Results and conclusion from the validation tests

Deliverable Number	D5.3
Deliverable Type	R – Document, Report
Dissemination Level	PU (Public)
Author(s)	Anna Sjörs Dahlman, Michael Brunner, Markus Rehmann, Stina Larsson, Paolo Pretto, Manuela Prior, Markus Amann, Mateo Jukic
Document Version & Status	V2.0 | 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	Name	Organisation
Anna Sjörs Dahlman	VTI	Paolo Pretto	VIF
Michael Brunner	RUAS	Manuela Prior	VIF
Markus Rehmann	RUAS	Mateo Jukic	VIF
Stina Larsson	VTI	Markus Amann	HRI-EU

Reviewers

Name	Organisation	Date
Thomas Weisswange	HRI-EU	2024-08-28
Felix Maier	MAR	2024-08-16

Change History

Version	Date	Name/Organisation	Description
0.1	2024-06-05	VTI	ToC
0.2	2024-07-27	VTI	First draft
0.3	2024-08-09	VIF	Study 5
1.0	2024-08-12	VTI, RUAS	First draft for internal review
1.1	2024-08-26	VTI, VIF	Final draft for internal review
2.0	2024-08-30	VTI, VIF	Final (revised) version

Table of Contents

List of Figures

List of Tables

1. Executive Summary

This deliverable presents the results of the small-scale validation tests conducted in WP5. This deliverable is linked to the work performed in T5.3, Development of infrastructure for co-simulation tests, and T5.4, Design and implementation of the validation tests. It relates to project objective 3, which is to develop suitable validation methods for assessing fluid, cooperative Human-Machine Interface (HMI) solutions. Three small-scale studies (study 5, 6, and 7) were conducted for assessing validation methods for the HEIDI fluid and cooperative HMI system. The objectives of the validation studies were to i) assess the performance of the validation methods, ii) provide early insight into how different users interact with this technology, and iii) guide the design of evaluation tests in WP7.

This deliverable includes descriptions of the methods used in the validation studies, results and conclusions from each study, and a general evaluation and conclusion regarding the suitability of various validation methods for further use in the evaluation of the final HEIDI HMI solutions. Initial exploratory testing of the iHMI and eHMI was performed in WP2 and WP3 to investigate the usability, safety and comfort of the first prototypes of the HEIDI fluid human-machine interface. The results of the exploratory studies informed the design of the final iHMI and eHMI prototypes used here, ensuring improved usability and efficiency of the interface.

The main objective of study 5 was the assessment of the revised version of the fluid internal HMI (iHMI) with regards to safety, comfort and usability and to investigate the validity and suitability of the developed methodologies and tools for the assessment of safety, comfort, and usability of adaptive fluid HMI. Based on the outcome of previous exploratory studies of the initial iHMI, a final version of the iHMI was developed and tested for regular, distracted and elderly drivers in a driving simulator. The study included 14 crossing scenarios along a 5 km urban route. A counterbalanced within-subject design was applied where HMI support (no HMI or different versions of the iHMI) and distraction were manipulated.

Study 6 aimed to simulate real-world pedestrian-vehicle interactions and assesses the effectiveness of different eHMIs in various scenarios involving different levels of pedestrian awareness and interaction dynamics. The experiment setup involved a pedestrian group consisting of two subjects with one additional virtual pedestrian that is computer-controlled. One subject wore a Virtual Reality (VR) headset, while the other did not, simulating a distracted pedestrian. Subjects were exposed to repeated external HMI (eHMI) stimuli to observe differences in crossing behavior, using a counterbalanced within-subject design. The condition variables include the type of eHMI (None, Short-Term, or Long-Term) and the presence of other pedestrians (on the same side or the other side).

The main objective of study 7 was to test validation methods for the cooperative HMI (cHMI) system using networked simulators (co-simulation) in regular drivers and regular and older pedestrians. Various validation methods were tested in the co-simulation setup with a first version of the cHMI system. Tests were performed using a networked VR headset and a driving simulator with augmented reality (AR) features. 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, a driver distraction task was used in 25 % of the pedestrian-driver interactions.

All three studies used both subjective and objective validation tools. Some of the validation tools were the same across the three validations studies whereas others were tailored for the objective of a specific study. The main differences between studies were the target group (drivers or pedestrians) and whether there was full control of the vehicle or the pedestrian behavior in the simulator scenario. Thus, study 5 focused on the experience of the driver, study 6 on the pedestrians and study 7 on the interaction between the driver and pedestrian. Subjective ratings were given via questionnaires and interviews. Objective data were mainly measures derived from simulator log data, motion capture and eye-tracking.

Since the focus of the studies was to evaluate the validation methods, a relatively small number of participants were recruited for each study. Due to the limited sample size, the interpretation of the results should be done with caution.

In study 5, the introduction of an iHMI improved regular drivers' behavior, leading to earlier braking and reduced pressure on the brake pedal compared to baseline conditions, especially when accounting for driver distraction. Regular drivers found the iHMI useful and efficient, preferring versions that adapt to attentional states, as they felt safer and more informed. Trust in automation and situational awareness remained high across all conditions, with participants appreciating the iHMI for its practicality rather than enjoyment. Objective measures showed no significant impact of vertical fluid iHMI on older drivers' behavior, although it may encourage slightly more cautious driving. Participants reported similar stress, usability, and trust across iHMI versions, with equal preference but a desire for more personalized interface options.

In study 6 both the Short-Term HMI (HMI1) and Long-Term HMI (HMI2) eHMIs demonstrated potential benefits compared to the baseline condition without an eHMI (HMI0). The objective measures suggest that both eHMIs resulted in faster crossing initiations and reduced Time-To-Resolve (TTR) for pedestrian-vehicle interactions. This indicates that the presence of eHMIs may facilitate more efficient decision-making processes for pedestrians when the vehicle is autonomous.

The results of study 7 indicated that the current version of the cHMI needs improvement to provide a real benefit to drivers and pedestrians. The effects of the cHMI were not as expected since the participants found it less clear who had priority when HMI was triggered. Time to resolve was longer indicating that the cHMI did not improve efficiency. It was found to be problematic when the eHMI communicated that the driver was intending to yield when the driver could, in fact, choose not to adhere.

Summarizing the conclusions from all three studies, it was evident that the HEIDI concept worked well when iHMI and eHMI were implemented separately and there was full control of one of the agents' behaviors. The cHMI needs further refinement to be able to handle ambiguous situations and non-adherence to the communicated message. Regarding the validation methods, the combination of subjective and objective measures provided a comprehensive evaluation of both the behavior and opinions of drivers and pedestrians. A set of suitable validation metrics to evaluate the effective safety and usability of the HEIDI HMI solutions were defined.

The validation studies in combination with experiences from the experimental studies preformed in WP2 and WP3 provided insights about which validation methods are suitable for the evaluation of the final HEIDI HMI solutions. Based on the results from the validation studies, a set of general recommendations are given for the evaluation tests in WP7. In short, both subjective and objective measures are recommended for use in the evaluation studies. The validation studies showed that a combination of questionnaires and interviews with participants provides a comprehensive view of their experiences. The objective and subjective validation metrics that are recommended for use in WP7 evaluation tests are presented in section 7.6.

Keywords: fluid human-machine interface, validation methods, simulator studies, driver distraction, older drivers

2. Objectives

This deliverable presents the results of the small-scale validation tests conducted in WP5. This deliverable is linked to the work performed in T5.3, Development of infrastructure for co-simulation tests, and T5.4, Design and implementation of the validation tests. It relates to project objective 3, which is to develop suitable validation methods for assessing fluid, cooperative HMI solutions. The connected project KPI is to define at least five concrete validation metrics to evaluate the effective safety and usability of the HEIDI HMI solutions. Three small-scale simulator tests were designed and implemented where drivers and pedestrians separately or simultaneously (via connected simulators) interacted with parts of the HEIDI system. The first study (Study 5) focused on testing the fluid internal HMI (iHMI) for elderly drivers and distracted drivers. The second study (Study 6) tested the external HMI (eHMI) for groups of pedestrians. The third study (Study 7) tested a first version of the cooperative HEIDI system (cHMI).

The objectives of the validation studies were to i) assess the performance of the validation methods, ii) provide early insight into how different users interact with this technology, and iii) guide the design of evaluation tests in WP7. In WP7, five studies will be performed to demonstrate and evaluate the final HEIDI HMI concepts. The validation methods should therefore be suitable for use in the following studies: Study 8 – a controlled real-world evaluation test of the HEIDI iHMI, Study 9 – a field demonstration of the iHMI and Osmotic Layer, Study 10 - a controlled real-world evaluation test of the HEIDI eHMI, Study 11 – a field demonstration of the eHMI, and Study 12 – a simulator evaluation test of the cooperative HEIDI HMI.

The validation studies were designed based on recommendations from the state-of-the-art report from T5.2, Validation methods of adaptive HMIs. The studies used the HEIDI HMI concepts developed in WP2 (internal fluid HMI), WP3 (external HMI) and WP4 (cooperative HMI). Initial exploratory testing of the iHMI and eHMI was performed in WP2 and WP3 to investigate the usability, safety and comfort of the first prototypes of the HEIDI fluid human-machine interface. The results of the exploratory studies informed the design of the final iHMI and eHMI prototypes used here, ensuring improved usability and efficiency of the interface. Study 7 was the first test of the cHMI and therefore also served as an exploratory study of the cHMI features.

The tests were conducted on driving and pedestrian simulators located at VTI, VIF and RUAS using scenarios representing relevant use cases defined in WP1. The adequacy of the methods was evaluated regarding aspects of technical and methodological nature. Various objective and subjective validation tools were defined and evaluated. The outcomes from the validation studies provide necessary inputs for prototyping (WP6) and evaluation (WP7) studies. A list of recommendations for the forthcoming studies in the HEIDI project was created and suitable validation metrics were defined. The outcomes will also be used in T5.5, Recommendations for Legal Framework & Standardization and provide input to WP8.

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 5 - small-scale test of fluid iHMI for regular, distracted and elderly drivers

The main objective of this study was the assessment of the revised version of the fluid iHMI (internal Human-Machine Interface) with regards to safety, comfort and usability. The initial version was tested in two exploratory studies that addressed regular (also distracted) and elderly drivers, respectively, as described in D2.3 “Initial simulator studies results”. Based on the outcome of the exploratory studies and on user-centered iterative design principles, the initial iHMI was revised, as described in D2.4 “revised fluid iHMI concept”. Then, the study described in this deliverable has tested the resulting final version of the iHMI for regular, distracted and elderly drivers. The specific changes of the current iHMI version are described in detail in section 3.1.2.

Similarly to the exploratory studies (Study 1 and Study 2, see D2.3), the present study (Study 5) aimed at identifying the benefits and drawbacks of the iHMI revised concept in a simulated environment. Moreover, the evaluation methods that were developed for the previous studies were refined accounting for the lessons learned and the experience gained with other exploratory studies carried out in other work packages and tasks (WP3, WP5). Therefore, another objective of the present study was to investigate the validity and suitability of the developed methodologies and tools for the assessment of safety, comfort, and usability of adaptive fluid HMI.

The outcome of the present study will inform the implementation of the physical prototypes in real demonstration vehicles, where the cooperative HMI can be demonstrated in predefined conditions (WP6, WP7).

3.1.1 Use case

This evaluation study investigated the use case 1, as defined in D1.2 “Use case definition”, which was devised to test and validate the iHMI – internal Human-Machine Interface –, i.e., the interface between the vehicle and the driver, depending on different configurations of crossing pedestrians. Specifically, 14 situations were implemented in the driving simulation, corresponding to 14 *crossing events* and occurring in predefined track *segments* (see below for more details), resulting from the combination of visible/occluded crosswalk marking, presence/absence of one or more crossing pedestrians, including children and crossing outside the designated crosswalk (“with offset”) (Table 3-1).

All the iHMI messages were designed to address specific user categories, as defined in D1.1 “Description of user needs”. In this study, a “**regular**” driver is considered a person aged between 20 and 50, with a valid driving license for at least three years, and with an average car usage of at least once a week. A “**distracted**” driver fulfils the same requirement of a regular driver. However, the person is engaged in a secondary non-driving related task (e.g., playing a game on a tablet) and is, therefore, in a distracted state. An “**older**” driver is a person aged above 70, with a valid driving license for at least three years, and with an average car usage of at least once a week.

Table 3-1: Study 5 – crossing events

Event No.	Event	Road side
1	crosswalk only	
2	crosswalk occluded - no pedestrian	right
3	crosswalk occluded - no pedestrian	left
4	crosswalk occluded	right
5	crosswalk occluded	left
6	pedestrian	right
7	pedestrian	left
8	pedestrian with offset	right
9	pedestrian with offset	left
10	pedestrian with offset in group	right
11	pedestrian with offset in group	left
12	child crossing - no crosswalk	right
13	child crossing - no crosswalk	left
14	pedestrian changes mind	right

Crossing event 14 "Pedestrian changes mind" was introduced in this study as additional challenge for the iHMI (and the driver) as it represents a not-so-uncommon situation of a pedestrian who first has the intention to cross but suddenly decides to step back and yield to the vehicle. A simplified graphical representation of all 14 crossing events is presented in Figure 3-1.

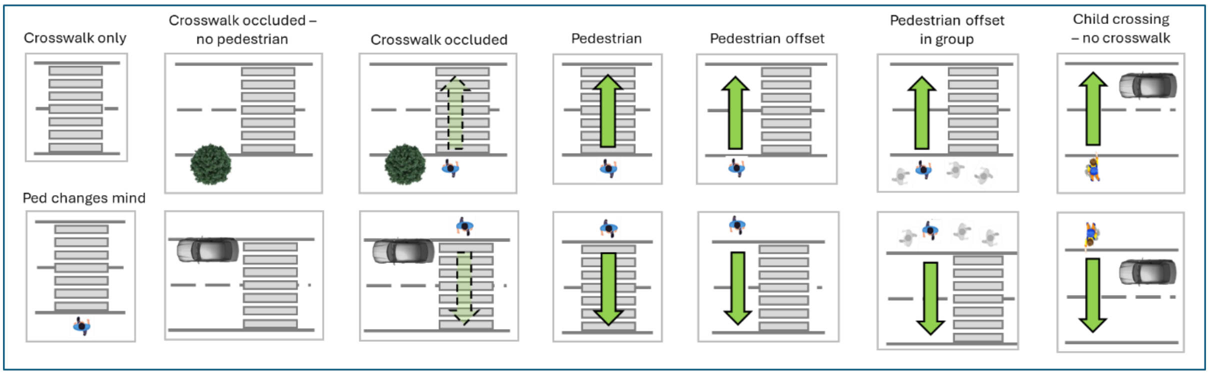


Figure 3-1: Study 5 – crossing events (graphical representation)

3.1.2 HMI solutions

The iHMI consists of several audio-visual components that can be activated based on a series of conditions and that allow multi-sensory communication to the driver. The visual component allows the display of warning icons that are over-imposed on the driving simulator screen, corresponding to the projection onto a virtual HUD (Head-Up Display). In addition, the icons can be repeated on a touch screen, which serves as an in-vehicle infotainment display that is used to carry out a distracting secondary task. An additional visual component is provided by an LED strip located at the bottom of the windscreen frame, horizontally connecting both front pillars of the vehicle. In some HMI variations the LEDs are activated to highlight the presence of the detected pedestrians. Those LEDs also "follow" the motion of the pedestrian across the road to facilitate the perception of the pedestrian's motion and direction. The audio channel consists of a 3D spatial sound system, which can reproduce auditory warning in form of "beep" sounds of voice messages. Figure 3-2 shows a schematic representation of all the iHMI components.

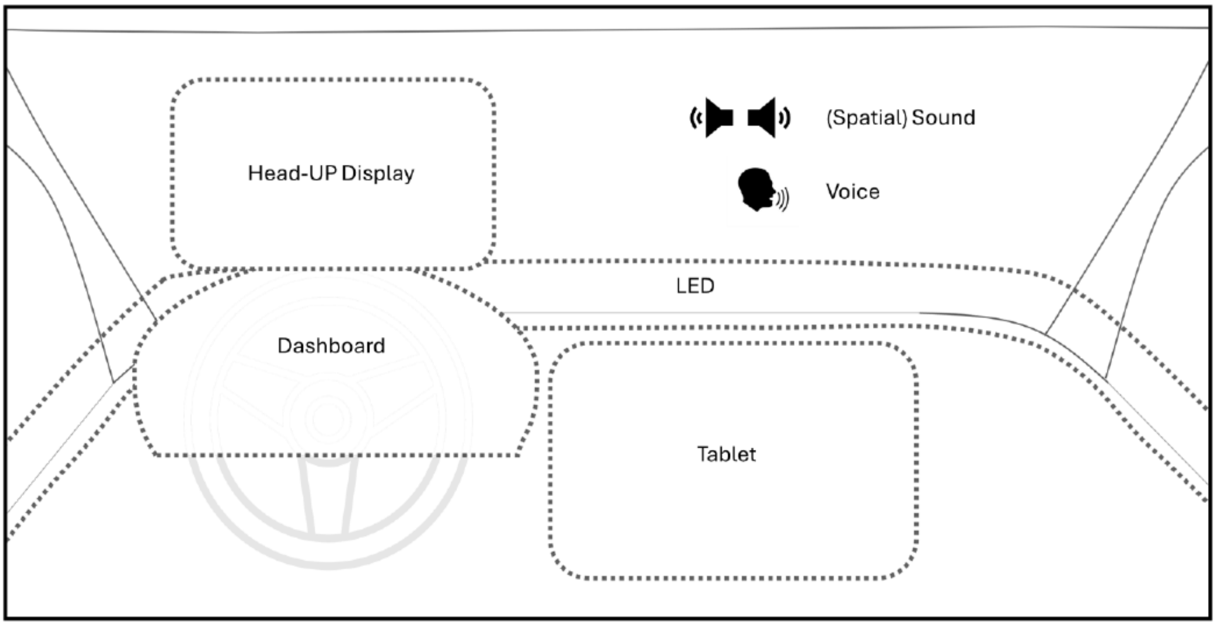


Figure 3-2: Study 5 – iHMI components for visual and auditory information

The iHMI implementation in this study consisted of five variations of the components described above. The different versions were tested on the same road and events to address specific user needs and to evaluate the efficacy of the implemented HMI features to adapt to the drivers' states and conditions.

The principles of **fluid interaction** that inspired the design of the iHMI, its components and logic, are explained in detail in D2.4. Below, a short summary of those principles is provided:

- The interface must adapt to drivers who are attentive to the road ahead, as well as to those who are temporarily distracted, i.e. engaged in a secondary activity on the vehicle infotainment display, or to the needs of elderly drivers (see D1.1 "Description of user needs"). Therefore, a fluid interface adapts to the continuously switching driver's state (based on the data retrieved by eye-tracking).
  - Only the information that is needed is provided when and where necessary, in a way that minimizes distraction and workload and maximizes understanding, acceptance and safe (re)actions. Therefore, fluid adaptiveness elicits different alertness levels according to the evolution of the situation over time and to the driver's actions.
- Messages to the driver can vary in timing, sensory modality, location and amount of information to be displayed, according to the two principles mentioned above.

Figures 3-3 to 3-7 show the final design of the developed iHMI for regular, distracted and older drivers, and how each iHMI acts on different alert levels. Alert levels are defined based on the time-to-collision (TTC) between vehicle and approaching pedestrian. The smaller the TTC gets, the higher the escalation level – saliency and urgency of the iHMI messages targeting the driver – becomes. In general, the iHMI escalation levels are described below, while specific design differences will be described afterwards, in the iHMI variations description.

- "**L0**" (Level 0) is only implemented in the HMI logics, and it is used to indicate that the pedestrian warning system is unavailable. Even though this level is outside the scope of the present study and, therefore, not further considered in this document, its design differed for regular and older drivers (see below).

- "**L1: Inform**" informs the driver of an approaching crosswalk. Since this information is present in the simulation database and could be available also in real navigation systems, the TTC is rather conservatively set to 10s. That is, the message is displayed 10 seconds before the vehicle reaches a crosswalk. As this is a low-risk situation with ample time for reaction, the level is only meant to inform the driver preventively
- "**L2: Warn**" escalates the message to the driver as it provides information of a detected pedestrian with recognized intention to cross with a TTC of 8 seconds. This message provides information about the roadside in which the pedestrian is located and may be accompanied by sound or voice messages and LED lights.
- "**L3: Alert**" is triggered when TTC is between 4-6 seconds and the driver has not reacted to the previous message (e.g., reducing the speed). This message always requires the driver's prompt intervention to reduce the speed. Therefore, the saliency of the visuo-auditory messages is maximized for all drivers, compatibly with their attentive states and abilities, and the "STOP" request is prompted.
- "**L4: Emergency**" notifies the driver that the automatic emergency response of the vehicle has taken over control and triggered an emergency stop. This level, similar to "L0", is only implemented in the iHMI but, as it falls outside driver's control, will not receive further consideration within the scope of this study.

The five iHMI variations under test can be distinguished by their compliance to fluid principles and target user:

- **Non-fluid iHMI for regular drivers**: this version is equivalent to the state-of-the-art systems that trigger a warning when an imminent collision is detected based on on-board sensors. Neither preemptive warning nor time-based escalation are provided (Figure 3-3). In case of system unavailability, a text message is shown on the dashboard and a short accompanying sound is played once. In case of imminent collision with a pedestrian, a HUD icon shows a pedestrian crossing and the explicit recommendation to stop 5 seconds before collision, together with LED lights and a specific sound signal. If the driver does not properly react to the recommended action, the automatic emergency response is triggered two seconds before collision and the car stops autonomously. A HUD icon with the indication of emergency stop is shown, together with a fully illuminated LED red strip at the bottom of the windshield and the voice message "Nothalt" ["emergency stop" in German].

	Non Fluid Regular iHMI *if driver is attentive / distracted:*
L0: System Failure (unable to detect pedestrian)	Pedestrian detection not available + **sound**
L1: Inform (approaching crosswalk)	
L2: Warning (location of detected pedestrian)	
L3: Alert (behavior recommendation)	TTC = 5s  + **LED** + **sound**
L4: Emergency (automatic braking)	TTC = 2s  + **LED** + **voice**

Figure 3-3: Study 5 – Non-fluid iHMI for regular drivers

- **Horizontal fluid iHMI for regular/distracted drivers**: this version works like the previous one but produces different messages based on the detected driver's attentional state. If the driver is attentive the message is triggered 5 seconds before the potential collision (TTC = 5s) and the HUD icon is accompanied by LED lights and a specific warning sound. If the driver is distracted (i.e., attending a secondary task on the infotainment display) the messages is triggered 1 second earlier (TTC = 6s) and the HUD icon is repeated and magnified on the infotainment screen, interrupting the secondary task (Figure 3-4).

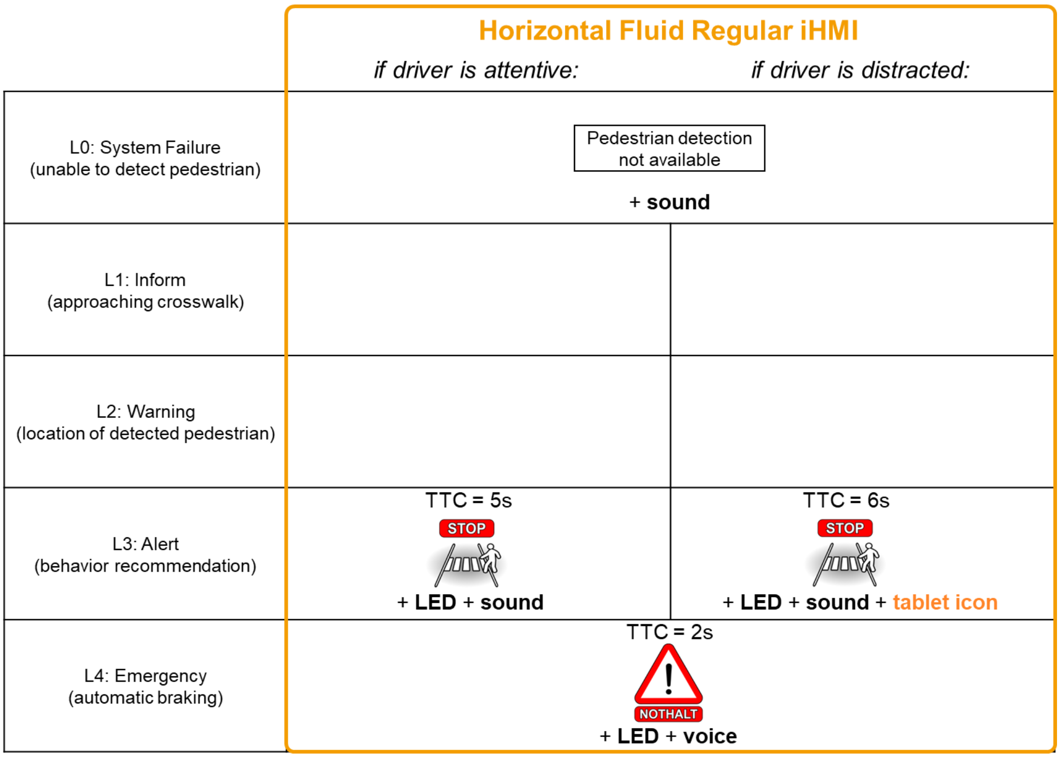


Figure 3-4: Study 5 – Horizontal fluid iHMI for regular drivers

- **Vertical fluid iHMI for regular drivers**: this version features all levels of escalation ("L0"-"L4", as described above), therefore adapting to the driver's reaction and the evolution of the situation over time but does not consider the driver's attentional state (Figure 3-5). The next escalation level is triggered only if the driver does not (or chooses not to) comply with the recommended action. "L1: Inform" is triggered when TTC = 10s and indicates whether crosswalk markings are present or not in a HUD icon. At TTC = 8s the "L2: Warn" message is triggered if a crossing pedestrian is detected. In "L2" the HUD icon also indicates from which side (left/right) the detected pedestrian is crossing. LEDs at the basis of the windshield are turned on from the side of the detected pedestrian to help the driver gain awareness. If the driver does not react to the situation within 3 seconds (e.g., does not reduce the speed) the warning is escalated to "L3". "L3" message depicts a HUD icon with the explicit recommendation to stop, together with a small yellow LED strip at the bottom of the windshield frame that highlights the position of the detected pedestrian across the street. A specific sound is also played at the same time.

