A Low-Tech Methodology for Understanding User Needs and Preferences for User-Interface Design for Mixed Reality in a Dynamic Motion Context

Dukalski, Radoslaw R.; Moore, Jason K.; Beek, Peter J.; Brazier, Frances M.

DOI
10.1145/3749385.3749400

Publication date
2025

Document Version
Final published version

Published in
Proceedings of the 1st Annual Conference on Human-Computer Interaction and Sports, SportsHCI 2025

A Low-Tech Methodology for Understanding User Needs and Preferences for User- Interface Design for Mixed Reality in a Dynamic Motion Context

Radoslaw R. Dukalski*
Delft University of Technology
Delft, Netherlands
r.r.dukalski@tudelft.nl

Jason K. Moore
Delft University of Technology
Delft, Netherlands
J.K.Moore@tudelft.nl

Peter J. Beek
Vrije Universiteit Amsterdam
Amsterdam, Netherlands
p.j.beek@vu.nl

Frances M. Brazier
Delft University of Technology
Delft, Netherlands
f.m.brazier@tudelft.nl

Abstract

This paper introduces a novel methodology for user interface prototyping of Mixed Reality applications for a dynamic motion context, namely race cycling. During lab sessions participants prototyped information provisioning in 3D-space. Their choices reflected a trade-off between cost to visual-field real estate and personal value of elected information. Information type, purpose, representation, location, size, and colour were analysed across participants. Participants preferred similar information positioning in the two investigated scenarios (descent, ascent) but included different types of information in each scenario. Heatmap visualisations revealed six preferred visual-field segments, highlighting the amount and types of information as well as segments kept empty. Balanced mock-ups of optimal layouts for descent and ascent are presented. Besides presenting a methodology for both data collection and processing – that is generally applicable by usability researchers both within and outside sports – this study provides specific insights for designers of user interfaces in road race cycling.

CCS Concepts

• **Human-centred computing – Visualisation – Visualization design and evaluation methods**; • **Human-centred computing – Visualisation – Empirical studies in visualisation**; • **Human-centred computing – Visualisation – Visualization techniques**;

Keywords

cycling, visualisation, methodology, guidelines, mixed reality, user interface, prototyping

*contact person

SportsHCI 2025, Enschede, Netherlands

ACM ISBN 979-8-4007-1428-3/2025/11
https://doi.org/10.1145/3749385.3749400

ACM Reference Format:
Radoslaw R. Dukalski, Jason K. Moore, Peter J. Beek, and Frances M. Brazier. 2025. A Low-Tech Methodology for Understanding User Needs and Preferences for User- Interface Design for Mixed Reality in a Dynamic Motion Context. In *Annual Conference on Human-Computer Interaction and Sports (SportsHCI 2025), November 17–19, 2025, Enschede, Netherlands*. ACM, New York, NY, USA, 34 pages. https://doi.org/10.1145/3749385.3749400

1 Introduction

Striving for performance enhancement is inherent to competitive sports, yet it often comes with risks to safety. Performance enhancement is dependent on data – its availability, applicability, and specificity regarding the goal the athlete is trying to achieve. While beneficial in preparation for competitive events, such information is most useful when becoming available only when relevant, i.e. as the activity unfolds. Real-time feedback systems have long been of interest in the Sports HCI domain, especially so as wearable and immersive technologies mature. Virtual and Mixed Reality (VR and MR) devices have undergone a tremendous development, both technologically and commercially. Their application in sports is regarded as the next game changer in performance enhancement by mixed reality (MR) engineers, athletes [1] and sports scientists alike [2], however developing effective applications remains challenging.

A challenge in designing for MR lies in resolving the implicit trade-off between the value of the displayed information, and the cost of the now obscured visual-field real estate. Road cycling, the focus of this paper, stands to gain a lot from just-in-time information provided through MR, but proves to be a difficult case for development. Races are long, environments are novel and unique, conditions change, and margins are thin, making strategic decisions crucial. While the dynamic environment poses an opportunity for a competitive edge, it makes the MR's careful value-cost balance a moving target, as value and cost change with circumstance. Design guidelines are needed to help designers address those needs in a fittingly dynamic manner.

The challenge is however further compounded, as the exploration of said MR user needs, and validation of potential solutions are impractical and dangerous. There is an inherent methodological challenge in effectively prototyping MR solutions for dynamic environments – e.g., adequate fidelity and elicitation of needs, sufficient contemplation, available tooling to express ideas, potential for iteration, and perhaps most crucially, safety. Riding a sharp

descent with pen and paper in hand would be ill-advised and ineffective, while conducting in-situ testing of best-guess prototypes is ethically questionable.

The methodology this study presents in practice, was conducted in a laboratory setting, where participants (n = 18) voluntarily explored and communicated their trade-off boundaries, without quantifying either the cost to the visual field or the value of the self-elected, desired information. Starting with a blank three-dimensional canvas enclosed by a panoramic print, participants positioned annotated acetate cut-outs in their visual field first for downhill (descent) and then uphill (ascent) contexts. Participants' choices concerning type, positioning, representation, purpose, and other parameters were documented and analysed in depth. The aggregated results provided insights into participants' preferences for visualisation in effective MR for road cycling. As a result, design guidelines are derived for (semi-)professional cyclists and other cyclists, in a form of apportioned ordered lists, placement recommendations, and compiled user interface mock-ups. The analysis focused on where to best position specific types of information, in what form, and the degree to which preferences in this regard are context dependent.

The contribution of this paper is two-fold in addressing both challenges: design-as-verb and design-as-noun. By documenting the methodology – namely the data collection process, data processing pipeline, and in providing final scenario-specific design guidelines – the paper is addressed to each, MR usability researchers and MR user interface design practitioners. Through its focus on a challenging case as is road cycling, the user study presented in this paper sheds light on the boundaries of the broader design challenge behind the application of context-relevant MR in similar dynamic environments.

2 Background and Related Work

Amongst the current grand challenges in Sports HCI [3] is a lack of understanding of how to design interactive technologies, especially in MR where the competition for visual real estate between the different reality layers (real and virtualised) is inherent. Further, designs should aim to minimise cognitive load by managing the amount and timing of information [4]. Designing effective user interfaces for Mixed Reality (MR, defined in 2.1) for dynamic environments such as road cycling presents challenges of performance, prototyping, usability, and safety. To this end, as explored in 2.2, involving the end-users in the design process at an early stage is necessary to identify valid user preferences [5]. Performance studies in road cycling are addressed in 2.3, while applications of MR interfaces in similar dynamic contexts are reflected on in 2.4. Lastly 2.5 touches on ethical considerations and dark patterns of MR. Knowledge gap is summarised in 2.6.

2.1 Mixed Reality Terminology

Augmented- and Mixed Reality (AR, MR) have been defined in diverse and often inconsistent ways across the HCI literature, as found by [6] where six distinct definitions of MR are identified, popular amongst them being the Reality–Virtuality Continuum [7]. Their seven-dimensional framework stresses the importance of making the operational definition explicit. This study adopts a vision-centric MR definition aligned with in-the-field augmented overlays delivered via stereoscopic smart glasses, with interaction constrained to visual attention and real-time perception.

2.2 Methodology for Prototyping of MR and VR User Interfaces

A considerable body of research addresses prototyping for MR and VR, with approaches spanning low-, medium-, and high-fidelity. A survey of prototyping techniques in [8], highlighting the trade-offs between realism and accessibility, and [9] lists barriers in authoring applications (e.g., lack of design guidelines). Low-fidelity methods like 360-degree paper templates [10], and physical-digital hybrid tools like ProtoAR [11] allow non-technical participants to contribute in early stages, but may lack spatial accuracy. Medium-fidelity approaches, such as 360theater [12], bridge this gap by adding video capture and interaction simulation. High-fidelity immersive authoring systems, such as XRDirector [13], support more precise spatial coordination, but require higher technical skill and development effort. Ultimately, immersive authoring systems [14, 15] and Wizard of Oz techniques [16–18] enable iterative, low-barrier testing before implementation. With their varied approaches, these methods demonstrate the value of tangible, context-rich elicitation environments for uncovering user needs. Lacking however are subsequent data processing pipelines, where gathered insights are largely qualitative and subject to interpretive bias.

2.3 Performance and Data in Road Cycling

Road bicycle racing performance is shaped by both physiological and contextual factors, and real-time feedback can influence strategy and safety under unique and continuously changing circumstances [19]. This involves varying power output, and by extension depleting and reenergizing the anaerobic energy capacity. Cyclists push uphill, while looking forward to a planned recovery moment while careening downhill on the edge between safety and minimizing time loss. This changing environment (uphill, downhill) alters the task, brings along context-specific challenges, and requires different types of information to compete effectively. Studies in professional cycling have quantified physiological race load [20, 21], explored delivery of optimal pacing strategies using adaptive feedback models [22, 23], and examined event-specific differences in physiological demands.

Compared to traditional bicycle computers, wearable MR has the potential to offer relevant, hands-free information while preserving an aerodynamic pose and forward gaze. Recent changes to regulations concerning race-day communication [24] resulted in increased reliance on onboard devices, thus creating new opportunities, as well as challenges, in delivering real-time information and feedback directly to the cyclist.

2.4 MR Interfaces for Cycling and other Dynamic and Safety-Critical Sports Contexts

Designing MR interfaces for motion-intensive sports requires balancing informational delivery with preservation of situational

awareness. Prior work has identified challenges in balancing information density, cognitive load, and timely delivery. Studies in running and jogging [25–27] demonstrate that minimal, well-placed visual cues can provide motivation and pacing benefits while reducing occlusion risks. Similar benefits were shown in motorcycling [28], where head-up navigation displays reduced reaction times without degrading primary task performance.

In commuter cycling, safety-focused research explored the application of MR in safety visualisations [29], hazard notifications [30], and target selection techniques for notifications [31]. In the data-rich sport of road cycling, decision making can be assisted [32] by providing Augmented Feedback [33] from external sources (e.g., peloton gap size, road conditions [34]) or improving the fidelity of the rider's senses (e.g., heartrate [35]). These works highlight the need to balance timely, context-specific information delivery with adequate consideration for situational awareness.

In Postma et al.'s [36] Sports ITech design space taxonomy, an MR application for road cycling would be positioned as an in-the-field, vision-based feedback system. Functionally, such system would serve the performance optimisation and situational awareness categories, supporting tactical decision-making during ascent and descent scenarios. This study's participatory, low-fidelity prototyping approach enables self-paced exploration of this design space, revealing spatial information placement in a laboratory setting while mimicking the context of competitive cycling. While high-fidelity VR sports training systems, such as VR4VRT for rowing [37] provide richer, real-time data integration, their complexity and hardware requirements can make them less suited to early-stage, participatory ideation, especially in safety-critical contexts such as road cycling where effective hardware simulation is challenging.

A significant body of work has investigated attentional-tunnelling and -blindness in MR and related head-up displays [38–40]. Findings show that augmentation can over-direct attention towards highlighted content, reducing detection of hazards, particularly under high workload. This reinforces the importance of preserving critical real-world visual zones. Related studies have explored mitigation strategies, including differentiated visual design, and timing content delivery with user movements [41].

Dynamic adaptation of displayed MR content has been addressed in walking and stationary settings using rule-based and optimisation techniques [42, 43]. Approaches include adjusting the level of detail, shifting elements to preserve forward visibility, and aligning presentation with the movement. While these systems operate in slower and more controlled outdoor domains such as pedestrian navigation, the underlying principles of minimising distraction, aligning with task demands, and preserving situational awareness, are directly applicable to MR interface design for cycling.

2.5 Ethical and Safety Considerations in MR Design

Recent work has emphasised the need to address potential harms in MR early in the design process. Possible misuse scenarios through low-fidelity prototypes were discussed in [44], while [45] highlights emerging safety, security, and privacy threats. Manipulative design patterns specific to MR were categorised in [46], noting that harmful effects may arise unintentionally from poor design choices, further necessating design guidelines for design practitioners.

2.6 Knowledge Gaps

As discussed above in 2.2 and 2.5, there is a need for a methodology which includes context-catered data collection, and a more rigorous data processing pipeline. In addition, as in 2.3, 2.4, and 2.5, there is a need for design guidelines for a safe and ethical MR design guidelines for (road) cycling.

3 Methods

This user study was designed to allow the extraction of user preferences from expert participants regarding the presentation of information in 3D space. The method relied on utility maximisation of the visual field in a prototyping challenge under a set of design constraints – giving participants free reign affording them the opportunity to make trade-offs, sacrificing valuable visual-field real estate in exchange for self-elected information, without quantifying the value of either.

To achieve the study's aims, prospective end-users participated in prototyping sessions in a lab environment. These sessions consisted of a structured sensitising interview, a warm-up exercise, and a two-part creative session during which participants prototyped a display interface in 3D-space using office stationery, outlined in detail below. The study was approved by the Human Research Ethics Board of Delft University of Technology before its conductance (TU Delft, LabServant ID: 2639).

3.1 Participants

Eighteen participants took part in the study (5 women, 13 men), comprising a mixture of experienced amateurs who themselves practice road cycling, and enthusiasts of the road cycling sport. Fifteen of them reported owning a racing bicycle (here: any bicycle with a drop-bar) and routinely collected data when cycling, while the other three participants owned city bicycles, occasionally collected data, and recreationally followed or attended professional races. Twelve participants regarded cycling as their main sport activity. There were no age selection criteria. Though age was not formally documented, participants were estimated to be between 20 and 40 years old. Recruitment was done through word of mouth and through local cycling associations in the university city. Majority of participants had a technical background. Participation was thus not compulsory, and not monetarily incentivised (although a small financial token of appreciation was granted afterwards; unknown in advance).

3.2 Experimental Design

For each participant a single two-hour one-on-one design session was held with the same facilitator. The participants were tasked with prototyping a 3D-interface for stereoscopic smart glasses (with audio capabilities) for use in cycling. During the prototyping phase they were given a free handle (i.e. free from current race regulations, see Discussion) in describing and visualising the information they desired, and how they wished this information to be delivered during an individual time trial race. The experiment compared two race scenarios (scenario 1 – downhill, and scenario 2 – uphill), to

Figure 1: Composite images of the Prototyping Station showing the prototyping space, with a semi-cylindrical 3D canvas, and video tablet, for descent (left) and ascent (right) (panorama images exported from Google Maps Street View)

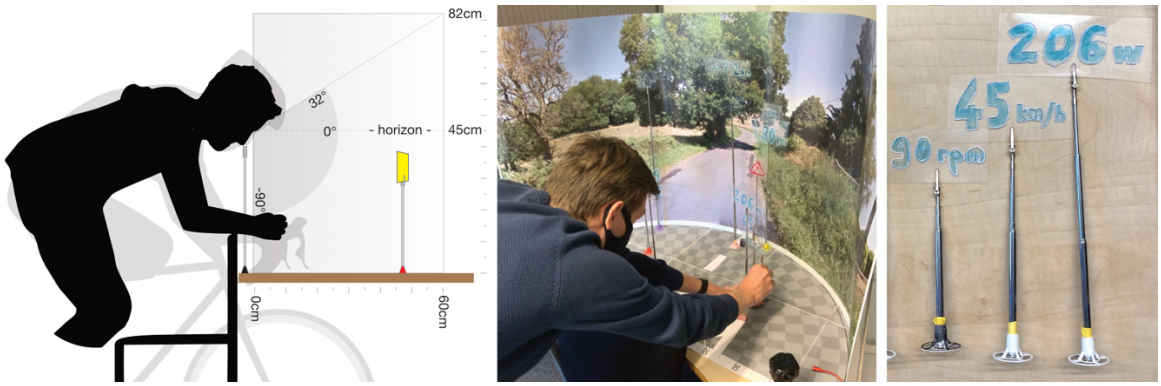

Figure 2: Side cross-sectional view of the Prototyping Station, with the kneeling position (black) and reference cycling position (grey); Photographs of a participant positioning stilts within 3D-space at the Prototyping Station (middle), and acetate examples on stilts (right)

examine the differences in what information the participants wished to have, and how they wished this information to be presented.

3.3 Experimental Setup

The prototyping session was conducted at two office tables, the Tooling Station and the Prototyping Station, between which the participants could move freely. The Tooling Station contained the tools to construct a prototype (acetate sheets, markers), while the Prototyping Station consisted of an enclosed semi-cylindrical volume (henceforth referred to as '3D canvas', see Figures 1 and 2) placed atop a desk, for positioning the constructed prototypes in context, reflecting the participants' design choices and preferences. The 3D canvas was delimited by an open-top structure, consisting of a semi-circular asphalt-grey base, and a vertical curved large format photographic print for each of the scenarios. Dimensions can be found in Figure 2. For panorama images (exported from Google Maps Street View), the two real-world locations were used (Figure 1), selected from past UCI race locations: a descent (Scenario 1; Gien, France; 47.6965033, 2.6446811), and an ascent (Scenario 2; Freibourg, Switzerland; 46.8010122, 7.1631485). The full lists of items available to the participants on both stations are listed in Appendix A.

For positioning prototypes in space, the Prototyping Station required assuming the critical cycling position (illustrated in Figure 2, left) with an arched back and neck, afforded by a reversed chair. This was favoured over an actual bicycle for comfort and access reasons. Prototypes were either placed on telescopic stilts with alligator clips (Figure 2, right), or laid flat on the table. A dedicated telescopic stilt was placed on the table immediately below the participant's head. Placing the nose on its tip (Figure 2) facilitated participants in calibrating the eyes to panorama's horizon level when positioning each acetate cut-out. The desk-scale of the setup placed the entirety of the 3D canvas space at arm's length.

3.4 Procedure

For each participant the procedure consisted of four steps (Orientation, Warm-up exercise, Task Introduction, and Prototyping), which are detailed below. Presented as a flowchart in Figure 3, each phase of the procedure is presented with the facilitator's inputs on the left (e.g., list of cycling related questions), the participant's actions in the middle (e.g., respond), and the participant's outputs on the right (e.g., responses).

3.4.1 Orientation. The session began with a structured interview, concerning personal and professional road bicycle racing experiences (Appendix B). This was conducted for orientation and sensitisation purposes only, without discussing the responses.

