# You're the One Whom I'm Talking To: The Role of Contextual External Human-Machine Interfaces in Multi-Road User Conflict Scenarios

YUMIN KANG, Gwangju Institute of Science and Technology, Republic of Korea
JEONGJU PARK, Gwangju Institute of Science and Technology, Republic of Korea
SEOKHYUN HWANG, University of Washington, USA
MINWOO SEONG, Gwangju Institute of Science and Technology, Republic of Korea
GWANGBIN KIM, Gwangju Institute of Science and Technology, Republic of Korea
SEUNGJUN KIM*, Gwangju Institute of Science and Technology, Republic of Korea

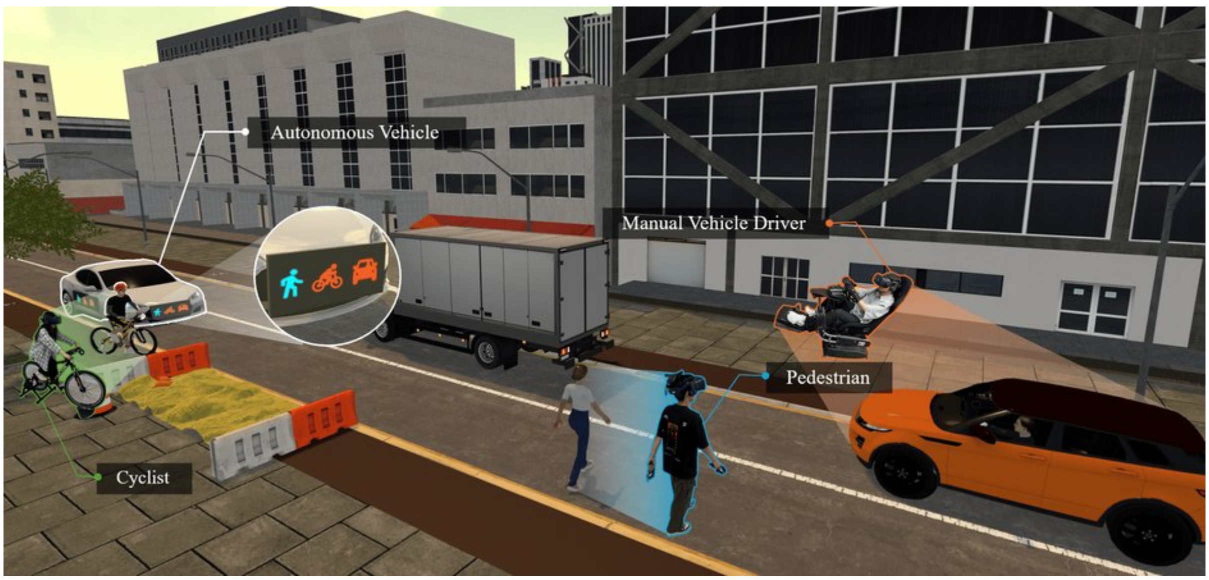


Fig. 1. Exploring crossing performance and subjective feelings of pedestrians, cyclists, and manual vehicle drivers during AV conflicts on a two-lane road. The AV used eHMI signals to convey its intentions, enabling road users to navigate their routes and make decisions based on the interpreted AV cues.

*Corresponding author.

Authors' Contact Information: Yumin Kang, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea, yuminkang@gm.gist.ac.kr; Jeongju Park, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea, jeongjupark@gm.gist.ac.kr; Seokhyun Hwang, University of Washington, Seattle, Washington, USA, seokhyun@uw.edu; Minwoo Seong, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea, seongminwoo@gm.gist.ac.kr; Gwangbin Kim, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea, gwangbin@gm.gist.ac.kr; SeungJun Kim, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea, seungjun@gist.ac.kr.


ACM 2474-9567/2025/9-ART94
https://doi.org/10.1145/3749473

As autonomous vehicles (AVs) become more prevalent, mixed-traffic environments involving pedestrians, cyclists, and manual vehicle drivers pose significant challenges for ensuring safe and effective interactions. External Human-Machine Interfaces (eHMIs) have emerged as a solution, particularly context-based eHMIs, which provide specific information such as *Whom*, *When*, and *Where*, showing potential for improving communication in complex scenarios. However, their impact on road user behavior and safety in interactions involving multiple road users remains insufficiently explored. This study addresses this gap by examining how contextual eHMIs affect crossing performance and subjective feelings during multi-user conflict scenarios. Using a virtual reality-based multi-agent simulation, 42 participants were equally divided into three groups—pedestrians, cyclists, and manual vehicle drivers—to make crossing decisions during interactions with an AV. Our findings demonstrated that providing contextual information in AV-multi-road user interactions significantly enhanced participants' crossing performance and improved their perceived safety, trust, and clarity. These findings highlight the potential of context-based eHMIs to facilitate safer and more intuitive interactions in mixed-traffic environments.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: autonomous vehicles, eHMI, scalability, context, multi-road users, pedestrians, cyclists, manual vehicle drivers, vulnerable road users, virtual reality, AV-multi-road user interaction

**ACM Reference Format:**
Yumin Kang, Jeongju Park, Seokhyun Hwang, Minwoo Seong, Gwangbin Kim, and SeungJun Kim. 2025. You're the One Whom I'm Talking To: The Role of Contextual External Human-Machine Interfaces in Multi-Road User Conflict Scenarios. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 9, 3, Article 94 (September 2025), 37 pages. https://doi.org/10.1145/3749473

## 1 INTRODUCTION

The adoption of automated driving technology is expected to enhance traffic safety by reducing collisions and facilitating smoother traffic flow [47, 119]. However, it also poses additional concerns outside the vehicle due to insufficient cues provided by autonomous vehicles (AVs). Traditionally, drivers incorporated explicit signals, such as eye contact and hand gestures, for proactive interaction with other road users [87]. Since AV occupants focus primarily on non-driving-related tasks and may not have vehicle control [74], non-AV road users face difficulties receiving explicit cues from AV drivers [3, 61]. As AVs increasingly share roads with human-driven vehicles [86, 112] and other road users [55], ensuring safe and efficient traffic interactions requires innovative communication solutions.

External Human-Machine Interfaces (eHMIs) have emerged as a potential solution to assist vulnerable road users (VRUs), such as pedestrians and cyclists, in understanding AV status and intent [53, 77]. Research has examined various visual interaction modalities, including LED light bands [35, 72], projections [41, 97], symbols [25, 84], and anthropomorphic designs [21, 22]. While these studies have contributed to improving AV-pedestrian interactions [71, 119, 128] and interactions with other road users [5, 16, 28, 107], they have primarily focused on individual user groups. Given the distinct physical and cognitive engagement levels across user types, validating these findings in multi-user scenarios, where multiple road users coordinate and interact concurrently, remains a critical need.

Indeed, road environments involve diverse user groups, including pedestrians, cyclists, and manual vehicle drivers, who constantly interact with each other and coordinate their actions. These stakeholders engage in varying levels of physical and cognitive demands [111], requiring effective communication to understand the current and future behaviors of others and adjust accordingly. Pedestrians require basic motor skills and wayfinding abilities [83, 133], while cyclists need situation awareness—perceiving environmental elements, understanding their significance, and predicting future states [45]—as well as physical endurance to control their vehicles [95, 111]. Cyclists often share the road with vehicles, encountering challenges such as overtaking and merging [14, 15]. In contrast, manual vehicle drivers navigate high-speed, dynamic environments that demand constant alertness [58]. Given these diverse physical and cognitive engagements, the effectiveness of eHMIs could vary across user types, highlighting the need for testing in multi-user scenarios.

The variability in eHMI research across road user types suggests that designs primarily intended for one group may not be equally effective for others. For example, eHMIs designed specifically for pedestrians might perform differently when applied to other road users, potentially leading to miscommunication [132] and compromising traffic safety and efficiency [77, 103]. While controlled scenarios that assume one-to-one interaction are valuable for understanding the fundamental AV-road user interaction, they do not capture the complexity of real-world scenarios, which typically involve multiple road users interacting simultaneously [40]. Thus, equal emphasis should be placed on multi-road user interactions when designing eHMIs to facilitate clearer communication.

The complexity of interactions between AVs and multiple road users requires contextual information to reduce ambiguity or misunderstanding. For instance, the multi-road user scenario raises ambiguity about the recipient of an intended message [40, 127], increasing uncertainty, and diminishing the effectiveness and trustworthiness of eHMIs [14, 46]. On the other hand, providing specific contextual information, such as the time or location at which an AV will stop enhances situational awareness, safety, trust, and crossing performance [36, 80]. Ultimately, it may facilitate efficient interactions and reduce potential accidents through improved communication clarity [36, 110].

Given the need to explore complex interaction scenarios with multiple road users [127, 132] and the importance of contextual information, investigating eHMIs that provide detailed contextual information becomes increasingly necessary. This is especially crucial in high-urgency [82] or complex scenarios, such as where road users move at varying speeds or directions, or face space-sharing conflicts. To address this research gap, we propose an eHMI that conveys contextual information to diverse road user groups, including pedestrians, cyclists, and manual vehicle drivers, to facilitate communication with AVs. We examine how these contextual eHMIs influence road user decision-making and suggest considerations for scaling eHMIs.

Specifically, we focus on providing contextual information using the *Kipling method* (5W1H), which has been used to facilitate human-vehicle interactions [24, 38, 65, 66, 115], particularly the elements of *Whom*, *When*, and *Where*. We designed an eHMI incorporating contextual elements and assessed its impact on the crossing performance, subjective feelings, and physiological variations of different road users through virtual reality (VR)-based multi-agent simulation (N = 42). To evaluate the effects of different context types relevant to their distinct characteristics or conflicts [89], we experimented with pedestrians, cyclists, and manual vehicle drivers interacting simultaneously with an AV.

We formulated the following research question to guide our assessment of the proposed eHMI, focusing on the unique characteristics of different road user groups:

**RQ.** How does providing specific contextual information (*Whom/When/Where*) in a yielding message affect crossing performance (onset time, duration), subjective feelings (perceived safety, trust, clarity), and physiological response (Electrodermal Activity (EDA)) of *pedestrians*, *cyclists*, and *manual vehicle drivers* compared to when no contextual information is provided?

Our contributions are summarized as follows:

(1) We presented an empirical evaluation of eHMI performance in multi-user scenarios involving pedestrians, cyclists, and manual vehicle drivers, each with distinct characteristics and conflicts.
(2) We investigated the impact of specific contextual information (*Whom*, *When*, *Where*) in one-to-many interaction scenarios and analyzed its performance in addressing potential in real traffic.
(3) We developed a scalable evaluation platform in a simulated environment, enabling real-time interactions that addressed the complexities of real traffic scenarios.

## 2 RELATED WORK

### 2.1 Towards Scalable AV-Road User Interaction

Implicit cues, such as a vehicle's acceleration or deceleration patterns, help road users interpret its intentions [2, 105]. For example, deceleration implicitly signals yielding, while maintaining or increasing speed may indicate proceeding [12, 122]. However, these inherent vehicle characteristics alone may not effectively convey AV intentions [34, 75]. Consequently, studies have explored the influence of AV kinematics within the eHMI context [17, 75].

Potential issues arising from conflicts between AVs and diverse road user groups that require interaction with AVs have been explored, such as pedestrians [29, 37, 126], cyclists [4, 16, 54], and manual vehicle drivers [91, 123]. *Pedestrians* are the most commonly considered VRUs in AV interactions [53]. To aid efficient decision-making for pedestrians, research has categorized information types—such as AV status, perception, intent, and advice—to understand what information pedestrians need from AVs and how it affects their crossing behavior, with particular emphasis on status, perception, and intent [46]. Studies have also explored implementing interfaces that mimic eye contact from human drivers [21], and proposed integrated vehicle interfaces to deliver unambiguous messages and prevent pedestrian confusion [125]. Moreover, various AV-pedestrian interaction concepts have been compared to identify the most effective designs for improving pedestrian safety and trust [25, 81].

*Cyclists*, the second-most studied VRUs, are classified as a non-motorized group [53]. Unlike pedestrians, cyclists are distinguished by their speed and dynamic characteristics [51, 52]. Given their higher exposure to accidents in motorized environments [18, 109, 130], cyclists are crucial interaction partners for AVs. Studies on AV-cyclist interaction have focused on developing a taxonomy of eHMI characteristics tailored to cyclists through participatory design processes [6] and identifying critical interaction scenarios [14].

*Manual vehicle drivers* are another major road user group that requires interaction with AVs, especially in scenarios such as bottlenecks, where negotiation with AVs is inevitable [89, 108]. Interactions between AVs and manual vehicle drivers have been explored in scenarios such as bottlenecks [78, 107] and highway platooning [28].

While previous studies have included diverse user groups, there has been comparatively little focus on road users other than pedestrians, especially cyclists who face a high incidence of traffic accidents [18, 90]. This highlights the need for further research on external cues in scenarios involving these user groups, particularly given their different interaction dynamics.

To reflect more realistic road characteristics, studies have explored factors for applying eHMIs to real traffic, such as time constraints, pedestrian groups [24], smartphone distractions, and the influence of other pedestrians [71]. To better understand the effect of such independent variables, scenarios involving a single road user have been primarily evaluated [6, 16, 32, 101, 121]. While this approach is feasible for testing individual factors, these findings could be further validated under more naturalistic conditions, such as scenarios with multiple vehicles and pedestrians [55]. The importance of such ecological validity has been raised within eHMI studies, especially calling for addressing scenarios involving diverse road user groups for eHMI scalability [30, 129].

In response, efforts have been made to examine eHMI impacts on pedestrian crossing in more realistic scenarios involving multiple pedestrians [25, 40]. In addition, simulators were developed to analyze interactions between an AV and two road users (i.e., a pedestrian and a manual vehicle driver) [11], and strategies involving a pedestrian, a manually driven vehicle, and an AV have been explored [55]. Despite the increasing complexity of interaction dynamics in real road scenarios, most studies have assumed one-to-one interactions, whereas actual traffic scenarios frequently involve many-to-many interactions. The concept of Distributed Situation Awareness (DSA) [120] emphasizes the necessity of shared understanding among agents to address the complexities of traffic environments. In the context of eHMIs, DSA principles may support simultaneous interactions between multiple road users and AVs by fostering shared understanding, particularly when eHMIs provide information tailored to the surrounding context or specific situations. Nevertheless, the role of eHMIs in promoting shared understanding across diverse road users remains underexplored, as also noted by Bazilinskyy et al. [10]. Similarly, Tran et al. [127] highlighted the need to address eHMI scalability issues in scenarios that better reflect real traffic conditions. In response to these challenges, our study aims to examine different eHMI designs for interactions where multiple users simultaneously engage with AVs, assessing their effectiveness under ecologically valid and scalable conditions.

## 2.2 Providing Contextual Information via eHMIs

In scenarios with multiple road users, if an AV merely provides information such as whether it will yield, without adaptively responding to the surrounding environment, road users may struggle to determine whether the AV is yielding specifically to them or communicating information intended for others. When the intended recipient of the AV's message is unclear, it can lead to confusion among road users who observe the message [33]. Hübner et al. [55] found that most participants misinterpreted AV messages when the intended recipient was unclear. Similarly, Dey et al. [40] argued that conveying intention without specifying the recipient in scenarios involving more than one nearby road user poses a safety risk. This underlines the importance of clearly indicating the intended recipient (*'Whom'*) to mitigate potential accidents and enhance safety [43, 132].

Road users often make decisions in conflict situations by weighing convenience, time efficiency, and safety [116]. This decision-making process can lead to the adoption of crossing strategies that benefit them more, such as crossing in areas without designated crosswalks or crossing diagonally to avoid latency [100]. Providing information about AV action timing (*'When'*) has been shown to improve road users' understanding or trust [40, 63]. This can aid in efficient decision-making and potentially reduce uncertainty regarding the AV's stopping point, enhancing road users' perceived safety and trust. In this context, an interface concept was proposed that conveys information about the deceleration state of AV, including its yielding point, through a dynamic circular symbol and light strips [38, 115].

In scenarios such as shared spaces, where signals and crosswalks are absent and vehicles share the same area [62, 93], providing only vehicle kinematics or simple yielding cues may be insufficient to ensure road user safety and traffic efficiency due to the absence of explicit human driver signals [48, 77]. Colley et al. [25] found that projecting the AV's yielding location onto the street increased trust and acceptance. This suggests that conveying specific contextual information about where the AV will yield (*'Where'*) is crucial in reducing uncertainty regarding vehicle kinematics, potentially improving both safety and traffic efficiency.

These findings directly shape the focus of our study on providing contextual information (*'Whom'*/*'When'*/*'Where'*) through eHMIs. While conveying abundant information may aid in user understanding, excessive information may induce cognitive overload, especially under conflict scenarios with physical, cognitive, and temporal demands. Our research targets this trade-off by incorporating these three key contextual elements into eHMIs in situations involving multiple road user types.

# 3 METHODS

This section outlines the methods including in our study, covering the study design, experimental setup, procedures, and participant details.

## 3.1 Study Design

This study aimed to explore the impact of contextual information beyond AV-yielding behavior using a VR-based multi-agent simulation. The experiment employed a mixed design with two independent variables and seven dependent variables, focusing on scenarios involving interactions between the AV and multiple road users in simultaneous conflict situations. To prevent participants from preemptively anticipating the AV's behavior, the AV alternated between yielding and non-yielding behaviors, allowing participants to experience both conditions. A balanced Latin square design was applied for counterbalancing, with all conditions presented in a randomized order.

*3.1.1 Independent Variables.* We set two independent variables: road user type and eHMI type. The road user type was a between-subjects factor, while the eHMI type was a within-subjects factor.

First, the 'road user type' comprised three groups: pedestrians, cyclists, and manual vehicle drivers. These groups were selected due to their differing mobility characteristics and high likelihood of being key communication targets for AVs in mixed-traffic scenarios [102, 106, 107].

Second, the 'eHMI type' was defined to assess the impact of contextual information on the crossing performance and subjective feelings of each road user. This factor had five categories: *No eHMI* (baseline), *No Context eHMI*, *Whom eHMI*, *When eHMI*, and *Where eHMI*. While providing advice grants them clearly defined rights [25], explicit guidance to road users may raise legal and ethical concerns [46] or lead to potential misinterpretation by other road users [33]. To mitigate these issues, we adopted a method in which the AV communicated its intent, enabling road users to make their own decisions. All experimental conditions, including contextual information (*Whom, When, Where eHMIs*), were conducted under consistent scenarios with uniform instructions, ensuring equal emphasis across conditions to avoid bias arising from setup or instructional discrepancies. For yielding behaviors, we followed the approach of Werner et al. [131] and Faas et al. [46], selecting cyan (RGB 0, 255, 255), a color that does not overlap with existing traffic signals. For non-yielding behaviors, we used orange (RGB 255, 90, 40) based on the work of Kuge et al. [70] and Rettenmaier et al. [107], as it similarly avoids associations with current traffic systems. The interface was installed on all four sides of the vehicle—front, left, right, and rear—to maximize visibility and ensure it could be easily observed from any direction. This placement allowed participants to focus solely on the impact of eHMI by minimizing the influence of other factors, such as viewing angles, thereby enabling them to engage in natural interactions. Furthermore, all context-based eHMI types (*No Context, Whom, When, Where*) utilized a symbol-based approach to clearly communicate the AV's intentions while minimizing cultural or individual biases [9, 23, 25]. A summary of all independent variables is shown in Fig. 2.

