Discovering Road Cyclists' Needs and Preferences for Mixed Reality User Interfaces Using Immersive Simulation

Radoslaw R. Dukalski[1](✉), Jason K. Moore[1], Peter J. Beek[2], and Frances M. Brazier[1]

[1] Delft University of Technology, Delft, The Netherlands
{r.r.dukalski,j.k.moore,f.m.brazier}@tudelft.nl
[2] Department of Human Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
p.j.beek@vu.nl

Abstract. Striking a balance between simulation realism and immersion is essential for the development and evaluation of a prototyping environment (set-up and methodology) for Mixed Reality Road bicycle racing applications. Multi-sensory immersion, safety, usability, and design process iterability are hard requirements for the proposed prototyping environment in which a tandem rides on a treadmill with a participant riding up-front wearing a Virtual Reality headset. In the virtual environment, with interactable interface artefacts in a 360-video sphere, participants' needs and preferences are dis-covered as participants configure their own 3D user-interface for Mixed Reality capable cycling glasses. Sessions that simulate a cycling race, focus on two distinct scenarios of ascending and descending. Two prototyping methods, while standing and while riding a bicycle, are evaluated and compared. Cyclists' preferences include types of information, their placement in the visual field, size, and colour, depicted in heatmap visualisations. These insights are essential for designers of Mixed Reality systems with real-time information for outdoor bicycle racing.

Keywords: Cycling · Visualisation · Mixed Reality · Augmented Reality · Virtual Reality · User Interface · Prototyping

1 Introduction

Simulators aim to recreate a real-world context or scenario, often for research or training purposes. In traffic research they have been used to study road behaviour [1], and in sports to enhance technique and performance. The fidelity and realism of the simulation drives task immersion and engagement, and by extension encouraging a truer recreation of an associated real-world task or more effective training, teaching or learning of the simulated task.

To recreate outdoor mobility with its changing circumstances and environments, indoor simulators rely primarily on visual immersion, through large projection and Virtual and Mixed Reality (VR, MR) devices. Deeper immersion can be beneficial and can

J. Y. C. Chen et al. (Eds.): HCII 2025, LNCS 16338, pp. 74–97, 2026.
https://doi.org/10.1007/978-3-032-12808-9_6

be achieved through additional means, e.g., vestibular, audio, tactile, control, decision, haptic (force feedback), wind, exertion, and role playing. Combining modalities in itself can be challenging as sensory conflict can lead to simulator sickness. Wintersberger et al. [2] documented the development of a tilting bicycle simulator, focusing on increasing realism and reducing simulator sickness without decreasing performance. Matviienko et al. [3] took a different approach and investigated cycling fidelity in a virtual reality cycling simulation and found that a tandem leads to a high level of realism with little simulator sickness. The VTI 'How we roll' cycling simulator relies on a mechanical construction with free rollers for natural cycling dynamics, and employs a mixed reality headset for visualisation of the virtual world [4].

In strategic sports such as professional road cycling, real-time decision making is crucial, with cyclists relying on cruising and pacing by continuously gauging various factors affecting their performance, like their energy reserves, aerodynamics, heart rate, terrain, and motivation. Road cycling, the focus of this research, is a sport requiring strength, endurance, and strategic use of finite resources towards achieving a competitive goal. This efficiency requires appropriate information delivered in real-time. This is conventionally done with instruments mounted on a handlebar, a solution limited by screen size and its location. Delivering information to an athlete navigating changing and often unfamiliar terrain can be a safety hazard. It can draw attention away from the road, which is especially dangerous when exhausted, or when going downhill at speed on a tire as thick as a thumb.

See-through display Mixed Reality (MR) technology has become more attainable and stands to support task performance by providing users in-context information in a desired location of the visual field. Applying MR has the potential to provide cyclists more information while keeping their gaze up. Simpler non-spatially aware Augmented Reality (AR) devices geared towards cycling are already commercially available (e.g., Everysight Raptop). Merkel et al. [5] explored the application thereof in motorcycling (using a large screen projection as backdrop) and showed it to have benefit on perception and recognition of real-time information. However, their wider adoption in the professional cycling world remains limited.

Application of AR in commuter cycling has been explored by Matviienko et al. [6], investigating ways in which a fictional self-driving bicycle can communicate its intent to its cyclist passenger. In this study, the researchers employed a tandem bicycle outside to simulate self-driving capabilities of a bicycle, allowing the participant to embody the role of a passenger, and to focus on the interface evaluation study.

In pursuit of performance gains road cyclists can use MR smart glasses to sacrifice visual field area in exchange for real-time information. To facilitate designers of such solutions there is a need for design guidelines. This study proposes a methodology to explore this design space adopting a participatory design approach, in which road cyclists are involved early in the design process. This process provides designers with valuable insights into their needs and preferences that the product seeks to accommodate [7]. Generative prototyping sessions reveal tacit and latent user knowledge [8]. Scenario-based design approaches distil preferences in specific contexts [9]. Our previous work, submitted for publication[1], has shown that in road cycling information needs to be different per scenario and that when designing for MR glasses participants showed preferences for sacrificing certain areas of the visual field in exchange for the value of the information offered in return. Effective extraction of user preferences is challenging due to safety, tooling, communication, and protocol. Asking a person their preferences for MR interfaces while cycling alongside them on an outside course is dangerous, and the collected data would be incomplete. Interacting with an MR interface while cycling was explored by Kosch et al. [10], but this study limited itself to a simple interaction of accepting a notification. Conversely, asking a person to design at a desk on paper is easy, effective, and familiar [11] but lacks context and the third dimension, limiting the validity of the findings. Intermediate spatial ‘paper prototyping’ of a cycling-focused user interface has been investigated by the authors (under review), with the intention of verifying the findings in a more visually immersive environment.

Prototypes are of higher value when designed in a context closely matching the use scenario, as it elicits similar needs to those experienced in situ. As acquiring and evaluating user preferences for MR design in-situ can be dangerous, lab-based virtual simulations offer an alternative [12, 13], creating a representative 3D environment approximating real-world conditions, for deep immersion. Lab-based VR simulation affords a controlled 3D environment where participants can explore spatial capabilities of MR applications. Maintaining deep immersion in a controlled environment prolongs contemplation of user needs, giving time for exploration, prototyping, and evaluation. To this end, participants should be freely able to iteratively switch between sensitisation, exploration and prototyping, and evaluation [14], further requiring a safe environment.

Striking a balance between simulation realism and immersion, as well as prototyping capabilities, usability, and safety implies trade-offs. This paper is focused on the development and evaluation of a design prototyping environment with a novel multimodal setup and methodology, created to simulate road cycling and collect valuable User Interface (UI) needs and preferences for MR in a controlled, immersive environment. With the ability to achieve deep road cycling context immersion, it has been designed to elicit relevant information needs from cyclists and provide them with tools to prototype and evaluate real-time information solutions in a controlled environment. Following a multi-phase protocol, participants are asked to conduct a prototyping session for two common road cycling scenarios, and to evaluate their experience with the prototyping environment.