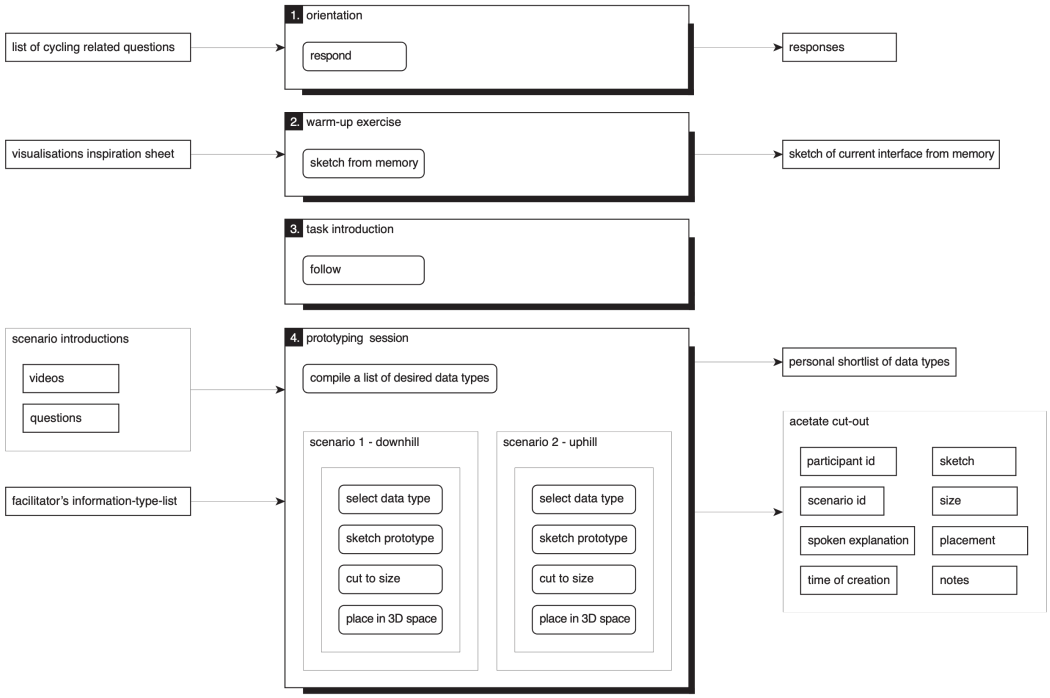

Figure 3: Flowchart of a four-phase experimental Procedure (1. Orientation, 2. Warm-up exercise, 3. Task introduction, 4. Prototyping). For each phase, facilitator's inputs (left), participant's actions (middle; e.g., 'respond'), and participant's outputs (right) are shown.

3.4.2 Warm-up exercise. During the warm-up exercise, participants were asked to recall and draw from memory their current interface, on which they rely during cycling. Following the warmup exercise, participants were presented with a three A4-page Inspiration Sheets (see Appendix C, and section 3.3) of various characteristics to consider during the prototyping session, with basic, illustrated examples for each. The examples contained were deliberately largely unrelated to cycling, introducing interface design elements with minimal priming. Participants were permitted to amend their sketches. Note that black and dark grey colour markers were not provided as to reflect the limitations of current see-through display technology. The resultant sketches were not analysed further.

3.4.3 Task Introduction. Participants were introduced to the design task. They were to assume the role of a road racing athlete, monitored by scouts and coaches within a team, during a 40-km Individual Time Trial race in France and Switzerland, which included the following segments: uphill, downhill, flat, uphill. Their design task was to focus on the second and fourth segments, without being constrained by budget considerations or data access regulations. Participants were then introduced to the prototyping tools found on the Tooling Station (outlined in Section 3.3), which they would use to embody their user interface.

3.4.4 Prototyping. The prototyping phase was conducted at both the Tooling Station and Prototyping Station (outlined in Section 3.3). It consisted of two scenarios, Scenario 1 - Descent followed by Scenario 2 - Ascent, always in that order, as the order could influence the overall race strategy and by extension the information needs. For each scenario, participants were first presented a looping video clip, showing the approximate twenty seconds approaching the panorama locations and ten thereafter (recorded from a coach car trailing a professional cyclist racing at the panorama location; as shown in Figure 1).

The choice for a third- in place of first-person view allowed the participant to observe the body posture of the cyclist in each scenario (with respect to speed, strain, and safety), to better embody them in their role. Participants were then asked questions pertaining to the observed speed, first impressions, and information needs throughout the video as well as upon its abrupt end at the panorama's location. The video looped throughout the session with the tablet to the side, as an enduring reference of the dynamic context.

The rest of the phase was a speak-out-loud prototyping session. In it, participants relied on a 'low-tech' method to bring their vision to life, creating prototypes (acetate sheets, office stationery) and placing it in 3D space (telescopic stilts with alligator clips). As visual reference and backdrop, an immersive panoramic view of the context environment was created by means of two panoramic photographs printed in large format atop a printed asphalt grey checkerboard base (Figure 1). In cases where participants indicated wanting to employ an audio cue, either as supplementary, or a sole carrier of data, they were instructed to verbalise the sound in words. Where the prototype was sound-only, annotated acetate was put aside.

Throughout the prototyping session, the facilitator assisted and encouraged thinking out loud. When approaching the completion of a scenario, the participants were asked to imagine their information needs 200 meters beyond or behind the panorama location. Upon completion of the first scenario, the locations were recorded, and without clearing the designed interface, the panorama backdrop was replaced to portray the second scenario. The looping video was changed to match, and as before the associated questions were posed again. With the understanding of the new context, the participants repeated the prototyping phase: amended their list of data types, if deemed necessary, and were invited to once again prototype, adding new acetate cut-outs and removing old, as desired. Upon completion of the second scenario the session was ended. The output of this phase is described in the next section.

3.5 Prototyping Output

The output of the sessions consisted of acetate cut-outs (Figure 3, bottom right; example in Appendix D) annotated associated parameters of participant id, scenario id, time of creation, an image of a cut-out piece of acetate sheet (see Figure 2), the size and placement on the table, and additional written and spoken notes. An acetate cut-out could serve multiple purposes but was considered single provided it was contained on a single cut out piece of acetate.

4 Data Processing and Analysis

This section describes the procedure of processing and analysing the data gathered during the creative sessions. During those sessions, participants made implicit trade-offs between the cost to their visual-field real estate and personal value of the elected information in determining if and where to place their acetate cut-outs, based on their needs and preferences,. The undertaken process of aggregating and transforming those raw insights into more accessible representations (e.g. heatmaps, mock-ups) is illustrated in the flowchart (Figure 4), and outlined further in the following sections. The results of the analysis can be found in the Results section (Section 5).

4.1 Data Preparation

An example of an unprocessed prototype is shown in Table D1. During clean up, prototypes were trimmed of padding, with the size measurement now reflecting the visually occupied area, instead of the size of the cut acetate sheet. During classification a sequential acetate-id was determined by combining the participant id and creation order (e.g., xx.yy).

Type classification and purpose classification were done manually (listed in Results). Type classification was based on user input, and consisted of type-metric and an optional type-qualifier. Type-metric denotes the core variable, encompassing physiological, mechanical, environmental or strategic metrics (e.g. distance, speed). An optional type-qualifier class further differentiated between e.g., distance-elapsed and distance-remaining. The prototype-id appended the type-metric to acetate-id (e.g., xx.yy.distance). Purpose classification was pre-defined, where four purposes were distinguished: information (no call to action), instruction (urging behaviour change for performance), warning (urging change for safety), and motivation (inspiring perseverance). Similar concepts were merged (e.g., elevation and altitude). Some prototypes received multiple classifications of type-metric, type-qualifier, or purpose.

Photographed sketches were used to classify representation, unit explication, and colour. The representation classification entailed number, graph, map, icon, and sound. Compound prototypes which combined different representations (e.g., graphs with numbers) were filed as belonging to each class simultaneously. Colour was expressed in words, grouped by common, single-word, approximant colours, similarly permitting for multiple colour values.

During Coordinate Determination photographed placement was used to determine numerical values for cylindrical coordinates of ρ (radial distance; collected as base checkerboard's ring number, rounded to the nearest half; with 0 as the radial grid's centre), φh (azimuth; collected as base checkerboard's radial sector numbers, rounded to the nearest half; with 0 as head-on), and z (height; rounded to the nearest centimetre, measured from the centre of the acetate sheet to the table surface). If participants did not employ the stilt and alligator clip and laid the prototype on the table, then the recorded coordinate was the centre of the sheet. Prototypes of representation: sound-only were assigned no size or position.

In addition, the cylindrical coordinate system used for recording locations was converted into polar coordinates (φh : azimuth or horizontal angle; φv : vertical angle; r : radius), with the polar origin (0°, 0°, 0°) at the participant's approximate gaze origin (Figure 2) towards the horizon on the road ahead (cylindrical coordinate: 0, 0, 45cm). The angle of the acetate sheet relative to the stilt or table surface was not recorded, and prototypes were approximated to be orthogonal to the polar origin. The measured size of acetate sheets was converted to 'visual field size' accounting for the placement proximity to the participant and expressed in degrees.

The parameters of the output of the Data Preparation phase are shown in Appendix E.

4.2 Aggregate Analysis

Part of the output from the Data Preparation phase (Section 4.1) served as input to the Aggregate Analysis phase, namely the type-metric, type-qualifier, purpose, representation, visual-field-size, position-polar, unit-explication, colour, scenario-id and participant id. This section focuses on frequency (incidence), as means of prioritising further analysis and synthesis.

4.2.1 Determine frequency. In the analysis, a single acetate featuring e.g., 'a graph of power-expected, power-current, and distance-remaining, is considered as a single prototype towards the count of type-metric: power as well as type-metric: distance. As this prototype serves information concerning power in two manners, it is considered as two separate power instances. The distinction is made to appropriately evaluate the absolute and relative contribution and classification of each acetate, with respect to, e.g., visual field occupation (prototypes), expressed information demand (instances), and embodiment (colour use on acetates, or placement thereof). The number of prototypes or instances belonging to each class within type-metric, type-qualifier, representation, purpose, unit explication, and colour was counted for each of the scenarios-ids. In this paper, the results are plotted on graphs or reported in plain text.

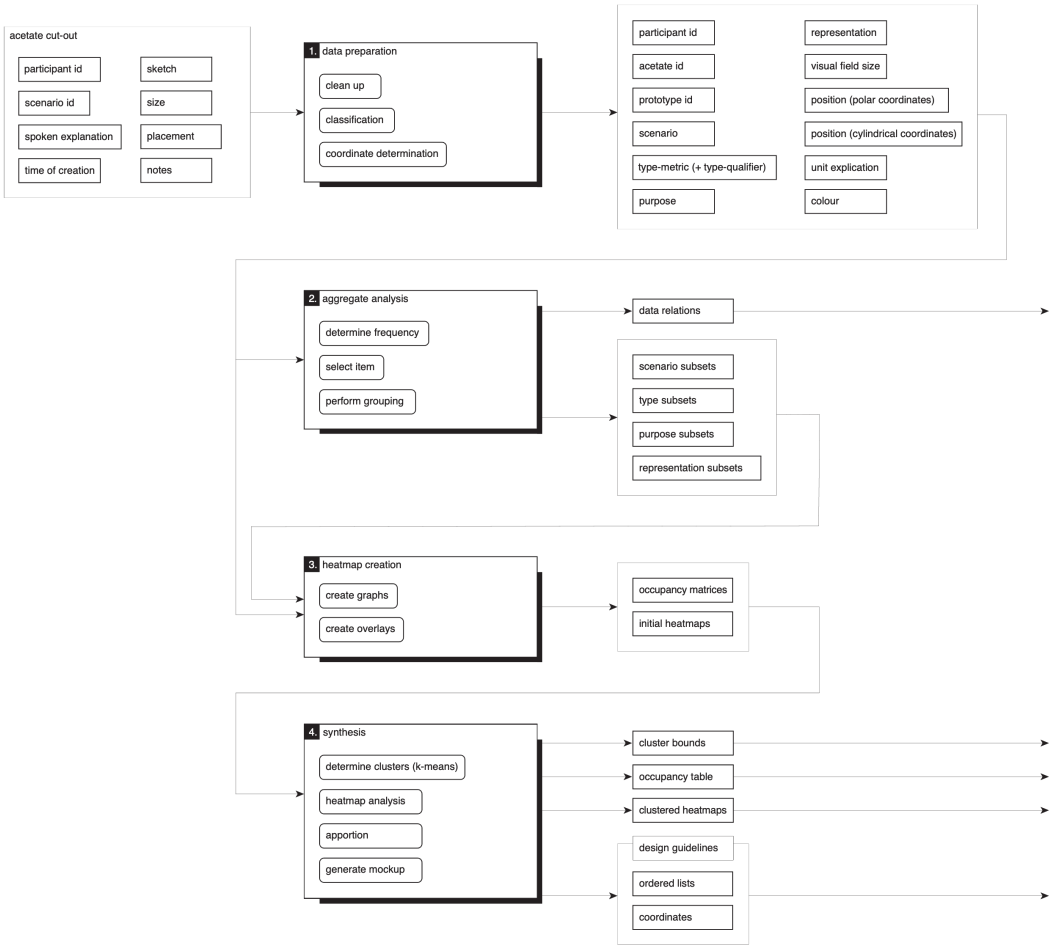

Figure 4: Flowchart of the four-phase data processing procedure (1. Data preparation, 2. Aggregate analysis, 3. Heatmap creation, and 4. Synthesis). For each phase indicating inputs (left), actions (middle; e.g., clean up), and outputs (right). Exit arrows (e.g., cluster bounds) denote an outcome discussed in Section 5.

4.2.2 Select item. A selection was made for further investigation. From representation all values were chosen, except for sound. From purpose, all classes were selected. For both type-metric, and type-qualifier, only the most frequent classes were retained for further analysis. A class was included if its frequency within a given scenario context (ascent, descent and both) was at least 50% of the frequency of the most commonly occurring class in the context in question.

4.2.3 Perform grouping. Based on the selection, the data was grouped and analysed as subsets based on various parameters: the entire set of prototypes; scenario (descent, ascent); type-metric and type-qualifier (based on user defined categories; see Results); representation (number, graph, map, icon); purpose (information, instruction, warning, motivation).

The output of the Aggregate Analysis phase consisted of graphs and plaintext of data relations, as well as chosen 'subsets' (sets of prototype-id values).

4.3 Heatmap Creation

Phase 1 and 2 provided the data needed for Heatmap Creation, namely position-polar, visual-field-size, prototype-id, scenario-id, as well as selected subsets. Heatmaps served to make the underlying spatial trends better interpretable and provided a foundational representation for further analysis and synthesis.

4.3.1 Create equirectangular projection graphs. Prototypes were mapped onto equirectangular projection graphs (x: -90 to +90; y: -90 through the horizon at 0 to 34; see Figure 2, with the y-axis bounds expanded to encompass all collected data), as centre points (polar angle values) as well as rectangles expanding outwards from their centre points (visual field size). Two graphs were created for each subset, one for each scenario (presented in pairs, descent scenario to the left and ascent to the right).

4.3.2 Create overlays. In the subsequent analysis, on each subset-specific graph, a grid was overlayed (62 by 90 squares, of size 2° by 2°) (example in Figure 7) and the number of prototypes present within

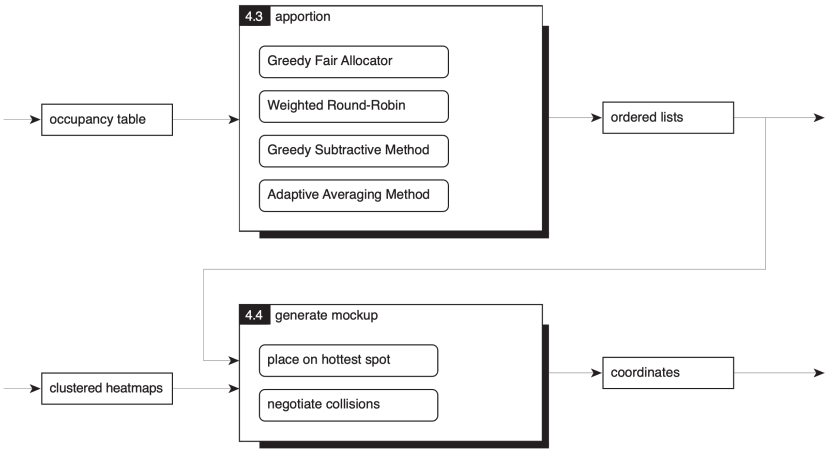

Figure 5: Detailed flowchart of a part of the Synthesis stage of the data processing procedure, with indicated inputs (left), actions (middle; e.g., clean up), and outputs (right). Exit arrows (e.g., cluster bounds) denote an outcome discussed in Section 5.

each grid square's bounds was counted, resulting in 'occupancy matrices'. These matrices determined the shade of the grid squares, resulting in subset-specific heatmaps. The outputs of this phase were a pair of 'global occupancy matrices' (for the entire data set, containing 'All prototypes'), as well as initial subset-specific heatmaps.

4.4 Synthesis

During synthesis, raw spatial data was interpreted into meaningful design insights and interface-level recommendations. The goal was to identify patterns and derive balanced UI mock-ups which reflect these patterns in a fair and structured manner. The inputs for this phase consisted of (a) subset-specific heatmaps and (b) global occupancy matrices. Those were used to determine clusters, analyse and visualise spatial tendencies through heatmaps, and lastly apportion interface zones according to collected patterns in a systematic manner. Based on this synthesis, mock-ups were generated to illustrate potential layouts for MR UIs, informed by participant design insights.

4.4.1 Determine clusters (k-means). The two global occupancy matrices were used to conduct weighted k-means clustering analysis [47, 48] (weighted by the hotness of the corresponding matrix cell; with ten randomly chosen centroids). This was done to partition the visual field into several sectors. The choice of k (number of clusters) was determined by the elbow method [49] and (weighted) silhouette analysis [50] for each matrix. Finally, a Voronoi diagram was drawn to visually separate the areas with what will be called 'global cluster bounds' from here on.

