Around the World in 60 Cyclists: Evaluating Autonomous Vehicle-Cyclist Interfaces Across Cultures

Ammar Al-Taie
School of Computing Science
University of Glasgow
Glasgow, United Kingdom
a.al-taie.1@research.gla.ac.uk

Andrii Matviienko
KTH Royal Institute of Technology
Stockholm, Sweden
andriim@kth.se

Joseph O'Hagan
School of Computing Science
University of Glasgow
Glasgow, United Kingdom
josephaohagan@gmail.com

Frank Pollick
School of Psychology and Neuroscience
University of Glasgow
Glasgow, United Kingdom
frank.pollick@glasgow.ac.uk

Stephen Anthony Brewster
School of Computing Science
University of Glasgow
Glasgow, United Kingdom
stephen.brewster@glasgow.ac.uk

Figure 1: Our cross-cultural user study was conducted in three cities in different countries. Cyclists used a novel AR simulator to cycle in real physical space and interact with moving virtual AVs projected onto the real world.

Abstract

Cultural differences influence how cyclists and drivers interact, affecting global autonomous vehicle (AV) adoption. AV-cyclist interfaces are needed to clarify AV intentions and resolve ambiguities when no human driver is present. These must adapt across cultures and road infrastructure. We conducted the first cross-cultural AV-cyclist user study across *Stockholm* (high segregation of cyclists from drivers), *Glasgow* (some segregation), and *Muscat* (no segregation). Cyclists used an AR simulator to cycle in physical space and experienced three holistic AV-cyclist interfaces. These integrated multiple interfaces into a larger ecosystem, e.g., a smartwatch synchronised with on-vehicle eHMI. Interfaces communicated AV location, intentions, or both. Riders from all cities preferred combined AV location and intention information but used it differently. *Stockholm* cyclists focused on location, validating intentions with driving behaviour. *Glasgow* riders valued both cues equally. *Muscat* cyclists trusted interfaces, prioritising intentions without relying on driving behaviour. These insights are key for global AV adoption.

CCS Concepts

• **Human-centered computing → HCI design and evaluation methods**.

Keywords

Autonomous Vehicle-Cyclist Interaction, Cross-Cultural Study, Augmented Reality

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713407

ACM Reference Format:
Ammar Al-Taie, Andrii Matviienko, Joseph O'Hagan, Frank Pollick, and Stephen Anthony Brewster. 2025. Around the World in 60 Cyclists: Evaluating Autonomous Vehicle-Cyclist Interfaces Across Cultures. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3706598.3713407

1 Introduction

Cycling is gaining worldwide popularity as people recognise its benefits for health and environmental sustainability [1]. However, encounters with motorised vehicles remain a significant safety concern [3]. According to the World Health Organisation, over 41,000 cyclist fatalities occur worldwide annually due to collisions [25]. Research has shown that cyclists must know driver intentions to clarify traffic ambiguities and safely navigate shared roads [2, 14]. Interpreting this information implicitly from vehicle driving and braking behaviours may be insufficient or misleading, so social interaction is common [9, 27]. This happens when drivers and cyclists explicitly communicate their next manoeuvres through social cues such as eye contact or hand gestures [2, 11, 27]. However, social cues may vary between cultures; for example, a raised hand may indicate a thanks gesture in one culture but a request to stop in another [19]. Different countries also have different levels of cyclist segregation from motorised vehicles, which could impact interaction behaviours [13]. Cyclists accustomed to more segregated cycle lanes typically only encounter vehicles at more controlled scenarios, such as intersections [38], so interaction may be slower-paced with more expressive or detailed signals, such as facial expressions [2]. In contrast, riders who frequently cycle in mixed traffic face more dynamic situations, such as lane merging, with faster-paced interactions on-the-move [24, 26]. Here, the exchanged cues may be less expressive; for example, road users may only establish eye contact to quickly validate they are aware of each other [2, 43].

As autonomous vehicles (AVs) join our roads, the social cues will disappear, and interfaces, such as external Human-Machine Interfaces (eHMIs) on vehicles [14], must compensate to provide clear communication [9, 14, 27]. Interfaces must also accommodate cultural variations in social interaction, and messages must be universally understood for the successful global integration of AVs [38]. Research on AV-cyclist interfaces has been predominantly conducted in single, mostly Western countries [4, 9, 24]. Different interfaces were proposed and evaluated for usability, including augmented reality (AR) glasses [46], vibrating bike handlebars [24], and eHMIs [6]. However, their effectiveness across broader cultural settings remains unclear. More recently, Holistic AV-cyclist Interfaces, or more simply 'holistic interfaces', were proposed [5]. These synchronise individual interfaces into a larger interconnected ecosystem to facilitate interaction collectively. For example, handlebar vibrations could be synchronised with pulsing eHMI animations, and AR glasses could project eHMI state onto the road for enhanced visibility [2, 5]. Holistic interfaces provide key advantages by making individual interfaces multimodal or extending the reach of displays, such as eHMIs, onto wearables [5]. However, they add further complications to cross-cultural problems by introducing a wider range of placements, modalities, and messages. Cyclists from different cultures may require different combinations of this large set of features, making the evaluation of holistic interfaces in single cultures insufficient [38]. So far, research on holistic interfaces has primarily been conducted in the UK [2, 5], where cyclists navigate a mix of segregated lanes and shared road spaces [2]. These studies suggested multimodal cues to communicate an AV's location and intentions simultaneously [5]. However, these insights may not apply universally. In countries like Sweden or the Netherlands, where cycling infrastructure includes highly segregated lanes [38], communicating an AV's location may be less critical since interactions usually happen in controlled scenarios with AVs in direct view [22]. On the other hand, in regions without cycle lanes and AVs may be in blindspots behind cyclists [2], AV location and intention information may be necessary for safe interaction [38, 43].

We explored cultural differences in how cyclists use holistic interfaces to identify culturally-inclusive placements, messages and modalities. We conducted the first cross-cultural AV-cyclist user study. The study was conducted across three cities, each in a different country to maximise the potential cultural differences [43]. Given the potential impact of physical cycling infrastructure on cycling behaviour [38, 43], we also ensured that each city had a different level of cyclist segregation from motorised vehicles: *Stockholm*, Sweden (highly segregated cycle lanes); *Glasgow*, UK (some segregated lanes, but cyclists ride in mixed traffic for parts of their journeys); and *Muscat*, Oman (no cycle lanes). Twenty cyclists from each city used an AR cycling simulator to evaluate three multimodal holistic interfaces communicating AV location, intention, or both. The simulator projected virtual buildings, traffic features, and AVs aligned with the real world. Cyclists interacted with moving AVs in three traffic scenarios while riding a real bike in real physical space. We found that cyclists across all cities preferred combined AV location and intention information but used it differently. *Stockholm* participants focused more on AV location and validated any intention signals with driving behaviour rather than fully trusting interface signals. In *Glasgow*, cyclists valued AV location and intention equally. In contrast, *Muscat* cyclists prioritised AV intention messages, placing more trust in interfaces and less need to verify the AV's behaviour visually. We contribute (1) the first cross-cultural AV-cyclist interaction study with participants using interfaces and interacting with AVs; (2) identification of novel cultural differences in cyclist perceptions of AVs and AV-cyclist interfaces; (3) insights into cycling behaviour in the Sultanate of Oman—a region where no prior automotive user interface or cycling user studies have been conducted; (4) novel design guidelines for culturally inclusive AV-cyclist interfaces. These findings are crucial for understanding how to integrate AVs globally across diverse traffic environments, ensuring that solutions are inclusive and effective worldwide.

2 Related Work

Autonomous vehicles will diminish the social interactions cyclists rely on to navigate shared road spaces safely [27]. Explicit communication of AV intentions is necessary to compensate for the loss of traditional driver cues [14, 27]. In real-world observations of autonomous shuttle-cyclist encounters, Pelikan [35] and Pokorny et al. [36] found that the lack of social communication between AVs and cyclists caused significant issues and ambiguities. For example, cyclists merged into lanes with a shuttle behind them without clear signals from the vehicle, leading to emergency braking and causing cyclists behind the shuttle to swerve into oncoming traffic, compromising their safety. These findings underscore the need for interfaces that explicitly communicate AV intentions to cyclists [9]. This was further emphasised by Hagenzieker et al. [20], who conducted a study with cyclists comparing photographs of human-driven vs autonomous vehicle-cyclist encounters. They found that cyclists were more confident in the awareness of human drivers due to eye contact, suggesting that AV-cyclist interfaces must compensate for the lack of traditional social cues to ensure predictable cycling experiences. A range of studies followed Hagenzieker et al. [20]'s work to design and evaluate novel AV-cyclist interfaces [9]. Interfaces consistently outperformed baseline *no interface* conditions in these studies [4, 14, 24, 32, 46, 48]. This further emphasises the need for AV-cyclist interfaces in future traffic. However, these must work across traffic cultures to be successful globally [14].

2.1 AV-Cyclist Interaction Across Cultures

Traffic culture impacts how cyclists perceive motorised vehicles and interact with drivers, directly influencing AV-cyclist interface design [43]. In a survey by Rodríguez Palmeiro et al. [38], cyclists from 15 Western countries evaluated photographs of AV-cyclist encounters at intersections. Cyclists from all countries felt more confident when AVs had a visible sign indicating they were self-driving, suggesting that AV-cyclist interfaces could be effective in these regions. However, perceptions varied based on the level of cyclist segregation from vehicles. Cyclists in countries with highly segregated cycle lanes, like the Netherlands and Denmark, had greater confidence in the AV's awareness and expected it to give them priority; AV-cyclist interfaces in these countries may need more focus on communicating intentions than awareness [2]. In contrast, cyclists from regions with less developed cycling infrastructure, such as North America, expressed lower confidence, citing concerns about AV reliability in shared roads, so clearer messages on both intentions and awareness may be needed [2, 6]. AV-cyclist interfaces should adapt to these varied expectations to ensure safety.

Similarly, Chataway et al. [13] conducted a survey to compare self-reported cycling behaviour between Brisbane and Copenhagen. Brisbane cyclists, with partially segregated lanes, exhibited a greater fear of traffic and felt less prioritised. Conversely, Copenhagen cyclists, benefiting from highly segregated lanes, were more likely to cycle while distracted. These findings indicate that differences in cycling infrastructure influence cyclist trust and expectations of vehicles. In our study, we chose regions with varying levels of cyclist segregation from vehicles to capture these differences and inform the design of interfaces adaptable to different traffic environments. Even countries with similar cycling infrastructure can exhibit differences in how cyclists behave. For example, Haustein et al. [22] conducted a survey comparing cycling behaviour in Stockholm and Copenhagen; both have highly segregated cycle lanes. Despite these similarities, cyclists in Copenhagen perceived themselves to have a higher priority in traffic and cycled more frequently than those in Stockholm. Additionally, Useche et al. [43] surveyed cyclists from 19 countries, revealing notable differences in road behaviour. For example, cyclists from African and Asian countries were more likely to take risky manoeuvres, while German cyclists were found to commit the most traffic violations. Therefore, we involved cyclists from three cities in different countries to maximise behavioural differences between our study locations.

To address cross-cultural challenges in AV-cyclist interface design, Al-Taie et al. [5] surveyed cyclists from 23 countries about their expectations of AVs. They found that most cyclists had not yet encountered AVs and could not define clear expectations without direct experience. This motivated our approach; rather than conducting a cross-cultural survey, we ran a user study with cyclists from different regions interacting directly with AVs, giving clear insights into cross-cultural perceptions and behaviours toward AVs and AV-cyclist interfaces. The authors [5] also asked cyclists to report their current devices to explore whether existing devices could be adapted for interaction or if new ones should be developed. Cyclists carried various devices, such as bike computers, phones and smartwatches. However, they strongly preferred interfaces placed on the vehicle or environment rather than on themselves. Berge et al. [11] interviewed cyclists in the Netherlands and found similar preferences for interface placements on the AV or environment.

In addition to cultural differences [13, 38], cyclists will encounter AVs in diverse traffic scenarios [8], such as roundabouts or lane merging, each with unique challenges [23]. Al-Taie et al. [2] studied human driver-cyclist encounters to inform AV-cyclist interface design. They first observed encounters in five scenarios with distinct traffic features and vehicle positions. Over 50% of encounters involved interaction, reiterating the need for AV-cyclist interfaces. However, the social cues and messages exchanged differed between scenarios. Therefore, interfaces should be versatile and adapt not only to different cultures but also to scenarios [4, 6]. In a second study [2], the researchers equipped cyclists with eye trackers and asked them to record their commutes to analyse their attention during vehicle encounters. Cyclist gaze behaviour also varied between scenarios. For instance, they focused more on traffic control features like road markings when available to anticipate driver intentions and assess right-of-way. Our study explored three scenarios with different levels of traffic control and AV positions to investigate the versatility of the evaluated interfaces.