2 Methods

Lab-based individual end-user design prototyping sessions were conducted (Fig. 1), consisting of a structured orientation interview, a warm-up exercise, and a two-part speak-out-loud creative session during which participants prototyped a display interface

[1] Dukalski, R.R., Moore, J.K., Beek, P.J., Brazier, F.M.: A Low-Tech Methodology for Understanding User Needs and Preferences for User-Interface Design for Mixed Reality in a Dynamic Motion Context. Manuscript submitted for review to the Annual Conference on Human-Computer Interaction and Sports (SportsHCI 2025). Proceedings to appear in the ACM International Conference Proceedings Series (ICPS).

in 3D-space using a virtual prototyping environment. This section introduces the study participants, and summarises the experimental design, procedure and setup.

2.1 Participants

Twenty-eight participants took part in the study (5 women and 23 men), consisting of a group of experienced road cycling amateurs and road cycling sport enthusiasts. All of them owned a racing bike and routinely collected data while road cycling. None of the participants reported participation in pro-level races. There were no age selection criteria nor was age recorded (participants were estimated to be between 20 and 40 years of age). Ethics approval was granted by the Delft University of Technology's Human Research Ethics Committee (TU Delft, LabServant ID: 2639). All participants consented before taking part and signed consent forms required for taking part in this study.

2.2 Experimental Design

For each participant a single two-hour one-on-one session was conducted with the same facilitator. The participants were tasked with prototyping a 3D-interface for binocular Mixed Reality smart glasses (with audio capabilities) for use in cycling. To this end, they used a multi-modal Virtual Reality environment, described and illustrated below.

During the prototyping phase they were given freedom (free from current race regulations, see Discussion) in describing and visualizing the information they desired, and how they wished it to be delivered during an individual time trial race. The experiment compared two race scenarios (scenario 1 – descent, and scenario 2 – ascent), to explore the differences in what information is desired, and the way in which it should be presented. Henceforth, 'prototype' refers to a single interactable instance of an interface artefact (e.g., a graph of speed, an icon of a heart, or a number for meters).

2.3 Procedure

Orientation. The session began with a structured interview, with questions concerning personal and professional road bicycle racing experiences. This session was held for orientation and sensitisation purposes, and not further discussed (question list in Appendix 1).

Warm-Up Exercise. During the warm-up exercise, participants were asked to draw from memory their current interface, on which they rely during road cycling. Following the warm-up exercise, the participants were presented with a three A4-page 'inspiration list' of various characteristics to consider during the prototyping session, with basic, illustrated examples of each (see Experimental Setup).

Design context introduction. As part of the design context introduction, participants were asked to assume the role of a road racing athlete, who is monitored by scouts and coaches within their team, taking part in a 40-km race in The Netherlands which included the following segments: uphill, downhill, flat, uphill. Their design was to focus on the second and forth segment, without being constrained by budget or access to data.

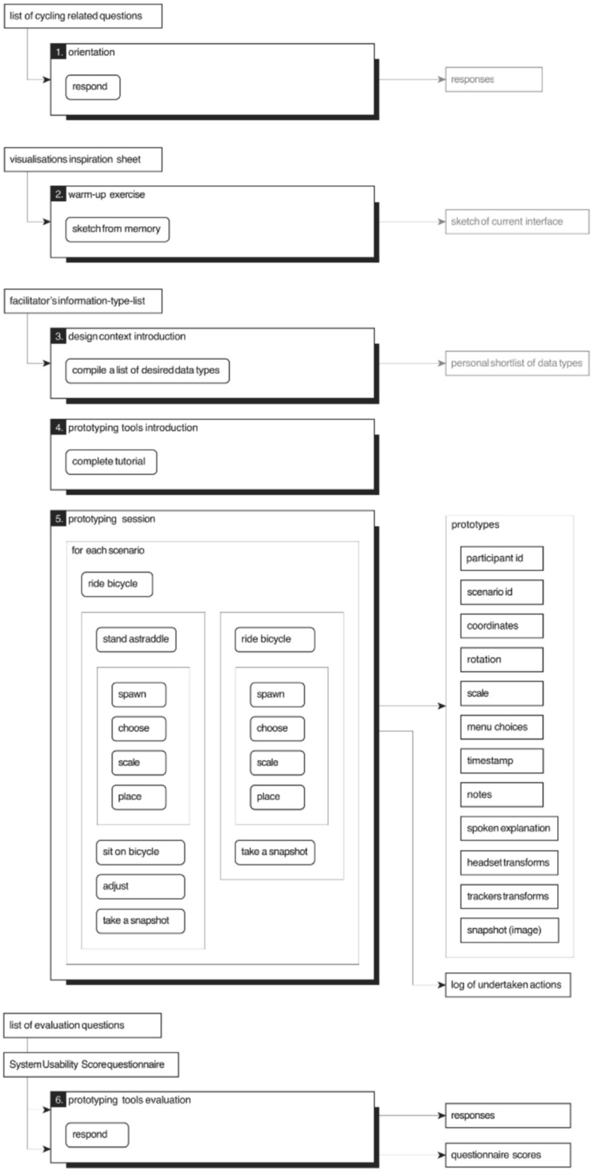

Fig. 1. Flowchart of the six-phase experimental procedure (1. Orientation, 2. Warm-up exercise, 3. Design context introduction, 4. Prototyping tools introduction, 5. Prototyping session, 6. Prototyping tools evaluation), with participant's output shown on the right.

Participants produced a shortlist of 'data types' (e.g., heartrate) which they wished to use in their prototype. Upon completing their list of desired 'data types', the facilitator presented and read out from the 'Facilitator's data-type-checklist' (see Experimental Setup) granting the opportunity to supplement their previously made list.

Prototyping Tools Introduction. The Session continued at the Physical Prototyping Environment (Fig. 3; outlined and portrayed in 'Experimental setup', below). Prior to providing the Virtual Reality headset, the virtual prototyping tools and environment were demoed with the help of a second controller and shown using the facilitator's virtual camera (Virtual Camera B) projected on the wall. The participants showed their grasp of the interactions required during the session: spawning a new prototype, navigating menus, making choices, grabbing and moving prototypes, scaling, deleting, and triggering the snapshot button used for capture. Safety full body harness gear was installed, safety protocol was explained, followed by a 'trust exercise' (30 s of treadmill cycling with eyes open, followed by 30 s of treadmill cycling with eyes closed). Sitting at the rear, the facilitator steered the tandem bicycle (front handlebar was decoupled), always holding onto the side rail with the other hand. The bicycle coasted 'forward' freely, with an occasional minor wobble. As the ceiling safety tether ensured that the bicycle did not ride off the back of the treadmill, pedalling was not required from either rider to remain on the treadmill.

Prototyping Session. In individual speak-out-loud prototyping sessions (illustrated in Fig. 2), Against the video backdrop within the Virtual Reality Environment (outlined and portrayed in 'Experimental setup', below), participants started with a blank slate, and used VR controllers to voluntarily spawn, move and scale interface artefacts. Using floating menus (Fig. 4), they would further specify parameters, e.g., information type, colour, purpose, or circumstance. The facilitator observed progress on a wall projection and aided when needed. Throughout the prototyping session, the facilitator assisted and encouraged thinking out-loud while watching over the progress using a secondary virtual camera (Virtual Camera B) projected onto the wall.

