Evaluating Cyclist Behavior in Different Roadway Designs through Immersive Virtual Environments and Psychophysiological Sensing

A

Dissertation

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

in partial fulfillment

of the requirements for the degree

Doctor of Philosophy

by

Xiang Guo

December 2022

APPROVAL SHEET

This

Dissertation

is submitted in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

Author: Xiang Guo

This Dissertation has been read and approved by the examing committee:

Advisor: Arsalan Heydarian

Advisor:

Committee Member: T. Donna Chen

Committee Member: Sara Lu Riggs

Committee Member: Jeffrey Woo

Committee Member: Nan Li

Committee Member:

Committee Member:

Accepted for the School of Engineering and Applied Science:

Jennifer L. West, School of Engineering and Applied Science

December 2022

Evaluating Cyclist Behavior in Different Roadway Designs through Immersive Virtual Environments and Psychophysiological Sensing

ABSTRACT

As a healthier and more sustainable way of mobility, cycling has been advocated by literature and policy. However, current trends in cyclist crash fatalities suggest deficiencies in current roadway design in protecting these vulnerable road users. The lack of cycling data is a common challenge in studying cyclists' safety, behavior, and comfort levels under different design contexts. The integration of human sensing technologies has greatly facilitated the human-centered design evaluation of road infrastructures. Another question of interest is cyclists' distraction and involvement in secondary tasks. Understanding cyclists' behaviors under the influence of distraction can provide evidence for interventions to address safety-related issues. This study will focus on cognitive distraction as it is related to the most frequently reported secondary tasks during cycling, such as listening to music or talking in the earphones.

To understand cyclists' behavior in different contextual settings, an Immersive Virtual Environment(IVE) bicycle simulator - Omni-Reality and Cognition Lab Simulator (ORCLSim) is developed with the ability to collect the following data: cycling performance (speed, steering, braking, acceleration and lane position), eye Tracking (gaze direction, fixation), physiological responses (Heart rate, head movement, hand acceleration), video recording and stated preferences surveys (subjective rating). With the ORCLSim system framework, this proposal aims to study the effect

of different roadway designs (external factors) and psychophysiological states (internal factors) on cycling behavior with the goals of (1). Capturing and analyzing cyclist behavior and psychophysiological responses within an immersive virtual environment. (2). Validate the use of an IVE-based bike simulator with multimodal data sources for cyclist study. (3). Evaluate alternative designs for cyclists with human-centered and data-driven methods. (4). Study the effect of cognitive distraction on cyclist behavior.

Specifically, in addition to the system setup integration and development, three experimental studies are conducted: (1). Benchmark the use of IVE for cyclist behavioral study. Cyclist behavior and psychophysiological responses will be compared between the real-world environment and IVE to validate what set of information in IVE is representative of the real world (n=6). (2). Evaluation of different design alternatives in the IVE. Three different roadway designs will be evaluated in the ORCLSim system framework to test if the cyclists' physiological response is different from their stated preferences and actual cycling behaviors (n=50). (3). Study the effect of cognitive distraction on cyclist behavior. Two types of cognitive distractions are tested: listening to music and talking on the phone with earphones. Both the standardized secondary task (Mock phone conversation task) and the actual secondary task (music listening) will be introduced in the experiment to test if the standardized secondary task can be applied to simulate cyclists' cognitive workload and the effect of cognitive distraction on cycling behaviors (n=75).

The results indicate (1). Most of the performance measurements have absolute validity, the IVE bike simulator can be further utilized for understanding cyclists' behaviors. (2). It is important to track physiological metrics to better understand how different settings may impact cyclists. Additionally, we showcased the importance of gaze tracking and heart rate data in capturing the behavioral response to different events or roadway settings. (3). Bicycle infrastructure can meaningfully impact cyclists' movement and psychophysiological responses. The protected bike lane design has the highest subjective safety rating, lowest cycling speed, and highest lateral distance to the vehicle lane, indicating the potential for safer bicycling behavior with lower speeds and increased separation from vehicles; cyclists focus their gaze on the cycling task more in the separate and protected bike lane scenarios; creating separate zones for bicyclists (whether separate bike lane or protected bike lane) has the potential to reduce the stress level, as indicated by decreased HR changes compared to the shared bike lane. (4). Differences are found in cyclists' adaptive behaviors with different types of cognitive distractions. Talking on the phone is rated as the most distracting scenario, cyclists would keep a lower speed with less input power and less head movement variation. While listening to music, cyclists would have a significantly higher speed, a lower standard deviation of speed, and higher input power. The introduction of bike lanes has the potential to stabilize the lateral lane position. Demographic information is also found to affect behavioral and psychophysiological responses.

Based on the findings and results of this dissertation, suggestions are made for policymakers and designers.

Dedication

Dedicated to my family for their love and support

Acknowledgments

This dissertation is an exciting yet challenging milestone of my Ph.D. journey. It would not have been possible to write this dissertation without the help and support of many people in my life.

First of all, I want to express my greatest appreciation to my advisor Professor Arsalan Heydarian for his encouragement, patience, and support during this journey. We started our work with the exploration of various projects and I have been always amazed by his knowledge of different research fields and insight into scientific problems. Professor Heydarian has given me plenty of freedom to navigate my research and life, and his high-level research taste and critical thinking ability always inspire me with appropriate advice and guidance on details.

I also would like to express my deepest thanks to my dissertation committee members, Professor T. Donna Chen, Professor Sara Riggs, Professor Jeffrey Woo, and Professor Nan Li for serving as committee members. Especially, I want to thank Professor T. Donna Chen for her nice suggestions and insightful advice for all the work in the ORCL lab. Without her support, I wouldn't be able to finish all the experiments in the dissertation.

I would like to extend my sincere thanks to Dr. Inki Kim for his support at the beginning of my research life at UVa and Professor Laura Barnes for her help during the transition period. In addition, I would like to thank Professor Vahid Balali and Professor Omidreza Shoghli for their advice on related research projects.

I'm especially thankful to Dr. Arash Tavakoli, whose research experience and knowledge helps me a lot in our cooperation. I want to thank Dr. Austin Angulo and Dr.

Erin Robartes for their great prior work in the ORCL lab. I would like to thank the following persons for their willingness to help during the experiments: Gabriel Witter, Andrea Saglio, Kunsh Singh, Sepehr Sabeti, and Hunter Ruebsamen. I am grateful to the members of the BRAIn Lab at UVA, especially, Mahsa, Alan, Siavash, Amir, Beatrice, Fengyi, and Reza, thank you for your suggestion in my research. I really enjoy working with all of you.

I also want to thank my friends, this journey would not be possible without their love and support. I want to thank Yichen Jiang for your patience in constantly listening to me and talking things out, for proofreading my papers, and for all your love and support over the years. I want to thank Erfan, Wen, Chang, Difan, Bruce, Xiafei, Junyu, Shan, Boyang, and many others for making my life in the US meaningful.

Finally, but most importantly, I would like to thank my parents, Qianjun Guo and Meisu Bao, you have always stood behind me, without your unconditional encouragement, I wouldn't be able to make it. I'm so proud of you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Cycling, as a traditional means of mobility, has gained more popularity in recent years. Bicycle mode share is rising in response to issues in modern cities, such as increases in traffic jams, land use, energy consumption, air pollution, climate change, and physical inactivity (Flusche 2012; Rupi and Krizek 2019). This trend has continued during the COVID-19 pandemic, as it is reported that cycling levels have significantly increased in many countries despite lockdowns and travel restrictions (Buehler and Pucher 2021). A study in Europe estimated that substituting a bike over a car just once a day reduces an average citizen's carbon emissions from transport by 67% (Brand et al. 2021). However, with a growing number of vehicles on the roads, there are more and more traffic conflicts between cyclists and vehicles. There has been an alarming increase in cyclist fatalities over the last decade. The National Highway Traffic Safety Administration (NHTSA)'s report shows that in the United States, the number of cyclists' fatalities has increased by more than 35% since 2010 (NHTSA 2021a). The report from the World Health Organization (WHO) also shows that more than half of global road traffic deaths (1.35 million each year) are among pedestrians, cyclists, and motorcyclists (Organization 2018).

Over the past couple of decades, the evaluation of roadway safety and design has been automobile-centric. The development of automotive technology has decreased the overall fatalities and fatality rate per 100 million Vehicle miles traveled (VMT) in traffic crashes. For example, in the United States, the fatality rate per 100 million VMT has decreased from 3.35 to 1.10 since 1975 (Administration et al. 2020). Many studies have been conducted to evaluate the impact of roadway design features on drivers' behaviors and safety, leaving out other roadway users such as cyclists. These trends indicate that the design of our roadways needs to be improved to be more inclusive for all users, especially for vulnerable road users such as cyclists (Rodriguez-Valencia et al. 2021).

To better inform roadway design, extensive datasets similar to the automobile-focused studies of the past are needed for cyclists. To develop robust cyclist-focused datasets, studies with both high ecological and internal validity are needed. Ecological validity refers to the extent an experimental environment matches with the real world, increasing the chances that the effects identified in an experimental environment generalize to real-world settings. Internal validity refers to the extent to which a cause-effect relationship is warranted in a study.

Various methods of collecting bicycle safety, risk, comfort, and behavior data have been utilized in the past. These studies spanned surveys, observational studies, naturalistic studies, and simulation studies. Subjective studies, such as surveys, provide measures of users and their perceptions of their environment but lack ecological and internal validity (Wynne, Beanland, and Salmon 2019). For example, interviews and surveys ask participants about their behaviors and comfort in a certain design context either by imagination or after an actual bike ride. However, the subjective response does not always reflect what the participant will do in a real-road setting and can suffer from hypothetical bias (Fitch and S. L. Handy 2018). Observational studies can record realistic changes within the environment and cyclists' responses in real-world conditions, but are unable to track cyclists' physiological changes. Naturalistic studies can provide information about realistic changes within the environment and cyclists' behavior with high ecological validity. Naturalistic studies can further record cyclists' responses in real-world conditions through different sensing modalities such as GPS, ECG, or mobile eye trackers (Rupi and Krizek 2019). However, these studies with lower internal validity are resource- and time-extensive, and have potential risks of injuries and fatalities for participants as they may be placed in dangerous roadway settings (e.g. distraction in high traffic density area) (Stelling-Konczak et al. 2018). Furthermore, naturalistic studies are influenced by many environmental factors that restrict the ability to fully isolate and understand the impact of independent variables, thus offering low internal validity, especially for physiological and behavioral factors (Fitch, Sharpnack, and S. L. Handy 2020; Teixeira et al. 2020). Thus far, the majority of cyclist studies rely on subjective and naturalistic data derived from real-world settings to assess participants' behavior and comfort in different traffic environments (Paridon et al. 2019).

Experimental studies provide an opportunity to evaluate the impact of safety-related conditions, infrastructure, and technology on cyclists as they can offer the ecological validity lacking in subjective studies and allow the researchers to control for external variables, unlike naturalistic studies, for greater internal validity. Experimental studies conducted through Immersive Virtual Environments (IVE) are an emerging approach that minimizes the hypothetical bias of subjective surveys while offering a controlled, low-risk, and immersive environment to evaluate the responses of cyclists to different roadway designs and conditions. The benefit of an Immersive Virtual Environment (IVE) is achieving high internal and ecological validity while also being cost-effective, and offering complete experimental control to replicate trials (Heydarian, Carneiro, et al. 2015; Heydarian and Becerik-Gerber 2017). Early

IVE lacked realism, which was primarily due to a lack of technological capability. Fortunately, IVE software and hardware platforms have significantly improved over the last few years with the release of high-end, commercially available head-mounted displays (HMD), and graphics cards capable of rendering highly detailed environments. Meanwhile real-life components can also be integrated into IVE studies. For instance, we can now integrate a physical bicycle along with haptic feedback with an IVE to create a more immersive setting for participants (Guo, Robartes, et al. 2021). Furthermore, as the level of immersion increases, we can expand these experiments to capture participants' physiological and psychological factors, which is a field of data that has historically been overlooked. Such data provides insights into how participants' behaviors and perceptions may change in contextual settings in different research fields (J. Kim, Yadav, Ahn, et al. 2019; G. Lee et al. 2020; Adami et al. 2021; Noghabaei, K. Han, and Albert 2021; Sharif and Oppenheimer 2021). With the increase of realism in IVE simulators and the development of low-cost ubiquitous sensors, IVE simulators have become promising tools for conducting highly realistic and immersive experimental studies (Ergan et al. 2019). In traffic safety studies, driving simulators have been widely applied to study drivers' behaviors, awareness (Soares, Ferreira, and Couto 2020, and psychophysiological states with multimodal data collection systems, such as eye trackers, electroencephalogram (EEG) and electrocardiogram (ECG) (Haufe et al. 2011; Akbar et al. 2017; Guo, Cui, et al. 2019; Baee et al. 2021). Some of the driver-related studies are conducted in IVE (Eudave and Valencia 2017). Meanwhile, for cyclists, only a few studies have applied physiological responses in IVE simulators. For example, cyclists' galvanic skin response had less peaks with a bike lane than with no bike lane condition (Cobb, Jashami, and Hurwitz 2021). In another cycling virtual reality study, electroencephalography (EEG) data shows its potential in a hybrid model framework as an indicator of the perceived risk of cyclists (Bogacz, Hess, Calastri, Choudhury, Mushtaq, et al. 2021). Overall, we still have very a limited understanding of cyclists' physiological responses in different roadway environments, especially in IVE studies. Some naturalistic studies have been conducted to evaluate cyclist's behavior and physiological responses in different contextual settings (Guo, Robartes, et al. 2021; McNeil, Monsere, and Dill 2015; Rupi and Krizek 2019; Teixeira et al. 2020). These preliminary studies revealed that psycho-physiological metrics (e.g., heart rate (HR), gaze variability, and skin conductance) are indicators of how participants' behaviors and perceptions may change in different contextual settings.