2.2 AV-Cyclist Interface Limitations

So far, no studies have focused on designing and evaluating AV-cyclist interfaces across cultures. Berge et al. [9] reviewed existing interfaces, including AR glasses worn by cyclists [47], eHMIs [45], and helmet-mounted audio [24]. However, these were primarily studied in single, mostly Western countries. A similar review by Dey et al. [14], which focused on eHMIs, echoed this finding, highlighting a significant research gap in the field. Interfaces designed and tested in one cultural setting may not generalise to others and could cause miscommunication when deployed elsewhere [13, 38]. Despite these challenges, studies consistently agreed that AV-cyclist interfaces must at least communicate basic vehicle intentions, typically using binary signals like AV-yielding or not yielding to the cyclist [4, 14, 24]. However, holistic interfaces can potentially communicate more comprehensive information using a wider range of modalities [5].

Despite cyclists preferring on-vehicle interfaces [3, 9], these solutions have limitations. Hou et al. [24] designed and evaluated interfaces for a lane-merging scenario, including AR glasses, eHMIs, and vibrating handlebars. In a Virtual Reality (VR) simulator study, cyclists merged lanes with an AV behind them. Results revealed that eHMIs are ineffective when the AV is behind the cyclist, leading to frequent shoulder checks. Al-Taie et al. [4, 6] reached similar conclusions. First, they conducted design sessions with cyclists and HCI researchers collaboratively sketching eHMIs on stationary vehicles [6]. Cyclists suggested features like visual and auditory cues to convey AV intentions and animations to indicate AV awareness. Designs were evaluated across five traffic scenarios in a VR simulator, followed by an outdoor Wizard-of-Oz study [4]; it became clear that eHMIs could become overloaded with information, leading to high cognitive demands on cyclists. They recommended that eHMIs stick to visual cues and limit their focus to communicating basic AV intentions. Like Hou et al. [24], they also found that eHMIs are ineffective when the AV was occluded or out of view in lane merging scenarios.

Both Hou et al. [24] and Al-Taie et al. [4] used VR simulators in their evaluations. While these display higher-fidelity visual displays compared to outdoor studies [39], they require cyclists to be stationary [49], significantly limiting the authenticity of the riding experience. To address this, Matviienko et al. [32] developed an AR simulator using HoloLens AR glasses to evaluate AR displays for AV-cyclist interactions. Cyclists wore the AR headset and cycled in real physical space to experience different displays around moving virtual AVs. AR displays were effective in facilitating interaction. However, the concern for added responsibility on cyclists remained, as not all cyclists may have access to AR glasses [11]. Nevertheless, the study effectively balanced the high-fidelity displays with authentic riding experiences. However, the HoloLens had limitations with a narrow field of view and translucent AR objects, reducing immersion. Newer simulators improved this with advancements in mixed-reality passthrough technology. For example, Aleva et al. [7] conducted a study with pedestrians using mixed-reality headsets that offered a wider field of view and clearer AR objects integrated with the real world. We applied this mixed-reality approach for cyclists in our study.

2.3 Holistic Interfaces Overcome Limitations, But Bring Cross-Cultural Challenges

Given cyclists' global preference for not needing to carry devices to safely share the road with AVs [3, 11] and the limitations of on-vehicle interfaces [24], Al-Taie et al. [5] recommended holistic interfaces. These integrate multiple interfaces into an interconnected ecosystem that uses different placements and modalities to enhance interaction. Holistic interfaces minimise the risk of conflicting signals cyclists may receive if individual interfaces were used independently without integration. They align with cyclist tendencies to rely primarily on vehicle or environment displays [2] while benefiting from additional, multimodal feedback from wearable or bike-mounted devices. For example, an eHMI could serve as the main display, while a smartwatch synchronised with eHMI animations could extend it to communicate AV intentions, even when the AV is out of view [5]. The authors conducted six design sessions where cyclists and HCI researchers collaborated around stationary vehicles, using props like helmets and bicycles to sketch holistic interface concepts and place them accordingly. Their findings stressed that holistic interfaces should communicate the AV's location and intentions, such as "an AV behind you is going to yield".

From this, they developed, but did not evaluate, a novel holistic interface that met participant needs and communicated a combination of AV location and intentions [5]. We used this design as a starting point for our evaluation. It was straightforward to separate these messages, allowing us to synthesise holistic interfaces communicating AV location, intention or both, suitable for our cross-cultural approach to identify the optimal messages between cultures [38]. Holistic interfaces remain largely conceptual and only explored in UK-based design sessions [5]. Given the diverse combinations of devices, cues, modalities, and placements holistic interfaces could employ, deploying them across cultures introduces significant complexity, especially if they communicate more than simple, binary AV intentions.

2.4 Motivation and Research Question

Traffic culture and cycling infrastructure significantly impact how cyclists perceive and interact with AVs [38]. Previous research has established the need for AV-cyclist interfaces to provide clear communication [35], but these have only been designed and tested within single countries [9, 14, 24]. To be effective globally, they must adapt to diverse cultural contexts [14, 38]. Cyclists will also encounter AVs in diverse traffic scenarios [2], such as intersections or lane merging, each presenting unique challenges [6]. Interfaces must be versatile to work effectively across scenarios [4, 23]. Recently, holistic interfaces were recommended [5]; these integrate individual interfaces into an interconnected ecosystem, allowing cyclists to rely on vehicle or environment displays while receiving additional, multimodal feedback from wearable or bike-mounted devices. However, holistic interfaces add complexity to cross-cultural challenges, introducing a large variety of device placements, messages, and modalities. So far, they have only been explored in UK-based design sessions [5] and were not evaluated with cyclists, so evaluation is a clear next step. Initial findings suggest they should use multimodal cues to communicate AV location and intentions, which may not apply across different cultures. Traditional evaluation within a single country is insufficient [13, 38]. Understanding cultural differences in how cyclists perceive holistic interfaces across different traffic scenarios is crucial for global adoption. Therefore, we ask:

RQ Considering holistic interfaces communicating *AV location, AV intentions or a combination of both*, how do cyclist perceptions and behaviours toward these interfaces differ across cultures and traffic scenarios?

3 Method

To answer the research question, we conducted a cross-cultural user study in *Stockholm* (highly segregated cycle lanes from vehicles), *Glasgow* (some segregation), and *Muscat* (no segregation). In each city, cyclists used our AR simulator to interact with AVs in three traffic scenarios. Participants tested a baseline *eHMI*-only condition and three holistic interfaces communicating AV location, intentions, and a combination. This study is the first to reveal cultural differences in cyclist use of AV-cyclist interfaces following direct interaction.

3.1 Interfaces

The evaluated interfaces are shown in Figure 2. Videos of each interface are supplementary materials. The interfaces were activated when an AV was within 15 metres of the cyclist. They were multimodal, each comprising an *eHMI*, AR glasses, a vibrating smartwatch and audio from a bike helmet. The same *eHMI* was used for all conditions. They worked as follows:

- ***eHMI***: We used the LightRing *eHMI* [4]. This was a cyan light bar around the vehicle to communicate that it is in autonomous mode and that all sensors function correctly. The lights pulse slowly in green (2 pulses per second) if the AV has recognised the cyclist and is yielding and flashes quickly in red (3 flashes per second) when not yielding. We chose LightRing because it was positively evaluated with cyclists in VR and real-world settings [4]. It was also shown to be easily integrated with other displays in a holistic interface; its simple AV-yielding/not-yielding signals are simple to translate across modalities [5].
- ***FullIntel***: We adapted Al-Taie et al.'s [5] holistic interface. This communicates the AV's location and intentions, making it a suitable starting point for answering the RQ. AR glasses display a traffic sign on the road scene; this explicitly specifies the AV location through text to avoid ambiguity (similar to road repair signs), e.g. "AV behind you". AR also augments all road markings within a 10-metre radius from the cyclist to flash red/pulse green in sync with the *eHMI*. Cyclists can check AV intentions on this large road surface through quick glances [4]. The smartwatch pulses/vibrates in sync with the *eHMI* to communicate AV intentions: two pulses per second if the AV is yielding and three if not. Vibrations are felt more strongly as the AV moves closer to communicate proximity. Spatial audio from the helmet loops a ringing noise if the AV is yielding and a beeping noise if not. It points toward the AV to communicate its location.
- ***Locator***: A variation of *FullIntel* that only communicates AV location so cyclists can check the *eHMI* for the AV's intentions. AR glasses display the traffic sign specifying AV location using text. The smartwatch vibrates a continuous tone that gets stronger as the AV moves closer to communicate proximity. Spatial helmet audio loops a sonar-like sound pointing to the AV. The same sound is played independent of AV-yielding intentions.
- ***Mirror***: Uses features from *FullIntel* to communicate AV intentions without location. AR glasses augment all road markings in a 10-metre radius from the cyclist to flash red or pulse green, synchronised with the *eHMI*. Smartwatch vibrations pulse in the same rhythm as the *eHMI*, but the pulse intensity remains consistent independent of AV proximity. Helmet speakers also loop a ringing noise if the AV is yielding and a beeping noise if not. However, the audio is not spatial.

3.2 Apparatus

We developed an AR cycling simulator using Unity (see Figure 3). It projected virtual objects around participants as they used a real bike to cycle in physical space while wearing a mixed-reality headset. This included AR buildings, road markings, and AVs, which were white Citroen C3 city cars. All AR objects were to-scale and aligned with the real world. The simulator was deployed on a Meta Quest 3[1] headset, which provided passthrough and depth-sensing features. This presented a wide field of view with a clear, coloured real-world background and opaque AR objects for enhanced immersion. The headset also supported spatial audio and controller haptics, suitable for multimodal interfaces. The left controller was attached to the participant's wrist to simulate smartwatch vibrations and was visually represented as a smartwatch in AR.

The AR simulator effectively balanced real riding with high-fidelity interface implementations [31, 32]. In contrast, VR simulators provide high-fidelity displays but require stationary riding that does not feel realistic, and outdoor Wizard-of-Oz studies provide real riding but use low-fidelity displays [4, 24, 29]. Our simulator used a *headset-only* setup for easy transport between cities, making it more practical than a VR simulator which would require additional hardware like a static bicycle trainer [49], or outdoor studies that require real vehicles and physical props to simulate obstacles [15]. It also allowed us to explore the interfaces with non-yielding AVs, while outdoor studies can only use yielding vehicles to preserve participant safety [4]; even then, AV driving behaviours are often inconsistent between trials as these typically take a Wizard-of-Oz approach with a hidden human driver [39]. Driving behaviours were identical between sessions in the AR simulator. The experiment was conducted indoors to maintain consistent environmental conditions across locations, such as weather, and to protect the headset from potential damage due to rain or snow. We rented a hall in each city. Each hall measured at least 25 metres in length and 12 metres in width. The flooring was dark grey to black, simulating tarmac and providing a contrasting background for the white AR road markings. Participants used common city bikes: a Van Moof S5 in *Stockholm*, a Giant Escape 3 in *Glasgow*, and a Scott Metrix 20 in *Muscat*. For safety, participants were also provided with helmets.

3.3 Study Design

The experiment had *City* as a between-subjects independent variable and *Interface* and *Traffic Scenario* as within-subjects variables. *Cities* were in different countries to maximise the differences between traffic cultures [43]. They were chosen for their distinct levels of cyclist segregation from motorised vehicles [38]: *Stockholm*, Sweden, has highly segregated cycle lanes, providing minimal encounters with motorised traffic. *Glasgow*, UK, has some segregated lanes, and cyclists navigate mixed traffic for parts of their commutes. *Muscat*, Oman, has no dedicated cycle lanes, requiring cyclists to share the road with motorised vehicles at all times. Research showed that cyclists will encounter AVs across a range of traffic scenarios, with different levels of traffic control and AV positions around a cyclist [2, 6, 9, 10]. Therefore, evaluating *Interfaces* in a single traffic scenario, such as an intersection, would be

