HEINZ NIXDORF INSTITUT
UNIVERSITÄT PADERBORN

Lukas Stratmann

Measuring Cyclists' Visual Attention and Interaction with an Interactive Real-Time 3D Simulation

Masterarbeit im Fach Informatik

31. Mai 2019

Please cite as:
Lukas Stratmann, "Measuring Cyclists' Visual Attention and Interaction with an Interactive Real-Time 3D Simulation," Master's Thesis (Masterarbeit), Heinz Nixdorf Institute, Paderborn University, Germany, May 2019.

Distributed Embedded Systems (CCS Labs)
Heinz Nixdorf Institute, Paderborn University, Germany

Fürstenallee 11 · 33102 Paderborn · Germany

http://www.ccs-labs.org/

Measuring Cyclists’ Visual Attention and Interaction with an Interactive Real-Time 3D Simulation

Masterarbeit im Fach Informatik

vorgelegt von

Lukas Stratmann

geb. am 26. Juli 1993
in Geseke

angefertigt in der Fachgruppe

Distributed Embedded Systems
(CCS Labs)

Heinz Nixdorf Institut
Universität Paderborn

Betreuer:	**Dominik S. Buse**
Gutachter:	**Falko Dressler**
	Ingrid Scharlau
Abgabe der Arbeit:	**31. Mai 2019**

Erklärung

Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde.
Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Declaration

I declare that the work is entirely my own and was produced with no assistance from third parties.
I certify that the work has not been submitted in the same or any similar form for assessment to any other examining body and all references, direct and indirect, are indicated as such and have been cited accordingly.

(Lukas Stratmann)
Paderborn, 31. Mai 2019

Abstract

In recent years, the safety of Vulnerable Road Users (VRUs) has improved much more slowly than that of motorized road users. A promising solution are Advanced Driver Assistance Systems (ADAS) for cyclists in the context of Vehicle-to-X (V2X) communication. To ensure their effectiveness and safety, these systems have to be tested psychologically without risking the lives of participants. Suitable simulators such as the Virtual Cycling Environment (VCE) can help with this task by letting cyclists interact with simulated networked traffic. So far, such simulators have rarely been used for VRUs in long scenarios with complex traffic. In this thesis, I evaluate the suitability of the improved VCE in an empirical study and I present a method for measuring cyclists' visual attention in different traffic densities. Results align with the literature in showing that higher-density traffic requires more attention. The VCE is therefore considered suitable for human-in-the-loop experiments. Acting on feedback from participants can further enhance this suitability, e.g., by improving the way other cars react to the cyclist.

Kurzfassung

Während sich die Sicherheit von motorisierten Verkehrsbeteiligten in vergangenen Jahren deutlich gebessert hat, haben ungeschützte Verkehrsbeteiligte wie Rad Fahrende weniger deutliche Fortschritte in ihrer Sicherheit erfahren. Die Entwicklung von Fahrerassistenzsystemen im Kontext von drahtlos kommunizierenden Fahrzeugen stellt einen vielversprechenden Lösungsansatz dar. Potenzielle Systeme können mit Simulatoren wie der Virtual Cycling Environment (VCE) getested werden, in der eine Radfahrerin oder ein Radfahrer auf einem stationären Fahrrad fahrend mit simuliertem vernetztem Verkehr interagieren kann. Bisher wurden derartige Simulatoren für ungeschützte Verkehrsbeteiligte selten für längere Szenarien mit komplexem Verkehr eingesetzt. In dieser Arbeit evaluiere ich die Eignung der verbesserten VCE in einer empirischen Studie und stelle eine Methode vor, um die visuelle Aufmerksamkeit von Radfahrerinnen und Radfahrern in verschiedenen Verkehrsdichtebedingungen zu messen. Die Ergebnisse hieraus stimmen mit der Literatur überein indem sie zeigen, dass eine hohe Verkehrsdichte mehr Aufmerksamkeit erfordert. Die VCE kann daher als geeignet für psychologische und künftige Assistenzsystem-testende Experimente angesehen werden. Die Umsetzung möglicher Anpassungen, die auf den Rückmeldungen der Versuchspersonen basieren, kann diese Eignung noch verbessern, z.B. durch ein noch realistischeres Verhalten der simulierten Autos.

Contents

Chapter 1

Introduction

According to a recent report by the European Union, 25 300 people in its member states died in traffic in the year 2017 [1]. While this number as a fraction of EU inhabitants was the lowest recorded thus far, much of the reduction in the number of fatalities is owed to improvements in the safety of motorized vehicles. In contrast, the number of fatal accidents among Vulnerable Road Users (VRUs) such as cyclists has reportedly decreased by a far lesser degree.

Drivers of modern cars are protected by seatbelts and crumpling zones, as well as Advanced Driver Assistance Systems (ADAS) such as collision avoidance systems. So far, a lot of research exists on ADAS in cars and, according to Bengler et al. [2], sensor networks are expected to be playing an increasingly important role in the future. Such Vehicular Ad-hoc Networks (VANETs) enable the development of assistance features that would be difficult without communication between vehicles. Examples for this include cooperative collision avoidance on highways [3] or at intersections [4]. Considering this potential, extending VANETs to include other road users like pedestrians or cyclists is a natural next step [5]. However, bicycles have less room and weight margins than cars for complex active safety features. It is therefore likely that typical solutions will require cyclists to cooperate by reacting to warning signals. For this reason, it is important to understand how cyclists react to warning signals and which types of signals are effective to what degree.

ADAS for cyclists require experimental validation and, before this, an understanding of cyclists' perception and behavior in traffic. However, conducting such experiments in the field is not a viable option if the focus of the investigation are dangerous traffic situations in which participants are at risk of coming to physical harm. Furthermore, field experiments have the potential downside that traffic conditions can be hard to reproduce between different participants, in follow-up experiments, or when trying to identify possible issues. Lastly, investigating VANET-enabled ADAS for VRUs would require a sufficiently large portion of road users to be equipped with the necessary hardware, which can be expensive. While these limitations do not justify giving up field experiments completely, it is often preferable to gather data from simulation. In particular, a so-called human-in-the-loop simulation in a Virtual Cycling Environment (VCE) such as the one presented by Heinovski et al. [6] allows a cyclist to navigate a virtual road network while sitting on a real but stationary bicycle. The risk of physical harm is therefore greatly reduced and, by virtue of simulation, experiments are comparatively easy to reproduce.

In previous work [6], we investigated how much time can be gained between a collision and the event of a bicycle and a car (or their respective occupants) recognising each other if bicycles are integrated in the vehicular communication. For this experiment, the VCE was used to record realistic mobility traces of cyclists without other cars on the road. Later, these mobility traces were then used for simulating many iterations with virtual traffic. The goal of the present thesis is to see how well the VCE is suited for psychological experiments in which cyclists have to interact with traffic and navigate through a much longer scenario. I do so by measuring participants' visual attention, based on a formal theory of visual attention, under conditions of low and high traffic density. Thereby, I explore the potential of the VCE as a tool for future psychological studies pertaining to the safety of VRUs, their perception of the environment, as well as their interaction with traffic or with ADAS. A portion of this thesis is dedicated to explaining the extensions made to the VCE that allow conducting this or similar other experiments.

An additional goal of this thesis is to improve upon the original VCE. This includes critical enhancements to the visual presentation of the virtual world and improving interactivity with the development of a more responsive steering sensor. In order to evaluate the VCE for performing psychological experiments with ADAS in the future, participants were queried about their subjective experience in the virtual environment. In particular, cyclists were given questionnaires about their perceived realism of the cycling experience and their experience of *flow* as an indicator for their presence within and their acceptance of the simulation. Additionally, in order to identify areas for future improvement, written feedback was also collected from participants.

In essence, the main contributions of this thesis can be summarized as follows:

- I present a method for measuring cyclists' visual attention using the extended VCE based on a formal theory of visual attention.
- I provide first insights into changes in the visual attention of cyclists when exposed to varying amounts of simulated traffic.
- I identify important aspects for participants' acceptance of such a VCE, in particular, the accuracy of steering controls and the predictability of how simulated cars will react to the presence of the cyclist.

Based on this thesis, we have submitted a workshop paper for the Workshop on ICT based Collision Avoidance for VRUs (ICT4VRU) 2019:
L. Stratmann et al., “Psychological Feasibility of a Virtual Cycling Environment for Human-in-the-Loop Experiments,” in *1st Workshop on ICT based Collision Avoidance for VRUs (ICT4VRU 2019)*, in peer-review, Kassel, Germany, Sep. 2019 [7].

Chapter 2

Fundamentals

The following chapter explains the foundations this thesis builds upon. This includes the existing Virtual Cycling Environment (VCE), which is to be extended and evaluated, as well as an overview of useful concepts for the psychological evaluation of an interactive simulation. To give some context for this thesis, Section 2.4 presents a summary of other people's efforts to build interactive cycling environments, to improve the safety of Vulnerable Road Users (VRUs), or to measure road users' attention in traffic.

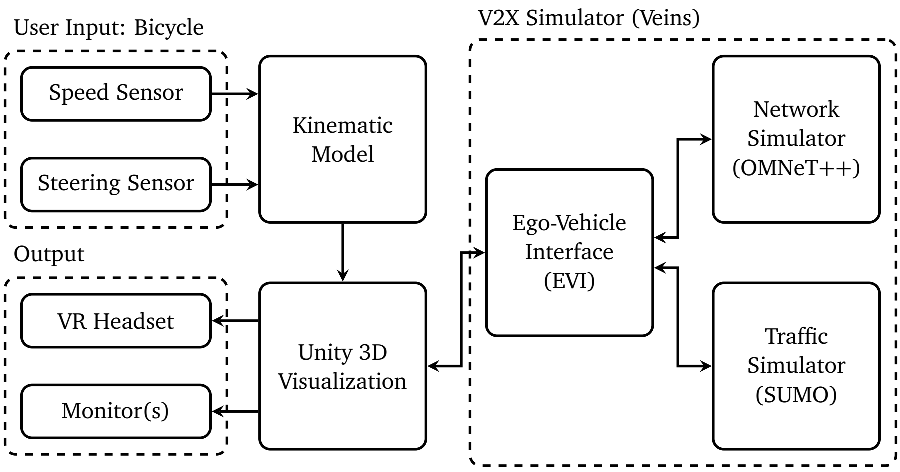


Figure 2.1 – Components of the original Virtual Cycling Environment (VCE), figure adapted from [6, Figure 2]

2.1 The Existing Virtual Cycling Environment

The Virtual Cycling Environment (VCE) is presented by Heinovski et al. [6] in close to its current form, with some improvements from this thesis already incorporated. As is shown in Figure 2.1, a user can interact with the simulation via a bicycle equipped with sensors for measuring the current velocity and steering angle. An off-the-shelf bicycle is propped up on a training stand and, unlike some of the solutions presented in Section 2.4, requires no major modifications like removing a wheel. Before I began working on this thesis, the steering angle was measured via a mechanism that was shipped as part of the training stand, which uses Bluetooth to transmit the steering angle values to the kinematic model. Similarly, the VCE used to rely on a cadence sensor provided by the manufacturer of the training stand for obtaining velocity readings. However, this solution proved to be too imprecise with new speed readings only arriving after intervals of at least 500 ms, which especially made it difficult to detect soon enough when the bicycle was stopped. Heinovski et al. [6] therefore mounted an infrared sensor to the bicycle, connected it to a Raspberry Pi single-board computer, and aimed the sensor at the spokes of the rear wheel. Attached to these spokes are tube-shaped reflectors that can be detected by the sensor. Rather than Bluetooth, the custom speed sensor communicates with the kinematic model via UDP over an IP network.

The component named *Kinematic Model* receives updates from the sensors in independent intervals. Within the kinematic model, the sensor data are used to update the bicycle position and orientation. This happens independently of the scenario and street layout currently shown in the visualization; in the kinematic model, the bicycle moves on an infinite plane. The position, velocity, orientation, and the current steering angle are forwarded to the visualization at a default rate of 10 Hz [6].

