State of the Art on validation methods for cooperative and adaptive HMI solutions

Deliverable Number	D5.2
Deliverable Type	R – Document, Report
Dissemination Level	PU (Public)
Author(s)	Ignacio Solis, Sara Nygårdhs, Anders Andersson (VTI)
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
Ignacio Solis	VTI		
Sara Nygårdhs	VTI		
Anders Andersson	VTI		

Reviewers

Name	Organisation	Date
Robert Leute	Marelli	2023-03-15

Change History

Version	Date	Name/Organisation	Description
1.0	2023-03-06	Ignacio Solis, Sara Nygårdhs, Anders Andersson (VTI)	Sent for review
Reviewed version	2023-03-15	Robert Leute (Marelli)	Review
2.0	2023-03-27	Sara Nygårdhs, Anders Andersson (VTI)	Final version, revisions made according to the review

Table of Contents

List of Figures

List of Tables

1. Executive Summary

The scope of this deliverable is to investigate the most commonly used methodologies for evaluating both internal and external adaptive human-machine interfaces (HMIs) in the literature. Another objective is to gather information on methodologies based on the use of co-simulation of multiple simulators, where different road users interact in shared virtual spaces. This work addresses the specific objective of T5.2 ("Validation methods for adaptive HMIs") and is aligned with HEIDI's strategic objective number 3 ("Develop suitable validation methods for assessing fluid, cooperative HMI solutions").

This report presents literature reviews for each of the three aforementioned areas: 1) Methodologies for evaluating adaptive internal HMIs, 2) Methodologies for evaluating external HMIs, and 3) Methodologies based on co-simulation for investigating interactions mainly between drivers and pedestrians. Additionally, various studies on internal HMIs for older drivers were analysed, as this is one of the target groups in the HEIDI project. For each study, general information was collected about the overall objectives and characteristics of the HMIs, as well as more specific information about methodological aspects such as test environments and scenarios used, participant groups, objective and subjective measures, and experimental designs. In total, 18 empirical studies on adaptive internal HMIs, 5 on specific internal HMIs for elderly drivers, 50 on external HMIs, and 7 on co-simulation were reviewed.

Keywords: State of the art, validation methods, adaptive, internal HMI, external HMI, co-simulation

2. Objectives

This deliverable aims to provide a representative overview of evaluation methods for adaptive HMI systems. In order to achieve this goal, the following specific objectives have been defined:

> (1) To carry out a thorough analysis of the existing methods for the evaluation of adaptive internal HMI systems.
>
> (2) To conduct an in-depth analysis of existing methods for the evaluation of external HMI (eHMI) systems, and
>
> (3) To collect information on studies that have used co-simulation methods to investigate the interaction between different road users, with a special emphasis on the interaction between a driver and a pedestrian.

These tasks address the HEIDI project's strategic goal "Develop suitable validation methods for assessing fluid, cooperative HMI solutions" (i.e., Objective 3 in the Grant Agreement). The literature review presented in this deliverable will serve as a foundation for identifying the best methods for evaluating the internal, external and cooperative HMIs developed in WPs 2, 3 and 4, respectively. These methods will in turn be tested and refined through the initial studies (WP2 and 3), small-scale studies (WP5) and employed for the evaluation studies (WP7) described in Annex I of the Grant Agreement (p.16).

3. Introduction

3.1 HMI solutions for vehicle-VRU interactions

One solution to reduce the high rate of traffic accidents in the world (i.e., 1.35 million deaths every year), is the use of active safety systems (i.e., systems that minimize the chances of collision; [1]). Specifically, among this group of systems, two have received significant attention. On the one hand, automated driving systems, capable of taking control of some (i.e., in the case of semi-automated levels) or all of the driving tasks (i.e., in the case of fully automated vehicles), minimizing the risk of accidents due to driver distraction, fatigue or sleep. On the other hand, human-machine interface (HMI) systems, which allow providing information and warnings to the driver to improve his/her performance or responses. More recently, with the rise of automation research, external HMIs (eHMIs) have also received attention as an effective solution to provide information to external users about the automation status or the vehicle's intentions, among others. Additionally, other systems, such as vehicle-to-vehicle systems or vehicle-to-infrastructure, driver and environment monitoring systems or advanced prediction algorithms, are also experiencing an exponential growth and will maximize the potential of HMIs and automated systems.

The effectiveness of internal and external HMIs has been widely studied in the literature, although in different contexts. For instance, several studies have emphasized the effectiveness of internal HMIs in mitigating sleepiness, fatigue, or distraction in drivers during manual or automated driving [2, 3]. Similarly, internal HMIs (iHMIs) have been shown to be effective in increasing situational awareness of drivers of automated vehicles and increase their ability to regain control in case of obstacles [4-6]. While iHMIs have been investigated in different levels of automation (from manual to autonomous level), eHMIs have been exclusively investigated in contexts of interaction between vulnerable road users (VRUs) and autonomous vehicles (AVs). This stems from the need to find ways of communicating between VRUs and AVs whose drivers may no longer be available [7]. To date, preliminary results indicate that eHMIs will be welcomed by VRUs, making them feel more secure and helping them make decisions on AV crossings [7]. However, these results are still preliminary and there are multiple aspects that have been investigated only superficially or not at all (e.g., different use cases, influence of vehicle behaviour, etc.).

70% of fatal VRU accidents occur in urban environments [1]. This is a clear indication that urgent measures are needed to facilitate interaction between different users in shared traffic spaces. In view of the promising results on the potential of iHMIs and eHMIs, a strategy to improve interaction between VRUs and vehicles, regardless of their automation level, would be to combine the potential of both to modulate the behaviour of drivers and pedestrians to prevent these collisions. Thus, instead of relying on one-sided communication strategies, cooperative HMIs (cHMIs), i.e., iHMIs cooperating with eHMIs, would make use of the information provided by internal and external sensors (i.e., about the driver state, the traffic situation, the current automation level and the pedestrian's characteristics and behaviour) to predict the course of the interaction and develop joint action strategies for safe, efficient and comfortable interactions. In turn, messages can be adapted to maximize their effectiveness, considering different scenarios and contexts.

Despite the fact that cHMIs have yet to be fully explored, the rapid pace of technological advancement suggests that these systems will become feasible in the near future. As such, the HEIDI project, as detailed in the Grant Application, aims to conceptualise, develop prototypes, and test a cooperative HMI system that is flexible enough to adapt to various situations, automation levels and user characteristics, with a particular emphasis on improving interactions between pedestrians and drivers of vehicles with different automation levels.

3.2 Validation methods for the HEIDI system

The validation of iHMI, eHMI, and cHMI systems is a prerequisite for their implementation in real vehicles. Unlike verification, which ensures that the system functions correctly and meets established requirements and standards, validation focuses mainly on ensuring that the system generates the desired outcomes. In the specific case of cHMIs, like the one to be developed in HEIDI, the validation process must demonstrate that these systems improve driver-vehicle-VRU interactions in terms of safety, efficiency and comfort.

To date, standardized validation methods for adaptive iHMIs or eHMIs do not currently exist, although organizations such as EuroNCAP are making efforts to integrate them in their assessment protocols. This implies that no methods for validating cHMIs are available either, which would require more complex and holistic approaches involving multiple users (e.g., drivers interacting with VRUs in different situations). On a positive note, the study of the effectiveness of adaptive iHMIs and eHMIs has grown in recent years, leading to various evaluation methods being proposed. At the European level, for example, these efforts are also evident in projects investigating adaptive HMIs and/or eHMIs (e.g., AIDE, Adaptive, ADAS&ME, InterAct, BRAVE, HoliDes, and MEDIATOR). Despite this, these methods are not fully consolidated and there are significant differences among studies in aspects such as the use cases and scenarios used, the type of participants, data collection techniques, and measures.

As defined in the Grant Agreement, one of the objectives of the HEIDI project is to propose innovative methods for the validation of cooperative HMIs, specifically in multi-user environments, where the interaction between drivers and pedestrians, modulated by the HEIDI cooperative HMI system, will be investigated. To achieve this goal, HEIDI will adopt a stepwise approach. The first step, the main purpose of this deliverable, is to conduct a state-of-the-art analysis of the different methods used to date to evaluate adaptive iHMIs and eHMIs. The aim of this step is to provide a representative review of the different methodologies employed. In a later phase, the different methodologies will be internally evaluated and adapted to the use cases defined in D1.2. Subsequently, different methods will be tested within WP2 during the development of the adaptive iHMI and within WP3 during the development of the eHMI. Finally, these methods will be combined to evaluate the validation methods of the cooperative HMI in a co-simulation environment.

4. Scope of the State-of-the-Art analyses

As mentioned, there are no studies where cHMIs, such as the one to be developed in HEIDI, have been tested. This implies that the validation methods to be designed in this project will have to build on the learnings of evaluation methods for adaptive iHMIs and eHMis. Moreover, the potential of the HEIDI cHMI cannot be evaluated without the involvement of several road users, i.e., driver and one or several pedestrians, interacting in concrete traffic situations. For this purpose, the use of co-simulation of simulators is an affordable, safe solution to analyse such interactions and with sufficient external validity to extrapolate the results to the real world.

As a preliminary step to T5.4 ("Design and validation of the implementation tests"), this deliverable presents three independent literature reviews carried out on three main areas, i.e., iHMI, eHMI and co-simulation methods, using the Scopus database as the main source, complemented by other sources such as Google Scholar or ResearchGate.

A brief description of the scope of the reviews in each of the areas is presented below:

- Adaptive iHMIs. One of the sub-components of the HEIDI system is the adaptive iHMI, capable of adapting the type, timing and mode of information to the driver's characteristics (e.g., age) and state (e.g., distracted or alert), to the level of automation (e.g., manual, semi-automated or fully automated), and to the interaction situation with other users. Hence, the focus of the state-of-the-art analysis carried out is on studies where adaptive iHMIs have been used. The studies include various automation levels, from manual to fully automated. As elderly drivers are included in the use cases of HEIDI but there is a lack of studies on adaptive iHMIs and elderly, some studies on non-adaptive iHMIs and elderly have also been included.
- eHMIs. Initially, the scope for this state-of-the-art analysis included all studies that have used eHMIs to improve VRU-vehicle interaction. However, as will be shown later, most studies focused on interactions between pedestrians and autonomous vehicles. Therefore, the scope of the state-of-the art analyses turned out to be more specific than planned as it does not include other VRUs (e.g., cyclists or two power-wheelers) and semi-automated vehicles.
- Simulator methods, where co-simulation (generally connected driving and pedestrian simulators) also provides a platform for research on interactions in a controlled environment. The literature on this topic is sparse and some insights and methods used in this young field are provided here.

4.1 Adaptive internal Human-Machine Interfaces

In the HEIDI project, adaptive HMIs are of interest. Adaptive HMIs can adjust to characteristics of the driver or traffic environment, such as when HMI modality or information is adapted to driver state or complexity of the traffic environment. In this part of the literature review, concerned with adaptive internal HMIs, the focus is on studies where the iHMI is adapted to the driver and where the driver can interact with the iHMI.

4.1.1 Paper selection and aggregation

A literature search was conducted in the literature database Scopus, where studies between 2012–2022 were included. Survey-only studies were excluded, because participants in these studies would have to infer things that they have not experienced. In the HEIDI project, the aim of this part of the deliverable is to investigate what methods have been used for evaluating adaptive internal HMIs. Keywords such as “internal HMI”, “adaptive” and “driver” were used in combination, resulting in a total of 228 references. The abstracts were reviewed to check which of the studies seem to fit within the scope of adaptive and internal HMIs. Of these, only studies that still fitted the scope after full-text review, or where there was enough significant information in the abstract, were included.

The final selection included 18 articles on adaptive internal HMIs. In Figure 4-1 the year of publication of the reviewed studies is visualised.

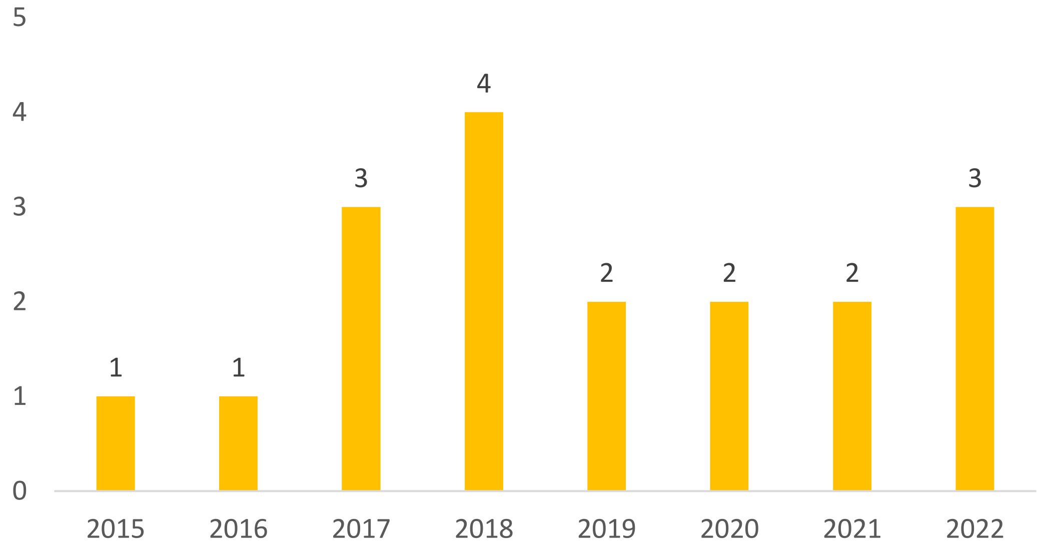

Figure 4-1: Year of publication of the reviewed studies on adaptive iHMIs.

Seven study characteristics corresponding to these articles were chosen to get an understanding of the studies and to be able to compare them. These characteristics were:

- Test environment
- Study objectives
- HMI characteristics
- Adaptation characteristics
- Sample characteristics and experimental design
- Scenarios used and constructs evaluated
- Measures used.

Table 4-1 shows a summary of the test environments used in the studies from the literature review on adaptive iHMIs.

Table 4-1: Aggregated test environments for adaptive iHMIs.

Reference	Driving simulator	Real road
Amparore et al. 2018; [8]	Desktop	
Biondi et al. 2017; [9]		Public roads
Boelhouwer et al. 2020; [10]	Fixed base, car mockup, 180°	
Coeugnet et al. 2021; [11]		Public roads, WOz
Feldhütter et al. 2018; [12]	Fixed base	
Galarza et al. 2017; [13]		Public roads

Large et al. 2019; [14]	Fixed base, car cabin, 270°	
Li et al. 2019; [15]	Moving base, car cabin, 180°	
Maag et al. 2015; [16]	Moving base, car cabin, 180°	
Manawadu et al. 2017; [17]	Desktop	
Manawadu et al. 2018; [18]	Desktop (forward, left, and right screens)	
Meiser et al. 2022; [19]	Not specified	
Musabini et al. 2020; [20]		Public roads, WOz
Perrier et al. 2022; [21]	Desktop	
Riyahi et al. 2021; [22]	Fixed base, car cabin (forward projection)	
Tanabe et al. 2022; [23]	Not specified	
Ueda & Wada 2016; [24]	Fixed base, car seat (forward projection)	
Wandtner 2018; [25]	Moving base, car cabin, 240°	

The larger part of studies on adaptive internal HMIs carried out has been performed in driving simulators. In a driving simulator, the experimenter has more control over the situation and has more degrees of freedom to decide which situations the participant should experience and when. Situations can be repeated and compared between participants. In addition, driving simulators have advantages in that the participant and the vehicle are not subject to real risks, although risky situations can be experienced. On the other hand, studies performed on real roads have all the elements necessary to resemble a real traffic situation, where the participant is exposed to real risks and can take decisions and perform behaviour as she or he would do in a more naturalistic way. Here, two out of four real-road studies on adaptive iHMIs used a Wizard of Oz (WOz) vehicle, where a hidden joystick was managed by a professional driver without the participant knowing, to simulate autonomous driving. For an adaptive iHMI to be used in real life, tests on public roads are necessary. However, all functionality and user acceptance tests that are rational to make before implementation in a real vehicle on real roads will probably benefit from the controlled yet versatile study environment of a driving simulator. Since none of the reviewed studies on adaptive internal HMIs included use of Virtual Reality (VR) headsets, only simulator studies and studies conducted in the real world are addressed in the following.

4.1.2 Study objectives and adaptive iHMI characteristics

Table 4-2 shows an aggregation of the study objectives, purpose and characteristics of the HMI and the adaptations in the studies.

The study **objective** is declared in the fourteen simulator and four real-world studies. The simulator study objectives can be summarised as investigating different interfaces and the impact of iHMI interventions on supporting drivers, on distraction, on acceptance, and in terms of efficiency. Similarly, in the real-world studies [9, 11, 13, 20] the study objectives also relate to the HMI.

In the studies, the **purpose** of the iHMI has mostly been to inform and trigger some kind of reaction from the driver (e.g., to improve driving or take-over). The information is sometimes a warning and sometimes encouraging a better (e.g., safer, with less emissions) driver behaviour. Promoting drivers to try out certain functionality is another of the iHMI purposes mentioned. In general, information from the iHMI should lead to increased situation awareness, which is a prerequisite for reduced driver distraction, in line with the HEIDI project objectives.

The most common **input modality** by which the driver can communicate with the iHMI in these studies is tactile. Examples of this is where the driver presses a button [10, 12, 16] or touches a touchscreen [17-19], but also by drivers' braking or steering operations [22, 24]. Another, but less common input modality is visual. Examples of visual input here is where the system scans the face of the driver [8], or where there is an interface for hand-gestures [17, 18]. In some study, auditory input as spoken commands [14] was also given.

