Light it Up: Evaluating Versatile Autonomous Vehicle-Cyclist External Human-Machine Interfaces

Ammar Al-Taie  
graham.wilson@glasgow.ac.uk  
Glasgow Interactive Systems Section,  
School of Computing Science,  
University of Glasgow  
United Kingdom

Graham Wilson  
eguan.freeman@glasgow.ac.uk  
Glasgow Interactive Systems Section,  
School of Computing Science,  
University of Glasgow  
United Kingdom

Euan Freeman  
stephen.brewster@glasgow.ac.uk  
Glasgow Interactive Systems Section,  
School of Computing Science,  
University of Glasgow  
United Kingdom

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

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

Figure 1: Cyclists encountering AVs with an eHMI in our two-stage eHMI evaluation: (A) in a VR cycling simulator; (B) in a real-world setting.

ABSTRACT

The social cues drivers exchange with cyclists to negotiate space-sharing will disappear as autonomous vehicles (AVs) join our roads, leading to safety concerns. External Human-Machine Interfaces (eHMIs) on vehicles can replace driver social signals, but how these should be designed to communicate with cyclists is unknown. We evaluated three eHMIs across multiple traffic scenarios in two stages. First, we compared eHMI versatility, acceptability and usability in a VR cycling simulator. Cyclists preferred colour-coded signals communicating AV intent, easily seen through quick glances. Second, we refined the interfaces based on our findings and compared them outdoors. Participants cycled around a moving car with real eHMIs. They preferred eHMIs using large surfaces on the vehicle and animations reinforcing colour changes. We conclude with novel design guidelines for versatile eHMIs based on first-hand interaction feedback. Our findings establish the factors that enable AVs to operate safely around cyclists across different traffic scenarios.

CCS CONCEPTS

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

KEYWORDS

Autonomous Vehicle-Cyclist Interaction, eHMI

ACM Reference Format:

1 INTRODUCTION

Cyclists share the road with motorised vehicles, encountering them in various traffic scenarios, such as intersections and roundabouts.
We test this hypothesis using video-recorded interactions with VUs, that were analyzed using qualitative methods. In our study, VUs were categorized into two groups: (1) pedestrians and (2) cyclists. The data was then analyzed to determine the most effective way to communicate the AV's presence to each group. Our results show that using different visual elements, such as colored lights and facial expressions, can help VUs better understand the AV's intentions and improve their safety. However, more research is needed to further refine these approaches and ensure that they are effective in real-world scenarios.
evaluated with pedestrians in VR [13] and outdoor crossings [12]. They also conducted an online survey to identify the colours and animation patterns the lights should use to communicate yielding at crossings [9]. Pedestrians ranked green as the best. The authors still advised using cyan, a neutral colour without a predetermined meaning, making messages unlikely to be misinterpreted as instructions from the vehicle. They acknowledged, however, that cyan signals may cause ambiguity and need to be learned. It is unknown whether this generalises to cyclists, and we addressed this gap by evaluating a variation of lightband catered to cyclists [3].

Berge et al. [4] reviewed cycling support systems, including AV interfaces. A minority were eHMIs, and only two (Tracker [11] and CommDisk [32]) could facilitate interactions anywhere around a vehicle. However, cyclists were not the target road users for these, and unique requirements, such as versatility, were not considered. The authors concluded that more work should go into designing and evaluating eHMIs catered for cyclists. We address this limitation by taking an iterative design approach to evaluate and refine eHMIs based on cyclists’ experience from interacting with eHMIs. Al-Taie et al. [1] used an in-the-wild approach to gather requirements for AV-cyclist interfaces (including eHMIs) based on current interactions with human drivers. First, they observed cyclist-cyclist encounters in multiple traffic scenarios. Then, they conducted a naturalistic study with cyclists wearing eye-trackers. Over 50% of encounters resulted in interaction, providing real-world evidence that AV-cyclist interaction must be facilitated. Most interactions happened when cyclists had right-of-way, and drivers should have yielded to them. The road users interacted differently between traffic scenarios, cementing versatility as a key issue separating cyclist interfaces from pedestrian ones and that evaluations of cyclist interfaces should cover multiple traffic scenarios, an approach we took in this paper.

Al-Taie et al. [3] then conducted design sessions with cyclists and AutomotiveUI experts to develop eHMIs around a real car. Each session resulted in two designs for a specific traffic scenario, e.g. a controlled intersection. The authors synthesised a taxonomy showing different eHMI features to combine overlapping ones for individual traffic scenarios and contributed the first set of versatile AV-cyclist eHMIs. These were a cyan light bar around the car using animation patterns to communicate AV intent and awareness of cyclists, a rooftop display communicating through emoticons and an on-road projection showing riders the AV’s intentions. See Section 3.2 for a detailed explanation of the eHMIs. We used these as a starting point in Stage 1 of the investigation as they address the most recent AV-cyclist eHMI requirements [1, 3, 4, 17] and are the only available designs which consider versatility [4]. However, they were never evaluated in practice; their usability, versatility and overall effectiveness are unproven. Evaluation is a critical next step in establishing their real-world suitability. Having cyclists use them in practice would reveal areas for refinement and uncover new requirements and guidelines for eHMIs. It would show, for the first time, how results with human drivers [1] generalise to interactions with AVs. It would also validate Al-Taie et al.’s [3] method for designing versatile eHMIs, opening doors to developing new ones. The designs use features, such as a cyan lightbar, that were effective with pedestrians [8]. Our evaluation would show how cyclists react to these and provide a starting point for synthesising eHMIs that function around multiple road user types.

2.1 Evaluating AV-Cyclist eHMIs

Cycling interfaces, such as augmented reality (AR) headsets [34], vibrating helmets [33] or eHMIs [8], are commonly evaluated in simulated or controlled outdoor environments. Hou et al. [18] evaluated five AV-cyclist interfaces, two of which were eHMIs, in a VR simulator. Participants cycled in the virtual world and merged lanes with an AV behind them across different interface conditions. They were asked about their confidence in performing the lane merging manoeuvre and the perceived usefulness of each interface. Shoulder checking and stopping behaviour (i.e. if the cyclist let the AV pass) were also measured. The authors found that having an AV-cyclist interface improved participant confidence and performance. eHMIs placed on specific car areas (e.g. windows or windscreen) did not perform well compared to ones using large surfaces, such as road projection, as they diverted cyclists’ attention from the road. However, this work only explored lane merging, so the versatility of their designs is unknown. Our paper widens the scope and evaluates eHMIs in five scenarios with different characteristics. A VR simulator also showed promise in examining these interfaces in a controlled setting without any practical, safety or environmental concerns, e.g. visibility issues for road projections [36]. It also helped the authors use SAE level 5 AVs (no human driver present in all traffic scenarios [31]) without any human driver and easily switch between high-fidelity interface implementations. We took this approach in the first stage of our investigation. However, we followed it up with a real-world evaluation of the eHMIs to understand the practical limitations of implementing the eHMIs and allow for more natural cycling behaviour.

Matvienko et al. [24] developed an augmented reality (AR) cycling simulator deployed on a Hololens2 to evaluate AV-cyclist interfaces that use AR. The authors only considered encounters at uncontrolled intersections, so how the interfaces performed beyond this scenario is unknown. They found that interfaces improved perceived safety and cycling performance as cyclists proceeded at the intersection with smaller gaps between them and the AV when an interface was used. The AR simulator helped trigger real cycling behaviours with participants cycling on a moving bicycle in physical space. We considered this approach; however, the field-of-view limitations of current AR headsets and the need to conduct the study in a dark indoor space motivated us to proceed with a two-stage investigation that used VR and outdoor space. A VR simulator allowed us to overcome any field-of-view and immersiveness issues in AR simulators without sacrificing any practical or environmental limitations on the eHMIs, and the outdoor study allowed participants to appreciate riding around real eHMIs mounted on a real car with greater ecological validity [36].

Some previous research has conducted outdoor evaluations of cycling interfaces; however, these are rarely conducted around moving cars. Vo et al. [33] evaluated a vibrating helmet that warned cyclists about nearby obstacles such as cars. Participants cycled on a 20m outdoor track. An experimenter controlled the helmet’s cues remotely via Bluetooth. Participants were asked to state the

---

8Hololens AR headset: microsoft.com/en-us/hololens
direction and proximity of an obstacle based on the helmet’s haptic cues. However, there were no real obstacles around the cyclist due to safety concerns; we used a real moving car in our investigation’s second stage to trigger natural responses from cyclists, allowing us also to measure cycling behaviour such as speed changes and shoulder checks. Matviienko et al. [23] conducted a two-step study evaluating cues to assist child cyclists in navigation. These were first explored in a screen/projector-based cycling simulator, followed by a test track study outdoors. This motivated us to take a similar direction, but we took an iterative design approach; our interfaces were revised based on participant feedback from the simulator evaluation before moving to the real-world study.