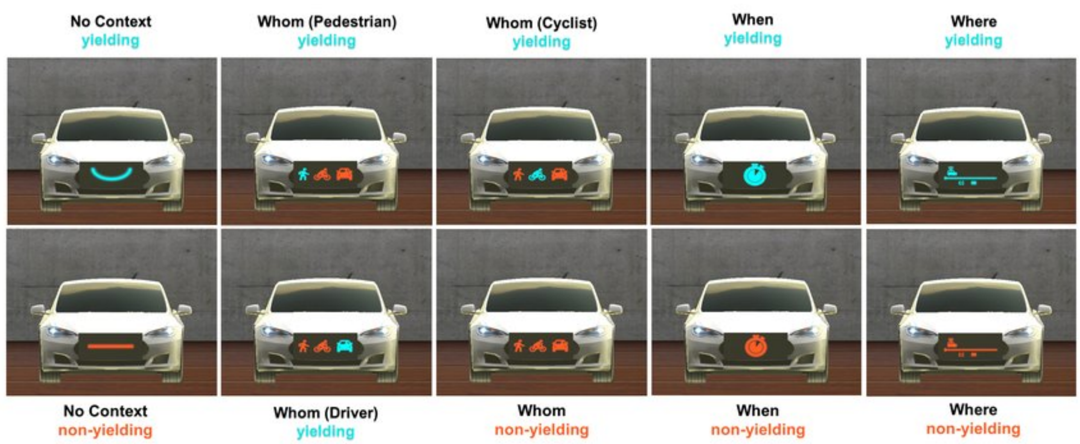


Fig. 2. Visual representations of eHMI types under yielding and non-yielding behaviors for different context types (*No Context*, *Whom*, *When*, and *Where*). Each panel illustrates the front display of the AV. The No eHMI condition is represented by the absence of any visual cue, whereas the No Context eHMI uses a mouth-shaped symbol to indicate yielding intent. Whom eHMIs are tailored to specific road user groups and differ by target icon, while When and Where eHMIs provide temporal and spatial information, respectively. All visualizations are shown from the perspective of road users.

***No eHMI.*** 'No eHMI' served as the baseline condition, in which no external display or information about the AV's behavior was provided. This setup aimed to evaluate how the absence of an eHMI affects road users' crossing behavior.

***No Context.*** ‘No Context’ provided only whether the AV would yield, without any contextual details. To avoid potential cultural and social variability associated with text-based indicators [68, 73], we adopted an anthropomorphic design approach to minimize such variability [124]. We designed a smiling mouth symbol for the yielding case and a neutral mouth symbol for the non-yielding [20] (see Fig. 2A and 2B).

***Whom.*** ‘Whom’ was designed to evaluate the effectiveness of explicitly indicating the target road user. The conceptual design of the *Whom eHMI* aimed to reduce ambiguity and promote immediate comprehension by clearly specifying which road user the AV would or would not yield to. Anthropomorphic symbols were adopted for all road users involved in the conflict, with the positions of the symbols fixed from left to right: pedestrian, cyclist, and manual vehicle driver (see Fig. 2C, 2D, 2E, 2F). In scenarios involving a single user group, only the relevant symbol was displayed to minimize visual clutter. If the AV did not intend to yield to any road user, all symbols were displayed in orange. The goal was to communicate the AV’s intentions without assigning authority, enabling road users to make contextually informed decisions. Scenarios where the AV yielded to all road users or more than one road user were excluded to prevent simultaneous crossing attempts, which could increase the risk of collisions due to conflicting trajectories.

***When.*** ‘When’ was designed to investigate the effectiveness of the eHMI in communicating the timing of the AV’s yielding behavior. The conceptual design of the *When eHMI* aimed to provide actionable temporal cues that could be easy to understand and apply. The interface displayed a timer symbol to indicate the AV’s yielding or non-yielding intention. The timer operated by showing a filled circle that gradually decreased in a clockwise direction, visually representing the passage of time (see Fig. 2G and 2H). Once the timer reached completion (i.e., the circle within the timer became completely empty), it signaled that the AV had come to a full stop and was yielding to the road user. It was designed to mirror natural time-based visual cues, enabling road users to predict the AV’s stopping time and plan their actions accordingly. By conveying both the onset and completion of the yielding behavior, the When eHMI provided clarity on the AV’s intended stopping behavior and timing.

***Where.*** ‘Where’ was designed to assess how effectively the *Where eHMI* communicated the AV’s precise stopping location to road users. The conceptual design of the *Where eHMI* aimed to provide clear spatial information through symbols and text indicating the remaining distance to the stopping point (see Fig. 2I and 2J). It consisted of three components: a dynamic vehicle symbol, a horizontal line, and text displaying the distance to the stopping point in meters. Dynamic animations highlighted the vehicle’s movement and stopping location, ensuring visibility even from a distance. The horizontal line represented the road, with the vehicle symbol moving from left to right above the line to depict the AV’s approach. The endpoint on the right side of the line marked the exact stopping point. When the vehicle symbol reached the endpoint (i.e. when it aligned with the rightmost point on the line) and the text displayed ‘0 m’, it signaled that the AV would completely stop and yield. The rationale behind the Where eHMI was to assist road users in making more informed crossing decisions by providing detailed and intuitive spatial information [9, 32]. To avoid providing prescriptive instructions, cues such as crosswalk symbols were omitted [25]. Instead, the design focused on conveying information about the AV’s actions, enabling road users to make decisions based on the AV’s signals as well as the broader traffic context.

*3.1.2 Dependent Variables.* We defined four objective measures and three subjective measures. The objective measures were: 1) crossing onset time (COT), 2) crossing duration (CD), 3) error rate, and 4) arousal state (EDA). The subjective measures included: 1) perceived safety, 2) trust, and 3) clarity. These variables were grouped into three categories: crossing performance (COT, CD), subjective feelings (perceived safety, trust, clarity), and physiological responses (EDA).

For the objective measures, we collected data on COT, CD, error rate, and EDA. COT was defined as the time between the eHMI display and the start of crossing, while CD measured the time participants spent crossing within their designated zones. These metrics were used to assess variations in crossing performance based on the contextual information provided by the AV. Error rate captured instances of misunderstanding the interface, leading to decisions that conflicted with the AV's intentions. EDA, which varies according to the arousal state induced by events encountered by road users [19, 76], was measured in real-time to evaluate participants' responses to the contextual cues.

For the subjective measures, we assessed perceived safety, trust, and clarity during decision-making in response to the contextual information provided. Participants rated their perceived safety using the Godspeed Questionnaire [8], which indicated their safety level when deciding to cross. Trust in the AV was assessed using the Trust in Automation scale [57], while clarity—an essential factor for scalable eHMI designs [40, 81]—was measured on an 11-point Likert scale based on Eisma et al. [44].

*3.1.3 Scenario.* To enhance the realism and relevance of the chosen traffic scenarios, we designed interactions that reflect typical road user conflicts while avoiding assumptions that might oversimplify real-world traffic dynamics. Accordingly, we selected a two-lane bidirectional road with a dedicated bicycle lane and no traffic signals, incorporating three conflict scenarios—*crossing paths*, *merging paths*, and *constrained head-on paths*—from the five types of space-sharing conflicts defined by Markkula et al. [89] (Fig. 3). These scenarios emphasize conflicts where multiple road users simultaneously attempt to occupy the same space, often in high-traffic environments with unclear right-of-way rules. Specifically, our scenarios included situations where pedestrian right-of-way is not always guaranteed, such as in regions without marked crosswalks, which are associated with higher pedestrian crash rates and reduced driver compliance [92, 113]. To simulate conditions where pedestrian priority is unclear, we modeled scenarios where pedestrians crossed in such areas [39, 125]. Each road user, moving at varying speeds and directions, was required to interact with the AV.

The AV began driving from a point 80 m away from the pedestrian, either providing no interface or offering contextual information (*'No Context'*, *'Whom'*, *'When'*, and *'Where'*). Contextual messages were presented starting 65 m [31] before the first interaction with the cyclist merging area. If the AV intended to yield, it stopped at the merging area and remained stationary until the target user(s) completed their crossing. If the AV did not yield and asserted its right of way (i.e., not yielding to any road user), it displayed a non-yielding message and continued driving. The AV's route is shown in white in Fig. 3A.

For *pedestrians*, their task was to interact with the AV approaching from their left (*crossing paths*) and to coordinate with cyclists and manual vehicle drivers. Pedestrians stopped at the designated crossing point to receive contextual messages from the AV. Based on the AV's message, pedestrians were prompted to decide whether to cross. The pedestrian route is indicated in light blue in Fig. 3B, with their perspective in Fig. 5A.

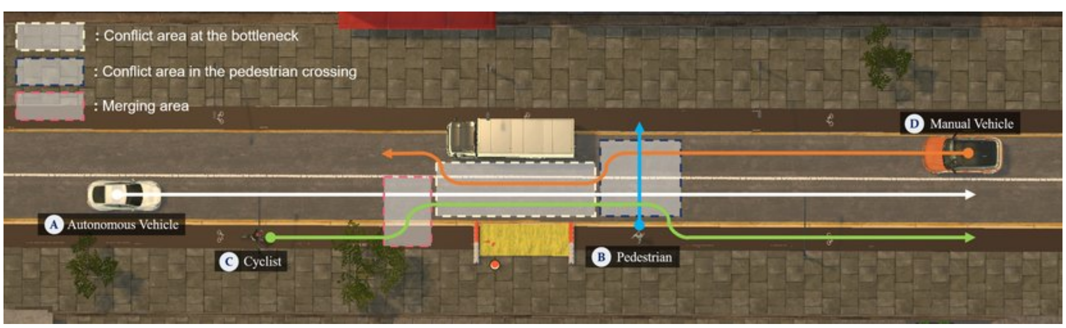


Fig. 3. Visualization of road user routes in the VR environment. Solid lines indicate the paths of each participant, while dashed lines represent areas of potential conflict involving other road users, including the AV. The positions of various road users, such as (A) an autonomous vehicle, (B) a pedestrian, (C) a cyclist, and (D) a manual vehicle driver are marked to illustrate their interactions within these scenarios.

For *cyclists*, merging scenarios are common on future roads and heavily influenced by cultural characteristics, making them crucial to consider [14]. Therefore, we adopted the 'blocked bike lane' scenario by Berge et al., incorporating a construction site on the dedicated bicycle lane [14] and a stopped truck in the manual vehicle driver's lane. This setup required all road users to interact with each other, reflecting the complexity of real-world traffic conditions. Consequently, the cyclist was required to navigate obstacles and merge into the AV's lane (*merging paths*) while coordinating with the manual vehicle driver and the pedestrian to pass through the bottleneck caused by the construction site and the truck. Starting 20 m from the construction site, cyclists stopped 10 m away to receive contextual messages from the AV. Based on these messages, cyclists decided whether to proceed. The cyclists' route is indicated in light green in Fig. 3C, and their perspective is shown in Fig. 5B.

The *manual vehicle drivers* interacted with the AV to navigate a narrow road blocked by a truck and a construction site (*constrained head-on paths*). They also coordinated with the cyclist and the pedestrian to pass through the bottleneck. Starting 40 m from the pedestrian, the manual vehicle driver stopped 2 m away to receive messages from the AV approaching in the opposite lane, then decided whether to cross based on the eHMI message. The manual vehicle drivers' route is indicated in orange in Fig. 3D, and their perspective is shown in Fig. 5C.

In the *Whom* condition, if the AV yielded to another road user (i.e., not the participant), participants were asked to wait until both the AV and the yielding target had passed before interacting with the remaining road users. This behavioral constraint was designed to isolate the effect of the Whom eHMI from other confounding factors and to help determine whether participant behavior reflected their understanding of the signal. As a result, road users other than the AV and the yield recipient were still responsible for assessing the situation and deciding whether it was safe to cross. In the *No Context, When, or Where eHMI* conditions, where the AV displayed a non-yielding signal, participants crossed after interacting with each other once the AV had passed. There was no intervention from the experimenter in the overall interactions among road users; participants interacted naturally to complete their tasks.

## 3.2 Experimental Setup and System Architecture for eHMI Evaluation in AV-Multi-User Scenarios

*3.2.1 Pedestrian Setup.* The pedestrian environment utilized a wireless *HTC Vive Pro Eye HMD* to enable unrestricted movement (Fig. 4B), allowing greater mobility within the interaction space. To enhance natural interaction, inverse kinematics (IK) was employed to simulate the pedestrian's walking motion, while hand movements were visualized using VIVE controllers to improve immersion.

*3.2.2 Cyclist Setup.* Cyclists were equipped with a *Deuter Trainer Cycle Roller* and an *Elite Sterzo Smart* to stabilize the bicycle (Fig. 4C). In the VR environment, two VIVE controllers were used to control the brakes and handlebar steering (Fig. 4D), while a VIVE tracker attached to the pedal detected pedaling motion. Bicycle speed was calculated based on the tracker's angular velocity relative to the gear center, following the method by Noren et al. [98], ensuring realistic cycling dynamics.

*3.2.3 Manual Vehicle Driver Setup.* Manual vehicle drivers used an *Atomic A3 2-DOF Racing/Driving Simulator* combined with a *Logitech G920 Driving Force Steering Wheel* to replicate real-world driving (Fig. 4E) [64, 134]. The driving simulator provided vehicle motion feedback, enabling participants to control the vehicle's direction with the steering wheel and adjust its speed using the pedal. Hand movements were captured using *Hi5 VR Gloves* and trackers [60], while upper-body movements were simulated with IK to facilitate natural interactions. A seated graphical representation of a human was included in the manual vehicle to visually represent the driver.

*3.2.4 Overall Environment and Server-Client Communication.* The VR-based multi-agent simulation environment was developed using Unity with a client-server architecture. It employed the KCP protocol, a lightweight transport layer protocol [117], to ensure reliable and low-latency synchronization of participant movements and positions

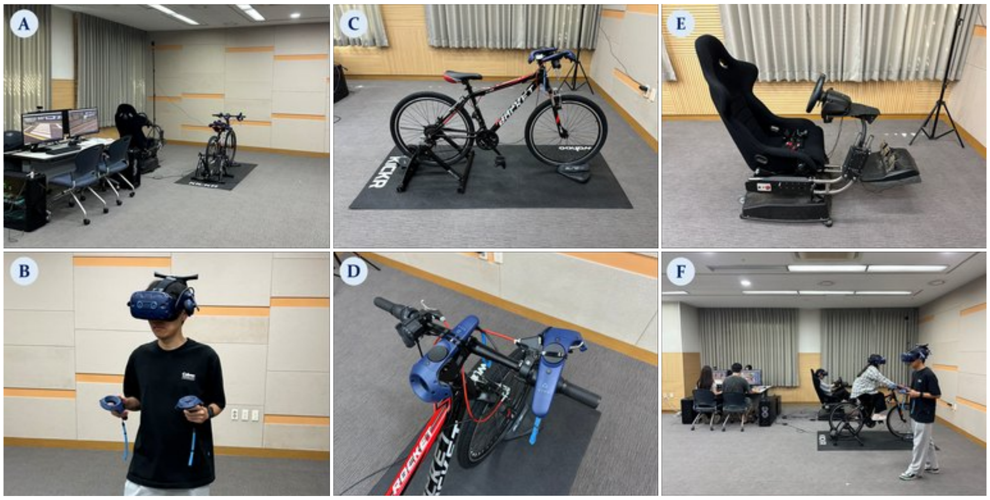


Fig. 4. Experimental setup for the evaluation platform used to assess the effectiveness of eHMIs in AV-multi-road user conflict scenarios. (A) Overview of the experimental setup. (B) Pedestrian setup, utilizing a wireless head-mounted display (HMD) to ensure a wider interaction space. (C) Cyclist setup, highlighting their position within the VR environment. (D) A Vive controller used to control the bicycle's direction and speed. (E) Manual vehicle driver setup, incorporating a driving simulator to replicate the driving environment for simulating realistic vehicle control in the VR simulation. (F) Photograph of the experiment in progress, capturing the real-time interaction of a participant with the VR system.

in real-time (Fig. 12 in Appendix). The experiment was conducted in a 5 × 6 m physical space, with participants wearing *HTC Vive Pro Eye* headsets. *Empatica E4 wristbands* worn on participants' non-dominant wrists, were used to collect EDA data. In the virtual environment, pedestrians, cyclists, and manual vehicle drivers were represented by distinct avatar models, while the AV was distinguished by the absence of a driver model.

In the system architecture, the pedestrian environment functioned as the server, with the cyclist and manual vehicle driver environments acting as clients. Participants' movements, including location and direction, were synchronized in real-time between the server and clients. This setup enabled all participants to observe each other's movements within the same virtual environment, facilitating realistic simulations of multi-road user conflicts.


Fig. 5. VR perspectives for each road user in the experiment: (A) pedestrian's viewpoint, (B) cyclist's viewpoint, and (C) manual vehicle driver's viewpoint. Each viewpoint was rendered in first-person to reflect the natural visual field and task context of each user group. Participants interacted with the simulated urban traffic environment from these perspectives while evaluating the behavior of AVs.

### 3.3 Procedure

Participants received a brief overview of the purpose of the study, the experimental setup, scenarios, and simulator usage, particularly for manual vehicle drivers and cyclists. To aid understanding, they were given an orientation that included an explanatory video and additional instructions. The video dynamically illustrated the components and meaning of each eHMI, as well as how each interface operated in context when mounted on the AV. They were informed that the signals were introduced as representations of the AV's yielding intention rather than to serve as explicit instructions. For instance, the Whom eHMI indicated which road user the AV intended to yield to. Participants then signed a consent form and completed a demographic questionnaire, which included questions on VR and driving experience (for drivers and cyclists). Before starting the experiment, participants familiarized themselves with the simulator through a 5- to 10-minute practice session until they felt comfortable. During the experiment, participants followed predetermined routes based on their assigned roles, stopping at designated points to observe the AV's eHMI and decide whether to cross. After each trial, participants completed a questionnaire about their subjective feelings as described in Section 3.1.2. The experiment consisted of 12 trials with variations in the AV's yielding behavior and type of contextual information. After completing all trials, participants participated in a structured interview to provide feedback. The entire session lasted approximately two hours, with ample breaks provided whenever needed.