4.4.2 Heatmap analysis. Resulting pair of 'global cluster bounds' were overlaid onto the initial heatmaps of various data subsets. The sectoring described by the global occupancy matrices was applied to subset-specific occupancy matrices. This determined subset-specific sector-occupancy values (percentage of heatmap's hotness falling within the sector). These values were compiled into a 'Heatmap Occupancy Table', listing the selected subsets. The outputs for this phase are global cluster bounds, clustered heatmaps, and the occupancy table.

4.4.3 Apportion and Generate Mock-ups. To generate a visual-field mock-up in a fair and proportional way, an ordered list is required (Figure 5). The ordered list determines the sequence in which items are placed in the visual field. Items of higher importance are given priority in claiming their optimal position and thus avoiding collisions (overlaps). Four methods of generating ordered lists are considered. Each apportion a fixed (integer) number of slots using (fractional) weights based on empirical frequency data (based on the Heatmap Occupancy Table, see Table 1). This disparity makes transparent justification of individual choices on said lists important, especially in design-facing research.

In all methods, based on abovementioned normalised weights (w_i), at each step, the item (i) with the highest effective weight (w_e) is selected from the "queue" and assigned a slot in the apportioned list. The target quota refers to the number of times an item is expected to appear, based on its normalised weight. The Greedy Fair Allocator method assigns each slot to the item currently furthest behind its target quota. The queue is re-evaluated at every step and sorted by effective weight ($w_e = \lceil w_i \rceil - c_i$). The Weighted Round-Robin method (as used in networking or CPU task scheduling) cycles through items in a queue sorted by original weight in descending order, allocating one slot per item in each pass, until their target quotas are met. In the Greedy Subtractive Method, at each step, the item with the largest remaining weight is selected ($w_e = w_i - c_i$). This method does not enforce quotas and can temporarily favour over-selected items if their original weight is sufficiently large. Adaptive Averaging Method dynamically adjusts each item's score based on how many times it has already been selected, using a simple average ($w_e = w_i / (c_i + 1)$). This has the effect of spacing repeated selections and smoothing distribution across the apportioned list. The ordered lists are visually presented as Gantt charts.

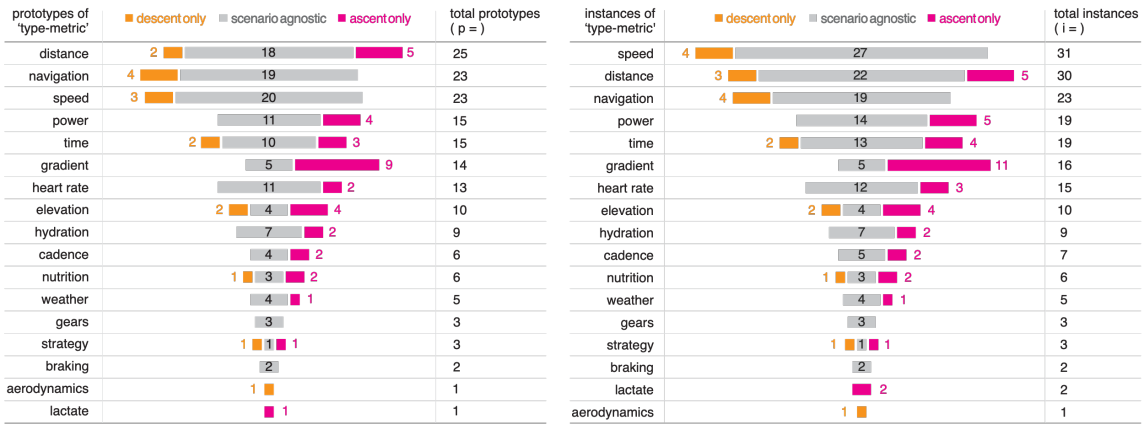

Figure 6: Total frequencies of prototypes (left) and instances (right) of various type-metric classes (left column) – divided into descent-only (orange, chart-left), scenario-agnostic (grey, chart-middle), and ascent-only (magenta, chart-right); sorted by total frequency (right)

To generate a mock-up, items from an ordered list are sequentially placed onto a shared map of the visual field. An item is represented by a circle of radius of 5° and snapped to a grid of 2°. An item is placed onto the location of the approximate centre of its specific heatmap's hottest point, unless that point is already occupied by another item. In case of multiple 'islands' of equally hot areas, the area in a sector of higher occupancy takes priority. The aforementioned rudimentary method of collision negotiation ensures items do not overlap and respect the priority.

5 Results

In this section, the processed results are reported in four sections. Section 5.1 (Data Relations) reports on broader relationships within the dataset. Section 5.2 introduces heatmaps and six sectors determined through clustering. Section 5.3 details how the sectors are occupied in different data subsets. Section 5.4 portrays in detail a selected number of subsets with overlaid sector bounds. Section 5.5 details apportioned ordered lists of instances, and mock-ups of their locations.

To allow comparison of the ascent and descent scenarios (Figure 6), the results are henceforth grouped into the following relational categories: scenario-exclusive (e.g., prototypes found only in ascent), scenario-specific (e.g., found in ascent or in descent, but not in both), scenario-agnostic (e.g., found in both ascent and descent), scenario-inclusive (e.g., found in either ascent or descent), or in-scenario (e.g., found in ascent). Prior terminology distinctions (acetate/prototype/instance) are made in Section 4.2.

5.1 Data Relations

Across the 18 participants 141 acetate cut-outs (henceforth: 'acetates') were created. Those contained 174 prototypes (16 descent-only, 36 ascent-only, 122 scenario-agnostic), which translated into 201 discrete type-metric class instances (18 descent-only, 42 ascent-only, 141 scenario-agnostic). Following the data preparation phase, there were 17 distinct type-metric values present, as several classes were merged (e.g., power/wattage was classified as power; other examples: elevation/altitude, speed/velocity, weather/wind/rain, navigation/ideal line/corner type). Participants' preferences for the type-metric of prototypes (and instances) differed per scenario, as shown in Figure 6.

For both prototypes as well as instances, Figure 6 shows the frequency of occurrence of each type, as found in both scenarios (scenario-agnostic), or in either scenario specifically, i.e. descent (left) or ascent (right). The graphs are sorted by total frequency in descending order. A single prototype type-metric (aerodynamics) was not used in the ascent scenario, while lactate was not used in the descent scenario. Classes of braking and gears were seldom used, but fully scenario-agnostic. No prototypes of class speed or navigation were ascent-specific, while none from power, gradient, heartrate, hydration, cadence, or weather were descent-specific. Class of gradient was the most scenario-specific, with 9 of 14 prototypes pertaining to ascent-only. Prototypes of classes distance, time, elevation, nutrition, and strategy were applied uniquely in each scenario.

The highest frequencies of type classes in either scenario were 31 (speed), in descent 31 (speed), and in ascent 27 (distance and speed). The type-metric classes which were selected for deeper investigation were ones with frequencies higher than half of the above totals (total or descent higher than 15.5, ascent higher than 13.5), thus: speed, distance, navigation, power, time, gradient, and heartrate. These became the list of type-metric subsets. Further breakdowns, relating to representation, colour, or placement can be found in Appendix F.

The Instance Frequency Table (Appendix G) shows the instance frequency of type-metric classes per participants, as well as the total number of discrete type-metric instances and acetates. Participants created an average of 7.8 acetates (range: 4-11), containing an average of 11.2 instances (range: 6-22). In the descent the average was 6.1 (3-9) with 8.8 instances (3-17), while in the ascent 6.8 acetates (2-11), with 10.2 instances (5-22). No type-metric class was used universally by everyone, with speed and navigation coming

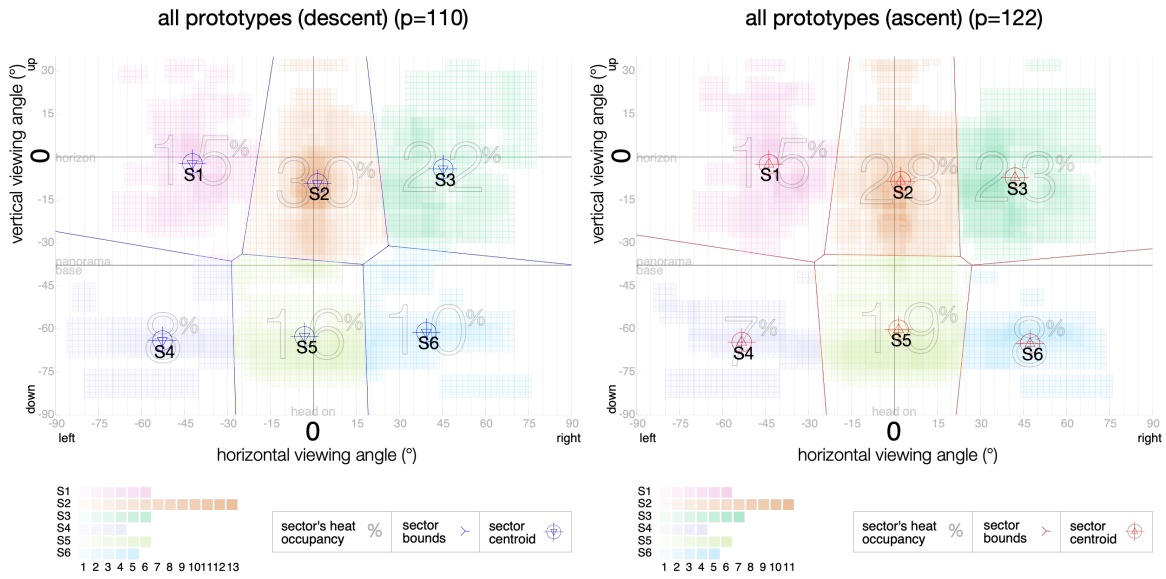

Figure 7: Heatmaps of prototype placements across the mapped visual field (x: −90° to 90°; y: −90° to 34°) for 'All prototypes' in two scenarios: descent (left), and ascent (right). Global cluster bounds for sectors are outlined, with sector centroids indicated (S1-S6). Sectors are visually differentiated by colour, and percentage heatmap occupancy is displayed for each sector. Mini histogram in the legend represent darkest points for each sector.

the closest. Finally, the average number of instances of each type-metric class per participant are shown, as well as the coefficient of variation (as a measure of participant congruence).

Following conversion of cylindrical coordinates to polar, parameter r denotes the radial distance of the centroid of the acetate cut-out away from the gaze point. Participants utilised this depth-of-field to variable effect, illustrated in Appendix H, with the mean radial distance measuring 435mm (shown as a dashed line). Some participants display tight clustering, while others exhibit broader dispersion.

5.2 Global Cluster Bounds

Two heatmaps (descent left, ascent right) are depicted in Figure 7, illustrating the preferred placement in the visual field (x-axis: φh, y-axis: φv) of all prototypes across all participants. Through its colour darkness, each shaded grid cell represents the number of prototypes overlapping within its bounds (occupancy matrix). With 18 participants taking part, both heatmaps contain the highest (darkest) value in the middle beneath the horizon (descent: 13; ascent: 11), with other hotspots (values: 6-7) to its left, right and below, separated from the central one by areas of lesser occupancy (values: 1-4). The heatmaps are most populated in sectors 2, 3, 5, and 1, and least used in 4. A mild bias towards the right can be discerned, as well as a tendency to place more prototypes in the top sectors. The heatmap pair is more populated in the ascent scenario.

As described in Section 4.4, the two global occupancy matrices were used to determine the global cluster bounds, through the application of weighted k-means clustering. To ensure comparability between the two scenarios, a cluster number of k=6 was chosen based on elbow and silhouette analysis for each scenario's data (Appendix I).

These resultant bounds were illustrated with Voronoi cells, which were overlaid on all the heatmaps (with blue lines for descent, and red for ascent), with centroids of each cell marked (points S1-S6, in blue and red; Appendix J). Furthermore, as a visual aid, the hue of heatmap grid cells was changed to correspond to Voronoi cell they fall into (e.g., sector 2 in orange). Same hue shading applies to the sectors consistently in consequent heatmaps. Figure 7 shows the heatmaps with 'all prototypes' for each scenario. The legend attached to each heatmap illustrates the heat measure for each of the sectors, i.e. the number of overlapping prototypes at that coordinate. Finally, the percentage of the heatmap's heat falling within the bounds of each sector is shown as a faint percentage overlay in each sector.

As a reference to the real-world context which influenced the participants' positioning of their prototypes, Figure 8 illustrates the resultant global cluster bounds as well as the sector centroids overlaid atop the respective panorama images. Comparing the two scenarios to each other, Figure J1 (left) in Appendix J visualises how the sector centroids differ, with a minor observable difference for sectors 3, 5 and 6. When considering scenario-agnostic acetates, Figure J1 (right) maps the change in locations of prototypes within the visual field (in φh, and φv), with radial distance r not adequately visualised on the 2D figure). Ascent placement is denoted with a square, descent by a circle, with a line connecting the two. Only five prototypes are observed to have crossed a sector bound. Illustrated in Appendix K are locations of prototypes' centre points, portraying the many (albeit likely smaller) prototypes in the upper sectors, and fewer (albeit likely compound) prototypes in the lower sectors.

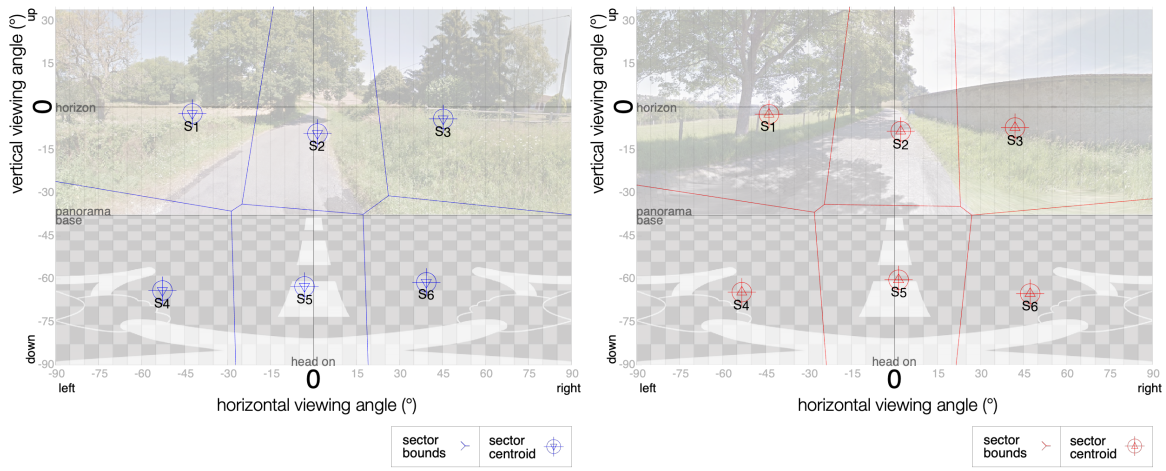

Figure 8: Equirectangular view of printed panoramas and bases, with global cluster bounds and centroids (descent left, ascent right)

5.3 Heatmap Occupancy Table

Based on the global cluster bounds, a Heatmap Occupancy Table (Table 1) was created with rows for various subsets, listing: prototype frequency, heatmap occupancy, and centre point count. Frequency refers to how many prototypes of a given subset were created by participants for each of the scenarios. The heatmap occupancy percentage is a measure of the specific heatmap's heat falling within the previously established sectors for each scenario. The centre point count is the absolute count of the prototype centre points falling within the sector. Occupancy values and centre point counts are presented as a two by three grid matching the approximate sector locations (S1-S3 in the top row, S4-S6 in the bottom row), and are shaded based on their percentage value in 10% increments. The table presents in rows select subsets of type-metric. The extended table with other subsets can be found in Appendix L. The legend (Appendix M) illustrates the relationship between the table and the heatmaps it is based on (for example: in the 'all prototypes' set, 122 prototypes were assigned to the ascent scenario, where sector 2 contained 28% of the heat, and 36 centre points).

5.4 Clustered Heatmaps

Presented is a small selection of the generated heatmap pairs (descent, ascent) for a subset of the collected data, where type-metric was navigation (Figure 9), containing a larger sample of prototypes and discernible patterns different from the entire dataset (Figure 7). Details on how to read the heatmaps can be found in Section 5.2, with remaining heatmaps found in Appendix N.

Most of the heatmap's heat is concentrated in sectors 2 (top-middle) and 5 (bottom-middle), where the prototypes are concentrated around the respective centroids. The highest value of a hotspot is 11 (descent) and 7 (ascent). There is a minor gap between said sectors. Sectors 6, 4, and 1, are sparsely used. There is a strong bias towards the centre horizontally. There is more heat found in the top sectors. The heatmap-pair is more populated in the downhill (23) than uphill (19) scenario, with the heat moving 'upward' to sector 2.

5.5 Ordered Lists and Mock-ups

As outlined in Section 4.4 (see Figures 4 and 5), the four apportionment methods are used to generate ordered lists. Each list defines the order of types-metric instances, which is then used to construct a visual-field mock-up. The order is significant as it determines priority and avoids collisions.

The Gantt charts in Appendix O illustrate the different resultant ordered lists, for sets: descent, ascent, and for scenario-inclusive (either scenario) layouts, as result of the chosen assortment method. Across methods, the lists contain the same number of items for each context, with 8 for descent, 10 for ascent and 11 for scenario inclusive. However, the lists differ in their number of unique items, prioritisation, and how repeated occurrences of instances are managed.

Using the ordered lists for descent and ascent, mock-ups were generated. For illustration purposes, presented in Figure 10 are mock-ups from the Adaptive Averaging Method (mock-ups of other lists are in the Appendix P). This figure shows that in the descent speed is placed as first, and again as fourth, while distance is second and seventh. The coordinates of each instance can be found in Appendix P.

6 Discussion

Building on the results reported in Section 5, this concluding section situates those findings and reflects on their implications. Section 6.1 reflects on the process, how participants prototyped their user interfaces, acknowledging methodological and practical constraints, and suggesting potential improvements to this approach. Section 6.2 reflects on how the design guidelines could apply to real use in context. Future Work (6.3) outlines promising extensions of this research. Lastly, Section 6.4 summarises the core contributions and research take-aways.