The predominant **output modality** is visual. For all accessible full-text studies, there was a visual component in the information from the iHMI to the driver. This is often materialised by a screen [8, 14] or a head-up display (HUD) [12, 16] but was also in the form of augmented reality (AR) overlay on the windscreen [10, 17] or LED flashes [22]. In some case the screen of a mobile phone was used [15]. In addition, acoustic signals were used for attention [12] or acceptance/rejection of driver input [17, 18] and verbal communication was used for information and warning [10, 11, 14, 21, 25]. Tactile output was also used, as haptic feedback at the steering wheel [16, 24], force-feedback opposing the steering and speeding commands given by the driver and additional vibration feedback if ignoring this force-feedback [17, 18].

For the studies to be included in this literature review for the HEIDI project, they needed to contain **adaptation**. In some studies, the system was adapted to the driver state, intention or driving style [11, 15], in other cases to the complexity of the environment/driving situation [9, 10, 13, 14, 16, 19, 22], and in some cases to both driver and environment [8, 25]. There were also cases where information of the iHMI was adapted to the type of automation mode that the vehicle was in [12] or to battery range of an electric car [20]. In some studies, the driver could choose his or her preferred input modality by themselves [17, 18]. In one study, adaptation of driver behaviour to different iHMIs was studied [21].

Table 4-2: Aggregation of study objectives, iHMI purpose and characteristics, and type of adaptation.

Reference	Study objective	iHMI purpose	iHMI input/output modality	Adaptation
Amparore et al. 2018; [8]	to illustrate the concept of “co-pilot” as an enabling technology for autonomous driving	prompt reactions	input modality: visual output modality: visual	suggestions of co-pilot and actions adapted to driver intention/distraction and complexity of environment
Biondi et al. 2017; [9]	investigate how naïve drivers interact with ACC and LKA	assist		HMI information adapted to traffic and traffic environment
Boelhouwer et al. 2020; [10]	investigate digital in-car tutor to support drivers in learning about and trying out their car automation	tutor, inform about system and its limitations	input modality: tactile output modality: auditory, visual	modality, timing, and duration of communication adapted to driving situation
Coeugnet et al. 2021; [11]	to assess an adaptive HMI to improve safety and quality of take-over in a level 3 automated vehicle	warn, prompt reactions	input modality: tactile, visual output modality: visual	HMI information adapted to detection of drowsiness or distraction
Feldhütter et al. 2018; [12]	to investigate the relevance of mode awareness and mode errors in the context of vehicle automation + examine influence of a cognitive-auditive and a visual-motoric NDRT and an adapted HMI	inform, prompt reactions	input modality: tactile output modality: visual and auditory	HMI adapts information displayed to automation mode
Galarza et al. 2017; [13]	ease driver interaction with infotainment system by adapting interface according to different levels of driving complexity	ease driver interaction with infotainment system in demanding scenarios and offer better user experience for low-demanding cases	input modality: tactile output modality: visual	type of infotainment system adapted to traffic scenario complexity
Large et al. 2019; [14]	explore how drivers’ behaviour and attitudes changed over a week, and impact of different HMI interventions	inform, prompt reactions	input modality: auditory output modality: visual, auditory	hand-over requests adapted to environment
Li et al. 2019; [15]	to investigate driver distraction when using Eco-Safe HMI	encourage an eco-safe driving style	input modality: vehicle-to-vehicle and vehicle-to-infrastructure (not directly from driver) output modality: visual	HMI information adapted to driver’s eco-friendly driving

Maag et al. 2015; [16]	investigate effectiveness, acceptance, and controllability of an advisory warning system using the haptic channel for warning the driver about upcoming hazards	warn, prompt reactions	input modality: tactile output modality: tactile or visual	HMI adapted to context in form of upcoming hazards
Manawadu et al. 2017; [17]	evaluate a multimodal interface system with three input modalities; touchscreen, hand-gesture and haptic to input tactical-level control commands (e.g., lane-changing, overtaking, and parking)	inform, prompt reactions	input modality: tactile, visual output modality: tactile, auditory, visual	modality adapted to driver’s desire
Manawadu et al. 2018; [18]	propose and evaluate tactical-level input method with a multimodal HMI for driver intervention in automated driving	inform, prompt reactions	input modality: tactile, visual output modality: tactile, auditory, visual	driver can choose input modality adapting to dynamics of driving environment
Meiser et al. 2022; [19]	to understand effects of adaptive HMI that change according to mental workload of driver	inform	input modality: tactile output modality: visual	information presented on display adapted to complexity of environment
Musabini et al. 2020; [20]	examine if acceptance of electric vehicle technology is eased by HMI proposing coping strategies	inform about battery range, where to charge, and what to do at battery breakdown	input modality: tactile output modality: visual	HMI information adapted to battery range
Perrier et al. 2022; [21]	compare usability of 3 HMIs for ACC	inform	input modality: tactile output modality: visual, auditory-visual	driver behaviour adapted to different HMIs
Riyahi et al. 2021; [22]	to explore if external stimulation can effectively improve drivers’ reaction and response capabilities	warn, prompt reactions	input modality: tactile output modality: visual	LED light box flash is adapted to hazard
Tanabe et al. 2022; [23]	investigate effects of robot HMI to inform drivers of emergencies when lane tracking assistance (LTA) disconnects	prompt reactions – prepare for overtaking	input modality: tactile output modality: tactile	driver response adapted to type of HMI
Ueda & Wada 2016; [24]	evaluate effectiveness of proposed HACC (haptic ACC) compared to conventional ACC	inform, prompt reactions	input modality: tactile output modality: tactile, visual	(driver can change preceding vehicle to another by adding torque to resist presented torque of the system)
Wandtner 2018; [25]	to evaluate different HMI concepts regarding their potential to facilitate drivers’ self-regulation in NDRT interactions and take-over performance, and assess user experience and acceptance of HMI concepts	inform, prompt reactions	input modality: tactile output modality: visual, auditory	Adaptive HMI 1 adapted icons, flashing and speech-output to situation and to mode of driving; Adaptive HMI 2 adapted icons, flashing and speech-output to situation and to mode of driving, but also to driver state by triggering notifications again in non-acute situations, and by automatic braking maneuver avoiding crash if driver did not respond to take-over request

Only a few of the adaptive iHMIs in simulator studies found in the literature search were adapted to the driver to some degree. Drivers' own driving style affected the information given by the iHMI [15] or the driver could choose which modality to use for HMI communication [17, 18] or the HMI was triggered to repeat notifications by non-responsive drivers [25]. In one simulator study only, the driver was monitored to detect driver distraction, which was considered as looking away from the road [8]. One of the real-world studies included an iHMI that was adapted to the driver and the purpose was then to detect driver drowsiness or distraction in relation to an automated vehicle [11], but the information about driver state was processed offline. In the HEIDI project driver state is very relevant as the cooperative HMI should be adapted to the human both inside and outside of the vehicle. In addition, both elderly and distracted drivers should be handled by the cooperative HMI.

4.1.3 Participant characteristics, experimental design and measures used

In Table 4-3, participant characteristics, the experimental designs, scenarios used, constructs evaluated, and measures used are aggregated.

The fourteen **simulator studies** in Table 4-3 were performed in driving simulators with 11–52 participants, with a median number of 30 participants. In most studies, a within-subject or mixed experimental design was used, and the most common types of factors were related to the iHMI, the traffic situation, or a specific task. In none of the studies participants older than 65 years were recruited. Participant experience of a system, or the lack thereof, was mentioned in a few studies [12, 25]. In the simulator studies, safety, efficiency, and trust were the major constructs evaluated, together with acceptance and situational awareness. The constructs have been evaluated in different types of scenarios, i.e., in different traffic environments (e.g., highways, rural roads, urban roads, intersections) and in different situations (e.g., roadworks, risky traffic events, take-over requests, support system malfunction). Many of the simulator studies [10, 14, 16] also aimed to estimate driver acceptance or usability of a system, which is hard to assess objectively. Consequently, objective measures have often been used together with subjective measures such as scales or interviews.

Two of the four **real-world studies**, of which one with 52 participants, were carried out as WOz studies, in which automated driving was simulated by use of a hidden joystick [11, 20]. The scenarios used in the real-world studies were in general selected to test the trust and acceptance of the system, and hence, subjective measures have mostly been employed. One of these studies [11] used objective measures, all physiological measures related to physical status and gaze. A possible explanation for not using more objective measures in real-world studies is the increased risk of severe consequences in real and complex traffic situations, in relation to simulator studies. Introducing measurement apparatus that might be intrusive could theoretically pose an additional risk to the driver.

Table 4-3: Aggregation of sample characteristics, experimental design, scenarios used, constructs evaluated, and measures used.

Reference	No. participants	Participant selection with respect to HMI	Experimental design	Scenarios used	Constructs evaluated	Measures used
Amparore et al. 2018; [8]	30		2x2 within-subjects: level of distraction, type of system	two-lane highway (preceding vehicle brakes unexpectedly, slow preceding vehicle with an incoming car on the left lane) + secondary task	safety	**objective**: no. of accidents, percentage of driving time with TTC of preceding vehicle less than 2 s, number of times driver makes a hard braking (deceleration of more than -8m/s2), average distance to preceding vehicle when user performs lane change, average TTC when driver starts to press the brake
Biondi et al. 2017; [9]	10	no prior experience with ACC and LKA	pre-post (before and after driving with ACC and LKA), paired sample t-tests	highway with ACC and LKA	trust, acceptance	**subjective**: questionnaires and scales, thinking-aloud
Boelhouwer et al. 2020; [10]	38	students or employees at University of Twente, (18–65ys)	2x3, between-subjects: tutoring method, session	highway w/wo fog, curved road, roadworks, jaywalker, pedestrian obstructed view, priority signs, unsignalized intersection, pedestrian crossing, road markings missing at highway with curved section, stationary car, emergency vehicle	acceptance, appropriate automation use	**objective**: collisions, correct take-over and reliance behaviour, TTC, deceleration rate, lateral acceleration **subjective**: questionnaire adapted from TAM
Coeugnet et al. 2021; [11]	52		within-subjects: 2 timings of take-over request (WOz)	unplanned and planned take-over requests (8 vs. 45 s) in automated mode	safety, trust	**objective**: eye tracking, heart rate, skin conductance, respiration **subjective**: interview
Feldhütter et al. 2018; [12]	49	23% no experience, 45% occasional use of system, 33% regular use of system	2x2, mixed: type of HMI, frequency of automation mode shift	three-lane highway and lane changes and NDRTs	efficiency (mode awareness)	**objective**: reaction time to automation error, proportion of undetected automation errors, proportion of false interventions
Galarza et al. 2017; [13]	15	high appreciation of technology	within-subjects: 4 complexities of HMI	complex and easy driving segments, different road types	user acceptance, usability, safety	**subjective**: scales for user acceptance and user-perceived safety
Large et al. 2019; [14]	52	drove regularly, possibly biased with high trust	2x2, between-subjects: level of feedback, advice given	take-over request	trust, acceptance, situational awareness	**objective**: time spent looking at roadway and different NDRTs, mean position of centre of vehicle after resuming manual control, number of mirror checks during advice, SDLP, time to driver readiness

						subjective: rating scales (trust, acceptance (TAM), situational awareness (SART))
Li et al. 2019; [15]	36	18–65 ys	4x3, within-subjects: type of info display condition, traffic situation	car-following, signalized intersection, stop-sign intersection	safety, efficiency	**objective**: eye tracking data: saccade duration (total and average), saccade frequency, glance duration (total and average), glance frequency; to areas of interest (front roadway/HMI/speedometer)
Maag et al. 2015; [16]	24	test driver panel of WIVW GmbH	3x2x2x2, within-subjects: type of HMI, dynamics of hazard, criticality of hazard, hazard object	urban road with one or two lanes per driving direction, intersections, parked vehicles etc. with 20 test situations with very low to high criticality (hazard motion, direction, type and position), and instruction to drive at 50 km/h and keep to right lane but can overtake slower vehicles	acceptance, efficiency, controllability, safety	**objective**: lateral distance between hazard and vehicle, maximum deceleration, speed difference, difference in steering wheel position, timing of lateral reaction **subjective**: interviews on preference, ratings of HMI usefulness (after each experienced hazard) and timing and criticality (for wrong warnings)
Manawadu et al. 2017; [17]	20		3x4, within-subjects: input modality, event	expressway (80 km/h) with lane closure due to roadworks, urban area (40 km/h) with abrupt braking of lead vehicle, sub-urban area (40 km/h) stopped vehicle and pedestrian incursion, parking lot (10 km/h) with person standing next to parking spot	safety, efficiency	**objective**: mean input error frequency, mean input time, information carried by different input modalities, use of input modality **subjective**: NASA-TLX, driving experience questionnaire
Manawadu et al. 2018; [18]	11	normal or corrected-to-normal vision	within-subjects: 3 takeover modes	roadworks in urban traffic environment	efficiency, usability	**objective**: vehicle position, speed, lane position, steering angle, pedal position, multimodal HMI input data, skin conductance, mean heart rate **subjective**: NASA-TLX, driving experience questionnaire
Meiser et al. 2022; [19]	35		2x3x2, mixed: type of HMI, task difficulty, task presence	straight segments or wide curves on country road (easy); and many tight curves and turns in a city (complex)	comfort, efficiency	**objective**: heart rate, heart rate variability, driving performance, secondary task performance, number of clicks, relative success with solving task subjective: UX questionnaire
Musabini et al. 2020; [20]	22	inexperienced AV drivers, 18–45 ys, living in a big city, no EV owner, one	between-subjects: 2 types of HMI (WOz)	electric car battery is low and finally empty	acceptance, trust	**subjective**: interviews (TIPI questionnaire, AttrakDiff scale)

		trip/week, >5000 km/year, familiar with ADAS				
Perrier et al. 2022; [21]	24	normal vision and hearing	2x3, mixed: system range, type of HMI	village road with 40 mph, motorway section with 70 mph	efficiency, usability, comfort, safety (distraction)	**objective**: eye-tracking data (glances towards HMI), total glance time at HMI before correct response (to set ACC to 55 mph for instance), response time **subjective**: scales (RSME, R-TLX, SUS); semi-directed interviews
Riyahi et al. 2021; [22]	30		3x3, within-subjects: type of warning, type of event	near-collision	efficiency	**objective**: steady-state visually evoked potentials (SSVEP) of brain signals, time to collision, reaction distance, collision rate
Tanabe et al. 2022; [23]	27		2x3, within-subjects: type of HMI, scenario	3 scenarios in absence of LTA due to bad weather	safety, trust	**objective**: driver fatigue, driver arousal (by near-infrared spectroscopy, NIRS), steering torque
Ueda & Wada 2016; [24]	12	males ca 20 ys, females ca 30-40 ys	2x3, within-subjects: type of system, scenario	straight one-way two-lane road (4 m wide lanes), 4 scenarios: forked road (preceding vehicle drove onto left branched road), false detection (system tries to follow passing vehicle), lane changing (stop following preceding vehicle and sometimes another vehicle drives in front of the ego vehicle), standard (follow on left lane of straight road)	efficiency	**objective**: SDLP (standard deviation of lane position), BPDP (brake pedal depression percentage), steering torque, velocity, acceleration timing (early, late), minimum TTC
Wandtner 2018; [25]	36	experience with smartphone or tablet devices	3x2x2x2, mixed: type of HMI, task, scenario, predictability of take-over situation	two-lane highway, drive on right lane 120 km/h or 80 km/h (at roadworks), with longitudinal and lateral vehicle guidance (HAD), broken down cars in own lane and overtaking vehicles in other lane, narrow zone and lane change to left, with take-over more and less acute + NDRTs	safety, efficiency, trust	**objective**: timing and quality of take-over, gaze data (glance duration), NDRT disengagement **subjective**: NASA-TLX, SUS, Likert scales

Subjective measures generally aim to assess the driver's perceived experiences. Examples of those used in the literature are:

- Questionnaires and scales. Questionnaires and scales used in the literature on adaptive iHMIs refer to driving experience [17, 18], user experience (e.g. UEQ) [19], user acceptance [10, 13], technology acceptance (e.g. TAM questionnaire) [10, 14], mental effort (e.g. RSME) [21], task load (e.g. NASA-TLX, R-TLX) [17, 18, 21, 25], usability (e.g. SUS) [21, 25] and situational awareness (e.g. SART) [14].
- Interviews. Interviews can allow for more nuanced answers to questions asked by the interviewer. In the literature on adaptive iHMIs interviews about preferences were conducted in two of the simulator studies [16, 21] and about opinions on system functionality [11] and state of mind of the driver [20] in real-world studies.
- Thinking-aloud, where the driver talks about his or her driving as it occurs, is a technique that was carried out in one of the real-world studies [9].

The **objective measures** can be divided into different categories, e.g., aspects related to the vehicle, driver performance, automation and take-over specific parameters, and physiological measures. In the following, measures are listed that could be assigned to these categories.

Vehicle

- Longitudinal control: Vehicle speed [16, 18, 24] and position along the road [18].
- Lateral control: Lane position [14, 18], standard deviation of lane position (SDLP) [14, 24], steering wheel angle [16, 18], steering torque [23, 24], lateral acceleration [10], lateral distance between hazard and vehicle [16] and timing of lateral reaction (comparing several HMIs) [16].
- Use of input modality [17] and multimodal HMI input data [18].