Other questions of interest are cyclists' distraction and involvement in secondary tasks. Distraction has been identified as one of the main reasons for traffic accidents. In the US, nine percent of fatal crashes, 15 percent of injury crashes, and 15 percent of all police-reported motor vehicle traffic crash in 2019 were reported as distraction-affected crashes. There were 566 vulnerable road users including pedestrians, pedalcyclists, and others killed in distraction-affected crashes (NHTSA 2021b). This data is from vehicle crash reports, and limited information about cyclists' distraction on the fatality is available. No only the limited data sources, but also the number of studies that have been published on distracted biking is small. As a result, our knowledge about the effect of distraction on cycling is insufficient. Previous studies have reported that distractions have a major prevalence among bike users and that they play a significant role in the prediction of the traffic crash rates of cyclists, through the mediation of risky behaviors (Useche et al. 2018). Studying cyclists' behaviors under the influence of distraction can provide evidence for interventions to address safety-related issues. Similar to the diving distraction, the cycling distraction can be categorized into three main types: Visual (taking the eyes off the road), Manual(taking the hands off the handlebar), and Cognitive (taking the mind off cycling). This study will focus on cognitive distraction as it is related to the most frequently reported secondary task during cycling, such as listening to music or talking in the earphones (Mwakalonge, White, and Siuhi 2014; Wolfe et al. 2016).

To understand cyclists' behavior in different contextual settings, with the development of an IVE bicycle simulator, this proposal aims to study the effect of different roadway designs (external factors) and psychophysiological states (internal factors) on cycling behavior with the goals of 1). Capturing and analyzing cyclist behavior and psychophysiological responses within an immersive virtual environment. 2). Validate the use of an IVE-based bike simulator with multimodal data sources for cyclist study. 3). Evaluate alternative designs for cyclists with human-centered and data-driven methods. 4). Study the effect of different types of cognitive distractions on cyclist behavior.

Specifically, the dissertation has three main research objectives (**RO**), and each research objective will be studied by an experimental study to answer the corresponding research questions (RQ):

- **RO 1**: Benchmark the use of an Immersive Virtual Environment (IVE) for behavioral study.
  - RQ 1: What set of information can we capture in the Immersive Virtual Environment (IVE) bike simulator?
  - RQ 2: Will the cyclists' cycling/Heart Rate/eye tracking behavior in the IVE be representative of the real world?
- **RO 2**: Evaluation different design alternatives in the IVEs

  - RQ 3: Can the proposed framework support the IVE bike simulator to study cyclists' behaviors Not different roadway designs?
  - RQ 4: Are the cyclists' physiological responses different from their stated preferences?

- **RO 3**: Study the effect of cognitive distraction on cyclist behavior
  - RQ 5: Can we use the standardized secondary task to simulate cognitive distraction during cycling?
  - RQ 6: What's the effect of different types of cognitive distraction on cycling behavior?

Chapter 2 introduces more details of literature reviews, the Chapter 3 describes the development of all components of the ORCLSim IVE bicycle simulator, the data collection and analysis methods which will serve as the foundations of all three studies. Chapter 4 introduces the benchmarking study between the real world and the IVE simulator to test the representative of the data collected from the proposed ORCLSim framework. Chapter 5 introduces the experiment designs, data analysis, and results for different roadway designs in the IVE. Chapter 6 the experiment about the effect of different types of cognitive distraction on cycling behavior and physiological response. Chapter 7 concludes the findings from all the studies and summarizes the future directions of the dissertation.

Chapter 2

Review of Literature

2.1 Types of Methods for Cyclist Study

Accident crash reports, stated preference surveys, observational studies, field test/-naturalistic studies, and simulator studies have been utilized by past researchers for cycling studies. Most of the accident crash reports are led by governments or international organizations, such as NHTSA and WHO (Organization 2018; NHTSA 2021a). The crash reports provide valuable real-world data, but they are post-accident analyses, the information for cyclists is often incomplete, which means limited behavioral or psychophysiological data is included.

Surveys have been widely used to study cyclists' behavior, particularly when faced with a lack of real-world data. Surveys, when composed carefully, can efficiently assess large populations of cyclists and have been used to study a wide variety of topics such as perceived safety/comfort (Parkin, Wardman, and Page 2007; Chaurand and Delhomme 2013; Abadi and Hurwitz 2018), route choice (Sener, Eluru, and Bhat 2009) and crash history (Robartes and T. D. Chen 2018; Poulos et al. 2012; H. Yang et al. 2019). For instance, (Chaurand and Delhomme 2013) studied the perceived risk of cyclists and drivers in certain interactions. Results suggest that perceived risk is higher for drivers compared to cyclists. Additionally, the perceived risk of cyclists is higher when interacting with a car than with another bike. Another study investigated the perceived level of comfort by cyclists near urban truck loading zones in varying conditions of truck traffic, bicycle lane marking type, and traffic signs (Abadi and Hurwitz 2018). Results indicate that the existence of trucks in the traffic is a significant factor in reducing cyclists' perceived comfort. Additionally, the study finds that women are generally more affected by the truck traffic than men. While these types of studies add significant value to our understanding of the effect of contextual settings on cyclists' perceived risk and comfort, they often are riddled with issues such as limitations on external validity. For instance, these studies may suffer from hypothetical bias, in which participants' responses to surveys may not reflect their real response in a naturalistic situation (Fitch and S. L. Handy 2018). For instance, Fitch and S. L. Handy 2018 reports that imagined ratings of comfort while biking may have a negative bias as high as 15% difference in comfort and safety when compared to real-world situations. Additionally, surveys and subjective measures generally cannot be used to understand the temporal dimension of the effect of certain contextual elements on cyclists. For instance, in the case of perceived comfort, it is not possible to understand the exact moment in which a bicyclist felt discomfort, or to what level the discomfort varies among different people and at different locations. On the other hand, physiological measures can be used as surrogate metrics to understand the time span of contextual elements' effects on the participants.

Observational studies can minimize the risk of hypothetical bias from stated preference surveys and provide a real-world assessment of bicyclists' responses in specific locations (Thompson et al. 2013). For example, an observational study conducted in Boston, United States, investigated distracted cycling behavior and reported the prevalence of two types of distractions: auditory (ear buds/phones in or on ears), and visual/tactile (electronic device or other objects in hand). Almost one-third of all bicyclists exhibited distracting behavior at four high-traffic intersections during peak commuting hours. The highest proportion of distracted bicyclists was observed during the midday commute (between 13:30-15:00) (Wolfe et al. 2016). Another observational study with 2187 cyclists in Germany shows 22.7% bicyclists are engaged in a secondary task such as wearing headphones or earphones (13.1%) or interacting with other cyclists (7.0%) (Huemer, Gercek, and Vollrath 2019). In observational studies, the collected data relies only on the behavioral responses that the observers can visually discern, without having the ability to manipulate different factors, such as traffic density or noise levels (Daniels, Nuyts, and Wets 2008). In recent years, the utilization of cameras has greatly increased the popularity of video-based observational studies. For instance, an observational study in China recorded 112 hours of video footage with 13,407 bicyclists riding shared bikes and 2061 riding personal bikes. Not wearing a helmet, violating traffic lights, riding in the opposite direction of traffic, not holding the handlebar with both hands, and riding in a non-bicycle lane are identified as top unsafe behaviors (Gao et al. 2020). Overall, observational studies can only evaluate bicyclists' behaviors in existing environments, and are unable to collect bicyclist's psycho-physiological responses. Psycho-physiological measures of bicyclists may be helpful for understanding the reason behind distraction, the length of the distraction, and to what level the stimulus affects the rider's decision-making. To have full control over design considerations, we need to evaluate how cyclists respond to different designs of roadways during the planning or design phase of projects.

The development of mobile sensing technologies has enabled mobility data collection from a variety of modalities. In this line of research, several naturalistic studies have investigated cyclists' behaviors and physiological responses, such as HR, heart rate variability (HRV) (Doorley et al. 2015), and gaze (Rupi and Krizek 2019). For example, the changes in ambient light levels can affect bicyclists' perception of the environment by changing their gaze reactions (Uttley, Simpson, and Qasem 2018). These studies show preliminary evidence on how infrastructure design is associated with cycling stress, although variations are found between different cities and road types (Fitch, Sharpnack, and S. L. Handy 2020; Teixeira et al. 2020; Guo, Robartes, et al. 2021). In Fitch, Sharpnack, and S. L. Handy 2020, through using a BodyGaurd II heart beat-to-beat interval measuring device, the HRV results from 20 female participants suggest that only local roads with no dividing yellow lines and low car speed and volume bicyclists provided less stress to the participants compared to collector (medium to high dividing traffic) and arterial roads (high volume, multi-lane). One of the limitations identified in the study is that environments with protected or separate bike lanes are not included, so the results cannot indicate how those designs might compare to the as-built designs (Fitch, Sharpnack, and S. L. Handy 2020). Lack of environmental control (e.g., traffic, weather, etc.) could be the main cause of the uncertainty, which undermines the interpretability of these results and suggests the need for further research. Additionally, only a limited number of human sensing devices can be applied in naturalistic settings, as many of these devices are intrusive. For example, the electroencephalogram (EEG) measurement devices are more capable for in-lab testing. This will have an effect on participants' behavior and safety, which may result in degraded data quality. Lastly, the potential risks of injuries and fatalities for participants trigger ethical concerns for naturalistic studies. For example, Stelling-Konczak et al. 2018 conducted a study in real traffic examining the glance behavior of teenage bicyclists when listening to music. The study was terminated after fourteen participants when the results indicated that a substantial percentage of participants cycling with music experienced a decrease in their visual performance (Stelling-Konczak et al. 2018).

2.2 Bicycle Simulator With IVE Study

Over the past decade, driving simulators, virtual reality (VR) technologies, and human sensing technologies have provided new insights into human behavior in different contextual settings, assisting in evaluating different design alternatives for roadways (X. Zou et al. 2021), buildings (Francisco et al. 2018), hospitals (Chías et al. 2019, and other civil infrastructure systems (H. Zou, N. Li, and L. Cao 2017; Noghabaei and K. Han 2020; Awada et al. 2021). Simulation methods utilizing IVE offer a low-cost, low-risk approach to studying the users' safety, perception, and behavior. Traditionally, real-world observation methods have been used to understand bicyclist and pedestrian behavior. These methods are often expensive, time-consuming, and unrealistic for studying naturalistic behaviors as they often require some level of unrealistic environmental control for the safety of test subjects. The improvements in IVE over the recent years have provided researchers, designers, and engineers with a way to evaluate alternative infrastructure design while providing high degrees of immersion. Novel, commercially available VR headsets offer a high degree of realism and immersion. Furthermore, environmental factors that may influence bicyclist behavior are highly controllable within IVE, allowing for replicable experimental trials.

The advancements in VR and bicycle simulation over the past decade have led to a rapid increase in its application among researchers, designers, and engineers to evaluate human responses to alternative infrastructure designs. The combination of IVE and instrumented physical bicycle simulators provides a high level of immersion and flexibility in experimental designs. Furthermore, it enables user engagement and allows subjective analysis of participants to better understand their behavior and preferences to the changes in simulated environments (Nazemi, Eggermond, Erath, and Axhausen 2018). For instance, Xu, Lin, and Schmidt 2017 was able to evaluate 30 participants' behaviors in an IVE, by designing a straight path with four sections of varying traffic conditions. Through this experiment, results suggest that the existence of a bike lane in low traffic conditions significantly improved cyclist lane-keeping performance (Xu, Lin, and Schmidt 2017). A more recent study compared cycling behaviors in an IVE between a keyboard-controlled bicycle and an instrumented bicycle where participants could pedal. The results indicated that there is more variance in the instrumented bicycle experiments in different measurements, such as speed, head movement, acceleration, and braking behaviors (Bogacz, Hess, Calastri, Choudhury, Erath, et al. 2020). Validation studies were also performed to compare the bicycling behavior between the IVE bicycle simulator and naturalistic studies (O'Hern, Oxley, and Stevenson 2017; Guo, Robartes, et al. 2021). Although there is a limited number of validation studies, promising results are shown in the validity of cycling performance such as lane position and speed (O'Hern, Oxley, and Stevenson 2017). However, most IVE-related studies are limited to observing cycling behaviors and preferences without exploring bicyclists' psycho-physiological responses.