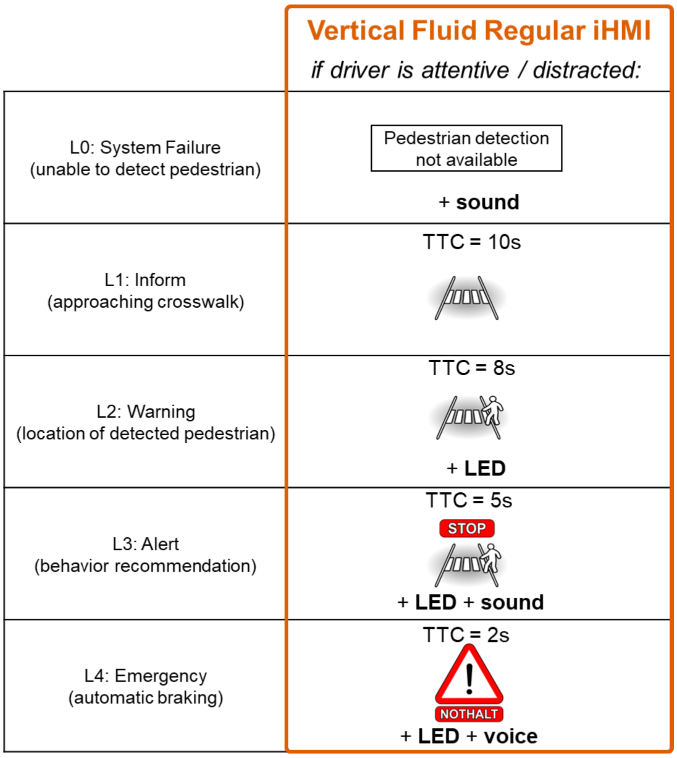


Figure 3-5: Study 5 – Vertical fluid iHMI for regular drivers

- **Fully fluid iHMI for regular/distracted drivers**: This version combines the previously described horizontal fluid iHMI and vertical fluid iHMI versions, as it accounts for both driver's attentional state and reactions over time, providing the full fluid features of the iHMI (Figure 3-6). With respect to the attentive driver messages, the message for distracted drivers in "L1: Inform" includes an additional sound, to help them gain awareness. Similarly, in "L2: Warn" a spatial sound (left/right) is added to the distracted driver's message. Finally, in "L3: Alert" the HUD icon is repeated and magnified on the infotainment screen and the message is triggered 1 second earlier.

	Fully Fluid Regular iHMI – *if driver is attentive:*	Fully Fluid Regular iHMI – *if driver is distracted:*
L0: System Failure (unable to detect pedestrian)	Pedestrian detection not available + **sound**	
L1: Inform (approaching crosswalk)	TTC = 10s	TTC = 10s + **sound**
L2: Warning (location of detected pedestrian)	TTC = 8s + **LED**	TTC = 8s + **LED** + **spatial sound**
L3: Alert (behavior recommendation)	TTC = 5s STOP + **LED** + **sound**	TTC = 6s STOP + **LED** + **sound** + **tablet icon**
L4: Emergency (automatic braking)	TTC = 2s NOTHALT + **LED** + **voice**	

Figure 3-6: Study 5 – Fully fluid iHMI for regular drivers

- **Vertical fluid iHMI for older drivers**: This version, similar to the vertical fluid iHMI for regular drivers, is adapted to the specific needs of elderly drivers, as described next. “L0” informs the older driver that the system is unavailable with an additional voice message: “Pedestrian detection system not available”. In “L1: Inform” the HUD icon with crosswalk is accompanied by a sound, similarly to the message for distracted drivers. The HUD icon in “L2: Warn” shows a pedestrian encircled by a red line and a voice message already provides the recommendation to reduce the driving speed. No LED or spatial sound is activated. Rather, a simple and loud sound is followed by the voice command to stop. If the older driver does not react as prompted, “L3: Alert” is triggered 4 seconds before collision. This may seem a contradiction, as for regular and distracted drivers this level is triggered 5 and 6 seconds before collision, respectively. However, elderly drivers have already received a first explicit behavioral recommendation in the previous level with a voice prompt. Therefore, differently from all other iHMI versions, the L3 message for older drivers represents the second explicit recommendation to stop, and it is therefore timed between the first warning (TTC = 8s) and the potential collision. Also, the HUD icon is simplified as compared to the iHMI L3 messages for regular and distracted drivers, as it only shows the word “STOP” with bigger, i.e., more easily readable, font. LED and spatial sound are not used for elderly drivers, as previous studies have shown that these features may induce confusion (D2.3, D2.4) (Figure 3-7).

	Vertical Fluid Older iHMI
L0: System Failure (unable to detect pedestrian)	Pedestrian detection not available + **sound** + **voice**
L1: Inform (approaching crosswalk)	TTC = 10s + **sound**
L2: Warning (location of detected pedestrian)	TTC = 8s + **sound** + **voice** **„Stop!“**
L3: Alert (behavior recommendation)	TTC = 4s STOP + **voice** **„Stop!“**
L4: Emergency (automatic braking)	TTC = 2s NOTHALT + **LED** + **voice**

Figure 3-7: Study 5 – Vertical fluid iHMI for older drivers

3.1.3 Research questions

The research questions listed below were formulated for study 5:

1. Does a fluid interface provide advantages compared to a non-fluid HMI?
2. Does a fluid interface provide disadvantages compared to a non-fluid HMI?

- How can the different HMI components be properly assessed?

3. Which methods are the most suitable to assess a fluid HMI as-a-whole?

However, due to the different characteristics of regular, distracted and older drivers, more specific research questions were formulated and investigated in two versions of study 5: study 5a focused on regular drivers, also in distracted state; study 5b focused on older drivers.

The specific research question of study 5a was:

5a. Which type of fluidity (vertical vs. horizontal) provides more benefits?

The specific research question of study 5b was:

5b. How do older drivers perceive the iHMI designed for regular drivers?

3.1.4 Methods

3.1.4.1 Participants

Overall, we had 23 participants taking part in study 5a. One participant had to be excluded from data analysis because of severe simulator sickness. From an additional four participants a few sessions had to be excluded because of technical problems while recording, so that in total there were 18 full datasets and four partial datasets. The gender ratio was balanced with 11 female and 11 male participants. With a minimum age of 18 years and a maximum age of 43 years we had an age range of 25 years, an average age of 27.64 (SD=6.5). Participants had an average possession of their driving license of 9.83 years (SD=6.92) and an average mileage of 10,472 km/year, 43% of which were driven on highway. People wearing glasses were excluded. From the 22 participants, four wore contact lenses during the study.

Study 5b was conducted with 9 participants. Three participants could not complete the study due to simulator sickness. Two of the remaining 6 participants were female and 4 male and the average age was 76.5 years with a range from 70 to 81 years (SD=4.51). The average possession of driving license was 58.83 years (SD=6.59) and all of them stated to drive multiple times a week. Participants indicated to drive on average 12,283 km/year, 27% on highways. To make sure participants could perceive the relevant aspects of the iHMIs and perform the driving task properly, questions about impairments regarding hearing, seeing or mobility were asked. Half of them were wearing glasses, none of them were wearing hearing devices and 5 out of 6 stated to not have hearing impairments.

The study was approved by the Virtual Vehicle Review Board (nr 2024-06) and adhered to the HEIDI ethical principles according to D5.1.

3.1.4.2 Simulator

The hardware and software setup used in this study has been described in detail in D2.3 – Section 5. A shorter description is reported here for convenience of the reader. The Simulator system, the HMI system, and a Data Acquisition system constitute the main components of the setup. The virtual vehicle, including the HUD, Dashboard, LED, Android-based tablet, sound system and simulated scenarios and pedestrians are controlled by the simulator setup, while the data Acquisition system works independently and is based on DataBeam data acquisition platform developed in-house. The HMI system is the core of the whole setup, as it receives and sends signals to all other components. In addition, the HMI system hosts the real-time driver's distraction detection algorithm AttenD [12].

The driving simulator consists of a vehicle cabin from VI-Grade, a 220-degree 5m-wide screen in front of the cabin, steering wheel with torque feedback, shakers providing engine and road vibrations. Figure 3-8 shows an outside view of the simulator.

Figure 3-8: Outside view of the driving simulator setup for study 5

Inside the cabin, the simulator is equipped with a monitor to show the dashboard (instrument cluster), and a surround sound system for environmental and engine sound. The SmartEye

Pro eye tracking system was mounted in front of the driver on either side of the steering wheel to track the driver's gaze and determine the attentive state in real-time. The HUD is simulated within the SCANeR Studio simulation. Depending on the type of driver and environment input, an icon is displayed on the windshield in front of the driver in the area that corresponds to the location of a real HUD. A horizontal strip of 173 LED is located at the bottom of the windshield and used to indicate pedestrians' position on the road. A 10" Android tablet is mounted in the area corresponding to an infotainment system to display a custom mobile application to serve as distracting task or to duplicate the HUD icon for a distracted driver (study 5a only). To emulate spatial (left/right) warning sounds in the different iHMI alertness levels, a smaller USB speaker system, comprised of separate left and right speakers, is also fitted to the simulator interiors and connected to the HMI system. All these elements are visible in Figure 3-9.

Figure 3-9: Driver's view of the driving simulator setup

3.1.4.3 Simulated environment

The simulated environment was created in the SCANeR Studio software and consisted of a two-lane, single carriageway urban road (Figure 3-9). The main road followed a roughly elliptical path about 5 km long, with no sharp curves (Figure 3-10). Drivers only drove in the outer elliptical road. Other vehicles were only present in the opposite direction as oncoming traffic. The speed limit was 50 km/h, except for some sections where construction sites were part of the track, with a speed limit of 30 km/h. Throughout the whole track the street was followed by a sidewalk which was consistently filled with pedestrians walking on it. They only appeared up to 150m in front of the car and disappeared up to 10m behind the driver. 50% of the simulated pedestrians were walking in the same direction as the driver. Pedestrians only crossed the street in the specific crossing events defined in Table 3-1. There were three categories of pedestrians, which differed in their crossing speed. Children were the fastest when crossing the street with 2m/s, regular adult were crossing with 1.5m/s and elderly pedestrians were crossing the slowest with 0.9m/s. One drive consisted of driving one round along the main road and took approximately 6 minutes. There were five versions of this track with a permutated order of the 14 crossing events. Each crossing event was assigned to a road segment in permutated order for each participant. When the car entered a segment, the secondary task was triggered, depending on the HMI condition (see 3.1.4.4 *Experimental design* and the description of the secondary task in 3.1.4.5.1*. Procedure*). Two of these track versions had a clockwise driving direction and three had a counterclockwise driving direction. In Figure 3-10, the superimposed thick blue line indicates the start and end points of the different track versions, while the thinner lines indicate the road segments in which the 14 events were displaced. Finally, several obstacles and construction sites were implemented to increase the driving task demand. In some of the construction sites, a lane change was required at low speed.

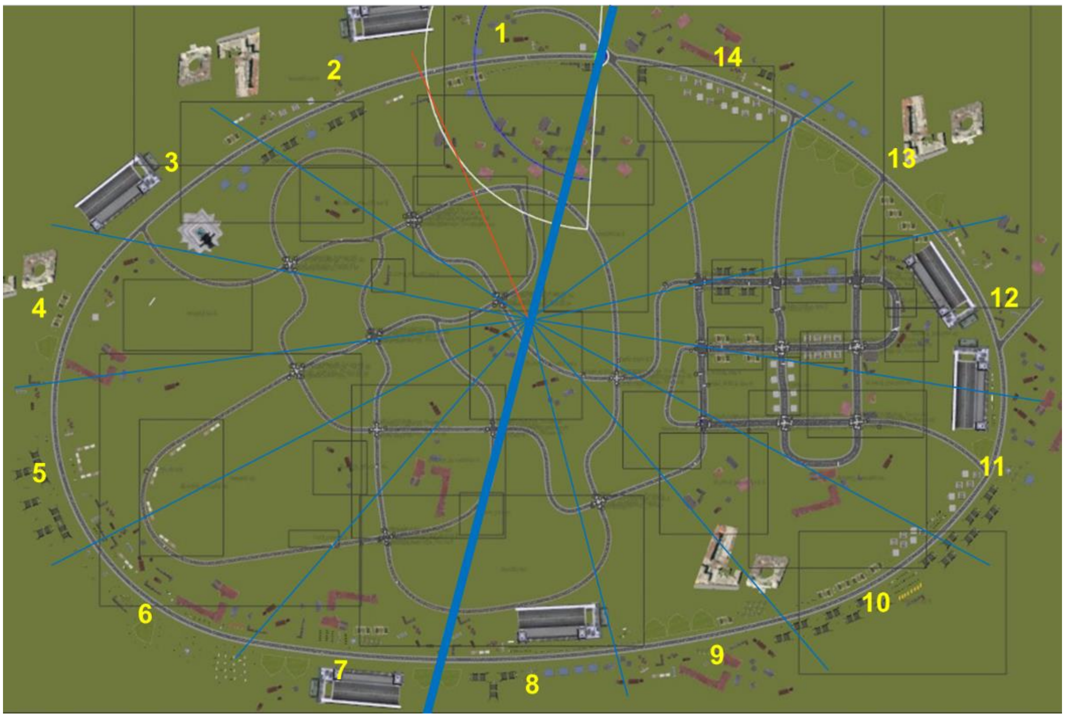


Figure 3-10: Track layout with segments and crossing events (1-14) for study 5

3.1.4.4 Experimental design

For both parts of study 5 – 5a for regular and distracted drivers and 5b for older drivers – a counterbalanced within-subject design was applied. Therefore, all participants were exposed to all the experimental conditions planned in the respective study part.

Study 5a had two independent variables:

- **HMI support** with 5 conditions: (i) *Baseline* with no HMI; (ii) *Non-fluid* with the Non-fluid iHMI for regular drivers; (iii) *Horizontal* with the Horizontal fluid iHMI for regular/distracted drivers; (iv) *Vertical* with the Vertical fluid iHMI for regular drivers; (v) *Full* with the Fully fluid iHMI for regular/distracted drivers.

  1. **Distraction** with two conditions: (i) *Attentive* with driver in attentive detected state; (ii) *Distracted* with driver in distracted detected state.

The distracted state was induced in 50% of the 14 crossing events. The order of the conditions was counterbalanced in both independent variables. Therefore, due to the within-subject design, each participant drove all the five HMI support conditions in a different order, approaching half of the events in a distracted state (also in a different order).

Due to age-related conditions, elderly participants are more prone to develop fatigue, discomfort and sickness in a driving simulator setup, as explained in D1.1 Description of user needs”. Therefore, following the rational explained in D2.3, it was decided to not induce distraction in older participants, and to provide them with a reduced number of conditions, resulting in a remarkably shorter study and minimized demand.

Study 5b consisted of the independent variable **HMI support** only, with three conditions: (i) *Baseline* with no HMI; (ii) *Vertical Older* with the Vertical fluid iHMI for older drivers; (iii) *Vertical Regular* with the Vertical fluid iHMI for regular drivers.

3.1.4.5 Procedures

The procedure of this study was slightly different for regular and older drivers. Therefore, detailed descriptions will be provided in separate sections.

3.1.4.5.1 Study 5a – regular drivers

Figure 3-11: Procedure for study 5a – regular drivers provides an overview of the study 5a procedure.

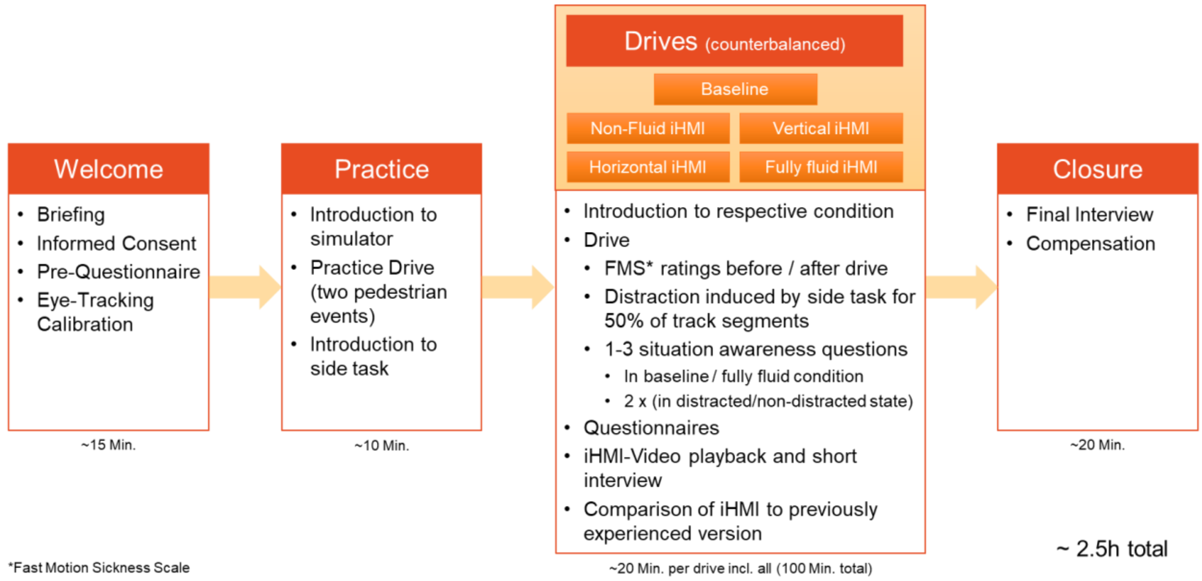


Figure 3-11: Procedure for study 5a – regular drivers

Participants were welcomed with a briefing about the purpose of the study and an explanation of the study procedure. They signed the informed consent and filled in a few questionnaires about demographic information. The eye-tracker was calibrated, and the simulator seat adjusted.

Participants were then introduced to the simulator and performed a practice session to familiarize themselves with the system, scenarios and crossing events. Then, they were introduced to the HMI as well as to the secondary task.

Regarding the HMI, participants were told that their vehicle is equipped with a system to detect crosswalks in advance and recognize pedestrians who are about to cross the road ahead in the proximity of their vehicle. It was explained that the system will provide hints and recommendations if a crosswalk or pedestrian is detected. They were informed that these hints can be visual, acoustic or a combination of both.

A secondary task to induce visual-cognitive distraction was designed for the participants of study 5a. The task mimics texting while driving. Pre-defined questions would pop up on the tablet mounted in the middle console and needed to be answered by the participant while driving, using the on-screen keyboard (Figure 3-9, bottom right). The questions, randomized for each participant, were either questions that could be answered with a single word (e.g., what is the current season?), or participants would be asked to type a displayed word or fill-in a missing word in a logical sequence (e.g., *"kindergarten, …, middle school"*, see Figure 3-12). When a question was answered, the next question would appear automatically. The secondary task was only triggered in 7 out of 14 track segments (see 3.1.4.3 *Simulated environment*). During non-distracted driving, the tablet screen would remain black.

Figure 3-12: Secondary task to induce distraction in regular drivers (study 5a)

The main task for the experimental session consisted of five drives along the predefined road (see 3.1.4.3 *Simulated environment* for details). Participants were instructed to drive safely, respecting the indicated speed limits and avoiding collisions with pedestrians. Also, they were instructed to promptly perform the secondary task for as long as it was displayed on the central tablet (Figure 3-12), keeping a safe driving throughout. The order of the drives was permutated across participants. Each drive was characterized by a different version of the iHMI, including a drive with no HMI support (see 3.1.4.4 *Experimental design*). Each drive lasted about 6 minutes. Before, during, and at the end of each drive, participants were asked to provide a rating on the fast motion sickness scale (FMS; [13]) to monitor the development of potential simulator sickness.

After each drive, participants were asked to fill in a questionnaire and shortly interviewed about what they liked/disliked of the tested HMI version. A short video of up to 30 seconds was shown after each drive to the participant. The video reproduced a crossing event and the sequence of signals for the specific version of HMI that was used in the drive. This provided all participants with the possibility of carefully evaluating the features of a specific HMI version in a common situation and with the correct driver's behavior, without the workload of the driving task and minimizing memory interference.

After all drives were finished, a semi-structured final interview was conducted. Overall, the study duration was about 2.5 hours, including short breaks in between drives. Participants were compensated for their participation.

3.1.4.5.2 Study 5b – older drivers

Figure 3-13 provides an overview of the study 5a procedure.

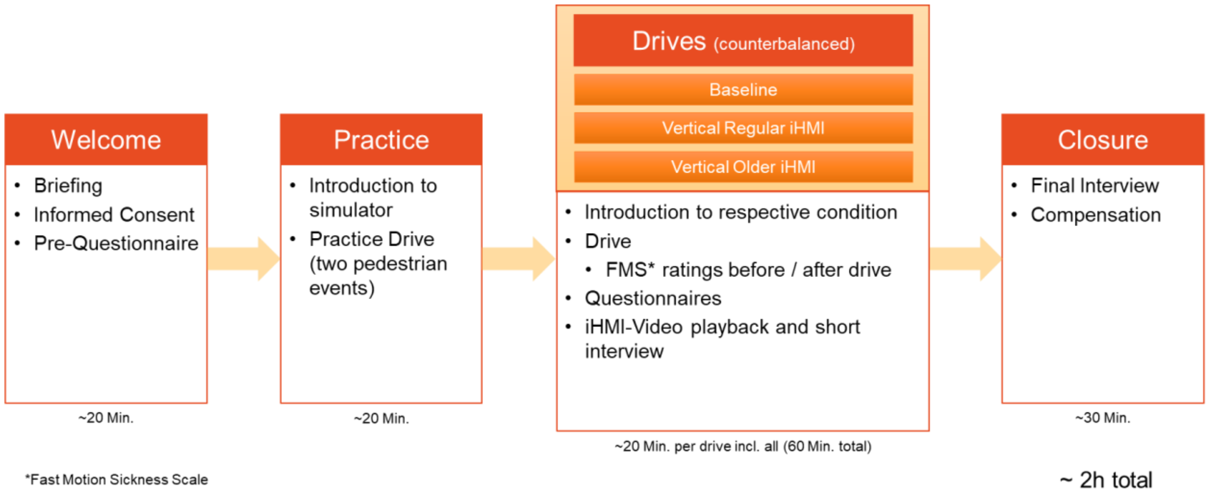


Figure 3-13: Procedure for study 5b – older drivers

Similarly to Study 5a, older participants were welcomed with a briefing about the purpose of the study and a brief explanation of the study procedure. They signed the informed consent, filled in a few questionnaires about demographic information and were accommodated in the simulator seat, which was individually adjusted.

Participants were then introduced to the simulator and performed a practice session to familiarize themselves with the system and the scenarios. Two crossing events were randomly selected to allow participants to familiarize with the situations.

Study 5b participants experienced a different set of HMIs, with more use of auditory messages, including verbal indications and recommendations. Therefore, it was important to ensure that these messages could be properly heard by all participants. The loudness of the sound system was then individually adjusted for each participant before the experimental drives. A sound sample was played, and its loudness manually adjusted according to each participant's feedback, until an appropriate level was reached.

Like in study 5a, participants were told that their vehicle is equipped with a system to detect crosswalks and recognize pedestrians who are about to cross. It was explained that the system will provide hints and recommendations if a crosswalk or pedestrian is detected. They were informed that these hints can be visual, acoustic or a combination of both.

The main task for the experimental session consisted of three drives along the predefined road (the same of study 5a). Participants were instructed to drive safely, respecting the indicated speed limits and avoiding collisions with pedestrians. One drive had no HMI support, whereas two drives were characterized by a different version of the vertical fluid iHMI: one developed for regular drivers, and one developed for older drivers (see *III* and *V* in 3.1.2 *HMI solutions*). The order of the drives was permutated across participants. Each drive lasted about 6 minutes. Before, during, and at the end of each drive, participants were asked to provide a rating on the fast motion sickness scale (FMS; [12]) to monitor the development of potential simulator sickness.

After each drive, the same questionnaires and interviews were carried out for all participants, as described in study 5a, with the support of short videos to be used as reference for their answers. Overall, the study duration was about 2 hours, including short breaks in between drives. Participants were compensated for their participation.

3.2 Study 6 - Small-scale test of eHMI for groups of pedestrians

This chapter provides a comprehensive exploration of the study, beginning with an examination of the use cases that form the foundation of the research. These use cases highlight the practical scenarios and contexts in which the study's findings are applicable. Following this, the chapter delves into the HMI solutions developed to address the identified needs and challenges, detailing their design and functionality.

The research questions section outlines the key inquiries driving the study, setting the stage for the investigation and analysis that follows. The methods section describes the approach and techniques used to conduct the research, ensuring a clear understanding of the processes involved.

The results section presents the findings of the study, supported by relevant data and analysis. This is followed by the conclusion, which summarizes the key insights and discusses the implications of the findings. The chapter concludes with a discussion of the limitations of the study and offers recommendations for future research and practical applications, providing a well-rounded perspective on the study's contributions and areas for further exploration.

3.2.1 Use case

This experiment combines three specific use cases, defined in deliverable "D1.2 Use Case Definition", to evaluate pedestrian-vehicle interactions:

- Use Case 2 - #4: Detection of relevant interaction partners on different roadsides.
- Use Case 2 - #5: Detection of relevant interaction partners on the same roadside.
- Use Case 2 - #6: Vulnerable Road User (VRU) group interaction.