Driver performance

- Handling hazards: Collision rate [22], number of collisions [8, 10], time to collision (TTC) [8, 10, 22, 24] and reaction distance (e.g. distance to collision subtracted by distance to driver's initial response) [22].
- Acceleration/deceleration: Number of times driver makes a hard braking (e.g. defined as deceleration of more than -8 m/s^2) [8], brake pedal depression percentage (BPDP) [24], maximum deceleration [16], deceleration rate [10] and acceleration timing (driver's intention to change lanes) [24].
- Driving task: Average distance to preceding vehicle when user performs lane change [8], response duration [21], mean input time (e.g. the time period from when the driver starts to move his or her hand to make an input to the point in time when the system accepts the input) [17], mean input error frequency (e.g. an input is given that the system cannot recognize) [17] and number of mirror checks during advice [14].
- Non-driving related tasks (NDRTs): NDRT performance (where the purpose of the task was to increase mental workload) [19].

Automation and take-over specific

- Automation and driver: Reaction time to automation error (e.g. time between appearance of lane change sign which the automation interpreted wrongly and intervention by driver in partially automated driving) [12], proportion of undetected automation errors [12] and proportion of false interventions [12].
- Timing: Timing of take-over [25].
- Quality: Quality of take-over [25], correct take-over and reliance behaviour [10], NDRT disengagement [25] and time to driver readiness [14].

Physiological

- Cardiac data: Heart rate [11, 18, 19] and heart rate variability [19].
- Respiratory data: Respiration rate [11].
- Skin conductance data: Skin conductance [11, 18].
- Eye tracking data [11]: Glance frequency [15], number of glances towards HMI [21], glance duration [14, 15, 21, 25], saccade duration [15] and saccade frequency [15].
- Hemodynamic brain activity: Near-infrared spectroscopy (NIRS) (to assess driver fatigue and driver arousal) [23].
- Electrophysiological brain activity: Steady-state visually evoked potentials (SSVEP) of brain signals [22].

Of the **vehicle** parameters, the speed and position of the vehicle are frequently used measurements [14, 16, 18, 24]. In a simulator, these parameters are easy to register and analyse in detail. The standard deviation of lane position (SDLP) [14, 24] is used as a measure of vehicle control. Additionally, the steering wheel angle [16, 18], steering torque [23, 24] and lateral acceleration [10] are used for estimating how much the driver turns and how hard he or she steers. Other measures that may be assigned to the vehicle category include timing of lateral reaction [16], use of input modality [17] and how much information is carried by different input modalities [18].

Connected to the vehicle parameters are aspects related to **driver performance**. The most critical measures used in the simulator studies are collision rate [22] and number of collisions or accidents [8, 10]. However, collisions are relatively rare in real traffic, and indicative measures such as time to collision (TTC) [8, 10, 22, 24], distance to preceding vehicle at driving manoeuvre [8], number of hard brakings [8], brake pedal depression percentage (BPDP) [24], and maximum deceleration [16] have also been used. In studies where the driver had a non-driving related task (NDRT), the performance of this task could be measured by response time and success of the NDRT [19]. Other measures used are acceleration timing [24], number of mirror checks [14], time spent at data input [17, 21] and input error frequency [17].

For studies where the mode shifts between manual and automated [12, 14, 18, 25], some specific **automation and take-over** measures have been used. Measures associated with driver performance and automation are reaction time and reaction distance to automation error, proportion of undetected errors and false interventions. Take-over measures include the timing as well as the quality of the take-over, which also covers a correct take-over behaviour. In addition, a correct reliance behaviour can be relevant to investigate. Other parameters mentioned in the simulator studies are disengagement in NDRT and time to driver readiness.

According to the definition in the APA dictionary a **physiological measure** refers to “any set of instruments that convey precise information about an individual’s bodily functions” [26]. Hence, physiological measures consist of a variety of measures, where some measurement methods require that equipment is applied directly on the body, while other methods take measurements from a distance. Heart rate, heart rate variability and skin conductance have been used to assess mental workload [18, 19]. Different types of eye tracking data have been used, such as glance measures to different targets [14, 15, 21, 25]. Glance frequency, glance duration, glances towards a specific area (such as the iHMI), the total glance time at a specific area, but also saccade duration are measures used in the simulator studies on adaptive iHMIs. These measures have predominantly been used to investigate distraction or situational awareness. Other physiological measures include signals from the brain. In this respect, previous studies used steady-state visually evoked potentials (SSVEP) for evaluating the effectiveness of a warning system [22], while near-infrared spectroscopy (NIRS) on the prefrontal cortex was used as indication of driver fatigue and arousal [23].

Among the HEIDI goals are to define valid methods to evaluate the safety, usability, and effectiveness of cooperative HMI solutions. In conclusion, from the studies on adaptive internal HMIs, vehicle measures of interest can be speed and position of the vehicle, including SDLP (for safety evaluation). Driver performance parameters such as collision rate in a driving simulator or more indicative parameters such as TTC (safety) could be useful. Physiological measures include eye tracking data such as glance frequency and glance duration (safety, usability, effectiveness) as well as heart rate variability, skin conductance (usability) and SSVEP (effectiveness).

4.1.4 Internal HMIs and distraction

Detecting driver distraction is an important part of the HEIDI project, which aims to take a holistic approach and predict both driver and pedestrian action. The definition of distraction can however be discussed.

The European New Car Assessment Programme (EuroNCAP) is a five-star safety rating system for comparing cars of different makes and models. EuroNCAP requires that all new cars should be equipped with warning as well as intervention apparatus for both long and short distractions [27]. A long distraction is by EuroNCAP defined as driver gaze away from the forward road to one consistent location of ≥3 seconds, while a short distraction consists of repeated glances away from the forward road view to one or multiple different locations, such as for a cumulative 10 seconds within a 30 second period when glancing at the forward road view for periods less than 2 seconds. To warn and intervene, the driver state needs to be monitored. For both long and short distraction, a warning requirement is that “when the vehicle is travelling at ≥ 20 km/h, a visual + (haptic and/or audible) warning must be issued immediately after the driver is classified as distracted”.

Considering the EuroNCAP definitions or other definitions of distraction (e.g. [28, 29]), it is important that driver state is monitored. In contrast to more steady driver states, such as age or physical impairments, distraction is a state that all drivers can be subject to and that needs to be detected rapidly. Of the references on adaptive iHMIs, only a few could be interpreted to deal with distraction. In one of the studies, visual activity was monitored together with physiological data of the driver’s attentional level by use of eye-tracking, heart rate, skin conductance and respiration measures [11]. This study aimed to improve safety and quality of take-over in a level 3 automated vehicle and was conducted as a WOz study on a real road with 52 participants. A study with the objective of investigating driver distraction when using an

HMI developed for eco-driving was conducted in a driving simulator with 36 participants [15]. Glance behaviour was used as indicator of distraction, and eye tracking data such as saccade duration (total and average), saccade frequency, glance duration (total and average), and glance frequency to areas of interest was analysed in different scenarios. Areas of interest were in this study defined as front roadway, iHMI and speedometer. Another simulator study, focussing on comparing usability of different iHMIs, used the iHMI as the main area of interest [21]. Glances towards iHMI, the total glance time at the iHMI before the correct response as well as the response time was analysed for 24 participants in two scenarios.

Judging from the literature found on adaptive iHMIs, studies where the driver is monitored for distraction are rare.

4.1.5 Internal HMIs and elderly

Due to age-related decline in sensory, cognitive and psychomotor functions, some older drivers limit their exposure to certain traffic situations, such as driving at night or in busy traffic [30, 31]. Since this self-regulation in addition limits their mobility, new vehicle technologies may be helpful. New systems can, however, be difficult to understand [32] and hence it is a good idea to support and tutor elderly drivers when needed, as proposed in the HEIDI objectives. Adapting intervention criteria to individual drivers as well as to driver state would be a good way forward.

Because of the lack of studies in the literature review attributed to adaptive iHMIs and elderly, some additional papers where older drivers had been recruited for studies on iHMIs were reviewed. Table 4-4 shows examples of studies on iHMIs and elderly drivers. Studies performed where iHMIs have been explored from a perspective of elderly drivers have for instance been concerned with visual impairments, in which auditory information of the iHMI has been evaluated [33]. In addition, the importance of display size, i.e., legibility, has been mentioned [34]. Cognitive impairments that might follow with age have been studied in relation to intersections [35] and to communicating proactive steering interventions to improve understanding [36]. For automated vehicles, take-over requests must be communicated clearly to be understood and acted correctly upon in due time, as a result of one of the studies [37]. None of the studies in Table 4-4 relates to an internal HMI adapting in real-time to whether the driver is young or old, has impairments or not. Overall, however, considering aspects of visual, auditory, and cognitive performance does not only refer to elderly drivers. All drivers benefit from iHMIs communicating information that is easily attained and understood.

Table 4-4: Examples of studies on elderly drivers and iHMIs.

Reference	Study objective	Participant selection	Test environment	Scenarios used	Constructs evaluated	Measures used
Hashimoto et al. 2009; [33]	evaluate HMI making it easy for elderly drivers to follow guidance instructions in parking assistance system	18 participants: ‘elderly drivers’	test track	parking assistance based on oral instruction	validity, acceptability	objective: vehicle operation (vehicle trajectory optimization, parking precision) subjective: questionnaire
Ito et al. 2019; [36]	improve understanding of proactive steering intervention for elderly drivers	15 participants: 65-75 ys; healthy drivers	driving simulator	proactive steering intervention to avoid a parked car in lane	safety, trust, acceptance, workload	objective: lateral deviation subjective: questionnaires
Li et al. 2019; [37]	evaluate effect of 3 HMI concepts based on older drivers' requirements, on takeover performance, workload and attitudes	39 old: 60-81 ys; 37 young: 20-35 ys	driving simulator	collision avoidance (other vehicle stationary in lane demands take-over request)	comfort, efficiency, safety, usability, acceptance	objective: reaction time, takeover time, indicator time, time-to-collision, acceleration, steering wheel angle subjective: NASA RTLX, 7-likert scales
Nakamura et al. 2014; [35]	designing iHMI based on the characteristics of cognitive workload of elderly driver when approaching an intersection	n.a.: ‘young and old’	driving simulator	approaching intersection while cognitively loaded		objective: driving manoeuvre, performance of dual-task, distance from intersection, type of turn subjective: subjective report
Nguyen et al. 2021; [34]	clarify difference in driver's state between normal and surprising situations	35 participants: 65-85 ys; valid driving license	driving simulator	gear shift intentionally reversed when driving out from a parking lot		objective: reaction time, driver reactions, heart rate, ECG data subjective: questionnaire

4.2 External Human-Machine Interfaces

One of the main sub-components of the HEIDI cooperative HMI system is the eHMI, which will be mostly developed within WP3. The ambition of the HEIDI project is for eHMIs to be adaptive to the characteristics and state of VRUs, with focus on pedestrians. As mentioned above, the literature on eHMIs is very recent, and does not contain studies on adaptive eHMIs. For this reason, the current state-of-the art analysis is entirely based on nonadaptive eHMIs.

4.2.1 Paper selection and aggregation

A literature search was conducted in Scopus where the following terms were combined using boolean operators OR and AND: eHMI, external human-machine interface, external interface, vehicle, AV, Automated vehicle, Autonomous vehicle, vulnerable road user, VRU, pedestrian, interaction, crossing, mixed traffic. This search was then complemented with other relevant papers identified during the revision. As a result, a list of 131 full texts of journal articles, conference papers and book chapters was generated. From this list, 90 were accessible for download which were screened for selection according to the following inclusion/exclusion criteria:

Inclusion criteria:

1. Studies analysing interactions between pedestrians and vehicles equipped with eHMIs. This includes video-based or test-tracks where participants did not actually cross in front of the vehicles.
2. Studies that involved manipulation of one or more factors and gathered subjective or objective data to assess interaction safety, efficiency, comfort and/or user experience.

Exclusion criteria:

1. Studies focusing only on interactions between VRUs and autonomous shuttles.
2. Studies looking into interactions involving vehicles without eHMIs.
3. Surveys on preferences for eHMIs designs.
4. Review articles or studies exclusively using qualitative methods (i.e., focus groups, interviews, etc.)

In total, 50 eligible papers were identified and analysed in detail, all published after 2018 (see Figure 4-2). Although this list may not be comprehensive, due to some articles being inaccessible or escaping our keywords, it can be regarded as sufficiently large to provide an overview of the most commonly used methodologies. Specifically, general information was collected from each study, i.e., authors and year, main goals and characteristics of the eHMIs analysed, and information on methodological aspects, i.e., characteristics of the sample of participants, test environments, scenarios, measures and experimental designs.

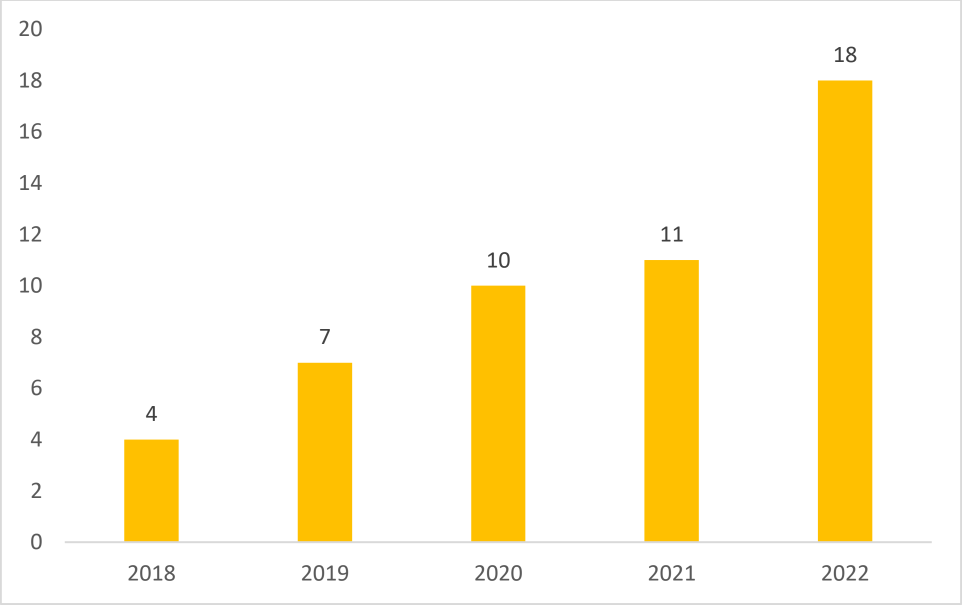

Figure 4-2: Year of publication of the reviewed studies on eHMIs.

4.2.2 Test environment

The vast majority of the studies reviewed were conducted in three types of test environments, described below:

- **Virtual Reality (VR) simulators**. Eighteen of the reviewed studies used simulators where participants were exposed to virtual scenarios in which they had to interact with vehicles equipped or not equipped with eHMIs. Sixteen studies used head-mounted sets, mainly HTC Vive Pro Eye equipment, while two others used Oculus Rift [38, 39]. Kaleefathullah et al. [40], by contrast, used a Cave Automatic Virtual Environment (CAVE) simulator, where scenarios were presented on external screens instead. The most common software engine used for the design of the scenarios was Unity, although others also used Unreal Engine (e.g., [41, 42]). In all studies, the relative position of pedestrians with respect to the autonomous vehicle was tracked, allowing the calculation of specific measures described in section 4.2.5.

- **First-person videos**. Eighteen of the reviewed studies used videos from a first-person perspective for the presentation of the experimental conditions. These studies can be categorised into those that used videos of real vehicles and environments, and those that used videos of virtual scenarios. Although in most studies the videos were viewed in the research laboratory, in other cases they were presented online (e.g., [43-45]. In such studies, participants are often presented with videos via head-mounted sets or on computer or TV screens (e.g., [43, 44]). The videos are typically 5-30 seconds long and depict common traffic interactions, usually in urban environments. Unlike in VR scenarios, the participants do not have to execute any crossings. As will be seen in the measures section (Section 4.2.5), these studies often use other ways to assess crossing intention, mainly based on button presses or responses to subjective scales.

- **Test-tracks.** Eight of the reviewed studies used test-tracks as test environments. In these environments, participants were exposed to "autonomous" vehicles on real, controlled roads. In almost all studies, the Wizard of Oz method was used (WOz), where the presence of the driver was hidden (mainly using car seat suits). In only one study [46], low-speed autonomous vehicles were used in an indoor test-track. As in the previous cases, the task of pedestrians was to decide whether or not to cross in the presence of an AV equipped with different eHMIs. In these cases, this decision was evidenced either by the crossing initiation, or by pressing a button.

4.2.3 Main goals

It should be noted that all revised studies shared a common overall objective, i.e., to analyse interactions between pedestrians and autonomous vehicles. Thus, although some studies included manual driving conditions (e.g., [47, 48]), no studies were identified including interactions with semi-automated vehicles. Besides this overall goal, studies also addressed specific objectives which can be conveniently categorised into the following groups:

1. **The effects of the eHMI type of information on pedestrian experience and decisions.** Arguably, one of the most researched eHMI messages is that of yielding intention. As can be seen in Table 4-5, practically all studies have included eHMIs that in different ways indicated their intention to yield. Comparatively, less attention has been paid to investigate the effectiveness of nonyielding messages to prompt non-crossing decisions. Other types of messages that have received large attention are those relating to the status of automation (e.g., manual or automated vehicle, [48-52]) or messages confirming pedestrian detection [39, 51, 53, 54]. Less frequently, several papers have analysed the effect of presenting information about specific vehicle manoeuvres (e.g., braking, starting, waiting [46]), estimated distance and/or time remaining to stop [53, 54] or time remaining to restart [48, 55].