Table 1: Heatmap Occupancy – Differences between two scenarios in 'prototype frequency', 'heatmap occupancy' and 'centre point count', for 'all prototypes' and across data subsets. Cells are shaded based on percentage value (see legend, Appendix M). For full detail see Section 4.3. (*heatmaps shown in Section 5.4)

	prototype frequency			heatmap occupancy (%)						centre point count					
				descent			ascent			descent			ascent		
subset	descent	ascent	either	S1 S4	S2 S5	S3 S6	S1 S4	S2 S5	S3 S6	S1 S4	S2 S5	S3 S6	S1 S4	S2 S5	S3 S6
all prototypes (entire set)	110	122	141	15 8	30 16	22 10	15 7	28 19	23 8	23 7	34 5	26 8	24 7	36 8	27 8
type-metric: distance	20	23	25	8 12	11 38	18 13	10 8	10 41	23 8	3 1	4 4	5 1	4 1	5 4	6 1
type-metric: speed	23	20	23	4 2	13 31	28 22	5 3	7 40	30 15	2 0	6 3	8 4	2 0	4 4	7 3
type-metric: navigation*	23	19	23	5 4	55 24	7 4	4 5	44 35	10 3	2 1	15 2	2 1	1 1	11 3	2 1
type-metric: power	11	15	15	18 3	3 54	12 10	13 3	24 48	10 3	5 0	1 3	2 0	5 0	3 4	3 0
type-metric: time	12	13	15	30 6	3 21	17 24	21 7	11 25	15 21	3 0	1 1	5 2	2 0	2 1	4 2
type-metric: gradient	5	14	14	10 0	3 51	29 6	12 2	21 36	28 0	1 0	0 2	2 0	3 1	4 2	3 0
type-metric: heart rate	11	13	13	17 21	6 45	0 11	17 18	10 47	4 5	3 4	1 2	0 0	3 3	3 2	1 0

(Table continued in Appendix L…)

6.1 Prototyping Methodology – Design as Verb

The methodology applied in this study relied on low-fidelity prototyping materials and data collection, while the subsequent data processing employed a rigorous, systematic, high-fidelity pipeline (classifications, scenarios, heatmaps, apportionment). Said pairing is rare, as most low-fidelity studies stop at qualitative themes, and spatial analysis is the domain of high-fidelity simulations. Here, it enabled safe, creative, and technology-agnostic exploration while producing quantitative insights.

This contrast naturally raises questions of resolution and generalisability of the gathered insights. The spatial trends revealed through heatmaps rest on structured aggregation and minor interpretive assumptions. However, the low-fidelity approach arguably encouraged broader and longer participation, and more spontaneous design choices driven by user need rather than technology push.

The challenge of applying MR in road bicycle racing is great. Owing to the fast-changing environment and perilous speeds, the use of MR alone is hazardous without safety-conscious design guidelines, in-situ prototyping is dangerous and impractical. Conversely, developing a high-fidelity cycling simulator is a complex puzzle, as vestibular forces are extremely difficult to faithfully recreate in a lab. The objectives of prototyping feasibility and simulation immersion are at complete odds. Here, a low-fidelity proxy was necessary, and only a careful methodological balance between ecological relevance and feasibility enables a bridge between participatory design and fair analysis. Future work with higher-fidelity tools (e.g., VR, MR) could always further refine or validate the observed patterns. The results should therefore be seen as indicative of conceptual preferences, and not tested performance.

The methodological balance did produce insights which are scenario-specific, and transferable. While developed for road bicycle racing, the protocol, the apparatus, embodied placement, and structured spatial data analysis could be adapted to other high-motion sports or other contexts where safety, situational awareness, and visual-field management are critical – e.g., crane operators, forklift drivers, or factory floor workers.

In practice, the addition of the adjacent, looping video to the static panoramic was designed to foster a stronger sense of immersion, by balancing spatial continuity with a wider temporal window of the environment. It remains unclear to what extent this setup approximated the demands of real-world cycling, or how it influenced placement of prototypes. Anecdotal evidence of participants repositioning prototypes when reminded to assume the cycling position during placement points to significance of the embodied body

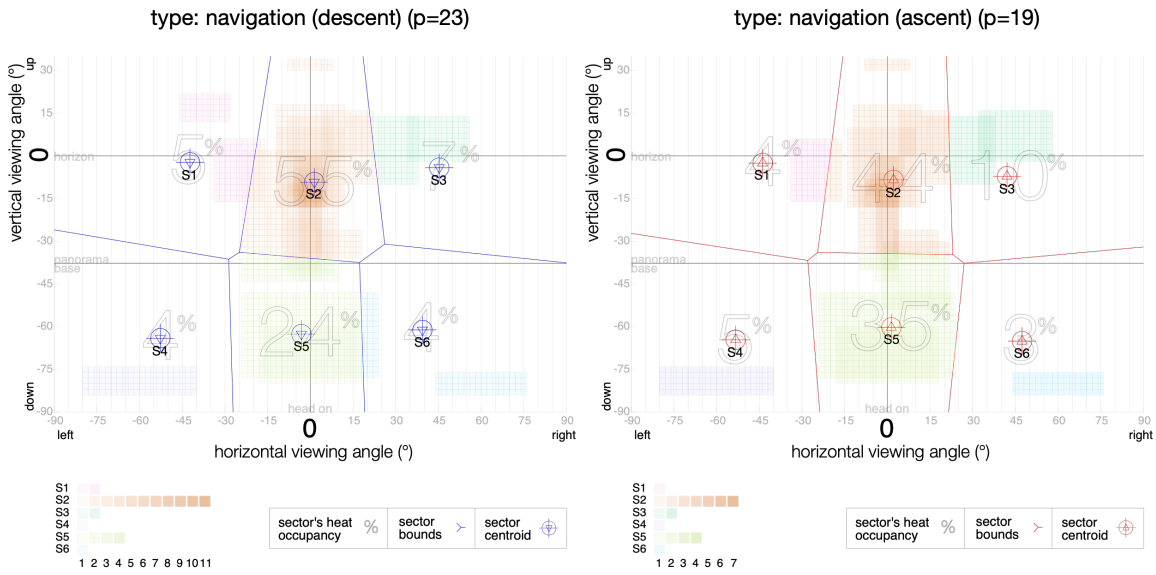

Figure 9: Heatmaps of prototype placements across the mapped visual field (x: −90° to 90°; y: −90° to 34°) for prototype of type-metric: navigation in two scenarios: descent (left), and ascent (right). Global cluster bounds for sectors are outlined, with sector centroids indicated (S1-S6). Sectors are visually differentiated by colour, with percentage heatmap occupancy displayed for each.

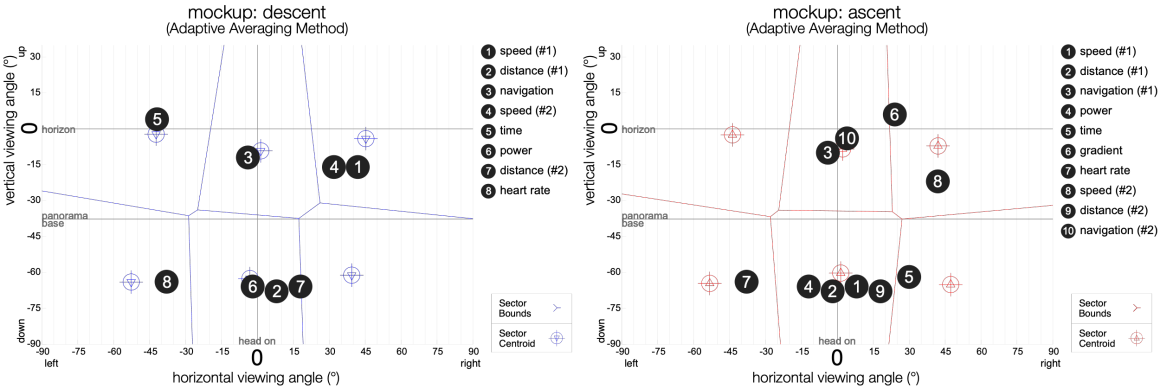

Figure 10: Mock-ups of locations of prototypes according to apportioned order of Adaptive Averaging Method, for descent in blue (above), ascent in red (below) (coordinates found in Appendix P). Also shown sector centroid coordinates of sectors 1-6 separated by global cluster bounds.

posture during prototyping. Future studies could investigate how a higher degree of environmental fidelity affects design choices.

Designing a dynamic interface with static artefacts (office stationery) is challenging. While it does allow the luxury of time with less fatigue compared to more technologically reliant methods, participants grappled with taking on the task of generalising their choices, and presenting or exploring varying states of a prototype (e.g., circumstance dependent colour changes). However, as colour was not the focus, the value of positioning guidelines stands.

During the sessions several future protocol improvements came to the fore. Participants who designed at a Prototyping Station first, estimated the desired size, and later 'scaled' the items by moving them closer or further within the 3D-canvas space. This made the position data of prototypes in terms of the 'proximity to the eyes' (focal length) of lesser value and was thus omitted. With respect to choosing a colour, participants prototyped on a separate, bright wood veneer desk, which may have adversely influenced the choice of colours. Participants often combined multiple classes in a single prototype acetate cut-out. Such compound prototypes overrepresent each classification in their respective heatmaps due to their combined size, and should have been split in advance.

The study possibly exhibits self-selection bias. Prospective participants uninterested in innovative technologies are less likely to enlist in a study enlisting help designing for smart glasses. In the

end, results did however show attendance of a few MR sceptics who used it sparingly. A challenge when inviting lay participants in design to prototype, is for them to step out of their own experience, and frame of reference, as illustrated by the paraphrased quote: 'I know what I mean I just did not draw it right.'. Participants were reminded of their expert position, not being time-bound, and of their contribution being appreciated, although it is unknown whether this lowered the artistic bar sufficiently. Nonetheless, the presented trade-offs of visual-field real estate remain valid, regardless of their level of polish.

Participants varied in experience and equipment. The use of a power meter is not prevalent amongst amateurs. As a stand-in for tracking active performance, instead, they rely on other metrics, such as speed or heartrate. This may have led to a splitting the frequency count of a hypothesised performance metric, and by extension affected the apportionment and subsequent mock-ups.

Participants seldom relied on sound (only 14 of out 141 prototypes) as a sole medium or complimentary to a visual prototype. However, it is unclear whether this was due to lack or shortcomings of tooling provided, insufficient focus in the study protocol, or lack of demand among participants.

6.2 Interpretation of the Results – Design as Noun

Despite the low-fidelity approach, the design guidelines generated are unusual in their quantitative grounding. The use of heatmap visualisations (e.g., Figure 7) revealed distinguishable visual-field sectors as preferred by participants, as well as their associated preferences for certain information types. The process moves seamlessly from design-as-verb to design-as-noun, that is from participatory design sessions conducted by usability researchers with expert-users to deployable design artefacts for design practitioners. The data processing pipeline is systematic, and transparent, allowing one to trace the origin of the outcome, while leaving sufficient room for interpretation but not inviting bias or ambiguity. Placement zones (sectors) derived from the aggregated preferences can be used as constraints in Adaptive User Interfaces, as explored in future work.

Across the ascent and descent scenarios, participants preferred similar information positioning but not always for the same type of information. For the scenario-agnostic prototypes, placement of prototypes was rarely adjusted between the two contexts. As regards the differences between scenarios, context-specific inclusion of additional information (e.g., elevation graphs, motivational cues) was observed. Whether this is due to the static imagery for panorama backgrounds, is unclear.

Data was not collected on the position of the road gaze point. Assuming a prioritised placement of information, and that it would fall equidistant from cluster centroids S1 (top-left), S3 (top-right) and S5 (bottom-middle), with this triangle straddling S2 (top-middle), the gaze point would fall at (-1.5, -16.0) in the descent, and (1.3, -16.9) in the ascent (see Appendix Q). The distance to each sector centroid is 45.9° in the descent and 44.4° in the ascent. Said gaze points, as well as sectors 1-6 are shown in side-view in context in Figure 11. Contemporary Mixed Reality glasses do not offer sufficient field of view (approximately 20-40°) to display such a range without cropping. In practice, this hardware limitation would require the participant to actively seek information in place of relying on their peripheral vision, rendering alert capabilities of these sectors void, and diminishing the value of the user interface.

The choice for a cylindrical panorama design, as 3D canvas and visual backdrop, may have impacted the positioning of prototypes, as the boundary of panorama and base closely matches the boundaries of sectors 1 and 4 as well as sectors 3 and 6 (Figure 8). Conversely, boundary of sectors 2 and 5 was not affected as much. The dimensions of the cylinder (and by extension the panorama-base boundary) were a balanced trade-off between the height limitation of available large-format printing (A0 spool width of 840mm), minimal panorama canvas proximity (further than arm's length), and the choice to maximise the use of the panorama for the varied backdrop of road surroundings (with the largely uniform road surface befalling to the panorama base). There are very few (7) prototypes placed above 30 degrees vertically. It is unclear whether this is a consequence of the panorama size chosen.

With respect to radial distance, Figure H1 shows many participants embracing the extra spatial dimension afforded to them by the hypothetical smart glasses they were tasked to design for. Some participants (e.g., 989, 679, 408) were further observed to have their prototypes clumped at designated radial planes, with the approximate gap of a single standard deviation (10-20cm). Changes in the radial distance between the two scenarios were not investigated.

As regards Apportionment and mock-ups, the Gantt chart (Appendix O) highlights differences in the allocation of slots between the four considered methods. For each method's lists mock-ups were made (Figure 10, and Appendix P). The ultimate choice for the method falls on a closer consideration of the appropriateness and 'fair fit' with the weights provided. The Greedy Fair Allocator ensures quick apportionment of weight-heavy items. Conversely the Weighted Round Robin risks prematurely including marginal items and only repeats weight-heavy items late. At which point the mock-ups are largely populated, and second instances of popular type-metric are unlikely to be left with many suitable hot spots from their respective heatmaps. The Greedy Subtractive Method and the Adaptive Averaging Methods attempt to balance the pacing of introducing second instances, continuously addressing a trade-off between a repeated instance of a popular type-metric versus a marginal one (for consideration, a hypothetical example of selecting three items from a list with type-metric: speed with a weight 1.6 and hydration with 0.501). Ultimately the Adaptive Averaging Method was seen as most fair with the weights provided and selected for Figure 10. Approaching the problem with multiple methods bridges the quantitative precision with design-level nuance.

The generated mock-ups illustrate the large differences as consequent of small margins, resultant from frequency- and heatmap analysis. Instances of speed and time come closer to the horizon on the descent, arguably best illustrating balancing of safety and performance. Placement of the gradient off-centre above the horizon on the ascent seems fitting, as it pertains to a dreaded piece of information about the height of the current climb. Finding navigation be the only occupant of sector 2, especially close to the gaze point, places it where it is most relevant – the road ahead. It can however be interpreted differently depending on the context (close to the action during a careening descent, disposable space during a

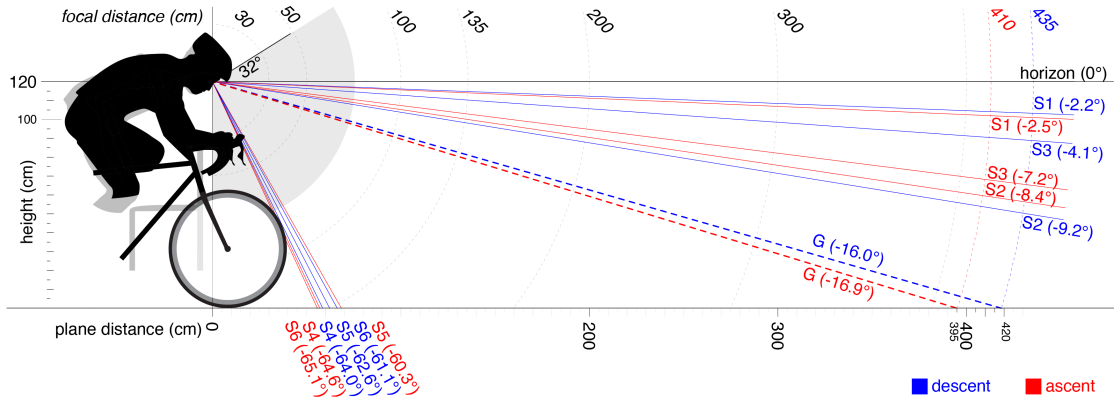

Figure 11: Side view of sectors 1-6 and 'gaze points' (G) for both ascent (red) and descent (blue). Also, see Figures 2, J1, Q1.

slow climb), and closer investigation is needed into the specifics of its application.

Lastly, although limited by the paper's 2D platform as a visualisation tool, the 3D mock-up can be completed by merging mock-ups' radial coordinates (Appendix P) with radial distance recommendations (Figure 11).

6.3 Future Work

The detailed insights gained in this study form a start towards a set of design guidelines for application of Mixed Reality in road cycling. A lab study with deeper immersion was warranted, that builds upon the present work, putting prototyping tools in the participants' hands. In this follow-up study [51], richer visual reference was provided in the form of VR dynamic backdrops of the scenario contexts, with added vestibular and tactile sensation achieved through participants riding a tandem bicycle on a treadmill. The comparison of the intermediate (heatmaps) and eventual (UI mock-ups) findings is in preparation, shedding light on the potential impact the tools have on the user insights obtained. Furthermore, a data driven design framework [52] is under development that enables prototyping UIs in a web browser, based on observation data from past studies (this paper). Improving upon the present methods, it is designed to negotiate a balanced use of visual real estate through force-driven mechanisms. Lastly, the resulting UI concepts will be evaluated by participants in the field.