Participants were given advance knowledge of the following protocol and phase order. For each scenario, the participants first experienced riding the bicycle in the virtual environment for 1–2 min (sensitisation), with the treadmill speed set to loosely resemble the speed shown in the looping video (15 km/h for descent, 8 km/h for ascent). The treadmill was then brought to a halt. Thereafter, participants first prototyped while standing astraddle the bicycle (phase: 'stationary'), and afterwards while riding the bicycle (phase: 'cycling'). For each prototype, participants were required to make at least a single choice in each second level menu (except for colour-secondary; outlined below). Upon completion of each prototyping phase (stationary, cycling), having indicated completion of the task, participants were instructed to trigger the snapshot button; upon clicking of which a 5-s countdown was shown with a message was displayed to assume a 'cycling position' appropriate to the scenario (assuming said position in the 'stationary' phase thus required boarding the stationary bicycle). A data snapshot was recorded to disk.

Having completed the first scenario, without clearing the designed interface, the video backdrop was replaced to portray the second scenario, and as before a brief sensitisation period followed. With a recent experience of the new scenario, the participants repeated the two prototyping phases. Participants amended their designs, adding new prototypes and editing or removing old prototypes as desired. Upon completion of the second scenario the prototyping session was concluded.

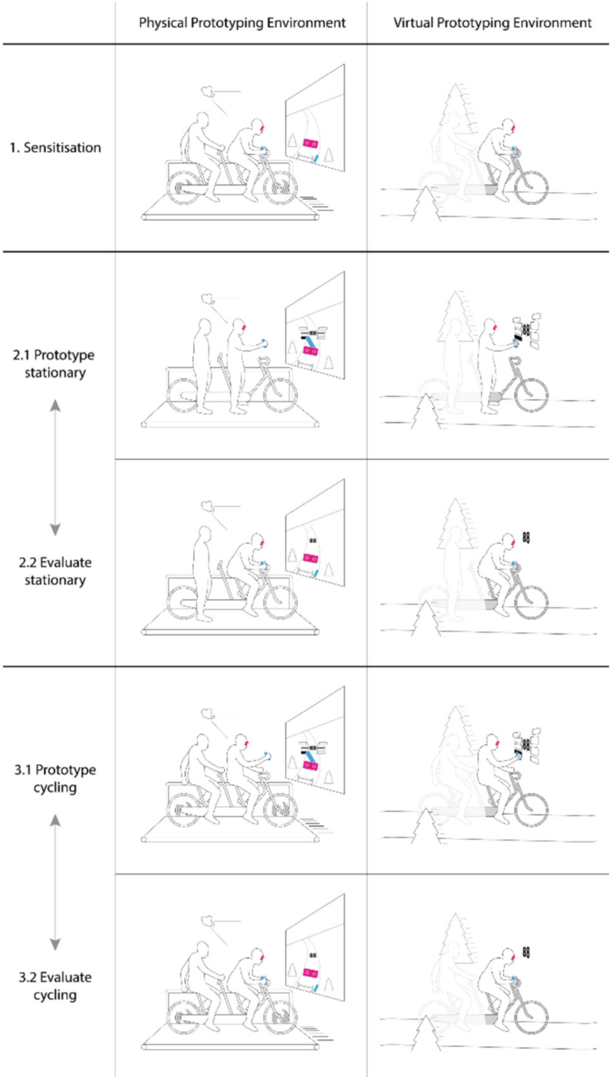

Fig. 2. Prototyping session phases for each scenario – 'Physical Prototyping Environment' (left), 'Virtual Prototyping Environment' (right).

Prototyping Tools Evaluation. Participants were asked to fill out the System Usability Score [15] questionnaire (with 'this prototyping environment, and not the prototyped interface' as the subject of evaluation) to quickly assess the usability, as well as to answer a series of evaluation questions to further specify experienced issues (Appendix 2).

2.4 Experimental Setup

The session was conducted indoors in Delft University of Technology's Bicycle Lab, at an 'Interview Table' and in the 'Physical Prototyping Environment' (which enabled the experience of the 'Virtual Prototyping Environment'), outlined and itemised in detail in Appendix 3.

The presented setup is a prototyping environment, that combines physical and virtual elements (Fig. 1). The physical environment consisted of a customised tandem bicycle riding on a treadmill (Fig. 3), with the participant riding up-front wearing a Virtual Reality (VR) headset, and facilitator riding and steering in the back. The tandem's front half was modified to resemble a racing bicycle, with an appropriate seat and seating position, and a (decoupled) racing handlebar. This provided participants with vestibular, tactile and haptic cycling sensations, with a semblance of control, without the responsibility of keeping the bicycle upright. The locations and orientations off the tandem bicycle as well as its front handlebar were tracked using compatible VR spatial trackers.

Fig. 3. 'Physical Prototyping Environment' – tandem on a treadmill, VR headset and controller.

Within the VR environment, participants were tasked with prototyping a UI for MR cycling glasses using hand-held VR controllers as input devices. The physical tandem's tracked data was mapped and represented by a virtual model. For each of the two scenarios, following a sensitisation (cycling with the treadmill on), two prototyping modes were tested: while standing astraddle on a stationary tandem, and while cycling with the treadmill running.

The Virtual Prototyping Environment was built using the Unity engine [16]. Three virtual cameras were used: one responsible for the participant's headset view (Figs. 6 and 7), second for the cycling facilitator view (Fig. 4) projected on the wall ahead of the tandem, and third with a side view (Fig. 5) matching the lab assistant's desk location.

At the start, a ten-meter diameter sphere was attached to the participant's virtual camera location. Pre-recorded 4K resolution 360-video footage from a real-world location (Figs. 4, 6, and 7) was used and mapped onto the sphere from within. The video's visual forward-flow was synchronised to the treadmill's forward direction using the base

Fig. 4. Facilitator's view (from Virtual Camera B) with the Facilitator's interface panel overlay; Participant is seen selecting a colour from the colour sub-menu.

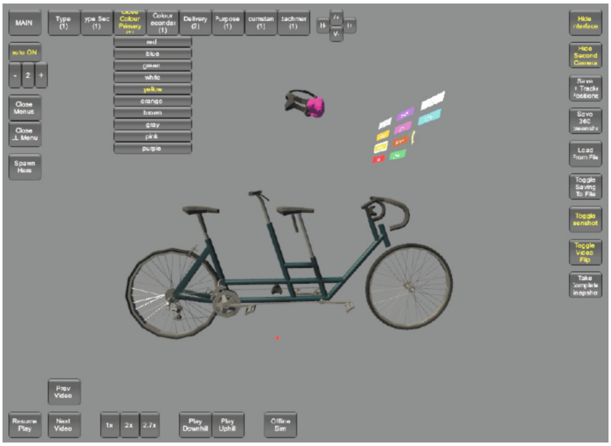

Fig. 5. Lab Assistant's view (from Virtual Camera C) with the Facilitator's interface panel overlay (background removed for illustration); Participant is seen selecting a colour from the colour sub-menu.

station locations (affixed to physical space and by extension, the treadmill). An individual time trial cycling race was simulated, with a video for a climb and for a descent. The spherical videos were captured from a bicycle traversing an outdoor road course, with the stretch of road (35 m elevation delta) completed in both directions. The location (52°19′30.5"N 6°24′59.6"E) was selected by an employee of a 'UCI WorldTour' professional cycling team as representative of a UCI race location (Union Cycliste Internationale, the cycling sport's governing body). The chosen two video clips represented a descent ('scenario 1'; starting at coordinated redacted, distance of 509 m, duration of 49 s), and an ascent ('scenario 2'; starting at coordinated redacted, distance of 228 m, duration of 80 s). The videos were captured using a 360-camera mounted to a bicycle frame at a cyclist's eye position. The descent video playback speed was adjusted to represent a capture taken while descending at 45 km/h, while the uphill video was made to represent a climb at 10 km/h. Each video was played on loop, fading through black, and was flipped horizontally upon each playback iteration.