2. **Effect of eHMI design on pedestrian perception, decisions and behaviour.** All studies used the visual modality for the eHMI messages through displays and LEDs. Some others also investigated eHMI strategies that, besides the visual display, incorporated vehicle kinematics as a channel for conveying information (e.g., [56, 57]), which can be considered a sub-type within the visual modality. Only two of the studies analysed eHMIs using auditory signals [58, 59]. Within the visual modality, the most investigated design aspects were: use of textual versus non-textual messages (e.g. "don't walk", [38, 60]), the location of the displays (e.g. grill, windshield top, etc; [61]), the perspective of messages (e.g. egocentric vs. allocentric, [62, 63]), the use of static vs. dynamic lights [64, 65], the colour of messages [66, 67] or the use of visual projections on the floor [68].

3. **The interplay between the vehicle behaviour and eHMI messages on pedestrian perception, decisions and behaviour**. A number of studies have noted that not only the type of message or the design of the eHMI is relevant for pedestrians when making crossing decisions. Kinematic aspects of the vehicle, such as speed, trajectory, and the timing and force of decelerations are also relevant, and indeed modulate the extent to which pedestrians comply with explicit eHMI instructions [38, 40, 50, 69, 70]. This has led to consider the inclusion of kinematic cues in the eHMI strategy, as discussed in the previous bullet point [56, 57, 71].

Besides the three main goals covered in the literature, various studies also looked at the influence of other factors. For example, [72] investigated the influence of other pedestrians on participants' decisions. Furthermore, others investigated the effect of vehicle appearance, i.e., size or model, reporting significant effects [38, 43, 73, 74]. Regarding individual-related aspects, various studies have investigated the influence of previous experience with eHMIs and malfunctions on pedestrian behaviour [40, 69, 75]. Similarly, the influence of age has also been investigated, albeit sparsely [65].

Table 4-5: Goals and eHMI characteristics of the revised studies.

Reference	Goal	eHMI message/s	eHMI modality
Bindschädel et al. 2022a; [57]	Effects of a two-step communication that uses an implicit cue at a long distance and subsequently an implicit or explicit cue at a short distance	Inform yielding intention	Visual and kinematic
Bindschädel et al. 2022b; [56]	Effects of active pitch motion and eHMI on pedestrian behaviour and experience	Inform yielding intention	Visual and kinematic
Bindschädel et al. 2021; [71]	Effects of eHMI presence and eHMI designs on traffic flow and road safety	Inform yielding intention	Visual
Burns et al. 2019; [46]	Effects of different eHMIs on pedestrians' attitudes and decision-making	Inform manoeuvring intention	Visual
Colley et al. 2022a; [72]	Effects of other pedestrians' behaviour, time pressure and prior experience on crossing decisions in front of AVs equipped with different eHMIs	Inform yielding/nonyielding intention and time to resume driving	Visual
Colley et al. 2022b; [43]	Effects of vehicle appearance, mode distinctiveness and passenger visibility on pedestrians	Inform yielding intention	Visual
de Clercq et al. 2019; [38]	Effects of AV yielding behaviour, AV size, eHMI design and eHMI timing on pedestrians' crossing decision	Inform yielding intention	Visual
Deb et al. 2018; [59]	Effects of visual and auditory eHMI features on pedestrians' behaviour and experience	Inform yielding intention	Visual and/or auditory
Dey et al. 2020a; [53]	Effects of different distance-dependent eHMIs on pedestrians' crossing decision	Inform status, intention, distance to stop, pedestrian recognition	Visual
Dey et al. 2020b; [50]	Effects of the eHMI under different vehicle dynamics on road-crossing decision	Inform yielding intention	Visual
Dey et al. 2022; [76]	Effects of explicit and implicit nonyielding communications	Inform nonyielding intention	Visual
Dey et al. 2021; [54]	Effects of scalable eHMIs for communications with two pedestrians	Inform yielding intention, pedestrian detection, estimated arrival time and estimated stopping point	Visual
Dietrich et al. 2019; [70]	Effects of eHMI timing with respect to different deceleration strategies	Inform yielding intention	Visual
Eisma et al. 2020; [77]	Effects of display location on crossing intentions and eye movements	Inform yielding/nonyielding intention	Visual
Eisma et al. 2021; [62]	Effects of eHMI perspectives (allocentric vs egocentric) on pedestrians' perception and decision under yielding and nonyielding situation.	Effects of eHMI location on crossing intention and eye movements	Visual
Epke et al. 2021; [39]	Effects of uni- or bi-directional communications between AVs and pedestrians on crossing behaviour and perception	Inform pedestrian detection	Visual
Faas et al. 2019; [67]	Effects of two eHMI colours indicating automation mode on pedestrians' opinions	Inform automation status	Visual
Faas et al. 2020a; [64]	Effects of eHMI designs (steady, flashing and sweeping light) on pedestrian comprehensibility	Inform yielding intention	Visual
Faas et al. 2020b; [51]	Effects of exposure to eHMIs informing status and intent on pedestrians' experience and crossing decision	Inform status or/and intention	Visual
Faas et al. 2020c; [78]	Effects of eHMI information type on pedestrians' crossing decision and experience	Inform status, pedestrian recognition and/or intention	Visual
Faas. et al. 2021; [49]	Effects of eHMI presence and driver state on pedestrians crossing decision	Inform automation mode	Visual
Ferenchak et al. 2022; [41]	Effects of eHMI presence on pedestrians' overall experience	Inform yielding intention	Visual
Guo at al. 2022a; [61]	Effects of different eHMI modalities and locations on pedestrians-AV interaction	Inform yielding intention	Visual
Guo et al. 2022b; [63]	Effects of different eHMI designs on crossing decision	Inform yielding intention	Visual

Haimerl et al. 2022; [58]	Effects of differnet eHMI desings (visual, auditory or both) on pedestrians with intellectual disabilities	Inform yielding intention	Visual and/or auditory
Hensch et al. 2020; [48]	Effects of different eHMIs and driver presence on pedestrian perception	Inform automation status, starting mode and yielding intention	Visual
Hensch et al. 2022; [52]	Effects of eHMI malfunctions on younger and elderly pedestrians' experience	Inform yielding intention	Visual
Hochman et al. 2020; [66]	Learning effects after interacting with AVs of different sizes equipped with different eHMI designs.	Inform intention	Visual
Holländer. et al. 2022; [68]	Compare eHMI with a SmartCurb concept	Inform yielding intention (eHMI) and safe crossing zones (smartcurb)	Visual
Joisten et al. 2021; [79]	Effects of culture, eHMI design and number of pedestrians on crossing behaviour and perception	Inform yielding intention	Visual
Kaleefathullah et al. 2022; [40]	Trust development during interactions with eHMI-equipped AVs showing different timings with respect decelerations	Inform yielding intention	Visual
Kooijman et al. 2019; [60]	Effects of eHMI designs on pedestrian decision and motion behaviour	Inform yielding/nonyielding intention	Visual
Lanzer et al. 2020; [44]	Effects of eHMI with different interaction styles (politeness) on pedestrian compliance, acceptance and trust	Inform nonyielding intention	Visual
Lau et al. 2022a; [73]	Effects of eHMIs in autonomous buses and autonomous cars with different richness levels and dynamic designs	Inform status, intention and perception of the AV	Visual
Lau et al. 2022b; [80]	Effects of interplay between eHMI and kinematics, as well as AV size on willingness to cross, safety and trust	Inform status and/or intention	Visual
Lee et al. 2021; [81]	Negative effects of eHMIs with different designs	Inform yielding intention	Visual
Löcken et al. 2019; [82]	Effects of six different eHMIs on crossing decision and pedestrians' experience. Effects of early vs late presentation of the eHMI	Inform yielding intention	Visual
Loew et al. 2022; [83]	Effects of different eHMIs design on pedestrians	Inform yielding intention and pedestrian recognition	Visual
Matsunaga et al. 2019; [84]	Effects of different eHMI messages and deceleration profiles	Inform yielding intention	Visual
Métayer et al. 2021; [85]	Effects of eHMI presence and designs on pedestrians' crossing decision and experience	Inform status and yielding intention	Visual
Nuñez Velasco et al. 2019; [47]	Effects of eHMI presence and vehicle characteristics on pedestrians 'experience	Inform yielding intention	Visual
Oudshoorn et al. 2021; [86]	Effects of different eHMi concepts on perceived safety	Inform yielding/nonyielding intention	Visual
Rodríguez-Palmeiro et al. 2018; [87]	Effects of driver attentiveness level in an AV on pedestrians' decision and experience	Inform status	Visual
Rouchitsas et al. 2022; [88]	Effects of different eHMIs with anthropomorphic designs on crossing behaviour and perception	Inform pedestrian recognition and intention	Visual
Sahaï. et al. 2022; [89]	Eye contact between drivers in AVs and cyclists/pedestrians/el-scooters	Inform yielding intention	Visual
Şahin et al. 2021a; [90]	Effects of eHMI timing with respect to deceleration	Inform yielding intention	Visual
Şahin et al. 2021b; [69]	Effects of malfunctions in eHMIs on pedestrians' trust	Not specified	Visual
Singer et al. 2022; [91]	Effects of different eHMI designs on pedestrians' behaviour and perception	Inform AV behaviour and intention	Visual
Song et al. 2018; [45]	Effects of eHMI message style on pedestrians' crossing decision and feelings	Inform intention	Visual
Wilbrink et al. 2021; [42]	Effects of static and dynamic eHMIs on pedestrians' experience and behaviour	Inform status, perception and/or intention	Visual

4.2.4 Scenarios

In this state-of-the-art analysis, the description of various scenarios was based on six key criteria: 1) Type of road layout (such as straight, turning, or parking), 2) Number of lanes, 3) Type of environment (urban or rural), 4) Visibility (clear or reduced due to low light, obstacles, etc.), 5) Presence of traffic rules (signalized or unsignalized), and 6) Presence of other pedestrians. This information is shown in Table 4-6.

The most frequently used scenario was a straight road with one or two lanes without traffic signals and without the presence of other pedestrians besides the participant (which was the case in 28 studies). In these scenarios, the participant was asked to decide whether to cross in front of an autonomous vehicle equipped with an eHMI, which would either be yielding or not yielding. The vehicle typically approached from the left side at speeds ranging from 10 to 50 km/h and presented different deceleration profiles along the way (i.e., different deceleration distances and forces). Although some studies only involved the AV, many others also included intermediate vehicles with varying gaps. These interactions usually took place under high visibility conditions and the participants were fully alert. In 14 of the studies, similar scenarios were used, but with added traffic signs (such as crosswalks, yielding signs on the road, or traffic lights) that prioritized pedestrian passage over the autonomous vehicle.

To a lesser extent, other studies considered low-speed scenarios, such as parking lots (which were the focus of 6 studies, such as [46] and [91]), or intersections where a turning AV interacts with a crossing pedestrian. Additionally, [72] included other pedestrian agents to investigate their impact on participants' decisions.

4.2.5 Measures

As depicted in Table 4-6, the majority of studies employed a range of measures to evaluate safety, efficiency, and user experience. These measures can be classified into two categories: subjective and objective. Subjective measures gather information about attitudes, opinions, or judgments explicitly provided by the participants during or after the interaction with the vehicles. On the other hand, objective measures directly observe and record pedestrian's state, attention, and behaviour. This category also encompasses measures in which the pedestrian explicitly indicates their decision to cross by pressing a button.

Subjective measures

The aim of the subjective measures was primarily to evaluate perceptions of safety, understanding of the vehicle's intentions, trust in the vehicle, acceptability, and other aspects related to the system's usability. Typically, these evaluations are carried out either during the experiment, for example, through post-trial scales, or after the experiment, through post-experiment interviews or questionnaires. The most commonly used subjective measures in the reviewed literature are outlined below:

- Likert scales administered after each trial or after the experiment. Such scales are used to quickly and quantitatively assess aspects such as comprehensibility of the vehicle's intention (e.g., [61, 66]) perceived safety (e.g., [49, 66]), trust [40, 79, 92], acceptance and user-experience aspects (i.e., usability, learnability, easy-to-use, etc.). Although to a lesser extent, mental workload has also been assessed in some studies [56, 72]. In multiple cases scales were designed ad hoc, while in others standardized scales were used (e.g., NASA-TLX for workload, Misery Scale for well-being, Van der Laan for acceptance or User Experience Questionnaire scales).

- Button press (e.g., [38, 46, 86]). Participants are required to keep a button pressed for as long as they feel safe and/or would be willing to cross. It provides a continuous measure of their experience through the interaction.
- Slider to indicate willingness to cross the road. As with the button press, this measure allows continuous assessment of pedestrians' willingness to cross the road throughout the interaction (e.g., [50, 76]). The main difference is that the participant moves the slider between a value of 0 (no willingness to cross or not safe at all) and a value of 100 (total willingness to cross or high safety perception), thus providing further information about the intensity of their willingness.
- Interviews. Structured and semi-structured interviews commonly followed up the completion of the experiment for gathering qualitative information about the participants' experience (e.g., [50, 76, 78]).
- Questionnaires. As with scales, questionnaires were administered at different moments of the experiment to assess aspects such as comprehensibility of the eHMI designs and vehicle intentions or participant trust on the vehicle. A widely used questionnaire was the User Experience Questionnaire [53, 76], and more specifically its short version (e.g., [58, 64, 73]).

Objective measures

Objective measures can be subdivided into those derived from pedestrian position, movements or actions (including button presses), and those derived from their visual attention.

Based on pedestrian movement and decisions.

This category includes measures obtained by tracking the pedestrian's position relative to that of the vehicle:

- Frequency of crossings. The number/percentage of times the participants decided to cross.
- Frequency of collisions. The number/percentage of interactions where a collision occurred.
- Crossing initiation time (CIT). Indicates the time at which the pedestrian makes the decision to cross in relation to the position of the vehicle at a particular time. Shorter CITs are associated with better understanding of the autonomous vehicle's messages and more efficient interaction. There are different approaches regarding the reference time point for the calculation of the CIT. The most commonly used are: time at which the trial starts, time at which the AV comes to a complete stop, time at which the AV initiates braking. In the latter case, negative CIT would indicate that the pedestrian decided to cross before braking. Studies also differ on the signal that indicates the crossing decision. The most commonly used are: the moment at which the pedestrian initiates the walking action (using head, arms or ankles as a reference), steps forward [87] or presses a button (most commonly used in test-track studies, e.g., [61, 62]).
- Crossing time (CT) and crossing speed (CS): These measures indicate the time taken by the pedestrian to complete the crossing from the moment he/she starts the crossing, as well as the speed of the crossing. It has been employed as a measure of traffic flow and road safety. Lower CT and higher CS indicate higher traffic flow and understanding of the AV messages, while higher CT and lower CS would reflect a strategy to increase safety in ambiguous situations [71].

- Clearing time (CleT): Time from the pedestrian's first step until he/she leaves the vehicle's path [93].
- Time to Arrival (TTA): Time a vehicle takes to arrive at the area, where a crossing pedestrians' trajectory overlaps with the one of an approaching vehicle [70].
- Post-Encroachment Time (PET): The time that elapses after the pedestrian leaves the point of encroachment, and the vehicle reaches it. It is a safety measure. At constant speeds, TTA equals PET [70].
- Deceleration to Safety Time (DST): The deceleration necessary for a vehicle to avoid collision with a pedestrian. It is a measure of the criticality and controllability of the situation [70].

Based on the visual attention of pedestrians

This category includes measures to indicate pedestrians' visual attention during the moment of interaction, obtained either from eye tracking systems or indirectly, from head movements:

- Head position: although not common, heading position or orientation has been used as a proxy for determining whether pedestrians were attentive to the arrival of the AV or not [48].
- Number and proportion of fixations to specific areas of interest (AOIs) (i.e., relative fixation durations, [57]). It is used as a proxy to determine the information requirements of the pedestrian during the different moments of the interaction. The most commonly used AOIs are: bumper, car side, hood, eHMI, windshield, road surface in front of the car, headlights, grill and wheels (e.g., [50, 71]).
- Fixation duration: indicates mean duration of the fixations. Longer times indicate a higher level of interest, relevance or complexity.
- Dwell time to AOIs: indication of the total time looking at a specific AOI. Longer times indicate a higher level of interest, relevance or complexity.
- Gaze spread: Used by Eisma et al. [77] to measure the spread of participants' visual attention during the interaction with an AV. Low dispersion indicates greater convergence among participants on the elements that are being visually attended to. High dispersion indicates greater difference between participants' attentional focus.

4.2.6 Participants

During the literature review, basic information was collected on the participant samples of the studies, namely, sample sizes, mean ages and variability. Since the HEIDI project intends to develop solutions for different types of users (mainly people of different ages and people with disabilities), we also noted studies where this type of sample was included.

In general, large samples of 30 or more subjects were used in most studies. To a large extent, the sample size varied according to the type of study. Video-based studies used larger samples (e.g., [73, 79, 80, 86]). In the case of Oudshoorn et al. [86] for example, 1141 individuals were recruited (this study is not represented in the graph so as not to distort it). On the other hand, the VR or test-track studies show greater variability, with studies involving few participants (18 in the case of [68] and [72]) to studies with a larger sample (> 50 in the case of [40, 56, 71]).

The vast majority of studies used young and mid-adult populations. The mean age of the samples ranged from 22 to 44 years. However, the variability of ages was also very different, with studies using age-homogeneous populations and others with more dispersed ages. Although in some studies the age range included younger people or older adults (e.g., [69, 73, 80]), very few analysed the age effect as factor, but mostly as a covariate. Specifically, only [52] and [84] compared older adults (>65 years) with other younger age groups. Additionally, [84] also included a sample of children aged 8 to 10 years. Finally, regarding the inclusion of populations with disabilities, only [58] conducted a study comparing 54 participants with disabilities to 56 participants without disabilities in their perception of different eHMI designs.