6.4 Distilled Findings

6.4.1 Involve your users in the design process early, no hardware or programming necessary. In a two-part creative session, focussing on uphill and downhill scenarios, participants prototyped an MR user interface using office stationery against static panoramic backdrops. When placing each acetate cut-out in 3D-space, participants wilfully made trade-offs between the value of the conveyed information and the cost to the visual-field real estate, sharing their personal preferences for what information they desired, how it is to be delivered, and its placement in 3D-space. Aside from involving end-users at a much earlier stage of the development process, the low fidelity approach offers numerous benefits like lower cost, access (hardware or location), no technical skill requirement, and possibility to conduct a user study remotely. In contrast, a high-tech equivalent might involve MR or VR hardware, an approach offering higher realism and immersion but introducing higher development overhead, as well as reduced safety, accessibility and iterability. Simulator sickness and discomfort should also be considered, and the impact thusly shortened creative session would have on the validity of the results. Nonetheless, this study revealed a range of classes (classified responses) for the parameters of type-metric and type-qualifier as well as a shortlist of colours embraced in prototyping, which can be used in development of a high-fidelity prototyping environment, where finite choice lists enable quicker prototyping.

6.4.2 There is room for safety, performance, and creativity. Results show participants rely on external real-time processed information to complete a cycling goal in a race, with regards to both safety and performance. There is however a tension between the real-world information and the augmentation layers, as evidenced by distinct visual-field areas being preserved, and others being favoured for sacrifice. To this end, heatmaps provide a useful means to analyse the design process and make it possible to process and present spatial preference data.

6.4.3 Most desired type-metrics are navigation, speed, and distance. Analysis shows differences and commonalities between participants' selection of desired information, with an average of 6.1-6.8 prototypes used at a time depending on the context, amongst them most commonly: distance, speed, and navigation, followed by collective performance-related metrics of power and heartrate. The exact placement of the individual instances depends on the type of data, its purpose, and how it is presented.

6.4.4 Context does matter. The choice of scenario had moderate impact on the choice of data types, tailored to the context's specific goals and challenges. While the centroid locations of the sector did not move much, the apportioned ordered lists differed for the two scenarios, and by extension the recommended layouts – illustrated in Figure 10, with speed and time portrayed in vastly different locations.

6.4.5 Mind the locations of all three dimensions. Analysis of the clustered heatmaps shows six sectors (illustrated with centroids and

boundaries in Figure 8), with their detailed utilisation per subset outlined in the Heatmap Occupancy Table (Table 1). The sectors form a grid of three columns and two rows, with centres approximately at 5 and 63 degrees vertically (down from the horizon), and -43, 2 and 45 degrees horizontally (from head-on direction). Sectors furthest away from the vanishing point (4 and 6; bottom-left and bottom-right) saw the least use. The gaps between sectors are most pronounced between top and bottom rows, especially during descent. Results show moderate experimentation with the radial distance (Figure H1) despite the limitations of the experimental setup.

7 Conclusions

The methodology presented in this user study focuses on the low-fidelity extraction and high-fidelity data analysis and processing of user preferences from road cycling athletes for user interface and information display in the application of see-through-display MR technology in road bicycle racing. User preferences are condensed in heatmap visualisations and mock-ups of optimal layouts for descent and ascent in road cycling races. Participants consistently prioritised inclusion of navigation, speed, and distance, and placed them in six distinct regions of their visual field. Figure 10 proposes distinct user-interface mock-ups for the two scenarios as a design guideline for future reference. This is achieved by a means of an innovative methodology for data acquisition, analysis, and processing. This methodology can be adapted for similar User Interface user studies aimed at other MR applications for assisted decision making while in motion in dynamic environments. Listed in Section 6.4 (above) are distilled design guidelines.

Acknowledgments

This study was funded as part of the 'Citius Altius Sanius' research programme (P16-28), financed by the Dutch Research Council (NWO). The authors would like to thank dr. Alexis Derumigny and dr. Laurens Rook for brainstorming assistance.

References

[1] Nedal Sawan, Ahmed Eltweri, Caterina De Lucia, Luigi Pio Leonardo Cavaliere, Alessio Faccia, and Narcisa Roxana Moşteanu. 2021. Mixed and Augmented Reality Applications in the Sport Industry. In Proceedings of the 2020 2nd International Conference on E-Business and E-Commerce Engineering (Bangkok, Thailand) (EBEE '20). Association for Computing Machinery, New York, NY, USA, 55–59. https://doi.org/10.1145/3446922.3446932

[2] Stefan Gradl, Bjoern M. Eskofier, Dominic Eskofier, Christopher Mutschler, and Stephan Otto. 2016. Virtual and Augmented Reality in Sports: An Overview and Acceptance Study. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (Heidelberg, Germany) (Ubicomp '16). Association for Computing Machinery, New York, NY, USA, 885–888. https://doi.org/10.1145/2968219.2968572

[3] Don Samitha Elvitigala, Armağan Karahanoğlu, Andrii Matviienko, Laia Turmo Vidal, Dees Postma, Michael D Jones, Maria F. Montoya, Daniel Harrison, Lars Elbæk, Florian Daiber, Lisa Anneke Burr, Rakesh Patibanda, Paolo Buono, Perttu Hämäläinen, Robby Van Delden, Regina Bernhaupt, Xipei Ren, Vincent Van Rheden, Fabio Zambetta, Elise Van Den Hoven, Carine Lallemand, Dennis Reidsma, and Florian 'Floyd' Mueller. 2024. Grand Challenges in SportsHCI. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 312, 1–20. https://doi.org/10.1145/3613904.3642050

[4] Buchner, J., Buntins, K., & Kerres, M. (2022). The impact of augmented reality on cognitive load and performance: A systematic review. Journal of Computer Assisted Learning, 38(1), 285-303. https://doi.org/10.1111/jcal.12617

[5] Schuler, Douglas, and Aki Namioka, eds., 1993, Participatory Design: Principles and Practices, Lawrence Erlbaum Associates, Hillsdale, NJ. https://doi.org/10.1201/9780203744338

[6] Maximilian Speicher, Brian D. Hall, and Michael Nebeling. 2019. What is Mixed Reality? In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). Association for Computing Machinery, New York, NY, USA, Paper 537, 1–15. https://doi.org/10.1145/3290605.3300767

[7] Milgram, P., Kishino, F. (1994). A taxonomy of mixed reality visual displays. IEICE TRANSACTIONS on Information and Systems, 77(12), 1321-1329.

[8] Nebeling, M., & Speicher, M. (2018, October). The trouble with augmented reality/virtual reality authoring tools. In 2018 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-Adjunct) (pp. 333-337). IEEE. https://doi.org/10.1109/ISMAR-Adjunct.2018.00098

[9] Ashtari, N., Bunt, A., McGrenere, J., Nebeling, M., & Chilana, P. K. (2020, April). Creating augmented and virtual reality applications: Current practices, challenges, and opportunities. In Proceedings of the 2020 CHI conference on human factors in computing systems (pp. 1-13). https://doi.org/10.1145/3313831.3376722

[10] Nebeling, M., & Madier, K. (2019). 360proto: Making interactive virtual reality & augmented reality prototypes from paper. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-13). https://doi.org/10.1145/3290605.3300826

[11] Nebeling, M., Nebeling, J., Yu, A., & Rumble, R. (2018). ProtoAR: Rapid physical-digital prototyping of mobile augmented reality applications. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 1-12). https://doi.org/10.1145/3173574.3173927

[12] Speicher, M., Lewis, K., & Nebeling, M. (2021). Designers, the stage is yours! medium-fidelity prototyping of augmented & virtual reality interfaces with 360theater. *Proceedings of the ACM on human-computer interaction*, *5*(EICS), 1-25. https://doi.org/10.1145/3461727

[13] Michael Nebeling, Katy Lewis, Yu-Cheng Chang, Lihan Zhu, Michelle Chung, Piaoyang Wang, and Janet Nebeling. 2020. XRDirector: A Role-Based Collaborative Immersive Authoring System. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376637

[14] Lee, G. A., Nelles, C., Billinghurst, M., & Kim, G. J. (2004, November). Immersive authoring of tangible augmented reality applications. In Third IEEE and ACM international symposium on mixed and augmented reality (pp. 172-181). IEEE. https://doi.org/10.1109/ISMAR.2004.34

[15] Blair MacIntyre, Maribeth Gandy, Steven Dow, and Jay David Bolter. 2004. DART: a toolkit for rapid design exploration of augmented reality experiences. In Proceedings of the 17th annual ACM symposium on User interface software and technology (UIST '04). Association for Computing Machinery, New York, NY, USA, 197–206. https://doi.org/10.1145/1029632.1029669

[16] Mast, Danica, Alex Roidl, and Antti Jylha (2023). "Wizard of Oz Prototyping for Interactive Spatial Augmented Reality in HCI Education: Experiences with Rapid Prototyping for Interactive Spatial Augmented Reality." Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3544549.3573861

[17] Wunsch, Matthias (2021). Evaluation of Interactive Systems for Cyclists Using Wizard of Oz Prototypes In-The-Wild. For: Cycling@CHI: Towards a Research Agenda for HCI in the Bike Lane. In CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '21 Extended Abstracts), May 8–13, 2021, Yokohama, Japan. ACM, New York, NY, USA.

[18] Lee, M., & Billinghurst, M. (2008). A wizard of Oz study for an AR multimodal interface. In Proceedings of the 10th international conference on Multimodal interfaces (pp. 249-256). https://doi.org/10.1145/1452392.1452444

[19] Atkinson, Greg, *et al.*, 2003, Science and cycling: current knowledge and future directions for research, Journal of sports sciences, vol. 21, no. 9, pp. 767-787. https://doi.org/10.1080/0264041031000102097

[20] Sanders, Dajo, Teun van Erp, and Jos J. de Koning, 2019, Intensity and Load Characteristics of Professional Road Cycling: Differences Between Men's and Women's Races, International journal of sports physiology and performance, vol. 14, no. 3, pp. 296-302. https://doi.org/10.1123/ijspp.2018-0190

[21] van Erp, Teun, Carl Foster, and Jos J. de Koning, 2019, Relationship Between Various Training-Load Measures in Elite Cyclists During Training, Road Races, and Time Trials, International journal of sports physiology and performance, vol. 14, no. 4, pp. 493- 500. https://doi.org/10.1123/ijspp.2017-0722

[22] Wolf, Stefan, Francesco Biral, and Dietmar Saupe, 2019, Adaptive feedback system for optimal pacing strategies in road cycling, Sports Engineering, vol. 22, no. 1, p. 6. https://doi.org/10.1007/s12283-019-0294-5

[23] Dukalski, R., Lukosch, S., Schwab, A., Beek, P. J., & Brazier, F. M. (2020, June). Exploring the effect of pacing plan feedback for professional road cycling. In *Proceedings* (Vol. 49, No. 1). MDPI. https://doi.org/10.3390/proceedings2020049058

[24] UCI. The appeal of road cycling: Findings from the public consultation (19 August 2019). Retrieved April 23, 2025 from https://www.uci.org/pressrelease/the-appeal-of-road-cycling-findings-from-the-public-consultation/1Dg381u3NmKgFeXISHjT4r

[25] Hamada, T., Okada, M., & Kitazaki, M. (2017, March). Jogging with a virtual runner using a see-through HMD. In *2017 IEEE Virtual Reality (VR)* (pp. 445-446). IEEE. https://doi.org/10.1109/VR.2017.7892371

[26] Hamada, T., Hautasaari, A., Kitazaki, M., & Koshizuka, N. (2020, March). Exploring the effects of a virtual companion on solitary jogging experience. In 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) (pp. 638-639). IEEE. https://doi.org/10.1109/VRW50115.2020.00170

[27] Hamada, T., Hautasaari, A., Kitazaki, M., & Koshizuka, N. (2022, March). Solitary jogging with a virtual runner using smartglasses. In *2022 IEEE conference on virtual reality and 3D user interfaces (VR)* (pp. 644-654). IEEE. https://doi.org/10.1109/VR51125.2022.00085

[28] Will, S., Wehner, T., Hammer, T., Merkel, N., Werle, A., Umlauf, I., & Neukum, A. (2022). Assessment of data glasses for motorcycle riders in a simulated lane change test. Transportation research part F: traffic psychology and behaviour, 89, 467-477. https://doi.org/10.1016/j.trf.2022.07.016

[29] Matviienko, A., Müller, F., Schön, D., Seesemann, P., Günther, S., & Mühlhäuser, M. (2022). BikeAR: Understanding cyclists' crossing decision-making at uncontrolled intersections using Augmented Reality. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1-15). https://doi.org/10.1145/3491102.3517560

[30] von Sawitzky, T., Grauschopf, T., & Riener, A. (2022). Hazard notifications for cyclists: comparison of awareness message modalities in a mixed reality study. In Proceedings of the 27th International Conference on Intelligent User Interfaces (pp. 310-322). https://doi.org/10.1145/3490099.3511127

[31] Kosch, T., Matviienko, A., Müller, F., Bersch, J., Katins, C., Schön, D., & Mühlhäuser, M. (2022). Notibike: Assessing target selection techniques for cyclist notifications in augmented reality. Proceedings of the ACM on Human-Computer Interaction, 6(MHCI), 1-24. https://doi.org/10.1145/3546732

[32] Ammons, Robert B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formulation, The Journal of general psychology, vol. 54, no. 2, pp. 279-299. https://psycnet.apa.org/doi/10.1080/00221309.1956.9920284

[33] Magill, Richard A., and David I. Anderson (2007). Motor Learning and Control: Concepts and Applications, McGraw-Hill, New York. ISBN13: 978-0-07-802267-8

[34] Andres, J., Kari, T., Von Kaenel, J., & Mueller, F. F. (2019). Co-riding with my eBike to get green lights. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (pp. 1251-1263). https://doi.org/10.1145/3322276.3322307

[35] Walmink, W., Wilde, D., & Mueller, F. F. (2014). Displaying heart rate data on a bicycle helmet to support social exertion experiences. In *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction* (pp. 97-104). https://doi.org/10.1145/2540930.2540970

[36] Postma, D. B., van Delden, R. W., Koekoek, J. H., Walinga, W. W., van Hilvoorde, I. M., van Beijnum, B. J. F., ... & Reidsma, D. (2022). A design space of sports interaction technology. *Foundations and Trends® in Human–Computer Interaction*, *15*(2-3), 132-316. http://dx.doi.org/10.1561/1100000087

[37] Robby van Delden, Sascha Bergsma, Koen Vogel, Dees Postma, Randy Klaassen, and Dennis Reidsma. 2020. VR4VRT: Virtual Reality for Virtual Rowing Training. In Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '20). Association for Computing Machinery, New York, NY, USA, 388–392. https://doi.org/10.1145/3383668.3419865

[38] Syiem, B. V., Kelly, R. M., Goncalves, J., Velloso, E., & Dingler, T. (2021). Impact of task on attentional tunneling in handheld augmented reality. In Proceedings of the 2021 CHI conference on human factors in computing systems (pp. 1-14). https://doi.org/10.1145/3411764.3445580

[39] Wickens, C. D., & Alexander, A. L. (2009). Attentional tunneling and task management in synthetic vision displays. The international journal of aviation psychology, 19(2), 182-199. http://doi.org/10.1080/10508410902766549

[40] Wang, Y., Wu, Y., Chen, C., Wu, B., Ma, S., Wang, D., ... & Yang, Z. (2022). Inattentional blindness in augmented reality head-up display-assisted driving. International Journal of Human–Computer Interaction, 38(9), 837-850. https://doi.org/10.1080/10447318.2021.1970434

[41] Liu, S., & Lindlbauer, D. (2024). TurnAware: motion-aware Augmented Reality information delivery while walking. Frontiers in Virtual Reality, 5, 1484280. https://doi.org/10.3389/frvir.2024.1484280

[42] Lindlbauer, D., Feit, A. M., & Hilliges, O. (2019). Context-aware online adaptation of mixed reality interfaces. In Proceedings of the 32nd annual ACM symposium on user interface software and technology (pp. 147-160). https://doi.org/10.1145/3332165.3347945

[43] Lages, W. S., & Bowman, D. A. (2019). Walking with adaptive augmented reality workspaces: design and usage patterns. In proceedings of the 24th international conference on intelligent user interfaces (pp. 356-366). https://doi.org/10.1145/3301275.3302278

[44] Jan Gugenheimer, Mark McGill, Samuel Huron, Christian Mai, Julie Williamson, and Michael Nebeling. 2020. Exploring Potentially Abusive Ethical, Social and Political Implications of Mixed Reality Research in HCI. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20). Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3375180

[45] Jan Gugenheimer, Wen-Jie Tseng, Abraham Hani Mhaidli, Jan Ole Rixen, Mark McGill, Michael Nebeling, Mohamed Khamis, Florian Schaub, and Sanchari Das. 2022. Novel Challenges of Safety, Security and Privacy in Extended Reality. In Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22). Association for Computing Machinery, New York, NY, USA, Article 108, 1–5. https://doi.org/10.1145/3491101.3503741

[46] Veronika Krauß, Pejman Saeghe, Alexander Boden, Mohamed Khamis, Mark McGill, Jan Gugenheimer, and Michael Nebeling. 2024. What Makes XR Dark? Examining Emerging Dark Patterns in Augmented and Virtual Reality through Expert Co-Design. ACM Trans. Comput.-Hum. Interact. 31, 3, Article 32 (June 2024), 39 pages. https://doi.org/10.1145/3660340

[47] Forgy, E.W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. Biometrics. 21 (3): 768–769. JSTOR 2528559.

[48] Lloyd, S. (1982). Least squares quantization in PCM. *IEEE transactions on information theory*, *28*(2), 129-137. https://doi.org/10.1109/TIT.1982.1056489

[49] Thorndike, R. L. (1953). Who belongs in the family?. *Psychometrika*, *18*(4), 267-276. https://doi.org/10.1007/BF02289263

[50] Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, *20*, 53-65. https://doi.org/10.1016/0377-0427(87)90125-7

[51] Dukalski, R.R., Moore, J.K, Beek, P.J, & Brazier, F.M, (2025) Discovering Road Cyclists' Needs and Preferences for Mixed Reality User Interfaces Using Immersive Simulation. HCI International 2025, VAMR2025, Gothenburg, Sweden. To appear in HCII 2025 - Late Breaking Papers. Forthcoming.