### 3.4 Participants

A total of 42 participants aged between 20 and 33 years (20 males and 22 females; $M$ = 24.6 years, $SD$ = 2.9 years) were recruited and equally distributed into three groups of 14: pedestrians, cyclists, and manual vehicle drivers. The demographic details for each group are as follows: pedestrians (5 males, 9 females; $M$ = 24.0 years, $SD$ = 1.6 years), cyclists (7 males, 7 females; $M$ = 24.4 years, $SD$ = 3.5 years), and manual vehicle drivers (8 males, 6 females; $M$ = 25.2 years, $SD$ = 3.3 years). All participants had a minimum corrected vision of 0.7 and no impairments affecting walking or driving. Manual vehicle drivers held valid licenses and drove more than 1.5 hours per week ($M$ = 5.38, $SD$ = 3.78), while cyclists rode bicycles for a similar duration ($M$ = 3.12, $SD$ = 2.19). All except seven of the participants had prior VR experience and none reported severe discomfort with VR or head-mounted display usage. The experiment was approved by the Institutional Review Board.

## 4 RESULTS

We conducted statistical tests on the quantitative data related to crossing performance, subjective feelings, and physiological response, and performed a qualitative analysis of the interview responses. Mixed-design ANOVA was employed to compare the quantitative data across different conditions, except for *clarity*. The *clarity* data, measured using a single-item Likert scale, was analyzed using the Aligned Rank Transform (ART) method to address the violation of normality assumptions. Post-hoc pairwise comparisons were conducted using the Bonferroni correction [42]. Mauchly's test of sphericity was applied to all data, and in cases where sphericity was violated, the Greenhouse-Geisser correction [50] was used to adjust the results accordingly. Normality was assessed for all statistical tests following Kline's method [69], with all data satisfying normality requirements (skewness between -2 and 2, kurtosis between -7 and 7). After confirming the normality assumptions, we assessed the reliability of each scale using Cronbach's alpha [88]. The results revealed that both the *perceived safety* scale and the *trust* scale demonstrated internal consistency ($\alpha$ = 0.715 and $\alpha$ = 0.795, respectively).

Unlike other eHMI types, the *Whom eHMI* was presented three times per participant. Specifically, each participant encountered the *Whom (Other-targeted) eHMI* twice to ensure that the AV yielded to each road user group in a balanced manner across sessions. However, in the *Whom (Other-targeted)* trials, participants were instructed to wait when the AV yielded to another road user, which may have constrained their natural responses.

To ensure fair comparisons across all eHMI conditions, only yielding trials were included in the statistical analysis. Accordingly, for the *Whom eHMI*, only the *Whom (Self-targeted)* trials—where the AV yielded directly to the participant—were analyzed.

## 4.1 Crossing Performance

Based on the descriptive statistics, we observed that when the eHMI indicated a yielding intention directed at the participant (i.e., *Whom (Self-targeted)*), it led to the shortest COT and shortest CD compared to other eHMI conditions for all road user groups, as shown in Fig. 6 and Table 1.

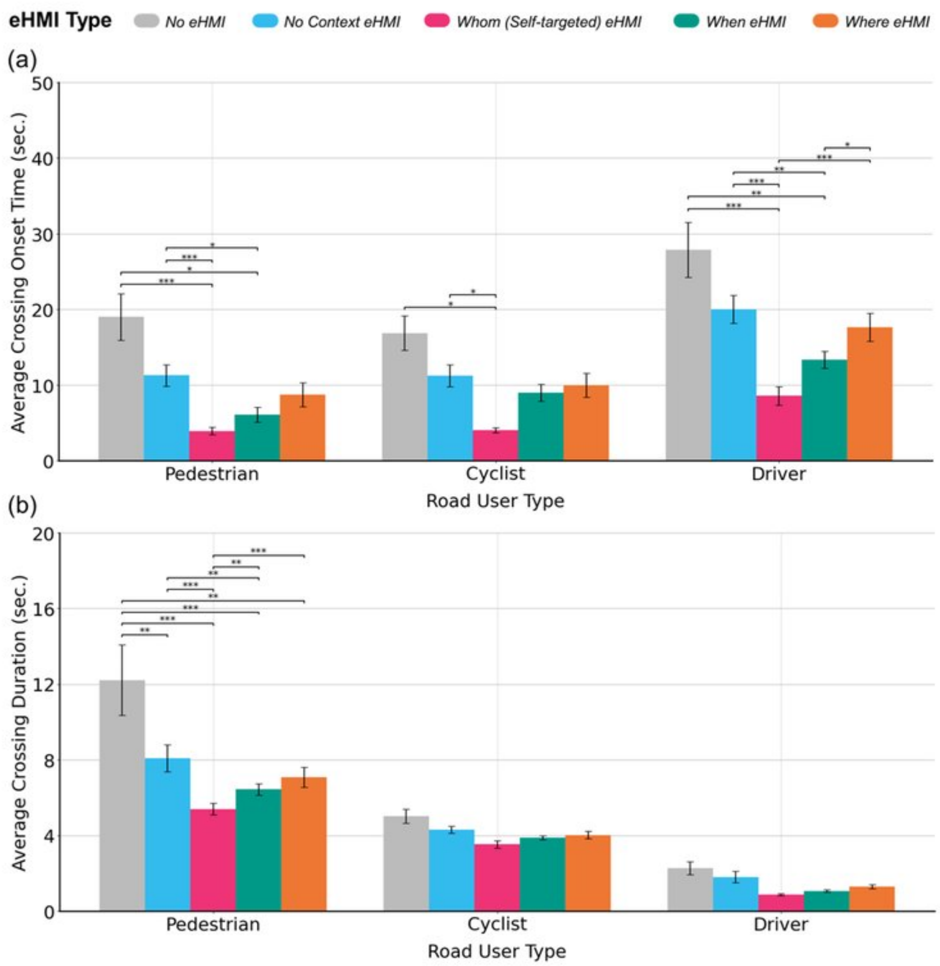


Fig. 6. (a) COT and (b) CD by road user type and eHMI type. The top bar plot shows the average COT for different road user types (*pedestrians*, *cyclists*, and *manual vehicle drivers*) under each eHMI condition (*No eHMI*, *No Context eHMI*, *Whom eHMI*, *When eHMI*, and *Where eHMI*). The bottom bar plot shows the average CD for the same road user types and eHMI types. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

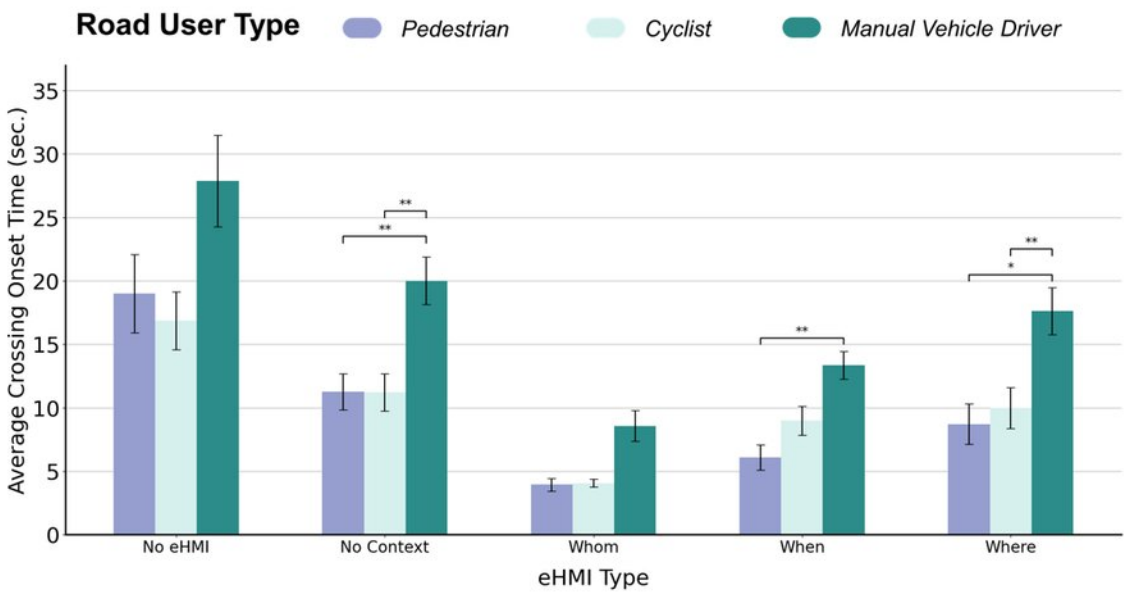


Fig. 7. Impact of eHMI types on COT across road user types. The bar plot illustrates the average COT across pedestrian, cyclist, and manual vehicle driver groups under different eHMI types (*No eHMI*, *No Context eHMI*, *Whom eHMI*, *When eHMI*, and *Where eHMI*). $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

***Crossing Onset Time (COT).*** The eHMI type had a significant main effect on the time it took for participants to begin crossing across all road user groups. Among all groups, the *Whom (Self-targeted) eHMI*, where the AV yielded to the participant, resulted in the shortest COT, particularly compared to the *No eHMI* and *No Context eHMI* (see Table 1).

The road user type had a significant main effect on the time it took for participants to begin crossing across *No eHMI*, *No Context eHMI*, *Whom (Pedestrian) eHMI*, *Whom (Cyclist) eHMI*, *When eHMI*, and *Where eHMI*. Additionally, eHMI type and road user type had a significant interaction effect ($F(5.316, 103.665) = 11.617$, $p < 0.001$). For pedestrians, both the *Whom (Self-targeted) eHMI*, where the AV yielded to pedestrians, and the *When eHMI* led to significantly shorter COTs compared to the *No eHMI* and *No Context eHMI*. Similar patterns were observed for cyclists. *Whom (Self-targeted) eHMI*, where the AV yielded to the cyclists, resulted in significantly shorter COTs compared to the *No eHMI*, *No Context eHMI*. For manual vehicle drivers, *Whom (Self-targeted) eHMI*, where the AV yielded to manual vehicle drivers, also showed significantly shorter COTs compared to the *No eHMI*, *No Context eHMI*, *Where eHMI*. In addition, the *When eHMI* led to significantly shorter COTs compared to the *No eHMI*, *No Context eHMI*, and *Where eHMI*.

Our findings emphasize the importance of providing clear contextual information through the eHMI, especially when specifying which road user the AV intends to yield to. The direct communication of yielding intent to the specific road user (*Whom eHMI*) consistently resulted in quicker decisions to cross, highlighting the role of context in facilitating faster and more confident road user interactions. The observed patterns are visualized in Fig. 6 (a), which illustrates the key significant effects. Additional details—including the statistical significance of the impact of eHMI types across road user types, the main effects, all significant pairwise comparisons, and the comparisons across road user types—are provided in Fig.7, as well as in Table 2 (a) and (b), Table 3, and Table 4 in Appendix B.

***Crossing Duration (CD).*** The eHMI type had a significant main effect on the duration of time participants spent crossing across all road user groups. Excluding *No eHMI*, the *No Context eHMI* consistently resulted in the longest CD across all road user groups, whereas the *Whom (Self-targeted) eHMI* led to the shortest CD (see Table 1). For pedestrians, *Whom (Self-targeted) eHMI*, where the AV yielded directly to the pedestrian, resulted in significantly shorter CDs compared to the *No eHMI*, *No Context eHMI*, *When eHMI*, and *Where eHMI*. Moreover, the *When eHMI* led to significantly shorter CDs compared to the *No eHMI* and *No Context eHMI*, while the *Where eHMI* led to significantly shorter CDs compared to the *No eHMI*. No significant differences were observed for cyclists or manual vehicle drivers.

Our findings on CD further emphasize the importance of delivering clear contextual information through eHMIs, particularly for pedestrians. Specifically, the direct communication of yielding intent to the direct road user (*Whom eHMI*) consistently resulted in shorter CDs. The overall trends are visualized in Fig. 6 (b), which highlights the significant effects. Further details on the statistical significance of the main effects and all significant comparisons are provided in Table 2 (c) and Table 3 (b) in the Appendix B.

***Error Rate.*** The error rate was calculated by logging instances where participants misunderstood whether the AV was yielding, based on the different types of contextual information, leading them to make decisions contrary to the AV's intentions. Descriptive statistics for the error rate are presented in Table 1.

No errors were observed when the AV did not yield, across all conditions. However, errors occurred in some conditions when the AV did yield. The highest error rate was observed in the *No Context eHMI* and *Whom (Cyclist) eHMI* conditions, where 2 out of 42 participants (4.76%) made errors in each case. In the *No Context eHMI*, a cyclist and a manual vehicle driver collided while attempting to cross based on the AV's signal. In the *Whom (Cyclist) eHMI*, a pedestrian and a manual vehicle driver made errors when the AV yielded. Additionally, in the *Whom (Pedestrian) eHMI* and *Whom (Driver) eHMI* conditions, a manual vehicle driver and a cyclist made errors respectively, each accounting for 2.38%. Notably, no errors were recorded in the *No eHMI*, *When*, or *Where eHMIs*, indicating that participants correctly interpreted the AV's message in those scenarios. Our findings highlight the need for unambiguous designs to minimize misunderstandings, particularly in scenarios where the yielding intent is directed at specific road users.

## 4.2 Subjective Feelings and Physiological Response

*4.2.1 Subjective Feelings.* We found that *No eHMI* led to the lowest subjective feelings of perceived safety, trust, and clarity among participants. Excluding *No eHMI*, the *No Context eHMI* resulted in significantly lower subjective feelings compared to eHMIs that provided contextual information (see Fig. 8).

Table 1. Descriptive statistics for COT, CD, and Error Rate by eHMI type and road user type. Mean and standard deviation (SD) values are reported for each combination of eHMI type (*No eHMI*, *No Context eHMI*, *Whom eHMI*, *When eHMI*, and *Where eHMI*) and road user type (*pedestrians*, *cyclists*, and *manual vehicle drivers*).

| eHMI Type | **COT** (sec.) | | | **CD** (sec.) | | | **Error Rate** | |
|---|---|---|---|---|---|---|---|---|
| | **Pedestrian** Mean/SD | **Cyclist** Mean/SD | **Manual Vehicle Driver** Mean/SD | **Pedestrian** Mean/SD | **Cyclist** Mean/SD | **Manual Vehicle Driver** Mean/SD | AV yielding | AV non-yielding |
| No eHMI | 19.00/11.49 | 16.88/8.50 | 27.87/13.53 | 12.21/6.95 | 5.03/1.42 | 2.28/1.27 | 0/42 | 0/42 |
| No Context | 11.28/5.29 | 11.23/5.45 | 20.03/6.98 | 8.09/2.66 | 4.30/0.69 | 1.82/1.18 | **2/42 (4.76%)** | 0/42 |
| Whom (Pedestrian) | **3.94/1.82** | 16.05/6.81 | 16.27/7.24 | **5.40/1.15** | 3.92/0.71 | 1.33/0.31 | **1/42 (2.38%)** | 0/42 |
| Whom (Cyclist) | 15.38/6.15 | **4.08/1.22** | 26.05/8.50 | 6.17/1.64 | **3.54/0.73** | 1.21/0.47 | **2/42 (4.76%)** | 0/42 |
| Whom (Driver) | 12.98/6.12 | 8.99/4.91 | **8.58/4.53** | 6.53/1.70 | 4.02/0.63 | **0.89/0.20** | **1/42 (2.38%)** | 0/42 |
| When | 6.10/3.75 | 9.00/4.32 | 13.35/4.13 | 6.44/1.12 | 3.90/0.41 | 1.08/0.27 | 0/42 | 0/42 |
| Where | 8.72/5.98 | 9.99/5.97 | 17.63/6.86 | 7.09/1.97 | 4.03/0.74 | 1.30/0.37 | 0/42 | 0/42 |

*Reference only: Pairwise comparisons including Whom (Other-targeted) eHMI conditions are reported here for transparency, but were excluded from the main statistical analysis.*

***Perceived Safety.*** The eHMI type had a significant main effect on perceived safety across all road user groups. Excluding *No eHMI*, the *No Context eHMI* consistently resulted in the lowest perceived safety across all road user groups. The interaction effect between eHMI type and road user type was not significant ($F(9.427, 183.833) = 0.959$, $p = 0.478$). Given this result, we examined the main effect of eHMI type. The *Whom eHMI* resulted in significantly higher perceived safety compared to the *No eHMI*, *No Context eHMI*, and *Where eHMI* across all road user types. In addition, the *No Context eHMI*, *When eHMI*, and *Where eHMI* each led to significantly higher perceived safety compared to the *No eHMI*.

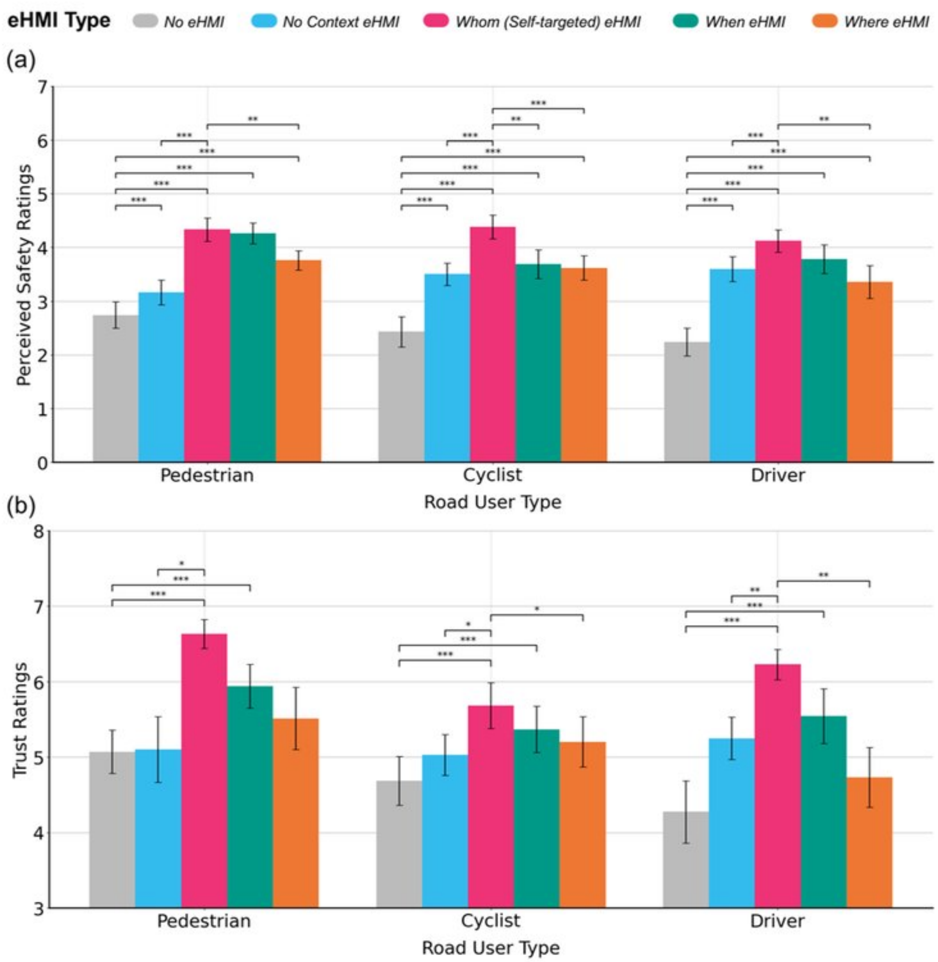


Fig. 8. (a) Perceived Safety and (b) Trust by road user type and eHMI type. The top plot shows perceived safety ratings and the bottom plot shows trust ratings, each across eHMI types (*No eHMI*, *No Context eHMI*, *Whom eHMI*, *When eHMI*, and *Where eHMI*) and road user types (*pedestrians*, *cyclists*, and *manual vehicle drivers*). $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Our findings emphasize the importance of providing specific contextual information, especially including *Whom* details, to enhance perceived safety among road users. Details on the statistical significance of the main effects and all significant comparisons can be found in Table 2 (d) and Table 3 (c) in the Appendix B.