The experiment setup, illustrated in Figure 3-14, involves a pedestrian group consisting of two subjects with one additional virtual pedestrian, that is computer-controlled. One subject wears a Virtual Reality (VR) headset, while the other does not, simulating a distracted pedestrian.

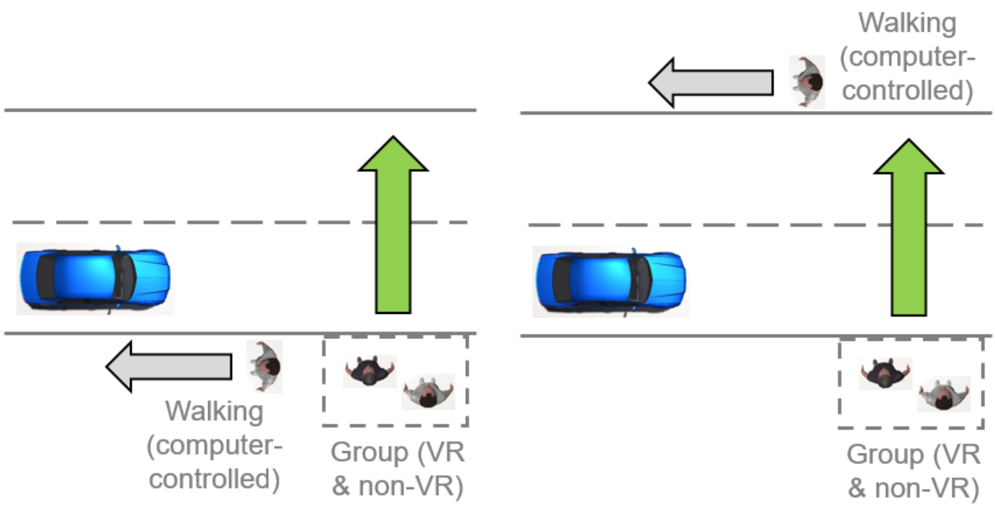


Figure 3-14: Adjusted use case of study 6.

The VR subject experiences a virtual environment, a virtual representation of their own body, and a virtual representation of the other subject's body. Meanwhile, the non-VR subject perceives the physical body of the VR subject and their own physical body but not the virtual environment.

Both subjects share the same motion capture space, allowing for both physical and non-physical interactions. The VR subject acts as the group leader, deciding when to cross the road, while the non-VR subject follows. Communication between the subjects can occur through body language or speech, encouraging natural group behavior.

This setup aims to simulate real-world pedestrian interactions and assess the effectiveness of different HMIs in various scenarios involving different levels of pedestrian awareness and interaction dynamics.

3.2.2 HMI solutions

To assess the effectiveness of the developed eHMIs, participants responded to a vehicle without any eHMI as a baseline, see Figure 3-15. This helps determine if an eHMI offers a beneficial effect in terms of earlier detection of the vehicle's intentions.

This setup was also used during training trials, as it closely resembles a real car and prevents distractions and contamination from an eHMI.

Figure 3-15: HMI 0 – No eHMI.

The Short-Term eHMI is designed to be highly feasible for current production and legislative standards. Its size and placement align with the prototypes outlined in the deliverable "D3.1 Concept – Pedestrian-Car HMI." The eHMI features two separate dot-matrix displays positioned at the front of the vehicle – one on the front left and the other on the front right.

The background color is black, with two colors available for active dots: white and amber. White is used for standard, non-urgent messages, while amber is used for messages with higher urgency. Amber was chosen for urgent messages because it is commonly associated with beacon lights that indicate danger or attract attention.

Figure 3-16: HMI 1 – Short-Term eHMI. (Left to right) message "slowing for pedestrian", "stopped" and "accelerating / not stopping".

The Long-Term eHMI, shown in Figure 3-17, was developed to enable directed communication with individual pedestrians at a higher spatial resolution. Unlike the Short-Term eHMI, this system allows the vehicle to address specific pedestrians on the same side of the road, which is crucial for multi-pedestrian scenarios as defined in deliverable “D1.2 Use Case Definition”.

Figure 3-17: HMI 2 – Long-Term eHMI. (Left to right) message “slowing for pedestrian”, “stopped” and “accelerating / not stopping”.

The Long-Term eHMI comprises three components on the vehicle. First, there is the main message display, a wide, high-resolution screen that wraps around the front of the vehicle. Second, a reference display is positioned above the main display to show an indication marker, signalling that a pedestrian has been detected and is being addressed. Lastly, a reference point is located in the middle of the hood, near the windshield.

When a pedestrian sees the indication marker on the reference display aligned with the reference point from their perspective, they can be confident that they have been detected and the vehicle is communicating with them. If the indication marker and the reference point are not aligned, it means the pedestrian has either not been detected or is not being addressed by the message on the main display, this can be seen in Figure 3-18.

Figure 3-18: Study 6 setup in simulation, with close pedestrian group.

If the relevant pedestrians are close, the display shows the message merged and centered between them. If the relevant pedestrians are further apart the message is shown for both pedestrians individually, as shown in Figure 3-19.

Figure 3-19: Study 6 setup in simulation, with separated pedestrians.

The red reference marker was included in the study, however preliminary tests have shown that subjects intuitively assume the centre of the vehicle as the reference point. The addressed angle would be slightly wrong, however in combination with the message following the movements of the pedestrian, identification still works very well.

3.2.3 Research questions

The first research question investigates whether the HMIs significantly influences the time-to-crossing initiation when used in the experiment. This metric is crucial as it measures the delay or hesitation pedestrians might experience before deciding to cross the road. An increase in time-to-crossing initiation could indicate a higher level of caution or uncertainty introduced by the HMI.

The second research question focuses on whether the HMIs significantly increase the average crossing speed of pedestrians. This metric is important as it reflects the confidence and decisiveness of pedestrians when crossing the road. An increase in average crossing speed could indicate that pedestrians feel more assured and less hesitant when interacting with vehicles equipped with the HMIs. However, it could also mean that pedestrians are forced to run.

The third question examines if the combined time taken by both the vehicle and the pedestrian to resolve the crossing situation is significantly reduced with the use of the HMIs. This combined time is a critical measure of efficiency and safety in pedestrian-vehicle interactions. A reduction in this time would suggest that an HMI facilitates quicker and more effective communication, leading to faster resolution of crossing scenarios.

The fourth question explores whether the HMIs significantly increase the trust pedestrians have in the vehicle during the experiments. Trust is a vital component of human-machine interaction, especially in safety-critical situations like road crossings. An increase in trust would suggest that pedestrians feel more secure and confident in the vehicle's actions and intentions when the HMI is employed.

The final research question assesses whether the HMIs significantly enhance pedestrians' understanding of the vehicle's future actions in the experiment. This understanding is crucial for ensuring safe and predictable interactions between pedestrians and vehicles. An improvement in this area would indicate that the HMI effectively communicates the vehicle's intentions, leading to better-informed decisions by pedestrians.

3.2.4 Methods

This chapter provides a detailed overview of the methodology employed in the study. It begins by introducing the participants and their tasks. This is followed by a description of the simulator utilized. Next, the simulated environment is outlined, detailing the virtual settings and scenarios created for the experiments. The experimental design section explains the structure and rationale behind the study, detailing the various conditions and variables considered.

The chapter then moves on to the procedures, offering a step-by-step account of the experimental process to ensure clarity and replicability. Finally, the validation tools and metrics used to assess the data are discussed, highlighting the methods and criteria for evaluating the study's outcomes. This structured approach provides a clear and thorough understanding of the research methodology, setting the stage for the subsequent analysis and discussion.

3.2.4.1 Participants

The small-scale test was conducted with two groups (i.e., 4 participants) due to practical limitations. Additionally, the technical complexity of the study required significant calibration effort to ensure accurate and reliable data collection. Each session lasted approximately three hours, which, combined with the extensive preparation and postprocessing required, further constrained the number of groups that could be feasibly tested.

VR subjects act as pedestrians intending to cross the street. They decide whether to cross before or after the approaching vehicle, observe the virtual bodies of other pedestrians, and see the eHMI. They also engage in implicit communication with their group partner through body language or speech. Non-VR subjects follow the VR subject and engage in implicit communication with their group partner through body language or speech.

All Participants are presented with a consent form that includes details about the study's content and tasks, the data collected and its usage, potential risks and adverse reactions, the voluntary nature of participation, and the right to withdraw or revoke participation. Participants must sign the consent form to participate.

3.2.4.2 Simulator

The RUAS Co-Simulation combines Virtual Reality (VR) and motion capture with a driving simulator to deliver an immersive experience to the participant. The participant acts as the pedestrian and interacts with the presented virtual environment that contains a road with the approaching vehicle. The vehicle is fitted with one of the described eHMI solutions and displays information to the pedestrian in an interactive fashion. The simulator generates ground truth data for evaluation and analysis of the experiment as well as for training machine learning models.

3.2.4.3 Simulated environment

The environment contains a straight road that is lined with urban buildings. On the side of the road are various props to increase the realism of the scene. The ego-vehicle approaches the interaction point between vehicle and pedestrian(s). At the pedestrian-vehicle interaction point, optionally a crosswalk can be activated. The scene has a clear and sunny sky.

Figure 3-20: General simulation environment (study 6 without crosswalk).

This scenario setup was chosen as it represents all Use Cases relevant for the external HMI (UC2), defined in deliverable “D1.2 Use case definition”.

When approaching vehicle decelerates for the pedestrians it starts at 30 km/h at 30 meters distance. At 30 Meters distance to the pedestrian the external HMI activates, if one is part of the experiment condition. The vehicle starts braking at 15 Meters distance to the pedestrian. At 3 meters distance the vehicle stops. The resulting time-to-collision is displayed in Figure 3-21.

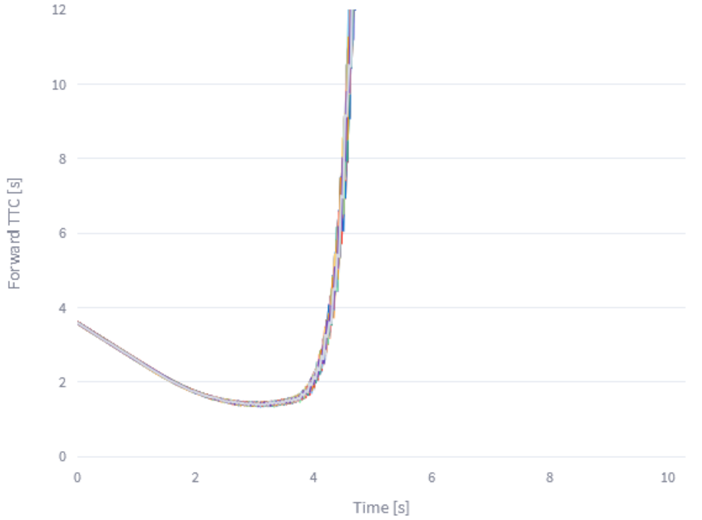


Figure 3-21: Time-to-collision of all experiment trials.

As the vehicles’ braking behavior is computer-controlled, all curves are very similar. When the vehicle stops, the TTC reaches infinity. Afterwards, the vehicle waits until the pedestrian has crossed and then accelerates at a constant rate.

3.2.4.4 Experimental design

Subjects are exposed to repeated eHMI stimuli to observe differences in crossing behavior, using a within-subject design that is counterbalanced, with the entire process lasting approximately three hours. The condition variables include the type of HMI (None, Short-Term, or Long-Term) and the presence of other pedestrians (on the same side or the other side).

A blinder activates before each trial for the VR subject, blocking the side of the road with the approaching vehicle. The blinder deactivates when the VR subject looks at a red dot on the blinder, roughly placed on the relevant approaching vehicle. To prevent early crossings that disregard the relevant vehicle, another vehicle appears directly behind the blinder and starts crossing immediately. The relevant vehicle sometimes does not stop for pedestrians, ensuring the VR pedestrian decides based on vehicle dynamics and potential HMIs. If the VR pedestrian crosses even when the car will not stop, an immediate stop has been added to the car, though this never triggered during trials. When the vehicle stops for the pedestrian, it waits until the pedestrian has left the lane before accelerating again. The trial ends after the relevant car passedthe crossing area and the pedestrian has crossed the road.

3.2.4.5 Procedures

Participants begin by reading and filling out the consent sheet, then dress in the motion capture suit and put on the VR headset. The actor is calibrated in the motion capture system and familiarized with the available MoCap area. They then undergo a training trial with a car (no eHMI), followed by running all conditions and trials (10 with a stopping car, 2 with a non-stopping car). After each trial, participants fill out questionnaires for the current eHMI condition. Finally, they remove the MoCap suit and engage in a final discussion. The study adhered to the HEIDI ethical principles according to D5.1.

3.3 Study 7 - small-scale test of the cooperative HEIDI system

Study 7 was a small-scale test of the cooperative HMI (cHMI). The main objective was to test validation methods for the cooperative HMI system using networked simulators (co-simulation). Various validation methods were tested in the co-simulation setup with a first version of the cooperative HMI system. Tests were performed using a networked VR headset and a fixed-base driving simulator with augmented reality (AR) features. The focus was on the interaction between a human driver and a pedestrian. The results will be used to further refine the validation methods for the evaluation of the final cHMI in study 12.

The methods used in study 7 are described in brief below. A more detailed description can be found in the internal report (Validation study 7 report).

3.3.1 Use case

The connected use case for study 7 is UC3. This UC mainly focuses on the interaction between a human driver and pedestrians. The experiments conducted during this study were used to test various validation methods for evaluating if the cHMI can facilitate cooperation in urban traffic. The experiments of study 7 have the underlying assumption that both interaction partners (ego vehicle and pedestrian) are human agents. In general, the use of the cooperative system is supposed to lead to a quicker and more transparent resolution of ambiguous traffic situations. In study 7, a subset of the experiments outlined for UC3 in D1.2 were tested. All simulator scenarios centered around an un-signalized pedestrian crossing. The selected interaction scenarios were experiment 1, interaction between a regular driver and a regular pedestrian at an unsignalized crossing, experiment 2, interaction between a distracted driver and a regular pedestrian at an unsignalized crossing, and experiment 4, interaction between a distracted driver and a regular pedestrian at an unsignalized crossing, with not-adhering driver. In addition, all experiments were tested with older pedestrians. Driver distraction was introduced in 25% of the scenarios. Adherence or non-adherence to the communicated HMI message was not manipulated in the study design. The drivers and pedestrians were free to act as they would normally do in real traffic. Therefore, experiment 4 had the same experimental design as experiment 2 and adherence was regarded as an outcome of the interaction.

3.3.2 HMI solutions

The cHMI messages were based on the iHMI and eHMI concepts developed in WP2 and WP3 but were modified to fit the cHMI logic developed in WP4. In study 7, the first version of the cHMI was tested. The HMI concept had the following features: The iHMI design consisted of 5 different messages (Figure 3-22). The first message describes the state “Off” when the system is not activated, and no HMI message is displayed. Message 1 and message 3 create a warning that there is a pedestrian nearby and these messages indicate that the recommended behavior implies that the driver should yield to let the pedestrian cross first. Message 1 is meant to be a subtle signal which is activated when the executed driver behavior suffices the recommendation and will lead to the desired outcome of the situation to approve the current behavior. Message 3 is an escalation of message 1 that is more prominent than message 1 and highlights the urgency to adapt the driver’s behavior when the driver does not show the recommended intention. The same reasoning holds for message 2 and 4 with the only difference that they are activated when the driver should cross the interaction zone before the pedestrian.

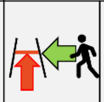
Signal number		0 – Off		1 – Approve		2 – Approve		3 – Inform/Warn		4 – Inform/Warn	
Symbol	Linguistic Interpretation	Empty state, no signal	System is off, no communication		Continue to let the pedestrian pass first		Continue to go		There is a pedestrian, you really should let them cross		There is a pedestrian, you really should go first

Figure 3-22: Symbolic and linguistic representation of the iHMI messages

The design of the eHMI consists of 3 messages (Figure 3-23). Similar to the iHMI design, message 0 describes the “Off” state of the system not displaying any message. Message 1 means that the driver has seen the pedestrian which is indicated by the eyes. Moreover, the chevrons that are pointing downwards indicate that the recommended behavior implies that the driver will brake for the pedestrian. The reasoning behind message 2 is similar to message 1 but referring to the recommendation that the driver will go first and expects the pedestrian to wait for the vehicle to cross.

Signal number		0 – Off		1 – Inform/Warn		2 – Inform/Warn	
Symbol	Linguistic Interpretation	Empty state, no signal	System is off, no communication		The driver is braking to let you cross because s/he wants you to cross		The driver is driving on because s/he thinks you stop

Figure 3-23: Symbolic and linguistic representation of the eHMI messages

The HMI was adaptive and fluid in the sense of communicating different messages depending on the context. The iHMI messages were presented in a head-up display and the eHMI messages were shown on displays on the front of the car (right and left), similar to studies 5 and 6, as shown in Figure 3-24.

Figure 3-24: Screenshot of the iHMI and eHMI

3.3.3 Research questions

Study 7 had the focus on the methods which will be used to validate the system in the larger scale studies at a later stage of the project. On the other hand, the RQs of this study also involve the behavior of the system and the respective reactions of the interaction partners The following RQs were derived with respect to both aspects.

- Which are the best methods to evaluate the benefit of a cooperative system compared to a system without any kind of HMI?
  - Which are the best methods to assess the timing and the effectiveness of i/eHMI communication?
  - Which are the best methods to measure and quantify cooperativity and cooperative behavior?
  - When and how should the cHMI communicate with human drivers efficiently without being distractive or annoying?
  - When and how should the cHMI communicate with pedestrians efficiently without being distractive or annoying?
- Do the validation methods work well for older people with and without cognitive disabilities?

3.3.4 Methods

Two test blocks were performed, one with cHMI and one without cHMI. Each test block had 8 interaction scenarios. The driver was either distracted or not, and the two actors were able to either adhere or not adhere to the communicated message. The study had a cross-over design where the order of test blocks with and without cHMI was counterbalanced between participants. 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. All participants answered questionnaires before, in between and after their test.

3.3.4.1 Participants

Twenty-four participants were recruited for the study, 12 drivers and 12 pedestrians (Table 3-2). Inclusion criteria were having normal or corrected-to-normal vision, agreed on wearing VR-glasses and headset and were available during office hours. Driver participants should drive on a regular basis and belong to the age range of 25-55 years. Pedestrians were divided into two groups, one aged 25-55 years and one group with participants older than 70 years of age. Since instructions and questions were given verbally an ability to hear good enough for conversations was required.

Table 3-2: Participants

	Total	Drivers	Pedestrian
N	24*	12	12
Age; Mean	46.4	41	51.8
Age; Std. Dev.	14.9	8.9	17.9
Gender; Female/Male	13/11	6/6	7/5
Years of education; Mean	16.4	16.8	15.9
Years of education; Std. Dev.	3.0	2.2	3.7

*2 participants did not complete both test-blocks due to nausea, therefore 2 tests were interrupted in advance.

Exclusion criteria for people interested in participating as pedestrians were the need walking aids and susceptibility to motion sickness. People with epilepsy were excluded as a security measure not to risk any seizures when being exposed to the virtual world.

We aspired to have an equal distribution between men and women. The final sample consisted of 13 women and 11 men, four of whom aged above 70 years.

The study was approved by the Swedish Ethical Review Authority (Dnr 2024-00861-01) and adhered to the HEIDI ethical principles according to D5.1.

3.3.4.2 Simulator

An overview of the co-simulation architecture is depicted in Figure 3-25. There were four main software modules: i) Multiplayer server; ii) Pedestrian visualization: iii) Driver visualization; and iv) VTI's driving simulation software. The first three modules were developed based on Unreal Engine 5.2, and they communicated with each other using ENet protocol version 1.3.17, which is a network communication layer implemented on top of User Datagram Protocol (UDP). The driver visualization module also communicated with VTI's driving simulation software, which simulates the ego vehicle. All visualization modules were run on a PC with 10th generation Intel® Core™ i9-processor with 32 GB RAM and equipped with NVIDIA GeForce RTX 3070 Ti.

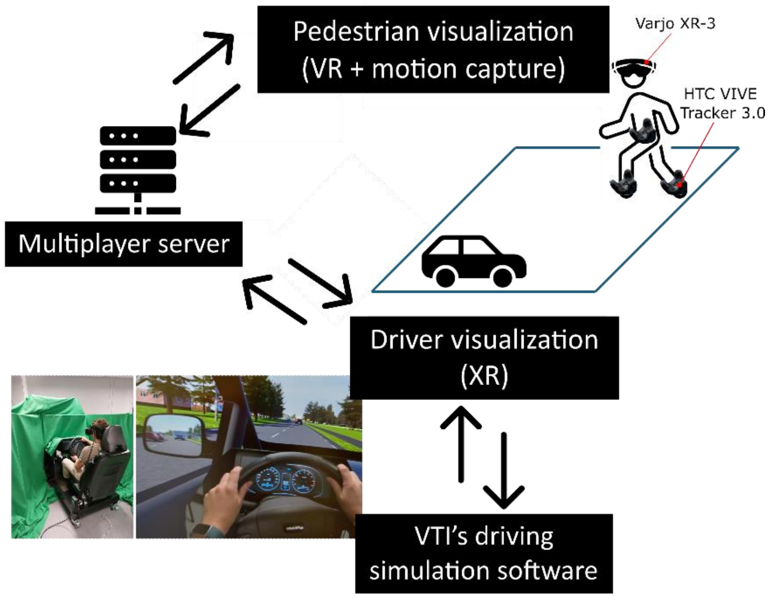


Figure 3-25: Overview of the co-simulation architecture

The VR setup for the pedestrian was a large room where 6.5 m x 4 m were prepared as walking space. The pedestrians were equipped with a Varjo XR-3 headset and three HTC VIVE tracker 3.0, one around the waist and one around each ankle of the pedestrian. LiveLinkXR and OpenXR plugins, which are openly available in Unreal Engine, are used to receive data from the headset and trackers in the pedestrian visualization module. The floor was flat and there was no stepping up or down on the sidewalks. The driving simulator had an XR setup that allowed the drivers to see an actual dashboard and their hands, while the area covered by the green screen was simulated. All participants wore headphones.

3.3.4.3 Simulated environment

The simulator environment had a standard urban design with sidewalks on each side of the road. It had eight 700m segments with the crossing scenario taking place in the last 170m of each segment. A screenshot of the crossing from the driver's and pedestrian's point of view is shown in Figure 3-24. The driving scenario was continuous, relatively straight, and had no hindrance. There was some traffic in the opposite direction in the sections between the crossing scenarios but no other vehicles or road users in the crossing sections. The driver continued to drive straight after each crossing scenario.

All simulator scenarios centered around a pedestrian crossing area. All crossings were un-signalized. For the pedestrians, smaller sections of road environment were created around the pedestrian crossing point. The pedestrians were able to cross back and forth. Road width was 6m in the crossing sections. There were eight crossing scenarios in each test block and the crossing scenarios was considered to start when the car reached a position 50 m from the pedestrian. The section of the road where the pedestrians were able to walk around and cross the road was defined as the interaction zone. The participants were able to see each other's avatars in the simulated environment. The pedestrians could move their heads, bodies and legs in the simulation, but no arm gestures could be used.

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 to the next crossing point along the route.

The driver performed a secondary task during crossing scenario 7 and 8 in each test block to induce visual-cognitive distraction. The purpose was to distract the driver and direct their attention towards a tablet in the center panel of the driving simulator. Circles with arrows in them were presented on the tablet and the driver was instructed to touch the circle with arrows pointing up or to press 'no' if there were no arrows pointing up.

3.3.4.4 Experimental design

The overall experimental design had participant type (driver vs. pedestrian) as a between-subjects factor and cHMI condition (cHMI activated vs. not activated) as a within-subjects factor. The order of the cHMI conditions was permutated for each pair of participants to avoid any sequence effects. The design was, however, different between drivers and pedestrians. For the drivers, the study had a counterbalanced 2x2 within-subjects design with two independent variables, cHMI condition (not activated vs. activated) and driver distraction (not distracted vs. distracted). That means that a driver performed a test block in each cHMI condition, and within a block, participants were asked to perform a secondary task during the last two segments of the drive (25 % of the pedestrian driver interactions) to induce distraction.

3.3.4.5 Procedures

Firstly, the participants were given information about the study, signed the informed consent sheet and completed a background questionnaire. The pedestrians then underwent a short cognitive screening (MoCA).
The tests started with a short training session after equipping the participants with VR headsets and sensors. Participants were informed that the aim of the study was to evaluate a new communication system in cars. In the very first trial, both the driver and pedestrian were naïve to the cHMI and in the second trial, the pedestrian was naïve. This meant that the participants were not given any explanations about the cHMI features and were instructed to interpret freely any symbols that appeared in the head-up display inside the vehicle or on the external displays of the car. After these initial tests it was clear that participants needed instructions to be able to evaluate the HMI. In the rest of the tests participants were given a short instruction about the HMI concept and the meaning of the symbols before the test block with the cHMI. The drivers were instructed that they were going to drive through a city and were requested to drive as they usually do. The pedestrians were informed they were going to be standing at the side of the road, and when rain started pouring down it was their cue to consider crossing the street in a manner they would normally do. The drivers were free to decide whether they wanted to yield to the pedestrian and the pedestrians could decide themselves if they wanted to cross the road before or after the car had passed.

After each test block, participants removed the headset and answered questionnaires about their experiences during the test. Lastly, a brief interview was performed with the driver and pedestrian together. The entire procedure took approximately 2 h for the drivers and 2.5 h for pedestrians.