[52] Dukalski, R.R., Beek, P.J, & Brazier, F.M, (2025) Data Driven Design Framework for Adaptive Mixed Reality User Interfaces for Road Cycling and Beyond. Submitted to SportsHCI 2025. https://doi.org/10.1145/3749385.3749414 Forthcoming.

Appendix A – Apparatus

Table A1: Tooling Station apparatus

Item	Notes
Table	Standard office table
Acetate transparency sheets	A3 and A4, clear, thick, 240 micron
Scissors	-
Telescopic stilts with flexible alligator clips	13-65cm length (seen in Figure 2)
Markers	approx. 100 water-colour markers, full colour spectrum, with black and dark grey colours removed
White paint correction-pens	
Tablet on a stand	playing the context-immersive video on loop (seen in Figure 1)
'Inspiration-sheet'	printed A4 sheets, (see Appendix B) 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
'Facilitator's Information type checklist'	printed A4 page with a printed list of line separated items: 'power, distance, time, heartrate, cadence, speed, energy, nutrition, hydration, navigation/map, altitude/elevation, weather, danger, lap/stage'

Table A2: Prototyping Station apparatus

Item	Notes
Table	Standard office table
Panorama (2x)	2x Semi-cylindrical 180° printed panorama photograph (Figures 1 and 2; 60cm radius, 82cm tall; 'horizon' at 45cm from the table surface), atop a printed semi-circular Base
Base	Semi-circular print representing the road surface – an asphalt-grey semi-circle, featuring a radial grid (12 rings of 5cm width, 36 radial sections of 5°) (with dashed line road markings, and a vector illustration of a bicycle handlebar gripped by gloves)
Chair	facing backwards, backed against the table (see Figure 2)
Nose calibration stilt	40cm tall (see Section 3.3 and Figure 2)

Appendix B – Question List

1. Did you cycle here to our meeting?
2. When's the last time you cycled?
3. Do you own a racing bike?
4. When's the last time you took your racing bike out?
5. How often do you race with it, in number of times per month?
6. How many km would you say you've cycled in the last 12 months?
7. What's the most times you've cycled on your racing bike in a month?
8. What's the most km you've cycled in 12 months?
9. What's your favourite route?
10. What's your dream route?
11. Do you watch races often, on TV (number of races per year)?
12. Which medium or TV channel do you follow races on? (local examples given, e.g. Eurosport, NOS, Sporza)
13. Have you attended a cycling race in person?
14. Are you a member of a cycling club or group?
15. Do you race as a team often?
16. Do you collect data when you cycle?
17. What data do you collect?
18. Which app do you use?
19. Which sensors?
20. Do you use a home trainer? Do you use Zwift or similar applications?
21. Do you have a bike computer/monitor/smartwatch to view the data back live?

22. Do you rely on this live data during race or training?
23. How do you rely on live data, motivation/goal setting/injury prevention?
24. Where do you stand on the use of data in cycling racing?
25. How do you see the future, 10, 20, 50 years, how will the sport of cycling look?
26. What will interfaces look like, the way that participants consume instructions, data, information

Appendix C – 'Inspiration Sheets'

Figure C1: 'Inspiration Sheets'

Appendix D – Example of Unprocessed Data

Table D1: Example of unprocessed data collected – field names (left), data (right)

Parameter	Data
Participant id	787
Scenario id	2
Spoken explanation	(redacted)
Time of creation	(redacted)
Sketch	
Size (height, width)	9, 13 (cm)
Placement: (cylindrical): descent (ρ, φ_h, z)	-, -, -
Placement: (cylindrical): ascent (ρ, φ_h, z)	17.5cm, 35°, 41cm
Notes	blue (with white outline) 7% current gradient, blue (white outline, green) altitude ahead

Appendix E – Output of Data Preparation

Table E1: Structure of the output of the Data Preparation phase: parameters of a processed prototype

Parameters	Notes
Participant-id	pre-assigned number
Prototype-id	format: (participant-id).number
Scenario-id	1 (descent) / 2 (ascent) / 3 (both)
Type-metric	e.g., cadence, distance, gradient, heartrate, power, speed
Type-qualifier	optional; e.g., elapsed, remaining, current, expected
Purpose	information / instruction / warning / motivation
Representation	number / graph / map / icon / sound / other
Centroid position (cylindrical): descent (ρ, φ_h , z),	ρ (distance; nearest 2.5cm from radial grid centre-point); φ_h (azimuth; nearest 2.5°
Centroid position (cylindrical): ascent (ρ, φ_h , z)	from 'head on'); z (height off the table) (nearest cm)
Centroid position (polar): descent (φ_h , φ_v , r),	φ_h (horizontal) and φ_v (vertical) in °; r (radial distance) in cm; gaze origin is at
Centroid position (polar): ascent (φ_h , φ_v , r)	(0,0,0)
Visual field size (polar): height	in °
Visual field size (polar): width	in °
Unit-explication	yes / no
Colour	single-word approximants, e.g., blue, brown, green, grey, orange, pink, purple, red, white, yellow

Appendix F – Results, Extended

As regards representation, amongst the 141 acetates, 70 were classified as number, 21 as graphs, 13 as maps, 51 as icons. In addition, 14 were classified as represented through sound. Unit explication was recorded in 69 acetates. In terms of purpose, 73 acetates were of information class, 33 instruction, 35 warning, and 39 motivation.

With respect to the use of colour, in 81 (60%) cases participants used a single colour, which did not change with respect to value or threshold. When multiple colours were used, different colours were used for power, heartrate and speed (type) classes and the instruction (purpose) class. As regards the frequency of use of colour, blue was used on 50 acetates, red in 47, green in 33, followed by white (27), orange (16), yellow (14), pink/purple (12; combined), brown (9), and light grey, which was used only once. Participants tended to use a single unifying 'neutral' or 'theme' colour: white, blue, orange, brown, pink, or purple. One participant used a rectangle with a background colour for reported emphasis, another participant employed white outlines on text and numbers for reported improved legibility, and yet another participant drew a multi-coloured digital 'rear view mirror'.

Among the 91 scenario-independent acetates, there were scenario-specific differences for placement (and by extension coordinates and visual field size), which was modified between the two scenarios by 7 participants on 24 acetates (26%). The participants used the telescopic stilts for 113 acetates, 9 of whom laid the 14 acetates flat on the table.

Appendix G – Instance Frequency Table

Table G1: Instance Frequency Table (scenario-agnostic) – Frequency of instances per type-metric class and participant, i.e. participant 378 created 22 instances on 7 prototype acetate cutouts, four of which were classified as type-metric: distance (rows sorted by total frequency per type-metric; columns sorted by total instances per participant; cells shaded by frequency)

		Participant id																		total	avg.	cv
		378	775	930	136	192	423	857	072	679	599	787	241	408	502	404	989	822	903			
Instances per type-metric class	speed	2	2	2	3	2	3	3	1	1	2	1	2	1	3		1	1	1	31	1.72	0.52
	distance	4	3		3	2	1	2	1	2	3	2	1	1	2		1	2		30	1.67	0.68
	navigation	2	1	1	1	1		1	3	2	2	1	1	1	1	1	1	1	2	23	1.28	0.52
	power	4			2	1	1		1	2		1	1	2		3			1	19	1.06	1.10
	time	2		3		2	2		2		1		1			4	1	1		19	1.06	1.15
	gradient	2	2	2	2		1	1			1	1		1	1		2			16	0.89	0.94
	heart rate	1		2		1	1	2	2	1		1		1	1		1	1		15	0.83	0.85
	elevation	3			1	1	1	1			1		1				1			10	0.56	1.41
	hydration					1		1		1		2		1	1			1	1	9	0.50	1.24
	cadence	2				1	1	1	1			1								7	0.39	1.56
	nutrition			1			1	1		1				1					1	6	0.33	1.46
	weather		1	1			1						2							5	0.28	2.07
	gears		2							1										3	0.17	3.09
	strategy		2		1															3	0.17	3.09
	braking					1			1											2	0.11	2.91
	lactate			2																2	0.11	4.24
	aerodyn.		1																	1	0.06	4.24
Total instances ®		22	14	14	13	13	13	13	12	11	10	10	9	9	9	8	8	7	6	201	11.2	0.33
Total acetates		7	10	10	9	9	7	3	11	9	9	10	8	8	6	5	6	10	4	141	7.8	0.29

Appendix H – Utilisation of Radial Distance

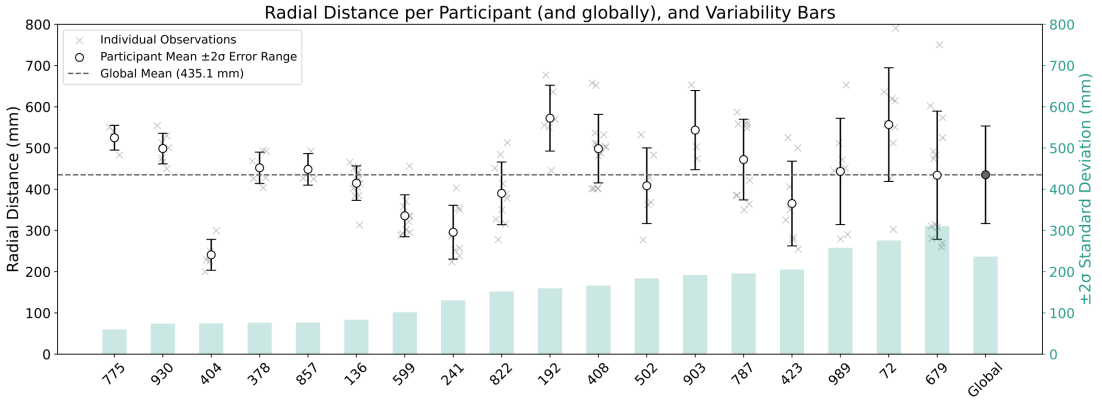

Figure H1: Utilisation of radial distance per participant; radial distance means per participant (with error bars) compared against the global mean radial distance (dashed); standard deviations of radial distance per participant (green bars) (sorted by standard deviation)

Appendix I – Plots: Elbow, Inertia, Silhouette

As seen in Figure I1, while the elbow was not pronounced, the inertia acceleration decreased after k=6, in both cases, while silhouette scores either peaked (ascent) or remained high (descent).

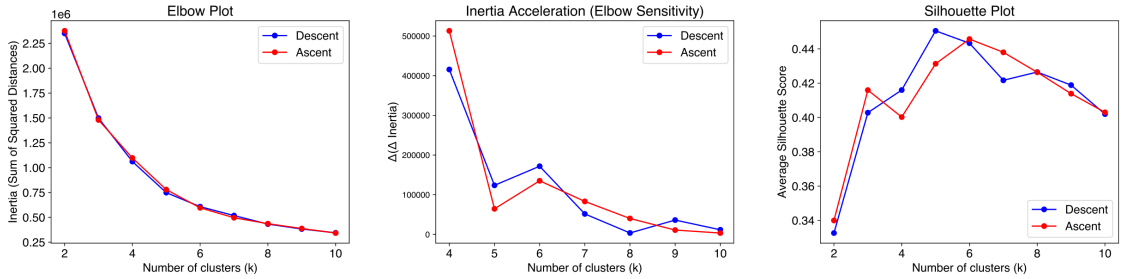

Figure I1: Elbow Plot (left), Inertia Acceleration (middle) and Silhouette Plot (right) for k-means clustering at different cluster counts, for descent (blue) and ascent (red) heatmap datasets

Appendix J – Global Cluster Coordinates

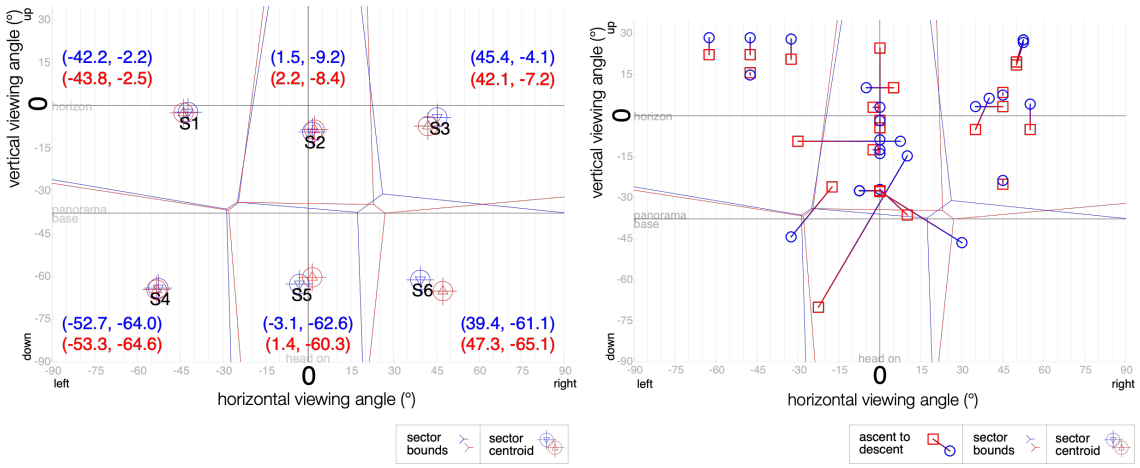

Figure J1: Sector centroid coordinates of sectors 1-6 separated by global cluster bounds (left) (descent in blue, ascent in red); Mapped locations of prototypes that have been moved between scenarios: ascent (red, squares) and descent (blue, circles).

Appendix K – Centre Points of Prototypes

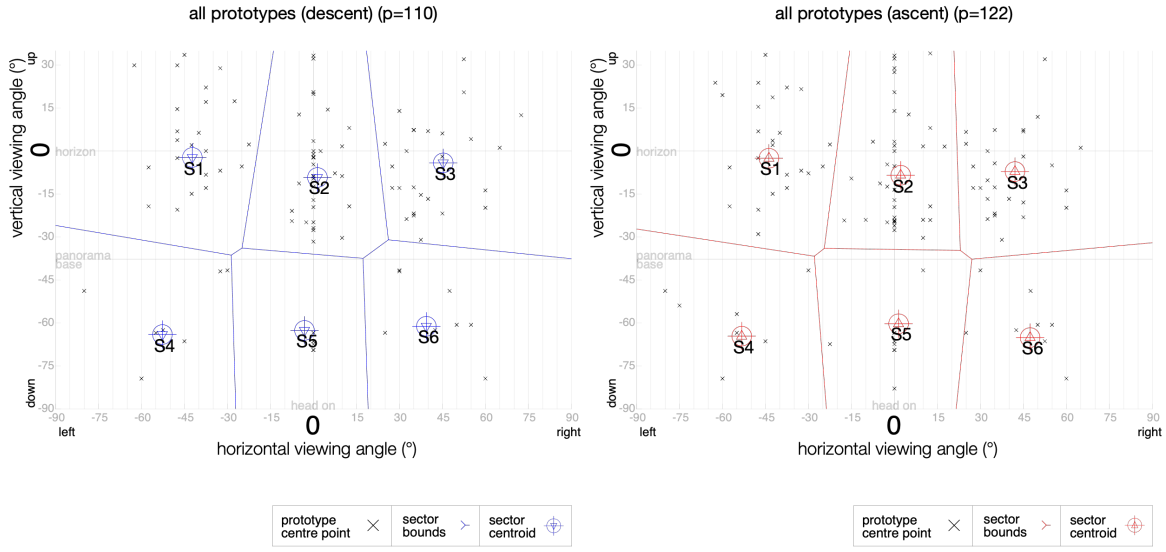

Figure K1: Centre points of prototypes with sector centroids of sectors 1-6 shown, separated by global cluster bounds (descent left, ascent right)

Appendix L – Occupancy Table (Continued)

Table L1: Continuation of Table 1

	prototype frequency			heatmap occupancy (%)						centre point location					
				descent			ascent			descent			ascent		
	descent	ascent	either	S1 S2 S3 S4 S5 S6			S1 S2 S3 S4 S5 S6			S1 S2 S3 S4 S5 S6			S1 S2 S3 S4 S5 S6		
(Table 1 continued)															
type-qualifier: current	30	37	38	19 9 27 7 25 13			20 14 31 4 26 6			9 4 9 2 3 2			10 9 12 1 3 1		
type-qualifier: remaining	18	23	25	15 15 20 13 33 4			16 20 24 7 33 0			4 4 4 1 3 0			5 8 4 1 3 0		
type-qualifier: alert	22	15	25	28 59 12 0 0 0			48 49 0 0 2 0			6 12 1 0 0 0			4 8 0 0 0 0		
type-qualifier: expected	12	13	14	16 24 16 0 29 15			10 39 10 0 34 7			3 5 2 0 1 1			3 6 2 0 1 1		
type-qualifier: elapsed	9	12	12	9 0 8 4 53 26			8 2 17 4 54 14			2 0 3 0 2 2			2 0 6 0 2 2		
type-qualifier: split	8	8	11	33 32 16 0 0 19			27 26 18 0 3 27			2 2 3 0 0 1			1 2 2 0 0 1		
representation: number	55	66	70	19 11 31 7 21 13			18 16 33 6 21 6			13 10 20 4 4 4			15 18 21 4 5 3		
representation: symbol	45	39	51	18 46 13 7 12 4			14 41 11 5 17 11			12 20 6 3 1 3			10 17 5 2 1 4		
representation: graph	11	19	21	5 14 28 15 28 10			6 20 35 8 25 6			1 1 4 2 2 1			2 5 7 2 2 1		
representation: map	12	9	13	0 57 6 0 34 3			0 49 5 0 46 0			0 9 1 0 2 0			0 5 1 0 3 0		
purpose: information	57	67	73	12 27 22 10 18 10			15 24 29 7 19 7			11 13 18 5 4 6			13 18 21 4 5 5		
purpose: motivation	25	34	39	17 12 18 10 32 11			13 27 16 6 33 5			6 4 7 1 4 1			7 10 6 1 5 1		
purpose: warning	32	26	35	14 47 11 8 12 8			12 35 8 11 25 9			7 17 3 1 1 1			5 12 2 1 2 1		
purpose: instruction	28	28	33	7 30 20 7 19 16			9 22 21 7 21 21			4 9 4 3 2 3			5 6 4 3 2 4		
scenario: agnostic	69	69	69	17 20 21 7 21 14			17 19 22 7 24 10			15 14 17 5 5 7			15 14 17 5 5 7		
scenario: specific	19	31	50	13 55 10 14 8 0			9 50 11 9 11 9			4 10 3 1 0 0			4 13 4 2 1 1		
scenario: specific / agnostic-moved	41	53	70	12 47 23 10 6 3			11 42 24 6 12 5			8 20 9 2 0 1			9 22 10 2 3 1		