The *Unity 3D Visualization* is an essential component of the VCE in the sense that it provides a cyclist with a rendering of their surroundings, of the bicycle itself, and of their position in the virtual world. Cyclists can view this virtual world either using a Virtual Reality (VR) headset as employed in the experiment of Heinovski et al. [6], or on ordinary computer monitors. The virtual world itself, including roads, buildings, and traffic lights, can be generated by the visualization component from the same files that are also used by the urban traffic simulator SUMO. Since SUMO road networks can be generated from OpenStreetMap data, i.e., real-world road networks, this feature presents a significant contribution to the realism of the VCE.

Due to the modularity of the VCE, the components presented so far already constitute a functioning interactive cycling simulation. However, an additional feature of the visualization component is the rendering of fellow traffic on the road in the form of cars. The fellow road traffic is simulated with the urban traffic simulator SUMO [8], which is designed as a time-discrete and space-continuous simulation [9]. In this case, this means that the state of all vehicles collectively changes in fixed, discrete intervals of simulated time. Since the VCE requires real-time interactivity, these discrete update intervals in simulated time must correspond to intervals in wall-clock time. Ensuring this is one of the tasks of the *Ego-Vehicle Interface (EVI)* component, which is signalled by the visualization component in 10 Hz ticks, upon which the EVI triggers the computation of vehicle states in SUMO for the next time step (cf. [10]). In return, the EVI receives the updated vehicle states from SUMO, filters out information about vehicles outside a configurable radius (e.g., 150 m), and forwards the updates to the visualization. Another task of the EVI is the synchronization of the position of the ego-vehicle (a bicycle in the case of the VCE) with the traffic simulation. Part of the contents of messages sent to the EVI in 10 Hz intervals by the visualization component is the position of the ego-vehicle. The EVI uses this information to insert the ego-vehicle/bicycle into the SUMO traffic simulation, and SUMO thereby obtains the necessary data to simulate the reaction of cars to the presence of the bicycle.

One feature that sets the VCE apart from other interactive cycling simulators in the literature (cf. Section 2.4) is the optional integration of simulated Vehicle-to-X (V2X) communication. As Buse et al. [10] show, the EVI is capable of synchronizing the V2X simulator Veins [11] with the ego-vehicle and with the road traffic simulation. In addition, Loewen et al. [5] provide an extension to classic Context Awareness Messages (CAMs) and Decentralized Environmental Notification Messages (DENMs) originally developed for cars such that they include relevant information for cyclists. So far, Heinovski et al. [6] have used the VCE's capability to record traces of cyclists traversing the virtual environment in order to simulate V2X communication separately and with a vastly larger number of iterations than would be feasible if every condition were to be tested with a human participant. Being able to incorporate V2X simulation into the VCE in real-time, however, promises the possibility to investigate the interaction of cyclists with realistically timed warning signals.

2.2 Visual Attention

Riding a bicycle requires paying attention to a miscellany of stimuli, which become apparent when considering the many subtasks a cyclist has to manage. Wierda and Brookhuis [12] characterise three levels for the task of cycling: On the *control level*, a cyclist must be able to balance the bicycle, maintain the desired speed, or control the bicycle to move in the wanted direction. On a higher level, the cyclist's task is to *maneuver* the bicycle, for example to avoid a deep puddle in their path, to make room for a car they hear approaching from behind, or to slow the bicycle down when pedestrians might step on the cycleway. The highest level is the *strategic level*, which involves longer-time tasks such as planning a route to the destination. The act of cycling therefore involves several senses, especially on the control and maneuvering level. For example, the sense of balance is important for controlling the bicycle, feeling and hearing the wind may contribute to judging one's current velocity, while hearing or smelling other vehicles may affect maneuvering. Nevertheless, vision seems like an especially important sense as it helps cyclists to see the condition and curvature of the path ahead, the location of obstacles, or other road users such as pedestrians or fellow cyclists. Consequently, the visual attention of cyclists is a worthy target of investigation.

For measuring visual attention, we can make use of the Theory of Visual Attention (TVA) developed by Bundesen [13]. According to the TVA, the information that an arbitrary stimulus x in the visual field belongs to some category i is encoded in a person's capacity-limited visual short-term memory. The processing rate $v(x,i)$ denotes the rate at which this encoding happens for a specific stimulus and category. The sum over all processing rates of all objects x and all categories i is called the overall processing capacity C:

$$C = \sum_x \sum_i v(x,i). \tag{2.1}$$

Tünnermann, Petersen, and Scharlau [14] and Krüger, Tünnermann, and Scharlau [15] have built upon the TVA to precisely measure certain aspects of visual attention, namely *prior entry* and *salience*. Prior entry here refers to the extent by which an attended stimulus is perceived earlier than another stimulus, for example when two letters appear on an otherwise empty screen and one of these letters appears in a spot to which the participant is currently paying attention. Salience refers to how feature contrasts such as a bright stimulus among a group of many dim stimuli can cause attention to shift to the bright stimulus, thus causing it to be perceived earlier. As a tool, the authors used so-called Temporal Order Judgment (TOJ) experiments.

The timeline of one TOJ trial as described by Krüger, Tünnermann, and Scharlau [15], typically out of several hundred, is shown in Figure 2.2. An order judgment is a binary choice between two stimuli called *probe* and *reference*; a person has to judge in which stimulus they noticed a change first. In the example shown in Figure 2.2, this change comes in the form of flickering, i.e., at first both stimuli are visible in the interval $[t_0, t_1]$ until each disappears for a brief amount of time (e.g., 80 ms [15]) one after the other. The length of this interval corresponds to the absolute value of the Stimulus Onset Asynchrony (SOA), which encodes whether the probe or the reference stimulus flickers first. If $\text{SOA} < 0$, this will be the case

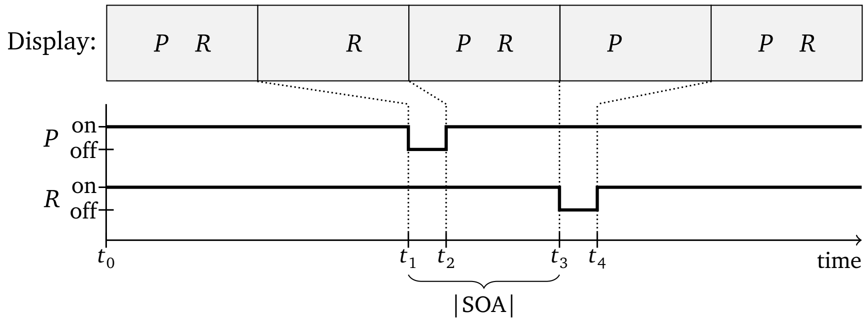


Figure 2.2 – Timeline of a Temporal Order Judgment (TOJ) trial in which a probe stimulus P and a reference stimulus R flicker one after the other. Note that the temporal order of P and R may be reversed and that the positioning of stimuli is arbitrary in this example. The participant's task is to judge whether P or R flickered first.

for the probe stimulus, the appearance of which can differ from the reference and background stimuli in salience experiments. If $\text{SOA} > 0$, the reference stimulus flickers first and both events will happen simultaneously if $\text{SOA} = 0$.

The idea behind such TOJ tasks is to present the same task under different conditions by varying the time between events and, for example, attempting to draw participants' attention to the probe stimulus by making it brighter. Based on the mistakes participants make, it is then possible to use the TVA to infer parameters of people's visual attention under the conditions of interest. How this can be done for the special case of cyclists' visual attention in different traffic densities is explained in Section 5.2.

2.3 Quality of Interactive Virtual Environments

The goal of this thesis is to assess the feasibility of the VCE for human-in-the-loop experiments, e.g., for evaluating ADAS in the future. Besides measuring cyclists' visual attention as a proof of concept, there are other factors that can be taken into account when it comes to estimating the quality of an interactive simulation. In a broad sense, the VCE should be realistic in such a way that experimental results obtained using the simulation correspond with a high degree of confidence to what the same experiment would yield in the real world. To make this outcome more likely, a bicyclist using the VCE should, to a sufficient degree, feel like they are riding a bicycle in the real world:

> "An increased sense of presence is often thought to magnify user effects (e.g., the extent to which user responses to virtual stimuli and virtual

> interactions resemble parallel responses to real-world counterparts) [...]." [16]

If this feeling can be quantified and intensified, the quality and the value of the VCE for human-in-the-loop experiments should improve.

A difficulty with quantifying a sense of presence is the multitude of definitions for this concept that differ in some key details, as well as the existence of related and, depending on the respective definitions, overlapping concepts like *immersion* (cf. [16], [17]). In some instances, for example, immersion is treated as synonymous with presence (e.g., [18]), while in others, immersion is defined as a property of the virtual environment and presence as the corresponding experience by one of its users [16]. Such differences have to be paid close attention to when comparing results about immersion or presence. Unified, general-purpose questionnaires such as the Igroup Presence Questionnaire (IPQ) [19] can help with comparability if two or more virtual environments were evaluated with the same questionnaire.

Another important concept is *flow*, which describes a balance between one's skills and the challenges faced in a task, e.g., in a video game. According to Nakamura and Csikszentmihalyi [20], the experience of flow is characterized by the following factors: intense and focused concentration, merged action and awareness, the loss of reflective self-consciousness, a sense of being in control of what happens next, a distortion of one's temporal experience, and experiencing the activity as intrinsically rewarding. Especially the factor *merged action and awareness* highlights the relatedness of flow and presence. If one's awareness is merged with one's actions in a task that is part of a virtual environment, it is possible to argue that this awareness is projected into the virtual environment, which is similar to a sense of being present. In addition to mere presence, however, flow also comes with the factors listed above. These factors may give indications about users' acceptance of a virtual environment in the sense that people are motivated to pursue the given task while being highly concentrated and feeling like they are in control. Especially for this thesis, which focuses on cyclists' visual attention, it is valuable to maintain participants' concentration and focus in a long-running task. Besides a balance of challenge and skill, another prerequisite for this aim is for the environment to present clear goals and to provide users with immediate feedback about their progress [20].

2.4 Related Work

In the literature, a number of different interactive bicycle simulators have already been developed besides the VCE [6], e.g., [21], [22]. The focus in these latter examples is mainly on correctly mapping the actions of the cyclist to the virtual bicycle according to a mechanical model, and in turn to provide realistic haptic feedback torque on the handlebar [22] or even to re-apply the motion of the simulated bicycle back to the physical bicycle [21]. However, there have not been many efforts in the past to use cycling simulators for conducting research on the safety of VRUs. Notable exceptions are presented by Matviienko et al. [23], who used a non-steerable interactive bicycle simulator to evaluate the effectiveness of a variety of warning signals for children, as well as the work presented by Plumert, Kearney, and Cremer [24], who use a virtual cycling environment with simulated traffic to study the behavior of children on bicycles. Heinovski et al. [6] used the VCE to measure the amount of time that can be gained before a possible collision by equipping bicycles with hardware for V2X communication. Experiments on people's attention in simulated traffic, however, so far appear to be confined to the realm of motorized vehicles.

The focus of the present thesis is on the visual attention of cyclists. Measuring people's attention in traffic is not a new idea and has been done with a wide variety of metrics. For example, Strayer, Drews, and Johnston [25] attempted to quantify their participants' attention, among other metrics, by collecting drivers' following distance to the car in front as well as the delays when performing braking maneuvers. These metrics are interesting as they are directly concerned with effects of distracted driving that can result in people being harmed. On the downside, this particular method is tailored to the situation of one car equipped with brake lights being followed by another, which is not easily transferable to bicycles. As another metric, the authors employed a memory task in which participants were asked to remember billboards shown during the experiment. Such a task would be easy to adapt for cyclists. As a between-subjects design, however, this paradigm would require a substantially larger number of participants the more conditions are to be tested.

A common alternative approach for measuring attention is the use of eye trackers, which can observe several different properties of how a person looks at a scene. For example, Strayer, Drews, and Johnston [25] additionally used the durations of eye fixations on billboards placed alongside the road and showed that participants did in fact look at the billboards they later failed to identify. According to the authors, this indicates that car drivers talking on the phone pay less attention even to the things they look at. As another example, Di Stasi et al. [26] collected the number, duration, and peak velocity of saccades, i.e., very quick eye movements, in a motorcycling simulation and compared them to participants' responses to a mental workload scale. Their results indicate that the peak saccadic velocity is a good indicator for participants' self-reported mental workload. In general, research shows that some eye tracking parameters are indeed suitable for measuring car drivers' mental workload under certain conditions, but no single property seems to be suitable under all circumstances [27].