Stage 2 of our investigation required us to convince participants that they were riding around a driverless car to trigger natural interaction behaviour. We took Rothenbacher et al.’s [29] Ghost Driver method to hide the driver in a car seat costume and produce the illusion that the car was autonomous. This was used in Wizard of Oz studies investigating AV-pedestrian interaction. For example, Dey et al. [12] used Ghost Driver to evaluate lightband; a car (with the eHMI) approached a real, closed-off, pedestrian crossing and participants indicated their willingness to cross using a handheld slider. The eHMI helped to resolve ambiguity; participants were more willing to cross when the car used an eHMI. We took a similar approach in the second stage of our investigation. We show, for the first time, how Ghost Driver can be used across multiple traffic scenarios to evaluate and compare multiple eHMIs with cyclists directly interacting with an ‘autonomous’ vehicle (as opposed to, e.g. using a slider), allowing us to gain novel insights into how cyclists interact with different eHMIs in a real-world setting.

2.2 Summary and Research Questions

eHMIs are a promising solution to facilitate AV-cyclist interactions necessary to navigate future traffic scenarios [1–3, 5]. However, there is no thorough evaluation of these interfaces and how riders interact with them across traffic scenarios with varying levels of traffic control [8]. Existing work has instead focused on requirements gathering [1, 3] and evaluating interfaces in a single scenario [18]. Work with pedestrians has evaluated eHMIs in VR and real-world settings [8, 12] and shown them to be good methods. However, cyclists interact with vehicles in many different ways than pedestrians. In this paper, we scale up these approaches to evaluate three eHMI designs across five traffic scenarios. We answer the following research questions:

RQ1 How versatile, acceptable and usable are eHMIs in terms of cyclist perception?

RQ2 How versatile, acceptable and usable are eHMIs in terms of cycling behaviour?

3 STAGE 1: EHMI EVALUATION IN A VR CYCLING SIMULATOR

Three eHMI designs were evaluated across five traffic scenarios, allowing us, for the first time, to test the versatility of eHMIs and bring them closer to real-world use.

3.1 Participants

We recruited 20 participants (4 Female, 16 Male; Mean Age = 29, SD = 6.6) through social media advertising. Ten cycled at least once a week, two at least once a month, five multiple times a year, and three once a year or less. Two participants had cycled in VR before. All had experience of riding in Glasgow (UK), on which our simulator was based. Participants were compensated with a £10 Amazon voucher.

3.2 Apparatus

The study used a Virtual Reality (VR) cycling simulator (see Figure 2) composed of a Giant Escape 3 size medium hybrid bicycle 3 mounted on a Wahoo Kickr Snap smart trainer 4. Similar to Hou et al. [18]’s simulator, the wheel-on trainer allowed cyclists to use the bike’s back brake in the virtual environment without any alterations to the bike. A Coospo Bluetooth speed sensor 5 attached to the back wheel hub controlled speed in VR. We used a Meta Quest Pro headset 6 to display the virtual world and measure gaze behaviour (using its eye-tracker) during the study. As in Hou et al. [18]’s setup, the headset’s left controller was attached to the handlebar to translate turn angles into the virtual world according to the controller rotation. The virtual environment was developed using Unity3D 2021.3.29f1; the EasyRoads3D package 7 was used due to its realistic textures and UK-like road infrastructure assets. An iPad was placed 60cm in front of the participant to combat simulator sickness and increase immersion by simulating headwind [25]. An iPad was used to complete post-condition surveys hosted on Qualtrics 8.

3.3 Implementing the eHMIs

We evaluated Al-Taie et al.’s [3] proposed eHMI designs (see Figure 3) as a starting point. This section explains how these were implemented in VR and why each concept was useful to evaluate. The interfaces were placed on a 2019 Citroen C3 3D model. The eHMIs used only visual cues to present information, avoiding potential real-world safety issues with cyclists wearing headphones or audio cues from the eHMIs masking other sounds in the environment (e.g. sirens). They work as follows:

- **Safe Zone**: Uses red (AV not yielding) and green (AV yielding) projections to communicate AV behaviour plus a bonnet display of Stop/Proceed (white arrow in a blue background) traffic signs synchronised with the projections. Previous work suggested that on-road projections suit cyclists. They cover a large surface area and are easy to spot through quick glances [1]. Projections using red/green were evaluated with cyclists for lane merging and were found usable [18]. However, how they perform beyond lane merging is unknown;

- **Emoji-Car**: Placed on the roof, a smaller surface still visible around the vehicle, so it suits cyclists’ requirements [1, 4]. This was important to compare with eHMIs using

---

3Giant Escape 3 bicycle: giant-bicycles.com/gb/escape-3
4Wahoo kickr snap: wahoofitness.com/devices/indoor-cycling/bike-trainers/kickr-snap-buy
6Meta Quest Pro: meta.com/gb/quest/quest-pro/
7EasyRoads3D: easyroads3d.com
8Qualtrics online survey platform: qualtrics.com
larger surfaces. It uses a cyan lightbar at the top, communicating AV state (cyan means the vehicle is in autonomous mode with sensors functioning correctly). The eHMI also uses emoticons, displayed on the front/sides/back, to communicate intent and awareness. Specific emojis were chosen by cyclists in Al-Taie et al.’s [3] design session; the display shows a cyclist emoji to communicate the AV’s awareness of the cyclist and that the AV is yielding to them. A lightning emoji conveys the non-yielding state as this was associated with speed and aggression [3]. The eHMI displays a blinking arrow on the front synchronised with the AV’s directional indicator (two blinks per second), allowing us to explore the feasibility for eHMIs to echo current vehicle signals; 

- **LightRing**: Repurposes Dey et al.’s [9] lightband design to suit cyclists. Evaluating this could provide a first step for eHMIs that work with multiple road users. The display includes an always-on cyan light around the AV, communicating AV state. Yielding intent is shown by repeatedly animating cyan lights: they stroke apart from the front centre (on the front bumper) when the car accelerates and move toward the front centre when decelerating (one stroke per second). This animation was effective with pedestrians [12]; evaluating LightRing would allow us to uncover cyclists’ perceptions of animations communicating yielding. The eHMI re-uses Dey et al.’s [11] approach in The Tracker [11] to communicate awareness; lights closest to the cyclist turn navy blue once the rider is detected; the blue segment grows wider around the AV as the cyclist moves closer. The interface echoes directional indicators; the lights on the relevant side blinking (two blinks per second) in amber.

- **No eHMI**: Baseline condition where no eHMI display was attached to the car. Some results with pedestrians suggested that driving behaviour alone, without an eHMI, may be enough to communicate intent at crossings [19]. However, it is unknown how these findings generalise to cyclists.

The eHMIs communicate different messages (e.g. awareness or autonomous mode) using different cues, such as emojis or animation. This allowed us to compare complete designs catered to cyclists’ expectations and needs and give valuable insights into the types of colour schemes, symbols, and animations eHMIs should use. The eHMIs started reacting to the cyclist when they were 20 meters away; this distance was determined through eight pilot tests.

### 3.4 Study Design

This within-subjects study had two independent variables: Scenario and eHMI. Participants used the VR simulator to interact with an SAE level 5 AV using each eHMI across five Scenarios: (1) Controlled Intersection, (2) Roundabout, (3) Uncontrolled Intersection, (4) Lane Merging and (5) Bottleneck (road users moving toward each other in a narrow lane). These often prompted human driver-cyclist communication [1]. Each scenario has different characteristics, e.g. traffic lights or AV position, allowing us to investigate eHMI versatility. Scenarios were grouped into four tracks, one for each eHMI condition. Each track was a straight 1km two-lane road. Riders cycled in the left lane and had the right of way at intersections and roundabouts. Tracks contained each of the five scenarios where the AV yielded, plus two additional ones where the AV did not yield (see Figure 4). These two were excluded from analysis and used to ensure cyclists paid attention and did not assume the car would always yield. Tracks had a random Scenario order. Participants navigated the seven scenarios, placed 100m apart, until the track’s end. All AVs in one track had the same eHMI. The eHMI sequence was counterbalanced using a Latin square.