It should be noted that some studies have also analysed cross-cultural differences by including participants from different countries. For example, [79] compared Chinese participants with German participants, while [91] included participants from South Korea, China, and the United States.

4.2.7 Experimental design

The most commonly used experimental design was the repeated measures design with 2 or 3 independent factors, and 2 to 6 levels each. However, a number of other studies also used mixed designs by including between-subject factors such as age group [84], type of vehicle communication [56], pedestrian position [46] or nationality [79], among others. Typically, studies included a baseline or control condition against which to compare the effect of the eHMI or of its different designs.

The number of experimental conditions varied widely from 3 or 4 (e.g., [41, 78]) to about 15-20 (e.g., [38, 71]), directly influencing the duration of the experiments. Often, each condition was presented more than once within the same block of trails or in different blocks, thus controlling for intra-individual differences. Moreover, yielding trials often were interspersed with nonyielding trials. to increase realism and prevent participants from predicting the vehicle's behaviour. In various studies, yielding behaviour was also used as an independent factor, meaning that pedestrians’ experience or behaviour was compared between yielding and nonyielding trials. To avoid learning or fatigue effects resulting from the numerous trials, the conditions are generally counterbalanced across participants.

Table 4-6: Aggregation of sample characteristics, experimental design, scenarios used, constructs evaluated, and measures used.

Reference	Test environment	Scenario/s	Constructs evaluated	Measures used	No. Participants	Experimental design
Bindschädel et al. 2022a; [57]	VR simulator	Straight one-lane road; Urban; Clear visibility; unsignalised; with and without other pedestrians	Safety and efficiency	**Objective:** Crossing initiation time (from the moment the AV started braking); Glance behaviour (absolute and relative fixation durations to different AOIs) **Subjective:** Scale for perceived safety	N= 30 Mean age: 26.53 (SD= 10.08, 23-59)	**Scenario "Pedestrian Group Behaviour":** 2x2 Within-subject: eHMI presence and group of pedestrians **Scenario "Timer":** Within-subjects: Presence of a timer on eHMI **Scenario "Non-yielding automated vehicles":** Within-subject: 3 exposures (learning effect)
Bindschädel et al. 2022b; [56]	VR simulator	Straight one-lane urban road; Clear visibility; unsignalised; no other pedestrians	Safety, efficiency and mental workload	**Objective:** crossing initiation time (starting from when the AV started braking), eye tracking measures (nr fixations, mean fixation time and nr of saccades **Subjective:** scale for perceived safety	N= 53 Mean age: 33.5 (SD=6.14, 22-56)	3x3 Mixed: vehicle type and communication type (between-subjects)
Bindschädel et al. 2021; [71]	VR simulator	Straight one-lane road; Urban; Clear visibility; unsignalised; no other pedestrians	Safety	**Objective:** Crossing initiation time (from the moment the AV started braking); **Subjective:** Scales for trust and acceptance	N= 51	2x5x2 Within-subjects: blocks, eHMI, scenario
Burns et al. 2019; [46]	Test-track (indoors, low-speed AV)	3 scenarios: Simulated parking, simulated turning into a lane; no traffic rules	Safety, usefulness, satisfaction	**Subjective:** press a button while feeling safe to cross.	N= 34	2x2 Mixed: pedestrian position and eHMI design
Colley et al. 2022a; [72]	VR simulator	Straight two-lane urban road: unsignalised; other pedestrians (in one of the scenarios)	Workload, comprehensibility and trust	**Subjective:** scales for Workload, predictability/understandability and trust	N= 18 Mean age: 31.28 (SD=10.53, 21-56)	Within-subjects: 3 scenarios with different eHMIs
Colley et al. 2022b; [43]	Videos on-line (virtual scenario)	Straight one-lane road; Clear visibility; unsignalised; no other pedestrians	Workload, comprehensibility and trust	**Subjective:** scales for Workload, predictability/understandability and trust	N= 59 Mean age: 35 (SD=11.82)	3x2x2 Within-subjects: vehicle appearance, mode distinctiveness and passenger visibility
de Clercq et al. 2019; [38]	VR simulator	Straight two-lane road; Clear visibility; unsignalised; no other pedestrians	Safety and efficiency	**Subjective:** Button press when feeling safe to cross; post-experiment questionnaire	N= 28 Mean age: 24.57 (SD= 2.63)	2x2x4x3 Within-subjects: Type of vehicle, yielding behaviour, eHMI design and timing eHMI-deceleration
Deb et al. 2018; [59]	VR simulator	Straight two-lane urban road; Clear visibility; unsignalised; no other pedestrians	Safety and efficiency	**Objective:** crossing duration and waiting time (not defined) **Subjective:** scales for evaluating eHMI features	N= 30 Mean age: males= 30.65(22-47), females= 31-62(18-47)	4x4 Within-subjects: visual eHMI features and auditory eHMI features

Dey et al. 2020a; [53]	Videos (real world + WOz)	Straight one-lane road; Clear visibility; unsignalised; no other pedestrians	Efficiency and UX	**Objective:** slider to inform willingness to cross **Subjective:** scales for UX	N= 26 Mean age: 24.7 (SD=5.2)	4x3 Within-subjects: eHMI information and yielding intention (including fake nonyielding)
Dey et al. 2020b; [50]	Test-track (WOz)	Straight one-lane road; Clear visibility; unsignalised; no other pedestrians	Efficiency and UX	**Objective:** slider to inform willingness to cross; glance behaviour (fixations to different AOIs) **Subjective:** interviews	N= 26 Mean age: 26.21 (SD=3.74)	2x4 Within-subjects: eHMI presence and deceleration pattern
Dey et al. 2022; [76]	Videos (real world)	Straight one-lane road; Clear visibility; unsignalised; no other pedestrians	Efficiency and UX	**Objective**: slider to inform willingness to cross **Subjective**: scale for UX	N= 25 Mean age: 30.52 (SD=16)	Within-subjects: eHMI presence/explicit/implicit
Dey et al. 2021; [54]	VR simulator	Straight one-lane road; Clear visibility; two pedestrians; crosswalk for one of the pedestrians	Efficiency and UX	**Objective**: button press to inform willingness to cross **Subjective**: scale for UX	N= 36 Mean age: 23.3 (SD=2.6)	Within-subjects: 4 eHMI designs
Dietrich et al. 2019; [70]	VR simulator	Straight two-lane road; Clear visibility; unsignalised; no other pedestrians	Safety and efficiency	**Objective:** Time to Arrival, Post encroachment time; deceleration to safety time. **Subjective:** open questions	N= 32	6x2 Within-subjects: Deceleration moment and eHMI presence
Eisma et al. 2020; [77]	Videos (virtual scenarios)	Straight two-lane road; Clear visibility; unsignalised; no other pedestrians	Efficiency	**Objective:** Gaze spread **Subjective:** Scales for clarity, button press to inform perceived safety to cross	N=61 Mean age = 23 (SD=1.8)	6x6 Within-subjects: 6 eHMi locations and 6 virtual environments
Eisma et al. 2021; [62]	Videos (real world)	Straight two-lane road, T-junction and an intersection; Clear visibility; unsignalised; no other pedestrians	Efficiency	**Objective:** Willingness to cross (button press Y or N), decision time; ocular measurements (pupil diameter and number of saccades) **Subjective:** scale for clarity of the AVs' intention	N= 103 Mean age: 23.3 (SD=2)	2x2x3 Within-subjects: eHMI perspective, yielding intention, memory task
Epke et al. 2021; [39]	VR simulator	Straight two-lane urban road; Clear visibility; crosswalk; no other pedestrians	Efficiency	**Objective:** Crossing intention (One step forward), frequency of hand gestures to the AV. **Subjective:** questionnaires during/after experiment	N= 26 Mean age= 26 (SD=3.7)	2x2x2 Within-subjects: opportunity to use hand gestures, recognition, yielding behaviour
Faas et al. 2019; [67]	Test-track (WOz)	Street-crossing (two-lanes) scenario and parking lot; Clear visibility; no other pedestrians; unsignalised	Efficiency	**Subjective:** questionnaires and structured interviews; Scales for visibility, saliency, discriminability, recognition, attractiveness, suitability, sense of safety and trust	N=59 Mean age=42.98 (SD=14.85)	2x2 Within-subjects: eHMI colour and traffic scenario
Faas et al. 2020a; [64]	Test-track (WOz)	Intersection; Clear visibility; unsignalised; no other pedestrians	Safety and efficiency	**Objective:** Crossing initiation time (start walking from when the AV starts braking), crossing duration. **Subjective:** scale of UX	N= 30 Mean age: 42.97 (SD= 15.37)	Within-subjects: yielding lights.
Faas et al. 2020b; [51]	Videos (real-world)	Straight two-lane urban road; Clear visibility; yield sign for AV; no other pedestrians	Efficiency and UX	**Objective:** Crossing initiation time (one step from when the AV starts braking). **Subjective**: scales for safety, trust, acceptance and UX	N= 34 Mean age: 41.5 (SD=15.8, 22-69)	Within-subjects: eHMI design (no, status and status+intent)
Faas et al. 2020c; [78]	Test-track (WOz)	Intersection and parking lot; Clear visibility; unsignalised; no other pedestrians	Safety, efficiency and UX	**Objective:** crossing initiation time (from when the AV fully stopped), crossing duration **Subjective:** scales for trust, perceived safety, UX, perceived intelligence and transparency**,** structured interview	N= 52 Mean age: 42.98 (SD=14.85, 18-66)	2x5 Mixed: traffic scenario and eHMI design

Faas. et al. 2021; [49]	Test-track (WOz)	Straight two-lane urban road; Clear visibility; crosswalk; no other pedestrians	Safety and efficiency	**Objective:** crossing initiation time (start walking from when the AV stopped) **Subjective:** scale for perceived safety	N= 65 Mean age: 43.1 (SD=15.6, 18-69)	2x3 within-subjects; eHMI presence, driver state
Ferenchak et al. 2022; [41]	VR simulator	Straight two-lane urban road; Clear visibility; crosswalk; no other pedestrians	UX	**Subjective:** post-trials scales for trust, comfort, safety, acceptance and understanding	N= 47 Mean age: 23.5 (SD= 5.54, 18-39)	Within-subjects: 4 eHMI designs
Guo at al. 2022a; [61]	Videos (virtual scenarios)	Straight one-lane urban road; Clear visibility; unsignalised; no other pedestrians	Safety and efficiency	**Objective:** Crossing initiation time (button press from when the AV appears), eye tracking measures (dwell time and number of fixations). **Subjective:** scale for clarity	N= 62 Mean age: 25.62 (SD=3.47, 19-37)	6x3 Within-subjects; eHMI modalities and eHMI locations
Guo et al. 2022b; [63]	Videos (virtual scenarios)	Straight two-lane urban road; Clear visibility; crosswalk; no other pedestrians	UX	**Subjective**: scales for UX	N= 90 Mean age: 37.24 (SD=2.44)	Within-subjects: 6 eHMI designs
Haimerl et al. 2022; [58]	Videos (virtual scenarios)	Straight two-lane urban road; Clear visibility; unsignalised; no other pedestrians	UX	**Subjective:** post-trial scales (using smileys)	N= 120, 54 with and 56 without intellectual disability. Mean age: With ID= 35.9 (SD 12.98) Without ID = 32.2 (SD=10.6)	2x4 Mixed: Intellectual disability and eHMI modality
Hensch et al. 2020; [48]	Real-road (WOz)	Unspecific: low-speed parking area	UX	**Subjective**: scales and closed-ended questions	N= 173 Mean age: 29 (SD= 10.6)	2x4 Between-subjects: driving condition and light signals
Hensch et al. 2022; [52]	Videos (real world)	Parking area; 2 scenarios: AV passing by or turning in front of the pedestrian; no other pedestrians	UX	**Subjective:** scales and questionnaires for UX	N= 19 young; 17 elderly Mean age: young=30.47 (SD=4.65); elderly=71 (SD=3.87)	3x2 Mixed: Failure moment and Age (between-subjects)
Hochman et al. 2020; [66]	Videos (virtual scenarios)	Straight two-lane urban road; Clear visibility; crosswalk; no other pedestrians	Safety and efficiency	**Objective:** Crossing initiation time (button press from when trial started), eye tracking measures (total fixation duration, total number of fixations. **Subjective:** scales for comprehension and open interview	N= 20 Mean age: 26 (SD=3, 21-34)	2x2x2x2x2x2 Within-subjects: eHMI presence, message type, modality, car size, color and car distance.
Holländer. et al. 2022; [68]	VR simulator	Straight two-lane urban road; Clear visibility; unsignalised; no other pedestrians	Safety and efficiency	**Objective:** Crossing duration, collisions. **Subjective:** scale for safety perception	N= 18 Mean age: 34(19-64)	3x3 Within-subjects: Traffic situation and communication solution (baseline, eHMI and *smartcurb)*
Joisten et al. 2021; [79]	Videos (virtual scenarios)	Straight two-lane urban road; Clear visibility; crosswalk; no other pedestrians	Safety and efficiency	**Subjective:** scales for trust and willingness to cross	N= 205; 126 Germans and 79 Chinese Mean age: Germans: 25.45 (SD= 4.7, 18-	2x2x3 Mixed design: Country (between-subjects), eHMI design and pedestrian group size

					57); Chinese: 34.08 (SD=10.47; 18-55)	
Kaleefathullah et al. 2022; [40]	VR simulator	Straight one-lane urban road; Clear visibility; unsignalised; no other pedestrians	Trust, comprehensibility	**Objective:** Crossing initiation time (start walking from when the trial starts) and pedestrian postion. **Subjective:** scales for perceived risk, comprehension and trust.	N= 60 Mean age: 24.4 (SD=4, 18-35)	2x2x2 Within-subjects: Yielding behaviour, eHMI presence and eHMI timing
Kooijman et al. 2019; [60]	VR simulator	Straight two-lane urban road; Clear visibility; crosswalk; no other pedestrians	Safety	**Objective:** forward gait velocity, moment of leaving curb, thorax angles. **Subjective:** fear scale	N= 24 Mean age: 25.4 (SD=2.5, 21-30)	3x3x2 Within-subjects: eHMI conditions, gap distance, yielding condition.
Lanzer et al. 2020; [44]	Videos on-line (virtual scenarios)	Straight two-lane urban road; Clear visibility;with and without crosswalks; no other pedestrians	UX	**Subjective:** scales for acceptance, affective state, compliance, fear, power distribution and trust	N= 90 Mean age: 22(18-60)	2x2x3 Mixed design: country (between-subjects), scenario and communication strategy
Lau et al. 2022a; [73]	Videos on-line (virtual scenarios)	Straight one-lane urban road; Clear visibility; unsignalised; no other pedestrians	UX	**Subjective:** scales for perceived safety, affectiveness and usability	N= 155 Mean age: 35.21 (SD=15.29, 16-77)	3x3x2 Mixed design: Information richness, eHMI dynamic design and vehicle size
Lau et al. 2022b; [80]	Videos (virtual scenarios)	Straight one-lane urban road; Clear visibility; unsignalised; no other pedestrians	UX	**Subjective:** scales for comprehensibility, safety and trust	N= 149 Mean age: 35 (SD=12.68, 19-71)	2x2x3 Within-subjects: yielding intention, vehicle size and eHMI presence/dynamic level
Lee et al. 2021; [81]	VR simulator	Straight two-lane urban road; Clear visibility;with crosswalk; no other pedestrians	Safety	**Objective:** Looks to the left road; number of stopping and collisions	N= 57 Mean age: 30.39 (29-45)	Between-subjects: 3 eHMIs
Löcken et al. 2019; [82]	VR simulator	Straight two-lane urban road; Clear visibility; unsignalised; no other pedestrians	Safety and UX	**Objective:** Crossing duration (i.e., seconds it took to cross after the AV started to break.) **Subjective:** scales for trust, safety and UX	N= 20 Mean age: 27 (SD=7.2)	6x2 Within-subjects: eHMI designs and Timing
Loew et al. 2022; [83]	Test-track (WOz)	Straight one-lane urban road; Clear visibility; unsignalised; no other pedestrians	Safety and efficiency	**Objective:** Crossing initiation time (start walking when the AV crosses a grid). **Subjective:** scales for acceptance and perceived safety.	N= 30 Mean age: 24.53 (SD= 2.37, 19-30)	Within-subjects: eHMI information (3 levels)
Matsunaga et al. 2019; [84]	Test-track	Straight two-lane urban road; Clear visibility; crosswalk; no other pedestrians	Efficiency	**Objective:** Button press when detecting yielding intention. **Subjective:** scale for certainty	N= 56; 27 non-elderly (18-64), 14, elderly (>65), 15 elementary school students (8-10)	4x2 Within-subjects: eHMI messages and deceleration
Métayer et al. 2021; [85]	Video game (Joystick)	Straight two-lane urban road; Clear visibility; crosswalk and traffic signs; no other pedestrians	Efficiency and UX	**Objective:** Crossing frequency. **Subjective:** scales for acceptance and UX	N= 49 Mean age: 41.02 (SD=12.03)	2x2x3 Within-subjects: eHMI presence, pedestrian crossing behaviour and eHMI design
Nuñez Velasco et al. 2019; [47]	360° Videos (real world)	Straight two-lane urban road; Clear visibility; unsignalised; no other pedestrians	UX	**Subjective:** scales for pedestrian behaviour, trust, perceived control	N= 55 Mean age: 24.9 (SD=3.5, 21-37)	2x2x2x2x2 Within-subjects: vehicle type, crossing facility, vehicle speed, gap size and eHMi presence