Having a closer look on the effect of traffic density on road users' attention, we see some research efforts in this area. Aforementioned Strayer, Drews, and Johnston [25] manipulated the traffic density between subjects in the first of their experiments and found that talking on the phone impaired people's car driving performance more strongly when the traffic density was high. Vlakveld et al. [28] conducted a field experiment in which they observed cyclists' mental workload and cycling speed in low and high complexity traffic situations while also comparing conventional bicycles to e-bikes. For measuring mental workload, the authors instructed their participants to react to blinking LEDs in their peripheral vision, called the Peripheral Detection Task (PDT), and used a combined metric of reaction time and hit rate. As the authors concede, one difficulty with such field experiments is the selection of appropriate traffic situations for VRUs that are neither too simple nor too dangerous: "The only safe way to investigate cyclists['] behaviour in real complex traffic situations probably is in a bicycling simulator."

It is this gap in research which this thesis will be focused on by presenting a methodology for measuring cyclists' visual attention in simulated traffic. This method is based on established research in the domain of visual attention and allows for the examination of arbitrarily complex traffic situations without the risk of physical harm.

Chapter 3

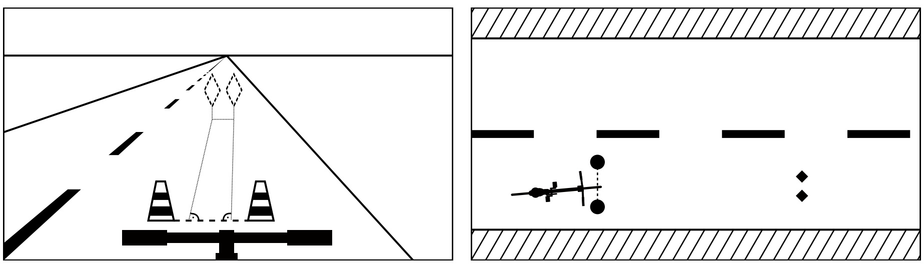
Concept

For the aim of measuring cyclists' visual attention, a concept for an experimental setup is required. As explained in section 2.2, a Temporal Order Judgment (TOJ) experiment evaluated based on the Theory of Visual Attention (TVA) presents a suitable tool for obtaining precise measurements of visual attention. How specifically this can be applied to the existing virtual cycling environment will be explained in the following.

A TOJ experiment is composed of a number of trials and the design of these trials should be considered carefully. A single temporal order judgment requires at least two separate events, e.g., a flicker or the appearance of a stimulus. In the experiments of Krüger, Tünnermann, and Scharlau [15], there are exactly two events for each trial distributed across a *probe* and a *reference* stimulus. What a stimulus looks like depends on the experiment. For example, simple shapes like circles can be used for measuring the effects of color, and letters can be used if participants are supposed to identify the shape itself. For Krüger, Tünnermann, and Scharlau [15], it is important that probe and reference look visually different in most of the trials, as the effect of those visual differences on visual salience is a primary metric. For measuring the effect of an increase in traffic on visual attention, however, we are interested in how changes in the environment around the stimuli affect cyclists' attention for the stimuli themselves. Therefore, the salience difference between probe and reference is of no concern, which is why they can safely assume the same appearance.

Regarding TOJ trials, the question that remains is how they are to be presented to cyclists in the VCE. One naïve option would be a two-dimensional overlay comparable to a virtual heads-up display, or alternatively rendering the stimuli on a virtual tablet screen mounted to the simulated bicycle handle. This would make the experiment somewhat more similar to the Peripheral Detection Task (PDT) (cp. Section 2.4). For performing an order judgment, buttons might be attached to the bicycle handle such that participants may report the left or the right stimulus. A potential downside of this approach is the detachment of the TOJ experiment from the virtual environment. While on the upside it would be relatively easy to transfer such an experiment to real roads, it is harder to incorporate the TOJ trials into a gamification concept (details below) if the surroundings are irrelevant for the task. Considering the presentation of stimuli on a screen attached to the handle, the positioning of the stimuli would furthermore present an issue. Ideally, the probe and reference stimuli should be approximately in the same parts of the visual field regardless of whether the traffic density is low or high. However, if stimuli are shown in the vicinity of the bicycle handle, the cyclist would be likely to look up from the TOJ task when other vehicles are present. This, in turn, may result in skewed results since people are more sensitive to flickering stimuli in their peripheral than in their foveal vision [29].

For the reasons discussed above, the presentation of TOJ stimuli will take the following form. Reminiscent of a stereotypical video game, the probe and reference stimulus will both be visualized as yellow gem stones hovering half a meter above the street (Figure 3.1a). Rather than pushing buttons, participants of the experiment will express their judgment of the temporal order by riding the virtual bicycle through that stimulus which they think flickered first. Participants can trigger the appearance of the next pair of stimuli by navigating through a pair of traffic cones as shown in Figure 3.1b. As this should lead cyclists to face roughly in a direction orthogonal to an imagined line between the cones (cp. Figure 3.1a), a pair of stimuli positioned on this orthogonal at a fixed distance from the cones should, on average, appear roughly in the center of participants' vision. Moreover, the event of riding the bicycle through the pair of cones is used to trigger the flickering of the stimuli. The distance between trials, the distance between cones and stimuli, as well as the timing between stimuli appearance and flickering has to be chosen carefully. Participants should have enough time to see the flickering of the stimuli happen, to form a decision



(a) The cyclist's point of view. The stimuli gem stones represented by dashed diamonds are not yet visible.

(b) Top-down view of the trial. Passing between the traffic cones (circles) triggers the appearance and then flickering of the stimuli (diamond shapes).

Figure 3.1 – Concept of one Temporal Order Judgment (TOJ) trial in the Virtual Cycling Environment (VCE)

about which stimulus flickered first, and then to react accordingly. At the same time, trials should not be too far apart in order to include a sufficient number of trials without the experiment becoming boring or fatiguing.

Tünnermann [30, Sec. 6.4.3] shows that it is possible to conduct TOJ experiments in dynamic environments such as a game and suggests that doing so "could be a tool for increasing the motivation of participants in attention experiments." Therefore, gamification is also an aspect of the experiment design. This should not only increase participants' motivation, but also their experience of flow in the simulation (cp. Section 2.3). According to Nakamura and Csikszentmihalyi [20], people can experience flow when they perceive their skills to be appropriate for the current set of challenges. TOJ tasks can be quite challenging by themselves, and since the experiment requires that participants make mistakes and trials are randomized, one's level of skill is often hard to judge. Embedding the TOJ experiment in a game, however, makes navigating traffic and collecting points the primary task from the perspective of participants, which can be assumed to provide a better judgment of skill. It also provides an opportunity to give clear and immediate feedback to participants, which constitutes another condition of flow [20]. In this case, feedback comes in the form of scores displayed on a virtual screen mounted to the handlebar. Since the scores are only visible for a short duration after collecting a stimulus, participants glancing down at the virtual display should not affect the results. 10 points are awarded for collecting the correct gem stone. To show participants that their order judgment was wrong, simply displaying a message *+0 points* would in principle suffice. However, the goal is to motivate participants, but the TOJ task is inherently difficult because the experiment requires participants to make mistakes. Therefore, 5 points will be awarded if the cyclist collects a wrong stimulus. In order to discourage reckless cycling and to incentivize adhering to traffic rules, collisions with cars result in the deduction of 20 points. The prospect of a favorable position in a high score list should provide motivation for participants to score as many points as possible.

The next important step is to plan the cyclists' route. It should be long enough to accommodate a sufficient number of trials, which for these TOJ experiments can be in the order of 90 trials or more per attention condition (e.g., per four conditions of stimulus orientation contrasts for a total of 360 trials [15]). On the other hand, it should not be so long that it will cause cyclists to lose their motivation to score as many points as possible. The plan for the experiment is to measure visual attention depending on traffic density as an independent variable. Traffic density is a within-subjects variable, which means that each participant of the experiment is exposed to each of two traffic density conditions: *low traffic* and *high traffic*. Therefore, the traffic density has to be different in different sections of the same scenario in a way that is predictable or that allows logging which condition a trial belongs to.

One straight-forward way to accomplish this would be to divide the route in half and to let each cyclist experience the *low traffic* condition in the first half and the *high traffic* condition in the second. However, this approach would introduce a likely confounding effect: assuming that participants perform worse in a TOJ task when fatigued [31], such an experiment would not allow to determine whether a lower level of visual attention was due to high traffic density or simply because the participant was getting tired. The inverse solution, where the participant starts in dense traffic and in the second half switches to less dense traffic, might result in the observation of a diminished effect for the same reason. Therefore, the traffic density conditions alternate frequently over the course of the experiment such that, in total, one participant experiences all conditions with approximately the same level of fatigue.

Before conditions of traffic density can be alternated over the course of an experiment, first a definition of traffic density is required. In this first experiment, only two conditions for traffic density are used: *low traffic* and *high traffic*. In the *low traffic* condition, no cars should be visible to the participant, and in the *high traffic* condition, some cars should be visible and in some way be relevant for the cyclist. A car that is relevant for the cyclist could be, for example, one that will intersect the cyclist's path at an intersection in contrast to a car coming from the front and turning right when the cyclist is also turning right. Having only these two conditions *low traffic* and *high traffic*, it is possible to check whether traffic density has an influence on visual attention at all, before possibly conducting more elaborate experiments in the future, for example regarding how many cars in what constellation influence visual attention to what precise degree. For the present experiment, the definition of traffic density can be more flexible. Indeed, it is possible to systematically vary traffic within the *high traffic* condition as a gamification feature in order to introduce difficulty gradients. Nevertheless, a balance must be struck between a minimum of cars to make a difference in comparison to the *low traffic* condition and having so many cars that the participant's view of the next trial is obstructed or simulator performance is degraded. In the first case, for example, if only one car were introduced for a *high traffic* segment of the experiment, this car might be gone before the participant had the opportunity to see all trials.

In short, participants of the experiment play a bicycling game with the goal to collect as many points as possible. They do so by traversing a route through city streets in which they encounter pairs of floating gem stones that flicker one after the other within each pair. If they collect that gem stone which flickered first, participants will gain the most points. At the same time, they have to watch out for cars that appear in some segments of the route (*high traffic*) but not in others (*low traffic*) and which can cost points if the cyclist collides with them. Details on the implementation are discussed in Chapter 4.

Chapter 4

Implementation

Based on the concept developed in Chapter 3, I will explain how my Temporal Order Judgment (TOJ) experiment is integrated in the Virtual Cycling Environment (VCE) and, in Section 4.1, how parts of this extension may be used for future experiments. In Sections 4.2 and 4.3, I will detail some of the general improvements made to the VCE.

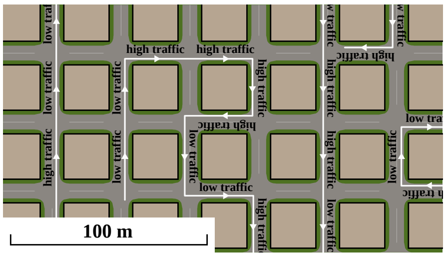
4.1 Experiment

This section describes the implementation of the conceptual TOJ experiment discussed in Chapter 3. An overarching goal of this particular implementation is to maintain some useful properties of the VCE, namely its modularity and the capability to quickly integrate and visualize new scenarios (cf. Section 2.1).