2.3 Measurement of Physiological Responses

Past studies have pointed out the utility of humans' physiological signals such as cardiovascular (e.g., HR), skin temperature, skin conductance, brain signals, gaze variability, and gaze entropy in understanding emotions, stress levels, anxiety, and cognitive load (Tavakoli, Kumar, Boukhechba, et al. 2021; H.-G. Kim et al. 2018; Lohani, Payne, and Strayer 2019). Apart from subjective studies, there are limited datasets including human physiological and psychological sensing (e.g. eye tracking, body tracking, and heart rate) for cyclists. It is crucial to assess participants' patterns of perception and reaction in certain contextual settings. Many traditional on-road studies have used accident statistics and road infrastructure data (e.g., roadside cameras) to evaluate the safety-related concerns of cyclists and pedestrians. To further study their perception and cognitive states, human sensing devices (e.g. physiology devices) have been shown to provide promising insights (Ridel et al. 2018; Trefzger et al. 2018; Ayres et al. 2015). There are practical concerns about the data collection of human sensing on real roads. First, safety, ethical and cost considerations prohibit large-scale on-road experiments (Stelling-Konczak et al. 2018). Second, the implementation of traditional human sensing devices (such as body trackers) is intrusive, which may affect the behavioral and perception ability of the participant, as well as the data quality on real roads (especially in high-speed scenarios) (Van Hoof et al. 2012). In light of these shortcomings, most IVE can handle the first limitation, as virtual environments provide a low-risk and cost-effective alternative to real settings. While the second shortcoming (monitoring perception and cognition) requires the integration of human sensing systems and ubiquitous computing into the IVE. The majority of existing IVE research in bicyclist and pedestrian studies has not utilized ubiquitous computing and human sensing techniques to monitor participants' behaviors and physiological states.

2.3.1 Eye-Tracking

Eye-tracking behaviors are usually found to be related to the process of cognitive resource allocation. Eye-tracking behavior is usually measured by optical eye-trackers, it has been widely used in studying users' visual perception and attention in different contexts. Different features such as blinking rate, saccade and fixation duration, gaze variability in different directions, stationary gaze entropy (SGE), and gaze transition entropy (GTE) are correlated to different states such as workload, stress level, and emotions (B. Shiferaw, Downey, and Crewther 2019). A fixation in eye patterns refers to maintaining the eye gaze on a specific location (Purves et al. 2001). At each fixation, the gaze is approximately stationary. The transitions between fixations are called saccades where the point of fixation changes rapidly to a new fixation point. The variation and sequence of fixations and saccades were shown to be correlated with human states such as stress and workload (May et al. 1990). In addition to fixation and saccade, gaze variability is an additional feature that refers to the standard deviation in gaze angles in both vertical and horizontal directions.

In general, two measures can be calculated for entropy. The first one is based on the definition of uncertainty associated with a choice (B. Shiferaw, Downey, and Crewther 2019). With more randomness in a system, the entropy also increases. This is calculated through Shannon's equation (Shannon 1948). In the gaze analysis research, this first entropy is referred to as the stationary gaze entropy (SGE), which shows the overall predictability of fixation locations and can be a proxy for the gaze dispersion (Shannon 1948). For a set of fixation locations in a sequence of eye movements, if we assign fixation locations to spatial bins of p_i, we can calculate the SGE as:

$$SGE = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{2.1}$$

Different studies have used SGE for human state analysis. For instance, SGE was used for detecting task demand, complexity, experience, workload, drowsiness, and being under the influence of alcohol (B. Shiferaw, Downey, and Crewther 2019). The second measure of gaze entropy is gaze transition entropy (GTE), which is the conditional entropy that takes into account the temporal dependency between different fixations. GTE is a measure of the predictability of the next fixation location given the current fixation location. For a sequence of transitions between different spatial bins of i and j, with a probability of p_{ij} the GTE can be calculated as:

$$GTE = -\sum_{i=1}^{n} p_i \sum_{j=1}^{n} p_{ij} \log_2 p_{ij} \tag{2.2}$$

Conceptually, for each specific combination of task demand and scene complexity, an optimal level of GTE exists (B. Shiferaw, Downey, and Crewther 2019). Deviation from the optimal GTE can provide information about changes in the human state. For example, an increase in stress, anxiety, and frequency of emotional episodes are associated with an increased level of GTE (relative to the optimum). While a decrease in the level of GTE (relative to the optimum) can be due to the usage of depressants such as alcohol (B. Shiferaw, Downey, and Crewther 2019).

In addition to gaze entropy, gaze data is usually analyzed based on the Area of Interest (AOI). People divert their attention away from the previous fixation to another, which reflects the changes in mental concentration through AOI. By mapping fixations with the AOI, it is possible to obtain a statistical description of key gaze parameters, including the fixation duration, fixation counts, and saccade counts. In transportation-related studies, road center is a frequently used AOI (Wang, Reimer, et al. 2014). The percentage of fixations in the road center AOI is referred to as the percentage of road center (PRC) feature (Guo, Jiang, and I. Kim 2020). PRC has been shown to increase with elevated cognitive demand (Engström, Johansson, and Östlund 2005).

For bicycle-related studies, eye-tracking behaviors have been analyzed in several real-world experiments. In a naturalistic study from Germany, 20 participants cycled at five defined test locations while wearing a mobile eye-tracking system. The outcome shows that spatially open locations are related to a higher level of perceived risk, and more cycling experience and greater familiarity with a location may lead to a more foresighted and focused gaze behavior (Stülpnagel 2020). Another naturalistic study in Italy investigated bicyclists' eye gaze behavior at signalized intersections. They collected this data through a mobile eye tracker from 16 participants in a 3-kilometer corridor. The results show that intersections that force bicyclists to merge with vehicle traffic yield notable differences in features of gaze behavior. For instance, when approaching intersections, the moment of the first fixation on the traffic lights occurs earlier for the case of no bike lane as compared to the case with a separate bike lane. Additionally, for inexperienced cyclists, intersections without a separate bike lane were associated with an increase in gaze variability and looking around (Rupi and Krizek 2019). The latest virtual reality headsets, such as the HTC VIVE Pro Eye, have integrated eye-tracking features, allowing IVE researchers to incorporate eye-tracking analysis within their studies. To our knowledge, no previous research has studied bicyclists' gaze behaviors in IVE with bicycle simulators.

2.3.2 Heart Rate

Electrocardiography (ECG) is a well-established method to record the electrical activity of the heart. A participant's heart rate (HR) and heart rate variability (HRV) can be measured using an ECG signal. Studies in different areas have used HR measures to analyze different human states in response to changes within an environment. While in medical applications HR is generally retrieved through devices such as Electrocardiography(ECG), wearable-based devices (e.g., smartwatches) generally use photoplethysmogram (PPG) technology. ECG measures the electrical activity of the heart through the application of contact electrodes, while PPG records the blood volume in veins using infrared technology. The blood volume measurement is then used to estimate the HR (i.e., beats per minute), and HRV (Lohani, Payne, and Strayer 2019; Tavakoli, Kumar, Guo, et al. 2021; Tavakoli, Boukhechba, and Heydarian 2021). HRV features are a set of signal properties that are calculated based on the beat-to-beat intervals in a person's HR, such as the root mean squared of the successive intervals (RMSSD) (Tavakoli, Boukhechba, and Heydarian 2020; H.-G. Kim et al. 2018). Both HR and HRV metrics are used in the literature for understanding the human state. In general, studies have shown that an increase in stress level is associated with an increase in HR, and a decrease in RMSSD features (Tavakoli, Kumar, Guo, et al. 2021; H.-G. Kim et al. 2018; Napoli et al. 2018). More specifically, in bicycling studies, an association has been found between perceptions of risk and HR (Doorley et al. 2015; Fitch, Sharpnack, and S. L. Handy 2020). For instance, a naturalistic study in Ireland showed that situations bicyclists perceive to be risky are likely to elicit higher HR responses. This study also found that busy roads and roundabouts without bike lanes were perceived as more dangerous compared to roads where cyclists are separated from traffic (Doorley et al. 2015).

2.3.3 Body Position

Body position influences leg kinematics and muscle recruitment for cyclists (A. R. Chapman et al. 2008). Some professional sensors can be used to build 3D body tracking by implementing multiple on-body receivers to study pedestrians' dynamics of indoor activity (Correa et al. 2016).

Recent development in computer vision has greatly reduced the cost of obtaining body movement data. For example, the OpenPose, an open-source real-time multi-person system, can jointly detect the human body, hand, facial, and feet key points on single 2D images (Z. Cao et al. 2017).

2.3.4 Electrodermal Activity (EDA)/ Galvanic skin response (GSR)

Electrodermal Activity (EDA) /Galvanic skin response (GSR) is the measurement of changes in the electrical conductivity of the skin. Sweat level is a physiological signal that indicates sympathetic nervous system activity. With an increased stress level, sympathetic nervous system activity increases, resulting in increased sweat gland activity. Higher EDA is indicative of physiological arousal due to increased stress levels. EDA is also sensitive to respiration and mental effort(Dawson, Schell, and Filion 2017). The main metrics of EDA include skin conductance level (SCL) and skin conductance response (SCR). SCL/SCR is higher during increased workload in dual-task relative to single-task driving, Higher SCL could also be indicative of lower levels of trust in the autonomous vehicles(Barnard and P. Chapman 2018). In the pedestrian movement experiment, a relationship between SCR amplitude and interpersonal distances is observed for quantitative measurement of social repulsive force(Zhao et al. 2019). EDA signals, after saliency detection analysis, are also feasible to evaluate built environments(J. Kim, Yadav, Chaspari, et al. 2020b). EDA is sensitive to many factors; therefore, we should be cautious when using EDA to interpret the findings. The experiments should be carefully designed for different control groups.

Table 2.1: IVE bicyclist simulator literature table. **-**: not included or not specified in the paper; ✓: included in the paper. **Visual Technology**: Subject viewed a *single screen/multiple screens or CAVE/head-mounted display(HMD)* as the visual source; **Agency of Movement**: *Stationary* - subject remained motionless or interacted via controller; *Dummy* - The subject was on a stationary bike but movements were not translated into VR; *Real-time* - subject movements were translated in VR. **Sound**: whether sound feedback was used. **Haptic**: Interaction with the environment through, vibration, resistance, etc. **Kinematic**: speed, steering, and direction data. **Movement**: body or head movements.

Report Information		Level of Immersion					Data Reported					
Author & Year	**Laboratory or Affiliation**	**Simulator Environment Setting**	**Visual Technology**	**Agency of Movement**	**Sound**	**Haptic**	**Participant Number**	**Kinematic**	**Movement**	**Eye Tracking**	**Physiological**	**Stated Preference**
(Van Veen et al. 1998	Max-Planck-Institute for Biological Cybernetics	Real-world (Tübingen, Germany)	Single Screen + HMD	Real-time	-	✓	-	✓	-	-	-	-
(Kwon et al. 2001	KAIST Bicycle Simulator	Real-world (KAIST Campus, Korea)	Single Screen + HMD	Real-time	-	✓	-	✓	✓	-	-	-
(Nikolas et al. 2016	Hank Virtual Environments Lab	Simulation	CAVE	Real-time	✓	✓	63	✓	-	-	-	✓
(O'Hern, Oxley, and Stevenson 2017	Monash University Accident Research Centre	Real-world (Monash University)	HMD	Dummy	-	-	30	✓	-	-	-	✓
(Xu, Lin, and Schmidt 2017	Intelligent Human Machine Systems Lab	Simulation	HMD	Stationary	-	-	30	✓	-	-	-	✓
(Kwigizile et al. 2017	Western Michigan University	Real-world (Western Michigan University Campus)	HMD	Real-time	-	✓	36	✓	-	-	✓	✓
(O. Lee et al. 2017	Delft University of Technology	-	HMD	Real-time	-	-	-	✓	-	-	-	-
(Stroh 2017	University of Iowa	-	CAVE	Real-time	-	-	-	✓	-	-	-	-
(Keler et al. 2018	Technical University Munich	Real-world (Munich, Germany)	Single Screen + HMD	Real-time	-	-	-	✓	-	-	-	-
(Sun and Qing 2018	ZouSim, University of Missouri	Real-world (Columbia, Missouri)	CAVE + HMD	Real-time	-	✓	-	✓	-	-	-	-
(Nazemi, Eggermond, Erath, Schaffner, et al. 2018	Future Cities Laboratory	-	HMD	Dummy	-	-	-	-	-	-	-	-
(Abadi, Hurwitz, et al. 2019	Oregon State University	-	Single Screen	Real-time	✓	-	48	-	-	-	-	✓
(Shoman and Imine 2020	IFSTTAR	-	CAVE	Real-time	-	✓	10	✓	-	-	-	✓
Current study	ORCL, University of Virginia	Real-world (Charlottesville, Virginia)	HMD	Real-time	✓	✓	-	✓	✓	✓	✓	✓

2.4 Summary of IVE simulators and data collection

To summarize existing literature in this area, we have categorized past IVE bike simulator studies with their IVE settings and data collection methods. Table 2.1 has been developed to better illustrate how the trends in technology, immersion, collected data, and analysis of cyclist research have changed over the last two decades. Note that for studies from the same research group, only the latest work is included.