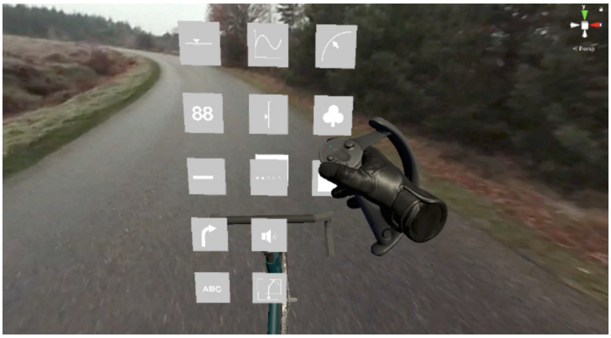

Fig. 6. Participant's VR view (from Virtual Camera A) – foreground: participant's controller shown navigating the representation sub-menu while customising a prototype, background: virtual bicycle and looping video

Using two VR controllers (mapped and visualised in the virtual environment), participants spawned each new instance of an interactable 'prototype artefact' in the location of the Virtual Reality controller, with a long press of a dedicated button. The prototype retained its spatial location relative to the bicycle frame until it was moved again and maintained its rotation independent of the bicycle rotation. Using the controller as a '3D-mouse', participants modified their placement by gripping the controller while hovering over the prototype's preview, scaled using a joystick (similarly during hover), and affected the other eight parameters (listed above) using a concealable two-level multiple-choice menu (using controller's trigger buttons). When shown, menus consisted of floating grey buttons on a 2D plane (Figs. 6 and 7). The eight first level menu buttons lead to designated second-level menus for each of the eight parameters, with button captions revealing the choices made thus far. First-level buttons for 'secondary' parameters (type-secondary, colour-secondary) were indented outwards for de-emphasis (Fig. 7, top left). Within each second-level menu the button order was randomised for each prototype. The prototype's preview visually reflected the choices made for colour-primary and delivery-means (e.g., a yellow arrow, a blue rectangle). Until said choices were made the prototype's visual preview displayed a random choice for each, rotating every one second. In case of multiple choices made (e.g., delivery-means: text and number, colour: red and blue), the last selected choice took precedence visually. Colour-secondary choices were depicted in a form of small swatches below the main preview graphic. Buttons in the second-level menus for colour were coloured to match. Proximity-based hover state mildly enlarged buttons. Buttons' selection state was made visually apparent with further scaling as well as a white border.

Once satisfied with their prototype, the participant used the VR controller to trigger either of the two snapshot buttons. Said buttons were locked to the participant's position, and positioned directly to either side of the participant, within arm's reach around hip height.

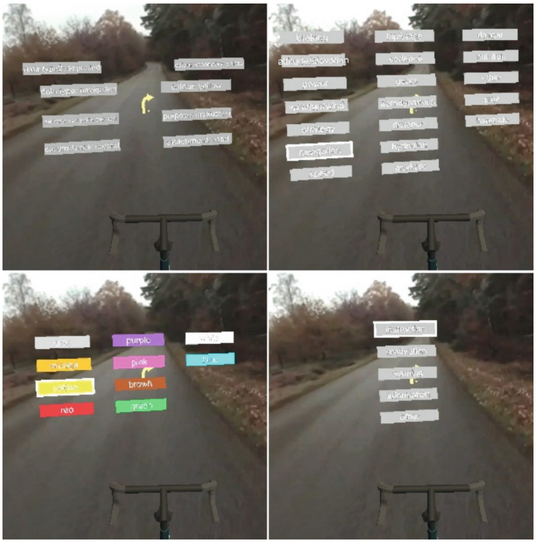

Fig. 7. Participant's VR view when navigating choice menus for an example prototype (top-left: first level menu, followed by example second level menus: type-primary, colour, purpose).

Lastly, illustrated in Figs. 4 and 5 is the Facilitator's interface panel – series of buttons overlayed atop the virtual camera views. The buttons were interactable from a computer using a mouse, allowing the facilitator (or lab assistant) to control video playback, manipulate prototype properties or trigger snapshots on behalf of the participant.

2.5 Prototyping Output

The output of the prototyping sessions (example in Fig. 8) consisted of the following set of data: choices made for each of the eight multiple-choice parameters (as outlined by first-level menu choices in Appendix 3), scale/size, as well as positions and rotations of each: prototype, VR base stations, the VR headset and both VR spatial trackers, as present at each moment of capture (four total: two scenarios in each of the two design phases).

Each participant placed prototypes relative to their distinct gaze point in virtual space. Therefore, when aggregating subsets (e.g., all heartrate-related prototypes), spatial data was used to synchronise the prototypes' positions and orientations as they were relative to their respective VR headset's position and orientation at the time of capture. Once synchronised to a single, unified headset position, the prototypes could then be illustrated in heatmap form (stubbed example in Fig. 9).

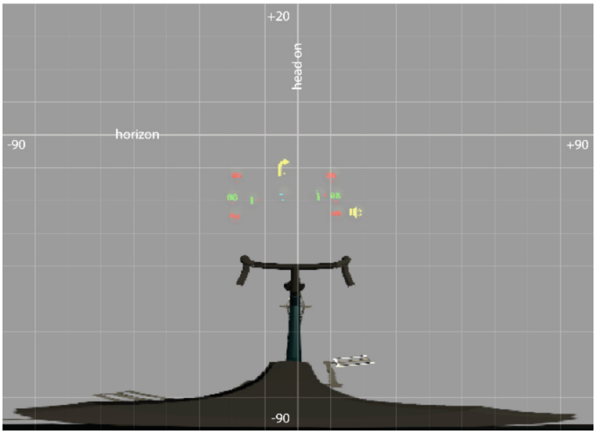

Fig. 8. Example interface of 11 prototypes relative to the bicycle, mapped as an equirectangular projection (backgrounds removed for illustration) from the participant's point of view.