4.1.1 Trials

While navigating the virtual road network, cyclists find pairs of red bumps in intervals of 9 m, separated by a 1 m gap, which mark the beginning of a trial. While the red bumps may easily be exchanged for traffic cones as suggested in Chapter 3, the advantage of lower-profile bumps is that they do not result in an unusual and possibly distracting impression like traffic cones that do not interact with cars driving over them. The reported distances are a result of trial and error; given the accuracy of the first iteration of the steering sensor Android app (cf. Section 4.2) and preliminary data from one participant, I found that using these parameters I would be able to complete a sufficient number of about 350 trials in about 40 min. Navigating the bicycle through such a pair of red bumps will trigger the appearance of a pair of yellow stimuli, called *probe* and *reference*, in front of the cyclist at a distance of 4.5 m from the pair of red bumps and separated by 0.6 m. In order to minimize potentially

(a) The Virtual Cycling Environment (VCE)



(b) Schematic of the street network and route

(c) Unity screenshot: *high traffic*

(d) Unity screenshot: *low traffic* with tutorial

Figure 4.1 – Experimental setup and screenshots of the scenario for the first block of the experiment

confounding effects due to a decrease in contrast between turn indicator lights of cars and the stimuli, I added an additional bright turquoise outline and a texture of black stripes. A screenshot of such a trial is shown in Figure 4.1c.

The experiment requires high accuracy in the timing of stimulus events since temporal order judgments should be difficult enough for participants to make mistakes. For this purpose, the Unity simulation is configured to run at a fixed frame rate of 60 Hz in synchronization with the monitor refresh rate. This makes it possible to precisely time the flickering of stimuli by simply counting frames after a cyclist rides through a pair of red bumps and thereby triggers the beginning of a trial. In a preliminary test, I logged the wall clock time for key events of the TOJ trials in the Unity game engine. The results show that the measured time difference between the flickering of the probe stimulus and the flickering of the reference stimulus deviates from the desired time difference (i.e., the SOA) by at most 1.42 ms ($N = 104$, $M = -0.05$ ms, $SD = 0.76$ ms). This is well below the configured duration of 16.67 ms of one frame and it indicates that the frame rate is sufficiently constant for the experiment. After a randomized delay with a duration between 15 and 22 frames, the stimuli of the current trial flicker one after the other with a delay of -5 to 5 frames corresponding to the SOA. A negative SOA indicates that the probe stimulus will flicker first.

4.1.2 Road Network and Route

Depending on the strength of the possible effect of traffic density on visual attention, several hundred TOJ trials may be needed for sufficient statistical power to detect this effect. Placing this many trials in the form of Unity game objects by hand would be tedious and would make it harder to make quick adjustments to the road network. For this reason, I automated the placement of trials along a predefined route. In a realistic road network such as a section of Paderborn imported from OpenStreetMap, many appropriately spaced trials can fit on a single SUMO lane (i.e., a stretch of road spanning between two junctions), which under most circumstances should make defining a route by hand preferable over placing trials manually. Somebody designing a new scenario for an experiment will therefore be able to create a configuration file in the YAML format which simply lists all SUMO lanes of the route. Additionally, adjustments can be made on a per-lane basis to the spacing between trials or to the offset of trials from the start or the end of a lane. The trials are then first inserted as points of interest into the SUMO road network, from which they can later be imported and converted to the corresponding Unity game objects along with the generation of the rest of the three-dimensional world.

As explained in Chapter 3, it is desirable to have a significant portion of cars in the *high traffic* condition be relevant for the cyclist. This, however, is harder to achieve if junctions are far apart and when, because of this, cyclists need not worry about yielding to cars. Finding a real-world route that meets the criterion of having many tightly spaced junctions is also not easy. For this thesis, I therefore chose to use a synthetic road network. With the help of a SUMO tool called *netgenerate*, a grid-shaped network can be created with arbitrarily though uniformly spaced junctions. This is shown schematically in Figure 4.1b. The full road network consists of 15 × 15 junctions placed 35 m apart with buildings added in a post-processing step.

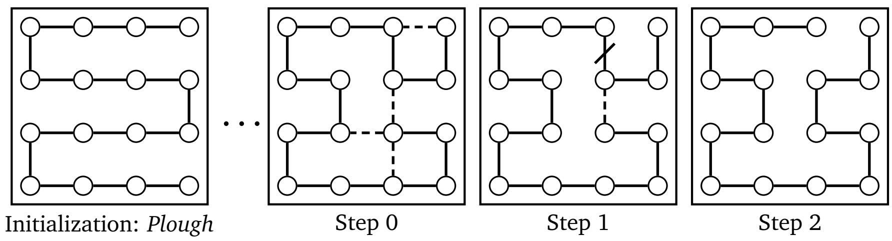


Figure 4.2 – Random Hamiltonian path generation, figure adapted from [32, FIG. 1 and FIG. 2]. In Step 0, dashed lines denote possible new edges. The dashed line in Step 1 is the edge to be created while the striked out edge is the one to be deleted.

A further convenience of a synthetic road network is the possibility to automate the generation of a route for the cyclist more easily. However, a rectangular circuit, a spiral, or a form of zig-zag route might be very predictable for the cyclist and could thereby hinder the goal of the gamification to preserve participants' motivation during the experiment. A randomized route would therefore be preferable. Oberdorf et al. [32] describe an algorithm for constructing and randomizing a Hamiltonian path over a lattice graph, i.e., a route that visits every node exactly once. As described in Algorithm 4.1 and as shown in the initialization step in Figure 4.2, the algorithm begins with a non-random configuration of a Hamiltonian path called *plough*. Afterwards, the initial configuration can be randomized by repeatedly applying what the authors call the *backbite move*. This is done by randomly choosing a potential new neighbor of either the current start node or the end node (Figure 4.2, Step 0; Algorithm 4.1, ll. 3f.). Using this neighbor to create a new edge will result in a cycle, which has to be fixed by deleting that edge of the selected neighbor which is part of this cycle and which is neither connected to the original start nor the original end node. Treating streets as edges and junctions as nodes, this algorithm can be applied to the synthetic grid road network.

Figure 4.3 shows some of the relevant classes for route generation and the automated placement of trials along such a route. These are also helpful for understanding the generation of traffic described in Section 4.1.3. The route randomization algorithm described above is implemented in the `GridEgoRoute` class, which inherits from `AbstractEgoRoute`. Alternatively, an `AbstractEgoRoute` can still be configured manually and loaded from a text file. For the placement of trials, it is necessary to iterate over all lanes of the route, which is made possible with the list of

---

Require: Lattice graph $G = (V, E)$ with vertices V and edges $E = \emptyset$, number of randomization steps M

1: Initialize G with the *plough* configuration (see Figure 4.2)
2: **for** $\_ = 0$ **to** $M - 1$ **do**
3:     $S = \{\{e, v\} \mid e \in V$ is an end point of the current Hamiltonian path, $v \in V$ is a potential new lattice neighbor of $v\}$
4:     Pick random edge $\{e', v'\} \in S$
5:     Insert $\{e', v'\}$ into E
6:     **for each** neighbor $v'' \neq e'$ of v' **do**
7:         **if** $\{v', v''\}$ is part of a cycle **then**
8:             Delete $\{v', v''\}$ from E
9:             **break**
10:         **end if**
11:     **end for**
12: **end for**

Algorithm 4.1 – Hamiltonian path randomization according to Oberdorf et al. [32]

`AbstractEgoRouteLane` instances assigned to each route. Not shown in Figure 4.3 is a specialized subclass for the lanes of a TOJ experiment, which contains additional parameters for the spacing of trials. Further subclasses exist for handling the simultaneous iteration over difficulty levels and lanes and for combining the functionalities of different `AbstractEgoRoute` subclasses. This will allow for some flexibility in the construction of future experiments, e.g., if the current form of difficulty levels or TOJ trials are not needed.

4.1.3 Traffic

In the experiment, cyclists will experience different traffic densities. To make this possible, traffic will have to be generated dynamically depending on the current experimental condition. For this purpose, I have extended the EVI with the ability to read a set of triggers from a specialized configuration file. Each trigger has a position in the road network and a radius. If, during the simulation, the ego vehicle comes within the radius of any one trigger, that trigger will be activated. For example, this trigger may then initiate the generation of traffic (optionally with a time delay) or activate signals for a specific vehicle. The EVI already uses SUMO's Traffic Control

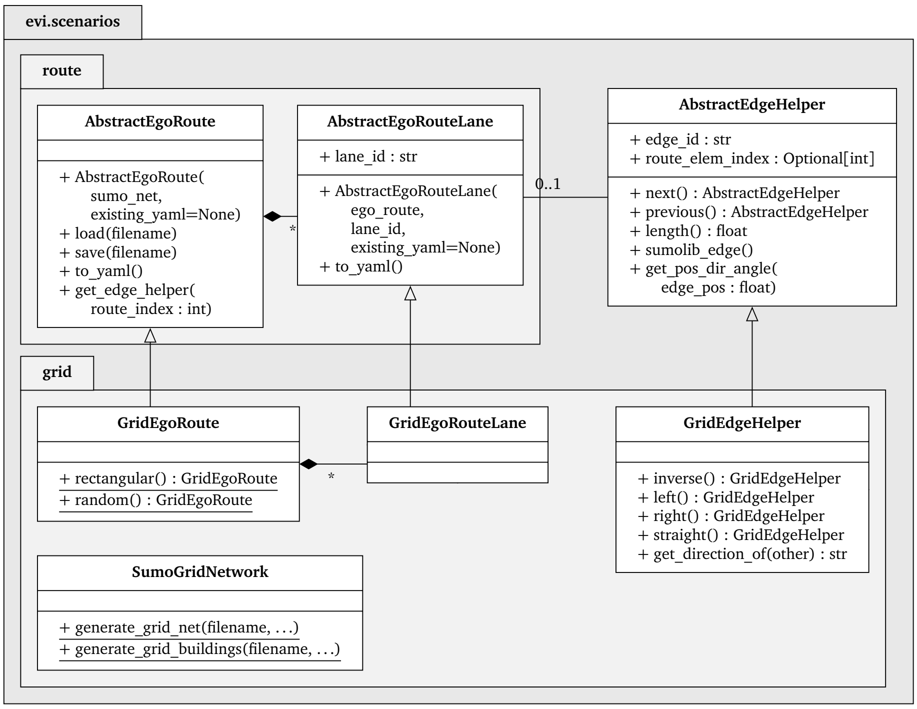


Figure 4.3 – UML class diagram of important classes for route and traffic generation

Interface (TraCI) for integrating the ego-vehicle into the traffic simulation. The dynamic triggers can act using the same TraCI connection.

Like the placement of experiment trials (Section 4.1.1) and the generation of the route (Section 4.1.2), the generation of traffic-generating triggers should be automated. Here again the class `AbstractEgoRoute` introduced in Section 4.1.2 and shown in Figure 4.3 is of great use. On the one hand it helps with traversing the cyclist's planned route and, in the case of the present thesis, place a traffic-generating trigger at the start of every lane that is assigned to the *high traffic* condition. On the other hand, each `AbstractEgoRouteLane` has access to an `AbstractEdgeHelper`, which in the case of a `GridEgoRoute` becomes a `GridEdgeHelper`. This helper class is useful for defining routes for other cars in the sense that it enables deviating from the cyclist's route via simple directives to move left, right, forwards, or backwards. This, in turn, makes it possible to precisely control how many cars will intersect the path of the cyclist and it can help ensure that traffic generated for the *high traffic* condition will not enter *low traffic* lanes. The latter aspect is especially important for the collection of clean data, i.e., participants should not see any cars when a stimulus pair of a *low traffic* lane flickers. At the start of a *low traffic* lane, this is avoided by not letting cars generated for a previous lane enter this new lane. At the end of the last *low traffic* lane before a *high traffic* lane, this is avoided by commencing the generation of traffic at the earliest when the cyclist enters the next junction.

4.2 Steering Sensor

An integral aspect of any interactive simulation is the user input interface. Using the VCE in its original form and especially before the replacement of the speed sensor (see Section 2.1), one quickly notices that the virtual bicycle does not react with sufficiently low latencies to pedaling or to turning the bicycle handle.

Several options can be considered for improving the steering experience. For example, Kooijman, Schwab, and Moore [33] attached a potentiometer to the head tube of a bicycle and connected it to the steerer tube with a belt. An advantage of this solution is the retained mobility of the bicycle. The original steering sensor solution of the VCE is integrated into the bicycle stand and thus would not allow collecting comparison data in field experiments. A disadvantage of the potentiometer solution is the more complex mounting mechanism, since attaching the belt might require disassembling parts of the bicycle and a stable physical attachment for the potentiometer would have to be developed (e.g., with the help of 3D printing). Early on during this thesis, a video project featuring the VCE called for a quick solution, which is why I chose an alternative approach.