4. Validation tools and metrics

The validations tools used in the validation studies were a combination of subjective and objective tools. Some of the validation tools were the same across the three validations studies whereas others were tailored for the objective of a specific study. The main differences between studies were the target group (drivers or pedestrians) and whether there was full control of one of the agents' behaviors. Study 5 focused on the experience of the driver, study 6 on pedestrians and study 7 on the interaction between the driver and the pedestrian. Subjective ratings were collected via questionnaires and interviews. Objective data were mainly derived from simulator log data, motion capture movement data, and eye-tracking. For the analysis in study 6, only the VR subject was considered, as only they could see the virtual environment and the vehicle with the eHMI.

4.1.1 Objective measures

Objective data was collected from the simulators. Based on the raw data (vehicle and pedestrian position, speed etc.) logged from the simulator runs, a number of outcome measures were calculated. The measures and their definition are listed in Table 4-1.

In study 5, the following objective measures were considered to describe the driving behavior in correspondence with the crossing events for both regular and older drivers: TTC at brake onset, speed at brake onset, pressure on brake pedal, standard deviation of lateral position, and number of collisions. The TTC at brake onset indicates the estimated time-to-collision when the brake pedal is pressed, in response to an HMI message, with at least 1N (only the last 100 meters were considered). The speed at brake onset refers to the speed of the vehicle when brake pedal is pressed. The (maximum) pressure applied to the brake pedal indicates how strong the reaction is. The measure of lateral behavior is included to consider the stability of the driving control. Finally, the number of collision with pedestrians works as direct indication of the achieved level of safety across the different version of HMI and in comparison to no HMI support. In study 5a only, the measures listed above were also compared between driving in attentive vs. distracted state to investigate how distraction can influence driving behavior and how it can possibly be counteracted by an appropriate HMI support.

In study 6, main objective measures consisted of Time-To-Crossing-Initiation between eHMI types, Time-To-Resolve (combined for pedestrian and vehicle) between eHMI types, and average pedestrian crossing speed between eHMI types, starting when the vehicle reached a distance to the pedestrians of 30m until it had passed the interaction area.

In study 7, additional measures were designed to evaluate the interaction between the driver and the pedestrian in the crossing scenarios. The measures were calculated in the section from 50 m before the pedestrian until the car passed the pedestrian, unless otherwise defined. It was recorded whether the driver yielded to the pedestrian and the pedestrian crossed before the car or if the pedestrian waited until the car had passed before crossing the road. To determine whether the participants adhered to the recommended behavior, adherence was also defined, and the data was labelled accordingly. It was assumed that a driver behaved according to the recommendation if the recommended iHMI message is message 1 or message 3 (the driver should yield to let the pedestrian cross first) and the driver shows yielding behavior latest after a maximum reaction time of 1.5s. Similarly, the driver reaction was assumed to be compliant if the recommendation included message 2 or message 4 (the driver should not yield and pass the pedestrian), and the driver reacted by either accelerating or keeping constant velocity latest after 1.5s reaction time. If the system was turned off (message 0), the recommendation compliance could not be determined as there was no recommended behavior anymore. The corresponding encoding is as follows: -1 means non-compliant reaction, 0 means that the system was turned off and that there was no recommendation, and 1 means a compliant reaction within the maximum reaction time.

Table 4-1: Main objective validation metrics

Measure	Unit	Definition	Study
Start Speed	km/h	The speed of the vehicle when entering the crossing scenario (50 m from the position of the pedestrian)	5, 7
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	5, 7
Distance at Brake	m	The estimated distance from the pedestrian the vehicle was at when the driver presses the brake with a force of at least 1N	5, 7
Standard Deviation of Lateral Position - SD	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	5, 7
Number of HMI communication signals	n	Number of times a i/eHMI message has been triggered	7
Maximum vehicle deceleration	m/s2	The maximum deceleration of the vehicle in the crossing scenario	7
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.	6, 7
Crossing speed	m/s	Average pedestrian crossing speed	7
TTR, time to resolve	s	Time starting when the car is 30 m away from the pedestrian and ends when the car has passed the pedestrian by 30 m	6, 7
Crossing Initiation delay (CID)/ Time-to-crossing-initiation	s	Time difference between vehicle entering the crossing scenario until the pedestrian starts crossing the road	6, 7
Accepted pedestrian crossing time-gap	s	Distance from the car to the pedestrian divided by current speed of the car at the Crossing initiation timestamp	7
Collisions detected	n	Number of times that the vehicle hits a crossing pedestrian	5
Maximum brake pressure	N	Maximum pressure applied to the pedal during braking	5
Crossing behavior	Categories 1/2	1 means the pedestrian crossed after the car had passed, 2 means the pedestrian crossed in front of the car	7
Adherence	Categories -1/0/1	Compliance to the first HMI message where, -1 means non-compliant	7

		reaction, 0 means that the system was turned off and that there was no recommendation, and 1 means a compliant reaction	

4.1.2 Subjective measures

Subjective measures were collected using questionnaires and interviews. The questionnaires used a combination of questions drafted specifically for the validation studies and previously used questions and validated instruments. Background questions included questions about age, gender, driving experience, general health, education etc. Questionnaire instruments used in the validation studies are listed in Table 4-2. In addition, questions about the general experience of the HEIDI HMI have been used. These included questions about how easy/difficult it was to perceive and interpret the cHMI on a scale from 1=very difficult to 6=very easy (study 7).

In study 5, after each drive with an iHMI version a short semi-structured interview was conducted which contained questions about what the participants' general impression of the system was, what they liked/disliked about the system, what they found particularly helpful/not helpful about the system, if they realized the adaptiveness to the attentive/distracted state of the *Horizontal fluid iHMI* and the *Fully fluid iHMI* versions (see II and IV in *3. HMI solutions*) and their opinion thereof, and eventually if they liked the experienced iHMI version more or less than the one experienced before. In the final interview participants compared the iHMI versions, picked their favorite one and explained why. They also were asked which version they found the most comfortable, the safest and the one that provided the best overview of the situation. Participants were invited to provide suggestions for improvement and were asked if and why they could/could not imagine using the iHMI in a real car. Regular drivers additionally received two to three self-constructed questions about their situational awareness. After a predefined event we asked them to stop for a short time to answer the questions verbally. They got asked if there was a crosswalk, if there was a pedestrian who wanted to cross the street and the side of the crossing pedestrian (if any). This happened once after a segment where the driver was attentive, and once in a segment where the driver was distracted by the secondary task for both the baseline and the fully fluid iHMI conditions.

Subjective measures in study 6 included the NASA Task Load Index (TLX) and the System Usability Scale (SUS) questionnaires.

Table 4-2: Subjective validation tools

Tool	Description	Output	Study
Mini-DBQ	Driver Behavior Questionnaire to measure aberrant driver behavior [1]]	Subscale scores for violations, errors and lapses (0-5)	7
ATI	Affinity for Technology Interaction, defined as tendency to actively engage in technology interaction [2]	Personal ATI score (1.0-6.0)	5, 7
EQ-5D-5L	Instrument to describe and value health [3].	Individual items (1-5) and health score (0-100)	5, 7
PBS	The Pedestrian Behavior Scale, a Self-reporting scale to measure	Violations, errors and lapses (1-6)	7

	pedestrians' injury risk behavior [4]		
SUS	The System Usability Scale, for assessing perceived usability [5]	Usability score (0-100)	5, 6, 7
SATI	SHAPE Automation Trust Index [6]	Trust score (0-5)	5, 7
UEQ-short	The user experience questionnaire, a measurement of subjective impressions of users [7]	Overall score and subscale scores for pragmatic and hedonic quality (-3 to 3).	5, 7
CTAM	Car Technology Acceptance Model, measures information technology acceptance in an automotive context [8]	Selected items (1-7)	5, 7
NASA-TLX	NASA Task Load is a multi-dimensional scale to measure workload estimates [9]	Subscale scores for mental-, physical- and temporal demand, performance, effort and frustration (0-21)	6, 7
VTI Acute stress scale	One-item stress scale to evaluate perceived stress [10].	(1-9)	5, 7
KSS	Karolinska sleepiness scale, a questionnaire to evaluate subjective sleepiness [11]	(1-9)	5, 7
DALI	Driving Activity Load Index	Overall score of task demand (effort, visual-, auditory-, temporary- demand, interference, situational stress)	5

In study 7, after each interaction the pedestrian was asked two questions that emerged in the virtual world. The first question was how safe they felt from a scale from 0 to 5 when crossing the road, where 0 represented very unsafe and 5 very safe. After they answered they were asked how certain they were about who had priority in the crossing, where 0 represented very uncertain and 5 very certain. The pedestrian said the answer out loud and the test leader entered the response using a keyboard. Study 7 also used a simulator evaluation questionnaire to get a better understanding of the participants experience of the co-simulation setup.

Semi-structured debriefing interviews were performed with the participants after the simulator tests in study 5 and 7. In study 7 the following questions were asked:

1. How did you experience this interaction/situation?
2. Do you remember if there was something special you thought about/noticed/wondered in the crossing situation (or any other you think of)?
3. Overall, did you perceive and understand the HMI?
4. What is your opinion on the virtual reality environment? Do you think you acted in the same way as you would have in the real world? Can you give an example?
5. Do you have any thoughts on the questionnaires?
6. Is there anything else you want to add or tell us?

5. Analysis

The analyses focused on evaluating the validation methods. Given the relatively small number of participants, the analyses were mainly descriptive.

In study 5, both objective and subjective measures are summarized with descriptive statistics. Selected parameters like, e.g., braking time, speed, and several items of the questionnaire data are analysed using inferential statistic (two-way ANOVA), t-test and Chi-square tests.

For study 6, statistical analysis was not conducted because the study was run with a small sample size.

In study 7, differences between drivers and pedestrians in questionnaire instrument scores were tested with t-tests. Differences in objective measures between crossing scenarios with and without HMI were tested with t-tests and two-way ANOVA with the factors HMI status (HMI triggered vs. no HMI) and crossing behavior (pedestrian crossed before or after the car). Differences between groups in subjective measures were tested with Mann-Whitney U-test and Chi-square tests. Differences within individuals in their experiences of the test blocks with and without cHMI was tested with Wilcoxon signed ranks test. Statistical analyses were conducted using IBM SPSS statistical software version 29.0. The alpha criterion was set to 0.05. No correction was used to compensate for multiple comparisons due to the exploratory nature of study 7.

6. Results

6.1 Study 5

Results about objective driving behavior and subjective impressions are provided for study 5a – regular drivers, and study 5b – older drivers, respectively.

6.1.1 Study 5a – regular drivers

6.1.1.1 Driving behavior

Driving behavior data was recorded during the study including the speed of the vehicle, the position on the track and the pressure on the brake pedal as well if the participants were distracted by non-driving related tasks (see for details).

Since the driving behavior data is in time series format the first step to process the data was to compute the above-mentioned variables per event, condition and participant resulting in a total of 1173 events. After removing extreme outliers and erroneous data a total of 1022 events which were used for analysis. The data of the single events were averaged for each condition and participant using the arithmetic mean to enable comparisons between conditions where each participant has one value for each variable per condition.

To compare the performance of the different HMIs, a two-way ANOVA with repeated measures for each variable was computed with the factors *HMI condition* (5 levels) and *Distraction* (2 levels). Three participants were excluded as data was missing in more than one condition resulting in 20 participants being analysed.

The results show a statistically significant difference in TTC at brake onset for the factor *HMI condition* ($F(76, 4) = 17.932$, $p < .001$) and *distraction level* ($F(19, 1) = 27.810$, $p < .01$) (Figure 6-1).

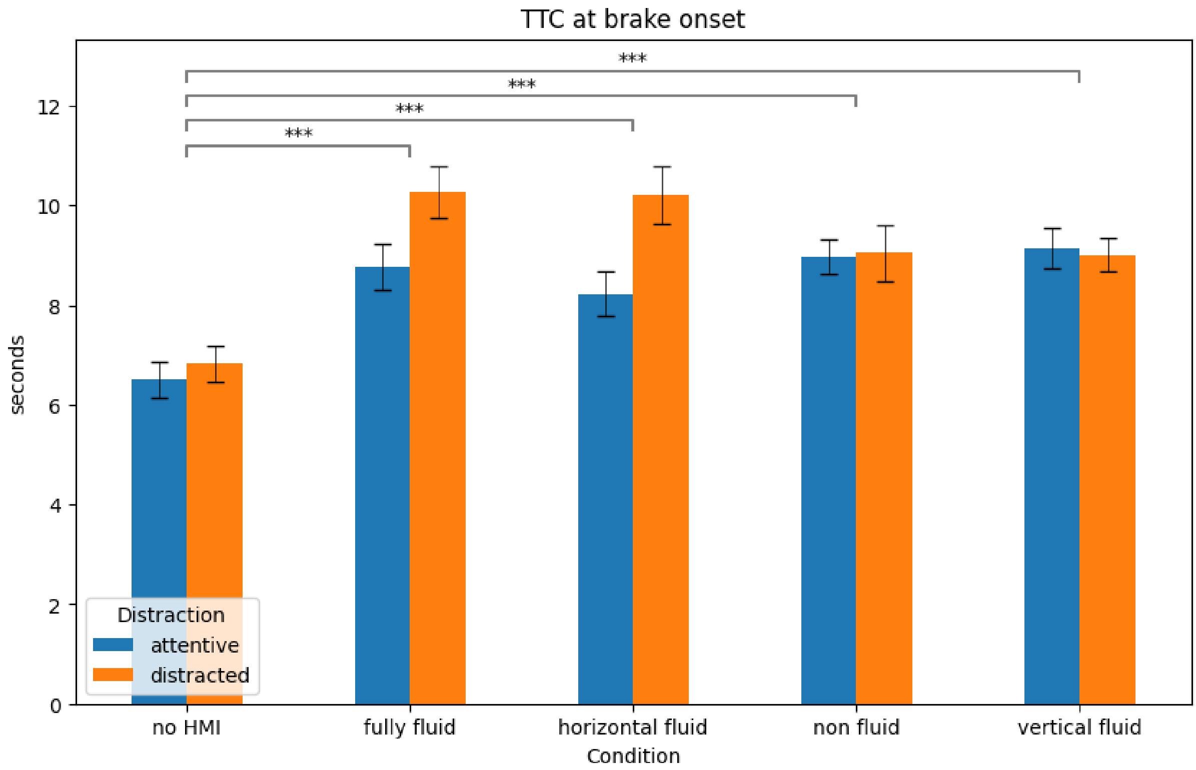


Figure 6-1: TTC at brake onset. *** = p < .001, error bars = standard error of the mean.

Group-wise comparisons show that the participants in the condition without HMI braked statistically significantly later compared to all other HMI conditions. It was also observed that in the fully fluid and horizontal fluid conditions the participants braked statistically significantly later when they were attentive than when they were distracted. This result is in line with predictions, as drivers in the distracted state received an HMI message 1 second earlier than in the attentive state (6s instead of 5s).

Regarding the speed at brake onset, the results suggest a statistically significant difference for the factor *HMI condition* ($F(76, 4) = 6.733$, $p < .001$) and *Distraction* ($F(19, 1) = 13.700$, $p < .01$) (Figure 6-2).

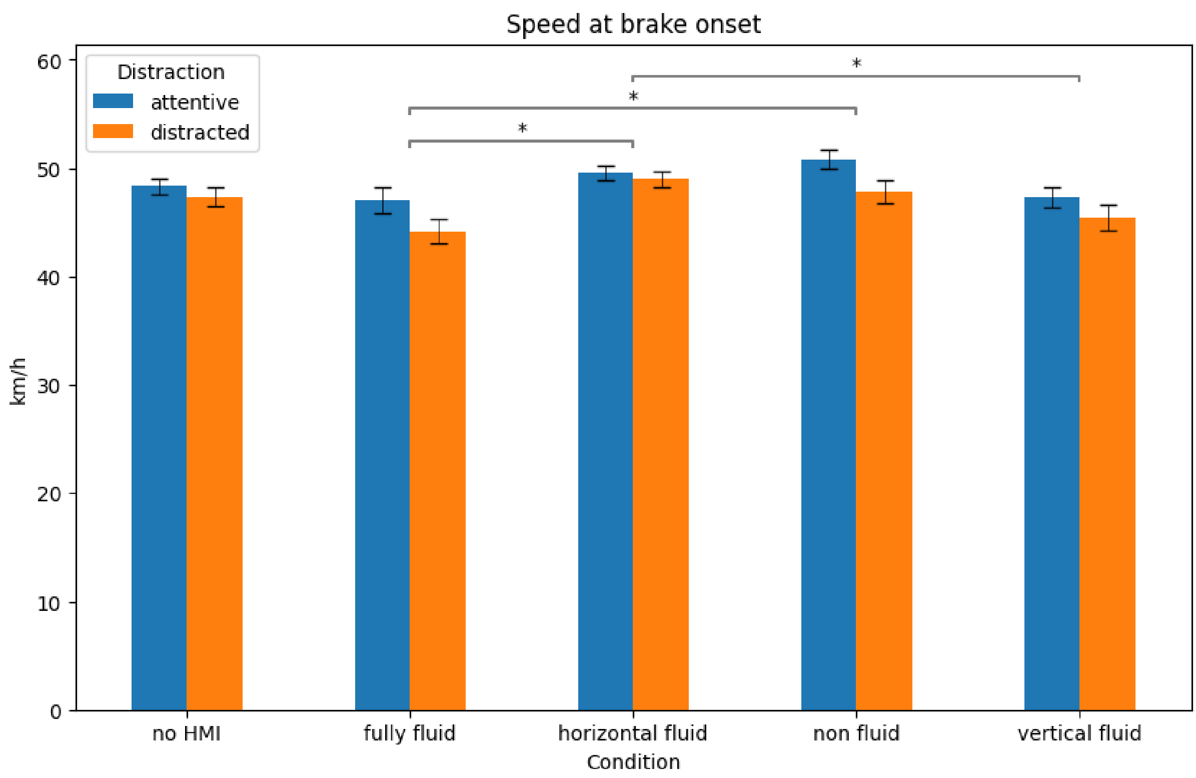


Figure 6-2: Speed at brake onset. * = $p < .05$, error bars = standard error of the mean.

Group-wise comparisons show that the participants in the Fully Fluid-HMI condition were statistically significantly slower at the brake onset compared to Horizontal Fluid-HMI and Non Fluid-HMI condition. Also, in the Vertical Fluid-HMI condition participants were slower at the brake onset than in the Horizontal Fluid-HMI condition. It was also observed that the participants were statistically significantly slower at the brake onset when they were distracted than when they were attentive. This result, consistently with the pressure on brake pedal (Figure 6-3), seems to indicate that distracted drivers are generally less engaged with the vehicle controls.

The results of maximum brake pressure analysis show a statistically significant difference for the factor *HMI condition* ($F(76, 4) = 12.535$, $p < .001$) and *Distraction* ($F(19, 1) = 8.502$, $p < .01$) (Figure 6-3).

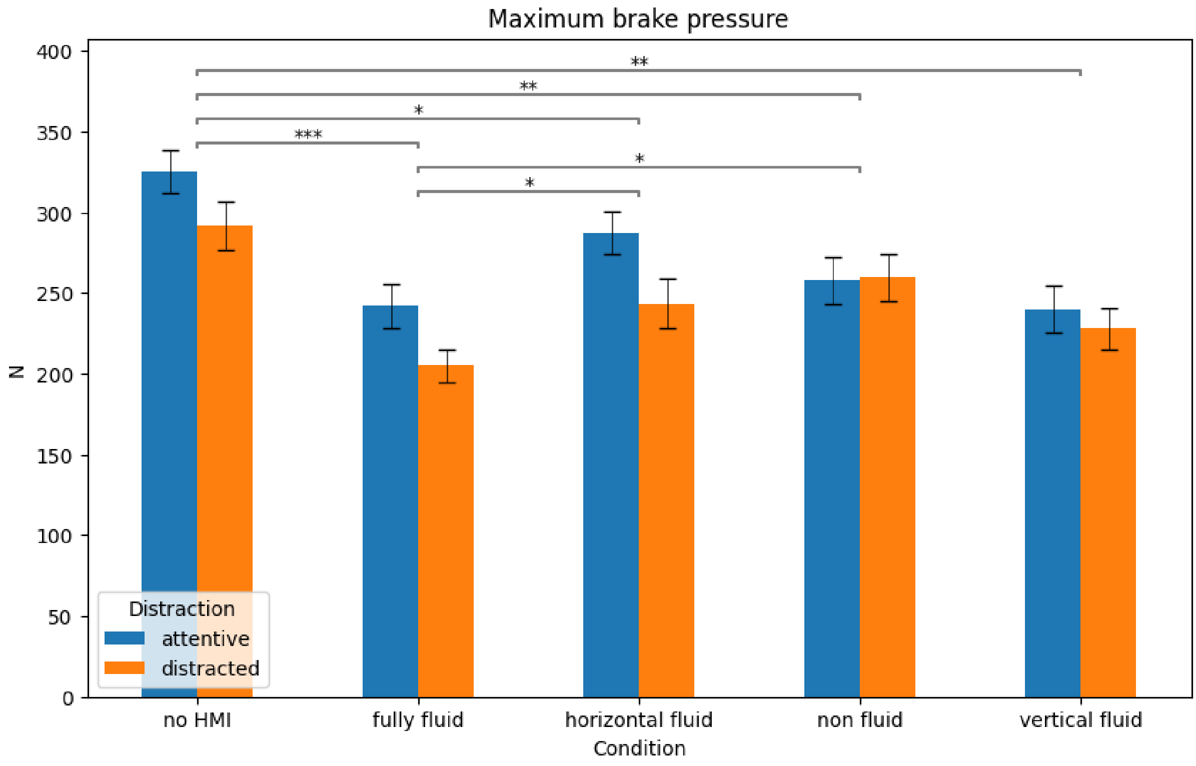


Figure 6-3: Maximum brake pressure. * = p < .05, ** = p < .01, *** = p < .001.

Group-wise comparisons show that the participants in the condition with no HMI braked statistically significantly harder compared to all HMI conditions. Also, in the Fully Fluid-HMI condition participants braked less hard than in the Horizontal Fluid-HMI and Non Fluid-HMI condition. It was also observed that the participants braked statistically significantly harder when they were attentive than when they were distracted. This result can be related to the fact that distracted drivers received the warning 1s earlier than attentive, which yields to approx. 14m additional distance to the pedestrian, prompting drivers to slow down with softer braking, as compared to attentive drivers.

A Chi-Square Goodness of Fit Test was conducted to determine if the distribution of collisions across the HMI conditions deviates from a uniform distribution. The results indicated that the distribution of collisions did not significantly differ from a uniform distribution, X^2 (4, N = 13) = 7.34, p = .119. This suggests that no specific HMI condition was statistically more likely to result in a collision (Figure 6-4). However, it must be observed that the number of collisions occurred without HMI support is more than double than the number occurred with any form of HMI support.

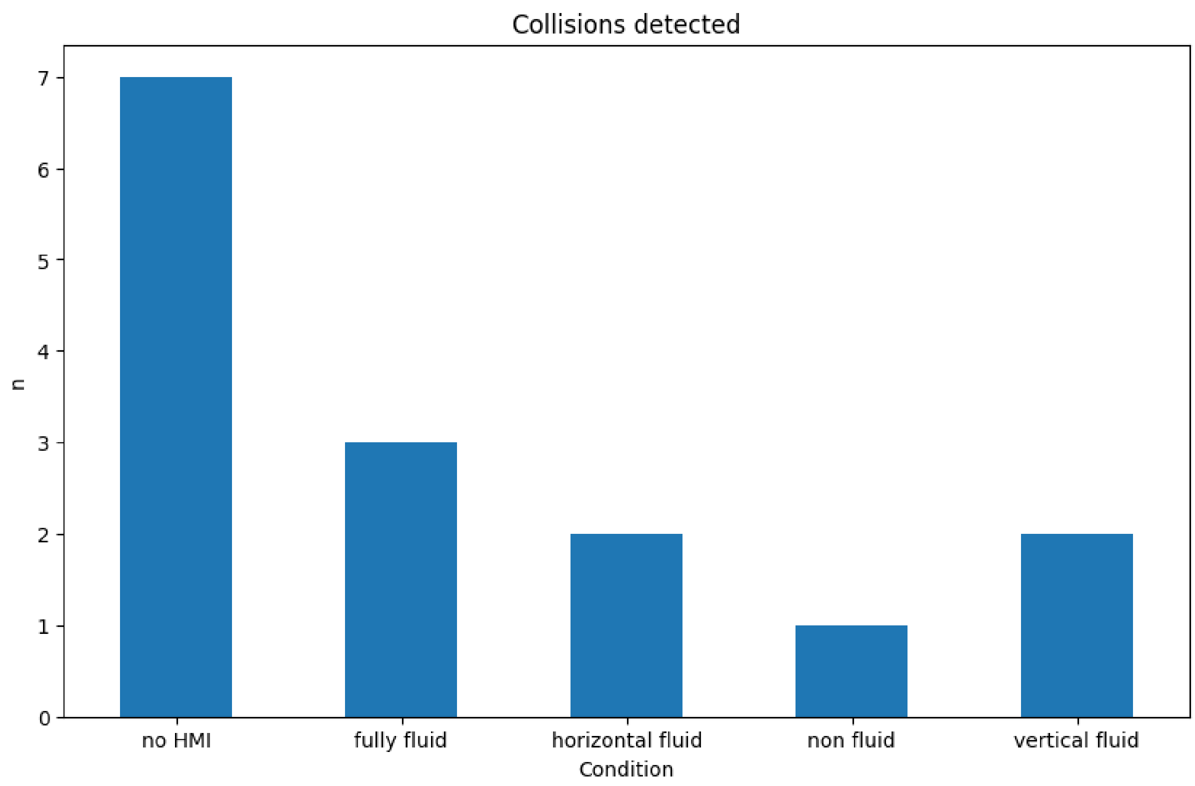


Figure 6-4: Collisions detected in the different HMI conditions.

6.1.1.2 Subjective impressions

6.1.1.2.1 Questionnaire Data

6.1.1.2.1.1 Pre-questionnaire

To get some impressions about participants' attitude towards technology interaction and risk behavior, in the pre-questionnaire we used the Affinity for Technology Interaction Scale (ATI) and the Short scale for risk-taking behavior (R-1).