Appendix M – Legend for Table 1

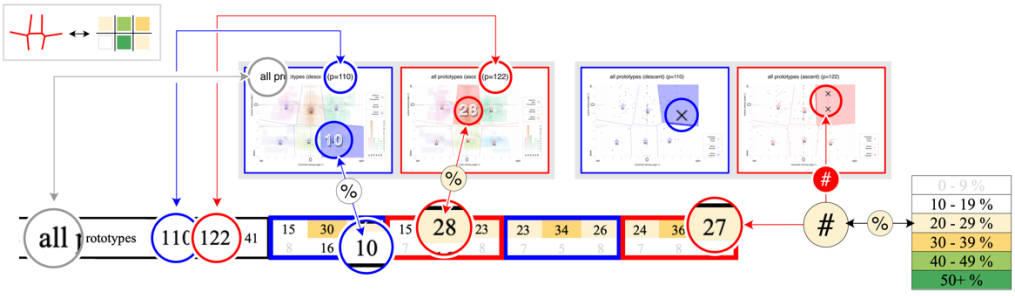

Figure M1: Legend for Table 1. Heatmap Occupancy

Appendix N – Heatmaps

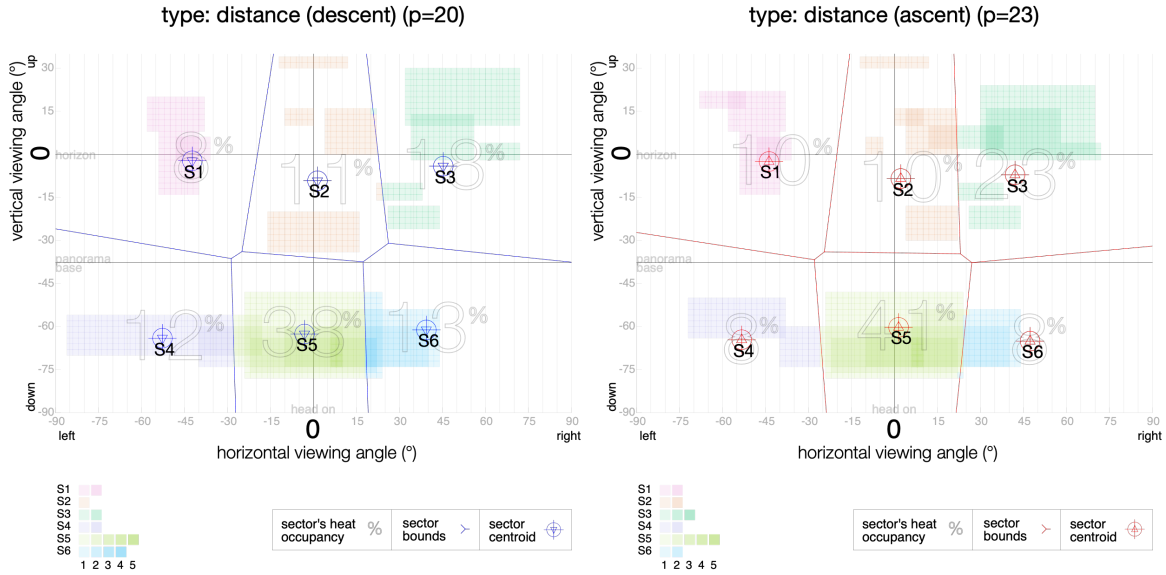

Figure N1: Heatmaps of type-metric: distance (descent, ascent)

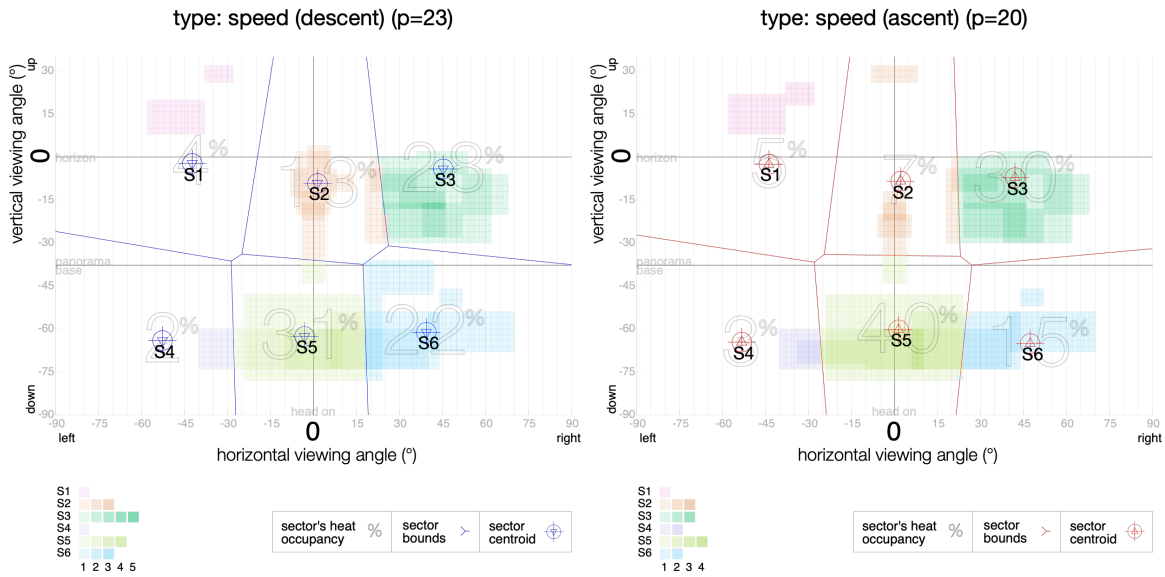

Figure N2: Heatmaps of type-metric: speed (descent, ascent)

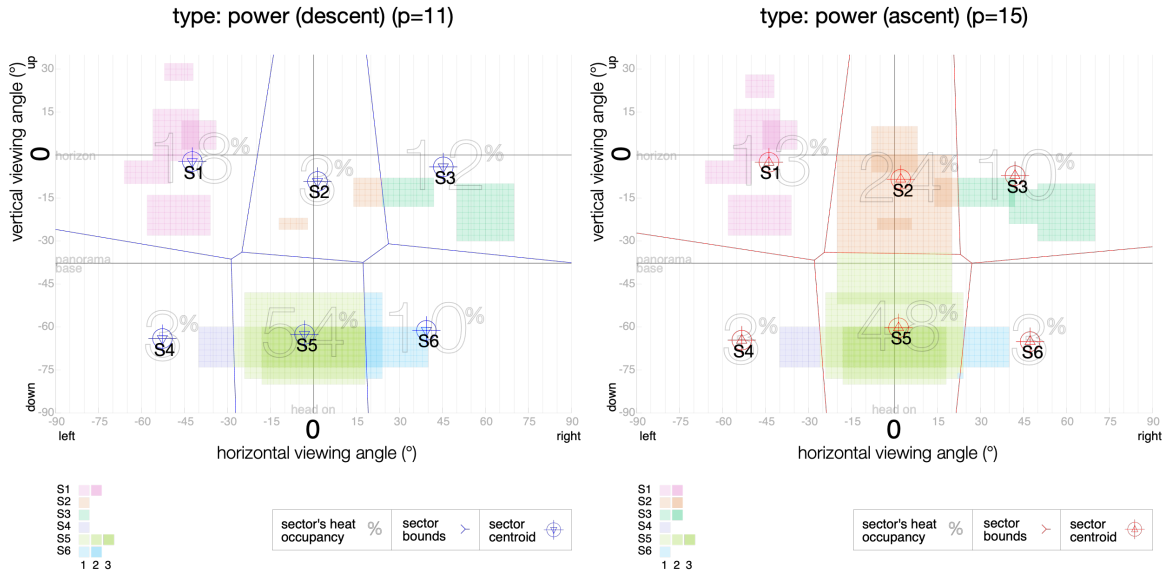

Figure N3: Heatmaps of type-metric: power (descent, ascent)

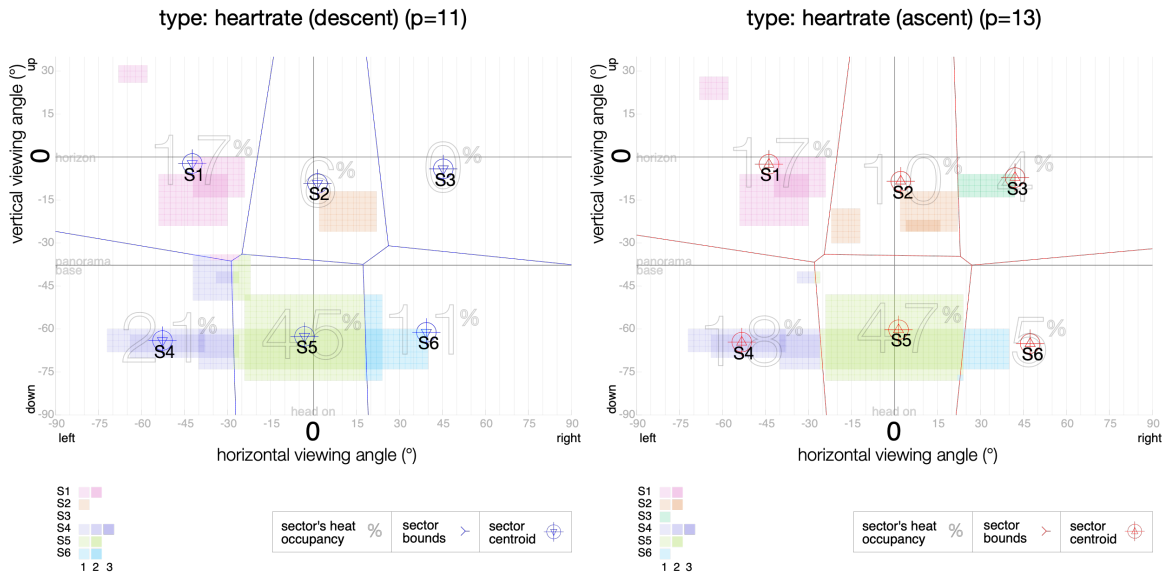

Figure N4: Heatmaps of type-metric: heartrate (descent, ascent)

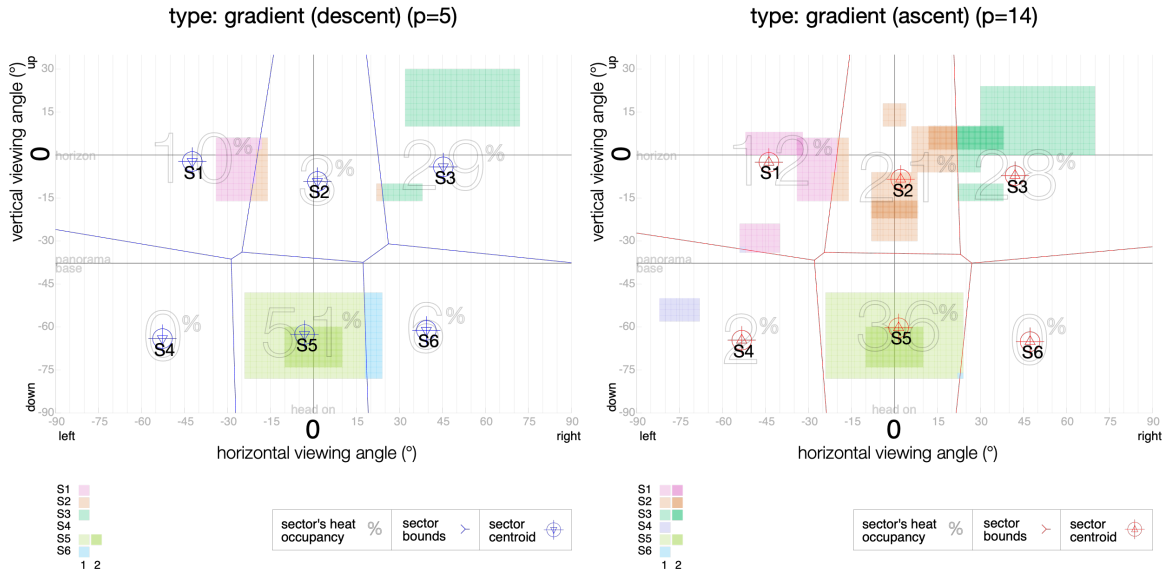

Figure N5: Heatmaps of type-metric: gradient (descent, ascent)

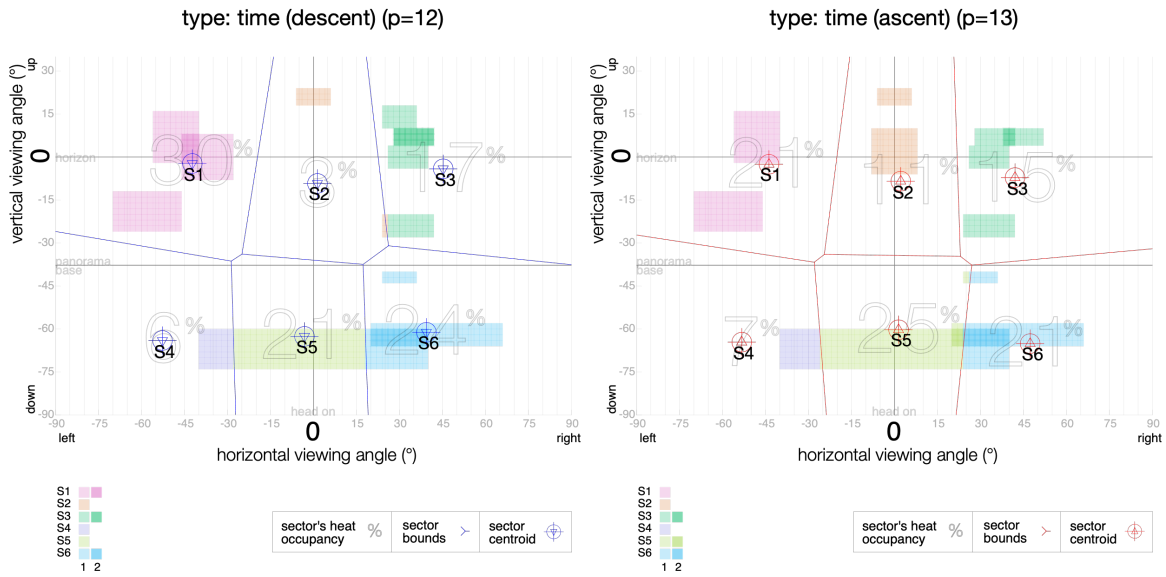

Figure N6: Heatmaps of type-metric: time (descent, ascent)

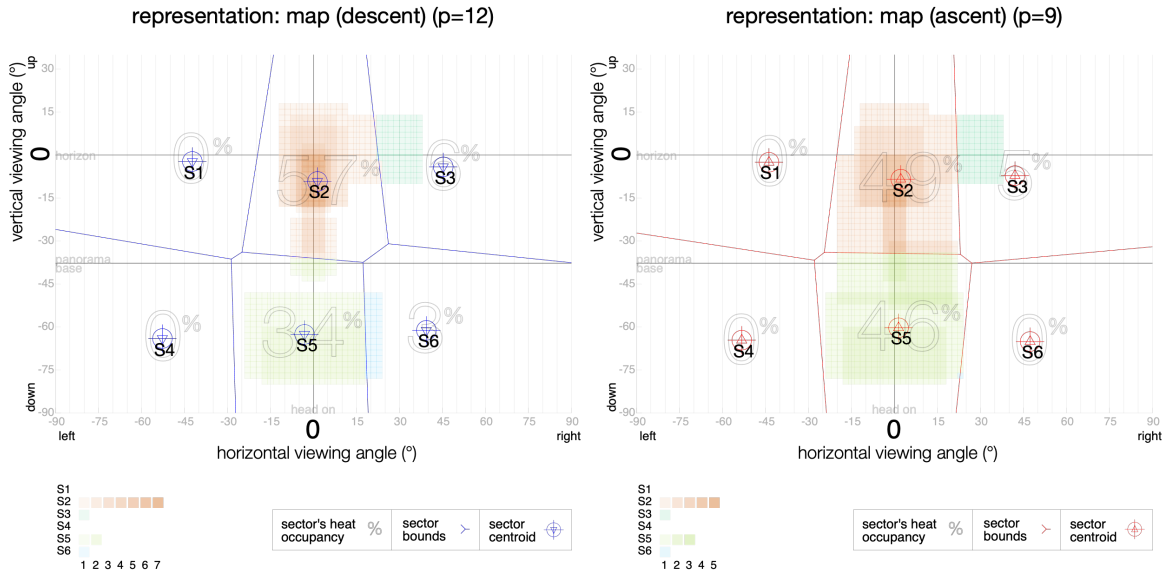

Figure N7: Heatmaps of representation: map (descent, ascent)

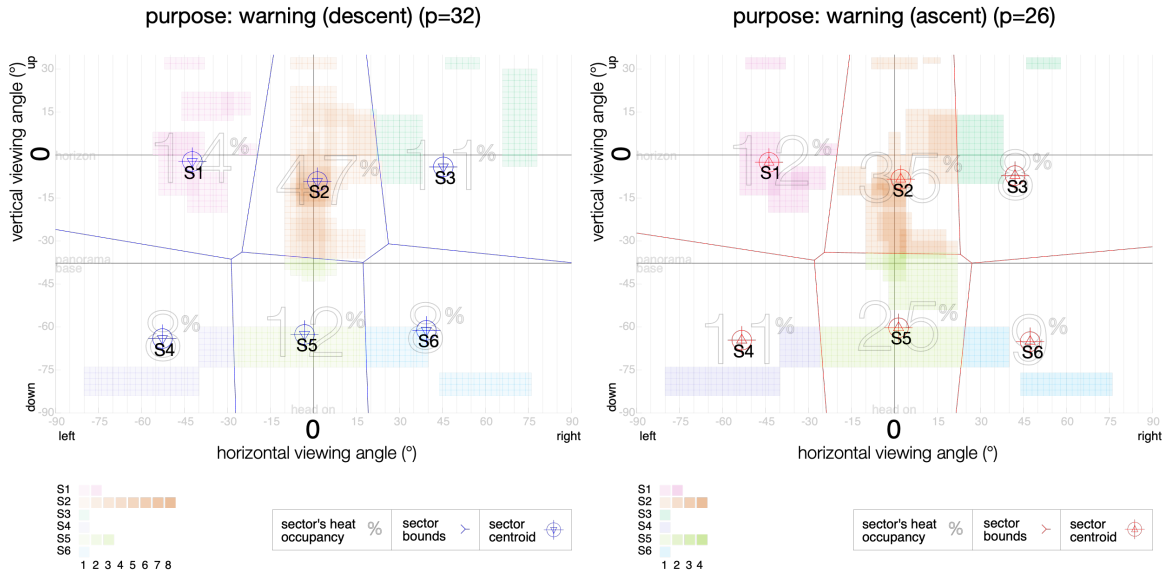

Figure N8: Heatmaps of purpose: warning (descent, ascent)

Appendix O – Alternative Apportionment Methods

For the descent list, the Adaptive Averaging Method places the second distance instance as the sixth item, whereas the Greedy Subtractive Method foregoes it altogether, in favour of introducing hydration as last on the list. In the ascent scenario, a similar difference is observed, trading hydration for a second instance of navigation.