Many modern smartphones have integrated sensors for detecting the Earth's magnetic field or for measuring the angular velocity of the phone. Such magnetometer and gyroscope sensors let software applications determine the current orientation of the phone, which is useful for compass, planetarium, or VR apps, as well as games. In this case, attaching a smartphone to the handle of the bicycle should allow us to obtain the steering angle so long as the bicycle itself remains stationary, which is the case for the VCE. The Android API[1] makes it easy to obtain the current azimuth of the device, i.e., the deviation angle from magnetic north in the horizontal plane. The thus obtained values can then be transmitted via UDP to the bicycle kinematics model (cf. Section 2.1).

Figure 4.4 was produced by recording traces of the values of both the original and the new steering sensor simultaneously while cycling in the VCE. It becomes clear that, over the course of this short test, new steering angle values virtually always arrive after at least half a second ($M = 0.65$ s, $SD = 0.27$ s) and occasionally take over an entire second. The apparent discreteness of the packet inter-arrival times of the original sensor suggests that at least one packet has been lost if the inter-arrival time is 1 s and at least two packets have been lost if it is 1.5 s, which indicates a packet loss of 29.9 %. In contrast, the new solution results in average packet inter-arrival times of only 0.05 s ($SD = 0.01$ s). With 99.94 % of all measured inter-arrival times lying below 100 ms, the interactivity of the VCE should be greatly improved. Delays introduced by wireless networking and routing in the local network

[1] https://web.archive.org/web/20190411205601/https://developer.android.com/guide/topics/sensors/sensors_position.html#sensors-pos-orient

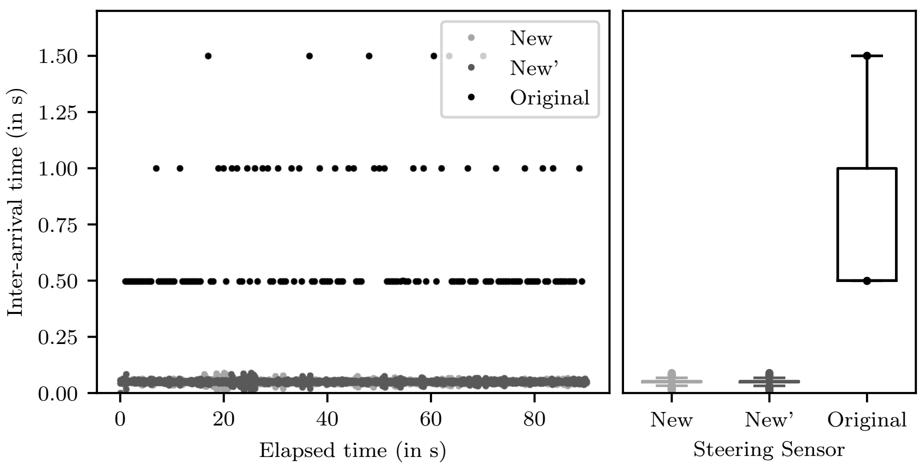


Figure 4.4 – Steering sensor packet inter-arrival times while cycling straight; recorded simultaneously for the original steering sensor, the first iteration of the smartphone app (*New*) used for the experiment, and the improved smartphone app (*New'*). The whiskers of the boxplots show the first and 99th percentiles.

will still have to be measured separately, but a different and more noticeable issue exists with the new steering sensor.

Besides latency, another important consideration is the accuracy of a sensor. As becomes apparent in Figure 4.5, the new solution produces far more noise than the original steering sensor. The figure shows the steering angles measured in the same run through the VCE in which data for Figure 4.4 was recorded. Since the bicycle was moving straight, steering angle values of the new sensor should roughly correspond to a constant value of 0 like the original solution, which exhibits a standard deviation of only 0.28°. In contrast, the standard deviation of steering angles measured with the smartphone app reaches 6.37°. (The initial spike is likely due to shaking when I first stepped onto the bicycle.) For cyclists in the VCE, these fluctuations become noticeable in the visualized bicycle handle randomly flickering left and right and in reduced steering precision. To some degree, this has been mitigated with an exponentially weighted moving average applied to sensor data within the smartphone app, where a user can configure the smoothing factor in a way that balances noise reduction and the delay induced by averaging. An alternative would be to apply sensor smoothing within the bicycle kinematics model, but since the smartphone app does not transmit all steering angle values that it measures, this solution would likely result in higher perceived delays.

Some time after the experiment (Chapter 5), it became clear that noise of the smartphone steering sensor could, in fact, be reduced to a large degree. Due to an oversight, I originally mistook the acceleration sensor for the gyroscopic sensor, which resulted in the steering accuracy relying fully on the accuracy of the smartphone's

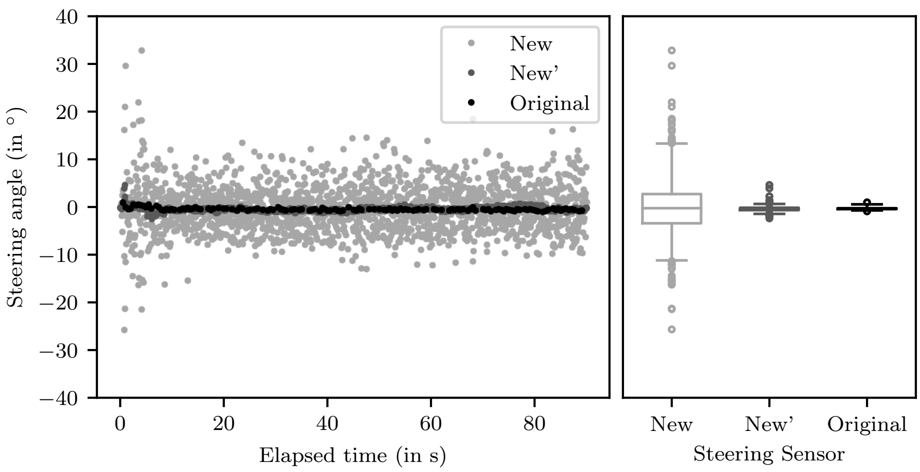


Figure 4.5 – Steering sensor values while cycling straight; recorded simultaneously for the original steering sensor, the first iteration of the smartphone app (*New*) used for the experiment, and the improved smartphone app (*New'*). The whiskers of the boxplots show the first and 99th percentiles.

magnetometer sensor. Switching to the correct composite sensor for obtaining a rotation vector from fused sensor data of the gyroscope and magnetometer, the noise could be greatly diminished. With a standard deviation of 0.44° and as can be seen in Figure 4.5, the improved smartphone steering sensor, by this metric, provides steering angles with almost the same consistency as the much less responsive original sensor.

4.3 3D Environment

In this section, I will discuss some of the visual improvements made to the VCE. In a meta-analysis, Cummings and Bailenson [16] come to the conclusion that image quality is a less important factor for the immersion in a virtual environment than the quality of the input method or a larger field of view. Nevertheless, it is clear that visual fidelity is far from irrelevant and given the level of detail in modern games and their popularity, prospective users of the VCE may have come to expect a certain level of visual realism. At the very least, the visual representation of the simulated environment should not be unnecessarily distracting for participants of the experiment. This could however plausibly happen due to visual artifacts that cause aspects of the virtual world to have an appearance contrary to real-world expectations. These will be detailed in Sections 4.3.1 and 4.3.2.

In addition to the visual changes presented in Sections 4.3.1 and 4.3.2, the following alterations were made:

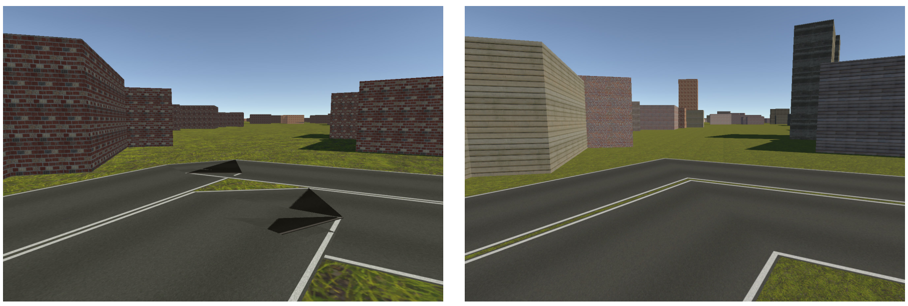
- As a replacement for a low-polygon-count bicycle model, Jan Tünnermann and I created a detailed model of the actual bicycle on the training stand of the VCE. The wheels, pedals, and the handle bar are animated.
- Turn signals of cars are now visible, allowing participants to predict where a car intends to go. Moreover, SUMO-supported signals for opening or closing the left or right car door can now be received and are visualized with the matching animation.
- Adopting the mechanism presented in Section 4.1.2 for placing TOJ trials on a predefined route, street signs can similarly be imported from SUMO points of interest. This includes priority and yield signs, but also a direction sign at each junction (similar to those described by Strayer, Drews, and Johnston [25] and Vlakveld et al. [28]) showing participants where to continue the experiment.
- A main menu was added to make it easier to configure network addresses, to select the type and controls of the ego vehicle, and to select a scenario.
- Users can now pause the simulation and will be shown a pause screen. This is also used as a gamification feature for presenting a participant's current

score when they complete a level or for showing a high score list when they complete a scenario.

4.3.1 Buildings

Buildings in the visualization component of the VCE are modeled as the vertically extruded shapes of their ground plan. So far, this means that buildings have no doors, windows, or non-flat roofs. This is in part owed to the fact that currently information about building height or roof type that is present in OpenStreetMap is not represented in the imported SUMO networks from which the visualization component generates the three-dimensional world. As Over et al. [34] note, an additional difficulty in the accurate representation of buildings lies in the way buildings are specified in OpenStreetMap, which often requires estimating the height of buildings merely from the number of floor levels. As a quick remedy to make the appearance of buildings slightly more diverse, I increased the variability in randomized building heights and I added four more textures besides the original brick material which is now randomly assigned to each new building.

However, a more pressing issue is the generation of building walls. The ground plans of buildings imported from SUMO do not specify vertices consistently in either clockwise or counterclockwise order. Because of this, it is not trivial to specify the vertices of the walls in clockwise order (from the perspective of an observer outside the building), which is required by Unity to make the walls face outside and to make them at all visible if backface culling is enabled. Previously, the VCE worked around this issue by rendering each building twice, flipping all faces in the second pass. This resulted in some walls being shaded dark despite them facing the sun, as can be seen in Figure 4.6a. A more efficient solution to this problem is possible by using the Surveyor's Area Formula [35], which provides a way of calculating the area of



(a) Before: Visual artifacts in roads and buildings **(b)** After: Fixed artifacts, more variation in building textures and heights

Figure 4.6 – Before and after comparison of roads and buildings in the Virtual Cycling Environment (VCE)

an arbitrary two-dimensional polygon, even if it is concave. For a polygon given by a list of vertices $(x_0, y_0), (x_1, y_1), \ldots, (x_{n-1}, y_{n-1})$ in counterclockwise order, the Surveyor's Area Formula is given as

$$A = \frac{1}{2}\left(\begin{vmatrix} x_0 & x_1 \\ y_0 & y_1 \end{vmatrix} + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \ldots + \begin{vmatrix} x_{n-2} & x_{n-1} \\ y_{n-2} & y_{n-1} \end{vmatrix} + \begin{vmatrix} x_{n-1} & x_0 \\ y_{n-1} & y_0 \end{vmatrix}\right) \tag{4.1}$$

If, however, the vertices are given in clockwise order, the value of A will be negative. In this case (without loss of generality), the vertex list can simply be inverted to obtain a consistent winding order for the ground plans of buildings. Additionally, this consistency allows the computation of correct surface normals. Even after fixing the winding order, sunlit walls all appear to have the same level of brightness irrespective of their angle to the sun. Using the above vertex list, let $(x_0, y_0), (x_1, y_1), (a_x, a_y), (b_x, b_y)$ denote the vertices of the first wall of a building. Before, it was not possible to tell whether (x_0, y_0) was the bottom left or the bottom right corner of a wall, again looking at a building from the outside. As a workaround, all normals were simply pointing in the global *up* direction. Now we can assume this vertex to be in the bottom left corner and use the right hand rule for calculating the normal vector:

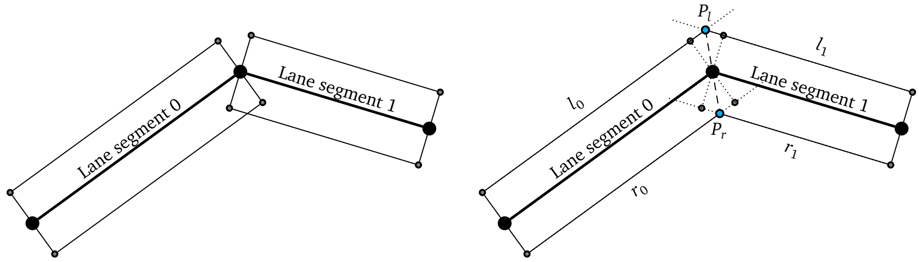
$$\begin{pmatrix} x_1 - x_0 \\ y_1 - y_0 \end{pmatrix} \times \begin{pmatrix} b_x - x_0 \\ b_y - y_0 \end{pmatrix} \tag{4.2}$$

Consequently, all face normals are now perpendicular to their corresponding wall and pointing outwards. As a result, walls can have different brightnesses depending on their angle to the sun as seen in Figure 4.6b.