The ATI consisted of nine items where participants had to rate how much they agree or disagree with specific statements from 1=completely disagree to 6=completely agree.

The R-1 consisted of one item, where participants had to estimate how willing they are to take a risk in general on a scale from 1=not at all willing to take risks to 7=very willing to take risks. Both, ATI and R-1 showed mid-scale values (ATI: mean=3.89, SD=0.32; R-1: mean=3.73, SD=1.61) indicating a moderate propension to risk and affinity with technology.

6.1.1.2.1.2 Questionnaires after all drives

After every drive questionnaires were used to assess driving activity load, perceived distraction, stress and sleepiness. After drives with an iHMI version additionally the system usability, perceived safety, user experience as well as trust in automation were evaluated.

Driving activity load was measured with the Driving Activity Load Index (DALI) which consisted of 6 items where participants rated the driving task regarding its demand in different aspects e.g. visual, auditive, stress, etc. on a scale from 0=low to 5=high (Figure 6-6). The average DALI scores over all items did not differ between the iHMI types. When having a more detailed look on the particular items, it can be seen that for the Baseline drive the General attention demand as well as Visual demand and Stress was perceived slightly higher than in the iHMI conditions (Figure 6-6). Opposing to that, for the Baseline drive the Auditory and Temporal demand were perceived as lower than for the iHMI conditions. The Interference of the secondary task with the driving task was rated nearly the same over all drives.

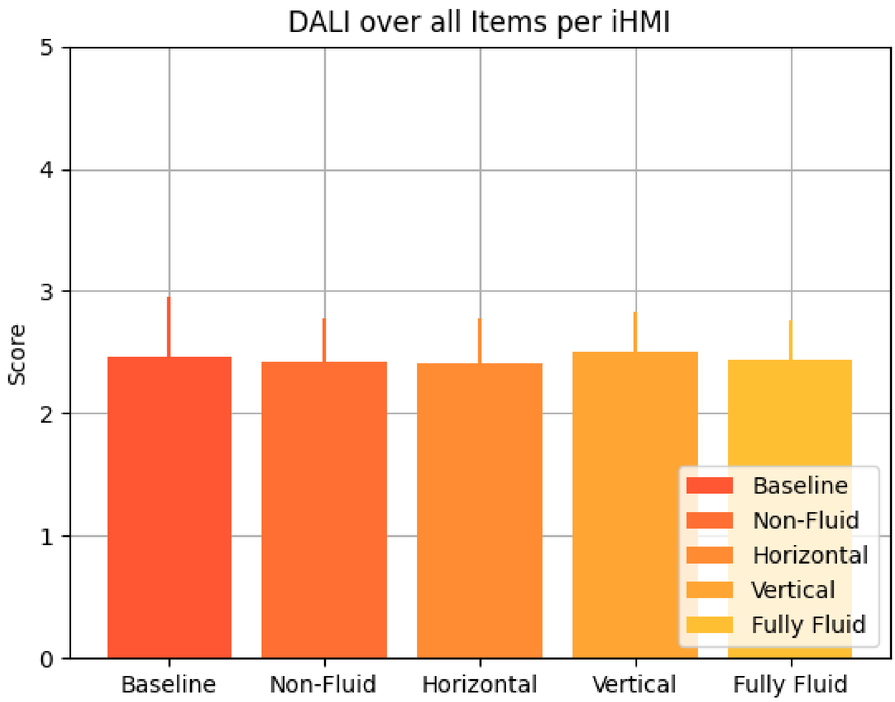


Figure 6-5: Total Driver Activity Load Index per HMI condition

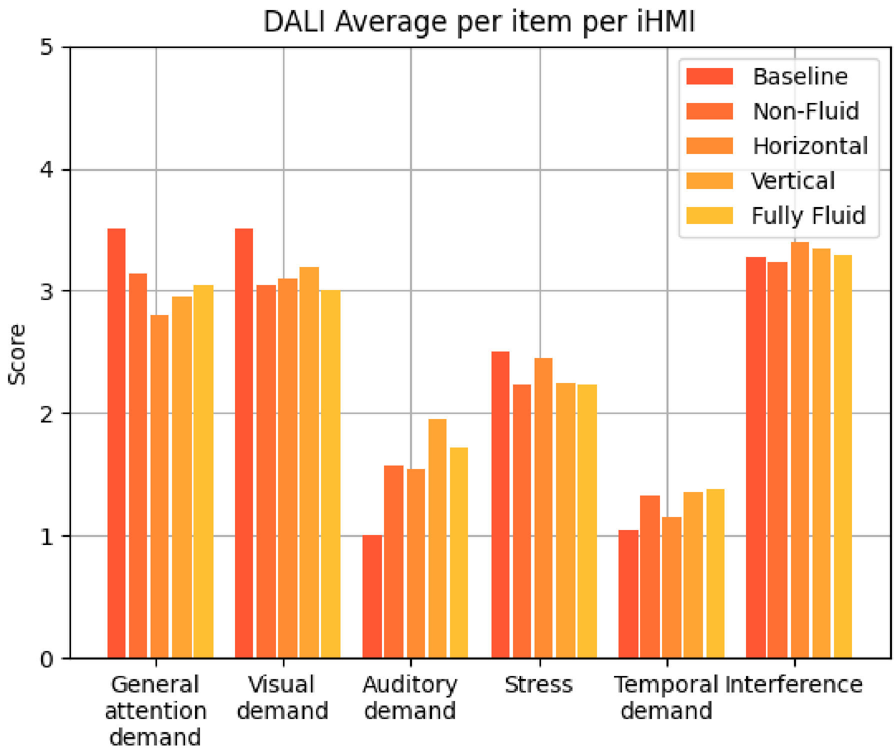


Figure 6-6: Individual Items of Driver Activity Load Index per HMI condition

The perceived distraction was measured with two self-constructed items about how good participants could concentrate on the drive and how distracted they felt in general during the drive. They answered these questions on a scale from 1=absolutely not to 7 totally. Participants felt similarly concentrated during the drives with different iHMI versions, for the Vertical and Fully Fluid iHMIs the distributions of answers were more spread. On average, the participants felt most concentrated during the Fully Fluid condition (Figure 6-7, left). Although the distraction during drive was perceived similarly for the different iHMI versions, for the Baseline and Vertical iHMIs the distribution of answers showed higher variation (Figure 6-7, right).

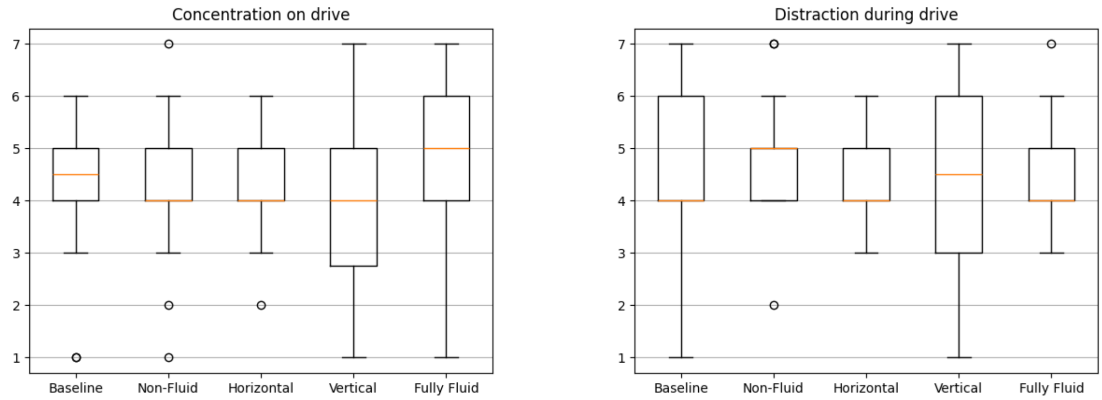


Figure 6-7: Concentration and distraction scores during drive per iHMI condition

Participants were also asked about their perceived stress during the drive by means of the VTI acute stress scale which consisted of one item and could be answered from 1=completely relaxed (feeling entirely calm and relaxed) to 9=extremely stressed (feeling very tense and under high pressure, on the verge of what I can handle). The spread of the perceived stress during the drive did not vary much, the average of perceived stress was rated similarly and in general as moderate, with the lowest average in the Fully Fluid condition (Figure 6-8).

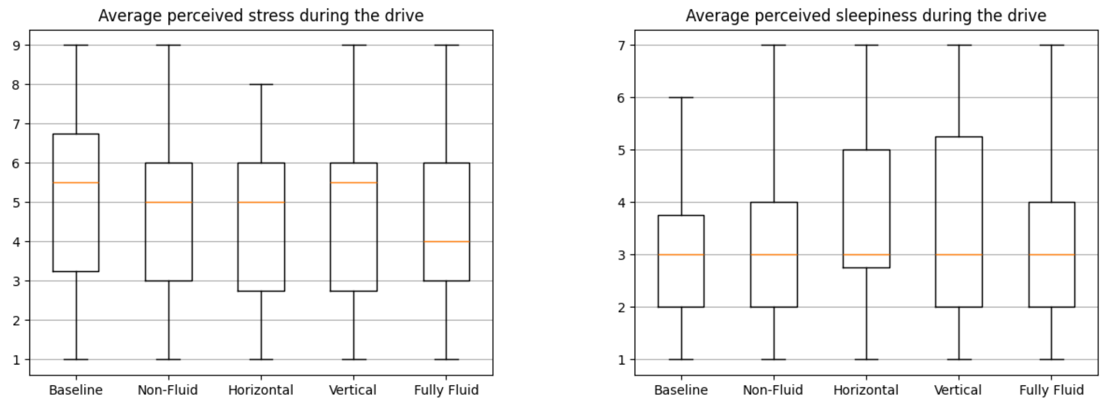


Figure 6-8: Perceived stress and sleepiness per HMI condition.

Also the level of sleepiness during the drives was measured with a single-item scale (KSS). Participants could estimate their fatigue on a scale from 1=extremely alert to 9= very sleepy, great effort to keep awake, fighting sleep. Participants on average felt equally sleepy during all drives. Overall, sleepiness was rated low (Figure 6-8).

6.1.1.2.1.3 Questionnaires about iHMIs

Assessment of usability, acceptance and perceived safety when using the iHMIs was measured by the System Usability Scale (SUS, 10 items), five self-constructed items and the subscale perceived safety from the CTAM (5 items). Participants rated statements from 1=absolutely do not agree to 5=absolutely agree. Regarding the aspects Usability, Acceptance and Perceived Safety, participants rated all iHMI versions approximately equal (Figure 6-9). Acceptance was rated quite high, whereas Usability and Perceived Safety received ratings around the middle sector.

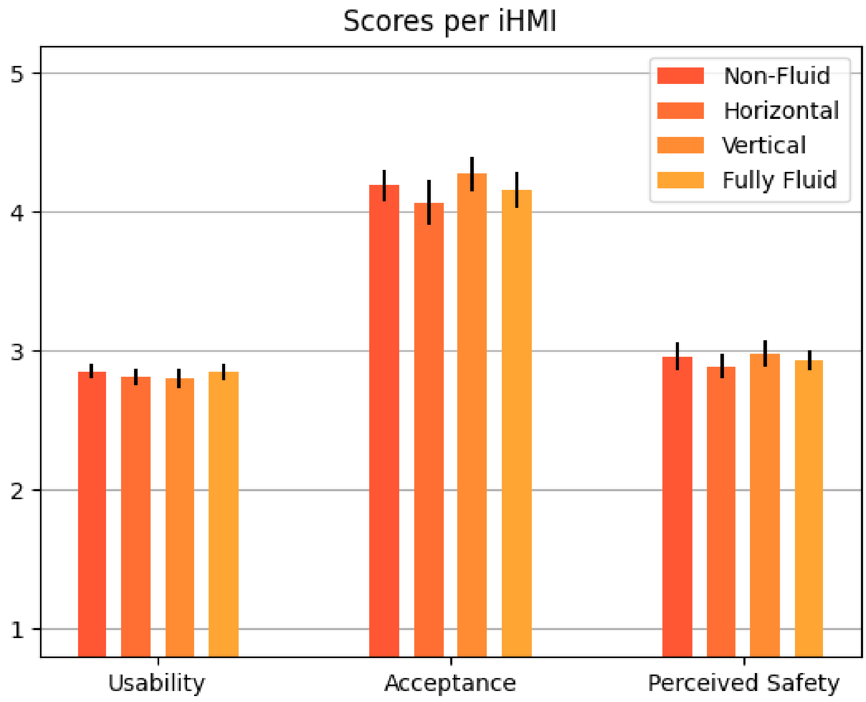


Figure 6-9: Usability, Acceptance and Perceived Safety of iHMIs

To capture the user experience the UEQ-short was used, with 8 items where each item consists of a word pair with opposite meanings. It is divided into the hedonic and the pragmatic subscale and items are scaled from -3 (most negative) to +3 (most positive), 0 represents a neutral answer. The different iHMI versions were rated similarly, with positive ratings in general, where the Pragmatic score was higher than the Hedonic score. It can be highlighted, that in both scales, Pragmatic and Hedonic, the Non-Fluid and Fully Fluid iHMI versions were rated slightly more positively than the Horizontal and Vertical versions (Figure 6-10).

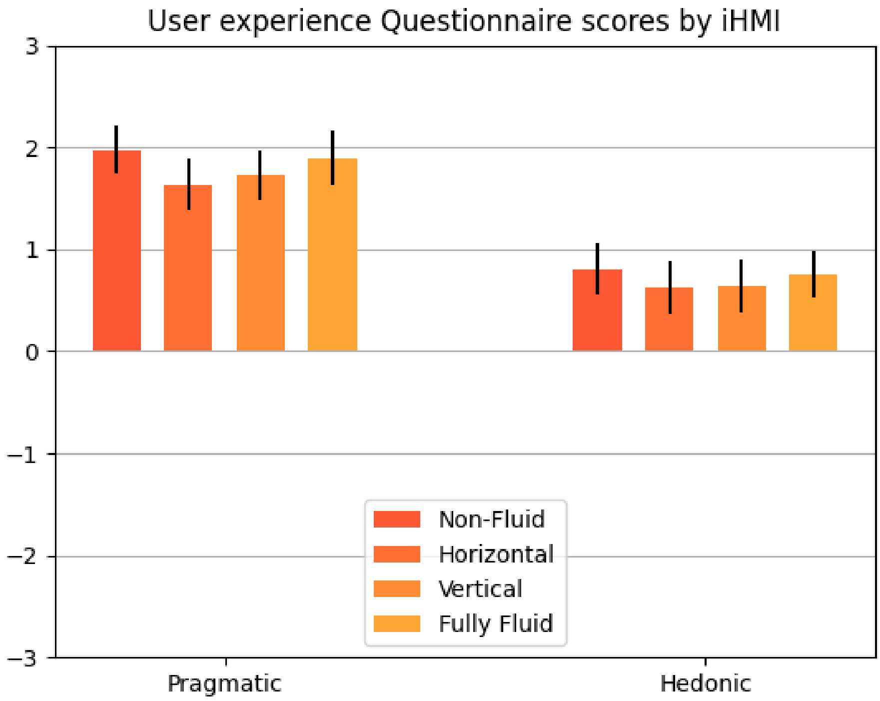


Figure 6-10: User Experience Questionnaire (UEQ-short)

The SHAPE Automation Trust Index (SATI) was used to assess human trust in automated systems. It consisted of 6 statements, where participants rated on a scale from 0=never to 6=always how often they found the statement to be true during the drive. Statements consider usefulness, reliability, accuracy, intelligibility, likability and ease of use. Trust in Automation was rated similarly high for all iHMI Versions (Figure 6-11). When having a more detailed look on the different items, it emerges, that while all iHMIs were rated as useful (Sati1), reliable (Sati2) and comprehensible (Sati4), the Vertical iHMI received the lowest ratings (Figure 6-12). The same applies to the perception of safe usage of the iHMI.

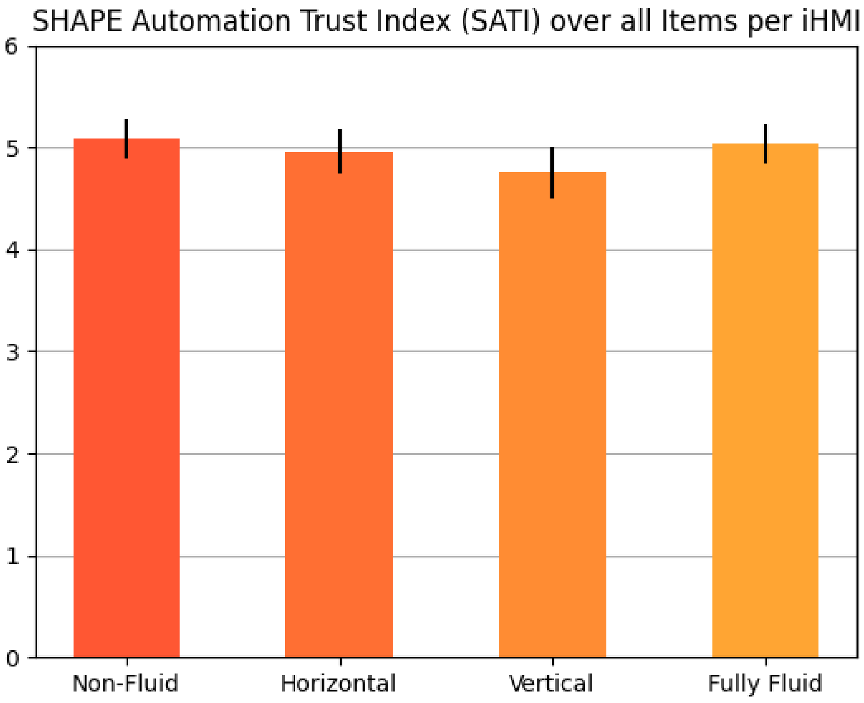


Figure 6-11: Total SHAPE Automation Trust Index per iHMI condition

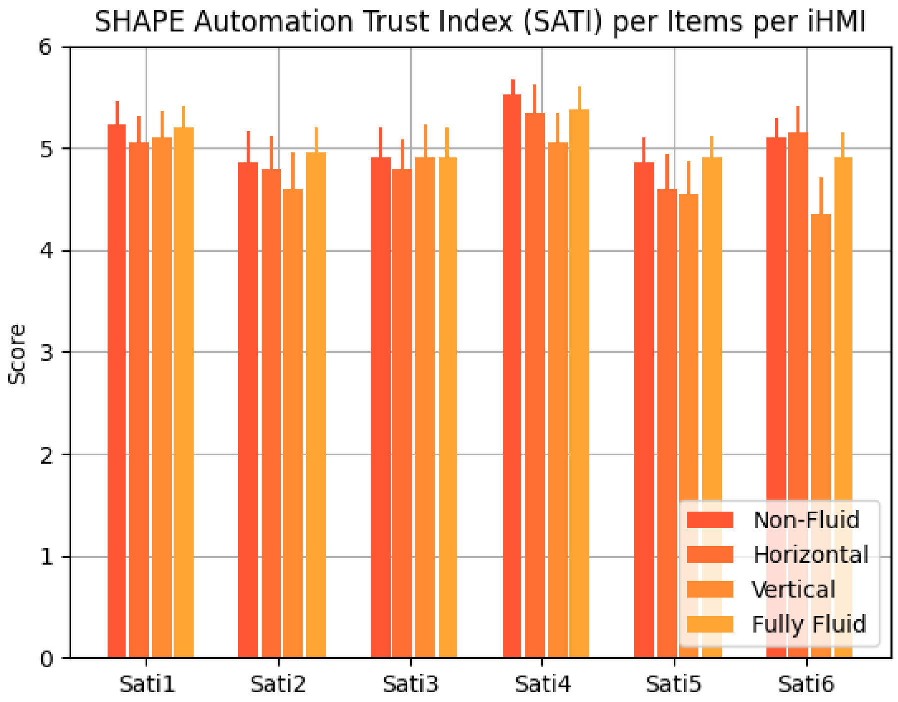


Figure 6-12: Individual Items of SHAPE Automation Trust Index per HMI condition

6.1.1.2.2 Situational Awareness

To assess how aware participants were of their surroundings, particularly of crosswalks and pedestrians, in the drives with baseline and fully fluid iHMI two events were selected to ask participants whether: there was a crosswalk in the previous event; if there was a pedestrian with the intention to cross and; and if yes, from which side the pedestrian was crossing. The events differed between the 5 different tracks and in each drive. One event was supposed to be in an attentive segment (i.e., without secondary task) and one in a distracted segment with secondary task. The events consisted of events with crossing pedestrians on crosswalk, crossing pedestrians offset to the crosswalk, crosswalks without crossing pedestrians and pedestrians crossing without crosswalks, randomized through the track randomization. Shortly after the selected event, participants were asked to stop the car on the road and to answer these 2-3 questions verbally. Participants could reach a score between 0 and 1 with question weights of 0.333 or 0.5 depending on the number of questions for the specific event.

Altogether situation awareness scores were quite high over all conditions and drives (Figure 6-13). For the baseline drive participants reached a higher score in the distracted condition, whereas in the drive with Fully Fluid iHMI version both situations, focused and distracted, show a similar score. When comparing the focused events, the scores in both drives were quite similar, but in distracted events participants scored higher in the baseline drive than in drives with Fully Fluid iHMI version.

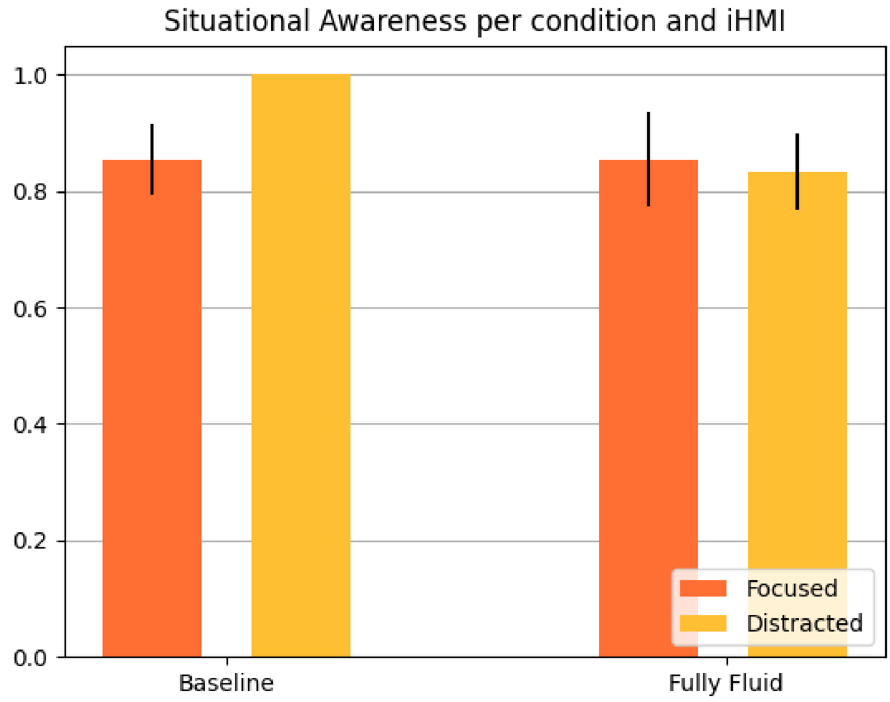


Figure 6-13: Situational awareness in attentive and distracted driving

6.1.1.2.3 Interview Data

After each drive with an iHMI version, a short interview was done as well as a final interview after the full session. The answers were written down on paper and transcribed to Excel. Statements were categorized and described through Pivot tables. Three participants were excluded from the analysis due to motion sickness or technical problems with the iHMI, which resulted in 20 interviews.

For 7 out of 20 people the preferred iHMI version was the Fully Fluid one, for 6 people the Horizontal one, and 4 each preferred the Vertical and Non-fluid ones. Most frequently mentioned reasons for preferences were on the one hand the additional Warning on the Tablet which interrupted the secondary task for Fully Fluid and Horizontal iHMI (8 out of 20), on the other hand the non-locked Tablet screen for Vertical and Non-Fluid Tablet (3 out of 20). A further advantage mentioned across all iHMI versions was the comfortable gradation of warnings, which did not cause sensory overload (10 out of 20). Three times the warning of crosswalks regardless of crossing participants or not was highlighted. Other positive aspects mentioned were the LEDs, perceived safety, adaptivity for distraction and reliability in terms of lowest number of false warnings. In general, participants reported that their preferred iHMI version also was the most comfortable, safest and the one which provided the best overview over the situation. The timing of warnings mostly was perceived as appropriate. When asking if participants could imagine using such a system as the preferred iHMI version in a real car, 15 answered with yes and 6 with no. Reasons for using it would be that it supports when a driver is distracted (5 out of 20), that it improves safety (8 out of 20) and that it is helpful when the accuracy of warnings is high (3 out of 20). Reasons why participants would not use it were sensory overload because of the high frequency of warnings and signals (3 out of 20) and a false sense of safety because of relying too much on the system (3 out of 20).

After each drive with an iHMI version we asked specific questions about the respective iHMI version. For the Fully Fluid iHMI version, 6 out of 22 highlighted the audio warnings to be especially helpful, another 6 found the additional warning and locking of the secondary task screen helpful, 7 liked the crosswalk warnings without pedestrians, 6 mentioned the LEDs, and 3 the warning style in general. Reasons for the highlighted elements were that they think it improves safety, gives more time to react to dangerous situations and it forces a driver to be focused.

For the Horizontal iHMI version similar reasons were mentioned. Nine out of 22 mentioned the audio warnings as especially helpful, 5 the LEDs in general, 6 the warning on the secondary task tablet and blocking of task, also the early warning and a comfortable design of warning levels was mentioned once each. Reasons for that were the exact same as for the Fully Fluid iHMI version.