Scenarios were modelled after video footage of cycling in the city of Glasgow [1]. UK traffic features were used. Lane Merging had obstacles requiring cyclists to enter from the right lane and exit from the left while merging lanes with a moving AV behind them. Bottleneck had parked cars on both sides; participants cycled in a narrow lane between them with the AV approaching, and one road user had to steer away. At intersections and roundabouts, the
AV accelerated to 30mph (standard UK speed in urban areas) when it was 50 meters from the cyclist and stopped 50cm behind the give-way line if yielding. It accelerated to 25mph in Lane Merging and decelerated to 10mph when yielding. The AV drove at 15mph in Bottleneck, steered to the left (between two parked cars) and stopped when yielding. The vehicle maintained speed when not yielding in all scenarios. Controlled Intersection had red lights for 30 seconds in the non-yielding condition. AVs used directional indicators in Roundabout and Bottleneck. We collected the following data:

- **Post-scenario questionnaire.** To measure the versatility aspect of RQ1, NASA TLX was used for an interaction’s workload, and five-point Likert scale (strongly disagree-strongly agree) questions asked: The AV was aware of my presence and I was confident in the AV’s next manoeuvre. These were derived from work showing that AV awareness and intent are key for AV-cyclist interaction [3, 22].
- **Cycling behaviour.** We addressed RQ2 by measuring speed (meters per second) and shoulder checks (Unity camera (head) Y-axis rotation > 45°; determined through eight pilot tests). These were logged every second while navigating each scenario. We collected gaze data: number of fixations on an area of interest (AOI); these cover vehicle (e.g. windscreen) and traffic control features (e.g. traffic lights, see Figure 6).
- **Post-track questionnaire.** We measured the acceptability aspect of RQ1 using the Car Technology Acceptance Model (CTAM) [26] and usability with the User Experience Questionnaire - Short Version (UEQ-S) [30]. These were previously used to evaluate cycling interfaces and pedestrian eHMIs [10, 35].

- **Qualitative data.** Post-study semi-structured interviews were used to contextualise the findings. Participants discussed and ranked each eHMI. They highlighted any points for improvement, discussed the different scenarios and identified ones that they felt needed/did not need eHMIs.

### 3.5 Procedure

Each participant answered a survey on their demographics and cycling experience. The experimenter then briefed them about the study and showed them videos of the eHMIs, familiarising them with the signals before the study. They were then familiarised with the simulator; the experimenter ensured the participant was comfortable with the bike gear and saddle height by riding for three minutes with no headset. Each participant practised between 7 and 15 minutes of virtual cycling in a car park environment before starting the experiment. A start menu was shown in VR before each track, and the experimenter informed the participant which track and eHMI to select using the right headset controller based on the Latin square. The experimenter then reminded the participant of the eHMI signals and turned on the fan. The participant started cycling and navigated through the scenarios until reaching the track end. The VR app paused after each scenario, and the experimenter read out questions from the post-scenario questionnaire; the participant verbally answered and unpause the app using the headset controller. After each track, the participant took off the headset and had a break while answering the post-track questionnaire on a tablet. This was done four times until they encountered all eHMI conditions. A semi-structured interview followed the experiment.
The study took approximately 90 minutes. The University ethics committee approved the study.

### 3.6 Results

We answer the RQs by reporting cyclist perceptions and behaviours toward each interaction. We also examine how visual attention changes with eHMI through eye-tracking and show how acceptable and usable each eHMI was. Non-significant post hoc results are included in the supplementary material for clarity.

#### 3.6.1 Post-Scenario Questionnaire

The data was not normally distributed (via the Shapiro-Wilk test); we conducted an Aligned-Rank Transform (ART) two-way ANOVA [37] exploring the effects of Scenario and eHMI on our outcomes. Post hoc tests between Scenario and eHMI pairs were conducted using the ART-C method [14].

**NASA-TLX Overall Workload.** Mean values are presented in Figure 5. We found a significant main effect of Scenario, with a small effect size ($F(4, 349.68) = 4.92, P < .001; \eta^2 = 0.05$) and a significant main effect of eHMI with a large effect size ($F(3, 349.78) = 23.49, P < .001; \eta^2 = 0.17$). There was no interaction ($F(12, 349.77) = 0.96, P = .486$). **Controlled Intersection** ($M = 3.9, SD = 3.55$), which had the most traffic control, was the least demanding **Scenario** to navigate; it was significantly less demanding than **Lane Merging** ($M = 5.66, SD = 4.69; P = .001$) and **Bottleneck** ($M = 5.62, SD = 4.34; P = .003$), these involved both road users moving and no traffic control. **Safe Zone** ($M = 3.1, SD = 3.2$) was the least demanding eHMI, it caused a significantly lower workload than **Emoji-Car** ($M = 5.08, SD = 4.1; P < .0001$), **LightRing** ($M = 5.73, SD = 4.25; P < .0001$) and **No eHMI** ($M = 5.75, SD = 4.14; P < .0001$). **Emoji-Car** also caused a significantly lower workload than **No eHMI** ($P = .03$), where cyclists had to infer yielding intent from driving behaviour. NASA-TLX subscale results followed a similar trend to the overall ones; **Controlled Intersection** required the lowest workload, **Safe Zone** outperformed the other eHMIs for all subscales. No interaction was found between **Scenario** and eHMI for all subscales. The detailed findings can be found as supplementary material.

**Confidence in AV Awareness.** A significant main effect of **Scenario** was found, with a small effect size ($F(4, 349.95) = 2.64, P < .05; \eta^2 = 0.03$) and a significant main effect of eHMI, with a large effect size ($F(3, 350.19) = 32.48, P < .001; \eta^2 = 0.22$). There was no interaction ($F(12, 350.11) = 0.93, P = .518$). No significant differences were found between **Scenarios**. Participants were least confident in the AV’s awareness when they did not receive explicit signals in the **No eHMI** condition ($M = 2.91, SD = 1.51$). Confidence was significantly lower around **No eHMI** than all other conditions: **Safe Zone** ($M = 4.26, SD = 0.97; P < .0001$), **Emoji-Car** ($M = 4.02, SD = 1.24; P < .0001$) and **LightRing** ($M = 3.73, SD = 1.1; P = .0001$). Similarly, **LightRing** caused significantly lower confidence than **Safe Zone** ($P < .0001$) and **Emoji-Car** ($P = .008$), despite it explicitly communicating awareness through colour changes.
Confidence in AV Intent. A significant main effect of Scenario was found, with a small effect size ($F(4, 350.03) = 3.79, P < .005; \eta^2 = 0.04$), and a significant main effect of eHMI with a large effect size ($F(3, 350.38) = 36.17, P < .001; \eta^2 = 0.24$). There was no interaction ($F(12, 350.23) = 0.83, P = .620$). Participants were most confident in the AV’s intent in the Bottleneck scenario ($M = 3.62, SD = 1.43$) when the AV was opposite them, moving at a slower speed. We found significant differences comparing Bottleneck with Lane Merging ($M = 3.05, SD = 1.54; P = .002$), where the AV was behind the cyclist (not in their field of view). Participants were most confident when Safe Zone ($M = 4.31, SD = 0.91$) was used. Safe Zone produced significantly higher confidence scores than Emoji-Car ($M = 3.65, SD = 1.4; P = .0008$), and LightRing ($M = 3.16, SD = 1.4; P < .0001$). Emoji-Car produced significantly higher scores than LightRing ($P = .02$). In contrast, participants were least confident in AV intent when there was no eHMI ($M = 2.65, SD = 1.47$), compared to Safe Zone ($P < .0001$), Emoji-Car ($P < .0001$) and LightRing ($P = .007$), where they had a display supporting them.

3.6.2 Cycling Behaviour. Data did not have a normal distribution, so we conducted a two-way ANOVA of Aligned-Rank Transformed (ART) data exploring effects of Scenario and eHMI on cycling behaviour, with post hoc comparisons using ART-C.

Cycling Speed. We found a significant main effect of Scenario with a large effect size ($F(4, 361) = 19.33, P < .001; \eta^2 = 0.18$), and a significant main effect of eHMI with a medium effect size ($F(3, 361) = 11.92, P < .001; \eta^2 = 0.09$). There was no interaction ($F(12, 361) = 0.64, P = .807$). Participants were fastest at Controlled Intersection ($M = 5.36m/s, SD = 1.37$), with no need for any right-of-way negotiation. They were significantly faster at Controlled Intersection than Roundabout ($M = 4.8, SD = 1.29; P = .03$), Uncontrolled Intersection ($M = 4.8, SD = 1.49; P = .01$) and Bottleneck ($M = 3.92, SD = 1.24; P < .0001$). They were slowest at Bottleneck, where the lane was narrower, so we found significant differences between Bottleneck and Roundabout ($P < .0001$), Uncontrolled Intersection ($P < .0001$) and Lane Merging ($M = 5.06, SD = 1.6; P < .0001$) (participants had to make a fast manoeuvre due to an AV moving behind them). Safe Zone ($M = 5.26, SD = 1.23$), which had simple signals covering a large surface, helped participants ride at higher speeds, compared to Emoji-Car ($M = 4.87, SD = 1.5; P = .05$), LightRing ($M = 4.33, SD = 1.49; P < .0001$) and No eHMI ($M = 4.69, SD = 1.5; P = .003$). In contrast, participants were slowest around LightRing, where they inferred yielding intent from eHMI animations, and were significantly slower than around Emoji-Car ($P = .0053$).