[1] Meta Quest 3: meta.com/quest/quest-3/

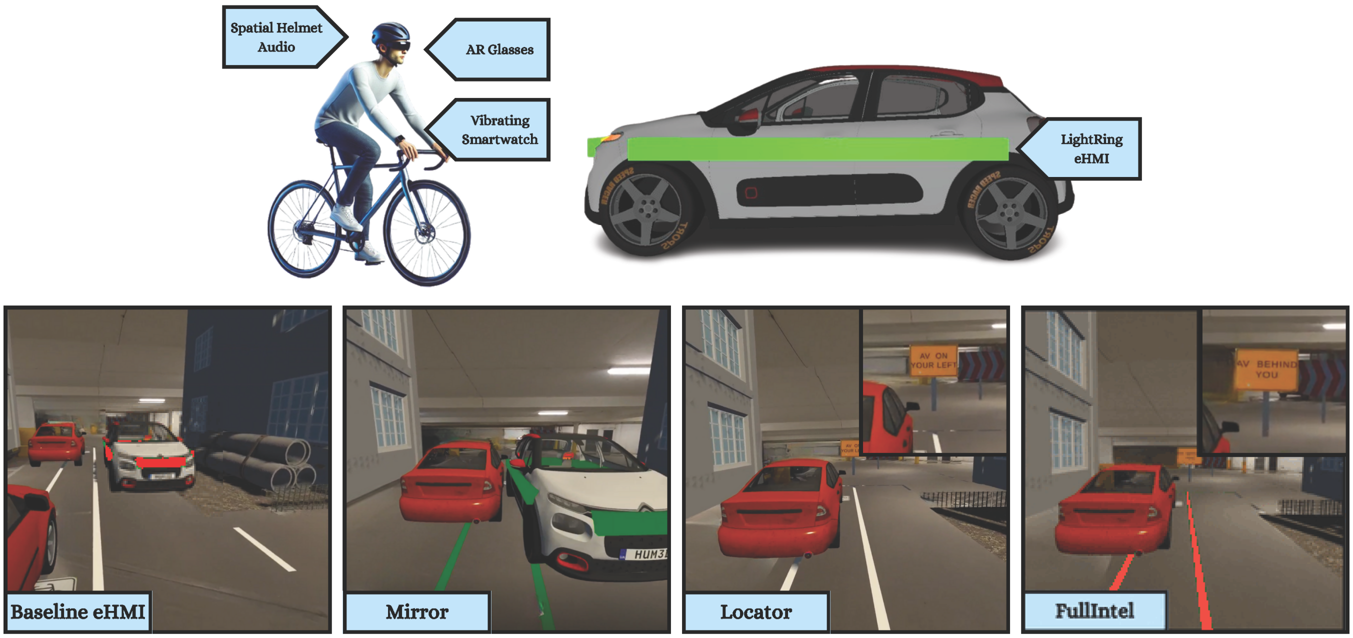

Figure 2: The evaluated interfaces. Top: Common components used in all of the interfaces. Bottom: Cyclist point-of-view in the AR simulator and the visual cues for each interface. Virtual objects were overlaid on the real-world environment. No additional displays were used in the *Baseline eHMI* condition (light band around the vehicle; green means it is yielding, red means not yielding). Road marking projections were synchronised with the eHMI for the *Mirror* condition, and a traffic sign communicated the AV's location for *Locator*. *FullIntel* combined both of these into one interface.

Figure 3: Study Setup. Left: Our headset-only approach with the left controller on the participant's wrist to simulate a smartwatch. Middle: A cyclist riding in real physical space with a mixed-reality headset. Right: The cyclist's point-of-view, with the real world (the indoor hall) overlaid with to-scale AR objects, in this case, the *FullIntel* interface in one of our scenarios.

insufficient [4]. We assessed *Interfaces* across three *Traffic Scenarios*, each with distinct features. This allowed us to explore cultural differences in navigating them and compare *Interface* versatility [4]. *Scenarios* were chosen from prior research on real-world motorised vehicle-cyclist encounters [2, 10], which showed these are common in everyday cycling routes [2, 8].

We set up our AR environment to facilitate the *Scenarios*; see Figure 4. It featured a 15-metre cycle lane beside a two-lane road leading to a three-way intersection. We used standard traffic features [8] for consistency between cities: the cycle lane had bicycle symbol road markings and solid white borders, while the intersection had dashed give-way line road markings. *Scenarios* were modelled according to real-world descriptions and video footage from prior work [2, 8, 37]. We placed obstacles to explore dynamic scenarios. These were red parked cars (to differentiate from white AVs) and road repairs in the lane farther from the cyclists on the two-lane road. All AVs were SAE level 5 vehicles [41]. *Scenarios* were mirrored in *Glasgow* for consistency with the city's right-hand drive traffic infrastructure. They were:

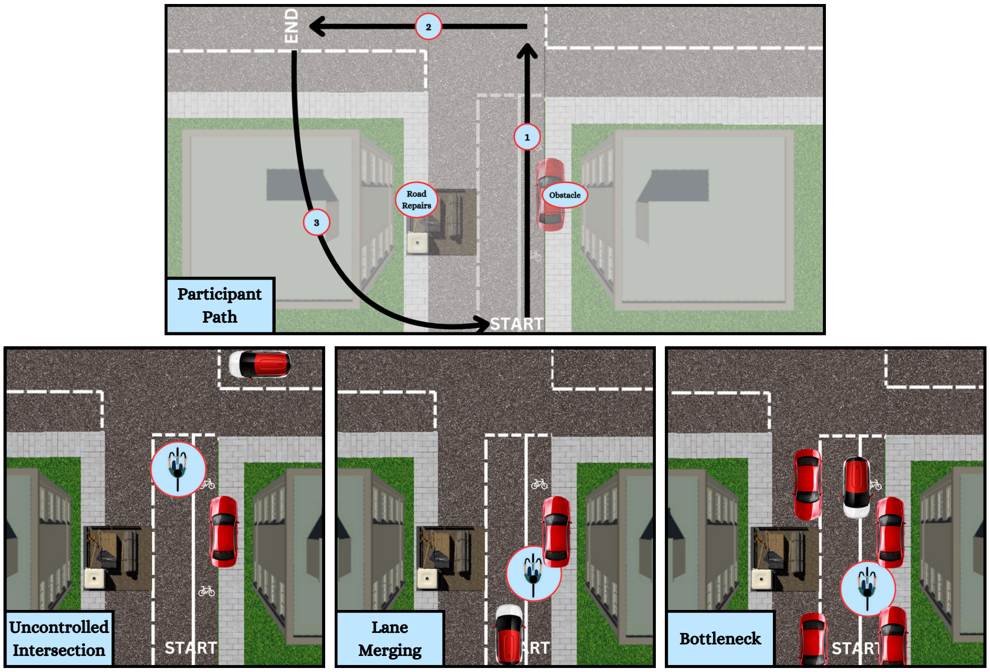

Figure 4: Top: The path participants had to navigate in the scenarios; any AR buildings disappeared after reaching the end road markings. Bottom: Our scenarios: Uncontrolled Intersection, Lane Merging and Bottleneck. These were mirrored in *Glasgow*.

- **Uncontrolled Intersection**: Cyclists approached the intersection. The AV approached from the right (left in *Glasgow*). It accelerated to 50 km/h when 50 metres from the cyclist and stopped 0.5 metres behind the give-way line if yielding. It maintained its speed when not yielding. This scenario involves stationary infrastructure, with give-way lines indicating right-of-way [4, 32, 47]. Intersections are commonly explored for AV-cyclist interaction [32, 33, 47], as they cause high collision rates [25]. This suggests that cyclists would benefit from knowing AV intentions to overcome ambiguity in this scenario [27]. For the first time, we explore intersections with holistic interfaces, across different cultural settings.
- **Lane Merging**: A parked car positioned 10 metres into the cycle lane represented an obstacle. This prompted cyclists to move from the cycle lane and merge lanes with a moving AV behind them. The AV drove at 40 km/h and decelerated to 15 km/h when yielding, maintaining its speed when not yielding. Cyclists previously ranked *Lane Merging* as one of the most challenging scenarios [3, 24]. In contrast to *Uncontrolled Intersection*, *Lane Merging* is dynamic; it can occur anywhere on the road due to obstacles or cycle lanes transitioning to mixed traffic [2, 24]. The primary challenge is that it requires cyclists to interact with moving vehicles behind them [24], triggering shoulder checks [2]. This allowed us to investigate *Interfaces* across different AV positions.
- **Bottleneck**: Parked cars on the cycle lane and adjacent road created a narrow lane between them. Due to road repairs, the AV was redirected to approach from opposite the cyclist [2, 37]. One road user had to steer away to avoid collision. The AV drove at 25 km/h. It steered to the left (right in *Glasgow*), between two parked cars, with the directional indicator active, and stopped when yielding. The AV continued straight, maintaining speed when not yielding. *Bottleneck* is common in urban areas [2, 3, 8, 37], with frequent road repairs and parallel parking setups. Urban areas may have narrower roads [8, 11, 14], making this scenario challenging as obstacles may be closer to cyclists, with little space to steer away [2, 8]. *Bottleneck* is slower-paced than the others, but requires fast decision-making from the cyclist [4], who should know the AV's intentions to make a decision [4, 8]. *Bottleneck* also allowed us to explore *Interface* versatility because in contrast to *Uncontrolled Intersection*, it has no traffic control, and unlike *Lane Merging*, the AV is in the cyclist's view.

3.3.1 ***Task and Measures***. Participants navigated each *Scenario* twice per *Interface* condition: once with a yielding AV and once with a non-yielding AV, to minimise learning effects and maintain focus [4]. We aggregated results from both yielding states for simplicity and clarity in our analysis. All scenarios had the same designated start and endpoints, marked by white road markings labelled *"Start"* and *"End"*. The start was at the beginning of the cycle lane, and the end was 5 metres into the left turn at the intersection (right for *Glasgow*). Participants cycled in a loop from the start to the end, then returned to the start (see Figure 4). Upon crossing the end road markings, any AR buildings disappeared, creating a clear path for participants to loop back to the start and begin the next *Scenario*.

This setup provided visual feedback indicating the *Scenario*'s completion, and the buildings reappeared once participants reached the start.

Scenarios were organised into four tracks, each corresponding to a different *Interface* condition. Each track involved six loops, one loop per *Scenario* and *AV-yielding state*. The order of *Scenarios* within a track was randomised. All AVs in a given track shared the same *Interface* condition, and the sequence of *Interface* conditions was counterbalanced using a Latin square design. We collected the following data:

- **Post Scenario Questionnaire:** After each *Scenario*, participants completed the NASA-TLX [21] to measure the workload experienced. Cyclists should know whether the AV is aware of them and understand its intentions [4, 6, 14], so they answered five-point Likert scale questions (Strongly disagree - Strongly agree): "I was confident the AV was *aware* of me" and "I was confident in the AV's *intentions*";
- **Post Track Questionnaire:** Answered after each group of scenarios to give an overall perspective on the *Interface*. Participants responded to the *Perceived Performance* and *Anxiety* dimensions of the *Car Technology Acceptance Model (CTAM)* [34], adapted for cyclists. They answered five-point Likert scale questions: "I *felt safe* when using the interface" and "I *trusted* the interface's messages." Participants also completed the *User Experience Questionnaire - Short Version (UEQ-S)* [40] to assess *Interface* usability;
- **Perceived Usefulness of Displays:** As part of the *Post-Track Questionnaire*, participants ranked the usefulness of cues for each *Interface* condition. These were: AR displays, eHMI, smartwatch vibrations, audio, and AV driving behaviour, into three categories: Not Useful, Somewhat Useful, and Very Useful;
- **Cycling Behaviour:** Logged by the simulator every 0.5 seconds. This included *Cycling Speed*, unity camera (player) movement in metres per second, and *Shoulder Checking*: '1' if the unity camera, which was on the participant's head, was rotated more than 90 degrees (determined through six pilot tests), and '0' otherwise. The simulator also logged any *Collisions* between the cyclist and AV;
- **Qualitative Data:** A short pre-study interview gathered participant comments on their experience riding in their city. Post-study semi-structured interviews provided additional context to the findings. Participants discussed and ranked each *Interface*.

3.4 Participants

We recruited participants who were either natives or residents (e.g., international students) of the respective cities through personal contacts and social media. To ensure familiarity with the city's traffic culture, natives had to have cycled in the city at least once a month over the past 18 months, while residents had to cycle in the city multiple days a week over the same period. The experiment included 60 participants, with 20 from each city: *Stockholm* (Natives = 17, Residents = 3; Male = 10, Female = 10; Mean Age = 29.5, SD =6.6), *Glasgow* (Natives = 18, Residents = 2; Male = 12, Female = 8; Mean Age = 30.2, SD = 5.5), and *Muscat* (Natives = 18, Residents = 2; Male = 14, Female = 6; Mean Age = 30.1, SD = 6.0). All participants were fluent in English, so no questionnaire translation was required. Participants were compensated with £10 gift vouchers to online stores in the region in their respective currencies.

3.5 Procedure

The same procedure was followed in each *City*: The participant arrived at the designated hall; they were briefed on the study and completed a demographics survey. They also participated in a short pre-study interview about their experiences cycling in the city. Next, the participant tested their comfort with the bike gear and saddle height by riding for 3 minutes without the headset; adjustments were made if needed. The experimenter calibrated the AR headset and secured the left controller around the participant's wrist to simulate the smartwatch. The participant practised riding loops in the AR environment for 7 to 15 minutes without any AVs to become familiar with the setup and equipment.

After practice, they removed the headset, and the experimenter selected the *Interface* and *Track* according to the Latin square. The experimenter explained that the participant would start encountering AVs, clarified the scenario setups, and described how the selected interface worked. The participant then wore the headset and began riding. After completing each *Scenario*, they stopped at the endpoint, where the experimenter read the *Post Scenario Questionnaire* for the participant to answer verbally. After each track, the participant removed the headset and answered the *Post Track Questionnaire* using a tablet, providing a break from the AR environment. This was repeated four times, allowing the participant to experience all the interfaces. Following the experiment, a semi-structured interview was conducted to gather additional insights. The entire study lasted approximately 80 minutes. The same experimenter conducted the study across the three cities, ensuring consistency. Relevant University Ethics Committees approved the study.