4.3.2 Roads

As shown in Figure 4.6a, roads in the original implementation of the VCE sometimes exhibit flickering artifacts. These are due to z-fighting, which in this case occurs when one segment of a road overlaps with another. Examples can be found where such overlaps are already present in the SUMO road networks before they are imported into Unity. However, this is not the case in situations like shown in Figure 4.6a. Instead, the issue is caused by a Unity extension for modeling roads via splines. While such an extension can certainly be useful for modeling smooth roads by hand in a 3D environment based on just a few control points, in our use case it sometimes fails to handle sharp corners imported from SUMO files properly. In addition, while this spline-based approach would make it possible to connect road vertices coming from SUMO via smooth curves, this feature is not activated in the original VCE. Rather, SUMO road vertices are always connected by straight road segments. This has the benefit that no additional overlaps of buildings and near-by roads can accidentally occur because of such interpolations. It will not easily be possible to fix all cases of buildings overlapping with roads, however. This is because in scenarios imported from OpenStreetMap, the widths of roads have to be derived from discrete road types [34].

Some alternative ways for modeling roads based on map data exist in the literature. For example, Vaaraniemi, Treib, and Westermann [36] present a geometric approach that generates a quad face (i.e., a rectangle) for each line segment of a road. As shown in Figure 4.7a, this leads to overlapping road segments and gaps. The authors therefore propose extruding each road segment and using a specialized shader to draw rounded, half-circle caps at its ends. While this is a valid solution for their use-case where roads are drawn with solid colors and z-fighting would be invisible, it is harder to apply when roads have asphalt textures or lane markings. Under the prerequisite that adjacency information must be available for road segments, Vaaraniemi, Treib, and Westermann [36] also mention another solution, illustrated here in Figure 4.7b: By computing the intersection point P_l of their left sides l_0 and l_1 as well as the intersection point P_r of their right sides r_0 and r_1, road segments can be connected without overlaps. In the case of the VCE, adjacency information is available since it is required by the SUMO traffic simulation. This latter solution is furthermore preferable because it does not present the restriction of having to draw roads with solid colors. Instead, the asphalt textures of the original VCE can be re-used.



(a) Road segments as rectangles (quad faces) **(b)** Connected road segments

Figure 4.7 – New road geometry, figure adapted from [36]

Chapter 5

Empirical Study and Evaluation

In an experimental study, our Virtual Cycling Environment (VCE) [6] is used to measure cyclists' visual attention in the context of different traffic densities. This chapter outlines details about the experimental study and elaborates on the research methods and setup of the experiment. Finally, I present and discuss the results from this experiment.

5.1 Method

The foundations for this empirical study are presented in Chapters 2 and 3, which lead to the improvements and extensions to the VCE presented in Chapter 4. Based on this work, the present section presents details about how the experiment is arranged and how it was conducted.

5.1.1 Participants

A total of 19 participants (3 female and 15 male, and one choosing not to answer) took part in the experiment. Their age ranged from 14 to 30 years and with the exception of 2 persons, all participants were students. The experiment was conducted according to the principles expressed in the *Declaration of Helsinki* and approved by the ethics committee of Paderborn University. Participation was voluntary and without monetary compensation.

5.1.2 Apparatus

The experiment is conducted within the VCE [6], which I adapted as detailed in Section 4.1. In order to provide a wide field of view and to avoid simulator sickness, participants observe the visualization through a triple monitor setup instead of VR glasses. A wide field of view is considered beneficial for the immersiveness of a virtual environment [16], and it allows participants to still see cars approaching an intersection from the left or right. As shown in Figure 4.1a, the bicycle is positioned in front of a desk with three 24" 1920 × 1200 monitors positioned 1.5 m from the handle bar, which all operate at a refresh rate of 60 Hz in sync with the constant refresh rate of the visualization. Further details about the VCE can be obtained from Section 2.1.

5.1.3 Procedure

The experiment was split into two blocks of three levels each. This is in hopes of countering fatigue by allowing participants to take breaks after predictable intervals. After ensuring participants' informed consent, they were given the chance to familiarize themselves with the simulator in a short four-streets-long tutorial level. The tutorial included banners above the first virtual street explaining the rules of the game. Further, the participants were told that they would have the chance to enter themselves into a high score list at the end of each of the two experiment blocks. After each level, participants saw a pause screen with their total score in the current block, as well as their score in the previous level and the maximal attainable score. The motivation for the following choice of questionnaires is explained in Section 5.1.4. The Flow Short Scale (FSS) was always filled out immediately after the end of the first three levels. Afterwards, participants continued with the last three levels. Once finished, demographics and survey data for the custom VCE-related questions and for the Igroup Presence Questionnaire (IPQ) were collected. Each of the TOJ experiment blocks of around 80 road segments took about 20 min to complete. In total, participants typically completed the experiment in less than 60 min.

5.1.4 Design

In the experiment, I measure participants' visual attention depending on traffic density as a within-subjects independent variable. The *low traffic* condition is defined as an absence of other road users and *high traffic* as an average of 3.6 cars appearing in the cyclist's visual field in every street of this condition, which varies slightly across the three levels of each block. The variation is kept low for this experiment (between levels, the average number of randomly produced cars differs by less than 1) in order to not have participants wait too long if they have to yield to a queue of cars at an intersection and also to keep the temporal gaps between cars sufficiently short.

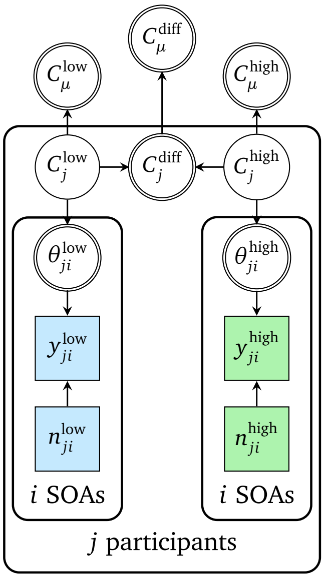
The hypothesis is that an increase in traffic density causes the allocation of more attentional weight to background stimuli such as the moving cars, which effectively reduces the capacity C available for the order judgment task. Given a participant j, we call the processing capacity allotted to the target stimuli of the task in the *high traffic* condition C_j^{high}, and C_j^{low} for the *low traffic* condition. C_μ^{high} and C_μ^{low} are the averages of these processing capacities over all participants. According to the hypothesis, we expect C_μ^{high} to be lower than C_μ^{low}.

After the experiment, participants receive a set of questionnaires on the features of the VCE with the aim to estimate the quality of the simulation as a whole: The Flow Short Scale (FSS) by Rheinberg, Vollmeyer, and Engeser [37], the Igroup Presence Questionnaire (IPQ) by Schubert, Friedmann, and Regenbrecht [19], and a collection of seven custom questions (cf. Table 5.1) specific to the VCE. Each questionnaire targets a different objective. In this case, the intention with measuring flow using the FSS is to get a sense of people's acceptance of the simulation and of their task regardless of whether they perceive the environment as realistic or not (see Section 2.3). Since an important goal of the VCE is to achieve a high degree of realism, the IPQ is included for its realness factor, which allows us to compare the perceived realism of the simulator with other publications. The custom questions, listed with results in Table 5.1, are intended to differentiate between several factors that might influence perceived realism, e.g., whether the behavior of other road users felt natural or to what degree the acts of accelerating and braking felt realistic. With the exception of one open-ended question, all custom questions allow responses on a 7-point scale ranging from *strongly disagree* (0) to *strongly agree* (6).

5.2 Results and Discussion

The experimental setup described in Section 5.1 yields the numbers of times participants collected the correct or incorrect gem stone of a stimulus pair under each condition. In order to estimate the visual processing capacities C_μ^{high} and C_μ^{low} for the high and low traffic density conditions, respectively, I used the Bayesian model depicted in Figure 5.1 based on the work of Tünnermann [30]. The variables n_{ji}^{high} and n_{ji}^{low} denote the number of trials for a given participant j and SOA i. Among these trials, y_{ji}^{high} and y_{ji}^{low} denote the number of times a participant reported the probe stimulus to have flickered first. Correspondingly, the function $P_{p^{\text{1st}}}(C, \text{SOA})$ is the probability that a given participant reports the probe stimulus to have flickered first. Based on the TVA, a concrete function can be derived that links the *probe first* report probability to TVA rate parameters v_p and v_r for probe and reference, respectively [30, p. 76]:

$$P_{p^{\text{1st}}}(v_p, v_r, \text{SOA}) = \begin{cases} 1 - e^{-v_p|\text{SOA}|} + e^{-v_p|\text{SOA}|}\left(\frac{v_p}{v_p+v_r}\right) & \text{, if SOA} < 0 \\ e^{-v_r|\text{SOA}|}\left(\frac{v_p}{v_p+v_r}\right) & \text{, if SOA} \geq 0. \end{cases} \quad (5.1)$$



$$
\begin{aligned}
C_j^{\text{low}} &\sim \text{Uniform}(0, 400)\\
C_j^{\text{high}} &\sim \text{Uniform}(0, 400)\\[4pt]
\theta_{ij}^{\text{low}} &\leftarrow P_{p^{\text{1st}}}\left(C_j^{\text{low}}, \text{SOA}\right)\\
\theta_{ij}^{\text{high}} &\leftarrow P_{p^{\text{1st}}}\left(C_j^{\text{high}}, \text{SOA}\right)\\[4pt]
y_{ij}^{\text{low}} &\sim \text{Binomial}\left(p = \theta_{ij}^{\text{low}}, N = n_{ij}^{\text{low}}\right)\\
y_{ij}^{\text{high}} &\sim \text{Binomial}\left(p = \theta_{ij}^{\text{high}}, N = n_{ij}^{\text{high}}\right)\\[4pt]
C_j^{\text{diff}} &\leftarrow C_j^{\text{high}} - C_j^{\text{low}}\\[4pt]
C_\mu^{\text{diff}} &\leftarrow \text{mean}\left(C_j^{\text{diff}}\right)\\
C_\mu^{\text{low}} &\leftarrow \text{mean}\left(C_j^{\text{low}}\right)\\
C_\mu^{\text{high}} &\leftarrow \text{mean}\left(C_j^{\text{high}}\right)
\end{aligned}
$$

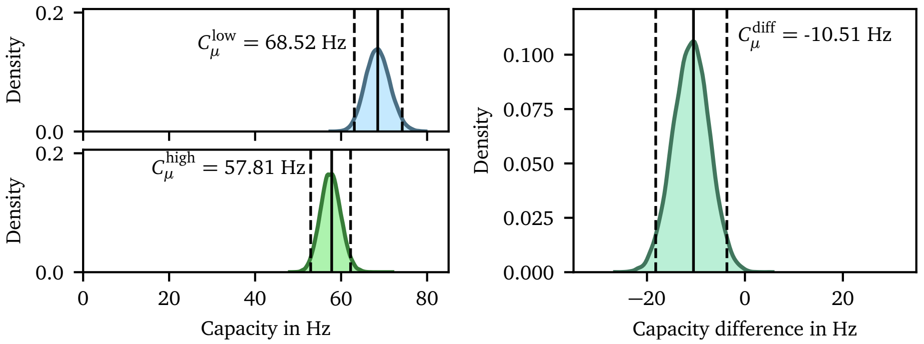
Figure 5.1 – Bayesian model: Shaded nodes denote observed variables, discrete variables are shown as rectangles, continuous variables as circles, and deterministic variables have a double border (cf. [38]).