***Trust.*** The eHMI type had a significant main effect on trust across all road user groups. Excluding *No eHMI*, the *No Context eHMI* consistently resulted in the lowest level of trust except for drivers, while the *Whom eHMI* led to higher levels of trust. Given that the interaction effect between eHMI type and road user type was not significant ($F(7.980, 155.617) = 1.136$, $p = 0.342$), we examined the main effect of eHMI type. The *Whom eHMI* resulted in significantly higher trust compared to the *No eHMI* and *No Context eHMI* across all road user types.

Our findings highlight the critical role of providing specific contextual information, including *Whom* details, as this consistently led to higher levels of trust compared to *No eHMI*. Details on the statistical significance of the main effects and all significant comparisons are provided in Table 2 (e) and Table 3 (d) in the Appendix B.

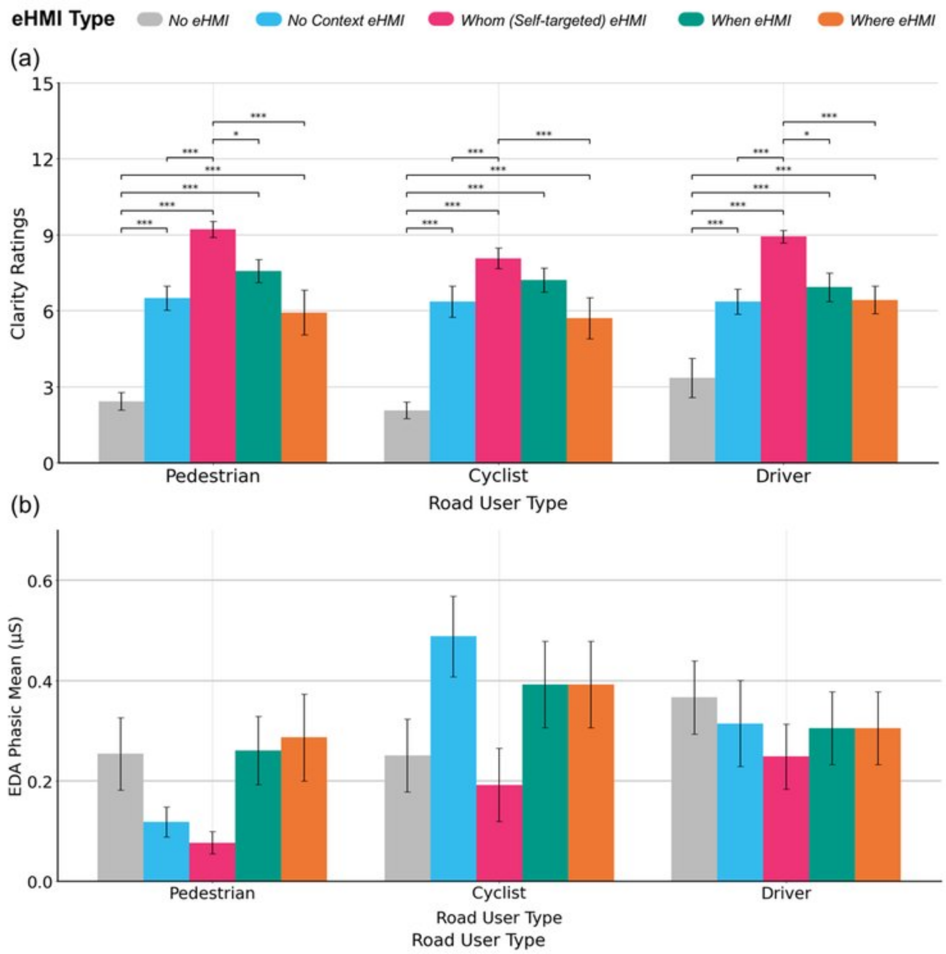


Fig. 9. (a) Clarity and (b) EDA Phasic Mean by road user type and eHMI type. The top plot shows clarity ratings and the bottom plot shows EDA Phasic Mean ratings, each across eHMI types (*No eHMI*, *No Context eHMI*, *Whom eHMI*, *When eHMI*, and *Where eHMI*) and road user types (*pedestrians*, *cyclists*, and *manual vehicle drivers*). $p < .05$, $^{**}p < .01$, $^{***}p < .001$.

***Clarity.*** The eHMI type had a significant effect on clarity across all road user groups. Among all conditions, the *Whom eHMIs*, particularly the *Whom (Self-targeted) eHMI* for the pedestrian and manual vehicle driver groups, led to the highest clarity across all road user groups. The interaction effect between eHMI type and road user type was not significant ($F(12, 182) = 0.360$, $p = 0.975$), so we examined the main effect of eHMI type. The *Whom eHMI* resulted in significantly higher clarity compared to the *No eHMI*, *No Context eHMI*, and *Where eHMI* across all road user types. In addition, the *No Context eHMI*, *When eHMI*, and *Where eHMI* each led to significantly higher clarity compared to the *No eHMI*.

These findings emphasize the importance of providing clear contextual information, especially *Whom* details, in improving the clarity of communication between AVs and road users. Details on the statistical significance of the main effects and all significant comparisons can be found in Table 2 (f) in the Appendix B.

*4.2.2 Physiological Response.* The raw EDA data were processed using Ledalab [13] to extract two key components—the phasic (short-term, fast-changing) and tonic (long-term, slow-changing). This preprocessing involved smoothing the data with an 8-window Gaussian filter and applying a 1st-order low-pass filter with a 0.2 Hz cutoff frequency. Continuous decomposition analysis was then performed to derive the continuous phasic and tonic activity from the EDA data.

To account for inter-individual differences in skin conductance, we applied min-max normalization to the phasic EDA data to ensure comparability across participants [67, 94]. The signal was scaled between 0 and 1, using each participant's minimum and maximum values as reference points. Our analysis focused on the phasic component, which reflects immediate physiological reactions following eHMI activation. For each condition, we analyzed the normalized phasic EDA responses during the 5 seconds after eHMI activation to assess participants' arousal [49, 118]. Specifically, we examined the mean value of EDA phasic data within this time window. This metric is commonly used to capture overall arousal levels, as it reflects the aggregated intensity of short-term sympathetic nervous system activation triggered by external stimuli [76, 114]. In addition, data from some participants were unavailable due to technical issues with the Empatica E4 wristbands. Specifically, the analysis included 8 cyclists, 11 pedestrians, and 12 manual vehicle drivers from the initial 14 participants per group.

The eHMI type had a significant main effect on EDA phasic mean values both for pedestrians and cyclists, while no significant effect was observed for manual vehicle drivers. For cyclists, *Whom eHMI*, where the AV yielded directly to the manual vehicle drivers, resulted in significantly lower EDA phasic mean value compared to *No Context eHMI*. Our findings indicate that different eHMI types lead to variations in event-related arousal levels, with cyclists showing significantly lower arousal in response to *Whom eHMI*. Details on the statistical significance are provided in Table 2 (g) and Table 3 (e) in the Appendix B.

### 4.3 Qualitative Feedback

We conducted structured interviews to gather their subjective opinions on preferences, communication efficiency, and comprehension through the contextual eHMIs. The qualitative feedback reinforced our quantitative findings, with participants expressing a clear preference for interfaces that provided intuitive, context-rich information. Overall, the *Whom eHMI* emerged as the most favored and easiest to comprehend across different road user groups, while the absence of eHMI cues was consistently deemed least effective. Although participants generally appreciated straightforward cues like *No Context eHMI*, those that offered richer contextual detail—when well executed—appeared to foster both trust and a stronger sense of engagement.

***Preferences.*** A total of 28 out of 42 participants selected *Whom eHMI* as their preferred condition due to its intuitiveness and ease of distinction (P3, P9-10, P14-15, P20, P22, P24, P27-28, P30, P37, P40-41), as well as its clarity in prioritizing decision-making (P1, P4-6, P7-8, P11, P19, P23). They also deemed it the most trustworthy (P13, P20, P26), and some noted that it provided a strong sense of interacting with the AV (P2, P20). Following that, 11 participants preferred *No Context eHMI* as it provided the AV's direct intent (P31, P33, P34-35, P38-39). Three participants preferred *When eHMI* for its straightforward (P12, P32) and culturally adaptable approach (P17), and appreciated the stopping time information for decision-making (P25). The distribution of preferences by road user

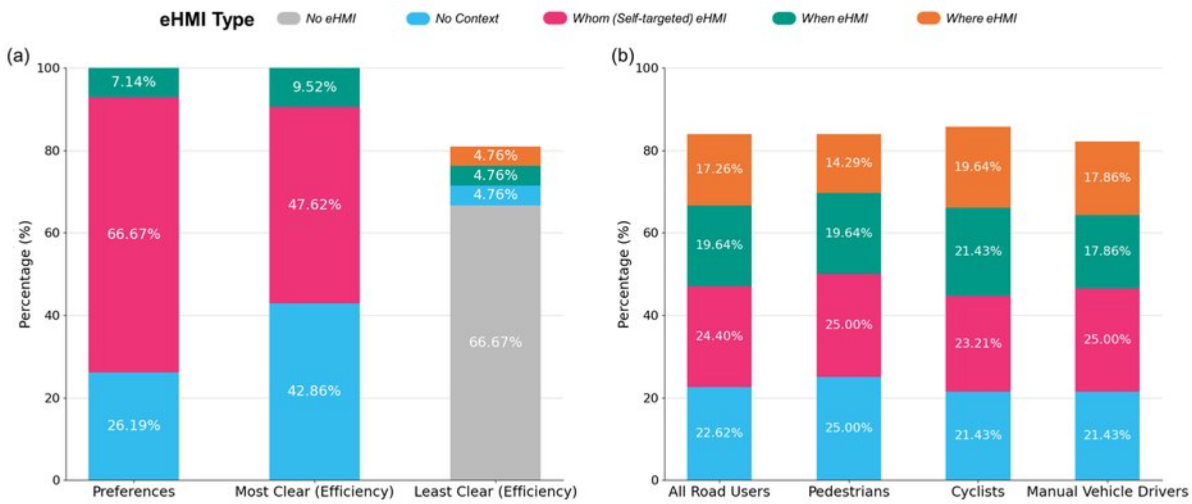


Fig. 10. Distribution of (a) preferences and communication efficiency among all road users, and (b) comprehension levels across different road user groups. In (a), each stacked bar chart shows the percentage distribution for each eHMI type across 'Preferences', 'Most Clear (Efficiency)', and 'Least Clear (Efficiency)' conditions. In (b), each stacked bar chart shows the percentage distribution for 'No Context', 'Whom', 'When', and 'Where' eHMI types. The comprehension of each eHMI type is normalized to a maximum of 25%, summing to 100%.

group is shown in the bar chart in Fig. 10 (a). This reflects a need for interfaces that balance intuitiveness with sufficient context-simple enough to be quickly understood, yet detailed enough to inform safe, efficient behavior.

***Communication Efficiency.*** When asked about clarity and effectiveness, participants consistently rated the *Whom eHMI* highest, with 20 out of 42 identifying it as the most effective, followed by 18 who chose *No Context eHMI* and four who favored *When eHMI*. Conversely, they viewed *No eHMI* as the least effective option. This pattern held across different user groups, though some variations emerged. For instance, pedestrians found *No Context eHMI* somewhat more effective than Whom, possibly due to its direct simplicity, while cyclists and manual vehicle drivers showed a stronger preference for *Whom eHMI*'s more contextual cues.

In other words, while the *Whom eHMI* generally excelled in communication efficiency, simpler interfaces like *No Context eHMI* still maintained appeal in certain groups (e.g., pedestrians), suggesting that different road users may weigh factors like complexity and detail differently. Ultimately, the results indicate that context-driven designs can enhance clarity, but a certain degree of simplicity may resonate with some users depending on their perspective or role in the traffic environment. The distribution of communication efficiency for all road users is shown in the bar chart in Fig. 10 (a).

***Comprehension.*** Comprehension patterns closely mirrored preferences and communication efficiency. The *Whom eHMI* achieved the highest comprehension across user types, enhanced by its resemblance to familiar traffic signals (P2-3, P10-13, P24, P36) and its clearly assigned yielding targets (P31). Although *No Context eHMI* was easy to understand (P24, P30-31), some participants struggled with its symbolic elements, like the "mouth" indicator, as these did not align well with established traffic conventions (P6, P18). For the When eHMI, participants generally grasped the timer-based logic, finding it culturally adaptable and intuitively representative of time (P11) because this is a commonly used approach for conveying temporal information [38]. However, some remained uncertain about whether they should wait for a complete stop before crossing (P20, P37). The *Where eHMI*'s combination of dynamic symbols and text supported comprehension for a few (P11, P24), but others encountered issues with information overload (P1, P15, P18, P40), text visibility (P13), or directional ambiguity—particularly cyclists (P14, P20), whose perspective sometimes clashed with the presented information. The distribution of comprehension by road user group is shown in the bar chart in Fig. 10 (b).

***Additional Feedback.*** Participants' general comments and suggestions shed further light on the conditions needed for effective eHMI deployment. Some noted that initial understanding could be challenging without prior explanation (P35, P37), while others wanted clearer post-yielding information, particularly for *Whom eHMI* (P11, P13, P33, P36). Several participants raised concerns specifically related to the contextual eHMIs, particularly *When* and *Where*. Several cyclists felt anxious or distracted when having to look sideways at the display (P8, P14, P32).

A few participants suggested using a color-coded system similar to traffic signals (e.g., green for yielding, red for not yielding) (P18, P34, P36). Other recommended incorporating alternative modalities (e.g., auditory signals or blinking animations) for users who might find visual interfaces challenging, such as those who are colorblind or visually impaired (P17, P32, P35, P41). Others suggested refining specific design elements in *No Context eHMI* incorporating a more advanced anthropomorphic interface such as facial expressions (P14), or using symbols commonly understood on roads (P18). For *Whom eHMI*, suggestions included providing additional information after the intended yield target had completed crossing, such as displaying an orange symbol to indicate non-yielding to all potential interaction partners (P36), or providing information on yielding priority (P34). For *When eHMI*, participants recommended including the remaining stopping time in text format, either as standalone numbers (P13, P15, P20-21, P37) or in addition to the timer (P14, P6). They also recommended using non-continuous displays, such as interfaces that indicate battery indicators (P36), or providing details on the duration of yielding (P19). Recommendations for *Where eHMI* suggested using either only dynamic animations (P8, P20, P40) or only text indicating the remaining distance (P1, P4, P9, P13, P21-22, P32).

Adjusting the interface's position—on the windshield (P3, P16, P19, P28), roof (P12), using a projection (P18), or delivering information directly to individual road users via personal devices or AR glasses (P6, P21)—could improve visibility and reduce cognitive load. In summary, participants envisioned more versatile, user-tailored designs that integrate familiar cues, multiple sensory channels, and carefully calibrated detail levels to better accommodate the diverse needs of all road users.

## 5 DISCUSSION

### 5.1 Effects of Contextual eHMIs

*5.1.1 Summary of Experimental Findings.* The following discussion addresses our research question by summarizing the key findings. The results demonstrate that contextual eHMIs significantly enhanced participants' crossing performance and subjective feelings across all user groups compared to non-contextual eHMI. Specifically, the *Whom eHMI* showed the best performance in COT, CD, trust, and clarity for all user groups. Furthermore, most participants expressed a preference for *Whom eHMIs*, reporting the highest effectiveness in communication and comprehension.

However, while the *Whom eHMI* exhibited the best performance in COT across all types of user groups, it is important to consider that a shorter COT may not always translate to improved traffic safety or reduced collision risks. For instance, participants may unconditionally accept the AV's yielding intention without fully assessing the surrounding context, potentially overlooking the rights of other road users or broader traffic conditions. Such insufficiently considered decisions could compromise safety [135], especially in high-traffic environments where right-of-way dynamics are complex. These findings emphasize the importance of designing eHMIs that promote both clear communication and an awareness of the broader traffic context.

Building on these considerations, the *Whom eHMI* also included a procedural constraint instructing participants to wait when the AV yielded to other road users. Despite this, the observed behavioral outcomes offer insight into how contextual eHMIs support participants in interpreting traffic situations. In the *Whom eHMI*, the instruction to wait was introduced to maintain consistency in how the signal was interpreted and to support more focused observation of participants' response timing based on the eHMI. While the scenario included a behavioral constraint to guide interpretation, the observed consistency in crossing behavior across user groups suggests that participants could reliably identify whether the AV's message was directed at them. These findings suggest that contextually targeted eHMIs have the potential to support road users in determining the relevance of AV signals in shared spaces involving multiple traffic participants. In comparison, participants in the *When* and *Where eHMIs* exhibited longer response times and may have prompted participants to pay closer attention to surrounding traffic conditions before making their own decisions. The absence of targeted information in these conditions likely increased uncertainty, leading to greater hesitation.

Finally, the *No eHMI* exhibited the poorest performance across all user groups in terms of COT, CD, perceived safety, trust, and clarity. Similarly, the *No Context eHMI* showed the second lowest performance in CD, perceived safety, and trust for pedestrians and cyclists. These outcomes suggest that the absence of contextual cues may have made it more difficult for participants to determine the AV's intentions, potentially contributing to delayed decision-making and reduced confidence during crossing. In summary of these results, contextual eHMIs appear to support more coordinated behavior and improved user experience by offering information that helps participants interpret the AV's yielding intentions.

*5.1.2 Interpreting the Effectiveness of Contextual eHMIs.* While contextual eHMIs consistently outperformed non-contextual alternatives across most behavioral and subjective measures, their superior performance may not be solely attributable to the presence of contextual information. Specifically, to convey information regarding the intended recipient (e.g., *Whom*), timing (e.g., *When*), or stopping location (e.g., *Where*), the contextual eHMIs employed symbolic and dynamic visual cues that may have enhanced interpretability. This raises the possibility that participants' improved crossing performance and higher ratings of subjective responses may have been influenced by both the richness of contextual information and the intuitiveness of the visual implementation.