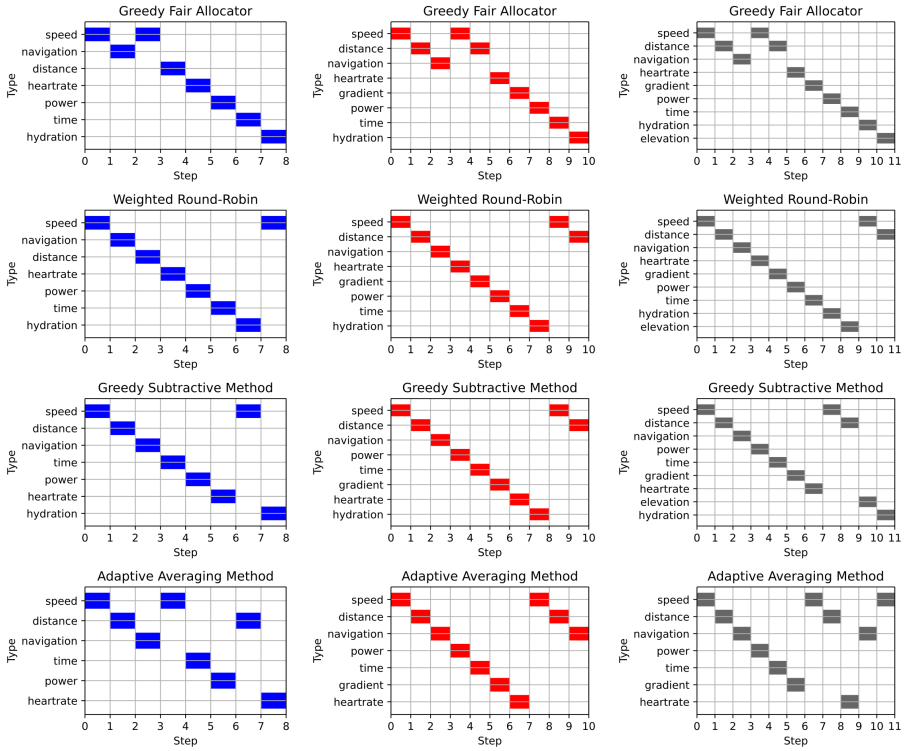

Figure O1: Gantt charts of ordered apportioned lists of type instances, as determined by four allocation methods, for descent (left, blue), ascent (middle, red), and scenario-inclusive (right, grey)

Appendix P – Mock-ups and Coordinates

(for Adaptive Averaging Method mock-ups, see Figure 10)

Table P1: Visual field coordinates of type-metric instances as apportioned for the descent mock-up by the Adaptive Averaging Method

item (i) (*type-metric*: class)	x coordinate (°)	y coordinate (°)	effective weight (w_e) (weight when picked)	initial weight (w_i)
1. speed (#1)	42	-16	1.560	1.560
2. distance (#1)	8	-68	1.258	1.258
3. navigation	-4	-12	1.157	1.157
4. speed (#2)	32	-16	0.780	1.560
5. time	-42	4	0.755	0.755
6. power	-2	-66	0.704	0.704
7. distance (#2)	18	-68	0.629	1.258
8. heartrate	-38	-64	0.604	0.604

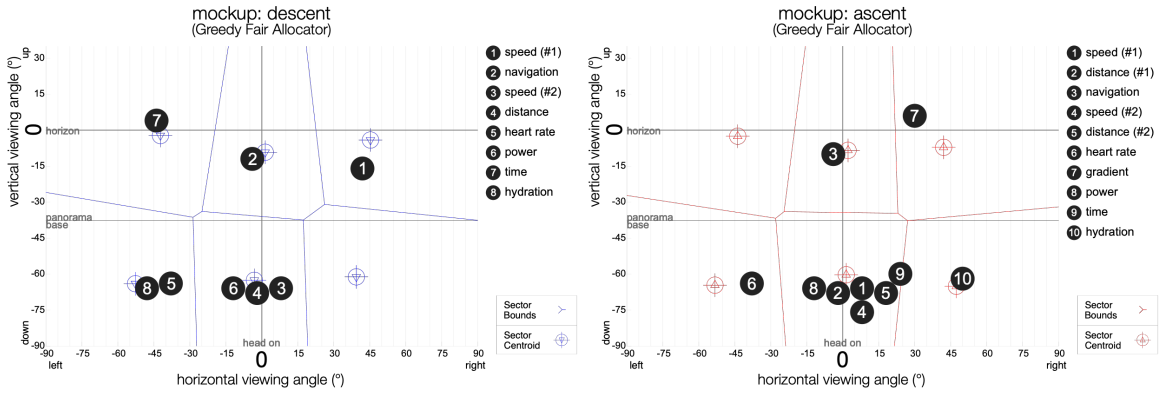

Figure P1: Mock-ups of locations of prototypes according to apportioned order of Greedy Fair Allocator, for descent in blue (above), ascent in red (below) (coordinates in a table below). Also shown sector centroid coordinates of sectors 1-6 separated by global cluster bounds.

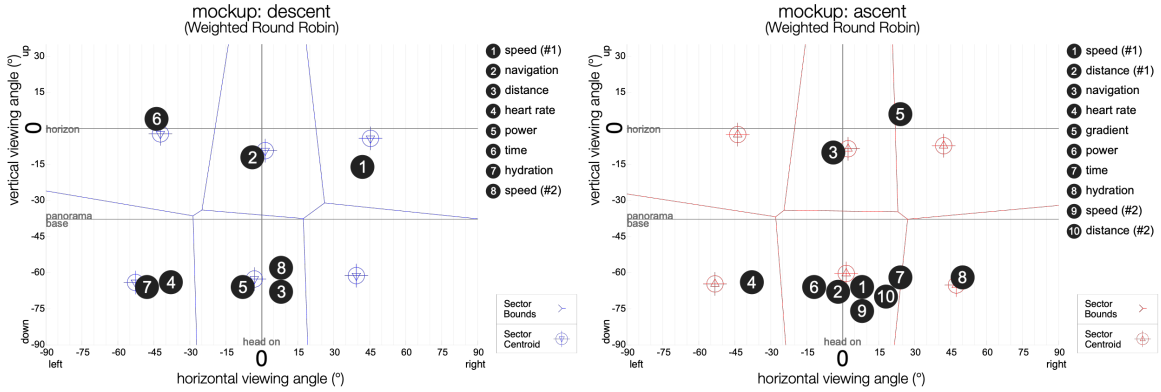

Figure P2: Mock-ups of locations of prototypes according to apportioned order of Weighted Round Robin, for descent in blue (above), ascent in red (below) (coordinates in a table below). Also shown sector centroid coordinates of sectors 1-6 separated by global cluster bounds.

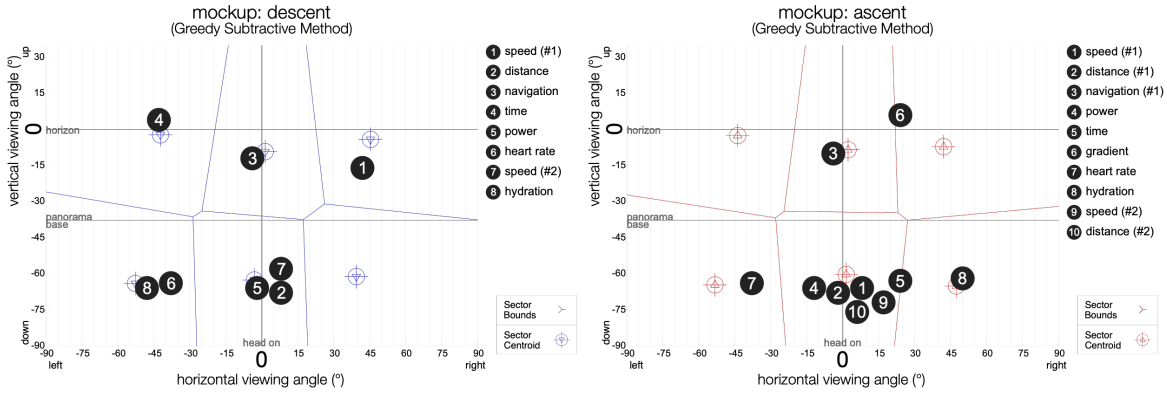

Figure P3: Mock-ups of locations of prototypes according to apportioned order of Greedy Subtractive Method, for descent in blue (above), ascent in red (below) (coordinates in a table below). Also shown sector centroid coordinates of sectors 1-6 separated by global cluster bounds.

Table P2: Visual field coordinates of type-metric instances as apportioned for the ascent mock-up by the Adaptive Averaging Method

item (i) (*type-metric*: class)	x coordinate (°)	y coordinate (°)	effective weight (w_e) (weight when picked)	initial weight (w_i)
1. speed (#1)	8	-66	1.476	1.476
2. distance (#1)	-2	-68	1.476	1.476
3. navigation (#1)	-4	-10	1.038	1.038
4. power	-12	-66	1.038	1.038
5. time	30	-62	0.929	0.929
6. gradient	24	6	0.874	0.874
7. heartrate	-38	-64	0.820	0.820
8. speed (#2)	42	-22	0.738	1.476
9. distance (#2)	18	-68	0.738	1.476
10. navigation (#2)	2	-2	0.519	1.038

Table P3: Visual field coordinates of type-metric instances as apportioned for the descent mock-up by the Greedy Fair Allocator

item (i) (*type-metric*: class)	x coordinate (°)	y coordinate (°)	effective weight (w_e) (weight when picked)	initial weight (w_i)
1. speed (#1)	8	-66	1.560	1.560
2. navigation	-4	-10	1.157	1.157
3. speed (#2)	8	-66	0.780	1.560
4. distance	-2	-68	1.258	1.258
5. heartrate	-38	-64	0.604	0.604
6. power	-12	-66	0.704	0.704
7. time	-44	4	0.755	0.755
8. hydration	-48	-66	0.352	0.352

Table P4: Visual field coordinates of type-metric instances as apportioned for the ascent mock-up by the Greedy Fair Allocator

item (i) (*type-metric*: class)	x coordinate (°)	y coordinate (°)	effective weight (w_e) (weight when picked)	initial weight (w_i)
1. speed (#1)	8	-66	1.475	1.475
2. distance (#1)	-2	-68	1.475	1.475
3. navigation	-4	-10	1.038	1.038
4. speed (#2)	8	-76	0.738	1.475
5. distance (#2)	18	-68	0.738	1.475
6. heartrate	-38	-64	0.820	0.820
7. gradient	30	6	0.874	0.874
8. power	-12	-66	1.038	1.038
9. time	24	-60	0.929	0.929
10. hydration	50	-62	0.492	0.492

Table P5: Visual field coordinates of type-metric instances as apportioned for the descent mock-up by the Weighted Round Robin

item (i) (*type-metric*: class)	x coordinate (°)	y coordinate (°)	effective weight (w_e) (weight when picked)	initial weight (w_i)
1. speed (#1)	8	-66	1.560	1.560
2. navigation	-4	-10	1.157	1.157
3. distance	8	-68	1.258	1.258
4. heartrate	-38	-64	0.604	0.604
5. power	-8	-66	0.704	0.704
6. time	-44	4	0.755	0.755
7. hydration	-48	-66	0.352	0.352
8. speed (#2)	8	-58	0.780	1.560

Table P6: Visual field coordinates of type-metric instances as apportioned for the ascent mock-up by the Weighted Round Robin

item (i) (*type-metric*: class)	x coordinate (°)	y coordinate (°)	effective weight (w_e) (weight when picked)	initial weight (w_i)
1. speed (#1)	8	-66	1.475	1.475
2. distance (#1)	-2	-68	1.475	1.475
3. navigation	-4	-10	1.038	1.038
4. heartrate	-38	-64	0.820	0.820
5. gradient	24	6	0.874	0.874
6. power	-12	-66	1.038	1.038
7. time	24	-62	0.929	0.929
8. hydration	50	-62	0.492	0.492
9. speed (#2)	8	-76	0.738	1.475
10. distance (#2)	18	-70	0.738	1.475

Table P7: Visual field coordinates of type-metric instances as apportioned for the descent mock-up by the Greedy Subtractive Method

item (i) (*type-metric*: class)	x coordinate (°)	y coordinate (°)	effective weight (w_e) (weight when picked)	initial weight (w_i)
1. speed (#1)	8	-66	1.560	1.560
2. distance	8	-68	1.258	1.258
3. navigation	-4	-10	1.157	1.157
4. time	-44	4	0.755	0.755
5. power	-2	-66	0.704	0.704
6. heartrate	-38	-64	0.604	0.604
7. speed (#2)	8	-58	0.780	1.560
8. hydration	-48	-66	0.352	0.352

Table P8: Visual field coordinates of type-metric instances as apportioned for the ascent mock-up by the Greedy Subtractive Method

item (i) (*type-metric*: class)	x coordinate (°)	y coordinate (°)	effective weight (w_e) (weight when picked)	initial weight (w_i)
1. speed (#1)	8	-66	1.475	1.475
2. distance (#1)	-2	-68	1.475	1.475
3. navigation	-4	-10	1.038	1.038
4. power	-12	-66	1.038	1.038
5. time	24	-64	0.929	0.929
6. gradient	24	6	0.874	0.874
7. heartrate	-38	-64	0.820	0.820
8. hydration	50	-62	0.492	0.492
9. speed (#2)	16	-72	0.738	1.475
10. distance (#2)	6	-76	0.738	1.475

Appendix Q – Gaze Point Estimation

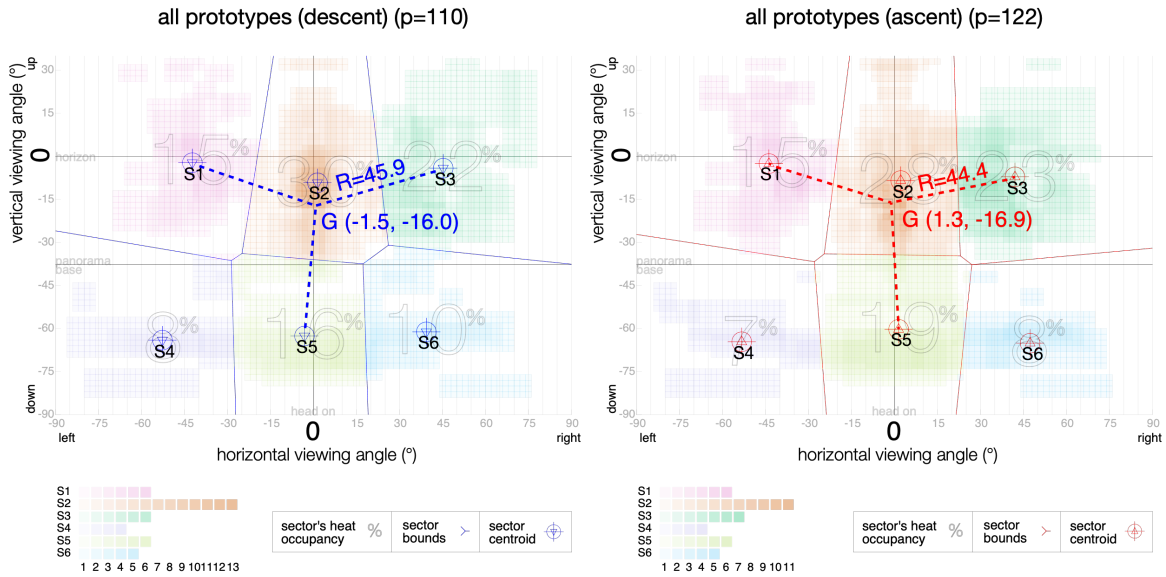

Figure Q1: Points equidistant to centroids of sectors 1, 3, and 5 for descent (left) and ascent (right). See Figure 8 for reference.

Difference between the sector centroids between scenarios is minor (Figure Q1). In the descent, sector 3 moves up and outwards, sector 6 moves inward towards the theorised gaze target, while sector 5 does the opposite. This can be a consequence of the road gaze point moving further forward due to a higher speed, combined with an attempt to improve road visibility immediately in front of the wheel. This can be also seen in the heatmap (Figure 7), with a clearer gap between sectors 2 and 5. In the descent, the road gaze point sits beneath sector's 2 hotspot, in a possible attempt to avoid it for safety due to a higher speed, while in the ascent the hotspot seemingly seeks the gaze point for motivation.