Shoulder Checks. Data were binary (1 shoulder check was conducted, 0 not); we analysed the mean number of shoulder checks for each eHMI in each scenario. We found a significant main effect of Scenario with a medium effect size ($F(4, 361) = 9.53, P < .001; \eta^2 = 0.10$) and a significant main effect of eHMI with a small effect size ($F(3, 361) = 3.94, P < .01; \eta^2 = 0.03$). There was no interaction ($F(12, 361) = 1.7, P = .064$). Shoulder checks were most likely at an Uncontrolled Intersection ($M = 0.04, SD = 0.08$), and we found significant differences with Controlled Intersection ($M = 0.12, SD = 0.05; P = .02$) and Roundabout ($M = 0.02, SD = 0.057; P = .002$). However, with the AV in front of the rider, they were least likely at Bottlenecks ($M = 0.11, SD = 0.03$). Shoulder checks were significantly less likely at Bottleneck than Controlled Intersection ($P = .03$), Uncontrolled Intersection ($P < .0001$) and Lane Merging ($M = 0.03, SD = 0.06$; $P = .0045$). Emoji-Car ($M = 0.03, SD = 0.07$) displayed icons on the AV’s roof, producing the highest likelihood of shoulder checks. Checks were significantly higher around Emoji-Car than Safe Zone ($M = 0.01, SD = 0.04; P = .006$), which projected colours on the road over a larger area, causing the least shoulder checks.

3.6.3 Gaze Behaviour. Figure 6 shows the effect of eHMI on participant gaze behaviours. We conducted a Chi-square test of independence to investigate the relationship between eHMI and fixation counts. Post hoc tests were performed using a Chi-Square test of independence with a Bonferroni correction. We found a significant association between the variables ($\chi^2(36, 10970) = 2187.8, P < .001$). Post hoc comparisons showed that participants relied more on traffic control with No eHMI as they fixated on traffic signs/lights and road markings more often than with Safe Zone ($P < .0001$), Emoji-Car ($P < .0001$) and LightRing ($P < .0001$). Results also showed that Safe Zone required less visual attention (fewer fixations on the eHMI display) than Emoji-Car ($P < .0001$) and LightRing ($P < .0001$).

3.6.4 Post-Track Questionnaire. Figure 7 shows the mean scale ratings. A Friedman’s test was conducted to investigate the impact of eHMI on the results. Pairwise post hoc comparisons were conducted using the Nemenyi test.
CTAM Overall Score. We found significant differences among the conditions ($\chi^2 = 19.675, \text{df} = 3, P < .001; \eta^2 = 0.2194$). Safe Zone (MD = 3.86, IQR = 0.64) was the most acceptable. It was significantly more acceptable than LightRing (MD = 3.21, IQR = 0.93; $P = .002$) and No eHMI (MD = 2.9, IQR = 1.02; $P = .0005$), which was the least acceptable. CTAM subscale results are presented in the supplementary material. Safe Zone was the most acceptable in all subscales, except for Perceived Safety, where there was no significant difference with the other conditions ($\chi^2 = 2.882, \text{df} = 3, P = .41017; \eta^2 = 0.0016$).

UEQ-S. Significant differences were found among the conditions ($\chi^2 = 29.793, \text{df} = 3, P < .001; \eta^2 = 0.3525$). No eHMI (MD = $-0.75, IQR = 1.28$) was the least usable. It was significantly less usable than Safe Zone (MD = 1.38, IQR = 1.5; $P < .0001$) and Emoji-Car (M = 0.69, IQR = 1.59; $P = .002$). Safe Zone was the most usable; it was significantly more usable than LightRing (MD = $-0.06, IQR = 1.53; P = .003$). We also found significant differences among eHMI Pragmatic Qualities ($\chi^2 = 27.454, \text{df} = 3, P < .001; \eta^2 = 0.3218$). Safe Zone (MD = 1.38, IQR = 1.5) had the highest qualities, these were significantly higher than LightRing (MD = $-0.63, IQR = 1.88; P < .0001$) and No eHMI (MD = $-0.38, IQR = 1.56; P = .0001$). LightRing had the lowest qualities, these were significantly lower than Emoji-Car’s (MD = 1, IQR = 2.75; $P = .005$). There were also significant differences between eHMI Hedonic Qualities ($\chi^2 = 25.569, \text{df} = 3, P < .001; \eta^2 = 0.297$). No eHMI (MD = $-0.63, IQR = 2.38$) had significantly lower qualities than Safe Zone (MD = 0.88, IQR = 2.19; $P = .0001$). Emoji-Car (MD = 0.75, IQR = 1.13; $P = .001$) and LightRing (MD = 0.63, IQR = 1.69; $P = .006$).

3.6.5 Qualitative Results. We report themes based on the post-study interviews. We conducted an inductive, data-driven, thematic analysis [6] of the interview transcripts (auto-transcribed by otter.ai and corrected by an author). Transcripts were imported into NVivo10. One author extracted 42 unique codes from the data. Two authors sorted these into three themes based on code similarity. This was iterative; disagreements were discussed, and codes were remapped until resolved. Themes with two or more overlapping codes were reassessed and combined when necessary. Participant eHMI rankings are visualised in Figure 8; participants ranked Safe Zone as the best and No eHMI as the worst.

Theme 1: eHMI colours. Participants spoke about their experiences with colour-changing eHMIs communicating intent. They were comfortable with Safe Zone using red and green: “I would go with conventional colours [...] They are easy to understand. I felt safer” - P14. They felt that the colours were distinguishable and unambiguous; “Red and green. Super, super intuitive. I understood very quickly what was going on” - P2. and preferred colour changes over animation: “LightRing would be my favourite if it used Safe Zone’s colours” - P16.

Theme 2: eHMI animations. LightRing’s animations communicating AV intent were hard to distinguish on the move: “I didn’t have time to concentrate on [the animation] while cycling” - P20. Some participants preferred animation to complement other distinguishable
signals: "I think it’s tough interpreting what the car will do through animation alone." - P18.

Theme 3: eHMI state distinguishability. Icons could help eHMIs be more detailed and explicit in their signals. However, participants could not easily differentiate between the emojis in Emoji-Car from a distance. For example, P21 said: "I spent more time trying to identify the emoji." and P13 said, "Interpreting emojis from far caused a lot of ambiguity". This could be because they are too detailed and share some features: "There is too much detail in the emojis, so I had to concentrate more. They are very similar. They both use yellow and have a similar shape." - P20.

3.7 Discussion and Design Changes

All designs were versatile; there were no interaction effects between Scenario and eHMI conditions in any result. This validates Al-Taie et al.’s [3] method for designing versatile eHMIs. However, we found areas where improvement was necessary; just proposing new designs based on cyclist expectations is insufficient, and first-hand interaction feedback needs to be part of an iterative design process. We used our findings to refine each design (see Figure 9).

Controlled Intersection required a lower workload. Cyclists relied on traffic lights, even when eHMIs were present: "I didn’t see the eHMI, I saw a green light and went." - P15. This was similar to findings with human drivers; cyclists fixated more on traffic lights than nearby cars [1]. AV position also impacted our results; participants experienced a higher workload, conducted more shoulder checks and were less confident in AV intent when it was behind them when Lane Merging, but were more comfortable when it was in front of them at Bottleneck, even though there was no traffic control.

Red/green signals were positively evaluated throughout. The colours were easy to recognise and distinguish; distinguishability is a key AV-cyclist eHMI feature due to the many fast-paced scenarios riders navigate. Most eHMIs use one colour and animation to communicate yielding intent [8]. However, we found that this hinders distinguishability and may not communicate enough information quickly. Our findings align with Hou et al.’s [18], where red/green signals performed well for lane merging scenarios. Red/green are also useful for pedestrians [8, 21], providing common ground for eHMIs accommodating multiple road user types.