4 Results

We answer the RQ by reporting the differences in cyclist experiences riding in each city. Next, we reveal the cultural differences in cyclist use of the interfaces through the Post-Track and Post-Scenario findings and cycling behaviours. Finally, we report themes from the post-study interviews to contextualise our quantitative findings.

4.1 Riding Between Cities

We summarise pre-study interviews on participant experiences riding in each city. Deductive thematic analysis [12] was conducted to answer the research question: *"What is participants' experience riding in each city?"* Interview transcripts, auto-transcribed by Otter.ai[2] and corrected by an author, were imported into NVivo[3] for analysis. One author identified 12 unique codes from the data. Two authors then collaboratively sorted these codes into themes for each city based on their relevance to the research question. This process involved discussing and resolving any disagreements through code remapping. Themes containing two or more overlapping codes were reassessed and combined when necessary.

[2]Otter.AI transcription software: otter.ai
[3]NVivo qualitative analysis software: lumivero.com/products/nvivo/

Stockholm. Cyclists ride on segregated lanes and rarely encounter vehicles: *"I have been biking to work for ten years. I rarely see cars"* (P5). *"I only see cars in the suburbs. They move at very low speeds"* (P10). Interactions mostly happen at stationary infrastructure, e.g., intersections, so they could also interpret intentions from expressive social cues and driving behaviour: *"Most of my encounters are at intersections or something, we negotiate through gestures. I expect the car to completely stop before doing anything"* (P11); *"I ensure the driver saw me through eye contact, and wait for them to stop"* (P13). Therefore, cyclists in *Stockholm* have slower-paced interactions with drivers, allowing the exchange of diverse social cues while interpreting intention from driving behaviour before deciding on the next manoeuvre.

Glasgow. Cyclists switch between segregated lanes and mixed traffic. This increases attentiveness and anxiety: *"It's great when you're segregated, then poof! There are many scary cars"* (P35); *"It makes me nervous. You must plan and know where the bike lane will end"* (P38). Cyclists encounter vehicles in a mix of slower-paced and dynamic scenarios: *"You see them at roundabouts, but I hate it when they overtake me"* (P22); *"You see cars throughout. Sometimes, they slow down, but they mostly speed past you"* (P30). These diverse encounters require cyclists to adapt their interaction behaviour: *"I wait for the car to stop and drivers to gesture, but when I'm in moving traffic, I don't have that luxury"* (P36). In summary, *Glasgow* cyclists encounter vehicles in scenarios with different traffic control levels but can expect drivers to give them right of way at intersections or roundabouts. The exchanged cues depend on the interaction scenario.

Muscat. Has higher-speed roads without cycle lanes: *"It feels like swimming with sharks"* (P51). Constantly riding in mixed traffic means interactions happen in more dynamic scenarios: *"Cars move quickly with little patience"* (P48); *"Cars can be behind you, in front of you; you must be assertive; there's no infrastructure to do this daily"* (P51). Interactions are fast-paced, leaving little room for expressive cues such as driver hand gestures: *"It happens so fast; cars don't really stop at intersections. You have to make decisions based on little information"* (P59). To summarise, *Muscat's* cyclists share the road with motorised vehicles at higher speeds. Vehicles rarely come to a complete stop, and cyclists must have fast-paced interaction to negotiate right-of-way quickly.

4.2 Post Scenario Questionnaire

Data were not normally distributed, as demonstrated via a Shapiro-Wilk test. We conducted a three-way Aligned Rank Transform (ART) ANOVA [50] to explore the fixed effects of *Interface*, *City*, *Scenario*, and their interactions on each *Post Scenario Questionnaire* subscale. The model included a random intercept for *Participant* to account for individual variability. *Post hoc* comparisons were performed using the ART-C method [18]. Mean values are in Table 1.

We report our factors' impact on the *Overall NASA-TLX Workload*, *Confidence in AV Awareness* and *Intentions*. A better-performing *Interface* would impose a lower workload and increase the cyclist's confidence in the vehicle's awareness and intentions. We found that *Interface* did not meaningfully influence the workload. However, confidence in AV awareness was increased when receiving cues through wearable devices, e.g. a smartwatch. This reassured cyclists that the AV had seen and communicated with them directly [16]. Confidence in AV intentions was increased when communicated outside the eHMI in *Mirror* and *FullIntel*. Regarding the *Scenarios*, participants consistently found *Uncontrolled Intersection* less demanding. This effect was more pronounced in *Muscat*, where pre-study interviews revealed that vehicles do not usually stop at intersections. Interestingly, *Stockholm* cyclists did not trust interfaces enough to be confident in AV awareness or intentions compared to those from other cities.

4.2.1 *Overall NASA-TLX Workload.*

There was no significant effect of *Interface* ($F(3, 627) = 2.1$, $P = .1$), or *City* ($F(2, 57) = 1.26$, $P = .29$). However, we found a significant effect of *Scenario* ($F(2, 627) = 30.98$, $P < .001$; $\eta^2 = 0.09$). Comparing *Scenarios* showed *Uncontrolled Intersection* required a significantly lower workload than the others ($P < .0001$ for all), and *Lane Merging* caused a significantly lower workload than *Bottleneck* ($P = .0001$). This could be due to the narrow lane and an AV driving toward the cyclist; participants had to make quick decisions in a tight space.

Interactions. We did not find interactions between *Interface* and *City* ($F(6, 627) = 1.78$, $P = .1$) or *Interface* and *Scenario* ($F(6, 627) = 1.6$, $P = .15$). However, there was a significant interaction between *Scenario* and *City* ($F(4, 627) = 2.38$, $P = .05$; $\eta^2 = .01$), with no interaction between the three variables ($F(12, 627) = 0.84$, $P = .6$). Comparing *Scenario* workloads in each *City* showed that *Uncontrolled Intersection* caused a significantly lower workload than *Bottleneck* in all *Cities* ($P < .01$ for all). However, in *Muscat*, *Uncontrolled Intersection* also caused a significantly lower workload than *Lane Merging* ($P = .0004$), possibly because cyclists are not accustomed to vehicles stopping at intersections, as discussed in the pre-study interviews. Hence, *Uncontrolled Intersection* was straightforward for them to navigate.

4.2.2 *Confidence in AV Awareness.*

There were significant effects of *Interface* ($F(3, 627) = 8.68$, $P < .001$; $\eta^2 = 0.04$), *City* ($F(2, 57) = 4.72$, $P = .013$; $\eta^2 = 0.14$), and *Scenario* ($F(2, 627) = 9.49$, $P < .001$; $\eta^2 = 0.03$). Between *Interfaces*, *eHMI* was significantly less effective than *FullIntel* ($P < .0001$), *Locator* ($P = .0046$), and *Mirror* ($P = .0047$). Therefore, receiving cues through wearables (e.g., smartwatches), indicating that the AV had seen them and presented messages to them, increased confidence in awareness. For *Scenarios*, participants were least confident with the vehicle out of view during *Lane Merging* ($P = .0005$ for all). Among the *Cities*, participants from *Stockholm* were significantly less confident than the others ($P = .02$ for all).

Interactions. There was a significant interaction between *Interface* and *City* ($F(6, 627) = 3.00$, $P = .007$; $\eta^2 = 0.03$). However, there were no interactions between *Interface* and *Scenario* ($F(6, 627) = 1.12$, $P = .349$), *Scenario* and *City* ($F(4, 627) = 1.52$, $P = .194$), or all three factors ($F(12, 627) = 1.17$, $P = .3$). Comparing *Interfaces* within each *City* revealed: In *Glasgow* and *Muscat*, *eHMI* caused significantly lower confidence than *FullIntel* and *Locator* ($P < .005$ for all). Additionally, *FullIntel* in *Glasgow* was more effective than *eHMI* and *Mirror* in *Stockholm* ($P = .003$ for both). Therefore, AV location cues improved cyclist confidence in AV awareness, particularly in *Glasgow* and *Muscat*. In contrast, communicating AV

	Interface				City			Scenario		
	FullIntel	*Locator*	*Mirror*	*eHMI*	*Stockholm*	*Glasgow*	*Muscat*	*Intersection*	*Bottleneck*	Lane Merging
Workload	8.04 ± 3.32	8.18 ± 3.15	8.36 ± 3.29	8.58 ± 3.41	8.95 ± 2.93	7.67 ± 2.55	8.24 ± 4.09	7.59 ± 2.95	9.02 ± 3.48	8.25 ± 3.39
Awareness Confidence	3.91 ± 1.13	3.73 ± 1.17	3.76 ± 1.12	3.51 ± 1.12	3.29 ± 1.05	3.97 ± 0.96	3.92 ± 1.26	3.82 ± 1.06	3.82 ± 1.11	3.54 ± 1.23
Intent Confidence	4.31 ± 0.88	4.02 ± 1.13	4.13 ± 1.05	3.78 ± 1.23	3.85 ± 1.01	4.10 ± 1.10	4.24 ± 1.15	4.34 ± 0.88	4.10 ± 1.04	3.75 ± 1.26

Table 1: Mean ± Standard Deviation values of the Post-Scenario Questionnaire subscales per Interface, City and Scenario.

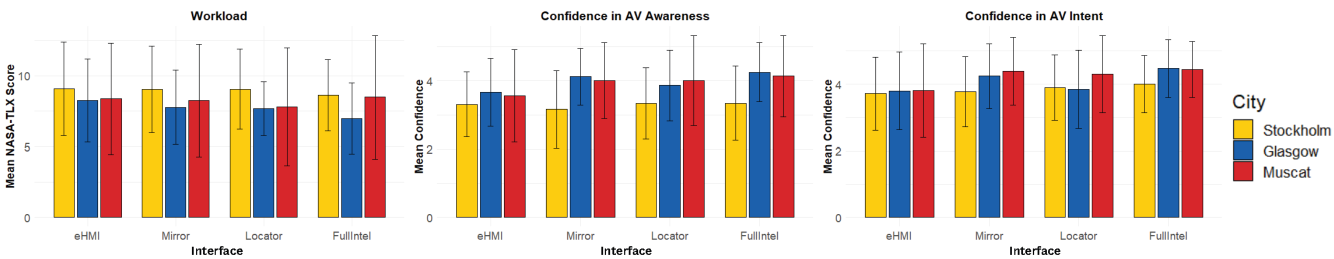

Figure 5: Bar charts showing means and standard deviations of Post-Scenario Questionnaire scores for each interface in each city. The left bar chart shows the workload imposed by interfaces in each City: Stockholm cyclists reported a higher workload for all interfaces. The middle bar chart shows participant confidence in AV awareness for each interface: Stockholm cyclists were least confident throughout. The right bar chart shows participant confidence in AV intentions for each interface: Muscat cyclists were the most confident in AV intentions throughout the interface conditions.

intentions without providing location information was particularly ineffective in *Stockholm*.

4.2.3 ***Confidence in AV Intent.*** We found a significant effect of *Interface* (F(3, 627) = 17.14, $P < .001$; $\eta^2 = 0.08$), but no effect of *City* (F(2, 57) = 2.83, $P = .067$). There was a significant effect of *Scenario* (F(2, 627) = 38.66, $P < .001$; $\eta^2 = 0.11$). Among *Interfaces*, *eHMI* was significantly less effective than *FullIntel* ($P < .0001$), *Locator* ($P = .001$), and *Mirror* ($P < .0001$). Similarly, *Locator* resulted in lower confidence than *FullIntel* ($P = .005$). Therefore, presenting redundant AV intention cues beyond the *eHMI* improved confidence. For *Scenarios*, participants felt most confident navigating *Uncontrolled Intersection* ($P < .0001$ for all); traffic control may have made AVs more predictable. They were also significantly more confident in *Bottleneck* than *Lane Merging* ($P = .003$), with the eHMI clearly in view.

Interactions. There were significant interactions between *Interface* and *City* (F(6, 627) = 3.10, $P = .005$; $\eta^2 = 0.09$), and *Interface* and *Scenario* (F(6, 627) = 8.12, $P < .001$; $\eta^2 = 0.07$). There were no interactions between *Scenario* and *City* (F(4, 627) = 0.28, $P = .894$) or the three variables (F(12, 627) = 1.16, $P = .312$). In *Glasgow*, *FullIntel* was significantly more effective than *Locator* and *eHMI* ($P < .0001$ for both). In *Muscat*, *eHMI* was less effective than all ($P < .0001$ for all). *eHMI* in *Stockholm* resulted in significantly lower confidence than *FullIntel* in *Glasgow*, and *FullIntel* and *Mirror* in *Muscat* ($P < .005$ for all). Therefore, in *Muscat*, any interface helped, but *Glasgow's* participants preferred AV intentions beyond the eHMI.