All contextual eHMIs used animated elements such as road user icons, dynamic timers, or spatial markers, that may have provided clearer visual affordances than the static smiley symbol (e.g., *No Context eHMI*), which does not align with conventional traffic communication standards. Participants who were less familiar with abstract or unconventional symbols may have found these visually grounded cues easier to interpret. Such an interpretation is supported by interview responses that emphasized the clarity and recognizability of context-enriched designs, particularly the *Whom eHMI*. In contrast, although some participants found the mouth-based cues in the *No Context eHMI* intuitive (P24, P30–31), others reported difficulties due to a lack of alignment with established traffic conventions (P6, P18). As a result, the observed advantages of contextual eHMIs may reflect not only the informational benefits of contextual content, but also the perceptual affordances of their specific visual implementations. These findings are consistent with prior research suggesting that the use of familiar and easily interpretable visual symbols can facilitate faster recognition and clearer understanding of their meaning [96].

*5.1.3 Effects of Whom eHMI.* The findings of the study indicate that providing specific information regarding the AV's yielding behavior toward the intended recipient is critical for enhancing decision-making efficiency, particularly in scenarios involving multiple road users. The *Whom eHMI* allows road users to clearly distinguish who is being yielded to and who is not, through distinct symbols and colors. This differentiation minimizes ambiguity, ensuring that road users can confidently interpret whether the AV's message is directed at them, reducing the potential for confusion [40, 41]. Such clarity is essential for ensuring safe and effective interactions in mixed-traffic environments. Incorporating *Whom* information into eHMIs can improve overall road user safety and contribute to a smoother traffic flow, aligning with the importance of clearly designating the intended recipient as discussed in [22]. Beyond supporting previous findings, our research advances the understanding of eHMIs by demonstrating how *Whom* information specifically addresses the complexities of multi-road user conflicts.

Beyond its informational function, the *Whom eHMI* may have also influenced user behavior by activating familiar visual cues. While it was designed to communicate the AV's yielding intention, some participants may have interpreted it not merely as a signal of intent, but as a behavioral suggestion—particularly when the symbol matched their own road user type. One possible explanation is that the design closely resembled elements from existing traffic control systems, such as pedestrian or bicycle icons used in traffic signals. These symbolic resemblance may have activated well-established mental models, leading participants to respond as if the AV were granting permission to proceed. Prior studies have examined the approach of using road user symbols as advisory cues and evaluated their impact on road user behavior [43]. Rather than treating this as a misunderstanding, we interpret it as evidence of how prior experience and learned conventions can shape the perceived meaning of new interfaces. Such findings suggest that future eHMI designs should consider how visual elements may activate mental models rooted in conventional traffic communication.

*5.1.4 Effects of When and Where eHMIs.* The relative performance of *When* and *Where eHMIs* were less distinct compared to *Whom eHMI*, suggesting that their impacts are less straightforward. We anticipate that several factors might have contributed to these findings.

First, When and Where eHMIs offer less clear information regarding crossing decisions, making them less effective compared to *Whom eHMI*. *When eHMI* provides temporal information on *when* the AV will stop; this may be less immediately relevant than the *Whom eHMI*, which explicitly identifies the yielding recipient. In interactions involving multiple road users with varying characteristics, temporal information might be perceived as less critical and clear in the immediate decision-making process than identifying the intended yield target. This is supported by our findings, where the *Whom eHMI* exhibited higher clarity and communication efficiency compared to *When eHMI*. Similarly, *Where eHMI* provides information on the yield location but might not offer immediate clarity regarding the intended recipient or action, contributing to its relatively ambiguous performance.

*5.1.5 Comparison with Previous Studies.* Two prior studies have investigated the potential scalability of eHMIs with contextual information for real traffic applications, similar to our research focus. Dey et al. [40] explored the impact of contextual information in conflict scenarios involving two pedestrians and an AV, while Colley et al. [25] examined eHMI effects in interactions with three pedestrians. Both studies emphasize the necessity of scalable eHMIs in situations with multiple road users and highlight the need for specifying the intended recipient of yield messages, aligning with our findings. However, our findings diverge from theirs in key areas. Dey et al. found that the *Where eHMI*, indicating the AV's stopping location, generally showed high eHMI performance, while Colley et al. reported similar results for the symbol-based *No Context* and *Where eHMIs*. In contrast, our findings suggest a somewhat different outcome, particularly regarding the role of *Whom eHMI* may stem from differences in the characteristics of the interaction partners and the variations in interface design elements used across studies. We interpreted these differences as arising from three main factors—interaction partner composition, eHMI design elements, and visual display modalities.

***Interaction Partner Composition.*** Both studies primarily focused on homogeneous pedestrian groups as interaction partners. Consequently, the results from these studies may differ from our findings, which involved conflict scenarios with diverse road users. These differences can be attributed to variations in the composition of the interacting parties. While these studies provided valuable insights into AV interactions with homogenous VRU groups, our research expands upon their scenarios by incorporating a more complex and diverse set of road users.

***eHMI Design Elements.*** In the case of *Where eHMI*, Dey et al. employed a projection-based animation that visualized the AV's deceleration and displayed a projected crosswalk aligned with the pedestrian's position, offering a clear, action-oriented cue. In contrast, our design conveyed the AV's stopping location and process, aiming to support road users' independent decision-making through a less directive approach. These design differences illustrate broader variations in how directly eHMIs communicate the AV's intent, ranging from projection-based cues that explicitly guide pedestrian behavior to symbolic and textual cues that are more open to interpretation. Such variation in communicative directness may have shaped how the effectiveness of the *Where eHMI* was perceived. Taken together, we suggest that the effectiveness of each approach be understood in the context of its intended use case. Such differences may thus reflect contextual distinctions rather than direct performance differences.

***Visual Display Modalities.*** These studies employed various visual display modalities, such as light bands, windshield displays, and street projections [40], with symbols and text integrated into these designs [25]. The diversity of these modalities may make it challenging to attribute specific effects solely to the eHMI context. In contrast, our study isolates the effects of contextual eHMI communication [46] by focusing on symbols and anthropomorphic elements that are less influenced by cultural differences [9] and designed to be visible from all directions, ensuring clarity across various road user groups.

Our study extends the existing body of research by comparing several eHMI contexts, such as *Whom*, *When*, and *Where*, addressing a crucial gap in understanding how eHMIs can support diverse road user interactions in complex traffic scenarios. These findings could offer broader implications for real-world traffic systems.

## 5.2 Perceived Safety and Trust in AV Communication

Our study revealed notable differences in perceived safety and trust across road user groups, distinct from patterns observed in crossing performance and other subjective feelings (i.e., clarity). Pedestrians and cyclists consistently rated the *No Context eHMI* lowest for both safety and trust, while manual vehicle drivers gave their lowest ratings to the *Where eHMI*.

These divergent perceptions are likely attributed to differences in how information elements were presented and accessed by each user group. For pedestrians and cyclists, the *No Context eHMI*'s minimal information about AV intentions may have failed to provide the explicit cues needed for confidence in the interaction. This aligns with Kamkar et al.'s findings [59] that road users experience greater feelings of safety when interfaces provide specific, direct explanations of AV behavior.

Manual vehicle drivers' lower perceived safety with the *Where eHMI* appears linked to challenges in information processing and visibility. The qualitative responses revealed that drivers struggled to interpret the *Where eHMI*'s small visual elements, often reporting information overload. Their positioning—being furthest from the AV and seated inside their vehicles—likely compromised their ability to clearly view and process the *Where eHMI*'s detailed spatial information. This visibility constraint may have created uncertainty about the AV's intended actions, undermining drivers' perceived safety during interactions.

Interestingly, although the additional spatial information of the *Where eHMI* clarified AV intentions, it did not translate into improved perceived safety or trust among drivers. These findings underscore the importance of considering multiple factors in eHMI design, such as visibility [75] and the potential for information overload [127], rather than relying on a single dimension—particularly when multiple road user groups are involved. Future developments should move beyond one-size-fits-all solutions and tailor eHMI information to the unique needs and constraints of different road users, ultimately supporting more effective and context-sensitive communication in real-world traffic environments.

## 5.3 Scalability of Contextual eHMIs

To explore the scalability of contextual eHMIs beyond simplified scenarios involving a single representative from each road user group, we considered scenarios with varied compositions of the same road user type. In our main setup, participants experienced a structured conflict involving one pedestrian, one cyclist, and one manual vehicle driver interacting with an AV on a two-lane road. To examine how contextual eHMIs perform under increased social complexity, we introduced three scenarios shown in Fig. 11, each involving multiple agents of the same road user type. The configurations included (1) two pedestrians alongside one cyclist and one manual vehicle driver, (2) one pedestrian and two cyclists with one manual vehicle driver, and (3) three pedestrians with one cyclist and one manual vehicle driver. In such cases, the current design of the *Whom eHMI*, which employs generalized icons (e.g., a single road user symbol), may lead to interpretive ambiguity.

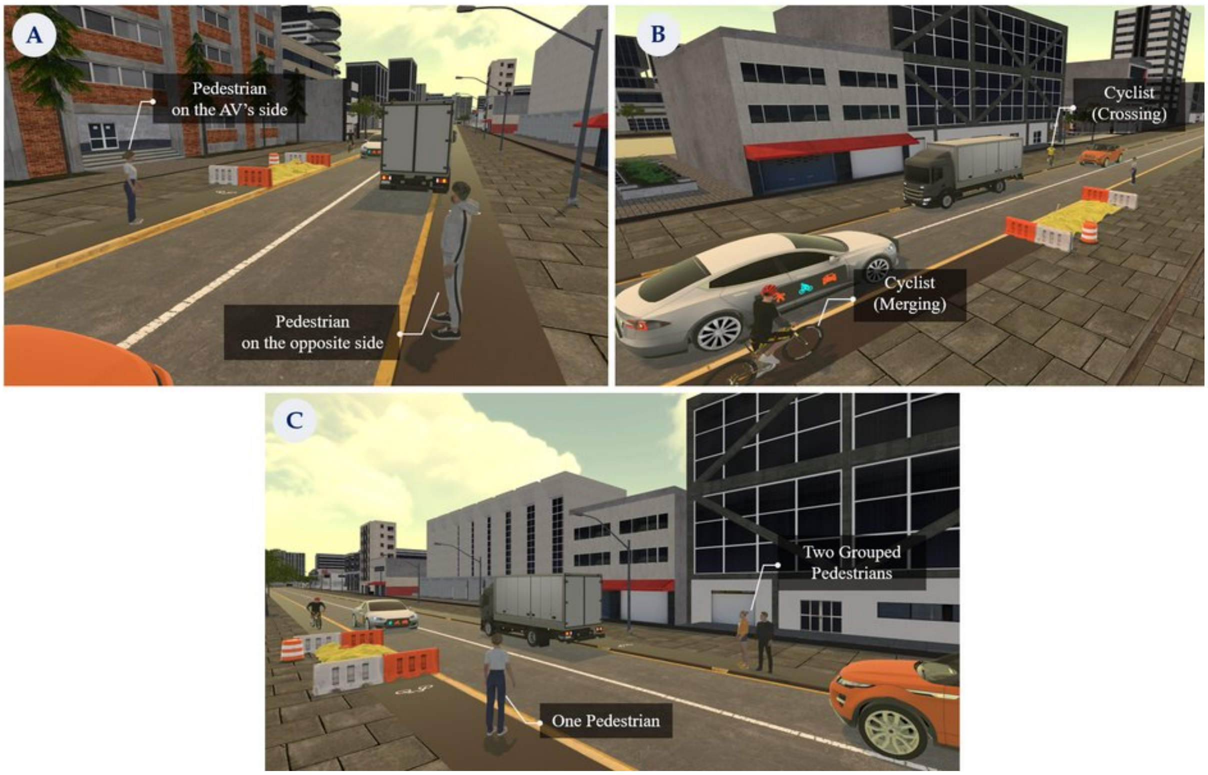


Fig. 11. Scalable scenarios involving multiple agents of the same road user type. (A) Two pedestrians, along with one cyclist, and one manual vehicle driver. (B) Two cyclists, accompanied by one pedestrian and one manual vehicle driver. (C) Three pedestrians, together with one cyclist and one manual vehicle driver.

One possible scenario involves two pedestrians approaching the AV simultaneously—one on the AV's side (as in the scenario used in our experiment) and the other on the opposite side. Both individuals may observe the pedestrian symbol and misinterpret the message as being directed at themselves, even when they are not the intended recipient [55]. In particular, if the pedestrian farther from the AV initiates crossing based on this misinterpretation, they may become vulnerable to a collision with a manual vehicle approaching from another direction. A comparable issue may arise with two cyclists—one approaching from the AV's side and intending to merge (as in the scenario used in our experiment), and the other approaching from the opposite side aiming to cross its path. Although their positions and intended movements differ, the generalized cyclist icon may not adequately distinguish between them, potentially leading to misaligned responses and safety concerns.

In a more complex case involving three pedestrians—one on the AV's side (as in the scenario used in our experiment) and two grouped together on the opposite side—it may be unclear to whom the symbol is directed, and whether it addresses an individual or the group as a whole. If the group on the far side assumes the AV is yielding to them, they may begin crossing and face a potential conflict with a manual vehicle approaching from another direction.

To address such challenges, contextual eHMI designs could be enhanced by explicitly distinguishing the intended recipient in scenarios involving multiple users of the same road user type. Possible improvements include directional cues such as arrows [107] or color-coded signals (e.g., cyan) displayed only on the vehicle surface facing the intended recipient. Additionally, road projections can help align the AV's signals with each user's relative position, thereby clarifying which user the message is intended for. In cases where vehicle-mounted displays may be insufficient, personalized feedback via mobile or wearable devices [128] or communication through connected infrastructure (e.g., V2X communication) [79] may enhance interpretability in complex traffic environments.

### 5.4 Finding a Balance Between Performance and Preference

Our results revealed a notable distinction between the quantitative performance and qualitative preferences of different eHMI designs. The *No Context eHMI* generally showed lower performance compared to the eHMIs incorporating contextual information across all road user groups. In particular, *No Context eHMI* received the lowest ratings among conditions with eHMI from pedestrians and cyclists in CD, perceived safety, and trust, while from drivers, it received the lowest rating in terms of CD. However, qualitative responses exhibited different trends. *No Context eHMI* was rated second-highest across all aspects discussed during interviews, including participant preferences, communication efficiency, and comprehension.

This discrepancy between performance and preference can be attributed to the varying amount of information provided by each eHMI design. While eHMIs with contextual information facilitated earlier crossing decisions, the simplicity of the *No Context eHMI* appeared to resonate more with users' preferences. Interview responses highlighted the familiarity of *No Context* (P39), its ease of comprehension (P33, 38), and the occasional distraction caused by contextual information (P1). These findings align with previous research suggesting that users often prefer less complex interfaces [99], and eHMI studies indicating that providing more information does not necessarily lead to better crossing decisions [46, 72]. While eHMIs may improve quantitative performance measures, these results highlight the need for balancing between performance and preference through intuitive design.

Our results also indicated some patterns in COT that reflect natural interaction behaviors among road users. For instance, pedestrians had faster COTs than drivers when AVs yielded to cyclists. Similarly, cyclists had faster COTs than drivers when AVs yielded to pedestrians. In both cases, the negotiation between road users not directly yielded to by the AV was left to their own interaction. Moreover, in the *When* and *Where eHMIs*, where only timing and location information were provided without specifying *'Whom'*, COTs for pedestrians and cyclists were lower than those for drivers. While potentially influenced by cultural norms, these findings suggest that when road users negotiated among themselves, they tended to prioritize the safety of more vulnerable participants. This natural tendency to protect vulnerable road users should be leveraged in eHMI design, with interfaces developed to reinforce and enhance these inherent protective behaviors.

Therefore, we argue that effective eHMI design requires balancing performance metrics with user preferences while respecting and reinforcing natural interaction patterns. Although eHMIs that convey richer contextual information can enhance road users' crossing performance, simpler eHMIs, such as *No Context*, which convey only basic AV-yielding intent, also hold significant value for user preference and comprehension.

In designing eHMIs for diverse road user groups in mixed-traffic scenarios, we recommend:

- Balancing information richness with simplicity to accommodate both performance requirements and user preferences.
- Respecting established interaction patterns, particularly those that naturally prioritize the safety of vulnerable road users.
- Designing eHMIs that encourage, facilitate, or reinforce natural protective behaviors while clearly communicating AV intentions.

By adhering to these principles, AV-to-multiple road user communication systems can enhance objective safety measures while providing an intuitive and user-preferred experience, ultimately supporting smoother integration of AVs into mixed-traffic environments.

### 5.5 Limitations and Future Work

First, our scenario featured a pedestrian crossing without a designated crosswalk to introduce real-world ambiguity and avoid predictable right-of-way dynamics. While this setup provides a basis for simulating space-sharing conflicts and varied road user interactions, it may not fully capture variations in traffic regulations and cultural contexts across regions. Recent studies emphasize that AV-road user interactions differ across regions, highlighting the importance of cross-country validation [1, 27]. The open-source nature of our study and tested scenarios can facilitate future research to extend the scalability of the proposed eHMI designs across different cultural contexts.

Second, it was conducted in a single controlled scenario, which may not fully capture the complexity of real-world traffic. Our study examined a controlled scenario involving one pedestrian, one cyclist, and one manual vehicle driver. While this setup enabled us to focus on the effects of contextual eHMIs, it may not adequately capture the dynamic and unpredictable nature of real-world traffic. Furthermore, although the current design of the *Whom eHMI* appropriately addressed only one road user group, its potential to communicate with multiple groups simultaneously was not investigated. Accordingly, scalability issues may arise when the AV needs to yield to more than one user within the same group, potentially complicating communication and reducing the eHMI's clarity. Future research should explore how the *Whom eHMI* functions when targeting multiple road users concurrently, to better understand its scalability and effectiveness in more complex, realistic scenarios. A qualitative discussion on these scalability concerns can be found in Section 5.3. Moreover, research should also include a wider range of scenarios, such as shared spaces [77], intersections [5], roundabouts [129], and unprotected turning events [7]. Expanding to include ambiguous scenarios, such as interactions with multiple AVs [82, 125], multiple users within each road user group [25, 40], or the presence of additional road users (e.g., other mobility users such as motorcyclists or wheelchair users), would help evaluate the scalability of eHMIs in dynamically rich traffic contexts. To support such extensions, future research may benefit from optimizing design parameters to match situational demands [26].