The opinions about the Non-Fluid iHMI version differed in the aspect of a non-locked secondary task screen, which 4 out of 22 found helpful. Also 9 participants mentioned the audio as positive aspect, 4 liked all warnings in general, 6 liked the aspects that indicated from which direction a pedestrian came and other reasons which were mentioned once each were that no warning came when a crosswalk without crossing pedestrian happened, a comfortable level of warnings and the early warnings.

About the Vertical iHMI again most liked the audio warnings (10 out of 22), 11 liked the elements that indicated from which side the pedestrian is going to cross, another 3 liked the warnings about crosswalks without pedestrians, 3 liked the “Emergency Stop” Warning and 2 mentioned the comfortable warning level. Reasons for mentioning these positive aspects were again the same as in the other iHMI versions.

For the question what the participants disliked about the specific iHMI version for the Fully Fluid version 6 out of 22 mentioned that they disliked nothing, 5 disliked that the secondary task screen was locked during warnings, 4 mentioned inappropriate warnings, 4 disliked the crosswalk warning when no one was crossing and each mentioned once were overstimulation with signals, insensibility of other pedestrians and detection of hidden pedestrians was not realistic.

Aspects of the Horizontal iHMI participants disliked were the locked screen of the secondary task tablet (4 out of 22), inappropriate warnings (4 out of 22) and overstimulation with signals (4 out of 22). Each once mentioned aspects were too early warnings, symbol is present too long, LEDs in general and the crosswalk warning without pedestrians crossing. Six participants disliked nothing in the design of this iHMI version.

A part of the Non-Fluid iHMI version participants disliked, was that there was no additional warning on the secondary task tablet when being distracted (6 out of 22). Two disliked the missing crosswalk warning when no one was crossing, 3 mentioned the warnings in inappropriate situations and 2 felt overstimulated by the signals. Six participants did not dislike any aspect of the iHMI.

For the Vertical iHMI 7 out of 22 participants shared no aspect they disliked. Three disliked warnings in inappropriate situations, another 3 did not like that there was no additional warning on the secondary task tablet, 2 felt overstimulated by the signals, 3 found it irritating that a crosswalk warning also came when no one was crossing and reasons mentioned once each were that they disliked relying too much on the system, symbols being present to long, and unrealistically early warnings.

Eventually participants were asked if they noticed the adaptivity of the two iHMI versions Fully Fluid and Horizontal and if they liked the idea of a system acting differently depending on their attention state. During the drive with the Fully Fluid iHMI version 12 did not notice the adaptivity, 10 noticed that the iHMI acted differently. From all 22 which experienced this version, 17 rated adaptivity as helpful in general. Reasons focused on the screen lock of the secondary task tablet and support to react better in dangerous situations. Concerns have been mentioned regarding sensitivity and realistic detection of distraction. Three participants were indifferent in answering this question. Only two participants found the adaptivity negative, because they cannot rely on a system which acts inconsistently.

For the Horizontal iHMI version 10 noticed the adaptivity and 12 did not. Altogether 17 liked the adaptivity, 4 did not like it and one was indifferent about that. Reasons were the same as for the Fully Fluid iHMI.

6.1.2 Study 5b – older drivers

6.1.2.1 Driving behavior

Since the driving behavior data is in time series format the first step to process the data was to compute the above-mentioned variables per event, condition and participant resulting in a total of 196 events. After removing extreme outliers and erroneous data, a total of 184 events were used for further analysis. Further, the data of the single events were averaged for each condition and participant using the arithmetic mean to enable comparisons between conditions where each participant has one value for each variable per condition.

To compare the performance of the different HMIs a one-way ANOVA with repeated measures for each variable was computed. Two participants were excluded as data was missing in more than one condition resulting in six participants being analyzed.

The results suggest no statistically significant difference between the HMI conditions in TTC at brake onset ($F(10, 2) = 1.411$, $p < .289$), speed at brake onset ($F(10, 2) = 0.704$, $p < .517$) or maximum brake pressure ($F(10, 2) = 1.843$, $p < .208$) (Figure 6-14 the y-axis has no units as different data were grouped together).

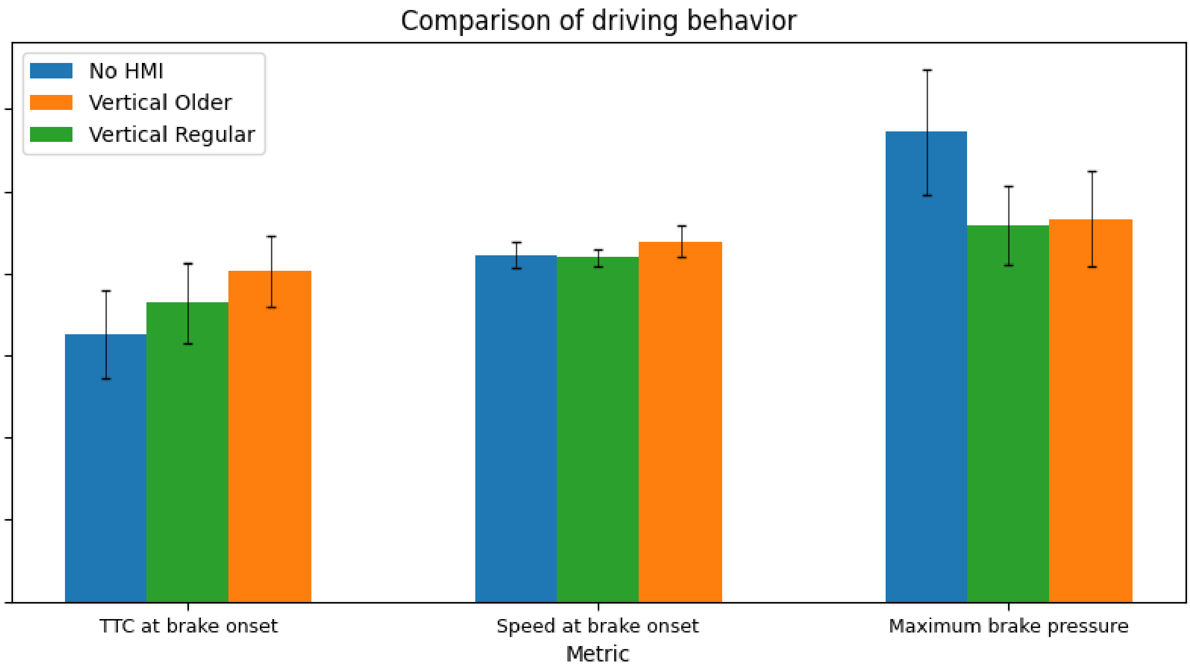


Figure 6-14: Objective measures of driving behavior for elderly drivers

A chi-square test of independence was performed to examine the relation between the HMI conditions and the number of collisions detected. The relation between these variables was not significant, X^2 (2, N = 184) = 1.12, p = .571. No condition was statistically significantly more likely to result in a collision.

6.1.2.2 Subjective impressions

6.1.2.2.1 Questionnaire Data

6.1.2.2.1.1 Pre-questionnaire

In the pre-questionnaire the same questionnaires were used as for the regular drivers (see 6.1.1.2.1.1), which targeted demographic data, affinity for technology interaction (ATI) and risk behavior (R-1). Participants had an average ATI score of 3.69 (SD=0.76) which depict rather mediocre affinity for technology interaction. Risk behavior was with a mean of 3.33 (SD=1.63) on the lower mediocre side. For older drivers additional questions about impairments were asked (e.g. seeing, hearing, mobility) to ensure participants were able to adequately perform the required tasks and perceive the all HMI signals.

6.1.2.2.1.2 Questionnaires after all drives

The same questionnaires as for the regular drivers were used, except for the DALI which was not used for the older drivers. Altogether, perceived distraction, stress (VSS) and sleepiness (KSS) were assessed (see 6.1.1.2.1.2). Participants rated their concentration during all drives nearly maximal (Figure 6-15, left). Conversely, the distraction was rated extremely low for all drives (Figure 6-15, right).

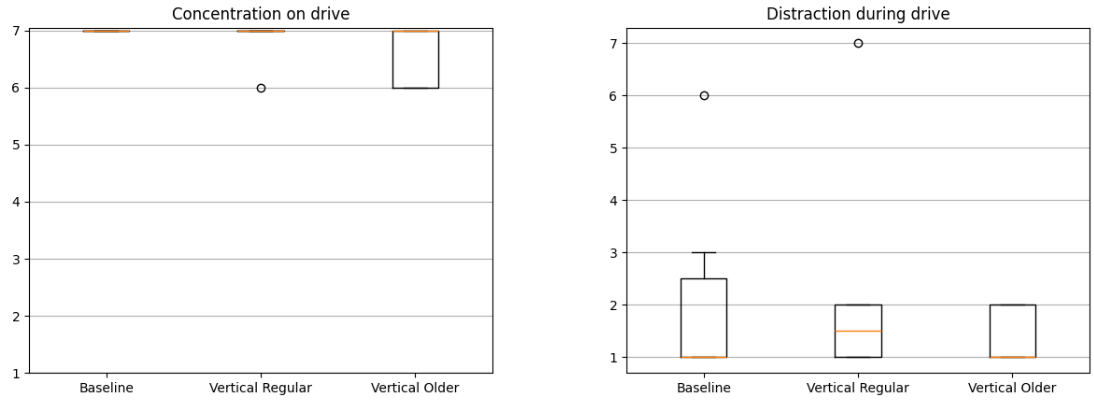


Figure 6-15: Concentration and distraction scores during drive per iHMI condition

Perceived stress was rated as rather low for all drives, with a higher perceived stress for the baseline than the drives with an iHMI version, the Vertical Older version causing the least stress (Figure 6-16, left). Perceived sleepiness was rated low with the Vertical Older iHMI version indicating the least perceived sleepiness (Figure 6-16, right).

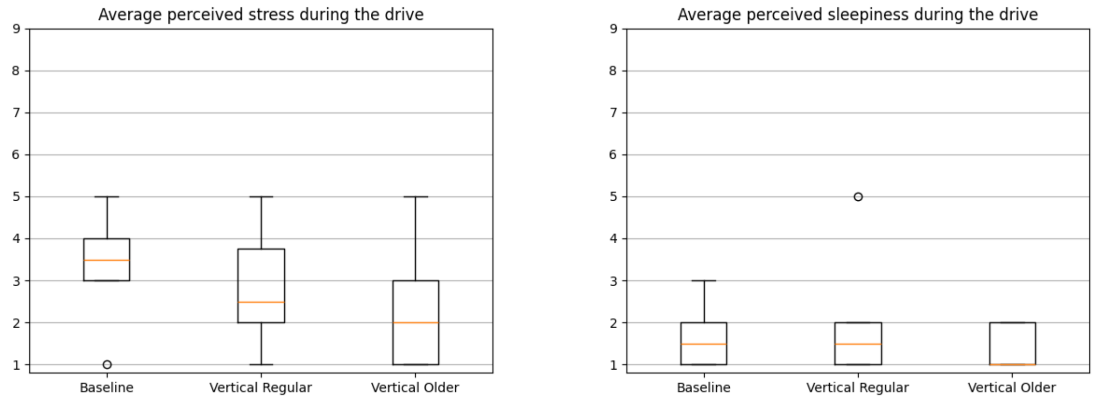


Figure 6-16: Perceived stress and sleepiness per HMI condition.

6.1.2.2.1.3 Questionnaires about iHMIs

To gather the impressions about the iHMIs the older participants got the same questionnaires as the younger ones, particularly the system usability scale (SUS), perceived safety (subscale from CTAM), user experience questionnaire (UEQ-short) and trust in automation scale (SATI). On the Usability scale, both iHMIs reached similar moderate ratings. Also the Perceived Safety was rated rather mediocre, with the Vertical Regular iHMI having a little higher ratings. Both iHMI versions, Vertical Regular and Vertical Older were rated rather high in acceptance with the Vertical Older iHMI receiving higher ratings than the Vertical Regular iHMI (Figure 6-17).

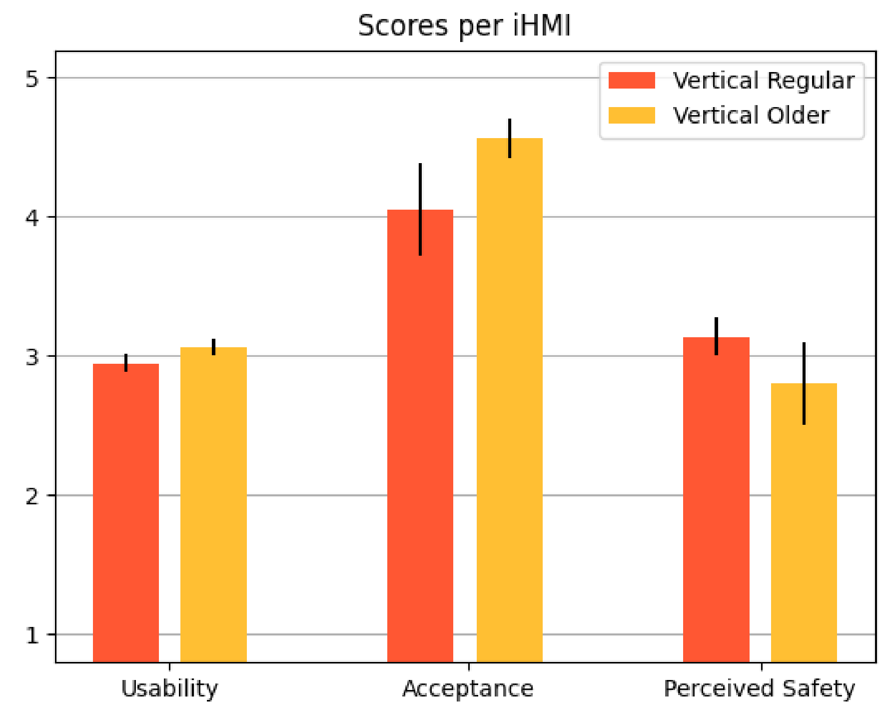


Figure 6-17: Usability acceptance and perceived safety scores.

Both iHMI versions obtained rather high Pragmatic and Hedonic scores in the User Experience Questionnaire, with the Pragmatic scores being higher than the Hedonic scores. Also, the Vertical Regular iHMI received higher ratings for the Pragmatic scores, whereas the Hedonic scores did not quite differ between the iHMI versions (Figure 6-18).

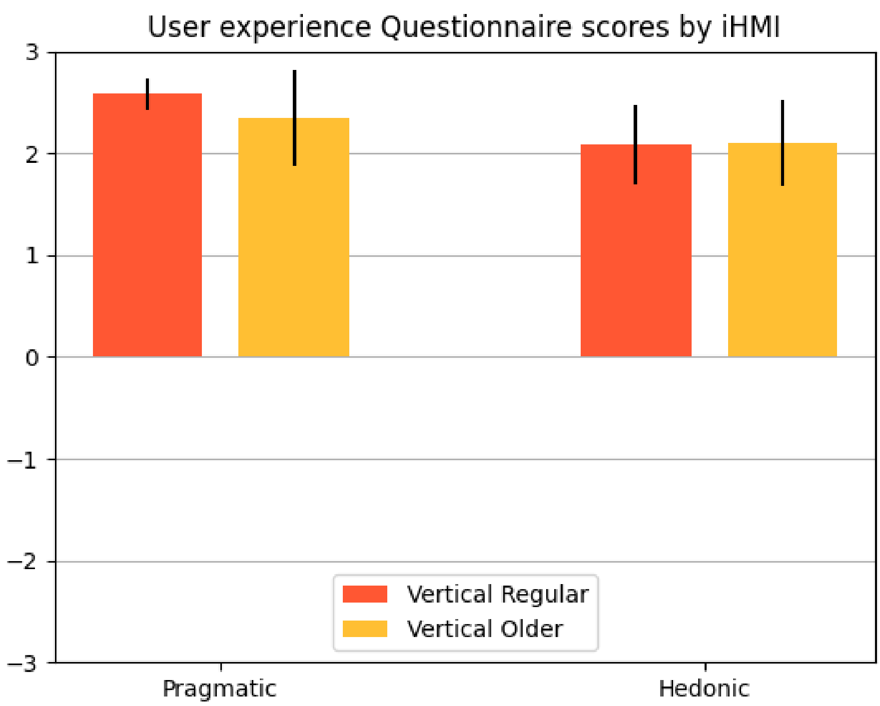


Figure 6-18: User Experience Questionnaire (UEQ-short)

In general participants showed a high level of trust in both iHMI versions, with the Vertical Older iHMI receiving a marginally higher score (Figure 6-19).

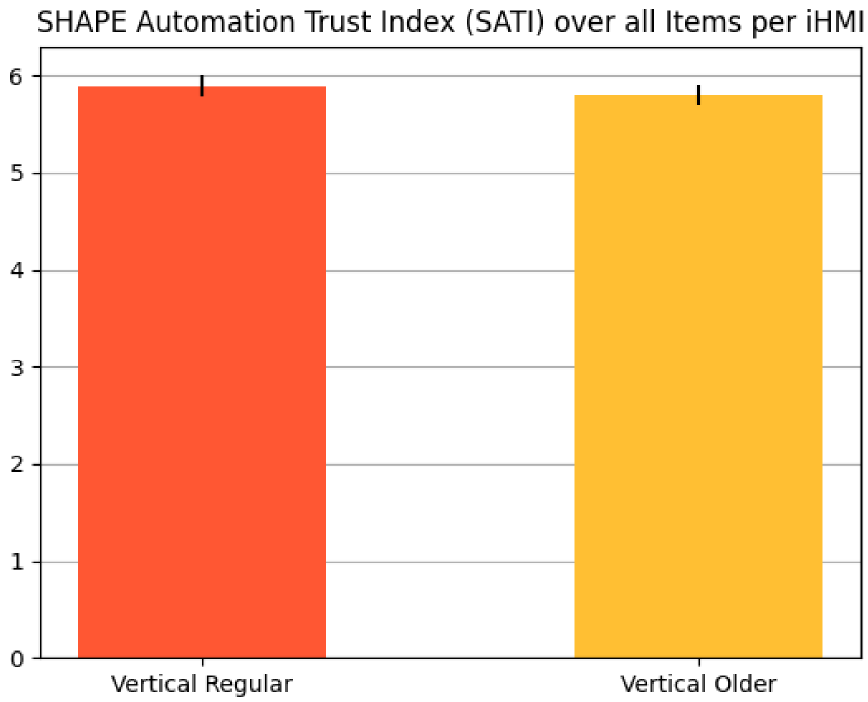


Figure 6-19: SHAPE Automation Trust Index for regular vs. older iHMI

6.1.2.2.2 Interview data

The interview procedure was the same as for the regular drivers. A short interview after each drive with an iHMI version and eventually a final interview was conducted.

In general, 3 participants preferred the Vertical Regular iHMI and 3 preferred the Vertical Older iHMI Version. Reasons for preferring the Vertica Regular iHMI were for all that the version was not overstimulating them and still warning sufficiently. The Vertical Older iHMI version was preferred because the participants felt more addressed with a voice telling them what is happening and they also liked the additional information the voice gave them about the situation. All of the participants could imagine, using such a system in a real car because they felt it increased the safety (4 out of 6). Two mentioned that the system must be adjustable in type of sound, which symbols to use, volume of sound, integration with other systems like navigation or speed limit displays.

Aspects participants liked about the Vertical Older iHMI were that the side of the pedestrian crossing was indicated (3 out of 6), and the crosswalk warning without a crossing pedestrian, the sound and informative content of the voice, and the early warning (once each). One participant indicated to feel completely suspicious about the system and to dislike everything about it; while another one liked everything about it.

About the Vertical Regular iHMI, participants liked the crosswalk warning without a crossing pedestrian most (4 out of 6). Two liked everything about it and one mentioned to prefer the sounds more than the voice.

Aspects the participants did not like about the Vertical Older iHMI were each mentioned once and included the LEDs, which were not found as helpful, the signals being annoying, the speech because it was irritating, the abruptness of the stop signal, and the unrealistic simulator environment. One disliked nothing.

In the Vertical Regular iHMI condition two participants disliked no aspect, 3 mentioned that they did not like the stop signal because it is surprising and abrupt, one did not like the LEDs due to being distractive.

All improvement suggestions were mentioned once and for the Vertical Older iHMI the following topics were addressed: A short text display about the side of the pedestrian would be enough, no voice in general, only visual warnings without any sound or voice, a more gently designed stop signal and adding a function to personalize the warning style and volume. Two participants could not mention anything they would want to be improved.

About the Vertical Regular iHMI participants wanted to improve the design of the stop warning to be more gente, one suggested combining it with the speed limit display, one would want to have only an acoustic signal for the direction of the pedestrian and also a more realistic simulation was requested. Three did not have any recommendations for improvement.

6.1.3 Comparison Regular vs. Older drivers (Vertical fluid iHMI)

As the iHMI version "Vertical fluid" was experienced by regular as well as older drivers, it appears useful to compare this condition between the two age groups regarding the iHMI ratings in terms of questionnaires and interview data.

For the Usability and Perceived safety, older drivers rated the Vertical Fluid iHMI a little higher but both age groups provided moderate scores. Acceptance was rated a little higher in general and younger drivers showed higher scores for Acceptance as older drivers (Figure 6-20).

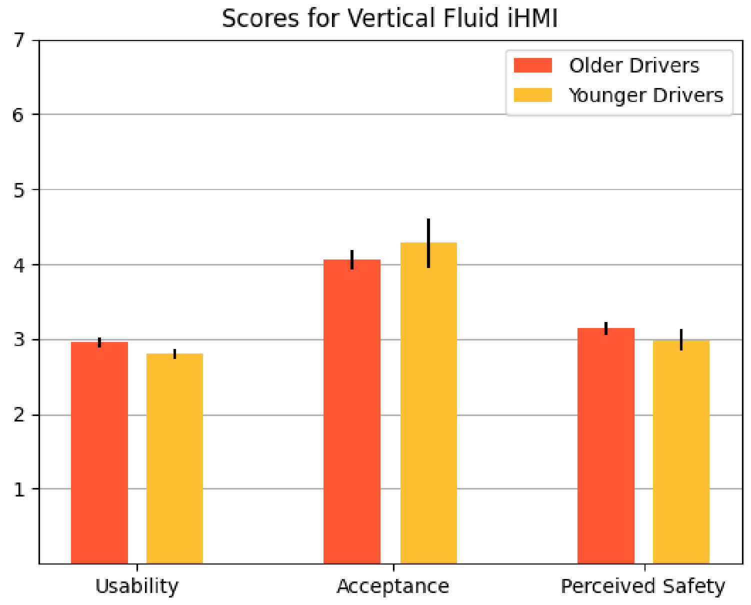


Figure 6-20: Usability, acceptance and perceived safety rating across age groups

In the User experience ratings, for both scales, Pragmatic and Hedonic, older drivers hat higher ratings than youngers and both age groups showed higher values for the Pragmatic than for the Hedonic scale (Figure 6-21).

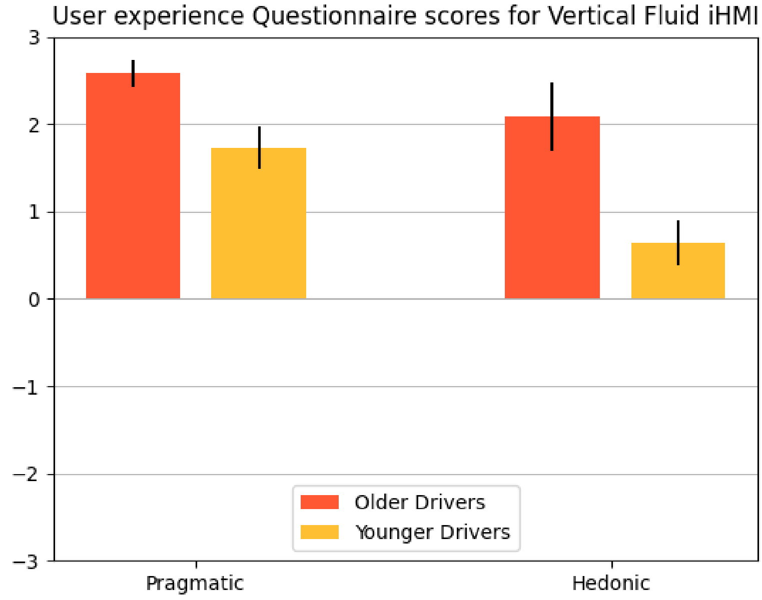


Figure 6-21: User experience rating across age groups

Both age groups seem to have rather high trust in the Vertical Fluid iHMI, with the older drivers reaching nearly the maximum trust level (Figure 6-22).

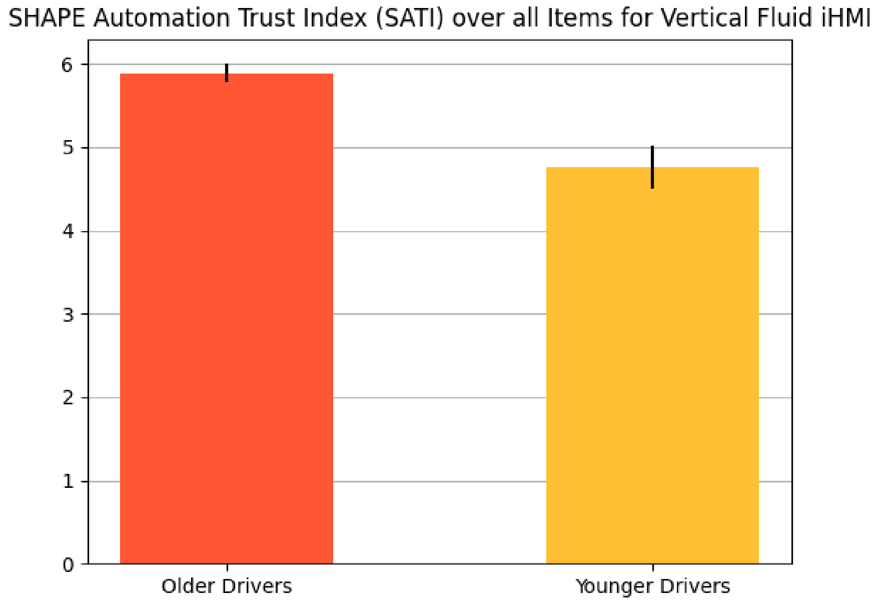


Figure 6-22: Trust index across age groups

6.2 Study 6

This section presents the findings from the study, which evaluated various HMIs in the context of pedestrian and vehicular interactions. The analysis covers several key measures: Time-To-Resolve for both cars and pedestrians, which indicates the duration required to complete a crossing manoeuvre; crossing initiation, capturing the moment pedestrians and vehicles begin their crossing; and average crossing speed for both pedestrians and cars, providing insights into the speed dynamics during the crossing. Additionally, subjective questionnaire measures reflect participants’ perceptions and experiences with the different HMIs. These metrics were assessed to understand the effectiveness and user experience of each HMI, aiming to enhance safety and efficiency in pedestrian-vehicle interactions.