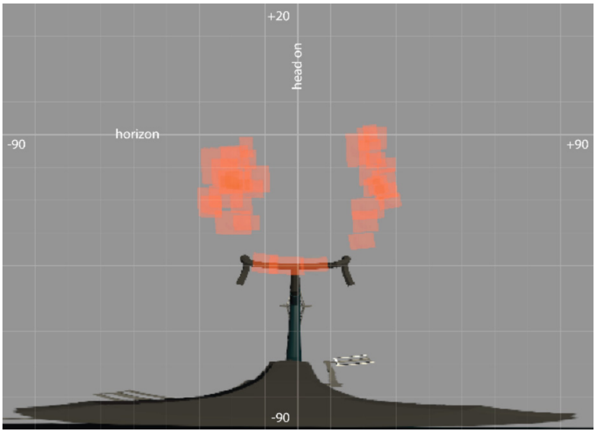

Fig. 9. Mock-up of a heatmap with stub data, mapped as an equirectangular projection from the participant's point of view.

2.6 Rationale

This section discusses in depth certain experimental design choices, where compromises had to be made between use-case immersion, safety, usability, and prototyping iterability. Additionally, consideration was taken to ensure the lab findings best translate to intended outdoor Mixed Reality application.

Design Process and Iterability. Participants were tasked with communicating wilful sacrifices of their visual field in exchange for desired information. Evaluating said choices required eliciting these real-time information needs, done by simulating race conditions (through role play) and dynamic circumstances (through tooling). Considering varied riding and training styles, a wide spectrum of information and visualisation choices was made available. Participants started with a blank slate to not introduce bias in the number of prototypes made, with the options' order randomised when asked to make the choices therein. To facilitate swift iteration between prototyping and evaluation, where possible, protypes' changes took immediate effect, and video kept playing in the background as participants contemplated their options.

The prototypes' menus presented a challenge, being an interface used towards prototyping an interface. To minimise priming with its design patterns, second level menus were varied and lacked symmetry, menu colours were removed entirely (opting for scale and border for communicating hover and active states), and the fidelity level was deliberately low. To encourage fast prototyping, lower the artistic bar and quell perfectionism, prototype previews were static and sketch-like in their appearance while some menu buttons had deliberate formatting issues with text spilling out of their containers.

Immersion and Realism. Deeper immersion was achieved with visual means (a moving video backdrop, matching 3D graphics), vestibular sensation, haptic sensation, as well as control and role immersion. Reproducing vestibular sensation was high on the wish list, however tasking the participant with the mental load of steering the bicycle, contemplating choices, prototyping, and communicating said choices would come at the inevitable cost of the quality of at least one of these much-desired actions. By employing a tandem bicycle, vestibular sensation was achieved through separation of the experience of riding a bicycle from the responsibility of steering it. This would come at the cost of realism due to loss of the feeling of control. To compensate, the participants were shown a reproduced virtual model of a tandem bicycle accurately mapped to the (tracked) physical tandem bicycle, including an independently tracked handlebar. This was not unlike the virtual reality controller in their hand: accurately reproduced, mapped, responsive, and haptic. To deepen role immersion, the front of the tandem had been modified to have physical artefacts resembling the ones found on racing bicycles. Although the handling was not a close match to a single-person bicycle (being less responsive to sideways sway due to two riders, especially when not synchronised in intention), it was deemed an improvement on the alternative: a static bicycle. The match between the vestibular sensation and the visual stimulus was at times broken, largely due to the use of pre-recorded video. The collected footage featured corners (while the treadmill was 'infinitely straight') and scenario-specific road gradient (which was not reflected in the treadmill's and by extension the bicycle's lack of inclination). It was, however, assumed that excluding corners from the video footage would come at the expense of its effect in eliciting real-life information needs, and inclining the treadmill would pose a safety hazard.

Safety. The prototyping session was conducted in a controlled environment without the usual change in circumstance found in road cycling. Steering safety was achieved with riding a tandem bicycle. Furthermore, the facilitator holding onto the rail rather than freely riding the treadmill was deemed sufficient.

The trust exercise was introduced to establish participant's comfort, and to filter out unsuitable participants early. In prior testing, riding on a treadmill with eyes closed lead to disorienting drift of perceived 'forward' direction and its match with the treadmill's direction, causing an unnerving feeling of veering off. Having visual stimulus (in 360 video) headed for a vanishing point eliminated this feeling.

The virtual prototyping tools relied on a single virtual reality controller, enabling participants holding onto the handlebar with the second hand. Adding the virtual model of the bicycle and tracking its position helped participants find the bicycle and handlebar when climbing on, cycling, and getting back down.

Front handlebar was decoupled to prevent participants who were too immersed from taking corners, though it was installed tightly enough to resemble its natural rotational inertia.

Usability. A choice for a rear-steered tandem affords participants ample space ahead, whilst enabling assuming a representative cycling position. During prototyping, upon changing a prototype's position, its location was coupled to bicycle's tracked origin, effectively making prototypes appear still despite the bicycle moving (an issue otherwise apparent when switching between the two design phases: standing and riding). The prototype's rotation however was decoupled from bicycle's rotation, otherwise as the bicycle rolled the prototypes' positions would wobble sideways with a magnified effect (considering their distance from the point of rotation, the ground). Interactable interface elements were made large to enable control despite motion. This ability to continue iterating while riding allowed maintaining vestibular and haptic immersion by not stopping the bicycle when iterating between prototyping and evaluation. The prototype's second-level properties of 'attached to' and 'circumstance' were not implemented as foreseen and about which participants were informed: all prototypes behaved as attached to 'frame', and circumstances were shown 'always' regardless of the respective menu choices. This conversely made the relevant prototypes easier to interact with, compared to fully implemented prototypes reappearing on a timer or moving past as if attached to the road.

Translation to Augmented Outdoor Cycling. From the menu choices of available colours, black was removed to reflect the see-through-display technology's limitation, as in some commercially available Mixed Reality glasses. The played video faded through black as it looped to minimise a sudden break in visual immersion and was mirrored horizontally to eliminate bias due to pre-recorded road's curvature and right-sidedness of the road. The video featured a single car coming from the opposite direction, left in the final cut to retain an element of surprise from unpredictable obstacles found outdoors, a feeling otherwise subdued by the repetitive indoor simulation.

Evaluation Method. The System Usability Score is a straightforward way of assessing system usability, and has been applied in over 1,200 studies, spanning hardware as well as software [17]. It has been shown to provide reliable results with low sample sizes [18] and is comparable to other standardised usability surveys [19]. The choice for a quick survey method stemmed from the anticipated (choice) fatigue following the prototyping session. As SUS is not a diagnostic tool, post-processing interviews were held to supplement the findings, returning the participants to their expert position, and leading the conversation.

3 Data Processing and Analysis

The prototyping setup and methodology were assessed with 28 participants, in three manners: time-stamped VR spatial and input log data, System Usability Score (SUS) survey [15], and a structured interview.

3.1 Log Data of the Prototyping Sessions

The Virtual Prototyping Environment logs to disk timestamped record of actions and interactions in the virtual environment, e.g., participants' spawning, moving, scaling, or pressing of buttons, as well as the facilitator's interactions through their interface. The time of the start of the prototyping process was determined by the timestamp of the participant's first spawn action. The duration of the four prototyping phases (1. Standing-downhill, 2. Cycling-downhill, 3. Standing-uphill, 4. Cycling-uphill) was determined using the timestamp of the participant electing to press the snapshot button. The duration of the sensitisation to the second scenario fell between phases 2 and 3 but was included towards duration of phase 3 as it was considered as part of contemplation prior to impending manipulation of prototypes.