Third, while the structured designs of the contextual eHMIs—such as symbolic icons, temporal indicators, and spatial cues—likely contributed to improved clarity, their specific visual features may have also amplified the effects observed in participant responses. In particular, some designs may have encouraged participants to rely on familiar traffic-related cues or pre-existing mental models, strengthening their reactions to certain eHMIs more than might occur with less conventional designs. This may be especially relevant for the *Whom eHMI*, which consistently showed the strongest performance. Its use of road user icons resembling familiar traffic signs likely activated prior knowledge and expectations [43, 85]. This made the interface more intuitively interpretable and behaviorally effective. Accordingly, the strong performance of contextual eHMIs in this study may partially reflect the influence of their visual design characteristics in addition to the semantic content of the contextual information. Future studies should investigate whether alternative visual designs of contextual eHMIs lead to similar effects across different road user groups and settings.

Fourth, participant feedback highlighted challenges with the *When* and *Where eHMIs*. For the *When eHMI*, participants reported difficulty understanding the timer-based mechanism. For the *Where eHMI*, issues such as information overload, poor text visibility, and the intuitiveness of distance indicators—especially for cyclists aligned with the AV's direction—were noted. Future studies should investigate ways to improve comprehensibility, visibility [75], and compatibility with distinct road user characteristics [55].

Finally, our study relied solely on visual cues for eHMI communication, excluding other modalities such as auditory [56] or haptic signals [104]. This could pose significant challenges in developing more inclusive and accessible AV communication strategies, especially for VRUs with visual impairments or color vision deficiencies. Multimodal eHMIs combining visual and auditory cues have been shown to outperform unimodal eHMIs in enhancing communication effectiveness [39]. Therefore, future research should explore incorporating auditory and haptic feedback, particularly in complex conflict scenarios.

## 6 CONCLUSION

Our study proposed a contextual eHMI to communicate the yielding intention of an AV to assist multiple road users with different characteristics. Using a VR-based multi-agent simulation (N=42), we assessed the crossing performance (onset time, duration, error rate) and subjective feelings (perceived safety, trust, clarity) of *pedestrians, cyclists, and manual vehicle drivers* across different eHMI types (*No eHMI, No Context eHMI, Whom eHMI, When eHMI, Where eHMI*). Structured interviews were also conducted to gather insights on user preferences, communication efficiency, and comprehension. Our results demonstrate the potential of contextual eHMIs to facilitate road users' decision-making in multiple road-user conflict scenarios. Notably, the *Whom eHMI*, which conveys the yielding intention of the AV specifically to the participant, had the most significant positive impact on the crossing performance, perceived safety, and clarity. Our study contributes empirical evidence on eHMI effectiveness in multi-road user scenarios, providing scalable design recommendations for different road user types and interaction contexts. A wider range of scenarios and more diverse distributions of road users should be explored to further validate the universal applicability of eHMIs in real-world traffic environments.

## Acknowledgments

This work was partly supported by the GIST-MIT Research Collaboration grant funded by the GIST in 2025, the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00343397), and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2025-16902996).

## References

[1] Leidy Marcela Barón Acela, Emanuel Sousa, Susana Faria, Raul Almeida, and Elisabete Fraga Freitas. 2024. Pedestrian crossing behavior in social groups: Exploring cultural contexts through a comparative study. *Case Studies on Transport Policy* (2024), 101333. doi:10.1016/j.cstp.2024.101333

[2] Claudia Ackermann, Matthias Beggiato, Luka-Franziska Bluhm, Alexandra Löw, and Josef F Krems. 2019. Deceleration parameters and their applicability as informal communication signal between pedestrians and automated vehicles. *Transportation research part F: traffic psychology and behaviour* 62 (2019), 757–768. doi:10.1016/j.trf.2019.03.006

[3] Sander Ackermans, Debargha Dey, Peter Ruijten, Raymond H Cuijpers, and Bastian Pfleging. 2020. The effects of explicit intention communication, conspicuous sensors, and pedestrian attitude in interactions with automated vehicles. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–14. doi:10.1145/3313831.3376197

[4] Ammar Al-Taie, Frank Pollick, and Stephen Brewster. 2022. Tour de interaction: Understanding cyclist-driver interaction with self-reported cyclist behaviour. In *Adjunct Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 127–131. doi:10.1145/3544999.3552531

[5] Ammar Al-Taie, Graham Wilson, Euan Freeman, Frank Pollick, and Stephen Anthony Brewster. 2024. Light it Up: Evaluating Versatile Autonomous Vehicle-Cyclist External Human-Machine Interfaces. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20. doi:10.1145/3613904.3642019

[6] Ammar Al-Taie, Graham Wilson, Frank Pollick, and Stephen Anthony Brewster. 2023. Pimp my ride: Designing versatile ehmis for cyclists. In *Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 213–223. doi:10.1145/3580585.3607161

[7] Hüseyin Avsar, Fabian Utesch, Marc Wilbrink, Michael Oehl, and Caroline Schießl. 2021. Efficient communication of automated vehicles and manually driven vehicles through an external Human-Machine Interface (eHMI): Evaluation at T-junctions. In *HCI International 2021-Posters: 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part III 23*. Springer, 224–232. doi:10.1007/978-3-030-78645-8_28

[8] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics* 1 (2009), 71–81. doi:10.1007/s12369-008-0001-3

[9] Pavlo Bazilinskyy, Dimitra Dodou, and Joost De Winter. 2019. Survey on eHMI concepts: The effect of text, color, and perspective. *Transportation research part F: traffic psychology and behaviour* 67 (2019), 175–194. doi:10.1016/j.trf.2019.10.013

[10] Pavlo Bazilinskyy, Patrick Ebel, Francesco Walker, Debargha Dey, and Tram Thi Minh Tran. 2024. It Is Not Always Just One Road User: Workshop on Multi-Agent Automotive Research. In *Adjunct Proceedings of the 16th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 268–272. doi:10.1145/3641308.3677400

[11] Pavlo Bazilinskyy, Lars Kooijman, Dimitra Dodou, and Joost CF de Winter. 2020. Coupled simulator for research on the interaction between pedestrians and (automated) vehicles. In *Proceedings of the Driving Simulation Conference Europe*.

[12] Matthias Beggiato, Claudia Witzlack, and Josef F Krems. 2017. Gap acceptance and time-to-arrival estimates as basis for informal communication between pedestrians and vehicles. In *Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications*. 50–57. doi:10.1145/3122986.3122995

[13] Mathias Benedek and Christian Kaernbach. 2010. A continuous measure of phasic electrodermal activity. *Journal of neuroscience methods* 190, 1 (2010), 80–91. doi:10.1016/j.jneumeth.2010.04.028

[14] Siri Hegna Berge, Joost de Winter, Diane Cleij, and Marjan Hagenzieker. 2024. Triangulating the future: Developing scenarios of cyclist-automated vehicle interactions from literature, expert perspectives, and survey data. *Transportation research interdisciplinary perspectives* 23 (2024), 100986. doi:10.1016/j.trip.2023.100986

[15] Siri Hegna Berge, Joost de Winter, and Marjan Hagenzieker. 2023. Support systems for cyclists in automated traffic: A review and future outlook. *Applied Ergonomics* 111 (2023), 104043. doi:10.1016/j.apergo.2023.104043

[16] Siri Hegna Berge, Joost De Winter, and Marjan Hagenzieker. 2023. User Interfaces for Cyclists in Future Automated Traffic. In *Companion Proceedings of the 28th International Conference on Intelligent User Interfaces*. 91–94. doi:10.1145/3581754.3584140

[17] Janina Bindschädel, Pauline Weimann, and Andrea Kiesel. 2023. Using eHMI, acoustic signal, and pitch motion to communicate the intention of automated vehicles to pedestrians: A Wizard of Oz study. *Transportation research part F: traffic psychology and behaviour* 97 (2023), 59–72. doi:10.1016/j.trf.2023.06.017

[18] Michael Branion-Calles, Thomas Götschi, Trisalyn Nelson, Esther Anaya-Boig, Ione Avila-Palencia, Alberto Castro, Tom Cole-Hunter, Audrey de Nazelle, Evi Dons, Mailin Gaupp-Berghausen, et al. 2020. Cyclist crash rates and risk factors in a prospective cohort in seven European cities. *Accident Analysis & Prevention* 141 (2020), 105540. doi:10.1016/j.aap.2020.105540

[19] Yi Cao, Xiangling Zhuang, and Guojie Ma. 2019. Shorten pedestrians' perceived waiting time: The effect of tempo and pitch in audible pedestrian signals at red phase. *Accident Analysis & Prevention* 123 (2019), 336–340. doi:10.1016/j.aap.2018.12.011

[20] Magnus Carlsson and Per Nilsson. 2016. *The Smiling Car-concept for autonomous cars.* Technical Report. https://semcon. com/smilingcar.

[21] Chia-Ming Chang, Koki Toda, Daisuke Sakamoto, and Takeo Igarashi. 2017. Eyes on a Car: an Interface Design for Communication between an Autonomous Car and a Pedestrian. In *Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications*. 65–73. doi:10.1145/3122986.3122989

[22] Xiang Chang, Zihe Chen, Xiaoyan Dong, Yuxin Cai, Tingmin Yan, Haolin Cai, Zherui Zhou, Guyue Zhou, and Jiangtao Gong. 2024. " It Must Be Gesturing Towards Me": Gesture-Based Interaction between Autonomous Vehicles and Pedestrians. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–25. doi:10.1145/3613904.3642029

[23] Michael Clamann, Miles Aubert, and Mary L Cummings. 2017. *Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles.* Technical Report.

[24] Mark Colley, Elvedin Bajrovic, and Enrico Rukzio. 2022. Effects of pedestrian behavior, time pressure, and repeated exposure on crossing decisions in front of automated vehicles equipped with external communication. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–11. doi:10.1145/3491102.3517571

[25] Mark Colley, Julian Britten, and Enrico Rukzio. 2023. Scalability in External Communication of Automated Vehicles: Evaluation and Recommendations. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 2 (2023), 1–26. doi:10.1145/3596248

[26] Mark Colley, Pascal Jansen, Mugdha Keskar, and Enrico Rukzio. 2025. Improving External Communication of Automated Vehicles Using Bayesian Optimization *(CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 80, 16 pages. doi:10.1145/3706598.3714187

[27] Mark Colley, Daniel Kornmüller, Debargha Dey, Wendy Ju, and Enrico Rukzio. 2024. Longitudinal Effects of External Communication of Automated Vehicles in the USA and Germany: A Comparative Study in Virtual Reality and Via a Browser. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 4 (2024), 1–33. doi:10.1145/3699778

[28] Mark Colley, Omid Rajabi, and Enrico Rukzio. 2024. Investigating the Effects of External Communication and Platoon Behavior on Manual Drivers at Highway Access. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–15. doi:10.1145/3613904.3642365

[29] Mark Colley, Marcel Walch, Jan Gugenheimer, and Enrico Rukzio. 2019. Including people with impairments from the start: External communication of autonomous vehicles. In *Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications: adjunct proceedings*. 307–314. doi:10.1145/3349263.3351521

[30] Mark Colley, Marcel Walch, and Enrico Rukzio. 2020. Unveiling the lack of scalability in research on external communication of autonomous vehicles. In *Extended abstracts of the 2020 chi conference on human factors in computing systems*. 1–9. doi:10.1145/3334480.3382865

[31] Hatice Şahin, Kevin Daudrich, Heiko Müller, and Susanne CJ Boll. 2021. Signaling Yielding Intent with eHMIs: the Timing Determines an Efficient Crossing. Association for Computing Machinery, New York, NY, USA. doi:10.1145/3473682.3480253

[32] Koen De Clercq, Andre Dietrich, Juan Pablo Núñez Velasco, Joost De Winter, and Riender Happee. 2019. External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions. *Human factors* 61, 8 (2019), 1353–1370. doi:10.1177/0018720819836343

[33] Joost de Winter and Dimitra Dodou. 2022. External human–machine interfaces: Gimmick or necessity? *Transportation research interdisciplinary perspectives* 15 (2022), 100643. doi:10.1016/j.trip.2022.100643

[34] Patricia R DeLucia. 2008. Critical roles for distance, task, and motion in space perception: Initial conceptual framework and practical implications. *Human Factors* 50, 5 (2008), 811–820. doi:10.1518/001872008X312297

[35] Debargha Dey, Azra Habibovic, Melanie Berger, Devanshi Bansal, Raymond H Cuijpers, and Marieke Martens. 2022. 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*. 307–318. doi:10.1145/3543174.3546086

[36] Debargha Dey, Kai Holländer, Melanie Berger, Berry Eggen, Marieke Martens, Bastian Pfleging, and Jacques Terken. 2020. Distance-dependent eHMIs for the interaction between automated vehicles and pedestrians. In *12th international conference on automotive user interfaces and interactive vehicular applications*. 192–204. doi:10.1145/3409120.3410642

[37] Debargha Dey, Marieke Martens, Berry Eggen, and Jacques Terken. 2017. The impact of vehicle appearance and vehicle behavior on pedestrian interaction with autonomous vehicles. In *Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications adjunct*. 158–162. doi:10.1145/3131726.3131750

[38] Debargha Dey, Marieke Martens, Chao Wang, Felix Ros, and Jacques Terken. 2018. Interface concepts for intent communication from autonomous vehicles to vulnerable road users. In *Adjunct proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications*. 82–86. doi:10.1145/3239092.3265946

[39] Debargha Dey, Toros Ufuk Senan, Bart Hengeveld, Mark Colley, Azra Habibovic, and Wendy Ju. 2024. Multi-Modal eHMIs: The Relative Impact of Light and Sound in AV-Pedestrian Interaction. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–16. doi:10.1145/3613904.3642031

[40] Debargha Dey, Arjen Van Vastenhoven, Raymond H Cuijpers, Marieke Martens, and Bastian Pfleging. 2021. Towards scalable eHMIs: Designing for AV-VRU communication beyond one pedestrian. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 274–286. doi:10.1145/3409118.3475129

[41] André Dietrich, J-H Willrodt, Karolin Wagner, and Klaus Bengler. 2019. Projection-based external human machine interfaces-enabling interaction between automated vehicles and pedestrians. In *Proceedings of the DSC 2018 Europe VR*.

[42] Olive Jean Dunn. 1961. Multiple comparisons among means. *Journal of the American statistical association* 56, 293 (1961), 52–64. doi:10.1080/01621459.1961.10482090

[43] Daniel Eisele and Tibor Petzoldt. 2022. Effects of traffic context on eHMI icon comprehension. *Transportation research part F: traffic psychology and behaviour* 85 (2022), 1–12. doi:10.1016/j.trf.2021.12.014

[44] Yke Bauke Eisma, Anna Reiff, Lars Kooijman, Dimitra Dodou, and Joost CF de Winter. 2021. External human-machine interfaces: Effects of message perspective. *Transportation research part F: traffic psychology and behaviour* 78 (2021), 30–41. doi:10.1016/j.trf.2021.01.013

[45] Mica R Endsley. 1995. Toward a theory of situation awareness in dynamic systems. *Human factors* 37, 1 (1995), 32–64. doi:10.1518/001872095779049543

[46] Stefanie M Faas, Lesley-Ann Mathis, and Martin Baumann. 2020. External HMI for self-driving vehicles: Which information shall be displayed? *Transportation research part F: traffic psychology and behaviour* 68 (2020), 171–186. doi:10.1016/j.trf.2019.12.009

[47] Daniel J Fagnant and Kara Kockelman. 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice* 77 (2015), 167–181. doi:10.1016/j.tra.2015.04.003

- [48] Yan Feng, Haneen Farah, and Bart van Arem. 2023. Effect of eHMI on pedestrian road crossing behavior in shared space with Automated Vehicles-A Virtual Reality study. In *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2038–2043. doi:10.1109/ITSC57777.2023.10421848
- [49] Bernd Figner, Ryan O Murphy, et al. 2011. Using skin conductance in judgment and decision making research. *A handbook of process tracing methods for decision research* (2011), 163–184.
- [50] Samuel W Greenhouse and Seymour Geisser. 1959. On methods in the analysis of profile data. *Psychometrika* 24, 2 (1959), 95–112. doi:10.1007/BF02289823
- [51] Marjan P Hagenzieker, Sander Van Der Kint, Luuk Vissers, Ingrid NL G Van Schagen, Jonathan De Bruin, Paul Van Gent, and Jacques JF Commandeur. 2020. Interactions between cyclists and automated vehicles: Results of a photo experiment. *Journal of Transportation Safety & Security* 12, 1 (2020), 94–115. doi:10.1080/19439962.2019.1591556
- [52] Julie Hatfield and Prasannah Prabhakharan. 2016. An investigation of behaviour and attitudes relevant to the user safety of pedestrian/cyclist shared paths. *Transportation research part F: traffic psychology and behaviour* 40 (2016), 35–47. doi:10.1016/j.trf.2016.04.005
- [53] Kai Holländer, Mark Colley, Enrico Rukzio, and Andreas Butz. 2021. A taxonomy of vulnerable road users for hci based on a systematic literature review. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–13. doi:10.1145/3411764.3445480
- [54] Ming Hou, Karthik Mahadevan, Sowmya Somanath, Ehud Sharlin, and Lora Oehlberg. 2020. Autonomous vehicle-cyclist interaction: Peril and promise. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12. doi:10.1145/3313831.3376884
- [55] Maximilian Hübner, Alexander Feierle, Michael Rettenmaier, and Klaus Bengler. 2022. External communication of automated vehicles in mixed traffic: Addressing the right human interaction partner in multi-agent simulation. *Transportation research part F: traffic psychology and behaviour* 87 (2022), 365–378. doi:10.1016/j.trf.2022.04.017
- [56] Christopher R Hudson, Shuchisnigdha Deb, Daniel W Carruth, John McGinley, and Darren Frey. 2019. Pedestrian perception of autonomous vehicles with external interacting features. In *Advances in Human Factors and Systems Interaction: Proceedings of the AHFE 2018 International Conference on Human Factors and Systems Interaction*. Springer, 33–39. doi:10.1007/978-3-319-94334-3_5
- [57] Jiun-Yin Jian, Ann M Bisantz, and Colin G Drury. 2000. Foundations for an empirically determined scale of trust in automated systems. *International journal of cognitive ergonomics* 4, 1 (2000), 53–71. doi:10.1207/S15327566IJCE0401_04
- [58] David Kaber, Sangeun Jin, Maryam Zahabi, and Carl Pankok Jr. 2016. The effect of driver cognitive abilities and distractions on situation awareness and performance under hazard conditions. *Transportation research part F: traffic psychology and behaviour* 42 (2016), 177–194. doi:10.1016/j.trf.2016.07.014
- [59] Tilak Kamkar and Andreas Riener. 2024. Be Here Now: Development and Evaluation of Scalable eHMIs to Enhance the Attention, Perception, and Safety of Inattentive Pedestrians through Targeted Communication in Complex Traffic Scenarios. In *Adjunct Proceedings of the 16th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 220–225. doi:10.1145/3641308.3685051
- [60] Seongjun Kang, Gwangbin Kim, and SeungJun Kim. 2023. Giant finger: A novel visuo-somatosensory approach to simulating lower body movements in virtual reality. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 233–242. doi:10.1109/ISMAR59233.2023.00038
- [61] Yumin Kang, SeongA Choi, Eunsol An, Seokhyun Hwang, and Seungjun Kim. 2023. Designing virtual agent human–machine interfaces depending on the communication and anthropomorphism levels in augmented reality. In *Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 191–201. doi:10.1145/3580585.3606460
- [62] Ioannis Kaparias, Michael GH Bell, Ashkan Miri, Carol Chan, and Bill Mount. 2012. Analysing the perceptions of pedestrians and drivers to shared space. *Transportation research part F: traffic psychology and behaviour* 15, 3 (2012), 297–310. doi:10.1016/j.trf.2012.02.001
- [63] Gwangbin Kim, Seokhyun Hwang, Minwoo Seong, Dohyeon Yeo, Daniela Rus, and SeungJun Kim. 2024. TimelyTale: a multimodal dataset approach to assessing passengers' explanation demands in highly automated vehicles. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 3 (2024), 1–60. doi:10.1145/3678544
- [64] Gwangbin Kim, Jieun Lee, Dohyeon Yeo, Eunsol An, and SeungJun Kim. 2023. Physiological Indices to Predict Driver Situation Awareness in VR. In *Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing*. 40–45. doi:10.1145/3594739.3610687
- [65] Gwangbin Kim, Dohyeon Yeo, Taewoo Jo, Daniela Rus, and SeungJun Kim. 2023. What and When to Explain? On-road Evaluation of Explanations in Highly Automated Vehicles. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 3 (2023), 1–26. doi:10.1145/3610886
- [66] Gwangbin Kim, Dohyeon Yeo, Taewoo Jo, Daniela Rus, and SeungJun Kim. 2025. Decoding Your Ride: What Information Builds Human Trust on Real Roads? *GetMobile: Mobile Comp. and Comm.* 29, 2 (June 2025), 5–9.
- [67] Seungju Kim, Jungseok Oh, Minwoo Seong, Eunki Jeon, Yeon-Kug Moon, and Seungjun Kim. 2023. Assessing the Impact of AR HUDs and Risk Level on User Experience in Self-Driving Cars: Results from a Realistic Driving Simulation. *Applied Sciences* 13, 8 (2023), 4952. doi:10.3390/app13084952
- [68] Young Woo Kim, Jae Hyun Han, Yong Gu Ji, and Seul Chan Lee. 2020. Exploring the effectiveness of external human-machine interfaces on pedestrians and drivers. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 65–68.