2.5 Cyclist Distraction

As mentioned before, we categorize cyclist distraction into three main types: Visual (taking the eyes off the road), Manual(taking the hands off the handlebar), and Cognitive (taking the mind off cycling). Recently distracted cycling due to the use of portable electronic devices has emerged which poses safety issues. Unlike distracted cycling, a large number of studies have focused on the effect of using mobile devices while driving, including reduced awareness of drivers' surroundings, increased reaction and braking times, increased incidences of collision, reduced vehicle speed, greater following variability, greater lateral variability, and reduced response time to the lead vehicle (Caird et al. 2008; Drews et al. 2009; Mwakalonge, White, and Siuhi 2014). Cyclists, much like drivers, also engage in secondary task activities while cycling such as listening to music, taking a photo or video, talking with other cyclists, writing text messages, or accessing social media (Terzano 2013; Wolfe et al. 2016; Young et al. 2020). The current state of knowledge on cyclist distraction is mostly retrieved from surveys or observational studies. For example, an observational study in New York City shows that headphone use is the most prevalent distraction among local cyclists (Ethan et al. 2016). However, observational studies are unable to track cyclists' physiological changes and get the details of secondary tasks (e.g., headphone use can be either music listening or talking on the phone). Surveys from different areas around the world have been collected to study cyclists' distracted behavior, listening to music or talking with earphones have been identified as the most prevalent distractions (Terzano 2013; Wolfe et al. 2016; Young et al. 2020). The limitation of the surveys, as discussed before, is the hypothetical bias, which undermines the realism of cyclists' behavioral responses. Cyclist distraction study in the IVE simulator can be a solution yet there are very limited existing studies.

The most frequent secondary tasks, both listening to music and talking with earphones can be categorized as the cognitive distraction. One of the main challenges in the quantitative analysis of cognitive distraction is the difficulty in measuring the workload needed for certain tasks. To understand the mechanism of distraction, a standardized secondary task with different levels of workload is required in the experimental study. To our knowledge, no prior studies have applied such methods for cyclist distraction. In other research fields, several standardized secondary tasks have been developed to simulate different levels of workload. For instance, the N-back task requires subjects to recognize items that have been presented in n-steps before. Since subjects must memorize the sequence of items in order to discover those repetitions that span multiple items, the task has strong validity for being a working memory task (Coulacoglou and Saklofske 2017). Depending on the type of stimulus, the N-back task can be either visual or cognitive, or a mixture of both. For visual distraction, more secondary tasks, like the Surrogate Reference task and Color Block Task are developed for different levels of visual secondary tasks to simulate visual distraction (Standardization 2016; Wang, Guo, et al. 2016).

Chapter 3

ORCLSim: A System Architecture for Studying Bicyclist and Pedestrian Physiological Behavior Through Immersive Virtual Environments

3.1 Introduction

Over the past couple of decades, the evaluation of roadway safety and design has been automobile-centric. Many observational, survey-based, naturalistic, and experimental studies have been conducted to evaluate the impact of roadway design features on drivers' behaviors and safety, leaving out other roadway users such as bicyclists and pedestrians. Furthermore, recent studies have highlighted the growth in the number of injuries and fatalities for vulnerable road users (National Highway Traffic Safety Administration 2020). National Highway Traffic Safety Administration (NHTSA) reported a 35% increase in pedestrian fatalities nationwide in the past ten years, and deaths of bicyclists in the United States reached all-time highs in 2018 and 2019.

These reports also indicate that although the overall number of vehicle crash fatalities is decreasing (which includes vehicle and non-vehicle related), the proportion of vulnerable road users (motorcyclists, pedestrians, pedal cyclists, and other non-occupants in the vehicle) fatalities has been increasing from 20% in 1996 to 34% in 2019 (Administration et al. 2020). These trends indicate that the design of our roadways needs to be improved to be more inclusive for all users, especially for vulnerable road users such as bicyclists and pedestrians (Rodriguez-Valencia et al. 2021).

Different factors, such as the speed limit, roadway design, and presence of large vehicles (e.g., trucks) have been shown to be associated with severe injury or fatality of bicyclists (P. Chen and Shen 2016). Additionally, the presence of intersections, traffic volumes, noise level, and physical segregation between bicyclists and vehicles have been shown to influence bicyclists' stress or comfort level (Teixeira et al. 2020; Rybarczyk et al. 2020; Cobb, Jashami, and Hurwitz 2021). Similarly, for pedestrians' safety, similar factors as bicyclists' are emphasized by researchers: pedestrian infrastructure, roadway design, traffic volumes, vehicle speed, and visibility of the road environment (Stoker et al. 2015). It is also found that bicycle paths, crossing surface material, street type, as well as the presence of nearby parked vehicles are associated with the number of pedestrian–vehicle conflicts from a naturalistic observation study (Cloutier et al. 2017).

To better inform roadway design, extensive datasets similar to the automobile-focused studies of the past are needed for bicyclists and pedestrians. To develop robust bicyclist and pedestrians-focused datasets, studies with both high ecological and internal validity are needed. Ecological validity refers to the extent an experimental environment matches with the real world, increasing the chances that the effects identified in an experimental environment generalize to real-world settings. Internal validity refers to the extent to which a cause-effect relationship is warranted in a study. Subjective, naturalistic, and experimental datasets can be utilized to tackle these issues. Subjective studies, such as surveys, provide measures of users and their perceptions of their environment but lack ecological and internal validity (Wynne, Beanland, and Salmon 2019). On the other hand, naturalistic studies can provide information about realistic changes within the environment and bicyclist and pedestrian behavior with high ecological validity, but these studies with lower internal validity are resource- and time-extensive and have potential risks of injuries and fatalities for participants. For example, a study in real traffic examining the glance behavior of teenage cyclists while listening to music was terminated when the results indicated that a substantial percentage of participants cycling with music decreased their visual performance (Stelling-Konczak et al. 2018). Furthermore, naturalistic studies are influenced by many environmental factors that restrict the ability to fully isolate and understand the impact of independent variables, thus offering low internal validity, especially for physiological and behavioral factors (Fitch, Sharpnack, and S. L. Handy 2020; Teixeira et al. 2020). Thus far, the majority of bicyclist and pedestrian studies rely on subjective and naturalistic data derived from real-world settings to assess participants' behavior and comfort in different traffic environments (Paridon et al. 2019; Shaaban, Muley, and Mohammed 2018).

Experimental studies provide an opportunity to evaluate the impact of safety-related conditions, infrastructure, and technology on bicyclists and pedestrians as they can offer the ecological validity lacking in subjective studies and allow the researchers to control for external variables, unlike naturalistic studies, for greater internal validity. Experimental studies conducted with virtual simulators can minimize the hypothetical bias of subjective surveys while offering a controlled, low-risk, and immersive environment that real-world experiments cannot guarantee. The benefit of an Immersive Virtual Environment (IVE) is achieving high internal and ecological validity while also being cost-effective, and offering complete experimental control to replicate trials (Heydarian, Carneiro, et al. 2015; Heydarian and Becerik-Gerber 2017). Early IVE lacked realism, which was primarily due to a lack of technological capability. Fortunately, IVE software and hardware platforms have significantly improved over the last few years with the release of high-end, commercially available head-mounted displays (HMD), and graphics cards capable of rendering highly detailed environments. Meanwhile, real-life components can also be integrated into IVE studies. For instance, we can now integrate a physical bicycle along with haptic feedback with an IVE to create a more immersive setting for participants (Guo, Robartes, et al. 2021). Furthermore, as the level of immersion increases, we can expand these experiments to capture participants' physiological and psychological factors, which is a field of data that has historically been overlooked. Such data provides insights into how participants' behaviors and perceptions may change in contextual settings in different research fields (J. Kim, Yadav, Ahn, et al. 2019; G. Lee et al. 2020; Adami et al. 2021; Noghabaei, K. Han, and Albert 2021; Sharif and Oppenheimer 2021). For example, pedestrians' distinct physiological responses (gait patterns, heart rate, and electrodermal activity) to negative environmental stimuli are reported from naturalistic ambulatory settings in a building (J. Kim, Yadav, Chaspari, et al. 2020a). With the increase of realism in IVE simulators and the development of low-cost ubiquitous sensors, IVE simulators have become promising tools for conducting highly realistic and immersive experimental studies (Ergan et al. 2019). In traffic safety studies, driving simulators have been widely applied to study drivers' behaviors, awareness (Soares, Ferreira, and Couto 2020), and psychophysiological states with multimodal data collection systems, such as eye trackers, electroencephalogram (EEG) and electrocardiogram (ECG) (Haufe et al. 2011; Akbar et al. 2017; Guo, Cui, et al. 2019; Baee et al. 2021). Some of the driver-related studies are conducted in IVE (Eudave and Valencia 2017). Meanwhile, for bicyclists and pedestrians, only a few studies have applied physiological responses in IVE simulators. For example, bicyclists' galvanic skin response had fewer peaks with a bike lane than with no bike lane condition (Cobb, Jashami, and Hurwitz 2021). In another cycling virtual reality study, electroencephalography (EEG) data shows its potential in a hybrid model framework as an indicator of the perceived risk of bicyclists (Bogacz, Hess, Calastri, Choudhury, Mushtaq, et al. 2021). For pedestrians, it is notable that older pedestrians spent more time focusing on the central area of the scene and even less so in the last five seconds before making the crossing decision in an IVE study (Tapiro et al. 2016).

In this study, we propose an IVE-based framework (ORCLSim) for supporting pedestrian and bicyclist research. The proposed framework integrates realistic visualizations from the real world in IVE along with a physical bicycle and a suite of passive sensing technologies, which enable the collection of physiological and behavioral responses of users. The goals of this paper are to: 1. identify research methods, trends, and gaps in knowledge related to bicyclist and pedestrian research in IVE; 2. present a novel framework for evaluating bicyclist and pedestrian behavioral changes through integrating human physiological sensing within IVE; 3. present a set of case studies to highlight how the proposed framework could be implemented to collect and analyze bicyclist and pedestrians' behavioral and physiological changes in different roadway conditions and designs.

3.2 Methodology

To address the existing knowledge gaps identified in the previous section, we introduce a new IVE-based framework - ORCLSim - where we can evaluate participants' behavioral and physiological responses in different simulated environments. This section provides details on the devices and processing techniques utilized in the proposed framework. To collect the multimodal data desired, multiple components are required to work in synchronicity within the IVE. The ORCLSim system architecture is shown in Figure 3.1, detailing all of the technology, software, communications network, and associated data flow. The details of the system framework will be discussed in this section.