However, due to the use of red/green in traffic lights, there is a risk of the signals being misinterpreted as instructions from the AV instead of its yielding intent. Nevertheless, participants were most confident in AV intent when the colours were used in all scenarios, and the signals performed well in negotiation-based scenarios, such as Bottleneck. Some examples in traffic show that the same colour can convey different meanings (e.g., amber for pedestrian crossings, traffic lights, directional indicators, hazard lights, and on-car blind spot warnings). Human drivers also use hand signals similar to instructive ones from traffic control officers (e.g., waving for ‘go’ [15]) to communicate their intentions to cyclists rather than instruct them [1]; this effect may extend to red and green. The signal’s perceived meaning may depend on its source, and our results captured this: "There is no rule telling me to stop. Even if it is red, the car will react to me if I go. A rule tells me to stop at traffic lights, and the lights communicate this rule." - P13. New traffic colours, such as cyan, could be a more effective approach to avoid red/green being misinterpreted [9], but these were not positively evaluated in our investigation. A longitudinal study with cyclists learning to interpret such signals might show different outcomes. However, our results showed they need suitable contrasts to be distinguishable and effective. A compelling area for future work is investigating suitable contrasting colours and comparing their performance with red/green eHMIs in different scenarios.

All refined designs incorporated red/green signals to enhance eHMI signal recognisability and distinguishability. We recognised the challenge for colourblind riders to differentiate between red and green, so we incorporated animations, patterns, or symbols into our designs to enhance accessibility. We drew inspiration from traffic lights using light positions (red-top and green-bottom) and animations (flashing amber) to convey meaning. Safe Zone was the most positively evaluated. The eHMI covered a large surface and used red/green signals to communicate intent. Al-Taie et al. [1] discussed the advantages of using the road as a design space.

Safe Zone led to fewer shoulder checks and reduced the workload. Eye-tracking data showed it was easily visible with quick glances; cyclists spent less time fixating on the eHMI than others. Participants did not pay much attention to the bonnet display used in Safe Zone. They were sometimes unaware of its presence; "There was something on the bonnet? I did not know" - P4. Therefore, we relocated the bonnet display to the roof and replaced the traffic signs with colours synchronised with the projected signals. This spread the signals throughout the AV area and emphasised the idea of having displays in cyclists’ peripheral vision, making it easier for them to process the colours and information. To accommodate
colour-blind cyclists, we incorporated patterns on the roof display, using vertical lines for green and crossed lines for red.

Cyclists were slower around Emoji-Car and performed more shoulder checks than Safe Zone. This could be due to the display being on the roof, which currently does not display any signals for interaction; participants were not used to this. They also paid greater attention to interpreting the icons than colours in Safe Zone; eye-tracking data supported this. Qualitative feedback indicated that participants had difficulty distinguishing between emojis, requiring a higher workload. They were also confused by the lightning emoji and suggested an icon more aligned with standard traffic symbols. “I can’t map lightning to anything meaningful” – P1. Therefore, eHMI signals must be easily distinguishable and understandable from a distance. Some participants incorrectly interpreted the top cyan light as a signal of the AV yielding, leading to potentially unsafe actions; P3 mentioned, “I saw the light on top and thought I could pass.” The blinking arrow echoing directional indicators proved redundant and ambiguous, as participants were unsure whether it instructed them to turn or displayed the AV’s turn direction.

We simplified Emoji-Car by keeping it focused on communicating the AV’s intent and awareness. The revised version used red triangles to communicate non-yielding (found in traffic signs, suggesting caution) and green bicycle symbols for yielding. We removed the cyan light and blinking arrow to avoid confusing riders, with the eHMI only communicating necessary information. To address colour-blind cyclists, we relied on icons to differentiate signals. We deviated from Hou et al.’s [18] findings where AV-cyclist interfaces placed on specific car areas did not perform well for lane merging, as we wanted to investigate roof-placed interfaces visible from around the vehicle, recommended by previous research [1,4]. This approach aimed to balance visibility and conformity to existing interface placements, such as taxi signs.

LightRing did not perform well; cyclists did not respond positively to a new colour (cyan) in traffic. Animations imposed a higher workload and were harder to distinguish than colours or icons. LightRing had a higher complexity; it incorporated features such as synchronising amber lights on the car’s side with directional indicators, navy blue lights to indicate awareness, and animations communicating intent. This proved a hurdle as cyclists preferred a more straightforward interface closer to Safe Zone. LightRing’s lights were changed to slowly pulse in green when the AV detects and yields to the cyclist and flash quickly in red when not yielding. Animations complement colour changes rather than being the primary source of information. It also helps colourblind cyclists distinguish between yielding conditions, as the animations (speed-based rather than directional) are easier to differentiate [9]. Flashing animations are used in traffic, e.g. some pedestrian crossing signs flash before changing state. LightRing still communicates AV state using cyan, as the signal changes are more apparent with animations and colours, it will not display multiple signals simultaneously, as in Emoji-Car. Here, cyan is used to communicate a new message not currently communicated by human drivers.

Overall, red/green was a useful colour scheme for eHMI to communicate easily distinguishable messages about the AV’s yielding intent across various scenarios. More complex messages, such as echoing a directional indicator, only added to the workload of using an eHMI. We adjusted all three designs based on cyclist feedback and behaviours observed in the simulator to evaluate a second iteration in a real-world setting.

Figure 9: Revised eHMI. Safe Zone and Emoji-car (1-2) yielding conditions from the front and side, and (3) non-yielding conditions. LightRing (1) the yielding condition, (2) non-yielding, and (3) the communication of AV state in cyan.
4 STAGE 2: WIZARD-OF-OZ EVALUATION WITH AN 'AUTONOMOUS' CAR

Cyclists encountered eHMIs presented on a real car across multiple traffic scenarios to evaluate the refined designs and, for the first time, explore how they may be realised through physical prototypes.

4.1 Participants

We recruited 20 participants (7 Female, 12 Male, 1 Non-Binary; Mean Age = 20.4, SD = 5.9) through social media advertising. Eleven cycled at least once a week, three at least once a month, two multiple times a year, and four once a year or less. Thirteen participants used their own bikes during the study. Five participated in the previous VR study. Participants were compensated with a £10 Amazon voucher.

4.2 Apparatus

We used a grey 2019 Citroen C3, the same car as in Stage 1. The driver wore sunglasses, black gloves and a car seat cover with holes for eyes and arms (see Figure 13). Participants never saw the driver, creating the illusion that the car was an SAE level 5 AV [29]. LED strips and an LED matrix were used to build the eHMIs (see Figure 10). They were plugged into the car’s USB port and controlled by an experimenter via an iPad over Bluetooth. The matrix was placed on the roof using a custom-built panel on a removable rack.11 The rack was present in all conditions; participants were told these were the AV’s sensors. Participants only encountered the car’s front or left side, so eHMIs were only visible in these directions. We used velcro on the car body to attach/detach eHMIs between conditions, white chalk to draw road markings on the ground, and traffic cones to represent obstacles (see Figure 12). Participants were given a Giant Escape 3 bicycle and a helmet if they did not have their own. The Tobii Pro Glasses 2 captured eye-tracking and head rotation (shoulder-checking) data. An iPhone 12 mini was placed on the handlebars to record speed using Cyclemeter.12

4.3 Implementing eHMIs

All eHMIs were placed on the Citroen (see Figure 10) and controlled by an experimenter standing outside the vehicle. They were activated when the car reached specific marked locations for each scenario. They worked as follows:

- **Safe Zone**: We did not use projections because the study was outdoors in daylight, so they were barely visible. This was determined through an early pilot test comparing the visibility of the LED matrix, strip and projection (see Figure 11). We experimented with different projectors, including a Dell 1100MP projector with a high (>11,000) lumens value, but the road projections were still not visible in daylight. We also tried using red/green ambient light through 15,000 lumens LED torches stuck under the car, but they also had minimal visibility. Eventually, we used an LED light strip stuck around the bottom of the front half of the car; this was attached with velcro. This approach brought the lights close to the road surface and still emphasised the concept of Safe Zone being in cyclists’ peripheral vision, especially when used with a roof display. The roof display (LED matrix) showed the red pattern seen in Figure 9 synchronised with red lights from the LED strip when the AV detected the cyclist but did not yield, and the green pattern with green LED lights when the AV was yielding.

---

11HandiWorld roof rack: handiworld.com/handirack/

12Cyclemeter iOS application: apps.apple.com/us/app/cyclemeter-bike-computer/id380595774
Figure 11: Pilot comparing visibility of an LED matrix on the roof, LED strips on the body and a projector on the front bumper.

Figure 12: The scenarios visualised on the study space. (A) Roundabout, (B) Uncontrolled Intersection, (C) Lane Merging and (D) Bottleneck. Traffic cones represent obstacles. Red flags represent endpoints.