3.2 System Usability Score Survey

Participants completed the System Usability Score (SUS) survey, with the prompt 'please evaluate this prototyping environment, and thus not your interface that you have just created'. It was completed by all participants, regardless of their successful or premature completion of the prototyping process. Individual scores were processed for average and standard deviations per survey question, to finally arrive at a final system score, as per SUS guidelines.

3.3 Post-Prototyping Interviews

Interviews conducted at the end of each session were recorded as audio and were transcribed with automated software. Word search was performed on timestamped transcripts to facilitate further processing. Affect was manually assessed during playback, focusing on the following categories: preference for prototyping while standing/cycling, prototyping process, immersion, realism, usability, safety, simulator sickness, and other. Participants' recommended ideas for improvement were tagged for consideration.

4 Results

Twenty-eight participants took part in the study, three of whom could not complete the prototyping session due to simulator sickness. The results of the evaluations are outlined below.

4.1 Log Data of the Prototyping Sessions

The 'prototyping session' consisted of four phases, with two phases (standing, cycling) for each of the two scenarios. Each pair of phases was preceded by a 90-s 'context sensitising ride'.

Log data tracked progress, culminating with participant-triggered snapshots, including position and orientation of the VR base stations, VR trackers, VR controllers and the VR headset, as well as the participants' gaze point, and UI artefacts' relative placement. The phases varied in duration, as shown in Table 1. The time spent on the four prototyping modes varied, dependent in part on their order (average: 1. Downhill-astraddle: 32:59, 2. Downhill-cycling: 2:43, 3. Uphill-astraddle: 9:23, 4. Uphill-cycling: 1:41). The 'total values' do not include the experiment protocol prior or after the 'prototyping session' part of the experiment. Three participants who did not finish the session because of simulator sickness were excluded from the calculation.

Table 1. Duration of prototyping phases as part of the prototyping session.

Phase	Average	St. Dev.	Minimum	Maximum
1. Stationary – scenario 1: downhill	32:59	12:48	13:39	71:54
2. Cycling – scenario 1: downhill	02:43	01:35	00:34	06:43
3. Stationary – scenario 2: uphill	09:23	06:09	00:07	23:54
4. Cycling – scenario 2: uphill	01:41	01:47	00:00	08:35
Total	49:04	13:37	30:06	88:59

4.2 System Usability Score Survey

Individual scores (Likert: strongly disagree 1–5 strongly agree) for each question (as part of the System Usability Score (SUS) survey [15]) can be found in Fig. 10. Odd-numbered questions required to rate a positively phrased statement, while even-numbered questions required to rate a negative one. The total system score was 71.2 (SD: 2.2). The three participants who did not complete the prototyping session did take part in the survey.

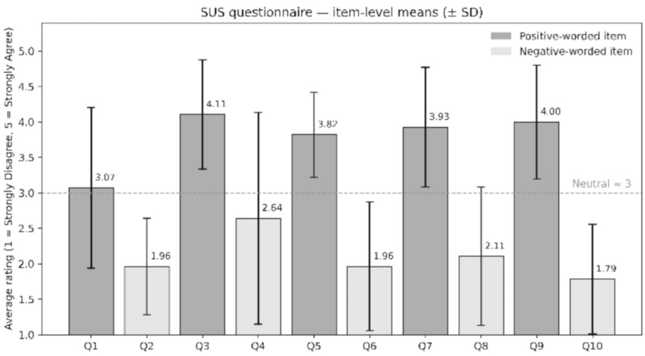

Fig. 10. System Usability Score evaluation value averages (with standard deviation); x-axis: question number, y-axis: reported score.

4.3 Post-Prototyping Interviews

Interviews were conducted with all participants, regardless of having completed the prototyping session. Despite the time duration differences reported above, when asked for their preference, 12 participants reported positive affect towards both stationery and riding modes, with 3 cases of preference only for stationary, 5 for riding, and 8 with no preference. Simulator sickness was experienced by 3 participants leading them to stop (all while prototyping while stationary), with a further 11 reporting having at some point experienced varying degrees of discomfort (reported primarily when stationary).

When asked interview questions (Appendix 2), participants expressed positive affect towards being able to iterate in 10 cases, and were divided on immersion (12 positive mentions, 21 negative), realism (+8, −11), safety (+4, −4), usability (+22, −26), and other factors (+2, −3). Forty eight improvement suggestions were provided by the participants, with respect to hardware (e.g., aero handlebars, pedalling resistance, handlebar turning resistance, inclining the treadmill), immersion (e.g., bicycle lean, wind sensation, road vibration, lack of feeling danger), video (longer, higher resolution, blurred when menu is open atop of it), virtual prototyping tools (e.g., dynamic delivery-means previews, colour opacity manipulation, prototype backgrounds, pre-made layouts), and protocol (e.g., starting prototyping in 2D and continuing in 3D, reversing order of phases).

5 Discussion

The discussion is three-fold: Sect. 5.1 interprets the empirical results through an HCI lens, Sect. 5.2 abstracts key design lessons, and Sect. 5.3 delineates limitations and avenues for future research.

5.1 Interpretation of Results

The timestamped log data revealed a vast discrepancy between the durations of the design phases (‘stationary’, ‘cycling’; Table 1). Amongst 28 participants, instances of simulator sickness (14) were reported to have primarily taken place during prototyping whilst 'stationary'. Whether this is a consequence of the duration discrepancy or the phase's conditions is unknown. Participants had prior knowledge of the protocol order, but whether they felt the autonomy to skip ahead to the 'cycling' phase, or whether they had an expectation of conditions improving, is unclear. The discrepancy in the durations is not reflected in the reported affect and preference for a prototyping phase as expressed during the closing interviews, where the two phases scored comparably both in single-preference (stationary: 3, cycling: 5) and bi-partial positive affect (12). Participants' shorter experiences in the 'cycling' design phases can be attributed to having had completed the scenario's interface with respect to information needs, with only the positioning remaining (supported by interviews). This was observed in both scenarios, with the second scenario being shorter still, attributed to a large overlap in information needs and the reuse of the previously made interface. Arguably, however short, had the 'cycling' phase been entirely superfluous, the average duration would have been approaching zero. Singular cases were observed of participants skipping the 'cycling' phase, supported by negative affect concerning 'safety' reported in the interviews thereafter.

Results of the System Usability Score questionnaire showed an above average (68) score of 71.2, considered 'acceptable' according to [20]. This is a high score considering the complexity of the system, however the novelty of the immersive experience could have positively affected the grading. Participation was self-elected and voluntary, and a token financial appreciation was not revealed until after, potentially leaving the personal satisfaction with one's prototyping 'output' a contributing factor in the evaluation. Furthermore, the prototyping environment is not a daily-use device, thus the grading is anticipated to have accommodated for minor (usability, safety) qualms in exchange for the novel experience, and the perceived value of the prototyping output.

The prototyping environment and methodology are a result of informed trade-offs between immersion, realism, safety, and usability (outlined in the Rationale). The results of post-prototyping interviews show varied affect in different categories. Despite the line of questioning encouraging reporting of negative aspects, the ratio of positive and negative affect remains in relative balance. The discussion which followed produced numerous recommendations for improvement, implementation of which however should consider their impact on the conflicting trade-offs areas.