6.2.1 Objective Measures

The following results are grouped by the three different HMIs. The two different condition variables where the computer-controlled pedestrian walked on the same or other side of the road were combined because the objective measures showed no differences between these scenarios.

The combined Time-To-Resolve (TTR) for both pedestrian and vehicle is shown in Figure 6-23. The combined TTR is calculated by starting when the vehicle enters a 30m range from the pedestrian and ends when both the pedestrian and the vehicle have left the intersection area.

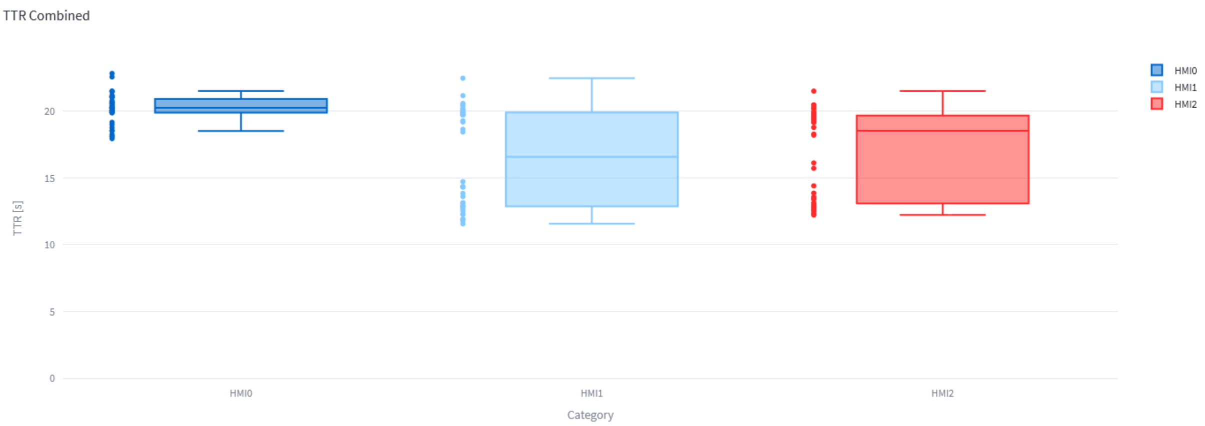


Figure 6-23: Time-To-Resolve of both car and pedestrians between eHMIs.

The crossing initiation time shown in Figure 6-24 is the time between vehicle entering a 30m range from the pedestrian and the pedestrian starting to cross the street.

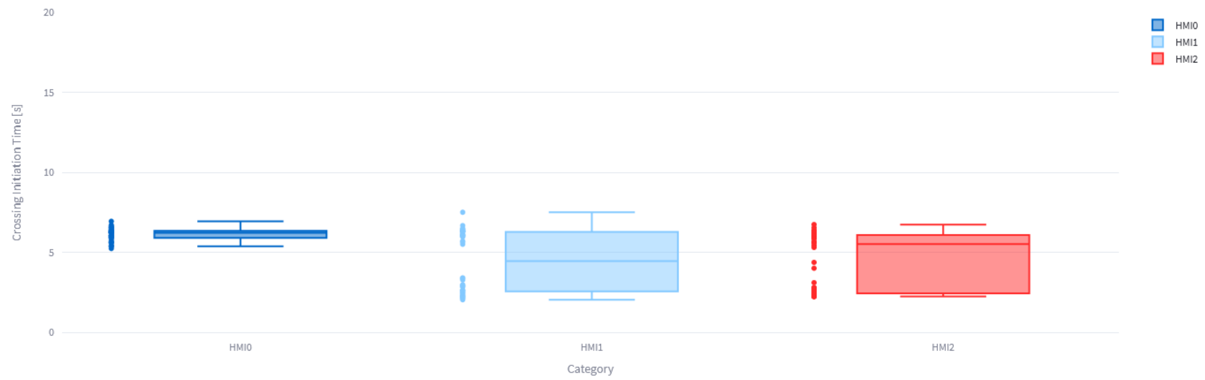


Figure 6-24: Crossing Initiation Time between eHMIs.

No HMI (HMI0) shows overall high Time-To-Resolve (TTR) and Crossing Initiation times. In contrast, both Short-Term HMI (HMI1) and Long-Term HMI (HMI2) sometimes have much lower TTRs and at other times are similar to HMI0. This suggests that the HMIs occasionally resulted in faster crossings. However, further analysis is needed to confirm this and to understand the implications for safety and user experience.

Figure 6-25 shows the distance to the vehicle at the time of the crossing initiation by the pedestrian.

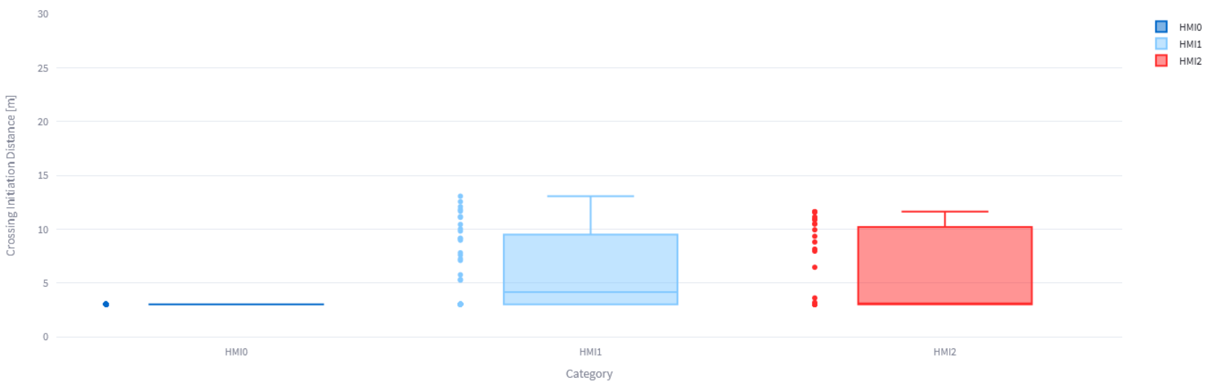


Figure 6-25: Crossing Initiation Distance between eHMIs.

Compared to Short-Term HMI (HMI1) and Long-Term HMI (HMI2), no HMI (HMI0) shows that all groups waited for the car to completely stop before starting to cross the street. While distances were similar between HMI1 and HMI2.

In Figure 6-26 the average pedestrian crossing speed is shown.

Figure 6-26: Average pedestrian crossing speed between eHMIs.

The pedestrian crossing speed stayed about the same for all HMIs, indicating that the HMIs did not result in running or strolling. The faster TTR thus only results from faster crossing initiations.

The average car speed for each eHMI across all experiments is shown in Figure 6-27.

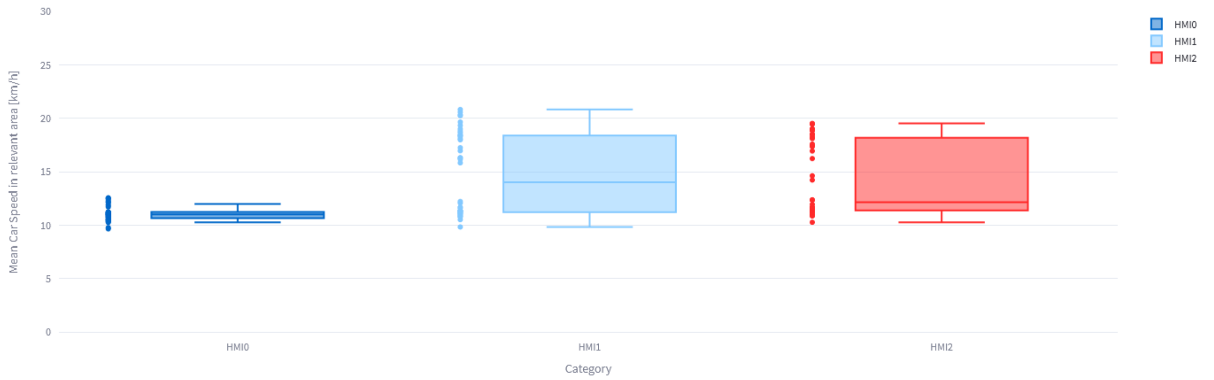


Figure 6-27: Average car speed between eHMIs.

Both eHMIs lead to an increased average car speed due to the pedestrians not waiting until the approaching vehicle has stopped. This indicates that the use of the proposed eHMIs can improve the traffic flow.

6.2.2 Subjective Measures

Only very few datapoints exist for the questionnaires, as such they only give an indication and require more subjects to give meaningful insights. However, the collected data with the NASA Task Load Index (TLX) questionnaire indicates, that mental, physical and temporal demand was very low (always below 20 of 100 points). Performance, Effort and Frustration was also low (always below 35 of 100 points) for both groups.

The System Usability Scale (SUS) indicated that both groups thought that the Short -Term and Long-Term HMI were easy to use and quickly learnable. Both HMI did not require much knowledge or support when using the systems and where not cumbersome to use, while also not introducing inconsistencies. Confidence was high for the Short-Term HMI while it was between high and middle for the Long-Term HMI. The Long-Term HMI was found to be unnecessarily complex when the computer-controlled pedestrian was walking on the other side of the road. Between the groups there was a large difference in whether they would like to use the system frequently.

6.3 Study 7

The main results of study 7 are reported below. More results can be found in the internal report (Validation study 7 report).

Two test blocks were aborted due to motion sickness and one test block was shortened to six interactions for one of the older participants. Both of the aborted test blocks were without HMI. In total, there were 78 crossing scenarios where one or more HMI messages had been triggered and 84 where no HMI was shown.

6.3.1 Objective measures

Simulator measures were analysed with a two-way ANOVA with the factors HMI status and crossing behavior (Table 6-1). Significant effects of HMI status were found for Max deceleration, Time to resolve, CID, and Distance at brake. Max deceleration was lower (harder brake), the distance at brake was shorter (closer to the pedestrian), Time to resolve was longer, and CID was longer when HMI was triggered. Significant effects of crossing behavior were found for Max deceleration, Crossing speed, and Time to resolve. Max deceleration was lower (harder braking), crossing speed was higher, and time to resolve was longer when the pedestrian crossed before the car. Significant interaction effects were found for Max deceleration and Time to resolve. It should be noted that some drivers did not brake at all in the crossing scenarios.

Table 6-1: Results from two-way ANOVA of simulator measures

Measure	HMI status		Crossing behavior		HMI*Crossing		
	F	p	F	p	F	p	df
Start speed	1.03	0.31	0.09	0.77	0.02	0.90	(1,158)
SD lateral position	0.58	0.45	0.57	0.45	2.42	0.12	(1,158)
Max deceleration	**92.57**	**<.001**	**271.58**	**<.001**	**75.00**	**<.001**	(1,158)
Crossing time	0.57	0.45	1.28	0.26	0.07	0.80	(1,157)
Crossing speed	1.30	0.26	**6.00**	**0.02**	0.39	0.53	(1,157)
Time to resolve	**4.99**	**0.03**	**133.52**	**<.001**	**6.11**	**0.01**	(1,158)
CID	**7.03**	**0.01**	0.01	0.92	1.16	0.28	(1,157)

Time to brake	1.78	0.19	2.58	0.11	0.08	0.78	(1, 53)
Distance at brake	**9.98**	**<0.01**	1.08	0.30	0.70	0.41	(1, 53)

To investigate the effect of the distraction task on driver behavior, we performed t-tests of a subset of the simulator measures. There were no statistically significant differences but there were some indications of driving slower and braking later and harder when distracted and the time to resolve slightly longer in the distracted crossing scenarios.

The previously explained classification of adherence allowed the analysis of driver behavior with respect to the communicated behavior recommendation. As can be seen in Figure 6-28, 25% of all reactions were non-compliant, in slightly less than 35% no iHMI message was shown, and in more than 40% of all cases, the driver reaction was compliant with the recommendation.

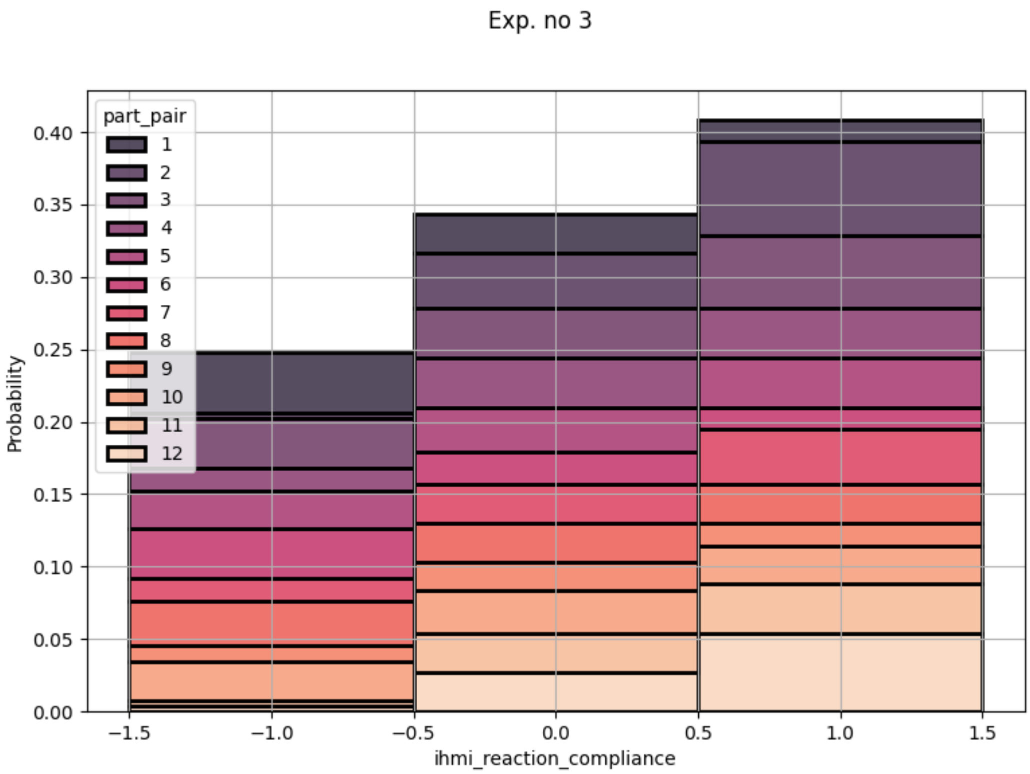


Figure 6-28: Proportion of compliant (value=1) and non-compliant (value=-1) driver reactions to an iHMI message normalized for the different participant pairs. The assumed maximum reaction time was set to 1.5s. The value 0 indicates that the iHMI was deactivated, hence, the subsequent driver behavior cannot be classified

In addition to the immediate driver behavior after an iHMI message had occurred, we analysed the drivers' reaction times to infer whether the driver reaction could have been realistically triggered by the recommendation or if the driver had possibly adapted the recommended behavior before the recommendation had occurred. Looking at the distribution of reaction times with respect to an iHMI message. The largest peak, with more than 140 occurrences, represented cases where the driver had already adapted the recommended behavior before the communicated message. Assuming that all compliant reactions within the first 0.5s after the iHMI message were not caused by the message itself, but an unaffected driver decision (as it represents unrealistic reaction times), allows to analyse the causality of the iHMI. Figure 6-29 shows that around 8% of all compliant driver reactions can be assumed to be caused by the iHMI.

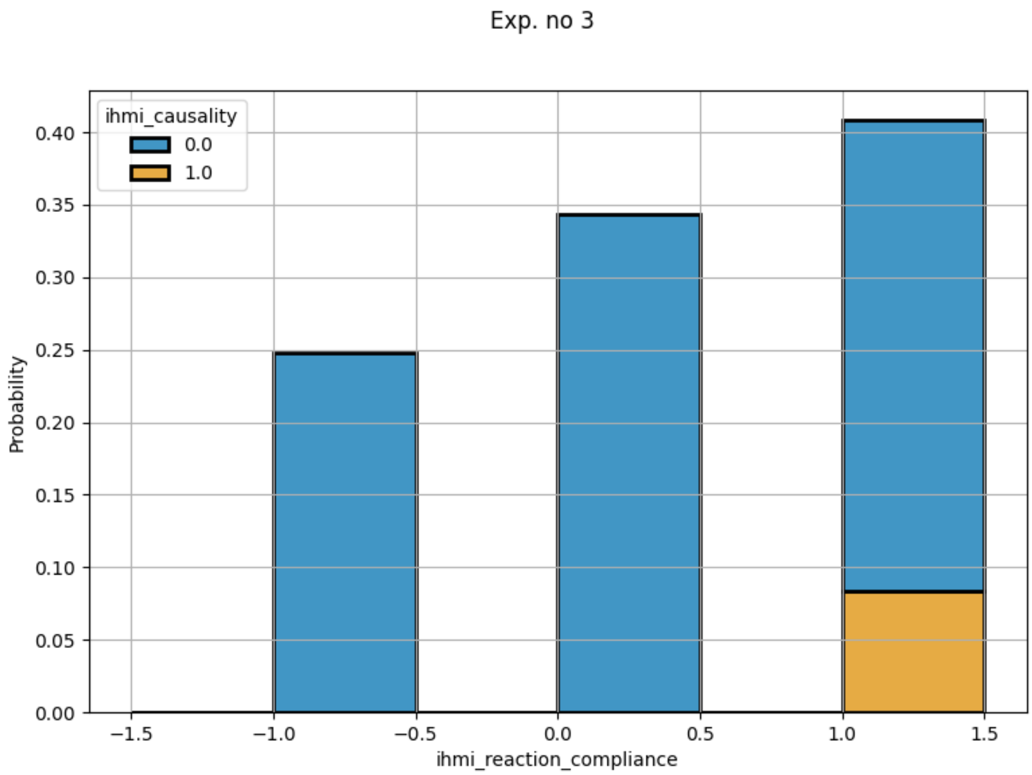


Figure 6-29: Proportion of compliant driver reactions to the iHMI. A communicated message is classified as causal (ihmi_causality=1.0) if the compliant driver reaction appeared within a time span of 0.5s – 1.5s after the iHMI message.

A similar analysis was done for the pedestrian crossing behavior. The classification of the pedestrian behavior was done with respect to the pedestrian position and its velocity. The pedestrian was classified as crossing if it reached a distance of 0.1m to the road at a velocity that exceeds an absolute value of 0.4m/s. This threshold allows to account for the noisy velocity data and filter smaller velocity values that occurred due to movement of the sensors. Most of the reactions that were compliant and that were not already initiated by the pedestrian can be found between 4s and 5s after the eHMI message showed up. This indicates a generally longer reaction time compared to the drivers. Approximately 5% of all compliant pedestrian reactions were assumed to be caused by the eHMI (Figure 6-30). Overall, this behavior explains why most of the interactions resulted in the vehicle passing first, even if the cHMI recommended that the pedestrian should cross first. The pedestrian adherence analysis shows that only a few pedestrians react within a short time with the correct behavior. This finding should be addressed in the next study in order to increase the acceptance of the cHMI by both participants.

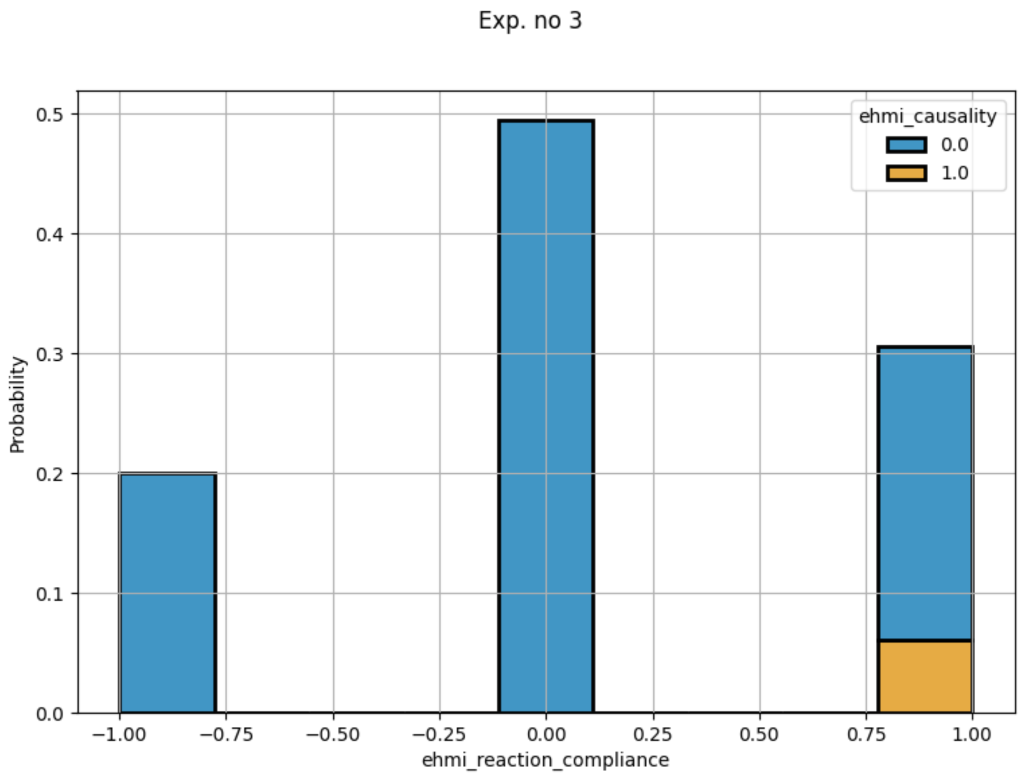


Figure 6-30: Proportion of compliant pedestrian reactions to the eHMI. A communicated message is classified as causal (ehmi_causality=1.0) if the compliant pedestrian reaction appeared within a time span of 0.5s – 5.0s after the eHMI message.

We analysed the drivers' gaze with respect to the pedestrian and the pedestrians' gaze with respect to the car (as proxy for looking at the eHMI). Table 6-2 shows that the percentage of time the drivers had their gaze on the pedestrian was similar in scenarios with and without cHMI. The pedestrians, however, had a significantly larger share of gazes towards the car in scenarios with cHMI triggered.

Table 6-2: Gaze behavior in crossing scenarios with and without cHMI

	HMI status	N	Mean	SD	t	p-value
Driver gaze on pedestrian (%)	No HMI	84	10.99	13.17	1.30	0.20
	HMI	78	8.39	12.15		
Pedestrian gaze on car (%)	No HMI	84	26.30	22.80	-5.20	<.001
	HMI	78	45.16	23.34		

There was no significant difference in drivers' gaze on pedestrians in scenarios with distraction task activated (11%) compared to scenarios without distraction (9%).

6.3.2 Subjective measures

Wilcoxon signed rank tests showed that there were no significant differences in the levels of perceived stress (Z=1.07, p=0.29) or sleepiness (Z=-0.05, p=0.96) in the test blocks with and without cHMI. Workload scores from the NASA-RTLX scale showed higher frustration scores after the test with cHMI (Z=-2.65, p<0.01).

The questions to pedestrians after each interaction showed that the participants felt relatively safe in the crossing situations (Table 6-3) but it was more unclear who had priority in the crossing scenarios where cHMI was activated compared to tests without cHMI (U=2570.5, p<0.01).

Table 6-3: Average ratings from the questions to pedestrians after each crossing scenario

	No HMI		HMI	
	Mean	SD	Mean	SD
How safe did you feel the situation was? (0 - very unsafe to 5 - very safe)	4.6	0.8	4.4	0.9
Was it clear to you who had priority? (0 - very unclear to 5 - very clear)	4.2	1.6	3.9	1.5

The participants rated how easy or difficult it was to see and interpret the cHMI messages on a scale from 1=very difficult to 6=very easy (Figure 6-31).