In this thesis, I am interested in the visual processing capacity that is allocated to the two stimuli of the TOJ task. The parameter C as calculated with Equation 2.1 is therefore only concerned with exactly two stimuli of the same appearance. Consequently, we can assume $v_p = v_r = \frac{C}{2}$, which leads to

$$
P_{p^{\text{1st}}}(C, \text{SOA}) = \begin{cases} 1 - \frac{1}{2}\exp\left(-\frac{C}{2}|\text{SOA}|\right) & \text{, if SOA} < 0 \\ \frac{1}{2}\exp\left(-\frac{C}{2}|\text{SOA}|\right) & \text{, if SOA} \geq 0. \end{cases} \tag{5.2}
$$

The model was estimated using the NUTS sampler [39] (20 000 iterations in four chains with the default parameters) with a given implementation in PyMC3 [40]. The priors on the C parameters were set to a uniform distribution from 0 Hz to 400 Hz. This is a conservative choice, given that typical values for C range from 50 Hz to 70 Hz [30].

The results indicate that the estimated visual processing capacity in the *high traffic* condition is lower than in the *low traffic* condition, cf. Figure 5.2a. This confirms the hypothesis that C_μ^{high} is lower than C_μ^{low}. The individual capacities C_μ^{high} and C_μ^{low} had their modes at 57.2 Hz and 69.0 Hz, respectively, and 95 % Highest Density Intervals (HDIs) of [52.5 Hz, 62.1 Hz] and [63.6 Hz, 75.3 Hz]. The 95 % HDI for the capacity difference C_μ^{diff} did not include a value of 0 Hz difference. Instead, the 95 % most credible values ranged from −19.4 Hz to −4.4 Hz with a mode of



(a) Top: *low traffic*, bottom: *high traffic*

(b) Difference between *low* and *high traffic* conditions

Figure 5.2 – Posterior distributions of visual processing capacity; the 95 % Highest Density Intervals (HDIs) are overlaid as dashed lines

−12.2 Hz, cf. Figure 5.2b. These results are in line with findings by Strayer, Drews, and Johnston [25] and Vlakveld et al. [28], who found an effect of traffic density and traffic complexity on mental workload.

In a next step, we have a closer look on the participants' acceptance of the virtual environment and their gamified task to collect the correct gem stones. The FSS yielded a mean flow score of 4.92 ($SD = 0.87$). To put this flow score into context, we can compare it to a number of reference activities by Rheinberg, Vollmeyer, and Manig [41, p. 32]. Of the 9 activities with a lower score, notable ones include watching TV or sleeping, while the 17 higher-rated examples include reading and social communication, but also household chores and doing office work.

Results from the evaluations indicate promising opportunities for improvements of the VCE. One possible reason for the disruption of the experience of flow is

Item		*M*	*SD*
• Using the simulator felt like riding a real bicycle.		2.58	1.46
• Steering felt realistic.		2.11	1.52
• Accelerating and braking felt realistic.		3.89	1.41
• The virtual bicycle behaved according to my expectations from experience with real bicycles.		2.63	1.38
• I felt confident in my ability to control the bicycle.		2.11	1.63
• Other road users behaved according to my expectations from the real world.		3.47	1.87
• I tried to adhere to traffic rules as much as in the real world.		3.74	1.52

Table 5.1 – Responses to Virtual Cycling Environment (VCE) related questions of the questionnaire (from 0, strongly disagree to 6, strongly agree)

a participant not feeling like they are in control [20]. Responses to the custom questionnaire items listed in Table 5.1 showed low ratings for participants' confidence in their control of the bicycle and for their judged realism of the steering mechanism. Written comments in response to the question *What should be changed in order to improve the realism or the usability of the cycling simulation?* (see Chapter A) show a similar picture: the aspect of steering was mentioned negatively 11 times. Another frequent comment (10 counts) was in regard to the behavior of other virtual road users, which frequently did not yield to bicyclists although traffic signs suggested otherwise. These results present a possible explanation for the outcome of the IPQ, which produced an average realness factor of 2.20 ($SD = 0.92$). This places the realness score of the VCE in the 58th percentile of valid submissions in the reference data given for the IPQ.[2]

As an implication of participants' feedback, improving the accuracy of the steering sensor and the behavior of simulated cars is likely to significantly improve the VCE in terms of usability and participants' perceived realism. Other possible improvements that are hinted at in the written feedback (Chapter A) include providing more variation in the set of simulated road users, e.g., by adding buses or pedestrians. Additionally, environmental sounds like noise from wind, tires, or from other cars could be integrated and more realistic and diverse buildings as well as foliage could be generated procedurally. One participant suggested placing 3D models of people on the seats of simulated cars. Taking this idea one step further, another visualization of a person may be seated on the virtual bicycle. If this virtual person can be persuaded to pedal and steer in synchrony with the real bicycle, this could significantly contribute to an increased sense of actually sitting on the virtual bicycle oneself [42]. Finally, many participants, verbally or in writing, complained about the bicycle being uncomfortable. A different type of saddle and handlebar could alleviate this issue.

In general, the VCE can be recommended for human-in-the-loop experiments with some constraints. An effect of traffic density on cyclists' visual attention has been demonstrated and the effect is in line with existing literature. This is encouraging in that it shows that the VCE can be used to obtain plausible data in human-in-the-loop experiments. However, the flow and realism scores lagged somewhat behind expectations. As mentioned above, this is likely due to the steering issues and the unpredictability of how simulated cars react to the bicycle. In this sense, the scores and the proportion of participants reporting these problems highlight their importance. Therefore, experimenters who plan to use the VCE in the future should think about the impact the issues might have on their results and consider addressing these issues first. On the upside, as explained in Section 4.2, the steering sensor accuracy has been improved significantly since the experiment. Furthermore, in

[2] Reference data was obtained from `http://www.igroup.org/pq/ipq/data.php` (2019-03-10)

comparison to steering accuracy and the behavior of simulated cars, none of the remaining issues brought forth by participants seem nearly as critically important for the suitability of the VCE for human-in-the-loop experiments.

Chapter 6

Conclusion

In this thesis, I evaluated the suitability of the VCE as a human-in-the-loop simulator for investigating future ADAS for VRUs. Based on a formal theory of visual attention, I presented a method for measuring cyclists' visual attention in simulation. By conducting an empirical study with multiple participants, I measured the impact of road traffic density on the visual attention of cyclists. The results indicate that high density traffic requires measurably more visual attention than low density traffic. This demonstrates the sensitivity of the VCE for measuring differences in visual attention and is in line with recent studies regarding the mental workload of road users [25], [28], thus suggesting the suitability of the VCE for in-depth psychological studies.

Furthermore, results from the questionnaires indicate that there is still room for improvements in the VCE. Since a common complaint of participants was the inaccuracy of the steering sensor, which since the experiment has much improved, it stands to reason that flow and perceived realism scores should be significantly higher in a hypothetical follow-up experiment. If, beyond this, the bicycle can be made more comfortable and virtual cars can be made to react more predictably to the presence of the cyclist, the most pressing concerns expressed by participants could already be resolved.

In future work, the VCE can now be used to test an actual ADAS for VRUs. In addition, further efforts should be made to increase confidence in the validity of future experimental data. While the visual attention experiment presented above aligns with results of other studies of road users, similar results can be expected if the TOJ paradigm were implemented in a less visually realistic virtual environment. Questionnaires on perceived realism as also employed in this thesis can provide further insights, but ultimately, it would be desirable to quantify the reality gap by comparing detailed real-world bicycling behavior with the behavior of cyclists in the VCE. Finally, since the magnitude of the effect of no traffic compared to a high traffic density on visual attention has now been established, future studies may explore the details of this effect. For example, how much does each additional car on the road contribute, is there indeed a difference between cars that are relevant for the cyclist and those that are not (cp. Chapter 3), and to what degree can the effect be mitigated by efforts to separate bicycle lanes from motor lanes or other changes in the configuration of roads and intersections? Eventually, such future work may contribute to significant improvements in the safety of cyclists and possibly VRUs in general.

List of Abbreviations

ADAS	Advanced Driver Assistance System
CAM	Context Awareness Message
DENM	Decentralized Environmental Notification Message
EVI	Ego-Vehicle Interface
FSS	Flow Short Scale
HDI	Highest Density Interval
ICT	Information and Communication Technology
IPQ	Igroup Presence Questionnaire
PDT	Peripheral Detection Task
SOA	Stimulus Onset Asynchrony
TOJ	Temporal Order Judgment
TraCI	Traffic Control Interface
TVA	Theory of Visual Attention
V2X	Vehicle-to-X
VANET	Vehicular Ad-hoc Network
VCE	Virtual Cycling Environment
VR	Virtual Reality
VRU	Vulnerable Road User

List of Figures

Bibliography

[1] EC, "Road safety in the European Union," European Commission, Report MI-AC-18-001-EN-N, Apr. 2018. DOI: 10.2832/060333.

[2] K. Bengler, K. Dietmayer, B. Farber, M. Maurer, C. Stiller, and H. Winner, "Three Decades of Driver Assistance Systems: Review and Future Perspectives," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 4, pp. 6–22, Oct. 2014. DOI: 10.1109/MITS.2014.2336271.

[3] S. Biswas, R. Tatchikou, and F. Dion, "Vehicle-to-Vehicle Wireless Communication Protocols for Enhancing Highway Traffic Safety," *IEEE Communications Magazine*, vol. 44, no. 1, pp. 74–82, Jan. 2006. DOI: 10.1109/MCOM.2006.1580935.

[4] C.-H. Chang, C.-H. Chou, C.-J. Lin, and M.-D. Lee, "A WAVE/DSRC-based intersection collision warning system," in *International Conference on Ultra Modern Telecommunications (ICUMT 2009)*, St. Petersburg, Russia: IEEE, Oct. 2009. DOI: 10.1109/ICUMT.2009.5345520.

[5] S. Loewen, F. Klingler, C. Sommer, and F. Dressler, "Backwards Compatible Extension of CAMs/DENMs for Improved Bike Safety on the Road," in *9th IEEE Vehicular Networking Conference (VNC 2017), Poster Session*, Torino, Italy: IEEE, Nov. 2017, pp. 43–44. DOI: 10.1109/VNC.2017.8275657.

[6] J. Heinovski, L. Stratmann, D. S. Buse, F. Klingler, M. Franke, M.-C. H. Oczko, C. Sommer, I. Scharlau, and F. Dressler, "Modeling Cycling Behavior to Improve Bicyclists' Safety at Intersections – A Networking Perspective," in *20th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM 2019)*, to appear, Washington, DC: IEEE, Jun. 2019.

[7] L. Stratmann, D. S. Buse, J. Heinovski, F. Klingler, C. Sommer, J. Tünnermann, I. Scharlau, and F. Dressler, "Psychological Feasibility of a Virtual Cycling Environment for Human-in-the-Loop Experiments," in *1st Workshop on ICT based Collision Avoidance for VRUs (ICT4VRU 2019)*, in peer-review, Kassel, Germany, Sep. 2019.

[8] P. Alvarez Lopez, M. Behrisch, L. Bieker-Walz, J. Erdman, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wießner, "Microscopic Traffic Simulation using SUMO," in *21st International Conference on Intelligent Transportation Systems (ITSC)*, Maui, HI: IEEE, Nov. 2018, pp. 2575–2582. DOI: 10.1109/ITSC.2018.8569938.

[9] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent Development and Applications of SUMO - Simulation of Urban MObility," *IARIA International Journal On Advances in Systems and Measurements*, vol. 5, no. 3&4, pp. 128–138, Dec. 2012.