Oudshoorn et al. 2021; [86]	Videos on-line (virtual scenarios)	Straight two-lane urban road; Clear visibility; unsignalised; no other pedestrians	Safety	**Subjective:** Button press while feeling safe to cross	N= 1141 Mean age: 37.4 (SD= 11.45)	6x3 Within-subjects: eHMI design and Vehicle state/intention
Rodríguez-Palmeiro et al. 2018; [87]	Test-track (WOz)	Straight two-lane track; Clear visibility; unsignalised; no other pedestrians	Safety, efficiency and UX	**Objective:** backward step as a proxy for critical gap (vehicle distance at the step moment9 **Subjective:** scales for driving mode, realism, speed perception. Post-experiment questionnaire	N= 24 Mean age: 24.5 (SD=2.95, 19-30)	2x3x3 Within-subjects: eHMI: driving mode, sign conditions, driver attention. 5x2x2 Within-subjects: appearance, stopping condition and approach direction
Rouchitsas et al. 2022; [88]	Videos (virtual scenario)	Straight one-lane urban road; unsignalised; no other pedestrians	Efficiency	**Objective:** Decide to cross or not (performance evaluated based on accuracy guessing AV's intention)	N= 30 Mean age: 33.1 (SD= 11.9)	2x2x7 Within-subjects: gender, gaze direction, facial expression
Sahaï. et al. 2022; [89]	Survey using images	Straight one-lane urban road; crosswalk; no other pedestrians	Efficiency	**Subjective:** determine whether eye contact would be necessary to cross (yes/no)	N= 462 Mean age: 43.25 (18-90)	4x5x4 Mixed design: Driver state, eHMI design and age (between-subjects)
Şahin et al. 2021a; [90]	Video game	Straight road; Clear visibility; unsignalised; no other pedestrians	User experience	**Objective:** Crossing initiation time (from the moment the vehicle appeared **Subjective:** Scales (comprehensibility, ease, trust, safety)	N= 20 Age range: 19-59	3x2 Within-subjects: eHMI timing and automation level
Şahin et al. 2021b; [69]	VR simulator	N.S.	Trust	**Subjective:** scales for pedestrian behaviour, pedestrian receptivity and scenario-based questionnaire	N= 7 Mean age: 41 (SD=25.5, 21-85)	2x2 Mixed: Lighting conditions and prior experience with a collision
Singer et al. 2022; [91]	VR simulator	3 different interactions in a parking area. AV coming from front, left and right; Clear visibility; Unsignalised	Safety and efficiency	**Objective:** pedestrians' position. **Subjective:** scales for intention recognition and perceived safety	N= 90 from China, S.Korea and US Mean age: S.Korea: 32.8 (SD= 6.5); China: 36.1 (SD=7.6); US: 40.7(SD=17)	3x3x3x3 Mixed: Nationality (between-subjects), vehicle position, vehicle state, side display
Song et al. 2018; [45]	Videos on-line (real-world)	Straight one-lane urban road; Clear visibility; unsignalised; no other pedestrians	Safety and efficiency	**Objective**: frequency of crossing decision **Subjective:** scales for preference	N= 125 Mean age: 29 (SD=12.26, 15-59)	Within-subjects: eHMI presence and communication style.
Wilbrink et al. 2021; [42]	VR simulator	Straight two-lane urban road; Clear visibility; unsignalised; no other pedestrians	Safety, Efficiency, comprehensibility	**Objective:** Button press when crossing decision, gaze behaviour (based on head movement). **Subjective:** scales for UX, feelings, safety perception	N= 62 Mean age: 33.19 (19-60)	3x3 Mixed: eHMI dynamic level and eHMI dynamic information

4.3 Co-simulation of drivers and pedestrians

From the presented literature in section 4.1 and 4.2 about adaptive internal or external HMI it is seen that driving and pedestrian simulators provide a valuable set of tools and methods for looking at human behaviours and interactions. One drawback of using only one of these simulators is that a behaviour model needs to be used for the other traffic participants limiting natural interaction. As these simulators become more available and connected, the idea to connect two such simulators to look at interactions between drivers and vulnerable road users has gained interest which can be seen from the found papers starting to appear [94-100].

For the HEIDI project there is an interest to look at what is important in the interaction between traffic participants such as pedestrians and vehicles. Co-simulation is an upcoming field of research and here we present the current state that can be useful for HEIDI.

4.3.1 Paper selection and aggregation

As a starting point a selection of articles was analysed to collect keywords describing a co-simulation study with a driving and pedestrian simulator. The following terms were selected for the literature search: multi-user, multi-driver, multi-vehicle, multi-agent, coupled, co-simulation, driver-pedestrian, human-human, linked, networked, and distributed. These terms were then used together with the following terms to narrow the search further: driver, driving, pedestrian, simulator, behavior, behaviour, and interaction. The literature search was performed in January 2023 using Scopus and resulted in a list of 477 full-text articles, conference papers, and book chapters. This list was then screened using the following criteria:

Included:

1. Experimental studies where test participants were used.
2. Experimental setups using at least one driving simulator co-simulated with at least one pedestrian simulator.
3. Both the participating driver and pedestrian are controlled by humans.
4. Experimental setups using at least one driving simulator co-simulated with at least one bike simulator.

Excluded:

1. Papers that only contained a technical description of a co-simulation setup and no methods were used.

In total, only 7 papers were found through Scopus where there was either direct access to the papers through Scopus or the papers could be accessed through the authors via ResearchGate. Their publication year is illustrated in Figure 4-3. The primary reason that so many papers were rejected was using a broad search that included many other topics.

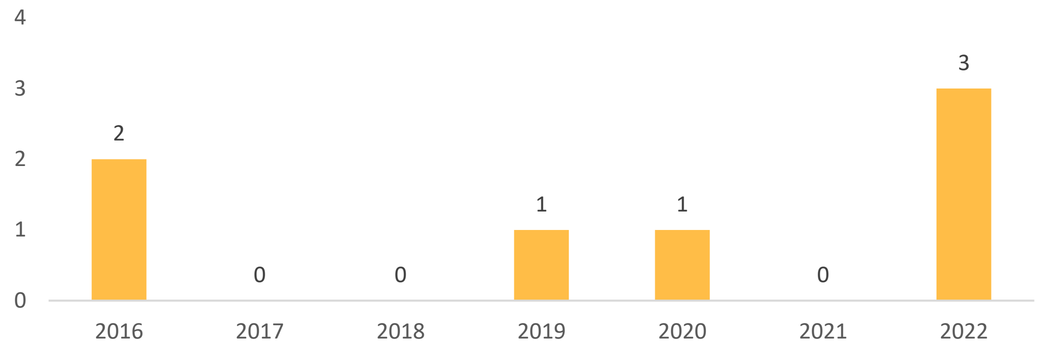

Figure 4-3: Year of publication of the reviewed co-simulation studies.

Only 7 papers are a small selection and some of them present the same study with a focus on different aspects. Grouping these studies based on the different research groups that performed them results in four different research groups. Here, the Institute of Ergonomics at the Technical University of Munich was the earliest lab to look at car to pedestrian interactions using a co-simulation setup [94], [95], and [96]. Some years later the Institute for Transport Studies at University of Leeds presented a study using a CAVE simulator connected to a driving simulator [97]. In 2020 the Department of Cognitive Robotics at Delft University of Technology started to look into co-simulations and in 2022 they presented a driver-pedestrian study [98]. Also from 2022, there are two publications from Traffic Engineering and Control at Technical University of Munich where a bike simulator was connected to a driving simulator [100], and [99].

Looking at this literature search we notice that there are further studies that aren’t captured. For example, a study performed at the Institute of Ergonomics at the Technical University of Munich published results already in 2015, see [101]. The reason for this paper not being found is a typo in the Scopus abstract. As this study uses the same or similar simulator setups and methods as presented in the other papers from the same research group, we chose not to continue the literature search and add these further papers but instead focus on the results from the found four research groups.

4.3.2 Test environment

Among the found papers there were only studies using simulators although studies performed in real environments were not automatically discarded. The used simulators were generally driver and pedestrian simulators. In one study a bicycle simulator was used [99, 100]. The simulators were connected as pairs using different network solutions. Pedestrian simulators used either a computer and keyboard solution, VR equipment, or a motion tracking system where the participant is dressed in a suit containing sensors at different body joints. The pedestrians in the virtual environment were mostly generating movements for the body and head. When a motion tracking system was used larger freedom of motion was possible and the pedestrians could move freely and use body language. One of the pedestrian simulators had a CAVE design where the pedestrian can move freely in a room where the surroundings are presented on the floor and the walls. This pedestrian was then represented in the graphical environment as spheres representing the locations of the sensors. The idea was to reduce the “uncanny valley” problem [102] while still give the pedestrian a large set of motion freedom.

Driving simulators ranged from a desk with three monitors and a touchpad interface where the passenger in the AV could control its decisions to smaller static simulators consisting of multiple screens presenting surrounding environment, a real car seat, steering wheel, and pedals. In between simulators used a desktop with gaming equipment for steering and pedals. None of the found studies used a larger moving base driving simulator although it was mentioned as future work.

The bike simulator used a stationary normal bike with added sensors for steering angle and rotational velocity of the rear wheel of the bike. In front of the bike a screen presented the surrounding world to the cyclist.

4.3.3 Participants

From the selected papers it was seen that the studies always were performed at different universities. When recruiting test participants for these studies it was common that participants for both the pedestrian and driving simulators were recruited in the vicinity of the university. As a result, the recruited participants were often young (mean age from the studies were: 31, 24, 33, 42, 24, and 23) and healthy (no reported disabilities). When describing the test participants, it was often done separately with one section for the drivers and another one for the pedestrians, although this was not always the case. Most of the recruited drivers had a valid driver’s license, but again not always (for example when operating a vehicle with a high degree of autonomy).

4.3.4 Scenarios

The found studies from the literature all present situations where different traffic participants need to interact to safely resolve a given traffic situation. Here, the pedestrian and driver interaction studies [94-98] used a scenario where a pedestrian wants to cross a road with usually one oncoming vehicle, although there were situations with more than one vehicle. The crossing could occur at a zebra crossing or an unmarked part of the road where the pedestrian would be clearly visible or a bit hidden by an occluding object. No scenario contained any traffic signals for how to prioritise the different traffic participants. In most of the studies the pedestrian could move its head and body to interact with the vehicle. One study [98] added a visual cone from the pedestrian (and the driver) to visualize gaze as one mean of interaction and in another study where the pedestrian was equipped with sensors to move freely, he or she could interact using body language. It was also common to include controlled pedestrians using a behavioural model for how to cross the road so that a driver wouldn’t know which type of pedestrian it would be [94-96]. In a similar way the vehicles were often either driven by humans or automated so that the pedestrian didn’t always know what type of driver it would interact with.

For the study containing a driving simulator and a bike simulator [99, 100], the scenarios were also crossing scenarios where the bike was either turning or going straight and by doing so crossed the path of the vehicle. In the scenario the bicyclist was generally free to bike wherever desired but was guided by a mobile phone application and the automated vehicle followed a specified path. In the automated vehicle in the same study a human was positioned as a passenger who could give instructions to the vehicle to change its behaviour. In this crossing scenario there were different ways for how the traffic was prioritized (for example, the AV could enforce the traffic rules or could be passive and yield to the bicyclist) so that the bicyclist and the passenger in the AV needed to interact to safely solve the traffic situations.

One specific aspect of co-simulation is the need for synchronisation of the participants prior to the point of interaction. Examples of how this was solved are: use a test leader as one of the drivers or pedestrians where the test leader can wait or get into position until the participant is in the right position, use an AV with a defined path with a speed controller so that it controls its movements according to another participant but still can be interacted with, or start the scenario very close to the point of interaction. It was also remarked in one study [94] that it is important that participants are given sufficient time to interact. If the time to interact is too short the resulting behaviour is more reactive than interactive.

4.3.5 Measures

All the studies found on co-simulation look at safety and as such many measures are like the ones for the studies on adaptive iHMI and eHMI. Generally, the co-simulation studies look at both subjective and objective data. Often the subjective data concerns self-assessment (both own performance but also how well understood the other traffic participant were), the perception of the used simulator (including extra systems if available, for example HMI solution), and questions regarding demographics. Often these questions were answered on a scale between 1-5 where commonly average values were reported.

Collected objective data generally tried to answer how critical the situation was. As such, examples of data collected were: vehicle and pedestrian velocities and accelerations [97]; time-to-arrival (TTA) [94]; time-to-collision (TTC) [96]; post encroachment time (PET) [96]; braking pressure [94]; deceleration to safety time (DST) [94, 96] with a 3 seconds margin; hold button as long as feeling safe [98]; and decision duration when making right of way decisions and if the decision was correct [99]. It was remarked that the measure DST was quite different between a computer-controlled bot and a human pedestrian in one study [94]. Different gaze measures from both the vehicle and pedestrian perspective were also evaluated, such as time-to-first-fixation [96], fixation-to-braking-time [96], gaze in bounding box (total time when driver looked at the pedestrian) [96], and gaze yaw angle [98].

The measures so far focus on one of the involved participants perspectives from the studies. There are two examples of measures that try to look at the interaction between the traffic users. The first ones are based on the cross-correlation coefficient (CCC) [96]. Considering both the vehicle and pedestrian velocities as time series, the correlation between them is evaluated. Thus, CCC describes the relationship between the driver and the pedestrian as -1 if they are perfectly negatively correlated, 0 if there is no relation between them and 1 if they are perfectly correlated. The time series are also shifted to see which shift or lag gives the best correlation. This can be seen as the time it takes for one participant to adapt to the other participant.

The other measures that investigate interaction were from cross recurrence plots [95]. Here, two time series from two systems are visualized in the same plot and analysed in the same phase space where the structures in these plots are analysed (vertical lines, diagonal lines, microstructure, and overall appearance).

It should be noted that it was reported that CCC gave a different time lag for a computer-controlled bot and a human. The conclusion was that the bots were dominant in the traffic situations by just walking out while the human pedestrians were more cautious letting the drivers be the dominant users. Thus, in this study there was a clear difference in the behaviour. Similar conclusions regarding computer-controlled bots and humans were also found in the study using cross recurrence plots.

4.3.6 Experimental design

One important aspect of co-simulation studies is the need for additional participants and personnel. Running two simulators at the same time usually require twice the number of participants. In most of the studies participants were either pedestrians or drivers in paired groups. No study shifted between drivers and pedestrians. Normally, a within-subject design was applied within each group, so for example all pedestrians would experience every situation with the oncoming driver. It is also mentioned that handling of two simulators is more complex and technically more challenging than handling only one [96].

When the scenarios were run the simulators were located at different locations. This was to prevent the participants to communicate in any other way than the one desired within the virtual environment. Here, participants could be prepared separately at the different simulators, or they could be briefed in the same room before going to the respective simulator. Unfortunately, the selected papers don’t provide details about how this was approached and thus, it is unclear how aware participants are of each other.

Another way to create a co-simulation was to use a test leader to control one of the simulators [94-96] (in the selected papers it was the pedestrian that was controlled by the test leader). This reduces the number of participants needed while still having a balanced within-subjects study and in every found study where a test leader controlled one simulator, a within-subject design was used.

5. Main findings and considerations for the HEIDI project

The main objective of the HEIDI project is to develop a fluid cooperative HMI that guarantees safe, smooth and comfortable interaction between pedestrians and drivers (Objective 3). In addition, another objective of HEIDI is to provide recommendations for methodologies to validate such technologies (Objective 4). To achieve these objectives, the implementation of appropriate and effective evaluation methods is required. Given the novelty of such systems, there are unfortunately no standardised validation methods available, yet. Instead, a wide variety of different methodologies and techniques have been applied across multiple studies, each with its own merits and demerits. Therefore, as a first step, it was decided to carry out a SOA review of different methods employed in the literature. In this respect, it is necessary to underline that this SOA did not cover the results of the studies with respect to the effectiveness of the different HMIs. This SOA will ideally serve as a basis for future discussions within the HEIDI project, namely in WP2 (iHMI), WP3 (eHMI) and WP4 (cooperative HMI), where different conceptualisations of internal and external HMIs will be tested.

Although the ultimate goal of the HEIDI project is to develop a holistic system where eHMI and iHMI communications respond to the same strategy, no literature exists that includes both sub-systems jointly. Instead, the different studies analysed the effectiveness of iHMIs or eHMIs independently. Moreover, there are still only few co-simulation studies where interactions involving various road users have been investigated. As mentioned, this type of setup is necessary for a holistic user evaluation of the overall HEIDI system (i.e., eHMI+iHMI). Such a constraint led us, therefore, to break down the SOA into three well-defined parts: 1) interaction between adaptive iHMIs and drivers, interaction between eHMIs and pedestrians, and co-simulation studies. In the following, we will discuss the main findings highlighted in the different reviews and emphasise aspects that should be considered in the HEIDI project by the responsible WPs for the development of the HEIDI system subcomponents.

5.1 Adaptive iHMIs

The main findings from the literature review on adaptive internal HMIs are that:

1. Only a few of the adaptive iHMIs found were adapted to the driver in real-time. Because of the NCAP requirements on new cars to not only detect driver distraction but also to intervene [27], driver monitoring will however be necessary in modern cars.
2. Most studies have been carried out in driving simulators. Driving simulators allow for control over events and situations and the internal HMI. Functionality that is not yet at hand in real cars can be tested within a safe environment. Using simulators also makes it easier to compare identical situations between participants, although most studies on adaptive iHMIs were carried out with a within-subject design, to evaluate a certain system.
3. The goal of many studies has been to evaluate iHMIs in terms of their efficiency and how well they are accepted and trusted by drivers.
4. Overall, the purpose of the iHMI has been to contribute to increased situation awareness.
5. The most common input modality used for a driver to communicate with the iHMI was tactile, such as pressing a button or touching a touchscreen. However, visual and auditory input have also been used.
6. The major output modality of the iHMI was visual. The visual component could be supplemented by other modalities, for example by auditory or tactile information.