3.2.1 Environment and Design Context

The IVE is developed based on a real-world location: the Water Street corridor in Charlottesville, Virginia. Water Street is well-trafficked by bicyclists, and has been identified by the Virginia Department of Transportation as a high-risk site for pedestrians, and is being considered for redesign by the city of Charlottesville as shown in Figure 3.2. The section of the corridor chosen for this experiment consists of four city blocks, with a 4% eastbound downhill in one of the road segments (road segment 1 in Figure **??**d), shared lane markings for bicycles in the east and westbound directions, a traffic signal at the intersection of East Water Street and 2nd Street SE, and a parking lane in the westbound direction. The IVE was developed on a one-to-one scale of the Water Street corridor based on technical drawings provided by the City of Charlottesville and in-field measurements (Figure 3.1-1). The textures - graphical images/skins laid atop 3D models to represent surface detail - used within

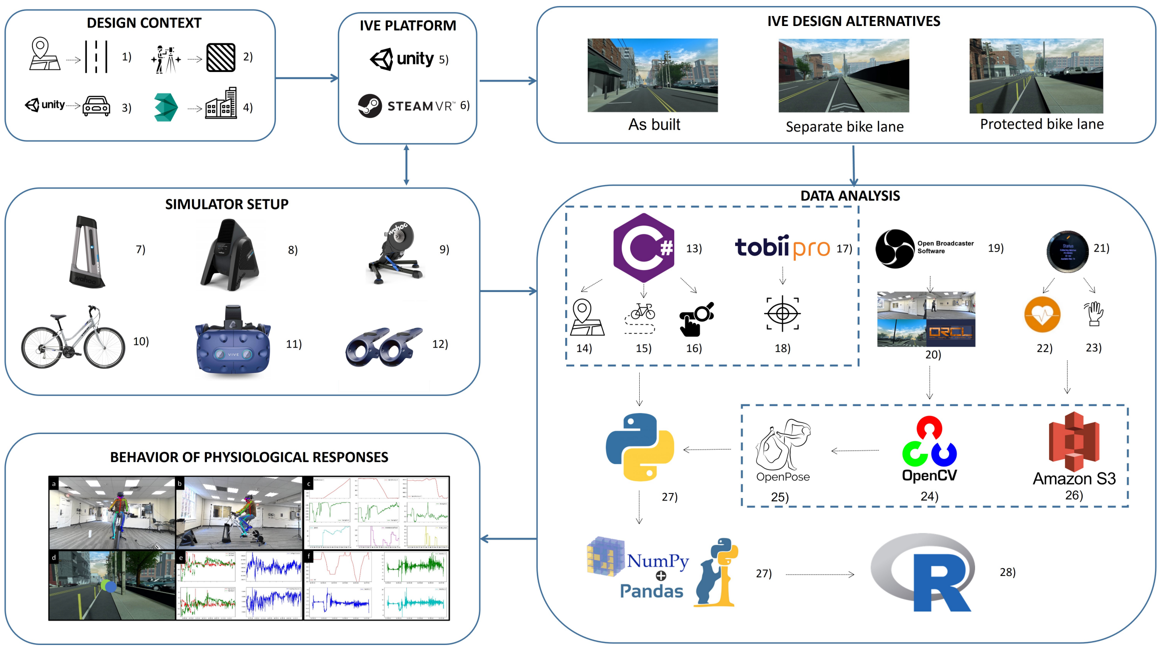


Figure 3.1: System architecture of ORCLSim framework. Design context: 1) Road geometry information from Google Maps; 2) Road texture from real-world measurement; 3) Vehicle modeling and traffic simulation in Unity; 4) Buildings modeling from 3DMax. IVE Platform: 5) Unity: 3D gaming engine; 6) SteamVR: integrating hardware with Unity. Simulator Setup: 7) Wahoo Kickr Climb: physical grade changes; 8) Wahoo Kickr headwind: headwind simulation by speed; 9) Wahoo Kickr Smart Trainer + ANT+: biking dynamics simulation; 10) Trek Verve physical bike; 11) HTC VIVE Pro Eye: VR headset with eye-tracking; 12) Controllers: steering and braking of the bike and pedestrian's interactions with the environment. Data Collection: 13) C# scripts in Unity to record: 14) Position, 15) Cycling performance and 16) Pedestrian's interactions (touch, click or press); 17) TobiiPro Unity API collects: 18) Eye tracking data; 19) OBS studio: records room videos and VR videos simultaneously as shown in 20); 21) Android smartwatch collects: 22) Heart rate and 23) hand acceleration data. Data Preprocessing: 24) Opencv: video and image processing; 25) Openpose: pose data extraction from videos; 26) Python: data cleaning, management and analysis; 27) Amazon S3: smartwatch data on the cloud

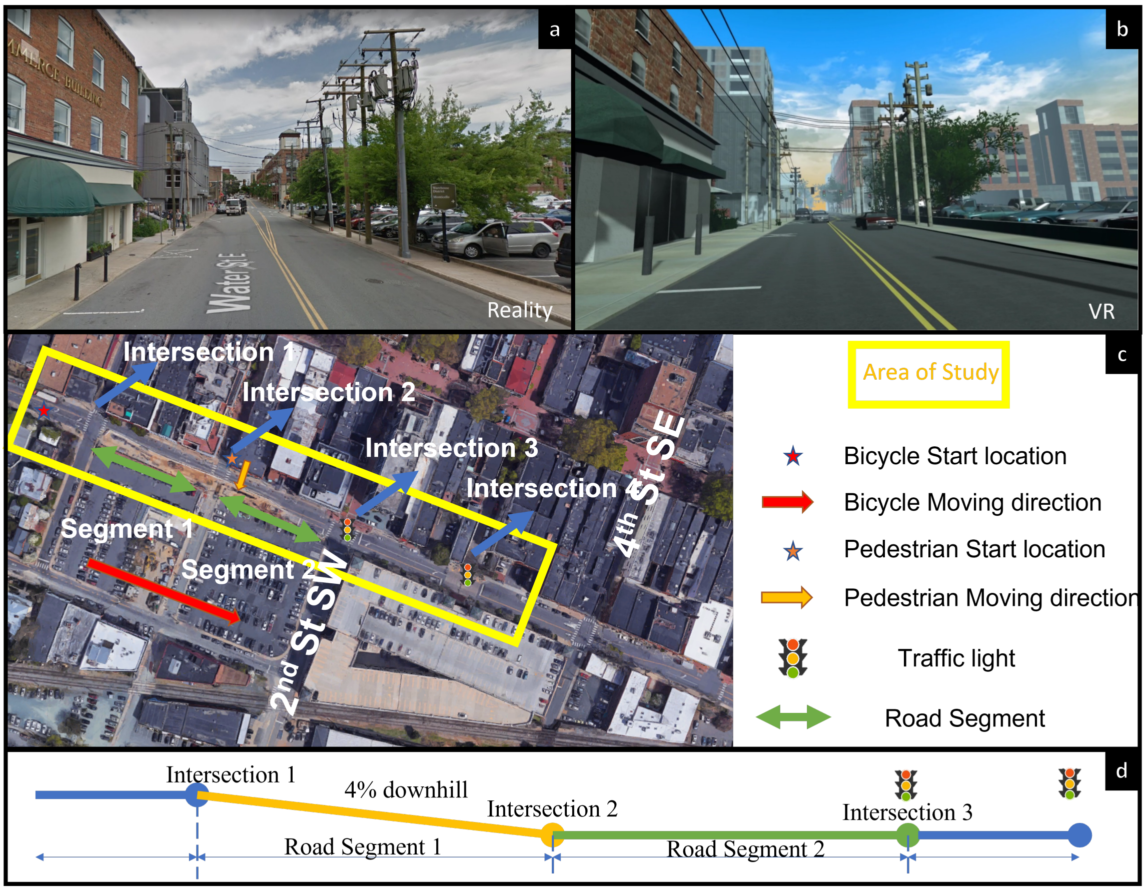


Figure 3.2: Comparison between (a) the real environment street map and (b) the IVE, (c) the map for the location of the real world environment, (d) profile of road geometry

the IVE were custom-made from high-resolution images taken on-site of the real-world surfaces appearing in the IVE so that all colors and surface details within the IVE represented the same of those in the real-world environment (Figure 3.1-2). Figure 3.2 (a) and (b) present the comparison between the real environment and the IVE created in Unity.

Real-world observations were conducted by installing four MioVision Scout cameras at the intersections of 2nd St SW, 1st St S, and 2nd St SE along Water St. The cameras captured video footage for two time periods: Tuesday 12 am to Thursday 11:59 pm on August 27-29 and September 3-5, 2019. Midweek days were chosen to avoid any abnormal fluctuations in traffic typically seen on Mondays and Fridays. Vehicle traffic within the IVE is generated based on the observed vehicle traffic during these periods. A cumulative distribution function (CDF) was developed for the observed headway gaps between vehicles in both directions of traffic. The resulting CDF was used to generate multiple, randomly weighted theoretical gap observations. The order of presentation of the gaps from the resulting theoretical distribution was randomized during each experiment to avoid any bias toward gap sizes or learning effects. Furthermore, four different car models were used within the IVE, and each car model was randomly chosen each time for vehicles generated within the IVE to further limit subject bias towards certain vehicles or learning effects (Figure 3.1-3). The buildings in the IVE are modeled individually in 3DMax and then imported into Unity (Figure 3.1-4). With all these methods aforementioned, we aim to maximize the immersion of the IVE.

3.2.2 IVE platforms

The IVE used in this framework is developed in Unity 3D game engine 2018 and runs through the SteamVR platform. High-end computing equipment is chosen so that development and testing would not be limited by computational performance. Two high-performance, factory overclocked Nvidia 1080Ti graphics cards run through Scalable Link Interface, an Intel Core i9-7920X CPU, 64 GB of DDR4 RAM at clock speeds of 3600MHz, and M.2 Solid State Hard Drives are installed in the lab computer to assure that environment rendering and stability, data collection speed, and information exchanges would not bottleneck at any component within the system while running SteamVR and Unity.

3.2.3 Simulator setup

This section will discuss the hardware components chosen for both simulators. Figure 3.3 demonstrates the appearances of both simulators. HTC Vive Pro VR headsets (Figure 3.1-11) with their accompanying controllers (Figure 3.1-12) are equipped in our simulators. The HTC Vive Pro is capable of running high resolutions (1440 x 1600 pixels per eye) and frame rates (90Hz), provides a wide field of view (110 degrees), has motion tracking and gaze tracking capabilities, and is compatible with SteamVR. Controllers can be used differently in bicyclist and pedestrian studies. For bicyclists, the spatial location of the controllers allows the system to detect turning movements. The braking action can be recognized by the squeezing value of the trigger keys on the controllers. For pedestrian studies, controllers can help the user to interact with objects in the virtual environment, as an extension of their hands.

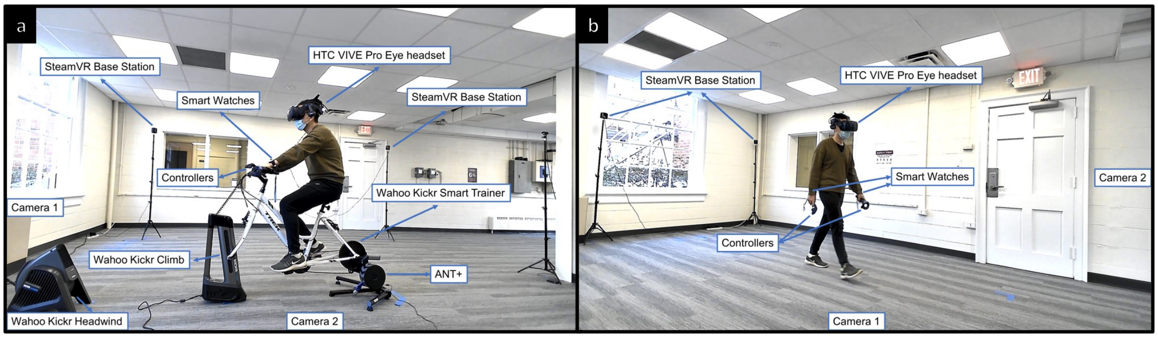


Figure 3.3: Appearance of the simulators, (a) Bicycle simulator; (b) Pedestrian simulator

Bicycle Specific Equipment

The following equipment has been specifically chosen for the bicyclist simulator:

- Wahoo Kickr Smart Trainer (Figure 3.1-9) - Power measurement system of +/- 2% for accurate, realistic resistance feedback. It has adaptive, real-time resistance based on road grade.
- Wahoo Kickr Climb (Figure 3.1-7) - Adaptive, real-time indoor bicycle grade simulator attached to the front fork of the bicycle that accurately raises or lowers the front end of the bicycle based on road grade. It is capable of simulating roadway grade in a range of -10% to +20%.
- Wahoo Kickr Headwind (Figure 3.1-8) - It provides an adaptive, real-time variable speed vortex fan capable of reaching wind speeds experienced by bicyclists on the road, providing tactile feedback based on bicyclist speed.
- ANT+ (Figure 3.1-9) - Wireless protocol used for communications between the Wahoo training equipment and desktop computer, capable of sending controller information between the two devices regarding speed, resistance, grade, and wind speed.

The Wahoo equipment collects the data necessary for research studies and provides haptic feedback to the users. Critically, through the use of ANT+ and a Unity asset, the Wahoo equipment is compatible with the computer hardware and environment software used in the simulators. An average physical Trek Verve bike (Figure 3.1-10) is adapted as the main body structure of the bicycle simulator.