Comparing *Interfaces* within each *Scenario* showed that *eHMI* in *Lane Merging* was significantly less effective than all other conditions ($P < .001$ for all). *Locator* during *Lane Merging* was significantly less effective than *Locator* at *Uncontrolled Intersection* ($P = .0037$), and *FullIntel* at *Uncontrolled Intersection* ($P < .0001$) and *Bottleneck* ($P = .007$); presenting location cues without intentions when the AV was out of view reduced participant confidence. Moreover, *Mirror* at *Bottleneck* was significantly less effective than *FullIntel* at *Uncontrolled Intersection* ($P = .014$); reiterating intentions was ineffective with the eHMI clearly in view.

4.3 Post Track Questionnaire

Data were not normally distributed, as assessed via a Shapiro-Wilk test. We conducted an ART two-way ANOVA to explore the fixed effects of *Interface*, *City* and their interactions on each *Post Track Questionnaire* subscale. The model included a random intercept for *Participant* to account for individual variability. *Post hoc* comparisons were performed using the ART-C method. Mean values are in Table 2.

A better-performing *Interface* would lower anxiety and raise scores for the other subscales. Cyclists consistently rated *FullIntel* as the best-performing *Interface*; *eHMI* was perceived as the most cumbersome. Participants from *Muscat* were the most receptive to AV-cyclist interfaces. They trusted displays more and found the interaction more usable than cyclists from other cities, regardless of the *Interface* used. This is likely due to their experience with fast-paced, dynamic interactions. *FullIntel* significantly reduced anxiety in *Glasgow*, addressing a key concern raised in pre-study interviews.

Perceived Cycling Performance. We found significant effects of *Interface* (F(3, 171) = 22.13, $P < .001$; $\eta^2 = 0.28$) and *City* (F(2, 57) = 3.67, $P = .032$; $\eta^2 = 0.11$). They had no interaction (F(6, 171) = 1.22, $P = .3$). Among *Interfaces*, *eHMI* produced significantly worse performance than all others ($P < .0001$ for all), and *FullIntel* produced significantly better cycling than *Locator* ($P = .0165$) and *Mirror* ($P = .003$). Therefore, additional cues alongside the eHMI improved cycling performance, especially when receiving a combination of

	Interface				City		
	FullIntel	*Locator*	*Mirror*	*eHMI*	*Stockholm*	*Glasgow*	*Muscat*
Cycling Performance	4.17 ± 0.84	3.95 ± 0.69	3.86 ± 0.83	3.31 ± 0.86	3.68 ± 0.68	3.78 ±0.83	4.01 ± 1.03
Anxiety	2.35 ± 0.55	2.48 ± 0.52	2.51 ± 0.60	2.69 ± 0.67	2.52 ± 0.54	2.51 ± 0.64	2.50 ± 0.61
Trust	4.08 ± 0.87	3.92 ± 0.85	3.93 ± 0.92	3.77 ± 1.00	3.73 ± 0.80	3.83 ± 1.00	4.23 ± 0.89
Perceived Safety	3.94 ± 0.82	3.7 ± 0.86	3.77 ± 0.82	3.44 ± 0.79	3.48 ± 0.69	3.75 ± 0.89	3.91 ± 0.87
Usability	0.99 ± 0.65	0.90 ± 0.62	0.93 ± 0.64	0.69 ± 0.7	0.65 ± 0.55	0.79 ± 0.61	1.2 ± 0.69
Cycling Speed	1.49 ± 0.43	1.46 ± 0.38	1.50 ± 0.39	1.53 0.41	1.56 ± 0.38	1.29 ± 0.30	1.64 ± 0.43

Table 2: Mean ± Standard Deviation values of the Post-Track Questionnaire subscales per Interface and City.

AV location and intent. There were also cultural differences; participants from *Muscat* rated their performance significantly higher than those from *Stockholm* (P = .01).

Anxiety. We found a significant effect of *Interface* (F(3, 171) = 4.29, P = .006; η^2 = 0.07), but no effect of *City* (F(2, 57) = 0.0004, P = .996). There was still a significant interaction between them (F(6, 171) = 2.83, P = .012; η^2 = 0.09). Among *Interfaces*, anxiety was significantly higher in *eHMI* than *FullIntel* (P = .0005). Comparisons for the interaction effect showed that in *Glasgow*, anxiety was significantly higher in *eHMI* than *FullIntel* (P < .0001). Therefore, communicating AV location and intent, in addition to the eHMI, significantly lowered anxiety in *Glasgow*.

Trust. We found a significant effects of *Interface* (F(3, 171) = 3.89, P = .01; η^2= 0.06), and *City* (F(2, 57) = 3.48, P = .037; η^2 = 0.11). There was no interaction (F(6, 171) = 1.85, P = .093). Among *Interfaces*, participants trusted *FullIntel* significantly more than *eHMI* (P = .0009); presenting redundant AV intent cues increased trust, but cyclists needed the AV's location to validate them. Between *Cities*, participants in *Muscat* trusted interfaces more than those in *Stockholm* (P = .017).

Perceived Safety. We found a significant effect of *Interface* (F(3, 171) = 10.97, P < .001; η^2 = 0.01), but no effect of *City* (F(2, 57) = 2.44, P = .1), and no interaction (F(6, 171) = 1.05, P = .39). Participants felt significantly less safe using *eHMI* than all others: *FullIntel* (P < .0001), *Locator* (P = .005) and *Mirror* (P = .0003). Therefore, placing cues exclusively on the AV hindered perceived safety.

Usability. We found significant effects of *Interface* (F(3, 171) = 7.08, P < .001; η^2 = 0.11), and *City* (F(2, 57) = 7.15, P = .002; η^2 = 0.20). There was no interaction (F(6, 171) = 1.48, P = .187). Between *Interfaces*, *eHMI* was significantly less usable than *FullIntel* (P < .0001), *Locator* (P = .0213), and *Mirror* (P = .004). Between *Cities*, participants in *Muscat* rated the usability significantly higher than those in *Glasgow* (P = .0251) and *Stockholm* (P = .001).

4.4 Perceived Usefulness of Displays

The frequency of each display within each usefulness category is presented in Figure 7. Data from the baseline *eHMI* condition were excluded as there were no additional displays. To explore the relationship between *Display* and *Perceived Usefulness*, a Chi-square test of independence was performed for each *Interface* and *City*. *Post hoc* tests were conducted using Chi-square tests of independence with Bonferroni correction. Results are shown in Table 3.

Between *Interfaces*, participants consistently found *AR* displays to be the most useful. They also rated *Vibration* as more useful than *Audio* cues throughout. *eHMI* and *Driving Behaviour* were important for inferring AV intentions in *Locator*. However, these were less valuable when AV intentions were beyond the eHMI in *FullIntel* and *Mirror*. Between *Cities*, participants from *Muscat* primarily relied on explicit visual cues, such as *AR* displays and *eHMI*, to make decisions. In contrast, participants from *Glasgow* and *Stockholm* also placed importance on AV *Driving Behaviour*.

4.5 Cycling Behaviour

We report the cultural differences in *Cycle Speed* and *Shoulder Checking* for each *Interface*. The number of collisions per *Interface*, *City*, and *Scenario* is visualised in Figure 8. Collisions predominantly occurred in dynamic scenarios with no traffic control, especially during *Lane Merging*. They were more frequent when AV intentions were not communicated outside the eHMI, i.e. in the *Locator* and *eHMI* conditions. To summarise cycling behaviour, participants were slower in *Locator*. They reduced their speed and took less frequent but longer shoulder checks to infer AV intentions through its eHMI or driving behaviour. *Glasgow* participants were notably slower and conducted more shoulder checks, which aligns with our findings on their higher anxiety. This cautious cycling resulted in the fewest collisions among cities.

Cycling Speed. We conducted the same analysis as the *Post Track Questionnaire*, investigating the impact of *Interface*, *City* and their interactions on *Speed*. We found significant effects of *Interface* (F(3, 658) = 3.77, P = .011; η^2 = 0.02), and *City* (F(2, 59) = 10.97, P < .001; η^2 = 0.27), but no interaction (F(6, 658) = 0.48,P = .822). Participants were significantly faster around *eHMI* than *Locator* (P = .001). *Glasgow* cyclists were significantly slower than the rest (P > .001 for all).

Shoulder Checks. Figure 8 shows the number of shoulder checks per *Interface* in each *City* and *Scenario*. Chi-square tests of independence were performed to investigate the likelihood of *Shoulder Checks* in different *Interface*, *City*, and *Scenario* settings. *Post hocs* were Chi-Square tests of independence with a Bonferroni correction.

We found a significant association between *Interface* and *Shoulder Check* (χ^2(3, 1045) = 24.6, P < .001). Participants were less likely to shoulder check in *FullIntel* than *Mirror* (P = .007) and *eHMI* (P < .0001). Shoulder checks were also unlikely around *Locator* than *eHMI* (P < .0001). There was a significant association

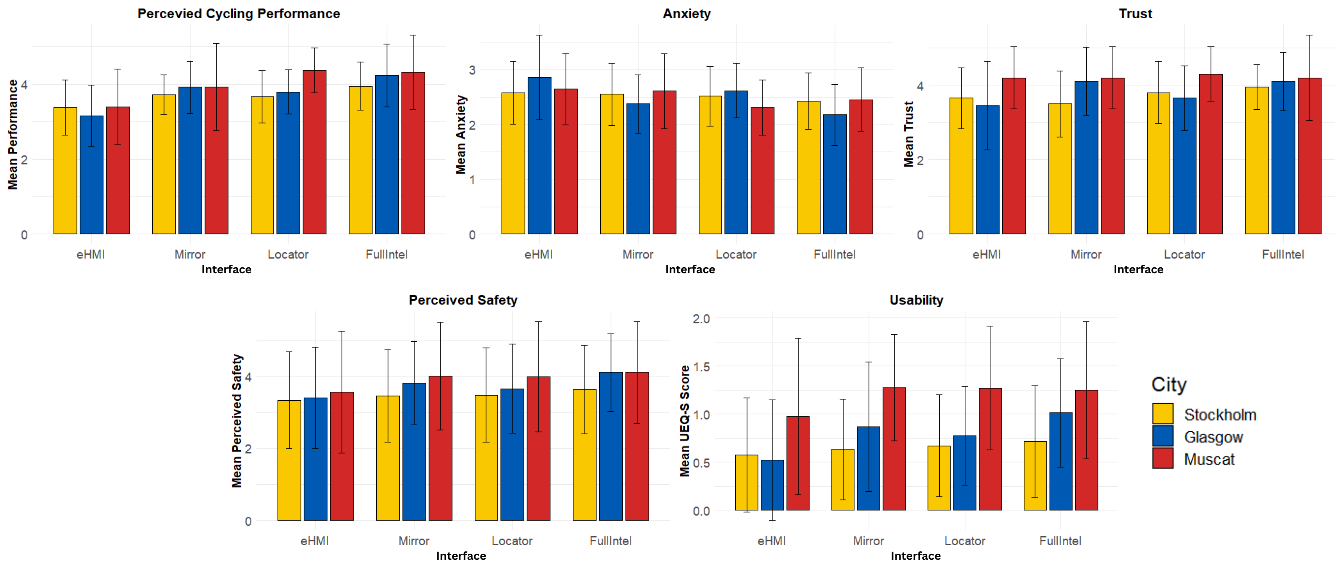

Figure 6: Bar charts showing means and standard deviations of Post-Track Questionnaire scores for each interface in each city. The bar charts from left to right illustrate Perceived Cycling Performance, Anxiety, Trust, Perceived Safety and Usability. Muscat cyclists were generally the most trusting of the interfaces. They also found them the most usable. Stockholm cyclists saw the least value in the interfaces.

	Chi-Square (χ^2)	P-Value	Significant Post Hocs	Post Hoc P
Interface				
FullIntel	$\chi^2(8, 295) = 94.08$	$P < .001$	AR > All Displays	$P < .0001$
			Audio < Driving Behaviour, eHMI	$P < .01$
Locator	$\chi^2(8, 288) = 57.294$	$P < .001$	AR > Driving Behaviour, Audio, Vibration	$P < .0005$
			Audio < eHMI	$P < .005$
Mirror	$\chi^2(8, 293) = 71.12$	$P < .001$	AR > All Displays	$P < .001$
			Audio < eHMI	$P = .0001$
City				
Stockholm	$\chi^2(8, 295) = 95.621$	$P < .001$	AR > All Displays	$P < .001$
			Audio < All Displays	$P < .05$
Glasgow	$\chi^2(8, 294) = 102.83$	$P < .001$	AR > All Displays	$P < .0001$
			Audio < All Displays	$P < .0001$
Muscat	$\chi^2(8, 285) = 48.198$	$P < .001$	AR > Driving Behaviour, Audio, Vibration	$P < .0005$
			eHMI > Driving Behaviour, Vibration	$P < .01$

Table 3: Perceived Display Usefulness for each *Interface* and *City*. The Significant Post Hocs column shows which displays were useful compared to others; > means more useful, and < means less useful.

between *City* and *Shoulder Check* ($\chi^2(2, 1045) = 39.3$, $P < .001$). Participants from *Glasgow* were more likely to shoulder check than the rest ($P < .0001$ for both). Finally, we found a significant association between *Scenario* and *Shoulder Check* ($\chi^2(2, 1045) = 125.6$, $P < .001$). Participants were most likely to shoulder check during *Lane Merging* ($P < .0001$ for all).