doi:10.1145/3409251.3411725
[69] Theresa JB Kline. 2005. *Psychological testing: A practical approach to design and evaluation.* Sage publications.
[70] Masaki Kuge, Hailong Liu, Toshihiro Hiraoka, and Takahiro Wada. 2024. An eHMI Presenting Request-to-Intervene Status of Level 3 Automated Vehicles to Surrounding Vehicles. In *2024 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2970–2975. doi:10.1109/IV55156.2024.10588772
[71] Mirjam Lanzer, Ina Koniakowsky, Mark Colley, and Martin Baumann. 2023. Interaction effects of pedestrian behavior, smartphone distraction and external communication of automated vehicles on crossing and gaze behavior. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18. doi:10.1145/3544548.3581303
[72] Merle Lau, Meike Jipp, and Michael Oehl. 2022. 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. *Accident Analysis & Prevention* 171 (2022), 106641. doi:10.1016/j.aap.2022.106641
[73] Merle Lau, Jieun Lee, Satoshi Kitazaki, Tatsuru Daimon, and Michael Oehl. 2024. Investigating Implicit and Explicit Communication of Highly Automated Vehicles in Japan: How do Light-band eHMIs affect Pedestrians' Willingness to Cross, Trust and Perceived Safety? *Accident Analysis & Prevention* 207 (2024), 107719. doi:10.1016/j.aap.2024.107719
[74] Seul Chan Lee, Sol Hee Yoon, and Yong Gu Ji. 2021. Effects of non-driving-related task attributes on takeover quality in automated vehicles. *International Journal of Human–Computer Interaction* 37, 3 (2021), 211–219. doi:10.1080/10447318.2020.1815361
[75] Yee Mun Lee, Ruth Madigan, Chinebuli Uzondu, Jorge Garcia, Richard Romano, Gustav Markkula, and Natasha Merat. 2022. Learning to interpret novel eHMI: The effect of vehicle kinematics and eHMI familiarity on pedestrian'crossing behavior. *Journal of safety research* 80 (2022), 270–280. doi:10.1016/j.jsr.2021.12.010
[76] Penghui Li, Yibing Li, Yao Yao, Changxu Wu, Bingbing Nie, and Shengbo Eben Li. 2021. Sensitivity of electrodermal activity features for driver arousal measurement in cognitive load: the application in automated driving systems. *IEEE transactions on intelligent transportation systems* 23, 9 (2021), 14954–14967. doi:10.1109/TITS.2021.3135266
[77] Yang Li, Hao Cheng, Zhe Zeng, Hailong Liu, and Monika Sester. 2021. Autonomous vehicles drive into shared spaces: ehmi design concept focusing on vulnerable road users. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. IEEE, 1729–1736. doi:10.1109/ITSC48978.2021.9564515
[78] Yang Li, Hailong Liu, and Barbara Deml. 2021. Manually driven vehicle encounters with autonomous vehicle in bottleneck roads: HMI design for communication issues. In *Proceedings of the 9th International Conference on Human-Agent Interaction*. 382–385. doi:10.1145/3472307.3484678
[79] Shiva Nischal Lingam, JCF de Winter, Yongqi Dong, Anastasia Tsapi, Bart Van Arem, and Haneen Farah. 2024. eHMI on the Vehicle or on the Infrastructure? (2024). https://resolver.tudelft.nl/uuid:e5d53e26-fd3c-4411-b10d-7f856b91fd8a
[80] Hailong Liu, Takatsugu Hirayama, Luis Yoichi Morales Saiki, and Hiroshi Murase. 2020. What timing for an automated vehicle to make pedestrians understand its driving intentions for improving their perception of safety?. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 1–6. doi:10.1109/ITSC45102.2020.9294696
[81] Andreas Löcken, Carmen Golling, and Andreas Riener. 2019. 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*. 262–274. doi:10.1145/3342197.3344544
[82] Andreas Löcken, Tamara von Sawitzky, Janine Bauer, and Andreas Riener. 2023. Exploring the Potential of eHMIs as Traffic Light Alternatives. In *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 99–104. doi:10.1145/3581961.3609885
[83] Glenn Lyons. 2020. Walking as a service–Does it have legs? *Transportation Research Part A: Policy and Practice* 137 (2020), 271–284. doi:10.1016/j.tra.2020.05.015
[84] Wei Lyu, Wen-gang Zhang, Xueshuang Wang, Yi Ding, and Xinyue Yang. 2024. Pedestrians' responses to scalable automated vehicles with different external human-machine interfaces: Evidence from a video-based eye-tracking experiment. *Transportation research part F: traffic psychology and behaviour* 103 (2024), 112–127. doi:10.1016/j.trf.2024.04.005
[85] Andreas Löcken, Philipp Wintersberger, Anna-Katharina Frison, and Andreas Riener. 2019. Investigating User Requirements for Communication Between Automated Vehicles and Vulnerable Road Users. In *2019 IEEE Intelligent Vehicles Symposium (IV)*. 879–884. doi:10.1109/IVS.2019.8814027
[86] Karthik Mahadevan, Elaheh Sanoubari, Sowmya Somanath, James E Young, and Ehud Sharlin. 2019. AV-Pedestrian interaction design using a pedestrian mixed traffic simulator. In *Proceedings of the 2019 on designing interactive systems conference*. 475–486. doi:10.1145/3322276.3322328
[87] Karthik Mahadevan, Sowmya Somanath, and Ehud Sharlin. 2018. Communicating awareness and intent in autonomous vehicle-pedestrian interaction. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12. doi:10.1145/3173574.3174003
[88] Paul Mallery and Darren George. 2000. *SPSS for windows step by step.* Allyn & Bacon, Inc.

[89] Gustav Markkula, Ruth Madigan, Dimitris Nathanael, Evangelia Portouli, Yee Mun Lee, André Dietrich, Jac Billington, Anna Schieben, and Natasha Merat. 2020. Defining interactions: A conceptual framework for understanding interactive behaviour in human and automated road traffic. *Theoretical Issues in Ergonomics Science* 21, 6 (2020), 728–752. doi:10.1080/1463922X.2020.1736686

[90] Lauren Meredith, Jordanka Kovaceva, and András Bálint. 2020. Mapping fractures from traffic accidents in Sweden: How do cyclists compare to other road users? *Traffic injury prevention* 21, 3 (2020), 209–214. doi:10.1080/15389588.2020.1724979

[91] Linda Miller, Jasmin Leitner, Johannes Kraus, and Martin Baumann. 2022. Implicit intention communication as a design opportunity for automated vehicles: Understanding drivers' interpretation of vehicle trajectory at narrow passages. *Accident Analysis & Prevention* 173 (2022), 106691. doi:10.1016/j.aap.2022.106691

[92] Meghan Fehlig Mitman, David R Ragland, and Charles V Zegeer. 2008. Marked-crosswalk dilemma: Uncovering some missing links in a 35-year debate. *Transportation research record* 2073, 1 (2008), 86–93. doi:10.3141/2073-10

[93] Simon Moody and Steve Melia. 2014. Shared space–research, policy and problems. In *Proceedings of the Institution of Civil Engineers-Transport*, Vol. 167. Thomas Telford Ltd, 384–392. doi:10.1680/tran.12.00047

[94] Lia Morra, Fabrizio Lamberti, F Gabriele Pratticó, Salvatore La Rosa, and Paolo Montuschi. 2019. Building trust in autonomous vehicles: Role of virtual reality driving simulators in HMI design. *IEEE Transactions on Vehicular Technology* 68, 10 (2019), 9438–9450. doi:10.1109/TVT.2019.2933601

[95] Iñigo Mujika, Bent R Rønnestad, and David T Martin. 2016. Effects of increased muscle strength and muscle mass on endurance-cycling performance. *International journal of sports physiology and performance* 11, 3 (2016), 283–289. doi:10.1123/ijspp.2015-0405

[96] Annie W.Y. Ng and Alan H.S. Chan. 2008. The effects of driver factors and sign design features on the comprehensibility of traffic signs. *Journal of Safety Research* 39, 3 (2008), 321–328. doi:10.1016/j.jsr.2008.02.031

[97] Trung Thanh Nguyen, Kai Holländer, Marius Hoggenmueller, Callum Parker, and Martin Tomitsch. 2019. Designing for projection-based communication between autonomous vehicles and pedestrians. In *Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications*. 284–294. doi:10.1145/3342197.3344543

[98] Hampus Norén. 2022. Developing a Virtual Reality Bicycle Simulator in Unity for Traffic Safety Research Integration.

[99] Michael I Norton, Jeana H Frost, and Dan Ariely. 2007. Less is more: the lure of ambiguity, or why familiarity breeds contempt. *Journal of personality and social psychology* 92, 1 (2007), 97. doi:10.1037/0022-3514.92.1.97

[100] Eleonora Papadimitriou, Sylvain Lassarre, and George Yannis. 2016. Introducing human factors in pedestrian crossing behaviour models. *Transportation research part F: traffic psychology and behaviour* 36 (2016), 69–82. doi:10.1016/j.trf.2015.11.003

[101] Vassilis Papakostopoulos, Dimitris Nathanael, Evangelia Portouli, and Angelos Amditis. 2021. Effect of external HMI for automated vehicles (AVs) on drivers' ability to infer the AV motion intention: A field experiment. *Transportation research part F: traffic psychology and behaviour* 82 (2021), 32–42. doi:10.1016/j.trf.2021.07.009

[102] John Parkin, Fiona Crawford, Jonathan Flower, Chris Alford, Phillip Morgan, and Graham Parkhurst. 2023. Cyclist and pedestrian trust in automated vehicles: An on-road and simulator trial. *International journal of sustainable transportation* 17, 7 (2023), 762–774. doi:10.1080/15568318.2022.2093147

[103] Praveena Penmetsa, Emmanuel Kofi Adanu, Dustin Wood, Teng Wang, and Steven L Jones. 2019. Perceptions and expectations of autonomous vehicles–A snapshot of vulnerable road user opinion. *Technological Forecasting and Social Change* 143 (2019), 9–13. doi:10.1016/j.techfore.2019.02.010

[104] Pooya Rahimian, Elizabeth E O'Neal, Junghum Paul Yon, Luke Franzen, Yuanyuan Jiang, Jodie M Plumert, and Joseph K Kearney. 2016. Using a virtual environment to study the impact of sending traffic alerts to texting pedestrians. In *2016 IEEE Virtual Reality (VR)*. IEEE, 141–149. doi:10.1109/VR.2016.7504697

[105] Amir Rasouli, Iuliia Kotseruba, and John K Tsotsos. 2017. Agreeing to cross: How drivers and pedestrians communicate. In *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 264–269. doi:10.1109/IVS.2017.7995730

[106] Amir Rasouli and John K Tsotsos. 2019. Autonomous vehicles that interact with pedestrians: A survey of theory and practice. *IEEE transactions on intelligent transportation systems* 21, 3 (2019), 900–918. doi:10.1109/TITS.2019.2901817

[107] Michael Rettenmaier, Deike Albers, and Klaus Bengler. 2020. After you?!–Use of external human-machine interfaces in road bottleneck scenarios. *Transportation research part F: traffic psychology and behaviour* 70 (2020), 175–190. doi:10.1016/j.trf.2020.03.004

[108] Michael Rettenmaier and Klaus Bengler. 2021. The matter of how and when: Comparing explicit and implicit communication strategies of automated vehicles in bottleneck scenarios. *IEEE Open Journal of Intelligent Transportation Systems* 2 (2021), 282–293. doi:10.1109/OJITS.2021.3107678

[109] Conor CO Reynolds, M Anne Harris, Kay Teschke, Peter A Cripton, and Meghan Winters. 2009. The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. *Environmental health* 8 (2009), 1–19. doi:10.1186/1476-069X-8-47

[110] Benjamin Ryder, Bernhard Gahr, Philipp Egolf, Andre Dahlinger, and Felix Wortmann. 2017. Preventing traffic accidents with in-vehicle decision support systems-The impact of accident hotspot warnings on driver behaviour. *Decision support systems* 99 (2017), 64–74. doi:10.1016/j.dss.2017.05.004

[111] Paul M Salmon, Kristie L Young, and Miranda Cornelissen. 2013. Compatible cognition amongst road users: The compatibility of driver, motorcyclist, and cyclist situation awareness. *Safety science* 56 (2013), 6–17. doi:10.1016/j.ssci.2012.02.008

[112] Anna Schieben, Marc Wilbrink, Carmen Kettwich, Ruth Madigan, Tyron Louw, and Natasha Merat. 2019. Designing the interaction of automated vehicles with other traffic participants: design considerations based on human needs and expectations. *Cognition, Technology & Work* 21 (2019), 69–85. doi:10.1007/s10111-018-0521-z

[113] Abbas Sheykhfard, Farshidreza Haghighi, Eleonora Papadimitriou, and Pieter Van Gelder. 2021. Analysis of the occurrence and severity of vehicle-pedestrian conflicts in marked and unmarked crosswalks through naturalistic driving study. *Transportation Research Part F: Traffic Psychology and Behaviour* 76 (2021), 178–192. doi:10.1016/j.trf.2020.11.008

[114] Yu Shi, Natalie Ruiz, Ronnie Taib, Eric Choi, and Fang Chen. 2007. Galvanic skin response (GSR) as an index of cognitive load. In *CHI'07 extended abstracts on Human factors in computing systems*. 2651–2656. doi:10.1145/1240866.1241057

[115] Timo Singer, Jonas Kobbert, Babak Zandi, and Tran Quoc Khanh. 2020. 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* 23, 4 (2020), 2904–2918. doi:10.1109/TITS.2020.3032777

[116] Virginia P Sisiopiku and Darcin Akin. 2003. Pedestrian behaviors at and perceptions towards various pedestrian facilities: an examination based on observation and survey data. *Transportation research part f: traffic psychology and behaviour* 6, 4 (2003), 249–274. doi:10.1016/j.trf.2003.06.001

[117] skywind3000. 2013. KCP: A Fast and Reliable ARQ Protocol. https://github.com/skywind3000/kcp. Accessed: 2025-01-24.