Pedestrian Specific Equipment

The HTC Vive Pro Eye headsets have been equipped with the HTC Vive Pro Wireless Adapters, which support a 6 x 6 m space for accurate tracking and operating on wireless communication. This HTC Wireless Adapter is chosen to meet the goals of the simulator as it eliminates the impact that wires hanging from the back of the headsets have on the users' ability to move around without getting tangled.

3.2.4 Data Collection

In this section, we will introduce details about the collected data from different data sources, including the data type and the frequency of data collection. Specifically, we first discuss the data streams exported from the Unity software, followed by the eye-tracking data points, and information extracted from the video recordings and smartwatches, as shown in Figure 3.1.

Unity

With the attached scripts written in C# programming language to the Unity scenario (Figure 3.1-13), we can extract the world position (in meters) and direction (unit vector) of each object in the virtual environment, including headset, controllers and other virtual objects such as vehicles. This data can decipher where the users are in the environment and their relative positions to other objects. The scripts also collect any input from the controllers. For example, the pulled trigger values (0 to 1) are the brake for the bike simulator. The frequency of Unity is generally around 30Hz. Additionally, the system timestamp is attached to the final Unity output data for time synchronization.

Eye tracking

The eye tracking features of HTC VIVE Eye Pro in Unity are from the integrated Tobii Pro eye tracker. The raw data extracted from the headset can be processed to track and analyze the eye movement, attention, and focus level of each participant. Details of the utilized eye tracking system, sample environment, and the code to extract the different data streams have been shared online (Guo 2021). The eye-tracking data is collected through Tobii Pro Unity SDK (Figure 3.1-17). It is integrated into Unity with C# scripts, and the data collection starts simultaneously with the running of the Unity scenes. The system timestamp is attached to the final output data. The output of Tobii Pro raw data is the 3D gaze direction, gaze origin, and pupil diameter. Pre-processing techniques are required to relate the eye tracker's coordinate system to the headset's position in a virtual 3D world (such as Unity). The frequency of eye tracking data is 120Hz.

Video Recording

The video recording system has three components: two video recordings from cameras capturing the body position of the participant and one screen recording of the participant's point of view in IVE. These videos are recorded simultaneously in OBS studio (Figure 3.1-19) with the same frequency (30Hz), resolution (1080p, 1920×1080), and system timestamp.

Smartwatch

Experiment participants wear two android smartwatches ((Figure 3.1-21, one for each wrist) that are equipped with the "SWEAR" app for collecting longitudinal data. The SWEAR app records heart rate (1 Hz), hand acceleration (10 Hz), audio amplitude (noise level, 1/60 Hz), and gyroscope (10 Hz) (Boukhechba and Barnes 2020). Both watches are connected to a smartphone via Bluetooth, and the smartphone and computer are on the same wifi network to make sure time is synchronized with the server before each experiment. All data from the smartwatches are stored on the local device and then uploaded to the Amazon S3 cloud (Figure 3.1-27) for future data extraction and analysis.

3.2.5 Data Pre-processing

All the data collection devices and platforms (except for the smartwatches) are connected to the local computer, allowing them to be synchronized with the computer's system time. Figure 3.4 shows the visualization of all the data collected in the simulator after time synchronization.

Information about each video source (frames per second, creation date, duration, height, and width) can be extracted from the singular video and be split into separate videos for each source (cameras 1 and 2) through the Opencv software (Figure 3.1-24). Two external video cameras in the lab capture each participant's body movement during the experiment. These recordings can be used to understand participants' movements and reactions. Furthermore, the body position data can be extracted from these videos using the OpenPose software (Figure 3.1-25). Figure 3.4(a) and (b) show the body position detection of the video recordings from OpenPose. Used in conjunction with the motion tracking of the VR headset and controllers, this video footage can help determine how participants react to their environments during experimentation for better behavioral analysis and event detection.

Combining the raw gaze direction from the eye tracking data with the video information of the point-of-view videos, it is possible to transform the 3D gaze direction into 2D videos to visualize what the participants are looking at in the IVE. As shown in Figure 3.4(d), the green and blue dots represent the direction left and right eyes are looking, respectively.

Data from the smartwatches are stored locally in the device during the experiment and then uploaded to the Amazon S3 cloud storage. After the experiment, the data can be downloaded for further analysis.

All the text data are transformed to .csv format for easier integration. All the physiological data are labeled corresponding to different road segments based on the time and position data from the unity output, where the road geometry information comes from.

3.2.6 Data Analysis

In this section, we discuss the change point detection algorithm applied to the HR as well as the gaze entropy, which is the basis of the event detection in our study. First, we discuss how the Bayesian change point (BCP) detection is applied to the

HR data. Similarly, for the gaze data, we discuss how gaze entropy can be calculated and used to identify the dispersion of gaze.

Bayesian Change Point Detection

Bayesian Change Point (BCP) detection methods are applied to detect abrupt changes in HR data. Change point analysis deals with time series data where certain characteristics undergo occasional changes. It is assumed that there is an underlying sequence of parameters partitioned into contiguous blocks of equal parameter values: the beginning of each block is a change point. Observations are then assumed to be independent in different blocks given the sequence of parameters (Barry and Hartigan 1993). A Bayesian approach to the change point problem can give uncertainty estimates not only for location but also for the number of change points.

Suppose we have a time series of HR data X, and we use $\rho = (U_1, ..., U_n)$ to indicate a partition of the time series into non-overlapping HR regimes where $U_i = 1$ means a change point happens at position i+1. To calculate the posterior distribution over partitions, we use the Markov Chain Monte Carlo (MCMC) method. We define a Markov Chain with the following transition rule: with probability p_i, a new change point at the location i is introduced. In each step of the Markov Chain, at each position i, a value of U_i is drawn from the conditional distribution of U_i given the data X and the current partition ρ. let b denote the number of blocks obtained if U_i = 0, conditional on U_j , for $i \neq j$. The transition probability p, for the conditional probability of a change point at the position $i + 1$, can be obtained from equation (3.1) (Barry and Hartigan 1993; Erdman and Emerson 2007):

$$\begin{aligned}\frac{p_i}{1-p_i} &= \frac{p(U_i = 1|X, U_j, j \neq i)}{p(U_i = 0|X, U_j, j \neq i)} \\ &= \frac{\int_0^\gamma p^b(1-p)^{n-b-1}dp \int_0^\lambda \frac{w^{b/2}}{(W_1+B_1 w)^{(n-1)/2}}dw}{\int_0^\gamma p^{b-1}(1-p)^{n-b}dp \int_0^\lambda \frac{w^{(b-1)/2}}{(W_0+B_0 w)^{(n-1)/2}}dw}\end{aligned} \tag{3.1}$$

Here B_0, W_0, and B_1, W_1 are the within and between block sums of squares obtained when $U_i = 0$ (with change point at location i) and $U_i = 1$ (without change point at location i), respectively. The two tuning parameters γ and λ can be calculated with MCMC. We use *bcp* package in R (Erdman and Emerson 2007) to implement the change point analysis. A similar approach has been utilized in a previous study to identify changes in driver's HR data in different roadway conditions (Tavakoli, Kumar, Guo, et al. 2021). The BCP output is time series data of the probability of change points.

Gaze entropy

Gaze entropy is a comprehensive measurement of visual scanning efficiency. The concept of entropy originates from information theory (Shannon 1948). It is only in recent years that entropy has gained growing attention from researchers attempting to quantitatively examine gaze behavior in naturalistic settings. By Shannon's equation of *entropy* and *conditional entropy* (Shannon 1948), there are two types of gaze entropy measures: *stationary gaze entropy* (SGE) and *gaze transition entropy* (GTE) (B. Shiferaw, Downey, and Crewther 2019). SGE measures overall predictability for fixation locations, which indicates the level of gaze dispersion during a given viewing period (Holmqvist et al. 2011). The SGE is calculated using Shannon's equation:

$$H_s(x) = -\sum_{i=1}^{n}(p_i)log_2(p_i) \tag{3.2}$$

Here $H_s(x)$ is the value of SGE for a sequence of data x with length n, i is the index for each state, p_i is the proportion of each state within x, it is assumed that fixation is an individual output of the gaze control system that makes spatial predictions regarding the location of subsequent fixations (B. Shiferaw, Downey, and Crewther 2019).

Gaze transition entropy (GTE) is conducted by applying the conditional entropy equation to 1st-order Markov transitions of fixations with the following equation:

$$H_c(x) = -\sum_{i=1}^{n}(p_i)\sum_{i=1}^{n}p(i,j)log_2p(i,j) \tag{3.3}$$

Here $H_c(x)$ is the value of GTE, p_i is the stationary distribution, same as equation (3.2), and $p(i,j)$ is the probability of transitioning from i to j. GTE provides an overall estimation for the level of complexity or randomness in the pattern of visual scanning relative to the overall spatial dispersion of gaze, where higher entropy suggests less predictability.

Specifically, to calculate the SGE and GTE, the visual field is divided into spatial bins of discrete state spaces to generate probability distributions. In this study, the fixation coordinates were divided into spatial bins of 100 × 100 pixels, followed by a previous study (B. A. Shiferaw et al. 2018). To get the trend of gaze entropy, it is calculated in a rolling window of five seconds (600 data points in raw gaze data streams). We also apply the BCP methods to calculate the change points in the two

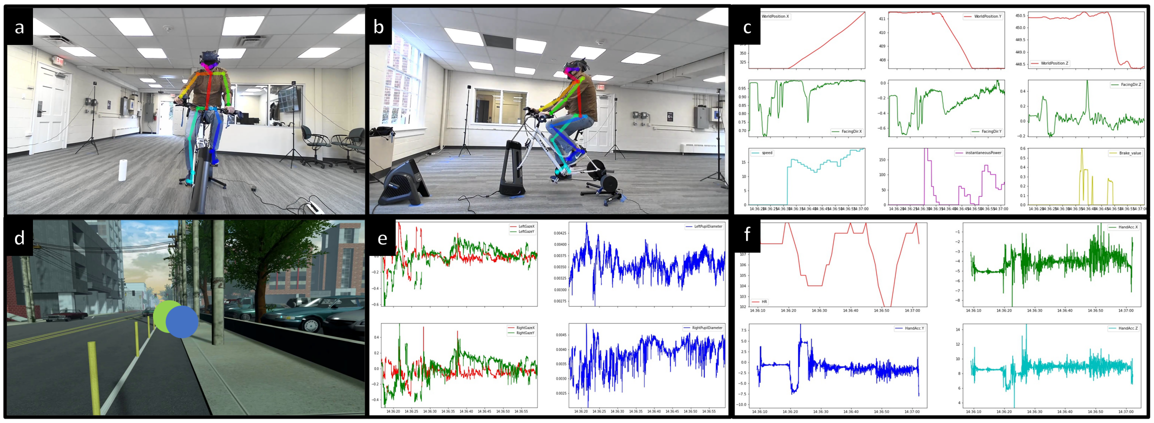


Figure 3.4: Example of data visualization output. (a,b) Video data from two cameras with body position detection; (c) position and controllers input data in VR ; (d) Field of view VR video recording with gaze mapping, green/blue dots indicate left/right eye fixation; (e) Eye tracking data, includes gaze direction and pupil diameter; (f) Heart rate and hand acceleration data from smart watch

gaze entropy values.

3.3 Case Studies

In this section, we present two case studies (one for bicyclists and one for pedestrians) from a pilot study of five participants to evaluate the proposed framework and highlight the importance of collecting physiological data, speed, and position data from participants. Using HR data, we demonstrate the relationship between the number of changes in HR data and the corresponding time frames with the changes in the environment (e.g., bicyclist arriving to an intersection) and/or the participant behaviors (e.g., when the pedestrian is ready to cross the street). The tasks in the IVE are different for each user type (bicyclists and pedestrians) in the case study. The bicyclists were asked to cycle eastbound along the corridor, as indicated by Figure 3.2. The pedestrian task was to cross the street using the crosswalk at intersection 2 whenever they felt it was safe to do so. More details about the bicyclists experiment can be found in our previous study (Guo, Robartes, et al. 2021). Furthermore, we show how the multimodal dataset can be utilized to detect pedestrians and bicyclists' state changes. The sample dataset is publicly accessible online (Guo, Robartes, et al. 2021). We first identify where abrupt changes happen in the HR readings and then identify the potential reasons behind the events that take place in a given time frame. To achieve this, the videos are manually annotated to identify an event or behavioral changes among participants. Then the timestamps of these event/behavior changes, as well as other physiological responses are compared to the time that we observe HR change points for each participant. Through this, we can show whether the effect of HR changes is consistent across different groups of participants.

The other physiological variables selected in the case study are: head movement direction, the position of the bicycle and pedestrian from Unity, gaze direction from the eye tracker, and the gaze entropy and its BCP probability from the gaze direction.