4.6 Post-Study Interviews

We identified themes based on post-study interviews using an inductive, data-driven thematic analysis [12]. The interview transcripts were auto-transcribed by Otter.ai and corrected by an author. Transcripts were then imported into NVivo for further analysis. One author initially extracted 15 unique codes from the data. Two authors then sorted these into three themes based on their similarities. This process was iterative, involving discussions to resolve disagreements and remapping codes until a consensus was reached.

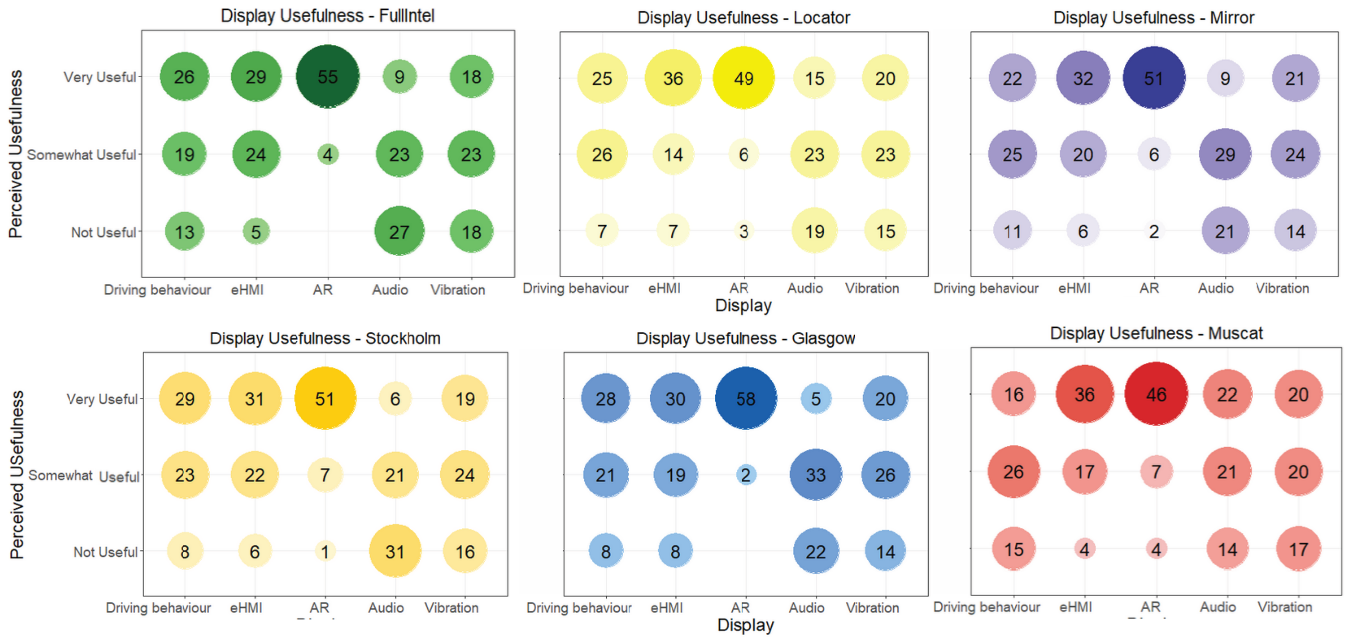

Figure 7: Dot plots showing the frequency of each Display within each Usefulness Category for each Interface and City. The larger dots illustrate a larger frequency. The top three dot plots show the frequencies for each interface, and the bottom three for each city.

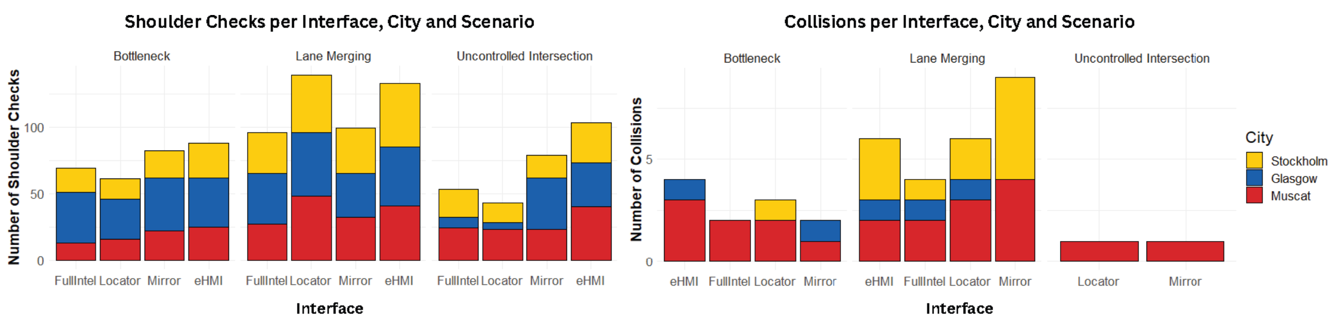

Figure 8: Barcharts illustrating the frequency of shoulder checks and collisions when navigating each Scenario, using each Interface in each City. Left: the number of Shoulder Checks conducted in each scenario per city and interface. Right: the number of Collisions in each scenario per city and interface.

Themes containing two or more overlapping codes were reassessed and combined when necessary. Figure 9 visualises the participant *Interface* rankings across the three cities. Participants in all cities ranked *FullIntel* as the best and *eHMI* as the worst. The extracted themes are as follows:

Theme 1: AR Simulator Immersion. Participants in all cities were immersed in the AR simulator: *"It just felt so real!"* (P18); *"I totally forgot the car was virtual; didn't expect myself to steer away from that car, but it felt too real."* (P31); *"Wow! I cycled like I do every day. I totally forgot these cars aren't real"* (P45); *"The threat felt completely real. I felt like I needed to avoid crashing, just as normal"* (P55). This suggests that the AR simulator was an effective way of testing cycling around AVs in the scenarios, and such simulators could be useful in HCI cycling research. Cyclists reported that they cycled as they normally do, also suggesting that AR simulators are effective in measuring real cycling behaviours.

Theme 2: Non-visual Interface cues. Participants generally appreciated non-visual cues: *"It's good that not everything is visual; it helps if I miss anything I can always rely on other signals"* (P19). Audio required careful interpretation to understand and differentiate: *"Spatial audio needs more focus, you need to carefully deduce where it's pointing"* (P30); *"I think with all the visual displays, I don't really have the headspace to process different sounds for vehicle intentions"* (P40). Vibration cues effectively communicated AV proximity and enhanced confidence in AV awareness: *"Vibration messages felt more personal to me, like the car was talking to me directly"* (P55).

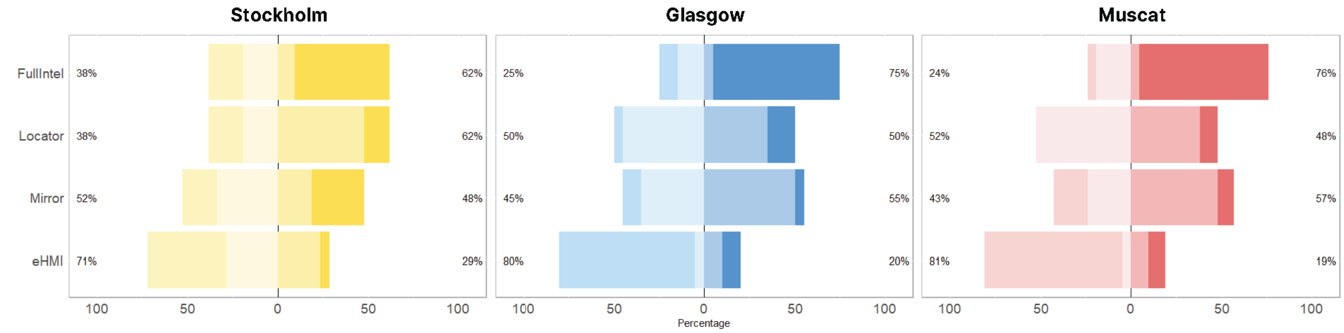

Figure 9: Participant rankings of the interfaces in each city. Less pronounced colours toward the left represent lower rankings, while brighter colours on the right are higher rankings. The left graph shows the rankings of Stockholm participants, the middle for Glasgow and the right for Muscat.

Theme 3: AR cues for AV location. Participants across cultures had varying preferences regarding the design of AR location cues. Text was universally appreciated: *"Text is great. It's unambiguous; if an arrow was used, I wouldn't be so sure if I should turn or the car"* (P10). However, opinions diverged on whether text should appear as a Head-Up Display (HUD) or a traffic sign. Some preferred a HUD for quicker recognition: *"I had to look for the sign, this would be confusing when there are many traffic signs on the road, a HUD would be faster to find"* (P35). Others favoured the traffic sign format: *"I like the traffic sign. It reminds me of signs for traffic lights or speed bumps. You can see them, like how I can sometimes see the car, but you get extra warning that there is a car approaching"* (P48). An adaptive approach displaying a HUD when AV is out of the cyclist's field of view, and a traffic sign otherwise could be beneficial.

5 Limitations and Future Work

While we explored cultural differences in interface use, our sample was limited to three cities. We maximised cultural diversity by selecting cities in different countries with varying levels of cyclist segregation from motorised vehicles. Our results show that future research should extend to more regions to enhance interface inclusivity. Additionally, we focused on three messages communicated to cyclists: AV intentions, location, or both, but other cultures might require different information. Future work could involve cross-cultural design sessions to identify cultural variations.

Participants experienced the interfaces implemented in the AR simulator. Even though the simulator enabled real cycling in physical space, and participants described it as immersive, we used the headset controller's haptics rather than a real smartwatch and spatial audio from the headset rather than a bike helmet. This maintained consistency and synchrony between displays, especially with the eHMI projected in the simulator. AVs were also projected by the headset. This may have influenced our results, and future work could attempt a Wizard of Oz approach [39], using remotely controlled vibrations from a smartwatch, spatial audio, and AR projections alongside a real vehicle with a hidden human driver [4, 29]. However, this may be challenging to synchronise across participants and cities. It will also prohibit any testing with non-yielding AVs to maintain participant safety [4]. We conducted our study indoors to control external factors, including weather, and maintain consistency between trials in each city. Future work should replicate the study outdoors, for example, by choosing specific seasons where differences in weather between cities are minimal. Finally, our study addressed one-to-one AV-cyclist encounters as a first step. While real-world interactions may involve multiple cyclists and AVs, investigating cultural differences in simpler encounters was essential to understanding perceptions of distinct AV positions and traffic control levels before scaling up to more complex many-to-many scenarios.

6 Discussion

Our research question was: "Considering holistic interfaces communicating *AV location, AV intentions or a combination of both*, how do cyclist perceptions and behaviours toward these interfaces differ across cultures and traffic scenarios?" Cyclists from all cities preferred *FullIntel*, communicating AV location and intention. This provided an adaptable solution; it communicated comprehensive information, and cyclists used different design aspects according to their traffic culture norms. Regarding the scenarios, our findings coincide with prior work that stationary infrastructure, e.g. *Uncontrolled Intersection*, is less demanding than dynamic scenarios, e.g. *Lane Merging* [4, 24]. *Lane Merging* is a challenging scenario [4], as the AV is out of the cyclist's view [24], making displays on the vehicle difficult to see. In line with earlier work, the *eHMI* alone did not suffice [2], but extending it with any holistic interface improved confidence and safety.

6.1 *Stockholm*: Interface Signals Must Be Validated With Driving Behaviour

Stockholm's participants were the most conservative toward interfaces. Pre-study interviews showed they were used to slower-paced interactions at stationary infrastructure by exchanging social cues with drivers before making decisions. They also typically waited until a vehicle completely stopped before proceeding. This was mirrored in our experiment. Participants expected expressive signals compared to binary AV intentions and for the AV to always prioritise them. This aligns with previous work where cyclists accustomed to highly segregated cycle lanes anticipated right-of-way [13], and led to lower interaction usability scores than participants from other cities. *Stockholm* cyclists also reported lower perceived cycling performance, likely due to our study's dynamic scenarios, such as *Lane Merging* and *Bottleneck*, which required them to adjust to interactions beyond the stationary infrastructure they are used to. In these situations, binary AV intention signals may be needed to make faster decisions [4]. However, *Stockholm* cyclists did not trust these on wearable devices; they also wanted AV location cues to locate the vehicle and verify any intention signals with the *eHMI* and driving behaviour. *Stockholm* participants remained sceptical about AV awareness, regardless of the interface, perhaps because they were used to regular eye contact with drivers. In contrast, prior work with UK cyclists found that changing *eHMI* states, i.e., yielding to not yielding, implicitly conveyed AV awareness [4].