- **Emoji-Car:** The LED matrix displayed three green bicycle icons (one on each side) if the cyclist has been detected and the AV will yield, and three red triangles resembling warning signs if not.
- **LightRing:** LED strips placed on the car’s left (2 meters long) and front (1 meter long) using velcro. LEDs were always on in cyan, showing the car was autonomous and not reacting to the cyclist. They changed to green pulsing slowly (1 per second) when the AV will yield and red flashing rapidly (2 per second) when not yielding.
- **No eHMI:** Baseline condition with no eHMI display present. All displays were removed from the car or switched off.

### 4.4 Study Design

A within-subjects design was used with Scenario and eHMI as independent variables. Participants cycled around a moving vehicle with the four eHMI in four scenarios: (1) Roundabout, (2) Uncontrolled Intersection, (3) Lane Merging and (4) Bottleneck. We excluded Controlled Intersection as Stage 1 showed it did not require an eHMI. The study commenced in a coned-off outdoor space (see Figure 12): a 60m straight road intersecting with a 50m road on the left. We drew lane-dividing lines replicating a two-lane road and used cones to mark participant start and endpoints. Participants cycled along the 60m road in all scenarios until they reached the marked endpoint, except in Roundabout, where they made a U-turn. They then cycled back to the start point. Like Stage 1, scenarios were grouped into tracks in random order. AVs used the same eHMI within a track. The eHMI sequence was balanced using a Latin Square.

The AV always yielded to maintain participant safety, but participants were shown both yielding and non-yielding states before each track and told that the AV might not yield. One driver was used for all sessions. They ensured a ≥1m distance from the cyclist, as the UK Highway Code advises. The driver accelerated to 20mph in Roundabout and Uncontrolled Intersection and stopped 50cm (marked using chalk) behind the give-way line. They drove at 15mph in Lane Merging and Bottleneck and decelerated (steered to the left in Bottleneck) according to the cyclist’s speed to yield. Directional indicators were used in Roundabout and Bottleneck. Measures were similar to those in Stage 1. Participants answered the same post-scenario and post-track questionnaires. We collected cycling speed (meters per second; logged every second in each scenario), shoulder-checking (Tobii Glasses’ Gyroscope Y rotation > 90°, determined through pilot tests), and eye-tracking data mapped to the AOIs using Tobii Pro Lab’s AOI tool13. A post-study interview was also conducted.

### 4.5 Procedure

Each participant met the experimenter in the outdoor space. They first answered a survey about their demographics and cycling experience. The experimenter briefed them about the study, instructed them about the different scenarios and showed the start and endpoints. Before each track, the experimenter showed the participant how the eHMI worked (with the lights on the vehicle) so that they were familiar with the signals before interaction. Those who did not

---

use their own bike ensured they were comfortable with the experiment bike. The experimenter checked they had appropriate safety gear, mounted the iPhone to the handlebars, and calibrated the eye-tracking glasses. The experiment started, and the participant moved to the starting point. They started cycling, and the driver started driving once they saw a thumbs-up from the experimenter. The experimenter controlled the eHMI to react to the rider at the appropriate moment. After each scenario, the participant returned to their starting point and answered the post-scenario questionnaire while the experimenter put the next scenario’s obstacles on the road. After each track, the experimenter switched the eHMIs as the participant answered the post-track questionnaire. The experiment ended once the participant cycled on all four tracks and experienced all eHMI conditions. This was followed by an interview with the same structure as the one from Stage 1. The University ethics committee approved the study.

4.6 Results

We report the results using the same structure as Stage 1. We start by reporting our post-scenario and cycling behaviour results, followed by findings from the post-scenario questionnaire (acceptability and usability) and qualitative feedback. Non-significant post hoc results are included in the supplementary material for clarity.

4.6.1 Post-Scenario Questionnaire. The data did not have a normal distribution, so we conducted an Aligned-Rank Transform (ART) two-way ANOVA exploring the effects of Scenario and eHMI on our outcomes. Post hoc tests between Scenario and eHMI pairs were conducted using the ART-C method.

Overall NASA-TLX Workload. We found a significant main effect of Scenario with a medium effect size \( F(3, 206.22) = 9.25, P < .001; \eta^2 = 0.12 \), and a significant main effect of eHMI with a large effect size \( F(3, 207.09) = 26.52, P < .001; \eta^2 = 0.28 \). There was no interaction \( F(9, 206.15) = 1.02, P = .422 \). Lane Merging \( (M = 7.6, SD = 3.2) \) caused a significantly higher workload than Roundabout \( (M = 6.22, SD = 2.57; P < .0001) \), Uncontrolled Intersection \( (M = 6.28, SD = 2.89; P = .0001) \) and Bottleneck \( (M = 6.7, SD = 2.78; P = .002) \). No eHMI \((M = 9.18, SD = 3.86)\) imposed a significantly higher workload than Safe Zone \( (M = 6.23, SD = 2.3; P < .0001) \), Emoji-Car \( (M = 6.5, SD = 2.36; P < .0001) \) and LightRing \( (M = 5.85, SD = 2.46; P < .0001) \). Subscale results are presented in the supplementary material. They follow a similar trend: Lane Merging was the most demanding to navigate, and No eHMI was the most demanding eHMI condition. No significant interaction was found between Scenario and eHMI for all subscales.

Confidence in AV Awareness. We found a significant main effect of Scenario with a small effect size \( F(3, 206.84) = 2.74, P < .05; \eta^2 = 0.04 \), and a significant main effect of eHMI with a large effect size \( F(3, 209.86) = 30.16, P < .001; \eta^2 = 0.3 \). There was no interaction \( F(9, 206.6) = 0.54, P = .846 \). Participants were least confident in the AV’s awareness of them when the vehicle was not in their field of view (Lane Merging \( (M = 3.7, SD = 1.12) \)). This produced significantly lower confidence scores than Roundabout \( (M = 4.03, SD = 1; P = .05) \), which has some set right-of-way rules. Participants were least confident in the AV’s awareness when they did not receive a signal from it; with No eHMI \((M = 2.56, SD = 1.18)\), the score was significantly lower than all other conditions: Safe Zone \( (M = 4.21, SD = 0.75; P < .0001) \), Emoji-Car \( (M = 4.24, SD = 0.72; P < .0001) \) and LightRing \( (M = 4.3, SD = 0.66; P < .0001) \).

Confidence in AV Intent. There was no significant main effect of Scenario \( F(3, 206.71) = 2.04, P = .109 \), but a significant main effect of eHMI with a large effect size \( F(3, 209.22) = 23.21, P < .001; \eta^2 = 0.25 \). There was no interaction \( F(9, 206.53) = 0.90, P = .526 \). As in the confidence in awareness scores, participants needed some explicit signal to be confident in AV intent: No eHMI \((M = 2.39, SD = 1.32)\) received significantly lower scores than Safe Zone \( (M = 4.01, SD = 0.85; P < .0001) \), Emoji-Car \( (M = 3.99, SD = 0.9; P < .0001) \) and LightRing \( (M = 4, SD = 0.85; P < .0001) \).

4.6.2 Cycling Behaviour. Data did not have a normal distribution, so we did a two-way ANOVA of Aligned-Rank Transformed (ART) data exploring effects of Scenario and eHMI on cycling behaviour, with post hoc comparisons using ART-C.

Cycling Speed. We found a significant main effect of Scenario with a small effect size \( F(3, 252.43) = 3.61, P < .05; \eta^2 = 0.04 \), and a significant main effect of eHMI with a small effect size \( F(3, 252.71) = 4.21, P < .01; \eta^2 = 0.05 \). There was no interaction \( F(9, 252.59) = 1.35, P = .211 \). Participants were slowest at Bottleneck \( (M = 6.42, SD = 2.06) \), cycling in a narrow lane. They were significantly slower at Bottleneck than Uncontrolled Intersection \( (M = 7.2, SD = 2.57; P = .04) \). Participants were significantly slower when there was No eHMI \((M = 5.98, SD = 1.86)\) than around eHMIs using large surfaces.
Figure 14: Mean overall NASA TLX workload per Scenario and eHMI in Stage 2.

and abstract signals: Safe Zone (M = 7.01, SD = 2.48; P = .02) and LightRing (M = 7.14, SD = 2.12; P = .009).

Shoulder Checks. Like Stage 1, data were binary; we analysed the mean number of checks. We found a significant main effect of Scenario with a large effect size (F(3, 149.38) = 8.71, P < .001; η² = 0.15), but no significant effect of eHMI (F(3, 154.5) = 0.97, P = .410). There was no interaction (F(9, 150.07) = 0.42, P = .923). Shoulder checks were most likely when the AV was behind the cyclist when Lane Merging (M = 0.1, SD = 0.14). They were significantly less likely when Lane Merging than at scenarios with more traffic control: Roundabout (M = 0.08, SD = 0.22; P = .04), and ones with the AV in front of the cyclist: Bottleneck (M = 0.01, SD = 0.02; P < .0001).