[118] Mel Slater, Christoph Guger, Guenter Edlinger, Robert Leeb, Gert Pfurtscheller, Angus Antley, Maia Garau, Andrea Brogni, and Doron Friedman. 2006. Analysis of physiological responses to a social situation in an immersive virtual environment. *Presence* 15, 5 (2006), 553–569. doi:10.1162/pres.15.5.553

[119] Yuanming Song, Qianni Jiang, Wenxiang Chen, Xiangling Zhuang, and Guojie Ma. 2023. Pedestrians' road-crossing behavior towards eHMI-equipped autonomous vehicles driving in segregated and mixed traffic conditions. *Accident Analysis & Prevention* 188 (2023), 107115. doi:10.1016/j.aap.2023.107115

[120] Neville A Stanton, Rebecca Stewart, Don Harris, Robert J Houghton, Chris Baber, Richard McMaster, Paul Salmon, Geoff Hoyle, Guy Walker, Mark S Young, et al. 2006. Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. *Ergonomics* 49, 12-13 (2006), 1288–1311. doi:10.1080/00140130600612762

[121] Lakshmi Devi Subramanian, Elizabeth E O'Neal, Nam-Yoon Kim, Megan Noonan, Jodie M Plumert, and Joseph K Kearney. 2024. Deciding when to cross in front of an autonomous vehicle: how child and adult pedestrians respond to eHMI timing and vehicle kinematics. *Accident Analysis & Prevention* 202 (2024), 107567. doi:10.1016/j.aap.2024.107567

[122] Matus Sucha, Daniel Dostal, and Ralf Risser. 2017. Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis & Prevention* 102 (2017), 41–50. doi:10.1016/j.aap.2017.02.018

[123] Xu Sun, Shi Cao, and Pinyan Tang. 2021. Shaping driver-vehicle interaction in autonomous vehicles: How the new in-vehicle systems match the human needs. *Applied ergonomics* 90 (2021), 103238. doi:10.1016/j.apergo.2020.103238

[124] Wilbert Tabone, Joost De Winter, Claudia Ackermann, Jonas Bärgman, Martin Baumann, Shuchisnigdha Deb, Colleen Emmenegger, Azra Habibovic, Marjan Hagenzieker, Peter A Hancock, et al. 2021. Vulnerable road users and the coming wave of automated vehicles: Expert perspectives. *Transportation research interdisciplinary perspectives* 9 (2021), 100293. doi:10.1016/j.trip.2020.100293

[125] Tram Thi Minh Tran, Callum Parker, Marius Hoggenmüller, Yiyuan Wang, and Martin Tomitsch. 2024. Exploring the Impact of Interconnected External Interfaces in Autonomous Vehicles on Pedestrian Safety and Experience. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17. doi:10.1145/3613904.3642118

[126] Tram Thi Minh Tran, Callum Parker, and Martin Tomitsch. 2021. A review of virtual reality studies on autonomous vehicle–pedestrian interaction. *IEEE Transactions on Human-Machine Systems* 51, 6 (2021), 641–652. doi:10.1109/THMS.2021.3107517

[127] Tram Thi Minh Tran, Callum Parker, and Martin Tomitsch. 2023. Scoping out the scalability issues of autonomous vehicle-pedestrian interaction. In *Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 167–177. doi:10.1145/3580585.3607167

[128] Tram Thi Minh Tran, Callum Parker, Yiyuan Wang, and Martin Tomitsch. 2022. Designing wearable augmented reality concepts to support scalability in autonomous vehicle-pedestrian interaction. *Frontiers in Computer Science* 4 (2022), 866516. doi:10.3389/fcomp.2022.866516

[129] Rutger Verstegen, Debargha Dey, and Bastian Pfleging. 2021. CommDisk: A holistic 360 eHMI concept to facilitate scalable, unambiguous interactions between automated vehicles and other road users. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 132–136. doi:10.1145/3473682.3480280

[130] Fred Wegman, Fan Zhang, and Atze Dijkstra. 2012. How to make more cycling good for road safety? *Accident Analysis & Prevention* 44, 1 (2012), 19–29. doi:10.1016/j.aap.2010.11.010

[131] Annette Werner. 2018. New colours for autonomous driving: An evaluation of chromaticities for the external lighting equipment of autonomous vehicles. *Colour Turn* 1 (2018). doi:10.25538/tct.v0i1.692

[132] Marc Wilbrink, Manja Nuttelmann, and Michael Oehl. 2021. Scaling up automated vehicles' eHMI communication designs to interactions with multiple pedestrians–putting eHMIs to the test. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 119–122. doi:10.1145/3473682.3480277

[133] Marjorie H Woollacott and Pei-Fang Tang. 1997. Balance control during walking in the older adult: research and its implications. *Physical therapy* 77, 6 (1997), 646–660. doi:10.1093/ptj/77.6.646

[134] Dohyeon Yeo, Gwangbin Kim, and Seungjun Kim. 2020. Toward immersive self-driving simulations: Reports from a user study across six platforms. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12. doi:10.1145/3313831.3376787

[135] Xiangling Zhuang, Changxu Wu, and Shu Ma. 2018. Cross or wait? Pedestrian decision making during clearance phase at signalized intersections. *Accident Analysis & Prevention* 111 (2018), 115–124. doi:10.1016/j.aap.2017.08.019


## A Code Availability

We provide the core scripts and an example Unity scene used in our study, which support the replication of key system functionalities, including multi-user synchronization and eHMI control logic. These materials are available at https://github.com/yuminkang-research/You-re-the-One-Whom-I-m-Talking-To.

## B Figures and Tables

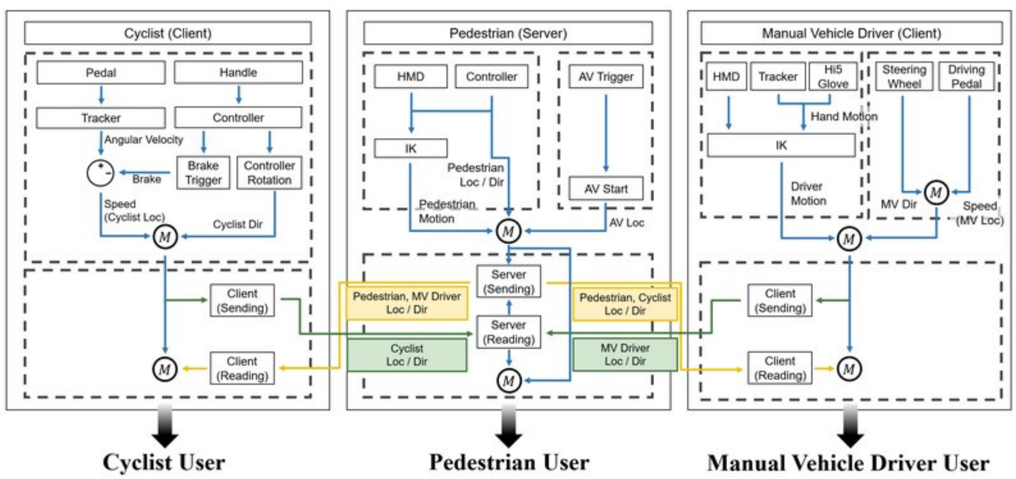


Fig. 12. System architecture for the overall environment and server-client communication among different road user types. Each panel illustrates the interaction flow and component structure for cyclist (left), pedestrian (center), and manual vehicle driver (right) users. The pedestrian client functions as the server, integrating motion and location data received from cyclist and driver clients and broadcasting synthesized data to all users to maintain synchronized multi-agent interaction. Arrows indicate the direction of data flow for motion, location, and direction information. Abbreviations used: MV (Manual Vehicle), IK (Inverse Kinematics), Loc (Location), Dir (Direction), and M (Merge).

Table 2. Main effects of eHMI type and road user type on crossing performance (COT and CD), subjective feelings (perceived safety, trust, clarity), and physiological response (EDA). Subsections (a), (c), (d), (e), (f), and (g) report the main effects of eHMI type within each road user group, while subsection (b) reports the main effect of road user type within each eHMI condition. $^{*}p$ < .05, $^{**}p$ < .01, $^{***}p$ < .001.

**(I) Crossing Onset Time (COT)**

**(a) Main effect of eHMI type for each road user type**

| Road User Type | F(df) | p-value |
|---|---|---|
| Pedestrian | F(2.270, 29.506) = 13.677 | <0.001*** |
| Cyclist | F(3.050, 39.650) = 16.808 | <0.001*** |
| Manual Vehicle Driver | F(2.226, 28.943) = 16.569 | <0.001*** |

**(b) Main effect of road user type for each eHMI type**

| eHMI Type | F(df) | p-value |
|---|---|---|
| No eHMI | F(2, 26) = 4.223 | <0.001*** |
| No Context | F(2, 26) = 10.804 | <0.001*** |
| Whom (Pedestrian) | F(2, 26) = 23.931 | <0.001*** |
| Whom (Cyclist) | F(2, 26) = 45.671 | <0.001*** |
| When | F(2, 26) = 10.577 | <0.001*** |
| Where | F(2, 26) = 9.339 | <0.001*** |

**(II) Crossing Duration (CD)**

**(c) Main effect of eHMI type for each road user type**

| Road User Type | F(df) | p-value |
|---|---|---|
| Pedestrian | F(1.324, 17.207) = 9.585 | 0.004** |
| Cyclist | F(2.413, 31.372) = 6.797 | 0.002** |
| Manual Vehicle Driver | F(1.360, 17.686) = 8.214 | 0.006** |

**(I) Perceived Safety**

**(d) Main effect of eHMI type for each road user type**

| Road User Type | F(df) | p-value |
|---|---|---|
| Pedestrian | F(6, 78) = 15.436 | <0.001*** |
| Cyclist | F(6, 78) = 9.499 | <0.001*** |
| Manual Vehicle Driver | F(3.672, 47.731) = 13.686 | <0.001*** |

**(II) Trust**

**(e) Main effect of eHMI type for each road user type**

| Road User Type | F(df) | p-value |
|---|---|---|
| Pedestrian | F(2.862, 37.205) = 6.867 | <0.001*** |
| Manual Vehicle Driver | F(2.808, 36.50) = 5.692 | 0.003** |

**(III) Clarity**

**(f) Main effect of eHMI type for each road user type**

| Road User Type | $\chi^2$(df) | p-value |
|---|---|---|
| Pedestrian | $\chi^2(6) = 54.117$ | <0.001*** |
| Cyclist | $\chi^2(6) = 47.737$ | <0.001*** |
| Manual Vehicle Driver | $\chi^2(6) = 49.544$ | <0.001*** |

**(IV) Physiological response (EDA phasic mean)**

**(g) Main effect of eHMI type for each road user typ**

| Road User Type | F(df) | p-value |
|---|---|---|
| Pedestrian | F(6, 60) = 2.261 | 0.023* |
| Cyclist | F(6, 42) = 2.552 | 0.034* |
| Manual Vehicle Driver | F(6, 66) = 0.750 | 0.612 |

*Reference only: Pairwise comparisons including Whom (Other-targeted) eHMI conditions are reported here for transparency, but were excluded from the main statistical analysis.*

Table 3. Mean Difference (MD) and Standard Error (SE) across different metrics: (a) COT, (b) CD, (c) Perceived Safety, (d) Trust, (e) Clarity, and (f) EDA Phasic Mean. Each row reports the MD between two eHMI types, the corresponding SE, and p-value. For (a) and (b), comparisons were conducted within each road user group based on significant interaction effects. For (c), (d), and (e), comparisons reflect the main effect of eHMI type; note that (e) was analyzed using the Aligned Rank Transform (ART) method and therefore reports estimated differences. (f) presents the significant comparison observed in the EDA analysis. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

**(a) Crossing Onset Time (COT)**

| Road User Type | eHMI Type | | MD | SE | p-value |
|---|---|---|---|---|---|
| **Pedestrian** | No eHMI | Whom (Pedestrian) eHMI | 15.068 | 2.542 | <0.001*** |
| | No eHMI | When eHMI | 12.907 | 2.875 | 0.013* |
| | No Context eHMI | Whom (Pedestrian) eHMI | 7.343 | 1.369 | <0.001*** |
| | No Context eHMI | When eHMI | 5.182 | 1.272 | 0.046* |
| | Whom (Pedestrian) eHMI | Whom (Cyclist) eHMI | -11.444 | 1.692 | <0.001*** |
| | Whom (Pedestrian) eHMI | Whom (Driver) eHMI | -9.039 | 1.818 | 0.003** |
| | Whom (Cyclist) eHMI | When eHMI | 9.283 | 1.468 | <0.001*** |
| | Whom (Cyclist) eHMI | Where eHMI | 6.663 | 1.621 | 0.041* |
| | Whom (Driver) eHMI | When eHMI | 6.879 | 1.184 | <0.001*** |
| **Cyclist** | No eHMI | Whom (Cyclist) eHMI | 12.807 | 3.065 | 0.034* |
| | No Context eHMI | Whom (Cyclist) eHMI | 7.156 | 1.713 | 0.034* |
| | Whom (Pedestrian) eHMI | Whom (Cyclist) eHMI | 11.976 | 1.692 | <0.001*** |
| | Whom (Pedestrian) eHMI | When eHMI | 7.052 | 1.387 | 0.002** |
| **Driver** | No eHMI | Whom (Pedestrian) eHMI | 11.606 | 2.542 | 0.010* |
| | No eHMI | Whom (Driver) eHMI | 19.294 | 3.315 | <0.001*** |
| | No eHMI | When eHMI | 14.519 | 2.875 | 0.002** |
| | No Context eHMI | Whom (Driver) eHMI | 11.449 | 1.647 | <0.001*** |
| | No Context eHMI | When eHMI | 6.674 | 1.272 | 0.001** |
| | Whom (Pedestrian) eHMI | Whom (Cyclist) eHMI | -9.791 | 1.692 | <0.001*** |
| | Whom (Pedestrian) eHMI | Whom (Driver) eHMI | 7.688 | 1.818 | 0.029* |
| | Whom (Cyclist) eHMI | Whom (Driver) eHMI | 17.479 | 1.754 | <0.001*** |
| | Whom (Cyclist) eHMI | When eHMI | 12.703 | 1.468 | <0.001*** |
| | Whom (Cyclist) eHMI | Where eHMI | 8.426 | 1.621 | 0.001** |
| | Whom (Driver) eHMI | Where eHMI | -9.052 | 1.329 | <0.001*** |
| | When eHMI | Where eHMI | -4.276 | 0.938 | 0.010* |

**(b) Crossing Duration (CD)**

| Road User Type | eHMI Type | | MD | SE | p-value |
|---|---|---|---|---|---|
| **Pedestrian** | No eHMI | No Context eHMI | 4.129 | 0.891 | 0.008** |
| | No eHMI | Whom (Pedestrian) eHMI | 6.809 | 1.175 | <0.001*** |
| | No eHMI | Whom (Cyclist) eHMI | 6.040 | 1.159 | 0.001** |
| | No eHMI | Whom (Driver) eHMI | 5.681 | 1.032 | <0.001*** |
| | No eHMI | When eHMI | 5.777 | 1.057 | <0.001*** |
| | No eHMI | Where eHMI | 5.127 | 1.063 | 0.005** |
| | No Context eHMI | Whom (Pedestrian) eHMI | 2.681 | 0.430 | <0.001*** |
| | No Context eHMI | Whom (Cyclist) eHMI | 1.911 | 0.465 | 0.041* |
| | No Context eHMI | Whom (Driver) eHMI | 1.552 | 0.296 | 0.001** |
| | No Context eHMI | When eHMI | 1.649 | 0.358 | 0.009** |
| | Whom (Pedestrian) eHMI | Whom (Driver) eHMI | -1.129 | 0.267 | 0.029* |
| | Whom (Pedestrian) eHMI | When eHMI | -1.032 | 0.224 | 0.009** |
| | Whom (Cyclist) eHMI | Where eHMI | -1.682 | 0.297 | <0.001*** |

**(c) Perceived Safety**

| eHMI Type | | MD | SE | p-value |
|---|---|---|---|---|
| No eHMI | No Context eHMI | -0.951 | 0.199 | <0.001*** |
| No eHMI | Whom (Pedestrian) eHMI | -1.872 | 0.199 | <0.001*** |
| No eHMI | Whom (Cyclist) eHMI | -1.991 | 0.190 | <0.001*** |
| No eHMI | Whom (Driver) eHMI | -1.808 | 0.195 | <0.001*** |
| No eHMI | When eHMI | -1.443 | 0.208 | <0.001*** |
| No eHMI | Where eHMI | -1.110 | 0.197 | <0.001*** |
| No Context eHMI | Whom (Pedestrian) eHMI | -0.920 | 0.158 | <0.001*** |
| No Context eHMI | Whom (Cyclist) eHMI | -1.040 | 0.171 | <0.001*** |
| No Context eHMI | Whom (Driver) eHMI | -0.857 | 0.140 | <0.001*** |
| Whom (Pedestrian) eHMI | Where eHMI | 0.761 | 0.169 | 0.001** |
| Whom (Cyclist) eHMI | When eHMI | 0.548 | 0.135 | 0.005** |
| Whom (Cyclist) eHMI | Where eHMI | 0.881 | 0.156 | <0.001*** |
| Whom (Driver) eHMI | Where eHMI | 0.698 | 0.154 | 0.001** |

**(d) Trust**

| eHMI Type | | MD | SE | p-value |
|---|---|---|---|---|
| No eHMI | Whom (Pedestrian) eHMI | -1.310 | 0.254 | <0.001*** |
| No eHMI | Whom (Cyclist) eHMI | -1.263 | 0.241 | <0.001*** |
| No eHMI | Whom (Driver) eHMI | -1.348 | 0.217 | <0.001*** |
| No eHMI | When eHMI | -0.940 | 0.190 | <0.001*** |
| No Context eHMI | Whom (Pedestrian) eHMI | -0.860 | 0.240 | 0.019* |
| No Context eHMI | Whom (Cyclist) eHMI | -0.812 | 0.239 | 0.033* |
| No Context eHMI | Whom (Driver) eHMI | -0.897 | 0.200 | 0.001** |
| Whom (Cyclist) eHMI | Where eHMI | 0.790 | 0.239 | 0.042* |
| Whom (Driver) eHMI | Where eHMI | 0.875 | 0.209 | 0.003** |

**(e) Clarity**

| eHMI Type | | Estimate | SE | p-value |
|---|---|---|---|---|
| No eHMI | No Context eHMI | -84.548 | 15.600 | <0.001*** |
| No eHMI | Whom (Pedestrian) eHMI | -175.060 | 15.600 | <0.001*** |
| No eHMI | Whom (Cyclist) eHMI | -158.810 | 15.600 | <0.001*** |
| No eHMI | Whom (Driver) eHMI | -161.452 | 15.600 | <0.001*** |
| No eHMI | When eHMI | -112.190 | 15.600 | <0.001*** |
| No eHMI | Where eHMI | -84.440 | 15.600 | <0.001*** |
| No Context eHMI | Whom (Pedestrian) eHMI | -90.512 | 15.600 | <0.001*** |
| No Context eHMI | Whom (Cyclist) eHMI | -74.262 | 15.600 | <0.001*** |
| No Context eHMI | Whom (Driver) eHMI | -76.905 | 15.600 | <0.001*** |
| Whom (Pedestrian) eHMI | When eHMI | 62.869 | 15.600 | 0.030* |
| Whom (Pedestrian) eHMI | Where eHMI | 90.619 | 15.600 | <0.001*** |
| Whom (Cyclist) eHMI | Where eHMI | 74.369 | 15.600 | <0.001*** |
| Whom (Driver) eHMI | When eHMI | 49.262 | 15.600 | 0.046* |
| Whom (Driver) eHMI | Where eHMI | 77.012 | 15.600 | <0.001*** |

**(f) EDA Phasic Mean**

| **Road User Type** | **eHMI type** | | **MD** | **SE** | **p-value** |
|---|---|---|---|---|---|
| **Cyclist** | Whom (Driver) | No Context | -0.348 | 0.093 | 0.031* |

*Reference only: Pairwise comparisons including Whom (Other-targeted) eHMI conditions are reported here for transparency, but were excluded from the main statistical analysis.*

Table 4. Statistical significance of road user type comparisons for each eHMI type on COT. The table reports pairwise differences in COT between pedestrians, cyclists, and manual vehicle drivers under each eHMI type (*No eHMI*, *No Context eHMI*, *Whom eHMI*, *When eHMI*, and *Where eHMI*. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

| **eHMI Type** | **Road User Type** | | **p-value** |
|---|---|---|---|
| No eHMI | Cyclist | Driver | 0.048* |
| No Context | Pedestrian | Driver | 0.003** |
| No Context | Cyclist | Driver | 0.003** |
| Whom (Pedestrian) | Pedestrian | Cyclist | < 0.001*** |
| Whom (Pedestrian) | Pedestrian | Driver | < 0.001*** |
| Whom (Cyclist) | Pedestrian | Cyclist | < 0.001*** |
| Whom (Cyclist) | Pedestrian | Driver | 0.002** |
| Whom (Cyclist) | Cyclist | Driver | < 0.001*** |
| When | Pedestrian | Driver | < 0.001*** |
| When | Cyclist | Driver | 0.034* |
| Where | Pedestrian | Driver | 0.003** |
| Where | Cyclist | Driver | 0.012* |

*Reference only: Pairwise comparisons including Whom (Other-targeted) eHMI conditions are reported here for transparency, but were excluded from the main statistical analysis.*