These insights highlight compelling directions for future research. First, AV driving behaviours were frequently interpreted as implicit intention signals in our study, but these may need to be more pronounced for better clarity. In *Stockholm*, AVs might require exaggerated braking and driving behaviours to facilitate quicker validation of interface signals, especially in dynamic scenarios where the AV does not completely stop. Second, the binary nature of AV intention signals is minimal compared to the more expressive negotiations *Stockholm* cyclists typically have. More detailed messages could improve their confidence in AV intentions and awareness. Holistic interfaces can support more detailed messages [5] with minimal adjustments to the current setup. For example, AR glasses could display a HUD showing a pair of eyes when the AV detects the cyclist and uses animated hand gestures as intention cues, complementing the *eHMI* rather than redundantly displaying its signals. This could improve global adoption because AR displays may be customised by their owner [42], whereas changing physical *eHMI*s on vehicles would be more challenging, especially given the success of binary signals in other cities.

6.2 *Glasgow*: Adaptive Interaction Behaviours; Location and Intention Signals Valuable at Different Points

Glasgow's participants regularly navigate stationary infrastructure and dynamic scenarios, leading to more cautious cycling. They performed more shoulder checks and cycled slower, giving them time to process and validate signals. This resulted in the fewest collisions among the cities. More cautious riding may be attributed to higher anxiety reported in pre-study interviews and post-track questionnaire scores. This aligns with previous research showing that transitioning from segregated cycle lanes to mixed traffic is one of the most challenging scenarios for cyclists [2, 3]. *FullIntel* proved particularly effective for *Glasgow* cyclists, significantly lowering their anxiety. Participants equally valued AV location and intention cues to adapt their interaction behaviours based on the scenario, much like their current interactions [2]. For location cues, like *Stockholm*'s participants, *Glasgow* cyclists relied heavily on observing driving behaviour. Location cues were beneficial for identifying AVs and validating signals at stationary infrastructure like intersections, where clear changes in AV driving behaviour are expected. Prior work analysing UK cyclist gaze behaviours in real traffic [2] aligns with our findings; cyclists interpreted social cues from drivers at intersections and then looked at the vehicle's bumper to verify the driver's message through braking behaviours. Therefore, *FullIntel* aligned with the more reflexive behaviours of *Glasgow* cyclists.

Intention signals were particularly useful in scenarios requiring faster decision-making, like *Lane Merging* [4]. Cyclists trusted interfaces more than their *Stockholm* counterparts, leading to greater confidence in AV intentions when these signals reiterated the *eHMI* through wearables. This contributed to a more pleasant experience during *Lane Merging*, which previous studies showed to be demanding due to the need to frequently shoulder check [24]. Unlike in *Stockholm*, providing personalised location cues, such as changing vibration intensity based on AV proximity, increased *Glasgow* and *Muscat* cyclists' confidence in AV awareness. This gave the sensation that the AV had detected and communicated with them directly. This is similar to prior work with holistic interfaces [5], suggesting that communication through wearable devices reduces ambiguity about with whom the AV is communicating. Ultimately, we found that *FullIntel* is suitable for *Glasgow* cyclists, but future work may consider adjusting AR location cues to address anxiety. Our qualitative results suggested that traffic signs may be more difficult to find, so these could be used for *Bottleneck* or *Uncontrolled Intersection* where the AV is in view, but additional visual location cues may be needed. For *Lane Merging*, AR glasses could display a HUD that is easy to spot so cyclists can quickly process *FullIntel*'s signals.

6.3 *Muscat*: Fast Paced Interaction Made Possible with AV Intention Signals

Muscat's participants were the most receptive to interfaces, likely due to their experience navigating fast-paced interactions without the benefit of cycle lanes. In pre-study interviews, they noted that vehicles rarely stop, even at intersections. Our findings confirmed this: *Muscat* cyclists found *Uncontrolled Intersection* significantly easier than those from other cities; vehicles in the study came to a complete stop. This may have contributed to *Muscat* cyclists' higher perceived cycling performance scores. Participants mentioned they are rarely prioritised on the road and often negotiate right-of-way assertively rather than relying on expressive social cues. This directly contrasts with *Stockholm* participants and emphasises the impact of segregated cycle lanes on cycling behaviours [38]. *Muscat* cyclists' assertive riding style was reflected in their cycling behaviour during the study, where they displayed more aggressive and sometimes reckless cycling, leading to the highest collision rate among all cities. This aligns with previous research showing that countries with emerging cycling infrastructure, especially in Asia, tend to exhibit more reckless riding patterns [43]. Unlike other participants, *Muscat* cyclists did not heavily rely on AV driving behaviour, likely due to vehicle braking being harder to differentiate on faster roads [2]. Their main focus was on quickly interpreting AV intentions. They placed significantly more trust in interfaces, and minimal *eHMI* cues were sufficient for them to make decisions. This prompted higher interface usability scores than *Glasgow* and *Stockholm* participants. Perhaps because the *eHMI* already provides more affirming information than their real-world interactions, and this led them to prioritise AV intention signals over location cues. Interestingly, location cues also improved *Muscat* cyclists' confidence in AV intentions because they were used differently than in other cities; cyclists located the AV faster to understand its *eHMI* rather than validating intention signals with driving behaviour.

This raises important questions about how AVs should behave in *Muscat*. If AVs consistently yield while cyclists continue encountering assertive human drivers, AVs may be seen as out of place and confusing. Cyclists may take advantage of their predictable yielding behaviour, particularly in dynamic situations with ambiguous right-of-way. This could result in negative experiences for AV passengers, reducing adoption. Conversely, making AVs more assertive could create unsafe conditions for cyclists, who already display more reckless riding. *Muscat* remains unexplored in the AutomotiveUI domain, and research must understand behavioural changes in AV deployment in the city; this should be longitudinal to see whether cyclists adapt to less assertive AVs over time or if AVs should instead employ more assertive driving. This is critical for safe global AV adoption. *Muscat* cyclists may be safer if AV intentions were only displayed after cyclists indicate their intentions; this may be a departure from requirements in other cities, but it would promote clear two-way communication between road users. For example, cyclists may gesture, use on-bike direction indicators [5] or novel input techniques [44] to send AVs their intentions; only then will the *eHMI* show its state, and *Mirror* or *FullIntel* will display this. This would make *Muscat*'s interactions more expressive and may reduce collisions.

6.4 Lessons Learned and Design Guidelines for Culturally-Inclusive Holistic Interfaces

Given the complexity and scale of our study, we contribute Lessons Learned (LL) from our experience that may help researchers design such experiments and interpret their findings. We also use our findings to contribute Design Guidelines (DG) for culturally inclusive AV-cyclist interfaces.

LL: Researchers should classify any potential influencing factors between cities. We identified physical infrastructure and vehicle driving behaviours as factors influencing cyclist perception and behaviours toward interfaces. Future work should identify other influencing factors, e.g., through pre-study interviews. Cross-cultural researchers should use these to classify their study locations, e.g., high versus low segregation from motorised vehicles. This would help them contextualise and interpret their findings and deduce trends in the data. In our study, we classified cities according to the level of cyclist segregation from vehicles, allowing us to identify key trends. For example, greater segregation and slower vehicle driving resulted in cyclists putting more value on AV location cues than intentions.

LL: Surveys are not always sufficient for cross-cultural interface evaluation. Cross-cultural studies are often conducted using online surveys [3, 13, 38], which are valuable for reaching a large, diverse participant pool across multiple cities and countries [3, 22]. However, we encourage researchers also to consider replicable user studies. This would allow them to explore situations that require physical interaction, such as evaluating multimodal interfaces in greater depth [4, 24]. For instance, our study incorporated interfaces with haptic cues, which are challenging to evaluate through surveys alone. User studies provide a more detailed understanding of user behaviours when interacting with interfaces [9, 14]. Quantitative measures such as cycling speed, shoulder-checking, and other behavioural patterns can be logged [17], offering insights into perceptive and behavioural differences across cultures. Researchers relying on online surveys should consider our previous lesson learned regarding classifying differences between cities or cultural settings. Online surveys are accessible globally; cultural distinctions may appear more nuanced, potentially complicating data interpretation.

LL: Researchers should utilise platforms enabling study replication. Researchers should strive for identical setups across cities to clearly isolate cultural factors. We adopted a headset-only approach, which was highly convenient to transport and made replication straightforward. It projected a complete urban environment, including props such as road markings, along with AVs and interfaces, meaning no other research equipment was required. Participants in all cities experienced identical AV behaviours and interface implementations. We also established specific requirements for indoor halls, including floor colours and hall dimensions, to minimise variability. We recommend that researchers centralise their hardware and adopt a similar approach wherever feasible.

DG: We recommend designers to build culturally-inclusive interfaces around a common baseline. The LightRing *eHMI* [4] employed simple binary AV intention cues understood across cities, promoting cultural inclusivity. *eHMIs* serve as a reliable failsafe for cyclists without additional devices [5], enabling cyclists to transition between cities with minimal effort if a consistent *eHMI* is adopted. LightRing's design was highly extendable: AR road projections, smartwatch vibrations, and auditory cues were synchronised with the *eHMI* in *FullIntel* and *Mirror*. This meant a minimal learning curve across displays and fostered trust, allowing cyclists to verify redundant signals easily. Designers could take a ground-up approach and anchor their holistic interfaces around a culturally inclusive baseline. This could be an eHMI or a broader display feature, such as an extendable animation pattern.

DG: We recommend designers to consider comprehensive 'middle-ground' interfaces. FullIntel was ranked as the best interface across the three cities. It was the most comprehensive and used multimodal cues to communicate the AV's intentions and location. This allowed cyclists from all cities to use the same interface but adapt the information to their needs; for example, those in *Stockholm* preferred AV location cues, but those in *Muscat* used AV intention cues. This suggests that designers could consider developing similar comprehensive solutions that give different features equal footing. This could enhance cultural inclusivity, with users adapting them according to their local needs.

DG: We recommend that any alterations to suit local norms be on wearable devices when a baseline display is used. Some cyclists may need additional information that could overload middle-ground solutions. For instance, cyclists in *Stockholm* could benefit from explicit cues confirming AV awareness. However, such cues might overwhelm *Muscat* and *Glasgow* cyclists if integrated into *FullIntel*. Designers could tailor awareness cues for *Stockholm* cyclists without negatively impacting users in other cities by using wearable devices. Altering baseline displays like the *eHMI* may undermine cultural inclusivity and extendibility. Disrupting universally understood signals, for example, changing 'green' to mean not-yielding, would reduce inclusivity. Instead, designers can provide additional signals via wearable or bike-mounted devices for local needs. These signals would only reach cyclists equipped with such devices [5, 9, 28–30]. Notably, vibration cues effectively reassured cyclists of AV awareness, especially when indicating proximity. Designers should explore non-visual cues further, as these can communicate specific information without being redundant to *eHMI*s. This approach allows tailored support without compromising the baseline's cultural inclusivity.

7 Conclusion

We conducted the first cross-cultural AV-cyclist interaction study with participants directly using interfaces and interacting with AVs. The study was conducted in *Stockholm* (high cyclist segregation from motorised vehicles), *Glasgow* (some segregation), and *Muscat* (no segregation). Cities were in different countries to maximise any cultural differences [13]. In each city, cyclists used our AR cycling simulator to ride in real physical space in an AR urban environment with moving AVs. They navigated three traffic scenarios to test three holistic interfaces, each incorporating an eHMI (conveying AV intentions), with multimodal cues communicating either AV location, AV intentions (redundant to the eHMI), or a combination of both. We evaluated holistic interfaces because they bring a wider range of modalities, messages and display placements than individual ones such as eHMIs [5], complicating the cross-cultural design problem. Our study revealed that cyclists from all cities preferred combined AV location and intention information, but their respective cultural norms significantly influenced their usage of this information. *Stockholm* cyclists, accustomed to slower-paced, expressive interactions at stationary infrastructure, favoured AV location cues to interpret driving behaviour and validate intention signals, as they were less trusting of the interfaces alone. *Glasgow* cyclists, who navigate a mix of stationary and dynamic scenarios, valued location and intention cues equally, adjusting their reliance on each depending on the context. *Muscat*'s participants, accustomed to fast-paced interactions in mixed traffic, placed greater trust in the interfaces and prioritised binary intention cues (yielding/not-yielding) to make quick decisions without verifying driving behaviour. Our findings provide the first insights into the design of culturally-inclusive AV-cyclist interfaces. This is vital to the successful global adoption of AVs [38].

Acknowledgments

This research received funding from the University of Glasgow's Mobility Scholarship and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (#835197, ViAjeRo).

