehen, Suchen...)? War die Aufgabe einfach oder komplex, leicht oder anspruchsvoll?
   [1 -- 20]
2. **Körperliche Anforderungen:**
   Wie viel körperliche Aktivität war erforderlich? War die Aufgabe entspannt oder anstrengend?
3. [1 -- 20]
4. **Zeitdruck:**
   Wie viel Zeitdruck haben Sie empfunden? War der Zeitrahmen für die Aufgabe angemessen oder stressig?
5. [1 -- 20]
6. **Leistungsniveau:**
   Wie zufrieden sind Sie mit Ihrer Leistung bei dieser Aufgabe? Haben Sie Ihre eigenen Erwartungen erfüllt?
7. [1 -- 20]
8. **Anstrengung:**
   Wie viel Anstrengung mussten Sie aufbringen, um Ihre Leistung zu erzielen?
9. [1 -- 20]
10. **Frustrationsniveau:**
    Wie frustriert, gestresst oder genervt waren Sie während der Aufgabe?
11. [1 -- 20]

ID:_______________

System Usability Scale

Einführung:

Der System Usability Scale (SUS) Fragebogen dient der Beurteilung der Benutzerfreundlichkeit eines Systems. Bitte bewerten Sie jede Aussage basierend auf Ihren Erfahrungen mit dem System.

Anweisungen:

Bewerten Sie jede der folgenden Aussagen auf einer Skala von 1 bis 5, wobei:
1 = Stimme überhaupt nicht zu
2 = Stimme nicht zu
3 = Neutral
4 = Stimme zu
5 = Stimme voll und ganz zu

Fragen:

1. Ich würde dieses System häufig verwenden.
   [1] [2] [3] [4] [5]
2. Das System ist unnötig komplex.
   [1] [2] [3] [4] [5]
3. Das System war einfach zu bedienen.
   [1] [2] [3] [4] [5]
4. Ich denke, dass ich die Unterstützung eines Technikers benötigen würde, um dieses System nutzen zu können.
   [1] [2] [3] [4] [5]
5. Die verschiedenen Funktionen in diesem System sind gut integriert.
   [1] [2] [3] [4] [5]
6. Es gibt zu viele Inkonsistenzen in diesem System.
   [1] [2] [3] [4] [5]
7. Die meisten Menschen würden den Umgang mit diesem System schnell erlernen.
   [1] [2] [3] [4] [5]
8. Das System ist sehr umständlich zu bedienen.
   [1] [2] [3] [4] [5]
9. Ich fühlte mich sicher im Umgang mit dem System.
   [1] [2] [3] [4] [5]
10. Ich musste viele Dinge lernen, bevor ich mit dem System arbeiten konnte.
    [1] [2] [3] [4] [5]

Teilnehmer ID: P______

	Stimme gar nicht zu (1)	Stimme nicht zu (2)	Neutral (3)	Stimme zu (4)	Stimme voll zu (5)
Die Erkennbarkeit des Auges ist brauchbar 	•	•	•	•	•
Die Erkennbarkeit der Chevrons ist brauchbar 	•	•	•	•	•
Die Größe des Auges ist brauchbar 	•	•	•	•	•
Die Größe der Chevrons ist brauchbar 	•	•	•	•	•
Die Helligkeit der Symbole ist passend	•	•	•	•	•
Die Auflösung des Auges ist brauchbar 	•	•	•	•	•
Die Auflösung der Chevrons ist brauchbar 	•	•	•	•	•
Das Auge ist logisch und verständlich 	•	•	•	•	•

	Stimme gar nicht zu (1)	Stimme nicht zu (2)	Neutral (3)	Stimme zu (4)	Stimme voll zu (5)

Die Chevrons nach oben sind logisch und verständlich 	•	•	•	•	•
Die Chevrons nach unten sind logisch und verständlich 	•	•	•	•	•
Der Unterschied zwischen Anhalten und Weiterfahren war mit Auge deutlich  **vs** 	•	•	•	•	•
Der Unterschied zwischen Anhalten und Weiterfahren war ohne Auge deutlich  **vs** 	•	•	•	•	•
Ich habe mich mit Auge sicherer gefühlt die Straße zu überqueren im Vergleich ohne HMI 	•	•	•	•	•
Ich habe mich ohne Auge sicherer gefühlt die Straße zu überqueren im Vergleich ohne HMI 	•	•	•	•	•
Das eHMI kann Missverständnisse zwischen Fahrern und Fußgängern reduzieren	•	•	•	•	•
Die Zeit zum Informieren war lang genug	•	•	•	•	•
	Stimme gar nicht zu (1)	Stimme nicht zu (2)	Neutral (3)	Stimme zu (4)	Stimme voll zu (5)
Ich sehe einen Nutzen im HMI	•	•	•	•	•
Ich sehe ein Risiko im HMI (Ablenkung, Blendung)	•	•	•	•	•

Ich konnte die Symbole von Anfang an gut verstehen	•	•	•	•	•
Ich konnte die Symbole zum Ende hin gut verstehen	•	•	•	•	•