[10] D. S. Buse, M. Schettler, N. Kothe, P. Reinold, C. Sommer, and F. Dressler, "Bridging Worlds: Integrating Hardware-in-the-Loop Testing with Large-Scale VANET Simulation," in *14th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2018)*, Isola 2000, France: IEEE, Feb. 2018, pp. 33–36. DOI: 10.23919/WONS.2018.8311659.

[11] C. Sommer, R. German, and F. Dressler, "Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis," *IEEE Transactions on Mobile Computing*, vol. 10, no. 1, pp. 3–15, Jan. 2011. DOI: 10.1109/TMC.2010.133.

[12] M. Wierda and K. A. Brookhuis, "Analysis of Cycling Skill: A Cognitive Approach," *Applied Cognitive Psychology*, vol. 5, no. 2, pp. 113–122, Mar. 1991. DOI: 10.1002/acp.2350050205.

[13] C. Bundesen, "A Theory of Visual Attention," *Psychological Review*, vol. 97, no. 4, pp. 523–547, Oct. 1990. DOI: 10.1037/0033-295X.97.4.523.

[14] J. Tünnermann, A. Petersen, and I. Scharlau, "Does attention speed up processing? Decreases and increases of processing rates in visual prior entry," *Journal of Vision*, vol. 15, no. 3, pp. 1–27, Mar. 2015. DOI: 10.1167/15.3.1.

[15] A. Krüger, J. Tünnermann, and I. Scharlau, "Measuring and modeling salience with the theory of visual attention," *Attention, Perception, & Psychophysics*, vol. 79, no. 6, pp. 1593–1614, Aug. 2017. DOI: 10.3758/s13414-017-1325-6.

[16] J. J. Cummings and J. N. Bailenson, "How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence," *Media Psychology*, vol. 19, no. 2, pp. 272–309, May 2016. DOI: 10.1080/15213269.2015.1015740.

[17] R. Skarbez, F. P. Brooks Jr, and M. C. Whitton, "A survey of presence and related concepts," *ACM Computing Surveys (CSUR)*, vol. 50, no. 6, 96:1–96:39, Nov. 2017. DOI: 10.1145/3134301.

[18] J. Psotka, "Immersive training systems: Virtual reality and education and training," *Instructional Science*, vol. 23, no. 5, pp. 405–431, Nov. 1995. DOI: 10.1007/BF00896880.

[19] T. Schubert, F. Friedmann, and H. Regenbrecht, "The Experience of Presence: Factor Analytic Insights," *Presence: Teleoperators and Virtual Environments*, vol. 10, no. 3, pp. 266–281, Jun. 2001. DOI: 10.1162/105474601300343603.

[20] J. Nakamura and M. Csikszentmihalyi, "The Concept of Flow," in *Flow and the Foundations of Positive Psychology*, Dordrecht, Netherlands: Springer, 2014, pp. 239–263. DOI: 10.1007/978-94-017-9088-8_16.

[21] D.-S. Kwon, G.-H. Yang, C.-W. Lee, J.-C. Shin, Y. Park, B. Jung, D. Y. Lee, K. Lee, S.-H. Han, B.-H. Yoo, et al., "KAIST Interactive Bicycle Simulator," in *IEEE International Conference on Robotics and Automation (ICRA 2001)*, Seoul, South Korea: IEEE, May 2001, pp. 2313–2318. DOI: 10.1109/ROBOT.2001.932967.

[22] A. L. Schwab and A. M. Recuero, "Design and experimental validation of a haptic steering interface for the control input of a bicycle simulator," in *ECCOMAS Multibody Dynamics 2013*, Zagreb, Croatia: ECCOMAS, Jul. 2013, pp. 103–110.

[23] A. Matviienko, S. Ananthanarayan, S. S. Borojeni, Y. Feld, W. Heuten, and S. Boll, "Augmenting Bicycles and Helmets with Multimodal Warnings for Children," in *20th ACM International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI 2018)*, Barcelona, Spain: ACM, Sep. 2018, 15:1–15:13. DOI: 10.1145/3229434.3229479.

[24] J. M. Plumert, J. K. Kearney, and J. F. Cremer, "Children's Perception of Gap Affordances: Bicycling Across Traffic-Filled Intersections in an Immersive Virtual Environment," *Child Development*, vol. 75, no. 4, pp. 1243–1253, Jul. 2004. DOI: 10.1111/j.1467-8624.2004.00736.x.

[25] D. L. Strayer, F. A. Drews, and W. A. Johnston, "Cell Phone-Induced Failures of Visual Attention During Simulated Driving," *Journal of Experimental Psychology: Applied*, vol. 9, no. 1, pp. 23–32, Mar. 2003. DOI: 10.1037/1076-898X.9.1.23.

[26] L. L. Di Stasi, V. Álvarez-Valbuena, J. J. Cañas, A. Maldonado, A. Catena, A. Antolí, and A. Candido, "Risk behaviour and mental workload: Multimodal assessment techniques applied to motorbike riding simulation," *Elsevier Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 12, no. 5, pp. 361–370, Sep. 2009. DOI: 10.1016/j.trf.2009.02.004.

[27] G. Marquart, C. Cabrall, and J. d. Winter, “Review of eye-related measures of drivers’ mental workload,” *Procedia Manufacturing*, vol. 3, pp. 2854–2861, Oct. 2015. DOI: 10.1016/j.promfg.2015.07.783.

[28] W. P. Vlakveld, D. Twisk, M. Christoph, M. Boele, R. Sikkema, R. Remy, and A. L. Schwab, “Speed choice and mental workload of elderly cyclists on e-bikes in simple and complex traffic situations: A field experiment,” *Accident Analysis & Prevention*, vol. 74, pp. 97–106, Jan. 2015. DOI: 10.1016/j.aap.2014.10.018.

[29] E. Hartmann, B. Lachenmayr, and H. Brettel, “The peripheral critical flicker frequency,” *Vision Research*, vol. 19, no. 9, pp. 1019–1023, 1979. DOI: 10.1016/0042-6989(79)90227-X.

[30] J. Tünnermann, “On the Origin of Visual Temporal-Order Perception by Means of Attentional Selection,” PhD Thesis (Dissertation), Faculty of Arts and Humanities, Paderborn, Germany, Jul. 2016.

[31] Á. Csathó, D. van der Linden, I. Hernádi, P. Buzás, and G. Kalmár, “Effects of mental fatigue on the capacity limits of visual attention,” *Journal of Cognitive Psychology*, vol. 24, no. 5, pp. 511–524, May 2012. DOI: 10.1080/20445911.2012.658039.

[32] R. Oberdorf, A. Ferguson, J. L. Jacobsen, and J. Kondev, “Secondary structures in long compact polymers,” *Physical Review E*, vol. 74, no. 5, pp. 051801-1–051801-9, Nov. 2006. DOI: 10.1103/PhysRevE.74.051801.

[33] J. D. G. Kooijman, A. L. Schwab, and J. K. Moore, “Some Observations on Human Control of a Bicycle,” in *7th International Conference on Multibody Systems, Nonlinear Dynamics, and Control*, San Diego, CA, USA: ASME, Sep. 2009, pp. 2021–2028. DOI: 10.1115/DETC2009-86959.

[34] M. Over, A. Schilling, S. Neubauer, and A. Zipf, “Generating web-based 3D City Models from OpenStreetMap: The current situation in Germany,” *Computers, Environment and Urban Systems*, vol. 34, no. 6, pp. 496–507, Nov. 2010. DOI: 10.1016/j.compenvurbsys.2010.05.001.

[35] B. Braden, “The Surveyor’s Area Formula,” *The College Mathematics Journal*, vol. 17, no. 4, pp. 326–337, Sep. 1986. DOI: 10.2307/2686282.

[36] M. Vaaraniemi, M. Treib, and R. Westermann, “High-Quality Cartographic Roads on High-Resolution DEMs,” *Journal of WSCG*, vol. 19, no. 1-3, pp. 41–48, 2011.

[37] F. Rheinberg, R. Vollmeyer, and S. Engeser, “Die Erfassung des Flow-Erlebens,” in *Diagnostik von Motivation und Selbstkonzept*, J. Stiensmeier-Pelster and F. Rheinberg, Eds., Göttingen, Germany: Hogrefe, 2003, pp. 261–279.

[38] M. D. Lee and E.-J. Wagenmakers, Bayesian Cognitive Modeling: A Practical Course. Amsteradam, Netherlands: Cambridge University Press, Jun. 2014. DOI: 10.1017/CBO9781139087759.

[39] M. D. Hoffman and A. Gelman, "The No-U-turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1593–1623, Jan. 2014.

[40] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck, "Probabilistic programming in Python using PyMC3," *PeerJ Computer Science*, vol. 2, e55, Apr. 2016. DOI: 10.7717/peerj-cs.55.

[41] F. Rheinberg, R. Vollmeyer, and Y. Manig, "Flow-Erleben: Untersuchungen zu einem populären, aber unterspezifizierten Konstrukt: Abschlussbericht," University of Potsdam, Potsdam, Germany, Tech. Rep., Oct. 2006.

[42] M. Slater, B. Spanlang, M. V. Sanchez-Vives, and O. Blanke, "First Person Experience of Body Transfer in Virtual Reality," *PLoS ONE*, vol. 5, no. 5, pp. 1–9, May 2010. DOI: 10.1371/journal.pone.0010564.

Appendix A

Detailed Results

Participants wrote the following comments in response to the question *What should be changed in order to improve the realism or the usability of the cycling simulation?*

- Adding other objects like trees or different buildings, also adding people walking
-
  - Lenkung hat zu viel geruckelt und war nicht responsiv genug
  - Autos sollten mich nicht von hinten ueberfahren
  - Brems- und Beschleunigungsverhalten ist zu responsiv
-
  - Make cars adhere to the right-of-way rules more properly
  - Make the rendering / terrain a bit more realistic
  - When accelerating fast on the physical bike, there sometimes was stuttering in the simulation
- einfachere Lenkung, Ton, abwechslungsreichere Haeuser, reales Fahrrad naeher an die Monitore, VR, unterschiedliche Autotypen
-
  - Bequemerer Sitz
  - Autos sollten sich mehr an Verkehrsregeln halten.
- The difficulty of turning the bicycle was much higher in the virtual world most probably because of the friction from the turning surface on which the front tires of the bicycle were resting. This was the only huge factor that made true immersion into the virtual world a little difficult.
-
  - more comfortable bike
  - waay better steering and speed controls
  - no driverless cars

  - cars should respect driving rules

- - teilweise zittert das virtuelle Lenkrad sehr
  - manchmal faehrt das virtuelle Fahrrad langsam weiter, obwohl man eigentlich steht
- Man sollte leichter lenken koennen, die Autos sollten sich besser an Verkehrsregeln halten
- - Die anderen Autofahrer sollten sich entsprechend der Schilder verhalten.
  - Insgesamt waere eine etwas kuerzere Experimentendauer schoener gewesen, da es keine grossen Aenderungen zwischen den Leveln gab
- Das Lenken ist teilweise sehr schwierig und tr[bt das virtuelle Erlebnis
- - Verkehrsregeln bei Autos einbauen
  - Keine Abstuerze
- Autos sollten Verkehrsregeln besser achten, Lenker manchmal schwammig
- I found it hard to steer the bicycle. In some situation I needed to move in a criss-cross pattern, which should have felt more smoothly as it felt right now.
- - Lenken muesste leichtgaengiger sein
  - Es sollte nicht moeglich sein, zwischen den Zielen hindurchzufahren
  - Man sollte umdrehen koennen
- - Sitzposition insbesondere Lenker,
  - Lenkergriffe nach einiger Zeit unbequem
  - Fahrzeugverhalten ist manchmal seltsam (Vorfahrtsregeln)
  - Neigung beim Fahrradfahren wichtig
- Sounds would further help to provide an atmosphere. I also miss the traffic behind me. Much of my experience in traffic has to do with somebody behind you waiting for you to cross the intersection. I could always wait far away from an intersection before the car traffic could recoqnize me and therefore avoid taking part in traffic. The short bursts of cars further distances me from reality as I began to expect only a bunch of cars and after that I could roam freely again.