4.6.3 Gaze Behaviour. Figure 15 shows a heat-map of cyclists’ fixations with each eHMI. We conducted a Chi-square test of independence investigating the relationship between eHMI and fixation counts. Post hoc tests were performed using a Chi-Square test of independence with a Bonferroni correction. We found a significant association between eHMI and fixation counts (χ²(30, 9263) = 2158.2, P < .001). Pairwise comparisons showed participants relied more on AV driving behaviour (by fixating on the bumper more often [1]), direction indicators and road markings when there was No eHMI than Safe Zone (P < .0001), Emoji-Car (P < .0001) and LightRing (P < .0001). Safe Zone required less attention than Emoji-Car (P < .0001) and LightRing (P < .0001) as participants fixated less often on the light displays. LightRing also required less attention than Emoji-Car (P < .0001); participants fixated more often on Emoji-Car’s roof display.

4.6.4 Post-Track Questionnaire. Figure 16 shows the mean scale ratings. A Friedman’s test was conducted to investigate the impact of eHMI on the results. Pairwise post hoc comparisons were conducted using the Nemeyni test.

CTAM Overall Score. We found significant differences among the eHMI (χ² = 24.929, df = 3, P < .001; η² = 0.29). No eHMI (MD = 2.85, IQR = 0.54) was significantly less acceptable than Safe Zone (MD = 3.72, IQR = 0.44; P < .0001), Emoji-Car (MD = 3.67, IQR = 0.44; P = .002) and LightRing (MD = 3.68, IQR = 0.22; P = .0003). Subscale findings are supplementary material. No eHMI was the least acceptable in all subscales, except for Perceived Safety (χ² = 9.94, df = 3, P < .05; η² = 0.0913) where there were significant differences between only No eHMI and LightRing (P < .05).

UEQ-S. Significant differences were found among the conditions (χ² = 34.5, df = 3, P < .001; η² = 0.41). No eHMI (MD = −1.06, IQR = 1.31) was significantly less usable than all other conditions: Safe Zone (MD = 1.5, IQR = 1.31; P < .0001), Emoji-Car (MD = 1.38, IQR = 1.75; P = .001) and LightRing (MD = 1.81, IQR = 0.94; P < .0001). We found significant differences in eHMI Pragmatic Qualities (χ² = 32.117, df = 3, P < .001; η² = 0.38). No eHMI (MD = −1.38, IQR = 2.06) had significantly lower qualities than all other conditions: Safe Zone (MD = 1.75, IQR = 1.75; P = .0001), Emoji-Car (MD = 1.38, IQR = 1.88; P = .008) and LightRing (MD = 2, IQR = .88; P < .0001). We also found significant differences in eHMI Hedonic Qualities (χ² = 18.484, df = 3, P < .001; η² = 0.2). No eHMI (MD = −1.13, IQR = 1.56), again, had significantly lower qualities than Safe Zone (MD = 1.5, IQR = 1.8; P = .012), Emoji-Car (MD = 2, IQR = 1.56; P = .005) and LightRing (MD = 1.63, IQR = 1.56; P = .001). Participants preferred explicit communication from the AV.

4.6.5 Qualitative Results. We used the same process as Stage 1. One author extracted 31 unique codes from the data. Two authors sorted these into three themes based on code similarity. eHMI rankings are visualised in Figure 17; participants ranked LightRing as the best eHMI and No eHMI as the worst.

Theme 1: eHMI placement. Placement is a key eHMI feature that enhances interface visibility [4]. Participants praised LightRing’s placement on the AV body; “You see better because you can kind of see the edge of the LED strip from wherever” - P8. Emoji-Car, which was placed on the roof, was harder to recognise: “Compared to LightRing, then you have to really look at the roof to see the emoji” - P18.

Theme 2: eHMI redundancy. Stage 1 showed participants wanted simple signals to communicate yielding, but subtle redundancy complementing colour changes was successful. Pulsing animations in LightRing successfully reinforced the AV’s yielding intent (“LightRing flashing drew attention to itself and different flashing speeds were easy to spot” - P4). Redundant messages presented on the top and bottom of the AV in Safe Zone were well received. For example, P20 said, “Always redundancy is better. The top and bottom displays accommodated that.”
4.7 Discussion

All refined eHMs maintained versatility in real-world settings. However, their differences with No eHMI became more apparent than in Stage 1. No eHMI received the lowest ratings across all metrics. Cyclists were slower and performed more shoulder checks. Visual attention was more spread out; they relied on more AOIs to infer AV yielding. This could be due to the study including a real vehicle, so cyclists may have felt less secure when encountering obstacles, and design improvements made the effect of having an eHMI more noticeable. However, the differences between the displays (and their performance) were less prominent. This general trend was also observed with the five participants who experienced the eHMs in Stage 1. Stage 2 results emphasise that eHMs communicating clear, easy-to-understand, and distinguishable messages significantly improve AV-cyclist interaction.

LightRing received better feedback in some metrics, with participants expressing a greater sense of safety and ranking it as the most preferred eHMI. One contributing factor was its communication of the AV’s state through cyan lights: “I liked the cyan colour telling me everything is fine” - P12. LightRing covered a larger AV surface; Stage 1 showed this was a desirable feature. The animations used in LightRing provided redundancy in conveying AV-yielding intentions, enhancing participant confidence. In comparison, Emoji-Car was not as well-received. Its roof placement drew significant attention from riders, as indicated by eye-tracking and speed data; riders were slower than with the other eHMs. The interface’s use of icons rather than just colours added complexity. For example, P16 noted, “Using emojis stops it from using the entire display space, and the colours were less apparent.” This finding aligns with Hou et al.’s [18], who noted that placing eHMs on specific AV areas could divert cyclists’ attention from the road. In comparison, participants could quickly infer signals from Safe Zone despite it being more abstract and relying solely on colour changes without icons or animations. The widespread distribution of lights (roof and car bottom) made them easier to locate through quick glances, as supported by eye-tracking data, which suggested that cyclists often looked at the car’s centre, not just the eHMI itself, to interpret the signals. According to our results, the design changes to enhance the visibility of Safe Zone have succeeded, even when no road projections were used.

Similar to Stage 1, Scenarios with more traffic control, e.g. Roundabout, needed lower workloads than spontaneous ones, e.g. Lane Merging. Participants had greater confidence in the AV’s intent
at Roundabout. This can be attributed to the well-defined right-of-way rules in the UK Highway Code, making interactions more predictable. Give-way lines indicated where the AV would stop, and the AV’s gradual slowing down gave cyclists more time to interpret implicit cues from driving behaviour. In comparison, Lane Merging was challenging and less predictable. Cyclists were in front of the moving vehicle and needed to conduct more shoulder checks. They had limited time to process signals while moving. Right-of-way was unclear; it was up to the AV to slow down and let them pass. The differences observed among the scenarios and how cyclists behave in them emphasise the challenges in achieving eHMI versatility. Despite this, we found that all scenarios would benefit from an eHMI; all required a higher workload when there was No eHMI.

Overall, Stage 2 showed a significant improvement in AV-cyclist interaction when eHMIs effectively convey clear, understandable, and easily distinguishable signals about the AV’s intentions. Using contrasting colours allowed for clearer communication, demonstrating the fundamental concept of having a simple two-state colour encoding to communicate yielding intent. LightRing demonstrated the benefits of communicating messages through colour changes and animation from all around the AV, while Emoji-Car faced challenges because it was more complex due to its placement and use of icons, which required more attention from cyclists, and Safe Zone effectively balanced abstract signals with visibility enhancements.

5 LIMITATIONS AND FUTURE WORK
All road infrastructure was based on the UK Highway Code, and participants were UK-based. Our methods should be replicated in different traffic cultures to see if the same solutions are still effective. We used a city car (Citroen C3), so it is unknown how our findings generalise to other vehicles, such as SUVs or buses. Our findings are based on initial interaction: some participants experienced the eHMIs in both stages, but future work should consider a longitudinal study giving cyclists more time and experience with each interface to see if the experience changes the performance of the eHMIs. The evaluation only considered cyclists as the eHMIs were designed for them. However, these must work with other road users, such as drivers and pedestrians. We identified factors that make eHMIs effective around cyclists and explored features previously used with pedestrians; future work can combine our results with those of other road users to design more inclusive interfaces.