3.3.1 Bicycle Pilot Study

In this experiment, after familiarization with the simulator and calibration for eye tracking and steering in a training scenario, the participants were asked to cycle eastbound in the simulated environment as indicated in Figure 3.2.

Figure 3.5 shows one participant's physiological responses from the pilot bike experiment. Using BCP, we are able to detect the moments when the underlying distribution of HR data changes in a short period of time. Figure 3.5b shows the overall time series of different physiological data. Figure 3.5b.I. shows the HR (blue) and the probability of detected change point events (red) during the whole experiment. In

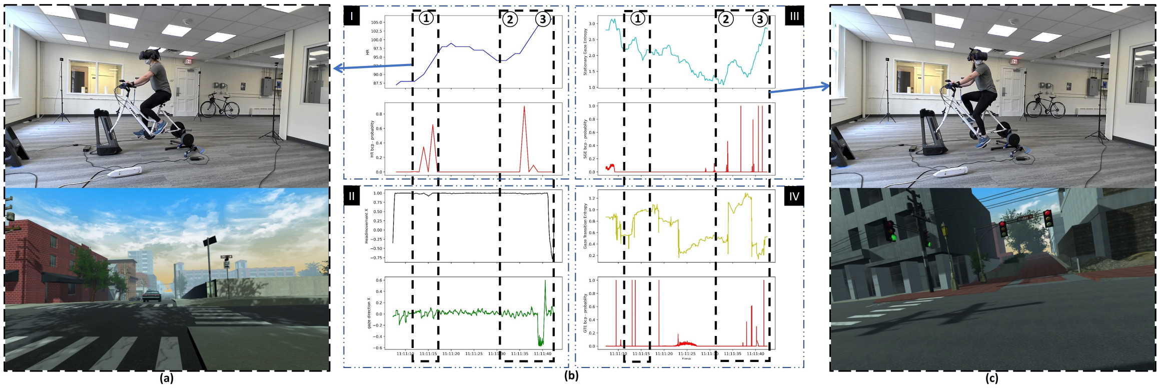
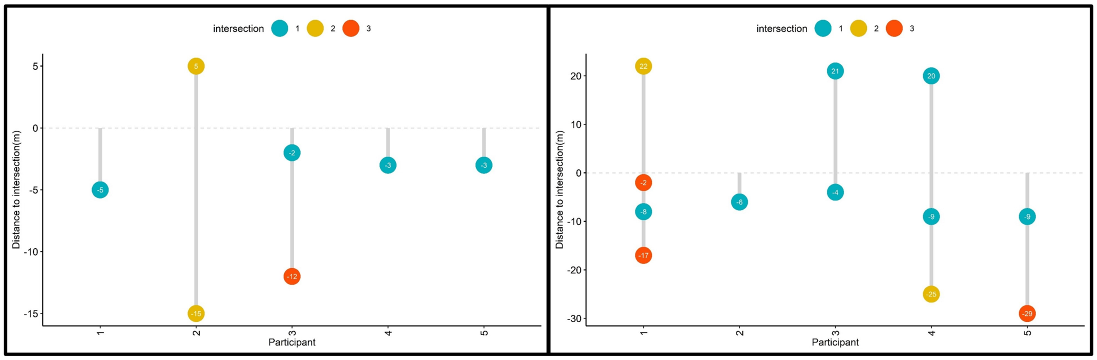


Figure 3.5: Bicyclist's HR change point analysis over time with other physiological responses. (a). Video snapshot of HR change point event 1. The bicyclist is approaching the intersection; (b). Visualizations of different physiological data: I. HR (blue) and HR BCP probability (red); II. horizontal head movement direction (black) and horizontal gaze direction (green); III. stationary gaze entropy (SGE) (cyan) and SGE BCP probability (red); IV. gaze transition entropy (GTE) (yellow) and GTE BCP probability (red); (c).Video snapshot of HR change point event 2. The bicyclist looked left behind to check if cars are approaching from behind.

addition to the HR data, Figure 3.5b.II. shows the head movement x (black) and gaze direction x (green), and the head movement in the x-axis indicates the head facing the direction from straight backward (-1) to straight forward (1). The gaze direction x indicates the gaze direction from left (-1) to the right (1). Figure 3.5b.III. shows the stationary gaze entropy (cyan) and BCP probability of SGE (red), Figure 3.5b.IV. shows the gaze transition entropy (yellow) and BCP probability of SGE(red).

Figure 3.5a and c show the corresponding screenshots for the two HR change points detected in Figure 3.5b-I. The first change point happens when the participant is approaching the first intersection on the road which does not have any traffic signals; at this time, the participant is also being passed by a vehicle on the left (Figure 3.5a). Meanwhile, the other physiological signals do not show abrupt changes except for minor peaks in GTE as shown in Figure 3.5b-IV. The second HR change point takes place when the participant is approaching the third intersection where there is a traffic signal. While crossing the intersection, a looking-around behavior is also observed as shown in (Figure 3.5c). As a result, we observe changes in both horizontal head and gaze direction (Figure 3.5b-II), a larger variance in SGE, and the change points detected from the SGE data points (Figure 3.5b-III). Similarly, we observe higher variance and more change points in GTE (Figure 3.5b-IV). Previous research suggests an increase in SGE associated with a higher GTE may reflect the influence of top-down interference on visual scanning, which results in a greater dispersion of gaze (B. Shiferaw, Downey, and Crewther 2019). In other words, increased SGE together with GTE indicates a higher visual or cognitive load in the experiment scenario for this participant. This case study indicates that the HR and gaze changes are sensitive to environmental changes as well as participant/bicyclist behaviors. It is also important to note that specific contextual factors (e.g., an intersection with or without a traffic signal) can trigger different physiological responses; therefore, it is important to collect and monitor different physiological data when conducting naturalistic or experimental studies of bicyclists.

To find the reason behind each event, all five of the participants' video recordings in the case study are manually analyzed. Figure 3.6 illustrates when the HR and gaze (use GTE as an example) change points happen for each participant. For HR, almost all the change points take place when participants are approaching an intersection within a distance of 15 meters, except for Participant 2. When Participant 2 was passed by a vehicle in intersection 2 with a very close lateral proximity, the HR went up immediately (no other participant in the pilot study had a car pass by them as closely). For gaze transition entropy, the change point generally happens earlier than the HR change point but follows a similar trend as the HR. Although the sample size is small, some of our observations from the case study include 1) among the



(a) (b)

Figure 3.6: Summary of where the HR (a) and gaze transition entropy (b) change point happen for bicycle experiment. Each dot indicates an HR/GTE change point. The x-axis indicates the participant number, the y-axis value is the distance to the intersection, and where a negative value indicates the distance the change point happens before the participant arrives at the intersection. For example, participant 2, has two HR change points at 15 meters prior to intersection 2 as well as 5 meters after passing through that intersection (a). And participant 2 has one GTE change point 6 meters prior to intersection 1.

five participants in the pilot study, there are more HR/GTE change points prior to reaching the first intersection. As it is the first intersection in the experiment, participants may feel more stress than when approaching other intersections, as they became familiar with the environment. This implies that in the early portions of VR experiments, participants still need some time to get adjusted to the IVE environment, even after a training scenario before the actual experiment. 2) The change points prior to intersections 2 and 3 take place farther from the intersections compared to the change points detection prior to intersection 1. This could be explained by two possible factors: first, the road segment after intersection 1 has a 4% downhill slope as shown in Figure 3.2d, where the participants' visibility of the road is limited until they get close or pass the intersection. Second, the roadway environments for intersections 2 and 3 are more complex. Intersection 2 is at the end of the downhill road segment and there is a lane shift after intersection 2, thus braking and right steering are needed

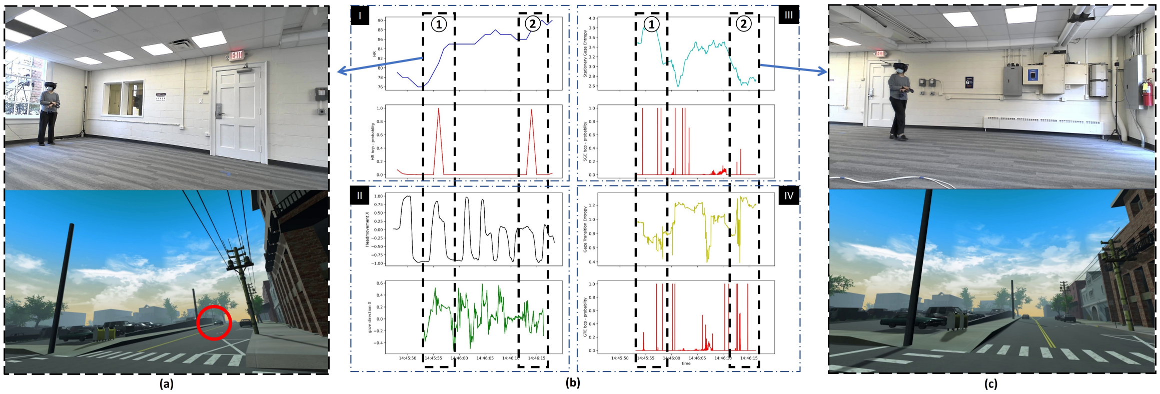


Figure 3.7: Pedestrian's HR change point analysis over time with other physiological responses. (a). Video snapshot of HR change point event 1. The pedestrian noticed the first approaching vehicle; (b). Visualizations of different physiological data: I. HR (blue) and HR BCP probability (red); II. horizontal head movement direction (black) and horizontal gaze direction (green); III. stationary gaze entropy (SGE) (cyan) and SGE BCP probability (red); IV. gaze transition entropy (GTE) (yellow) and GTE BCP probability (red); (c). Video snapshot of HR change point event 2. The pedestrian is crossing in the eastbound lane, just after taking a look at the approaching vehicle in the lane.

before they enter intersection 2. Intersection 3, as indicated before, includes a traffic signal. Although participants are told the signal will always be green during the experiment, their physiological (HR and gaze entropy) data still showed a distinct response at this intersection.

3.3.2 Pedestrian Pilot Study

The pedestrian pilot study is conducted at intersection 2 in the same IVEs 3.2 with the pedestrian simulator, where participants can walk freely as they would in real life to cross a crosswalk. As explained before, the eastbound lane has randomly-generated vehicles with different gaps. At the beginning of the pilot study, participants are asked to wait until the first vehicle passes before they can cross using the crosswalk. Once the first car passes, whenever they feel safe, they may cross the road.

Similar to the bicycle case study, we extract the physiological data with the HR change point analysis results for one of the participants as shown in Figure 3.7. The definition of the data is the same as the bicycle case study. The first change point happened when the pedestrian noticed the first approaching vehicle, as indicated by the red circle in Figure 3.7a. A larger variance in SGE (Figure 3.7b.III) and GTE (Figure 3.7b.IV) is observed at the same time. An increase in SGE associated with lower GTE is likely indicative of distraction (such as the first approaching vehicle in this case). The second change point happened during the crossing in the eastbound lane, just after the participant looks at the approaching vehicle in the lane (Figure 3.7c). During the change point event, only a larger variance in GTE (Figure 3.7b.IV) is observed, while SGE remains at a low level (Figure 3.7b.III). A reduction in SGE when GTE is increasing reflects top-down interference whereby the viewer focuses on specific items within the visual scene. In this case, the participant is looking straight to the other side of the road after the last look at the approaching vehicle in the lane, trying to cross the crosswalk quickly. In addition, after the pedestrian starts crossing, the range of horizontal head movement is smaller than before crossing (Figure 3.7b.II). This indicates that once they make the decision to cross, they will not observe the surroundings (e.g., incoming vehicles) as much as they do before crossing.

Table 3.1 shows the video annotation details for the pedestrian experiment. A total number of 7 HR change points are identified across the participants. There are three main categories within HR change points: two HR change points are detected when participants noticed the first approaching vehicle, two HR change points are identified when participants cross the crosswalk right after the first vehicle passes, and three HR change points are detected when participants are crossing in the vehicle-approaching (eastbound) lane. Similar to the bicycle pilot study, these change points correlate to the changes in the contextual setting, such as a vehicle approaching a crosswalk. These findings indicate why it is important to collect participants' physiological responses when conducting pedestrian studies. Although our preliminary findings show there exists a correlation between HR and gaze change points to the time that certain events take place in the environment, analysis of a larger group of participants is needed to verify the findings.

Table 3.1: Summary of pedestrian HR change point

Description of HR change point category	Included Participants
Noticed the first approaching vehicle in the initial position	1,5
Start crossing after the first vehicle passed from the initial position	3,4
Crossing in eastbound lane after looking at the approaching vehicle	1,2,5

3.4 Discussion and Conclusion

In this paper, we have developed a system architecture (ORCLSim) for VR simulators to capture physiological and behavioral changes in bicyclist and pedestrian studies. Specifically, the aim of this study is to determine (1) what metrics and set of information are needed to monitor bicyclists and pedestrians' behavioral changes, (2) what devices are available, and how different hardware and software packages can be integrated in IVEs to conduct similar studies, (3) how the multimodal data can be processed for observing the changes in physiological responses given different contextual settings, and (4) showcase how the proposed framework can be implemented by presenting two case studies for bicyclists and pedestrians.