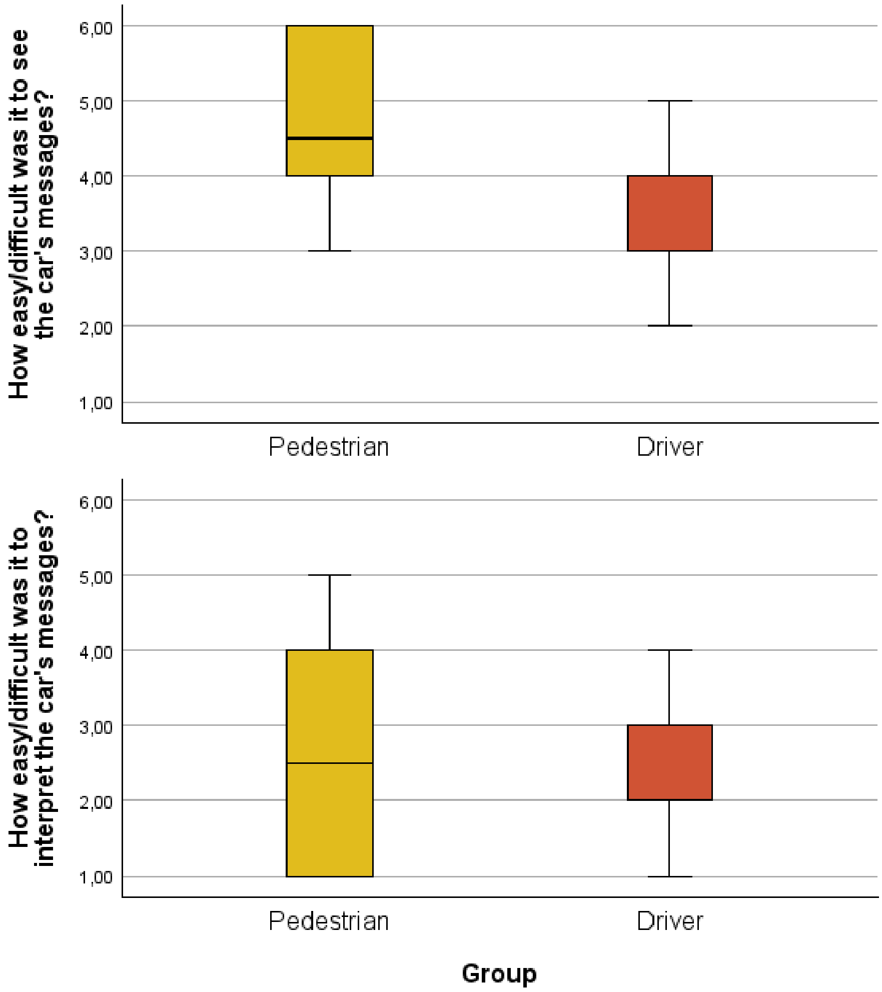


Figure 6-31: Mean ratings of how easy/difficult it was to see and interpret the cHMI on a scale from 1=very difficult to 6=very easy

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. Figure 6-32 provides an overview of the average ratings for each statement. In general, the cHMI was not rated as particularly helpful or safety enhancing, but it was not considered overly dangerous either.

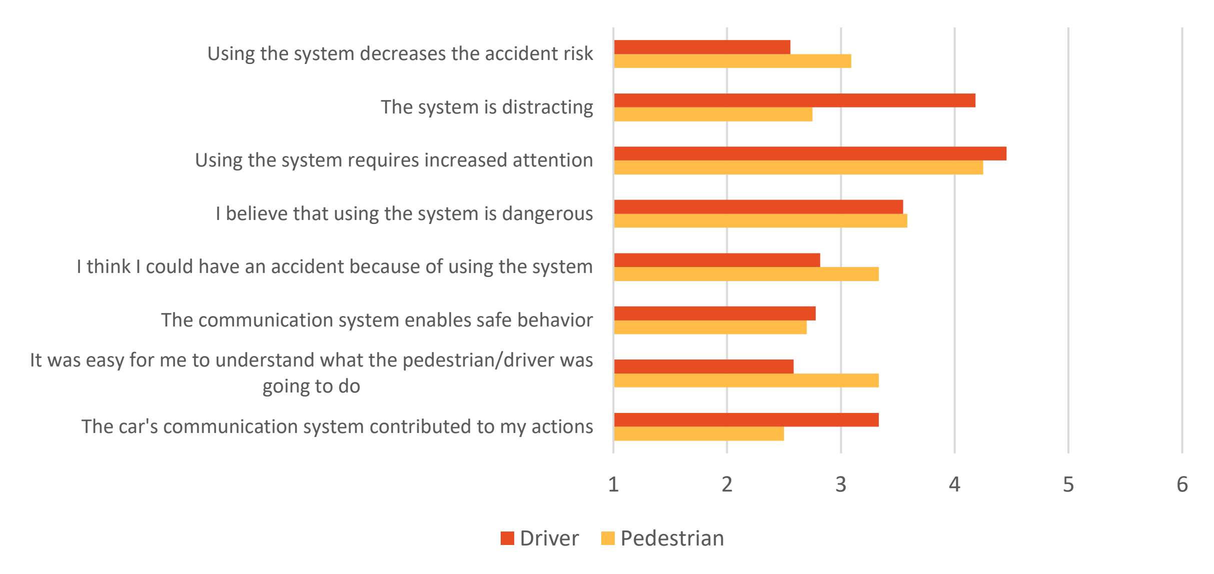


Figure 6-32: Average ratings of different statements about the cHMI on a scale from 1=do not agree at all to 6=totally agree

The questionnaire scores for SUS and SATI showed that the cHMI system had relatively low usability and trust (Table 6-4). Mean SUS scores were well below what is generally considered above average usability (SUS = 68). Mean trust scores on the SATI scale ranging from 0 to 5 were around 2 for both drivers and pedestrians.

Table 6-4: Questionnaire scores from the cHMI evaluation

	Pedestrian, mean (SD)	Driver, mean (SD)
SUS	51.0 (18.8)	57.9 (16.3)
SATI	2.0 (0.9)	1.9 (0.8)
UEQ Overall	-0.1 (1.2)	-0.1 (0.8)
UEQ Pragmatic	-0.6 (1.3)	-0.6 (1.0)
UEQ Hedonic	0.5 (1.3)	0.4 (1.0)

The results from the short UEQ showed that the cHMI system was experienced as relatively poor, with somewhat better ratings for hedonic quality compared to pragmatic quality. 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. The simulator experience was evaluated with a set of questions with response options ranging from 1 to 7. As seen in Table 6-5, the pedestrians scored the simulated environment as more realistic than the drivers.

Table 6-5: Questionnaire scores from the simulator questionnaire.

Question	Drivers (Mean/SD)	Pedestrians (Mean/SD)
How similar was this experience compared to walking or driving in reality? (1- not similar, 7- very alike)	3.8 (1.3)	5.0 (1.2)
How realistic did you perceive the milieu? (1- not realistic, 7- very realistic)	3.8 (1.3)	5.0 (1.1)
To what degree did you experience nausea? (1- to a very small degree, 7- a high degree)	3.1 (2.3)	2.2 (1.8)

If you experienced nausea, did that affect your driving or behavior? (1- very little, 7- very much)	3.7 (2.3)	2.7 (2.4)
How did you experience the visuals in the headset? (1- very stuttering, 7- very even)	6.3 (1.1)	6.2 (1.5)
Was the headset uncomfortable? (1- very uncomfortable, 7- not at all)	5.5 (1.4)	4.6 (2.2)

In the interviews, some of the participants acting as drivers raised questions about the timing of the cHMI. “It [the iHMI] changed very close [to the pedestrian] and if I would have stopped, I would have stopped earlier, because this would have been a hard brake...”. The participants also raised questions about the size and the placement of the iHMI, and the need for such a system. Multiple participants were worried the iHMI may take focus from the surroundings and therefore leading to a less safe driving behavior.

Some of the drivers mentioned they needed some time to get used to the iHMI and to learn to understand it. When asked about sufficient time in the training block they all agreed they had had enough time before the tests started.

The eHMI was prominently considered easy to see, although not as easy to understand or interpret. Some participants had trouble perceiving if the arrows/chevrons pointed upward or downward. The first participants who were completely naïve to the cHMI said that they ignored the cHMI, rather than trying to figure it out. They expressed that some instructions about the meaning of the HMI symbols would be necessary for understanding.

7. Conclusions

The main focus of the three studies was to evaluate the validation methods. Therefore, a relatively small number of participants was recruited for each study. As these are preliminary results from a limited sample size, the interpretation of the results should be done with caution. The validation studies in combination with experiences from the experimental studies preformed in WP2 and WP3 provided insights about which validation methods are suitable for the evaluation of the final HEIDI HMI solutions. Conclusions from the three studies and general conclusions and recommendations are presented below.

7.1 Study 5

7.1.1 Study 5a – regular drivers

Driving behavior was positively affected by the introduction of an iHMI, as compared to baseline driving conditions. Indeed, without iHMI, participants initiated the braking more than 1,5 seconds later than with an iHMI present. Accordingly, they applied significantly more pressure on the brake pedal. There was no statistical difference in those responses between iHMI versions. These results seem to indicate that any iHMI that adequately informs drivers of approaching pedestrians would be beneficial. Interestingly, drivers in distracted states braked, on average, even earlier than drivers in attentive state when the car was equipped with an iHMI that could account for the distracted state (horizontal fluid and fully fluid iHMI versions). Distracted drivers also applied less pressure on the brake pedal. Taken together, these results suggest that the choice of triggering the warning message one second earlier when driver's distraction is detected was effective in enhancing the safety margins. Finally, the occurrence of collisions, despite the lack of statistical evidence, was remarkably reduced in all the iHMI conditions as compared to baseline driving without the help of iHMI.

In the questionnaires, participants indicated on average a moderate mental workload across all conditions, including the one without iHMI support. However, at a deeper analysis, the results highlight how a drive without HMI support was perceived as slightly more demanding in terms of attentional and visual load, compensated by a reduced temporal and auditory demand, as compared to a driver supported by any type of HMI. This is to be expected, given the nature of the HMI signals that involve specific sequences of auditory clues, otherwise absent in the baseline driving conditions. In general, there is some evidence showing that the fully fluid iHMI helps drivers feel slightly more focused on the driving task, with neglectable influences on perceived stress and sleepiness. The user experience seems more defined by pragmatic aspects, rather than hedonic qualities of the interfaces. These results seem to indicate that the iHMI were perceived for their usefulness, efficiency, and ease of use, rather than for their "joy of use". All the iHMI versions tested showed no differences in their capacity to facilitate trust in automation, which was expressed by a high trust score (5 out of 6 points on average). Also, situational awareness was not affected by the presence/absence of the iHMI support, and was high (above 80%) in all measured conditions. This result seems to indicate that the chosen driving task and events were well within the abilities of the study participants.

From the interviews it can be concluded that 65% of participants preferred the fully fluid and horizontal fluid iHMIs, i.e., the systems that adapt the messages to the driver's attentional state, providing as arguments the additional warning displayed on infotainment screen (50%) and the comfortable escalation of the warnings (15%). The preferred iHMI versions were also indicated as the most comfortable, safe and informative. Overall, participants indicated that all elements of the iHMI – LED, sound, icons – were perceived as helpful in providing more time to react to the situations and in supporting the driver to focus on the driving task. This results are in line with the objective measures reported above and related to a safer driving behavior in the presence of an iHMI.

7.1.2 Study 5b – older drivers

Objective measures did not show statistically significant evidence of the impact of vertical fluid iHMI on driving behavior. However, at an observational analysis, the results seem to point in the same direction as reported for regular drivers, suggesting a slightly more cautions driving style.

From the responses to the questionnaires, it appears that driving without iHMI support results in a marginally higher stress, which remains however rather low. No differences result in terms of perceived distraction, sleepiness or efforts. All results related to the comparison between the regular and the older vertical iHMI versions show no differences in terms of usability, acceptance, perceived safety and trust.

In the interviews, participants appreciated the precise information provided by the system in relation to the side of the crossing pedestrians. Conversely, they expressed the wish to personalize the interface with regards to the voice messages and the graphical layout of the visual icons, often reported as too prominent. Overall, both iHMI versions were equally liked by participants (50% of preference each).

7.2 Study 6

The study examining the effectiveness of external Human-Machine Interfaces (eHMIs) for pedestrian-vehicle interactions has yielded promising initial results, despite the limited sample size. Both the Short-Term HMI (HMI1) and Long-Term HMI (HMI2) eHMIs demonstrated potential benefits compared to the baseline condition without an eHMI (HMI0).

The objective measures suggest that both eHMIs resulted in faster crossing initiations and reduced Time-To-Resolve (TTR) for pedestrian-vehicle interactions. This indicates that the presence of eHMIs may facilitate more efficient decision-making processes for pedestrians. Notably, with both eHMIs, pedestrians sometimes initiated crossing before the vehicle came to a complete stop, suggesting increased confidence in understanding the vehicle's intentions.

The average pedestrian crossing speed remained consistent across all conditions, indicating that the eHMIs did not induce unsafe behaviors such as running or strolling. This consistency in crossing speed, combined with faster initiation times, resulted in improved overall traffic flow, as evidenced by the increased average car speeds for both eHMIs.

Subjective measures, although limited due to the small sample size, provided encouraging feedback. Participants reported low mental, physical, and temporal demands across all conditions. Both eHMIs were perceived as easy to use and quickly learnable, with high confidence reported for the Short-Term HMI.

While these initial findings are promising, it is crucial to interpret them cautiously due to the study's limitations. Further research with a larger sample size is necessary to validate these preliminary results and draw more definitive conclusions.

The Long-Term eHMI was found to be unnecessarily complex when the computer-controlled pedestrian was walking on the other side of the road, potentially causing confusion. Additionally, the Long-Term HMIs slightly smaller messages may have been affected by resolution limitations of the VR headset, potentially impacting the clarity and effectiveness of the Long-Term HMI with higher spatial resolution.

7.3 Study 7

In this study, a first version of the cHMI was tested that did not include all features of the full HEIDI concept. The overall impression from the participants was that the current version of the cHMI needs improvement to provide any real benefit to drivers and pedestrians. The participants' comments were both about the appearance of the HMI messages and the timing of messages. The effects of the cHMI were not as expected since the participants found it less clear who had priority when HMI was triggered. Time to resolve was longer indicating that the cHMI did not improve efficiency.

In the interview, many participants said that the cHMI did not affect their behavior but from the simulator data we could see that both the drivers' braking behavior and the pedestrians' crossing initiation and gaze behavior was different when the cHMI was triggered.

It was problematic when the eHMI communicated that the driver was intending to yield when the driver could, in fact, choose not to adhere. Although the fluid cHMI is able to change the external message to reflect the driver's behavior, the change in message to the pedestrian might be too late for the pedestrian to see and process. The eHMI should thus not communicate that the driver is intending to yield in lower levels of automation (SAE 0-2) if it is uncertain if the driver is going to adhere. In order to leave the participants more time and space to react to the HMI, the resolution tracking should be revised for the next study considering the insights and results from this study.

No large effects of the distraction task were seen in driving behavior or gaze, probably due to the limited number of crossing scenarios with driver distraction. The task itself worked well.

Regarding the objective measures, some of these, such as time to resolve vary a lot depending on whether the pedestrian crossed before or after the car. Therefore, crossing (or yielding) behavior should be taken into consideration when analysing these parameters.

In study 7, it was not possible to define cHMI elements or the tablet presenting the secondary task as AOI in the eye tracking data. Future studies should track gazes on iHMI and eHMI elements to enable analysis of how much the drivers and pedestrians look at the cHMI. Tracking of gazes on the secondary task would also be helpful in analysing the effects of driver distraction.

Differences between younger and older pedestrians were not tested statistically due to the small number of older participants. The purpose here was instead to do a general evaluation of whether the validation methods and procedures worked for older people. In general, the protocol worked quite well for older people but there are concerns about increased risk of VR induces sickness or dizziness in this group.

7.4 General conclusions about validation methods

A combination of subjective and objective measures was used in the studies which provided a comprehensive evaluation of both the behavior and opinions of drivers and pedestrians.

To address potential VR resolution issues, future studies should utilize higher-resolution VR technology to ensure that the eHMI (especially the Long-Term HMI's more detailed display) is clearly visible and effective. Additionally, extending the research to real-world testing with prototype vehicles equipped with the eHMIs would help validate the findings in real-world scenarios and assess their effectiveness under various environmental conditions.

Longitudinal studies should be designed to allow for longer-term exposure to the eHMIs, evaluating their effectiveness and user acceptance over time. Cross-cultural studies should also be conducted to ensure the eHMI messages are universally understandable and effective across diverse cultural backgrounds.

In study 7, the first test round was performed with naive participants. It became clear that naive participants ignored the cHMI, rather than trying to figure it out, and therefore the decision was made to give the participants short information about the symbols they were about to see.

It was clear that there is a benefit of using both subjective and objective measures since the self-reported behavior and objectively measured behavior did not always match. Objective measures and surveys seem best suited to capture behavior change related to HMI activation whereas interviews provided more insights about the understanding and interpretation of HMI messages. A conclusion from study 7 was that group interviews may be well suited to discuss the HMI and gather reflections and insights about the system, and not as well suited when reflecting on one's own behavior.

The questionnaire instruments included in the study had several similar items, which could be frustrating to the participants. It is recommended to avoid asking too many or redundant questions. Some of the questionnaire items and scales are designed for a specific type of user or for either drivers or pedestrians which can make them difficult to answer. The SUS was perceived as difficult to answer by some pedestrians in study 7 as it refers to usage of a specific system. Looking at eHMI messages on a passing vehicle was not considered 'using a system' by them. Background questionnaire data such as driver and pedestrian behavior can provide useful information about the participants to enable detection of bias or support the interpretation of results.

Questions appearing in the simulation directly after an interaction proved effective to gather the experience of the pedestrians in study 7. A similar approach could be used also in real-life tests if both question and answers are done verbally.

7.5 Study limitations and lessons learned

All studies were small-scale studies with a relatively small number of participants which limits the possibility to generalize the results.

Summarizing the conclusions from all three studies, it was evident that the HEIDI concept worked well when iHMI and eHMI were implemented separately and there was full control of one of the agents' behavior. Testing the cHMI with a driver and pedestrian who were free to act as they found appropriate in study 7 revealed problems that had not been encountered in study 5 and 6 and in earlier exploratory studies. This was partly due to testing a limited version of the HEIDI cHMI but also highlighted the importance of being able to handle non-adherence to the communicated message. In addition, the cHMI test scenarios were without a zebra crossing, which might also impact the appraisal of the situation.

Study 6 was conducted with only two groups, significantly limiting the statistical power and generalizability of the results. Due to the small sample size, the planned statistical analyses could not be performed, limiting the ability to draw statistically significant conclusions.

The studies were conducted in a simulated environment, which may not fully capture the complexities and variabilities of real-world pedestrian-vehicle interactions. Furthermore, the limited exposure time to each HMI condition may not have allowed for a comprehensive assessment of long-term usability and effectiveness.

In study 5 there was no evidence for a different situation awareness between distracted and attentive drivers. This seems attributable to the method used for assessing the situational awareness, as the induction of distraction was confirmed by eye tracking data and the accomplishment of the secondary task.

The HMI signals used in study 7 were not exactly the same as the ones used in study 5 and 6. This was mainly because the messages to be communicated according to the decision logic were not the same, especially for the iHMI messages. In study 7, all crossing scenarios were at an unsignalized crossing, meaning that it was not mandatory for the driver to let the pedestrian cross first. Depending on the pedestrian behavior, the message communicated to the driver was sometimes to continue driving to avoid a deadlock situation. This type of message was not used in study 5. Furthermore, the cHMI in study 7 was a simplified version of the HEIDI concept and should be refined based on the results from the validation studies. Since the auditory signals were appreciated by the participants in study 5, it is recommended to add auditory signals in addition to the HUD for the cHMI.

The experiences from the data collections highlight the importance of thorough pilot testing to ensure that the procedures are safe, accepted, and understood by the participants. Some of the built-in safety measures in the pedestrian VR-world were possible to pass through. Additional restraints could be considered to ensure that the pedestrians do not walk outside of the designated area.

When performing tests with multiple participants it is important to separate the participants in the breaks between test blocks, because some participants were eager to discuss their behavior with other participants and find mutual solutions on the interaction.

7.6 Recommendations

Based on the results from the validation studies, the following general recommendations are given for the evaluation tests in WP7.

- Future research should focus on conducting larger-scale studies to enable robust statistical analysis and increase the reliability and generalizability of the findings.
- Use of both subjective and objective validation methods is recommended.
- Cross-over designs are preferred over between-subject designs since individual differences are large.
- Thorough pilot testing with the full experimental protocol is highly recommended, including a check of logged data.
- It is not recommended to have completely naïve participants in the evaluation of HMI features since there is a risk they will ignore the messages if they do not understand the meaning of them.
- To lower the burden on participants, it is recommended to only pick the most relevant questionnaire instruments and items.
- Short questions directly after a crossing scenario can be a good way of capturing immediate reactions.

- In simulator studies, susceptibility to motion sickness should be screened for in the recruitment process and driving scenarios should be designed to minimize motion sickness, in particular for older adults.

Both subjective and objective measures are recommended for use in the forthcoming studies. The validation studies showed that a combination of questionnaires and interviews with participants provided a comprehensive view of their experiences. Table 7-1 and Table 7-2 below present the validation metrics that are recommended for use in WP7 evaluation tests.

Table 7-1: Objective measures recommended for use in WP7 evaluation

Metric	Unit	iHMI	eHMI	cHMI
Start Speed	km/h	X		X
Time to Brake	s	X		X
Distance at Brake	m	X		X
Standard Deviation of Lateral Position - SD	m	X		X
Number of HMI communication signals	n			X
Maximum vehicle deceleration	m/s2	X		X
Crossing time	s		X	X
Crossing speed	m/s			X
TTR, time to resolve	s		X	X
Crossing Initiation delay (CID)	s		X	X
Crossing behavior	categorical			X
Adherence	categorical			X

Table 7-2: Subjective measures recommended for use in WP7 evaluation

Questionnaire tool	Metric	iHMI	eHMI	cHMI
Demographic data	Age, gender, driving experience	X	X	X
Mini-DBQ	Subscale scores for violations, errors and lapses (0-5)			X
ATI	Personal ATI score (1.0-6.0)			X
EQ-5D-5L	Individual items (1-5) and health score (0-100)	X		X
PBS	Violations, errors and lapses (1-6)			X
SUS	Usability score (0-100)	X	X	
SATI	Trust score (0-5)	X		X
UEQ-short	Overall score and subscale scores for pragmatic and hedonic quality (-3 to 3).	X		X
CTAM	Selected items (1-7)	X		X
NASA-TLX	Subscale scores for mental-, physical- and temporal demand, performance, effort and frustration (0-21)	X	X	X
VTI Acute stress scale	(1-9)	X		X
KSS	(1-9)	X		X

In addition, for the usage of eHMI solutions on the road, industry stakeholders and regulatory bodies should work towards standardizing eHMI designs and communication protocols, ensuring consistency across different vehicle manufacturers, as intended in deliverable "D5.5 Best practices for design and development of HMI systems" and deliverable "D7.3 Finalized recommendations for legislation proposal and standardization".

By addressing these recommendations, future research can build upon the initial findings of these studies and contribute to the development of effective, safe, and universally understood eHMIs for pedestrian-vehicle interactions.

8. Abbreviations

Term	Definition
ATI	Affinity for Technology Interaction
cHMI	Combined Human-Machine Interface
CID	Crossing initiation delay
CIT	Crossing initiation timestamp
CTAM	Car Technology Acceptance Model
DALI	Driving Activity Load Index
DBQ	Driver Behavior Questionnaire
eHMI	External Human-Machine Interface
EQ-5D-5L	Instrument to describe and value health
FMS	Fast motion sickness scale
HEIDI	Holistic and adaptivE Interface Design for human-technology Interactions
HUD	Head-Up-Display
iHMI	internal Human-Machine Interface
KSS	Karolinska sleepiness scale
LED	Light-Emitting Diode
NASA-TLX	NASA Task Load index
PBS	The Pedestrian Behavior Scale
PU	Public
R	Document, Report
SATI	SHAPE Automation Trust Index
SD	Standard deviation
SUS	The System Usability Scale
TTC	Time-To-Collision
TTR	Time to resolve
UEQ	The user experience questionnaire
VR	Virtual reality
VRU	Vulnerable road user
VSS	VTI acute stress scale
WP	Work Package

9. References

[1] Martinussen, L.M., et al., *Short and user-friendly: The development and validation of the Mini-DBQ.* Accident Analysis & Prevention, 2013. **50**: p. 1259-1265.

[2] Franke, T., C. Attig, and D. Wessel, *A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale.* International Journal of Human–Computer Interaction, 2019. **35**(6): p. 456-467.

[3] EuroQol Research Foundation. *EQ-5D-5L User Guide*. 2019 September 2019 [cited 2024; Available from: https://euroqol.org/publications/user-guides.

[4] Granié, M.-A., M. Pannetier, and L. Guého, *Developing a self-reporting method to measure pedestrian behaviors at all ages.* Accident Analysis & Prevention, 2013. **50**: p. 830-839.

[5] Lewis, J.R., *The System Usability Scale: Past, Present, and Future.* International Journal of Human–Computer Interaction, 2018. **34**(7): p. 577-590.

[6] Dehn, D.M., *Assessing the Impact of Automation on the Air Traffic Controller: The SHAPE Questionnaires.* Air Traffic Control Quarterly, 2008. **16**(2): p. 127-146.

[7] Schrepp, M., A. Hinderks, and J. Thomaschewski, *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., 2017.

[8] Osswald, S., et al., *Predicting information technology usage in the car: towards a car technology acceptance model*, in *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2012, Association for Computing Machinery: Portsmouth, New Hampshire. p. 51–58.

[9] Hart, S.G. *NASA-task load index (NASA-TLX); 20 years later*. in *Proceedings of the human factors and ergonomics society annual meeting*. 2006. Sage publications Sage CA: Los Angeles, CA.

[10] Dahlman, A.S., et al., *Validation of a one-item acute stress scale for driving tasks.* Advances in Human Factors of Transportation, 2024. **148**(148).

[11] Kaida, K., et al., *Validation of the Karolinska sleepiness scale against performance and EEG variables.* Clinical Neurophysiology, 2006. **117**(7): p. 1574-1581.

[12] Kircher, K., & Ahlström, C. (2017). The Driver Distraction Detection Algorithm AttenD. In M. A. Regan, J. D. Lee, & T. W. Victor, Driver Distraction and Inattention (1. Aufl., S. 327–348). CRC Press. https://doi.org/10.1201/9781315578156-23.

[13] Keshavarz, B., & Hecht, H. (2011). Validating an efficient method to quantify motion sickness. Human factors, 53(4), 415-426.