5.2 Lessons Learned

The proposed indoor multi-modal motion-based Virtual Reality cycling experience simulation has been demonstrated to work. The prototyping environment was shown to be very promising in facilitating deeper context and role immersion, eliciting in participants representative need for real-time information in a safe manner. It further provided them with prototyping tools to iteratively explore, communicate, and evaluate their proposed solutions to their information needs.

The use of the tandem gave the researchers the ability to safely provide participants with deeper immersion with vestibular and haptic stimulus. By taking over steering duties, it also afforded ample time and attention for the participants to think things through and explicate their interface choices clearly and thus provide richer insights.

The use of a secondary virtual cameras enabled researchers to comfortably follow the course of action and provide facilitation.

The deep visual immersion was made possible with 360 videos and a spatially matched virtual bicycle model. The use of video in favour of a computer-generated virtual environment was reported as preferable and more realistic despite the lower resolution. Aligning the video's horizon and vanishing point and the virtual 'forward' direction to match the physical forward direction was essential to maintain immersion. Further mirroring the video on playback iteration eliminated bias. The virtual bicycle model was mapped using spatial trackers and further served as a useful mobile reference origin point in the virtual environment. Participants experienced going 'downhill' and 'uphill' without changing the treadmill's inclination and taking corners while riding straight in reality. Participants were free to move relative to the bicycle and were immersed to a point of physically leaning into the corner, which inadvertently and ironically negatively impacted the immersion with the facilitator's pushed to counterweighting safety measures. A sense of control over the bicycle was afforded through a decoupled handlebar (mapped to a virtual handlebar) and an ability to pedal.

Participants used a virtual reality controller to create, position, and modify prototype interface artefacts. Iteration between prototyping and evaluation was facilitated, with participants reporting positive affect toward prototyping while stationary as well as while riding the bicycle with the treadmill running. Simulator sickness was reported in 3 cases, and discomfort in a further 11 cases, primarily when prototyping while stationary. Despite it, participants remained in the virtual environment for an average of 39 min. The trade-offs between realism, immersion, safety and usability were balanced. The proposed setup and methodology are of use to other VR experiences mediating use scenarios involving motion, and it would be of interest to see it applied in other applications.

5.3 Limitations and Future Research

The selection criteria for participation (road bicycle ownership and experience) as well as hardware mobility limitation have led to a less diverse and by extension smaller population sample, one which is likely disproportionately prone to interest in novel technology. The immersion and usability of this prototyping environment should be evaluated with a more diverse sample.

Simulator sickness affected a small portion of the participants (3 of 28). Amongst assumed causes are visual-vestibular mismatch, long exposure, and individual sensitivity. Future work should consider pre-screening methods, and monitoring sickness with a structured questionnaire.

The order of prototyping methods (stationary, cycling) was not randomised, as the 'stationary' phase was deemed a necessary steppingstone towards the more (physically) challenging 'cycling' phase. Future research should explore reversing the order.

The data collection was not conducted anonymously, and the method used (System Usability Score) is subjective, potentially giving a distorted view of participant's experience. Another evaluation method should be added in the future to support gaining a more objective understanding of the immersion and task load.

User Interface. As part of the virtual prototyping environment, adding higher fidelity previews of prototypes (opacity, dynamic and responsive delivery-means previews, fully implemented circumstance and attachment visualisation) can help improve communication, evaluation, and affect the choices made as part of the prototyping output.

Experimental Setup. Future work should explore implementation of further immersive modalities: vestibular (allowing the participants to lean and take corners to match the video), control (letting participants steer), and exertion (pedalling resistance to match the scenario and task).

6 Conclusions

An indoor Virtual Reality cycling experience simulator, consisting of a tandem riding on a treadmill, was built for iterative prototyping and evaluation of Mixed-Reality real-time information interfaces for outdoor use. Prototyping sessions performed with 28 participants compared two multi-modal approaches, prototyping while 'stationary' and while 'cycling'. Findings indicate that designing MR interfaces for cycling implies trade-offs between simulation realism, immersion, prototyping capabilities, safety, and usability. Following the prototyping sessions, participants were interviewed to evaluate their experience and preference. Twelve participants reported positive affect towards both prototyping methods, with each having specific strengths. Prototyping while cycling in a virtual world was shown to be viable. Deep immersion fosters richer insight into user needs and preferences but can bring along practical and ergonomic constraints. Our approach strives to balance these factors, providing enough realism to evoke authentic user experience in a safe environment.

The presented setup and methodology provide insights into cyclists' needs and preferences, supporting the development of MR UI for use in motion solutions that provide a competitive edge without endangering riders. Although developed for cycling, our multi-sensory approach, that follows a domain-agnostic protocol, offers a structured yet flexible means to develop MR solutions for other mobile applications, including running, driving and flying. Facilitating an iterative design process, our approach helps designers refine the essential aspects of MR UI in motion.

Acknowledgments. A This study was funded as part of the 'Citius Altius Sanius' research programme (P16–28), financed by the Dutch Research Council (NWO).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Appendix

Appendix 1. Orientation Question List

1. Did you cycle here to our meeting? When's the last time you cycled?

2. Do you own a racing bike? When is the last time you took your racing bike out?
3. How often do you race with it, in number of times per month?
4. How many km would you say you've cycled in the last 12 months?
5. What's the most times you've cycled on your racing bike in a month?
6. What's the most km you've cycled in 12 months?
7. What's your favourite route?
8. What's your dream route?
9. Do you watch races often, on tv (number of races per year)?
10. Which channel do you follow races on?
11. Do you see races in person?
12. Are you a member of a cycling club or group?
13. Do you race as a team often?
14. Do you collect data when you cycle?
15. What data do you collect?
16. Which app do you use?
17. Which sensors?
18. Do you use a home trainer, do you use Zwift?
19. Do you have a device to view the data back live?
20. Do you rely on this live data during race or training, and how?
21. Where do you stand on the use of data in cycling racing?
22. In the future, 10, 20, 50 years, how will the sport look?
23. What will interfaces look like, the way that participants consume instructions, data, information?

Appendix 2. Post-prototyping Interview Questions

1. What did you like? What did you dislike? What were you missing?
2. Which method did you prefer, designing while stationary (treadmill not running), or while cycling (treadmill on)?
3. Where there times when your immersion was deeper, or when it was broken?

Appendix 3. Experimental Setup

Interview Table.