7. There was a spread of scenarios in the literature review, where both traffic environment and situation have been altered.

Some aspects that should be considered in the HEIDI project as a consequence of the results from the literature review on adaptive iHMIs are as follows:

1. Driver monitoring: Monitor driver state in real-time. In the HEIDI project the cooperative HMI should be adapted to humans both inside and outside of the vehicle and adaptation to driver state is vital. Distraction is one driver state that should be observed. A recommendation is to use a system for eye tracking, which corresponds to the EuroNCAP requirements of driver gaze [27].
2. Participant drivers: Include elderly drivers. To fulfil objective 1 of HEIDI, where iHMIs should be developed especially for elderly drivers by providing support and tutoring where needed, elderly drivers should logically be part of the participant group. The definition of elderly needs to be thought through, but previous studies on adaptive internal HMIs have only used participants of up to 65 years of age, and the EuroNCAP [27] requires that people of at least 80 years of age have been tested for driver state monitoring assessment.
3. Experimental conditions: Using a within-subject design for evaluating the HEIDI HMIs would make it possible to compare different solutions. A mixed design, where drivers are recruited in two or more groups, would allow for comparisons between different groups, e.g., middle aged and elderly drivers. Depending on the goal of the study, a pure within-subject design or a mixed design should be chosen.
4. Scenarios: The scenarios in HEIDI should reflect the objectives and allow for evaluating the cooperative HMI concept for the specific groups and situations as described in the use cases that will be defined within the project (D1.2).
5. Use of measurements: Both objective and subjective measurements should be considered in the HEIDI project. While objective measurements are easy to quantify and calculate, subjective measurements indicate qualitatively whether a system is accepted or not, by use of scales, questionnaires, or interviews. Among objective measurements used in previous studies, eye-tracking data such as glance frequency and glance duration are of interest to assess distraction and situational awareness. It is important that some safety measure or indicator is used in the simulator studies, e.g., time-to-collision (TTC), and physiological measures such as heart rate variability and skin conductance could potentially be used for assessing mental workload and in the end the usability of the HMI.

5.2 eHMIs

As derived from the literature review, research on the effectiveness of eHMIs has accelerated in the past few years, continuing an exponential trend. Over this short period a number of promising results have been reported on the use of eHMIs to improve vehicle-VRU interactions, especially for pedestrians. Furthermore, a variety of methodologies with various scenarios, experimental designs or measures, among others, have been explored. Yet, a standardised approach to validate such technology is still lacking and a number of aspects still need to be considered. The review has identified some relevant aspects which are described below:

1. Almost the entire literature has focused on analysing interactions between pedestrians and fully autonomous vehicles equipped with eHMIs. However, no attention has been devoted to the potential of eHMI solutions to improve the VRU and semi-automated vehicle interaction (SAE levels 2-4). In these vehicles, eHMI messages can be triggered automatically by the vehicle (and reported to the driver) but also activated/deactivated by the driver himself, thus modifying the course of the interaction. Since semi-automated vehicles will dominate and coexist in the traffic system in the coming decades, it is necessary to develop methods that evaluate the effectiveness of eHMIs in vehicles with different levels of automation.
2. Studies differ substantially in dynamic aspects of the vehicles presented to the VRUs during the interactions. These differences are mainly related to the speed of the vehicle or the force and timing of its decelerations. In this respect, other work has shown that the implicit signals derived from the behaviour of the vehicle is itself informative of its intentions, even modulating the effectiveness of the eHMIs. This suggests that the validation of future eHMIs should take into account the variety of dynamics that the vehicle may exhibit in the situation for which it has been designed.
3. The vast majority of studies have analysed signalised or unsignalised crossing situations between a pedestrian and an AV in perfect visibility conditions. While this type of situation is common, it probably does not represent the multiplicity of interaction scenarios that can occur in mixed traffic environments. To increase the external validity of the validation tests, additional types of representative interactions need to be considered. To do so, it is necessary to identify such use cases and scenarios, and to investigate them in experimental contexts.
4. Much has been invested in researching ways of communicating yielding intentions in AVs. But rarely have ways of communicating nonyielding intention been explored and compared. Including this analysis is especially important in confusing situations where the vehicle decelerates for different reasons (e.g., to increase safety in crowded areas or adapt speed to the car in front), providing unclear information about yielding intention. Therefore, it is important that test protocols for the validation of eHMIs not only look at the effectiveness in communicating yielding intention, but also at non-yielding intention.
5. Most studies have been carried out on populations of young adults and mid-adults. In contrast, children, older people or people with disabilities, who may face greater difficulties in interacting with vehicles in general, and eHMIs in particular, have been largely excluded from the literature. Since these groups represent a large part of the population, it is important that they are represented during the eHMI design and validation. As such, future validation methodologies should include participants of different ages and conditions where the eHMI effectiveness is tested.
6. Typically, studies have used participants who were previously trained in the eHMI at hand, who had been warned of the approach of automated vehicles, and who were highly attentive. In real traffic, it is unrealistic to expect all pedestrians to meet these conditions. Therefore, evaluation methods should be designed to assess the effectiveness of eHMIs in communicating vehicle intentions to all VRUs regardless of their awareness of the eHMI or their attentional level.
7. Different studies have demonstrated the utility of different subjective and objective measures to assess the safety and effectiveness of eHMIs (See Table 4-6). However, in many cases these measures differ across studies, affecting their comparability. In other cases, similar measures have been defined or calculated differently. For instance,

crossing initiation time or crossing decision time are the most commonly used objective measures. However, studies differ on the moment from which the time is computed (e.g., when the trial starts, when the AV brakes, etc.), or on the action signalling the decision time (e.g., step forward, step backwards, button press.). The measures are therefore calculated differently across studies. Future validation tests should make consistent use of reliable and well-defined measures that allow the effectiveness of different eHMI systems to be tested and compared.

8. Objective measures of the safety and efficiency of AV-pedestrian interactions are still very limited. However, there are other complementary objective measures that have not been used and may well complement the widely used crossing initiation/decision times. Some examples used by Dietrich et al. (2020) are Time-to-Arrival (TTA), Post-encroachment time (PET), Deceleration to safety time (DST). Future validation tests will be more complete if multiple measures are included that report on different aspects of the interaction (e.g., safety, efficiency, controllability).

Based on the literature review, the following considerations will be made during the HEIDI project:

1. As reflected in the literature, for a holistic user evaluation of eHMIs, multiple levels need to be carefully considered, primarily, safety, efficiency and user experience aspects. Following this, HEIDI will provide a final list of useful and well-defined measures to assess these aspects in future validation methods.
2. The objective of the HEIDI project is to improve the interactions between VRUs and vehicles, regardless of their level of automation. Since most of the literature has focused on fully autonomous vehicle-pedestrian interactions, new methodologies will be created or existing methodologies will be adapted, covering vehicles with a lower level of automation.
3. In relation to the previous point, the evaluation methods developed in HEIDI will strongly rely on co-simulation infrastructures, whereby interactions between drivers of vehicles with different levels of automation and one or more pedestrians can be analysed in a safe and reliable way. As it has been seen, co-simulation studies in this context are very scarce, so one of the objectives of HEIDI will be to design, test and propose methodologies that accommodate this type of analysis.
4. The influence of vehicle kinematics will be considered in several crucial tasks of the project. On the one hand, in WP3 for the design of more effective eHMIs where vehicle testing is part of the eHMI strategy. On the other hand, in WP5, for the design of validation methods where the effectiveness of eHMIs is analysed under different vehicle kinematic profiles.
5. The literature reviewed does not include studies on adaptive eHMIs, i.e., eHMIs that adapt the information to the situation, the level of automation and especially to the type (e.g. elderly, disabled, child) or state of the pedestrian (e.g. attentive, distracted) or the number of pedestrians. Since the purpose of HEIDI is to develop fluid and adaptive HMI systems, it will also be necessary to develop evaluation methodologies adapted to them.

5.3 Co-simulation

Summarizing the main parts from the literature study based on co-simulations:

1. There are very few studies performed using co-simulation strategies for looking at driver to pedestrian interaction. It is known that there are more performed studies than presented here but knowledge about how to perform and analyse results from a co-simulation is sparse.
2. Used simulators ranged from desktop simulators to VR and CAVE solutions. No moving base driving simulators were used.
3. Participants were almost always recruited in vicinity of universities, thus the participants were mainly young people. No studies considered elderly or disabled people, neither as drivers nor pedestrians.
4. The used scenarios were generally non-signalled crossing situations. In these situations, it is important to give the drivers and pedestrians enough time to interact so that interactions are studied instead of reactions. It is also important to make sure that the drivers and pedestrians meet at the designated spots. This synchronisation of participants needs to be considered if natural interaction is to be investigated.
5. Measures used in co-simulations are generally very similar to those used in other HMI studies. Measures for looking specifically at the interaction between the driver and pedestrian are few, although the few methods used indicate that there is a difference between a computer-controlled avatar and a human.
6. Setting up a study using several simulators almost certainly requires more participants (with two simulators you roughly get twice as many participants) and test leaders (probably one test leader for each simulator that greets the participant and prepares them). It also requires several systems to operate at the same time increasing the technical complexity of the simulation.

The following recommendations are considered for HEIDI specifically targeting co-simulation:

1. **Use at least one measure investigating the interaction between the road users.** The indication is that there is a significant difference between human and computer-controlled avatars which would be interesting to investigate further. Especially, as there are several studies that uses computer-controlled avatars, it would be interesting to get more insight into what that implies. Consider mixing human participants with both computer-controlled pedestrians as well as automated vehicles.
2. **Make sure that participants interact with each other through the simulators.** The participants shouldn’t be able to see or hear each other from outside their simulators and should interact (not react) within the virtual environment. Considerations should also be made to the instruction before the study starts, e.g., should the participants be aware of upcoming human-human interaction or be unaware and just act.
3. **Consider using a test leader as one of the participating road users.** Depending on the availability of needed test participants, for example elderly or disabled, it can be difficult to recruit them. As such, this consideration aims at making the experimental design more robust or provide a fallback plan towards participants becoming sick or unable to attend. The task leader can also help guide/enforce interaction in pre-defined spots in the virtual environment.

6. Conclusion

The objective of this deliverable was to gain an understanding of the state of the art on methodologies used to analyse the effectiveness of adaptive iHMIs and eHMIs, as well as on co-simulation-based methodologies. To this end, three independent literature reviews were carried out and the results are summarised in Chapter 5. In addition to the specific conclusions derived from each of the reviews, some overarching conclusions were derived that will be considered in the coming stages of the HEIDI project.

Firstly, it is important to stress that assessment methods should not be based solely on objective parameters of safety and/or efficiency. Holistic evaluations, where the user experience of interacting with HMIs is taken into account, are necessary. The deployment of this technology will depend to a large extent on widespread societal acceptance. Evaluating all these aspects requires the use of key performance indicators (KPIs), many of which are not yet well established.

The literature has also shown that very little is known about adaptive iHMI systems, and no literature exists for adaptive eHMIs, a crucial aspect of the HEIDI system. Since adaptive systems are more sophisticated and complex than non-adaptive ones, their implementation must be justified. To this end, validation of these systems needs to analyse their effectiveness both absolutely (i.e., how they improve the safety and efficiency of interactions) and relatively (i.e., how they generate superior effects to non-adaptive systems).

Finally, the literature review highlighted the scarcity of studies where co-simulation has been used for the analysis of pedestrian-driver/vehicle interactions. The use of co-simulation-based methodologies, however, is considered as a more ecologically valid solution to understand the actual dynamic interactions between pedestrians and vehicles both manually and with different levels of automation.

All in all, the reviews presented in this document have provided a good understanding of the state of the art in adaptive iHMIs, eHMIs and co-simulation. Equally, relevant methods and measures have been detected that will be investigated and adapted during the different tests planned in HEIDI. Nevertheless, it seems clear that there is ample room for the development and testing of new methods. This will be done through coordination between WP5 (i.e., in charge of the development of validation methods) and WPs 2 (i.e., conceptualisation, development and testing of iHMIs), WP3 (i.e., conceptualisation, development and testing of eHMIs) and WP4 (i.e., conceptualisation and development of the cHMI).

7. Abbreviations

Term	Definition
ACC	Adaptive Cruise Control
ADAS	Adaptive/Advanced Driver Assistance System
ADS	Automated Driving System
AOI	Area Of Interest
AR	Augmented Reality
AV	Autonomous Vehicle
BPDP	Brake Pedal Depression Percentage
CAVE	Cave Automatic Virtual Environment
CCC	Cross-correlation coefficient
cHMI	Cooperative HMI
CIT	Crossing Initiation Time
CleT	Clearing Time
CS	Crossing Speed
CT	Crossing Time
D	Deliverable
DST	Deceleration to Safety Time
eHMI	external Human Machine Interface
EuroNCAP	European New Car Assessment Programme
FCW	forward collision warning
HACC	Haptic Adaptive Cruise Control
HEIDI	Holistic and adaptivE Interface Design for human-technology Interactions
HMI	Human Machine Interface
HUD	Head-up Display
iHMI	internal Human Machine Interface
KPIs	Key Performance Indicators
LKA	Lane Keeping Assist
LTA	Lane Tracking Assistance
NASA-TLX	NASA Task Load Index
NDRT	Non-Driving Related Task
NHTSA	National Highway Traffic Safety Administration
NIRS	Near-infrared Spectroscopy
PET	Post Encroachment Time
PU	Public
R	Document, Report
RSME	Rating Scale Mental Effort
R-TLX	Raw Task Load Index
SART	Situation Awareness Rating Technique

SDLP	Standard deviation of lane position
SOA	State of the Art
SSVEP	Steady-state Visually Evoked Potentials
SUS	System Usability Scale
T	Task
TAM	Technology Acceptance Model
TTA	Time To Arrival
TTC	Time To Collision
UX	User Experience
VR	Virtual Reality
VRU	Vulnerable Road User
WOz	Wizard of Oz study
WP	Work Package

8. References

1. World Health Organization, *Global Plan for Decade of Action for Road Safety 2021–2030.* 2021.
2. Horberry, T., et al., *Human-centered design for an in-vehicle truck driver fatigue and distraction warning system.* IEEE Transactions on Intelligent Transportation Systems, 2021.
3. Meng, F., et al., *Dynamic vibrotactile signals for forward collision avoidance warning systems.* Human factors, 2015. **57**(2): p. 329-346.
4. Wulf, F., et al., *Recommendations supporting situation awareness in partially automated driver assistance systems.* IEEE Transactions on Intelligent Transportation Systems, 2014. **16**(4): p. 2290-2296.
5. Hirano, T., J. Lee, and M. Itoh. *Effects of Auditory Stimuli and Verbal Communications on Drivers' Situation Awareness in Partially Automated Driving*. in *2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*. 2018. IEEE.
6. Richardson, N.T., et al. *Conceptual design and evaluation of a human machine interface for highly automated truck driving*. in *2018 IEEE intelligent vehicles symposium (IV)*. 2018. IEEE.
7. Habibovic, A., et al., *Communicating Intent of Automated Vehicles to Pedestrians.* Front Psychol, 2018. **9**: p. 1336.
8. Amparore, E.G., et al. *Adaptive artificial co-pilot as enabler for autonomous vehicles and intelligent transportation systems*. in *ATT@ IJCAI*. 2018.
9. Biondi, F., et al. *Partial-autonomous frenzy: Driving a level-2 vehicle on the open road*. in *Engineering Psychology and Cognitive Ergonomics: Cognition and Design: 14th International Conference, EPCE 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II 14*. 2017. Springer.
10. Boelhouwer, A., et al., *Supporting drivers of partially automated cars through an adaptive digital in-car tutor.* Information, 2020. **11**(4): p. 185.
11. Coeugnet, S., et al. *A user-centered approach to adapt the human-machine cooperation strategy in autonomous driving*. in *Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021) Volume III: Sector Based Ergonomics*. 2021. Springer.
12. Feldhütter, A., C. Segler, and K. Bengler. *Does shifting between conditionally and partially automated driving lead to a loss of mode awareness?* in *Advances in Human Aspects of Transportation: Proceedings of the AHFE 2017 International Conference on Human Factors in Transportation, July 17− 21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8*. 2018. Springer.
13. Galarza, M.A., T. Bayona, and J. Paradells, *Integration of an adaptive infotainment system in a vehicle and validation in real driving scenarios.* International journal of vehicular technology, 2017.
14. Large, D.R., et al. *A Longitudinal simulator study to explore drivers' behaviour in level 3 automated vehicles*. in *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2019.
15. Li, X., et al., *Effects of an in-vehicle eco-safe driving system on drivers’ glance behaviour.* Accident Analysis & Prevention, 2019. **122**: p. 143-152.
16. Maag, C., et al., *Car Gestures–Advisory warning using additional steering wheel angles.* Accident Analysis & Prevention, 2015. **83**: p. 143-153.
17. Manawadu, U.E., et al. *A multimodal human-machine interface enabling situation-Adaptive control inputs for highly automated vehicles*. in *2017 IEEE Intelligent Vehicles Symposium (IV)*. 2017. IEEE.
18. Manawadu, U.E., et al. *Tactical-level input with multimodal feedback for unscheduled takeover situations in human-centered automated vehicles*. in *2018 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. 2018. IEEE.