Stage 1 participants were not moving in physical space with no real obstacles around them. This could impact results such as perceived safety. There were also some rendering limitations with the Meta Quest Pro. Some displays were not clearly rendered from a distance due to the headset resolution. The Quest Pro has a similar resolution and higher pixel density than common VR headsets (e.g., Quest 2), so limitations apply across similar simulators. We overcame some of Stage 1’s limitations in Stage 2. However, we did not conduct the study on real roads due to safety concerns. Future work should evaluate the eHMIs on real roads; this is important, as scenarios may be more complex. We focused on one-to-one encounters, but real scenarios may have multiple cars or cyclists, so eHMIs must be scalable. We chose single interactions to give baseline knowledge that others could extend to more complex ones. We focused on versatility, covering a broader range of scenarios. Both issues must be resolved for AVs to work effectively. Our eHMIs communicate yielding intent through colour changes; we hypothesise that they are already scalable, as they broadcast what the AV will do rather than what other road users should do.

Participants did not see the driver under the car seat costume in Stage 2 and behaved as they would around an AV. This was evident in the results; No eHMI significantly under-performed compared to when an eHMI was present. Qualitative feedback also emphasised this: “It’s so hard with no driver in the car!” - P3 and “I couldn’t see anything! There were no signals.” - P11. Due to safety concerns, the vehicle always yielded in Stage 2 and moved at a maximum speed of 20mph. Participants were still shown the non-yielding conditions and told that the vehicle might not yield to them before each scenario. However, how our results will generalise to faster, non-yielding AVs is unknown.

6 OVERALL DISCUSSION AND GUIDELINES
Our investigation provided insights into several eHMI features, such as animation, icons, colour and placement. We answered the RQs by measuring cyclist perception and behaviour towards the eHMIs. The interfaces were versatile, but needed adjustment from initial designs suggested in previous research [1, 3], demonstrating the necessity of evaluating design ideas. Stage 2 showed that all the interfaces improved the interaction experience despite scenario differences. They emphasised that eHMIs must communicate simple messages that are easy to understand and differentiate through quick glances to be acceptable and usable. We achieved this by utilising large surfaces, such as the vehicle’s body or road around it, to display simple red/green signals about its intent. We contribute novel design guidelines synthesised from our results and use them
as headlines for discussion. We show how our findings compare to ones with human drivers and highlight contrasting points from those with other road users.

*eHMI are key AV-cyclist interaction facilitators.* Social cues from human drivers help riders safely plan their next manoeuvre [1, 3, 22]. We found that this also applies to AV interactions. Unlike findings with pedestrians [12], driving behaviour is insufficient to facilitate interaction, and eHMIs are an encouraging replacement for current human cues. Cyclists were more confident in the AV’s intent and awareness when eHMIs were present: they conducted fewer shoulder checks and were more comfortable riding at higher speeds. Qualitative findings reinforced this: ‘You definitely need eHMIs. When there was no intervention, I had no idea and no control. I felt unsafe’ - P2 (Stage 2).

*eHMI level of detail could depend on traffic control level.* eHMIs were most helpful in scenarios such as Lane Merging and Bottleneck; right-of-way was ambiguous with little traffic control to help cyclists. This aligns with previous work observing cyclists around human drivers; interaction was most likely in these settings [1]. Participants still saw the value of eHMIs in more controlled scenarios. Designers could consider eHMI level of detail here. For example, messages may be displayed later in controlled scenarios to allow riders to infer AV intent from driving behaviour first.

**Versatile eHMIs can use the same signals between scenarios.** This solves a key AV-cyclist interaction problem; cyclists expressed concerns about learning different signals to interact in different scenarios [3, 5]. A scenario-independent language communicating the AV’s intent using abstract (not-yielding/yielding) signals was enough for cyclists to safely navigate all scenarios independent of traffic control level. This differs from findings with human drivers, who used different signals in different scenarios, e.g., hand gestures communicating intent at uncontrolled intersections and facial expressions communicating awareness at roundabouts [1]. It is also a positive step toward synthesising eHMIs that accommodate multiple road users; communicating intent using binary signals was also effective with pedestrians [8].

*eHMIs should explicitly communicate intent and implicitly communicate awareness.* Awareness and intent are considered to be two separate messages AVs should communicate [3, 8, 22]. Abstracting them to one message is effective. For example, Safe Zone and the updated LightRing changed to green when the AV yielded, and the cyclist was detected, so awareness was communicate implicitly. In contrast, separating the messages overwhelmed cyclists. For example, LightRing’s first iteration communicated intent using animation and awareness through colour changes, and the updated Emoji-Car communicated awareness through bicycle symbols and intent by having the symbol green; riders were slower and fixated more on these signals compared to the abstract ones; more effort was required to process them.

*eHMIs should be placed on large surfaces on the AV’s body.* Previous work suggested eHMIs could be placed on areas, such as the roof, to be viewable anywhere around the vehicle [1, 3, 4, 32]. We found that is not enough; cyclists preferred signals quickly viewable at a glance. This requires using large surfaces on the vehicle’s body, as seen in LightRing. Utilising large surfaces supports eHMI versatility; we found the AV’s position relative to cyclists impacts workload, confidence and cycling behaviour. While road projections were well received by cyclists both in Stage 1 and previous work [1, 18], our experience implementing this in outdoor settings proved difficult and expensive with current technology. LEDs placed on the car are more feasible and easily seen in sunny environments. Road projections may also have different effects across varying road surfaces (e.g., gravel), so placing the displays on the car may be more practical.

**Colour changes are critical to message distinguishability.** Cyclists must easily distinguish between messages and understand their meaning. Colour is a primary distinguishable feature. It can be supported by animation or icons, but using these alone did not work; they added to the workload, and cyclists fixated on them more to determine yielding intent. Designers should ensure that any colours used contrast each other and are easily distinguishable from a distance; future work should identify potential colour pairs. AV-pedestrian research also showed colour takes precedence over animation [9]. Designers should consider this overlap when developing eHMIs that work for multiple road users.

**Echoing vehicle signals could confuse cyclists.** This contradicts previous research suggesting eHMIs should help cyclists interpret vehicle signals (e.g., if directional indicators on the front/back are not visible) [1, 3]. The two approaches tested in Stage 1 (LightRing and Emoji-Car) increased workload and confused cyclists (‘does the arrow [in Emoji-Car] mean I should turn or is the car turning?’ - P3). Designers should develop eHMIs that communicate the AV’s yielding intent without overlapping current vehicle signals. Cyclists must distinguish whether the message is from the vehicle signals or eHMI, this could be achieved using distinct placements, colours or animation patterns for novel eHMIs.

**eHMI signals should not significantly depart from the current traffic vocabulary.** Cyclists wanted eHMIs to have a minimal learning curve and blend in with the current traffic vocabulary; this could be through colour (using red/green resulted in a lower workload and higher confidence in AV awareness and intent), animations (where flashing, similar to the ones found in crossings, required cyclists to fixate less than when stroking was used) or any icons used, as our participants could not map lightning emojis to non-yielding behaviour. Therefore, designers should ensure that cyclists are familiar with some aspect of the eHMI to avoid a significant learning overhead or misinterpretation of the messages.

Using these guidelines, we formed a new eHMI design (see Figure 18). A two-strip light band; the top displays red lights when the AV detects a cyclist but will not yield. The bottom shows a green light when the AV is yielding. Separating them into top/bottom makes it easier for colourblind riders to distinguish. This way, designers can use animations (e.g., flashing or progress bar) to communicate messages with different detail levels between scenarios.

7 CONCLUSION

We conducted a two-stage investigation comparing three AV-cyclist eHMIs to test their versatility, acceptability and usability. First, we
assessed each interface in a VR cycling simulator ($N = 20$) across five traffic scenarios with varying traffic control levels. Cyclists preferred eHMs using red/green signals to communicate AV intent. Second, based on the results, we refined all three eHMs and compared them outdoors in a Wizard-of-Oz study ($N = 20$). Participants cycled around a Ghost Driver ‘autonomous’ moving car with real implementations of the eHMs. Cyclists preferred interfaces using large surfaces surrounding the vehicle viewable through quick glances, with animation supporting colour changes. Our results contribute insights into how cyclists respond to real eHMs across various traffic scenarios, and how evaluating design suggestions leads to improved and more valid eHMs. We combined findings from both stages and developed novel guidelines for AV-cyclist eHMs. These offer valuable recommendations on how eHMs can mitigate potential ambiguities and conflicts related to space-sharing with AVs. Our research paves the way for safer and more pleasant interactions between cyclists and AVs in diverse traffic scenarios.

ACKNOWLEDGMENTS

This research received funding from the University of Glasgow Excellence Bursary Award and the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (#835197, ViAJeRo). We thank the University of Glasgow security team for providing us with space to conduct a user study on university grounds.

REFERENCES


[29] Dirk Rothenbücher, Jami 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/10.1145/2809738.2809755