References

[1] Gov UK Active Travel. 2024. Cycling Factsheet, England: 2022. , 1 pages. https://assets.publishing.service.gov.uk/media/64e71a4f20ae890014f26d32/cycling-factsheet-2022.pdf

[2] Ammar Al-Taie, Yasmeen Abdrabou, Shaun Alexander Macdonald, Frank Pollick, and Stephen Anthony Brewster. 2023. Keep it Real: Investigating Driver-Cyclist Interaction in Real-World Traffic. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–15. https://doi.org/10.1145/3544548.3581049

[3] 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*. ACM, New York, NY, USA, 127–131. https://doi.org/10.1145/3544999.3552531

[4] Ammar Al-Taie, Graham Wilson, Euan Freeman, Frank Pollick, and Stephen 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*. ACM, New York.

[5] Ammar Al-Taie, Graham Wilson, Thomas Goodge, Frank Pollick, and Stephen Brewster. 2024. Bike to the Future: Designing Holistic Autonomous Vehicle-Cyclist Interfaces. In *15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, Stanford.

[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*. ACM, New York, NY, USA, 213–223. https://doi.org/10.1145/3580585.3607161

[7] Thomas K. Aleva, Wilbert Tabone, Dimitra Dodou, and Joost C. F. de Winter. 2024. Augmented reality for supporting the interaction between pedestrians and automated vehicles: an experimental outdoor study. *Frontiers in Robotics and AI* 11 (1 2024). https://doi.org/10.3389/frobt.2024.1324060

[8] 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 (1 2024), 100986. https://doi.org/10.1016/j.trip.2023.100986

[9] 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 (9 2023), 104043. https://doi.org/10.1016/j.apergo.2023.104043

[10] Siri Hegna Berge, Joost De Winter, and Marjan Hagenzieker. 2023. User Interfaces for Cyclists in Future Automated Traffic. *International Conference on Intelligent User Interfaces, Proceedings IUI* (3 2023), 91–94. https://doi.org/10.1145/3581754.3584140

[11] Siri Hegna Berge, Marjan Hagenzieker, Haneen Farah, and Joost de Winter. 2022. Do cyclists need HMIs in future automated traffic? An interview study. *Transportation Research Part F: Traffic Psychology and Behaviour* 84 (1 2022), 33–52. https://doi.org/10.1016/j.trf.2021.11.013

[12] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[13] Elijah Steven Chataway, Sigal Kaplan, Thomas Alexander Sick Nielsen, and Carlo Giacomo Prato. 2014. Safety perceptions and reported behavior related to cycling in mixed traffic: A comparison between Brisbane and Copenhagen. *Transportation Research Part F: Traffic Psychology and Behaviour* 23 (3 2014), 32–43. https://doi.org/10.1016/j.trf.2013.12.021

[14] Debargha Dey, Azra Habibovic, Andreas Löcken, Philipp Wintersberger, Bastian Pfleging, Andreas Riener, Marieke Martens, and Jacques Terken. 2020. Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transportation Research Interdisciplinary Perspectives* 7 (9 2020), 100174. https://doi.org/10.1016/J.TRIP.2020.100174

[15] Debargha Dey, Andrii Matviienko, Melanie Berger, Marieke Martens, Bastian Pfleging, and Jacques Terken. 2021. Communicating the intention of an automated vehicle to pedestrians: The contributions of eHMI and vehicle behavior. *IT - Information Technology* 63, 2 (6 2021), 123–141. https://doi.org/10.1515/ITIT-2020-0025/MACHINEREADABLECITATION/RIS

[16] 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*. ACM, New York, NY, USA, 274–286. https://doi.org/10.1145/3409118.3475129

[17] Debargha Dey, Francesco Walker, Marieke Martens, and Jacques Terken. 2019. Gaze Patterns in Pedestrian Interaction with Vehicles: Towards Effective Design of External Human-Machine Interfaces for Automated Vehicles • Human-centered computing ˜ User studies • Human-centered computing ˜ Field studies. (2019). https://doi.org/10.1145/3342197.3344523

[18] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. *UIST 2021 - Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology* 15, 21 (10 2021), 754–768. https://doi.org/10.1145/3472749.3474784

[19] Kevin Gildea and Ciaran Simms. 2021. Characteristics of cyclist collisions in Ireland: Analysis of a self-reported survey. *Accident Analysis & Prevention* 151 (3 2021), 105948. https://doi.org/10.1016/J.AAP.2020.105948

[20] Marjan P Hagenzieker, Sander Van Der Kint, Luuk Vissers, Ingrid N L G Van Schagen, Jonathan De Bruin, Paul Van Gent, Jacques J F Commandeur, Ã Marjan,

and P Hagenzieker. 2019. Journal of Transportation Safety & Security Interactions between cyclists and automated vehicles: Results of a photo experiment Interactions between cyclists and automated vehicles: Results of a photo experiment. *Journal of Transportation Safety & Security* 12, 1 (2019), 94–115. https://doi.org/10.1080/19439962.2019.1591556

[21] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology* 52, C (1 1988), 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[22] Sonja Haustein, Till Koglin, Thomas Alexander Sick Nielsen, and Åse Svensson. 2020. A comparison of cycling cultures in Stockholm and Copenhagen. *International Journal of Sustainable Transportation* 14, 4 (2 2020), 280–293. https://doi.org/10.1080/15568318.2018.1547463

[23] 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*. ACM, New York, NY, USA, 1–13. https://doi.org/10.1145/3411764.3445480

[24] 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*. ACM, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376884

[25] Riaz(WHO) Ishrat. 2018. Global Status Report on Road Safety 2018: Summary. *World Health Organization* 1 (2018), 20. http://apps.who.int/bookorders.

[26] Alessandra Mantuano, Silvia Bernardi, and Federico Rupi. 2017. Cyclist gaze behavior in urban space: An eye-tracking experiment on the bicycle network of Bologna. *Case Studies on Transport Policy* 5, 2 (6 2017), 408–416. https://doi.org/10.1016/J.CSTP.2016.06.001

[27] G. Markkula, R. Madigan, D. Nathanael, E. Portouli, Y. M. Lee, A. Dietrich, J. Billington, A. Schieben, and N. Merat. 2020. Defining interactions: a conceptual framework for understanding interactive behaviour in human and automated road traffic. *Theoretical Issues in Ergonomics Science* 21, 6 (11 2020), 728–752. https://doi.org/10.1080/1463922X.2020.1736686

[28] Andrii Matviienko, Swamy Ananthanarayan, Shadan Sadeghian Borojeni, Yannick Feld, Wilko Heuten, and Susanne Boll. 2018. Augmenting bicycles and helmets with multimodal warnings for children. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, New York, NY, USA, 1–13. https://doi.org/10.1145/3229434.3229479

[29] Andrii Matviienko, Swamy Ananthanarayan, Stephen Brewster, Wilko Heuten, and Susanne Boll. 2019. Comparing unimodal lane keeping cues for child cyclists. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*. ACM, New York, NY, USA, 1–11. https://doi.org/10.1145/3365610.3365632

[30] Andrii Matviienko, Jean Baptiste Durand-Pierre, Jona Cvancar, and Max Mühlhäuser. 2023. Text Me if You Can: Investigating Text Input Methods for Cyclists. *Conference on Human Factors in Computing Systems - Proceedings* (4 2023). https://doi.org/10.1145/3544549.3585734

[31] Andrii Matviienko, Florian Müller, Dominik Schön, Régis Fayard, Salar Abaspur, Yi Li, and Max Mühlhäuser. 2022. E-ScootAR: Exploring Unimodal Warnings for E-Scooter Riders in Augmented Reality. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. ACM, New York, NY, USA, 1–7. https://doi.org/10.1145/3491101.3519831

[32] Andrii Matviienko, Florian Müller, Dominik Schön, Paul Seesemann, Sebastian Günther, and Max Mühlhäuser. 2022. BikeAR: Understanding Cyclists' Crossing Decision-Making at Uncontrolled Intersections using Augmented Reality. In *CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–15. https://doi.org/10.1145/3491102.3517560

[33] Ali Mohammadi, Giulio Bianchi Piccinini, and Marco Dozza. 2024. Understanding the interaction between cyclists and motorized vehicles at unsignalized intersections: Results from a cycling simulator study. *Journal of Safety Research* 90 (9 2024), 306–318. https://doi.org/10.1016/j.jsr.2024.05.007

[34] Sebastian Osswald, Daniela Wurhofer, Sandra Trösterer, Elke Beck, and Manfred Tscheligi. 2012. Predicting information technology usage in the car. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '12*. ACM Press, New York, New York, USA, 51. https://doi.org/10.1145/2390256.2390264

[35] Hannah R.M. Pelikan. 2021. Why Autonomous Driving Is So Hard: The Social Dimension of Traffic. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, New York, NY, USA, 81–85. https://doi.org/10.1145/3434074.3447133

[36] Petr Pokorny, Belma Skender, Torkel Bjørnskau, and Marjan P. Hagenzieker. 2021. Video observation of encounters between the automated shuttles and other traffic participants along an approach to right-hand priority T-intersection. *European Transport Research Review* 13, 1 (12 2021), 59. https://doi.org/10.1186/s12544-021-00518-x

[37] 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 (4 2020), 175–190. https://doi.org/10.1016/J.TRF.2020.03.004

[38] Ana Rodríguez Palmeiro, Marjan P Hagenzieker, A Rodríguez Palmeiro, S van der Kint, MP Hagenzieker, Inlg van Schagen, and JCF de Winter. 2018. *Cyclists' expectations when encountering automated vehicles: results of an international photo-based questionnaire How should automated vehicles communicate with other road users? View project Traffic Psychology View project Cyclists' expectations when encountering automated vehicles: results of an international photo-based questionnaire*. Technical Report. 2018 pages. https://www.researchgate.net/publication/344217593

[39] Dirk Rothenbücher, Jamy Li, David Sirkin, Brian Mok, and Wendy Ju. 2015. Ghost driver. In *Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, New York, NY, USA, 44–49. https://doi.org/10.1145/2809730.2809755

[40] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence* 4, 6 (2017), 103. https://doi.org/10.9781/IJIMAI.2017.09.001

[41] Society for Automotive Engineers. 2021. SAE J3016 - Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles.

[42] Tram Thi Minh Tran, Callum Parker, and Martin Tomitsch. 2023. Scoping Out the Scalability Issues of Autonomous Vehicle-Pedestrian Interaction. *ACM International Conference Proceeding Series* (9 2023), 167–177. https://doi.org/10.1145/3580585.3607167

[43] Sergio A. Useche, Francisco Alonso, Aleksey Boyko, Polina Buyvol, Isaac Castañeda, Boris Cendales, Arturo Cervantes, Tomas Echiburu, Mireia Faus, Zuleide Feitosa, Javier Gene, Adela Gonzalez-Marin, Victor Gonzalez, Jozef Gnap, Mohd K. Ibrahim, Kira H. Janstrup, Arash Javadinejad, Irijna Makarova, Rich McIlroy, Miroslava Mikusova, Mette Møller, Sylvain Ngueuteu-Fouaka, Steve O'Hern, Mauricio Orozco-Fontalvo, Ksenia Shubenkova, Felix Siebert, Jose Soto, Amanda N. Stephens, Raquel Valle-Escolano, Yonggang Wang, Ellias Willberg, Phillip Wintersberger, Linus Zeuwts, Zarir H. Zulkipli, and Luis Montoro. 2022. Cross-culturally approaching the cycling behaviour questionnaire (CBQ): Evidence from 19 countries. *Transportation Research Part F: Traffic Psychology and Behaviour* 91 (11 2022), 386–400. https://doi.org/10.1016/j.trf.2022.10.025

[44] Jochem Verstegen and Pavlo Bazilinskyy. 2024. Slideo: Bicycle-to-Vehicle Communication to Intuitively Share Intention to Turn with Automated Vehicles. https://doi.org/10.54941/ahfe1005210

[45] 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*. ACM, New York, NY, USA, 132–136. https://doi.org/10.1145/3473682.3480280

[46] Tamara von Sawitzky, Thomas Grauschopf, and Andreas Riener. 2021. "Attention! A Door Could Open."—Introducing Awareness Messages for Cyclists to Safely Evade Potential Hazards. *Multimodal Technologies and Interaction* 6, 1 (12 2021), 3. https://doi.org/10.3390/mti6010003

[47] Tamara von Sawitzky, Thomas Grauschopf, and Andreas Riener. 2022. Hazard Notifications for Cyclists: Comparison of Awareness Message Modalities in a Mixed Reality Study. *27th International Conference on Intelligent User Interfaces* 22 (3 2022), 310–322. https://doi.org/10.1145/3490099.3511127

[48] Tamara Von Sawitzky, Philipp Wintersberger, Andreas Löcken, Anna-Katharina Frison, and Andreas Riener. 2020. Augmentation Concepts with HUDs for Cyclists to Improve Road Safety in Shared Spaces. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3383022

[49] Philipp Wintersberger, Andrii Matviienko, Andreas Schweidler, and Florian Michahelles. 2022. Development and Evaluation of a Motion-based VR Bicycle Simulator. *Proceedings of the ACM on Human-Computer Interaction* 6, MHCI (9 2022), 1–19. https://doi.org/10.1145/3546745

[50] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963