19. Meiser, E., et al. *In-Vehicle Interface Adaptation to Environment-Induced Cognitive Workload*. in *Adjunct Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2022.
20. Musabini, A., et al., *Influence of Adaptive Human–Machine Interface on Electric-Vehicle Range-Anxiety Mitigation.* Multimodal Technologies and Interaction, 2020. **4**(1): p. 4.
21. Perrier, M.J., T.L. Louw, and O.M. Carsten, *Usability testing of three visual HMIs for assisted driving: How design impacts driver distraction and mental models.* Ergonomics, 2022: p. 1-22.
22. Riyahi, P., A. Eskandarian, and C. Zhang, *A brain wave-verified driver alert system for vehicle collision avoidance.* SAE International journal of transportation safety, 2021. **9**(1): p. 105-122.
23. Tanabe, H., et al. *Effects of a Robot Human-Machine Interface on Emergency Steering Control and Prefrontal Cortex Activation in Automatic Driving*. in *Engineering Psychology and Cognitive Ergonomics: 19th International Conference, EPCE 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26–July 1, 2022, Proceedings*. 2022. Springer.
24. Ueda, S. and T. Wada, *A Haptic Communication Method for A Preceding Vehicle Following System.* International Journal of Automotive Engineering, 2016. **7**(3): p. 99-105.
25. Wandtner, B., *Non-driving related tasks in highly automated driving-Effects of task characteristics and drivers' self-regulation on take-over performance*. 2018, Universität Würzburg.
26. American Psychological Association. *APA Dictionary of Psychology*. 2023 January 31, 2023]; Available from: https://dictionary.apa.org/physiological-measure.
27. Euro NCAP, *Euro NCAP Assessment Protocol - Safety Assist Safe Driving*. 2022.
28. Ahlström, C., G. Georgoulas, and K. Kircher, *Towards a Context-Dependent Multi-Buffer Driver Distraction Detection Algorithm.* IEEE Transactions on Intelligent Transportation Systems, 2022. **23**(5): p. 4778-4790.
29. Young, R.A., *A Tabulation of Driver Distraction Definitions [working document].* 2012.
30. Blanchard, R.A. and A.M. Myers, *Examination of driving comfort and self-regulatory practices in older adults using in-vehicle devices to assess natural driving patterns.* Accident Analysis & Prevention, 2010. **42**(4): p. 1213-1219.
31. Charlton, J.L., et al., *Characteristics of older drivers who adopt self-regulatory driving behaviours.* Transportation Research Part F: Traffic Psychology and Behaviour, 2006. **9**(5): p. 363-373.
32. Novotný, S. and P. Bouchner, *Elderly drivers vs. IVIS and ADAS--Results from a set of driving simulator studies.* Advances in transportation studies, 2011(24).
33. Hashimoto, N., S. Kato, and S. Tsugawa, *Parking Assistance System Based on Oral Instruction.* IEEJ Transactions on Industry Applications, 2009. **129**(2): p. 222-227.
34. Nguyen, T.T., et al., *Driver state detection based on cardiovascular system and driver reaction information using a graphical model.* Journal of Transportation Technologies, 2021. **11**(02): p. 139.
35. Nakamura, T., T. Daimon, and T. Oda. *Fundamental study of in-vehicle information provision based on cognitive workload of elderly driver when approaching an intersection*. in *21st World Congress on Intelligent Transport Systems: Reinventing Transportation in Our Connected World, ITSWC 2014*. 2014.
36. Ito, T., T. Shino, and M. Kamata, *Information Sharing to Improve Understanding of Proactive Steering Intervention for Elderly Drivers.* International Journal of Intelligent Transportation Systems Research, 2019. **17**: p. 18-31.
37. Li, S., et al., *Evaluation of the effects of age-friendly human-machine interfaces on the driver’s takeover performance in highly automated vehicles.* Transportation research part F: traffic psychology and behaviour, 2019. **67**: p. 78-100.
38. de Clercq, K., et al., *External Human-Machine Interfaces on Automated Vehicles: Effects on Pedestrian Crossing Decisions.* Hum Factors, 2019. **61**(8): p. 1353-1370.

39. Epke, M.R., et al., *I See Your Gesture: A VR-Based Study of Bidirectional Communication between Pedestrians and Automated Vehicles.* Journal of Advanced Transportation, 2021: p. 1-10.
40. Kaleefathullah, A.A., et al., *External Human-Machine Interfaces can be misleading: An examination of trust development and misuse in a CAVE-based pedestrian simulation environment.* Human Factors, 2022. **64**(6): p. 1070-1085.
41. Ferenchak, N.N. and S. Shafique, *Pedestrians' Perceptions of Autonomous Vehicle External Human-Machine Interfaces.* ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg, 2022. **8**(3).
42. Wilbrink, M., et al., *Impact of External Human–Machine Interface Communication Strategies of Automated Vehicles on Pedestrians' Crossing Decisions and Behaviors in an Urban Environment.* Sustainability, 2021. **13**(15).
43. Colley, M., C. Hummler, and E. Rukzio, *Effects of mode distinction, user visibility, and vehicle appearance on mode confusion when interacting with highly automated vehicles.* Transportation Research Part F: Traffic Psychology and Behaviour, 2022. **89**: p. 303-316.
44. Lanzer, M., et al., *Designing Communication Strategies of Autonomous Vehicles with Pedestrians: An Intercultural Study*, in *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2020. p. 122-131.
45. Song, Y.E., et al., *External HMIs and Their Effect on the Interaction Between Pedestrians and Automated Vehicles*, in *Intelligent Human Systems Integration*. 2018. p. 13-18.
46. Burns, C.G., et al. *Pedestrian decision-making responses to external human-machine interface designs for autonomous vehicles*. in *2019 IEEE Intelligent Vehicles Symposium (IV)*. 2019. IEEE.
47. Nuñez Velasco, J.P., et al., *Studying pedestrians' crossing behavior when interacting with automated vehicles using virtual reality.* Transportation research part F: traffic psychology and behaviour, 2019. **66**: p. 1-14.
48. Hensch, A.-C., et al., *Effects of a light-based communication approach as an external HMI for Automated Vehicles - a Wizard-of-Oz Study.* Transactions on Transport Sciences, 2020. **10**(2): p. 18-32.
49. Faas, S.M., V. Stange, and M. Baumann, *Self-Driving Vehicles and Pedestrian Interaction: Does an External Human-Machine Interface Mitigate the Threat of a Tinted Windshield or a Distracted Driver?* International Journal of Human–Computer Interaction, 2021. **37**(14): p. 1364-1374.
50. Dey, D., et al., *Communicating the intention of an automated vehicle to pedestrians: The contributions of eHMI and vehicle behavior.* it - Information Technology, 2020. **63**(2): p. 123-141.
51. Faas, S.M., L.-A. Mathis, and M. Baumann, *External HMI for self-driving vehicles: Which information shall be displayed?* Transportation Research Part F: Traffic Psychology and Behaviour, 2020. **68**: p. 171-186.
52. Hensch, A.C., et al., *The Effect of eHMI Malfunctions on Younger and Elderly Pedestrians' Trust and Acceptance of Automated Vehicle Communication Signals.* Front Psychol, 2022. **13**: p. 866475.
53. Dey, D., et al., *Distance-Dependent eHMIs for the Interaction Between Automated Vehicles and Pedestrians*, in *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2020. p. 192-204.
54. Dey, D., et al., *Towards Scalable eHMIs: Designing for AV-VRU Communication Beyond One Pedestrian*, in *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2021. p. 274-286.
55. Colley, M., et al., *Towards Inclusive External Communication of Autonomous Vehicles for Pedestrians with Vision Impairments*, in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020. p. 1-14.
56. Bindschädel, J., I. Krems, and A. Kiesel, *Active vehicle pitch motion for communication in automated driving.* Transportation Research Part F: Traffic Psychology and Behaviour, 2022. **87**: p. 279-294.

57. Bindschädel, J., I. Krems, and A. Kiesel, *Two-step communication for the interaction between automated vehicles and pedestrians.* Transportation Research Part F: Traffic Psychology and Behaviour, 2022. **90**: p. 136-150.
58. Haimerl, M., M. Colley, and A. Riener, *Evaluation of Common External Communication Concepts of Automated Vehicles for People With Intellectual Disabilities.* Proceedings of the ACM on Human-Computer Interaction, 2022. **6**(MHCI): p. 1-19.
59. Deb, S., L.J. Strawderman, and D.W. Carruth, *Investigating pedestrian suggestions for external features on fully autonomous vehicles: A virtual reality experiment.* Transportation Research Part F: Traffic Psychology and Behaviour, 2018. **59**: p. 135-149.
60. Kooijman, L., R. Happee, and J.C. de Winter, *How do eHMIs affect pedestrians' crossing behavior? A study using a head-mounted display combined with a motion suit.* Information, 2019. **10**(12): p. 386.
61. Guo, F., et al., *A Video-Based, Eye-Tracking Study to Investigate the Effect of eHMI Modalities and Locations on Pedestrian–Automated Vehicle Interaction.* Sustainability, 2022. **14**(9).
62. Eisma, Y.B., et al., *External human-machine interfaces: Effects of message perspective.* Transportation research part F: traffic psychology and behaviour, 2021. **78**: p. 30-41.
63. Guo, J., et al., *External Human-Machine Interfaces for Autonomous Vehicles from Pedestrians' Perspective: A Survey Study.* Sensors (Basel), 2022. **22**(9).
64. Faas, S.M. and M. Baumann, *Yielding Light Signal Evaluation for Self-driving Vehicle and Pedestrian Interaction*, in *Human Systems Engineering and Design II*. 2020. p. 189-194.
65. Othersen, I., et al., *Designing for automated vehicle and pedestrian communication: Perspectives on eHMIs from older and younger Persons.* Proceedings of the Human Factors and Ergonomics Society Europe, 2018. **4959**: p. 135-148.
66. Hochman, M., Y. Parmet, and T. Oron-Gilad, *Pedestrians' Understanding of a Fully Autonomous Vehicle's Intent to Stop: A Learning Effect Over Time.* Front Psychol, 2020. **11**: p. 585280.
67. Faas, S.M. and M. Baumann, *Light-Based External Human Machine Interface: Color Evaluation for Self-Driving Vehicle and Pedestrian Interaction.* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2019. **63**(1): p. 1232-1236.
68. Holländer, K., et al., *Take It to the Curb: Scalable Communication Between Autonomous Cars and Vulnerable Road Users Through Curbstone Displays.* Frontiers in Computer Science, 2022. **4**.
69. Şahin, H., et al., *An Exploration of Potential Factors Influencing Trust in Automated Vehicles*, in *Human-Computer Interaction – INTERACT 2021*. 2021. p. 364-367.
70. Dietrich, A., M. Tondera, and K. Bengler, *Automated vehicles in urban traffic: The effect of kinematics and eHMIs on pedestrian crossing behavior*, in *Road Safety and Simulation Conference*. 2019: University of Iowa.
71. Bindschädel, J., I. Krems, and A. Kiesel, *Interaction between pedestrians and automated vehicles: Exploring a motion-based approach for virtual reality experiments.* Transportation Research Part F: Traffic Psychology and Behaviour, 2021. **82**: p. 316-332.
72. Colley, M., E. Bajrovic, and E. Rukzio, *Effects of Pedestrian Behavior, Time Pressure, and Repeated Exposure on Crossing Decisions in Front of Automated Vehicles Equipped with External Communication*, in *CHI Conference on Human Factors in Computing Systems*. 2022. p. 1-11.
73. Lau, M., M. Jipp, and M. Oehl, *One Solution Fits All? Evaluating Different Communication Strategies of a Light-based External Human-Machine Interface for Differently Sized Automated Vehicles from a Pedestrian's Perspective.* Accid Anal Prev, 2022. **171**: p. 106641.
74. Dey, D., et al., *Pedestrian road-crossing willingness as a function of vehicle automation, external appearance, and driving behaviour.* Transportation research part F: traffic psychology and behaviour, 2019. **65**: p. 191-205.

75. Faas, S.M., A.C. Kao, and M. Baumann, *A Longitudinal Video Study on Communicating Status and Intent for Self-Driving Vehicle – Pedestrian Interaction*, in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020. p. 1-14.
76. Dey, D., et al., *Investigating the Need for Explicit Communication of Non-Yielding Intent through a Slow-Pulsing Light Band (SPLB) eHMI in AV-Pedestrian Interaction*, in *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2022. p. 307-318.
77. Eisma, Y.B., et al., *External human–machine interfaces: The effect of display location on crossing intentions and eye movements.* Information, 2020. **11**(1).
78. Faas, S.M., A.C. Kao, and M. Baumann, *A Longitudinal Video Study on Communicating Status and Intent for Self-Driving Vehicle - Pedestrian Interaction*, in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020. p. 1-14.
79. Joisten, P., et al., *Communication of Automated Vehicles and Pedestrian Groups: An Intercultural Study on Pedestrians' Street Crossing Decisions*, in *Mensch und Computer 2021*. 2021. p. 49-53.
80. Lau, M., M. Jipp, and M. Oehl, *Toward a Holistic Communication Approach to an Automated Vehicle's Communication With Pedestrians: Combining Vehicle Kinematics With External Human-Machine Interfaces for Differently Sized Automated Vehicles.* Front Psychol, 2022. **13**: p. 882394.
81. Lee, J., T. Daimon, and S. Kitazaki. *Negative effect of external human-machine interfaces in automated vehicles on pedestrian crossing behaviour: A virtual reality experiment*. in *Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021) Volume III: Sector Based Ergonomics*. 2021. Springer.
82. Löcken, A., C. Golling, and A. Riener. *How should automated vehicles interact with pedestrians? A comparative analysis of interaction concepts in virtual reality*. in *Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications*. 2019.
83. Loew, A., et al., *Go Ahead, Please!—Evaluation of External Human—Machine Interfaces in a Real-World Crossing Scenario.* Frontiers in Computer Science, 2022. **4**.
84. Matsunaga, N., et al. *Effect of the external human machine interface (eHMI) of automated vehicle on pedestrian's recognition*. in *26th International Display Workshops, IDW 2019*. 2019. International Display Workshops.
85. Métayer, N. and S. Coeugnet, *Improving the experience in the pedestrian's interaction with an autonomous vehicle: An ergonomic comparison of external HMI.* Appl Ergon, 2021. **96**: p. 103478.
86. Oudshoorn, M., et al., *Bio-inspired intent communication for automated vehicles.* Transportation research part F: traffic psychology and behaviour, 2021. **80**: p. 127-140.
87. Rodríguez Palmeiro, A., et al., *Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment.* Transportation Research Part F: Traffic Psychology and Behaviour, 2018. **58**: p. 1005-1020.
88. Rouchitsas, A. and H. Alm, *Ghost on the Windshield: Employing a Virtual Human Character to Communicate Pedestrian Acknowledgement and Vehicle Intention.* Information, 2022. **13**(9).
89. Sahaï, A., et al., *Crossing the street in front of an autonomous vehicle: An investigation of eye contact between drivengers and vulnerable road users.* Front Psychol, 2022. **13**: p. 981666.
90. Şahin, H., et al. *Signaling Yielding Intent with eHMIs: the Timing Determines an Efficient Crossing*. in *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2021.
91. Singer, T., et al., *Displaying the Driving State of Automated Vehicles to Other Road Users: An International, Virtual Reality-Based Study as a First Step for the Harmonized Regulations of Novel Signaling Devices.* IEEE Transactions on Intelligent Transportation Systems, 2022. **23**(4): p. 2904-2918.
92. Löcken, A., et al., *Accessible Automated Automotive Workshop Series (A3WS): International Perspective on Inclusive External Human-Machine Interfaces*, in *14th*

International Conference on Automotive User Interfaces and Interactive Vehicular Applications. 2022. p. 192-195.

93. Lee, Y.M., et al., *Designing cooperative interaction of automated vehicles with other road users in mixed traffic environments D.3.1 Cooperation and Communication Planning Unit Concept. InterACT D6.1 Methodologies for the Evaluation and Impact Assessment of the InterACT Solutions, 1.* 2019.
94. Lehsing, C., T. Benz, and K. Bengler, *Insights into interaction-effects of human-human interaction in pedestrian crossing situations using a linked simulator environment.* IFAC-PapersOnLine, 2016. **49**(19): p. 138-143.
95. Lehsing, C., M. Fleischer, and K. Bengler. *On the track of social interaction-A non-linear quantification approach in traffic conflict research*. in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. 2016. IEEE.
96. Lehsing, C., et al., *Effects of simulated mild vision loss on gaze, driving and interaction behaviors in pedestrian crossing situations.* Accident Analysis & Prevention, 2019. **125**: p. 138-151.
97. Sadraei, E., et al. *Vehicle-pedestrian interaction: A distributed simulation study*. in *Proceedings of the driving simulation conference. Antibes, France*. 2020.
98. Mok, C.S., P. Bazilinskyy, and J. de Winter, *Stopping by looking: A driver-pedestrian interaction study in a coupled simulator using head-mounted displays with eye-tracking.* Applied ergonomics, 2022. **105**: p. 103825.
99. Lindner, J., et al. *A mobile application for resolving bicyclist and automated vehicle interactions at intersections*. in *2022 IEEE Intelligent Vehicles Symposium (IV)*. 2022. IEEE.
100. Lindner, J., et al. *A coupled driving simulator to investigate the interaction between bicycles and automated vehicles*. in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. 2022. IEEE.
101. Lehsing, C., A. Kracke, and K. Bengler. *Urban perception-a cross-correlation approach to quantify the social interaction in a multiple simulator setting*. in *2015 IEEE 18th international conference on intelligent transportation systems*. 2015. IEEE.
102. Mori, M., K.F. MacDorman, and N. Kageki, *The uncanny valley [from the field].* IEEE Robotics & automation magazine, 2012. **19**(2): p. 98-100.