1. Paper and pens
2. 'Inspiration-sheet' with basic, non-cycling related, illustrated examples of various visualisation characteristics pertaining to: representation, use of sound, trigger, purpose, duration, dimming, opacity, animation, orientation, fill, scale, amounts, colour, frequency, permanence, precision, units, freshness, and emphasis
3. Facilitator's 'information type checklist' - 'power, distance, time, heartrate, cadence, speed, energy, nutrition, hydration, navigation/map, altitude/elevation, weather, danger, lap'
4. System Usability Score (SUS) questionnaire

Physical Prototyping Environment

1. PC, monitor, mouse
2. Projector, projection screen
3. Virtual Reality Headset (HTC Vive Pro 2)
4. Grip-sensing VR Controller (Valve Index)
5. Spatial tracking system (HTC Vive Base Station 2) (4×)
6. Treadmill (260 × 115 cm belt surface; padded side railings; speed: 15 km/h descent, and 8 km/h for ascent)
7. Tandem bicycle
   a. Steering coupled at the back, decoupled up front
   b. Brakes disengaged
   c. Front saddle - racing model, high seat position
   d. Front handlebar - drop handlebar, no brake levers
   e. VR Spatial Trackers (HTC Vive Tracker 2)
      (1) Spatial Tracker A - attached at the side of the bicycle's frame, low
      (2) Spatial Tracker B - attached to the front handlebar
8. Safety equipment – full body harnesses, ceiling tethers

Virtual Prototyping Environment (Built Using Unity [16])

1. Virtual Camera A (following headset) (headset view, Fig. 6)
2. Virtual Camera B (following Spatial Tracker A position, 2 m higher, always pointing ahead) (projector view; Fig. 4)
3. Virtual Camera C (looking at the bicycle from the side) (side monitor view, Fig. 5)
4. 10-m diameter sphere, with a looping 360 video texture (fading through black) (choice of two videos, one for each scenario)
5. Virtual tandem bicycle model (following spatial tracker A's position and rotation, with its handlebar and front wheel following Spatial Tracker B's rotation)
6. Virtual glove models and controller models (following respective controllers and hand gestures) (2×)
7. Interactable 'Prototype' artefact
   a. Grabbable handle changing location, rotation
   b. Prototype preview
   c. Interactable choice menu: (following handle's position, rotation facing VR headset's posi-tion)
      (1) Data type-primary (19): aerodynamics, altitude/elevation, braking, cadence, danger, distance, gears, gradient, heartrate, hydration, lap/stage, navigation, nutrition, power, speed, strategy, time, weather/wind, other
      (2) Data type-secondary (16): best, current, elapsed/completed, expected, future, gap, global average, stage, limit, max, min, past, remaining/left, running average, total, other
      (3) Delivery-means (13; portrayed graphically): boolean-count-indicator, arrow, rectangle, graph, horizontal linear indicator, horizontal-linear-bar, map, number, radial indicator, sound, symbol, text, vertical linear indicator
      (4) Circumstance (7): always, context, on demand, periodical, location, threshold, other

(5) Colour-primary (10), Colour-secondary (10): blue, brown, green, grey, orange, pink, purple, red, white, yellow
(6) Purpose (5): information, instruction, motivation, warning, other
(7) Attached to (3): frame, helmet, road

d. Virtual Text display of prototype's id number (only visible to Virtual Camera B and C)

8. Snapshot buttons (directly left and right of VR headset position, within arm's reach) (2×)
9. Facilitator's interface panel (Figs. 4 and 5; only visible to Virtual Camera B and C), interactable from computer using a mouse: Video playback controls, Pro-totype choice controls, Menu controls, Snapshot controls, Logging controls and indicators

References

1. Martinez Garcia, D. et al.: Parameter tuning of a bicycle simulator for a realistic riding behaviour and motion perception, Proceedings of the Driving Simulation Conference 2022 Europe, Driving Simulation Association, Strasbourg, France (2022)
2. Wintersberger, P., Matviienko, A., Schweidler, A., Michahelles, F.: Development and evaluation of a motion-based VR bicycle simulator. Proc. ACM Hum. -Comput. Interact. Vol. 6 (MHCI), Article 210 (2022) https://doi.org/10.1145/3546745
3. Matviienko, A., Hoxha, H., Mühlhäuser, M.: What does it mean to cycle in Virtual Reality? Exploring Cycling Fidelity and Control of VR Bicycle Simulators. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 879 (2023). https://doi.org/10.1145/3544548.3581050
4. Kircher, K., Lindman, M., Eldijk, J.V., Weman, J.: Removing barriers to cycling on rural roads. VTI rapport, ISSN 0347–6030; 1220A. (2024) urn: nbn:se: vti: diva-21334
5. Will, S., et al.: Assessment of data glasses for motorcycle riders in a simulated lane change test. Transport. Res. F: Traffic Psychol. Behav. **89**, 467–477 (2022). https://doi.org/10.1016/j.trf.2022.07.016
6. Matviienko, A., Mehmedovic, D., Müller, F., Mühlhäuser, M.: "Baby, You can Ride my Bike": Exploring maneuver indications of self-driving bicycles using a tandem simulator. Proc. ACM Hum. -Comput. Interact. **6**, MHCI, 188 (2022). https://doi.org/10.1145/3546723
7. van der Bijl-Brouwer, M., van der Voort, M.: Establishing shared understanding of product use through collaboratively generating an explicit frame of reference. CoDesign **10**(3–4), 171–190 (2014). https://doi.org/10.1080/15710882.2014.963125
8. Visser, F.S., Stappers, P.J., van der Lugt, R., Sanders, E.B.N.: Contextmapping: Experiences from practice. CoDesign **1**(2), 119–149 (2005). https://doi.org/10.1080/15710880500135987
9. Rosson, M.B., Carroll, J.M.: Usability Engineering: Scenario-Based Development of Human-Computer Interaction. (2001). https://doi.org/10.5555/2821587
10. Kosch, T. et al.: NotiBike: Assessing target selection techniques for cyclist notifications in augmented reality. Proc. ACM Hum. -Comput. Interact. **6**, MHCI, 197 (2022). https://doi.org/10.1145/3546732
11. Snyder, C.: Paper prototyping: The fast and easy way to design and refine user interfaces. Morgan Kaufmann, San Francisco (2003)
12. Lacoche, J., Villain, E., Foulonneau, A.: Evaluating usability and user experience of ar applications in VR simulation. Frontiers in Virtual Reality. Volume 3–2022 (2022). https://doi.org/10.3389/frvir.2022.881318

13. Grandi, J.G., Cao, Z., Ogren, M., Kopper, R.: Design and simulation of next-generation augmented reality user interfaces in virtual reality. 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). 23–29 (2021). https://doi.org/10.1109/VRW52623.2021.00011
14. Kolb, D.A.: Experiential learning: Experience as the source of learning and development. FT press (2014)
15. Brooke, J.: SUS-A quick and dirty usability scale. Usability evaluation in industry. **189**(194), 4–7 (1996)
16. Unity Technologies. Unity Real-Time Development Platform. https://unity.com
17. Bangor, A., Kortum, P.T., Miller, J.T.: An empirical evaluation of the system usability scale. Intl. Journal of Human-Computer Interaction. **24**(6), 574–594 (2008). https://doi.org/10.1080/10447310802205776
18. Tullis, T.S., Stetson, J.N.: A comparison of questionnaires for assessing website usability. Usability professional association conference. **1**, 1–12 (2004)
19. Sauro, J.: A practical guide to the system usability scale: Background, benchmarks & best practices. Measuring Usability LLC, Denver (2011)
20. Bangor, A., Kortum, P.T., Miller, J.T.: Determining what individual SUS scores mean: Adding an adjective rating scale. J. Usability Stud. **4**(3), 114–123 (2009)