Previous studies on bicyclists' and pedestrians' responses to changes in contextual settings highlight the advantages of controlled experimentation, especially in IVEs. In this paper, we demonstrated that it is important to track physiological metrics to better understand how different settings may impact bicyclists and pedestrians. Additionally, we showcased the importance of gaze tracking and heart rate data in capturing the behavioral response to different events or roadway settings. These measurements may indicate the impact of the stress levels as well as the cognitive load on the participants. In the case study, our initial findings from the five participants indicate that physiological data is sensitive to road environment changes or real-time events, especially for the change in heart rate and gaze behavior. In the presented framework, we use Bayesian Change Point (BCP) detection method to detect abrupt changes in physiological data. First, we use the HR change point to identify any potential events, then the video annotation results can help to get a better understanding of the causes behind each event. The findings are further verified by two measurements of eye tracking data: stationary gaze entropy (SGE) and gaze transition entropy (GTE). The dynamical changes in the eye tracking data also support the observations from the video annotation. For the presented bicycle case study, most change points happen prior to the intersections, while the eye tracking change points usually happened earlier than the HR change points. The increased SGE and GTE along with abrupt changes in HR indicate where the participants feel stressed in the environment, which is observed to be at the beginning of the experiment and when participants reach the intersection with a traffic signal. The physiological changes in the pedestrian case study are indicative of critical behavior during the crossing, such as observing the first approaching vehicle or the moment before crossing. Furthermore, the differences between individual participants' physiological responses also emphasize the importance of building personalized models for different groups of people. Although these preliminary findings are promising, we need to further examine whether these change points are observed when the number of participants is increased for both case studies.

We have open-sourced the system set-up document, code example, and sample dataset for the research community. The integration between the presented devices and software platforms along with the data processing method provides the foundation to support IVE experimental studies where we can identify the impact of different roadway designs on bicyclist and pedestrian behavioral and psychological changes. This presented system architecture can be used to study bicyclists' and pedestrians' behaviors that may be affected by their perception ability and cognitive state, which may also be influenced by different road design conditions. Furthermore, it makes the development of a VR simulator simpler and more robust since many of the modules are flexible and scalable to different systems and improvements. For example, the smartwatch system can be replaced by more recent and advanced wearable devices that can collect different data streams; or the video recording systems can integrate more event or activity detection through computer vision based techniques.

3.4.1 Limitations and Future Work

While useful in addressing many of the gaps in virtual simulation research, IVE has some limitations. Many of the reports included in Table 2.1 indicate that portions of their subject pool's data had to be disregarded due to the motion sickness participants experienced while in VR.

Furthermore, the removal of risk within the IVE may also be perceptible to participants – IVE experimentation relies heavily on subject immersion and while the environment may look, feel, and behave real, the knowledge that one is still in a risk-free virtual space where physical injury is not possible still exists. It is up to the realism of the IVE to suspend a subject's disbelief in the environment sufficiently to overcome this knowledge, which varies from person to person. Additionally, a subject's familiarity with VR technology could have an impact on their behavior and how they perceive the IVE - someone who has played multiple games within VR may be aware that if they were to collide with an object, they would not physically be affected.

With our system architecture ORCLSim, future IVE researchers can apply any physiological data collection modules to their IVE simulators to study the vulnerable road users' behaviors, perception abilities, and cognitive states in different contextual settings. Additionally, more physiological responses may be included in the system with off-the-shelf sensors, such as Electrodermal activity (EDA) and Electromyography (EMG). Future IVE can be improved to increase immersion and tackle more complex research problems. A feature that would greatly improve the ease of development of an IVE would be the integration of development platforms, such as Unity or Unreal Engine, and commercially available transportation simulation software, such as Synchro and Vissim. Through such integration, roadway segments, objects, vehicles, vehicle behaviors, and traffic networks could be simulated more realistically and robustly within IVE. Furthermore, more robust platforms for integrating multiple users into IVE would vastly improve immersion and realism. Instead of modeling vehicles to interact with bicyclists and pedestrians, having another subject controlling a vehicle via a driving simulator and interacting with a bicyclist and a pedestrian would provide valuable insights into the interactions between road users, unlike anything that has been done before.

Chapter 4

Benchmarking the Use of Immersive Virtual Bike Simulators for Understanding Cyclist Behaviors

4.1 Introduction

As a result of increasing numbers of single-track roadway travelers (e.g., bicycles, scooters, etc.), roadway design complexities, and varying traffic densities, bicyclist fatalities are rising, with deaths of pedal cyclists in the United States reaching 846 in 2019, close to the highest number of deaths in recorded history (National Highway Traffic Safety Administration 2020). Many contextual factors such as roadway design, physiological states, ambient lighting/noise level, traffic density, and cycling workload (uphill/downhill physical effort) can significantly impact cyclists' safety (Coyle et al. 1991; Zahabi et al. 2011).

However, there exists a lack of robust data sources on the safety and comfort of vulnerable road users, especially cyclists (Zeile et al. 2016). In traffic safety and specifically for driver-related studies, driving simulators have been widely adopted with a range of benefits compared to on-road studies by creating a safe and controllable environment to simulate different traffic scenarios cost-effectively. Simulator-based methods have been applied to study drivers' behaviors, awareness, and psychophysiological states. Driving simulators have been validated with a variety of data (Wynne, Beanland, and Salmon 2019). One approach to experimental studies relies on the construction of a testbed in which participants can interact with their environment, imitating a naturalistic environment. Virtual Reality (VR) simulation is a promising approach for infrastructure evaluation that avoids the high cost of test bed construction. The benefit of Immersive Virtual Environment (IVE) experiments is to achieve high internal and ecological validity while also being cost-effective and offering complete experimental control to replicate trials (Heydarian and Becerik-Gerber 2017).

Although IVE has a range of benefits, such tools have been mainly utilized for design improvements in indoor settings and have not been validated for transportation simulation, especially cyclists-safety studies. In the past few years, there has been an increasing number of studies on the application of IVE for transportation simulation. For the results of IVE studies to be meaningful, the cyclists' behavior between the real world and IVE must be consistent to a certain degree (O'Hern, Oxley, and Stevenson 2017). Higher fidelity of the IVE simulator usually can provide a more realistic experience, but this does not necessarily translate to a greater ability to replicate the specific task or behavior of the users (Wynne, Beanland, and Salmon 2019). Therefore, IVE simulators need to be validated against a set of key performance measures to assess the correlation between results (Wang, Mehler, et al. 2010). Traditionally IVE simulator validation studies have relied on cycling performance such as speed, lane keeping, and deviation in the lateral position, while the effectiveness of physiological sensing has not been validated. For example, cycling behavior and risk perception using behavior was significantly different between the non-immersive and immersive scenarios (Bogacz, Hess, Calastri, Choudhury, Erath, et al. 2020). Another study assessed the speed, lane position, and speed reduction on approach to intersections of cyclists both in IVE and on-road. Apart from subjective responses and direct cycling performance (e.g., speed, lane position), objective approaches such as gaze variability, head/body movement, and heart rate variability can be used to assess the safety of cyclists in different contextual settings. Furthermore, very few studies have collected multimodal physiological data of cyclists in naturalistic experiments (Zeile et al. 2016).

The goal of this study is to benchmark and validate the use of an immersive virtual bike simulator to evaluate cyclist behaviors in a naturalistic on-road experiment with its representative IVE. Through experimental studies, we can identify if there exist any significant differences in cyclist performance metrics between the two types of environments. To achieve this, through integrating the latest low-cost human sensing devices, we have built a multimodal human sensing data collection system to track cyclists' gaze, heart rate, pose, and head movement to better contextualize cyclist performance and behaviors. With the preliminary results from a pilot study, we can get insights about which measurement is consistent between IVE and real road environments.

4.2 Methodology

4.2.1 Experimental design

To achieve the identified research objectives, we conducted a pilot study in which we benchmarked participants' cycling behaviors and performance in a real-life environment and the corresponding IVE setting. The benchmark study has a within-subjects design to control for variance between subjects. The chosen corridor for this study was Water Street between 2nd St SW and 4th St SE in the city of Charlottesville, Virginia (Figure 3.2). Water Street is well-trafficked by bicyclists, and has been identified as high risk for vulnerable road users, and is being considered for a redesign. The section of the corridor chosen for this experiment consists of two city blocks, with a 4% downhill grade in the eastbound direction, shared lane markings in both directions, a traffic signal at the intersection of East Water Street and 2nd Street SE, and a parking lane in the westbound direction.

To collect data about the existing operating characteristics of the chosen corridor, video footage was collected along the corridor for two weeks. With permission from the City of Charlottesville, four cameras were set up along the selected Water Street corridor and video footage was recorded from August 27th to August 29th and September 3rd to September 5th in 2019, resulting in 144 hours of video recording. Peak traffic hours (7:00 - 9:00 AM and 4:00 - 6:00 PM) of the video footage were reviewed, and corresponding traffic volumes were recorded. These peak traffic volumes were used to determine the traffic flow in the design of the corresponding IVE settings. Figure 4.1 A-1 shows the google maps view of water street and Figure 4.1 A-2 is the corresponding IVE setting.

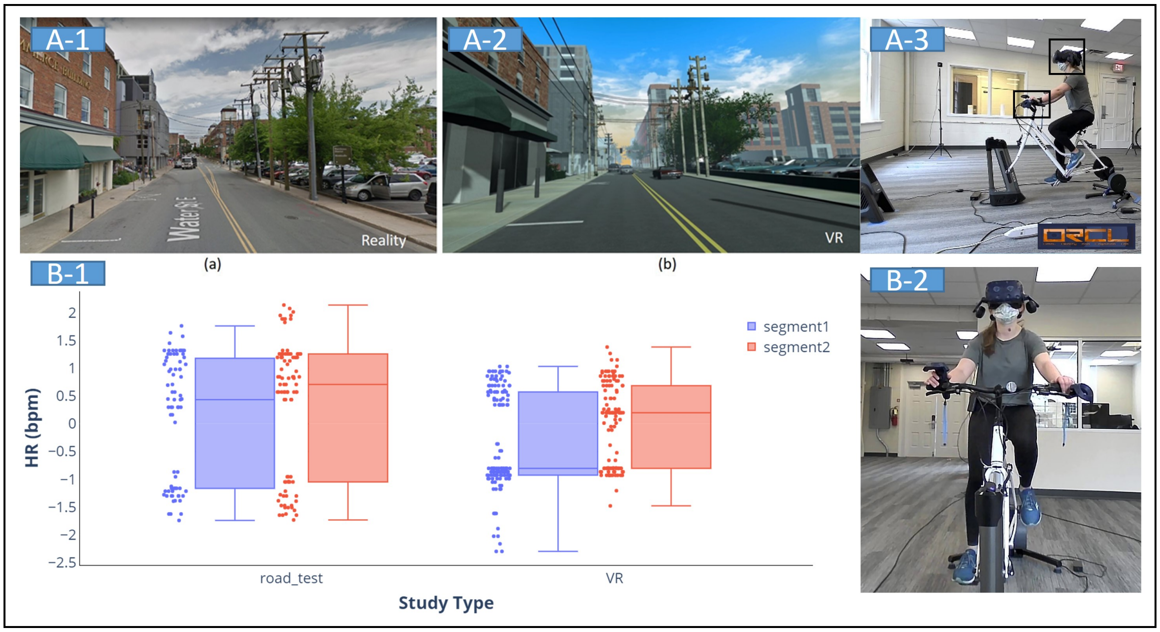


Figure 4.1: Comparison of real road(A-1) and IVE (A-2) environments; Video collection system of IVE bike simulator(A3-A4); Heart Rate Distribution of experiment(B-1)

4.2.2 Data collection

The data collection framework and system architecture of measuring cyclists' behaviors and physiological sensing is shown in Figure 4.2. For the IVE bike simulator, HTC Vive Pro Eye headsets with their accompanying controllers have been selected for tracking interaction in the IVE. This VR headset has an integrated Tobii Pro eye tracker with movement tracing capabilities, which works seamlessly with the headset as compared to traditional eye tracking systems (either screen-based or eye-tracking glasses). The spatial location of the controllers (attached to the handlebars) allows the system to detect turning or braking.

The VR environments were designed and programmed in Unity, and run through the SteamVR platform. With Unity C script, all the movement of the headset and controllers were tracked and extracted. An instrumented Trek bicycle was implemented

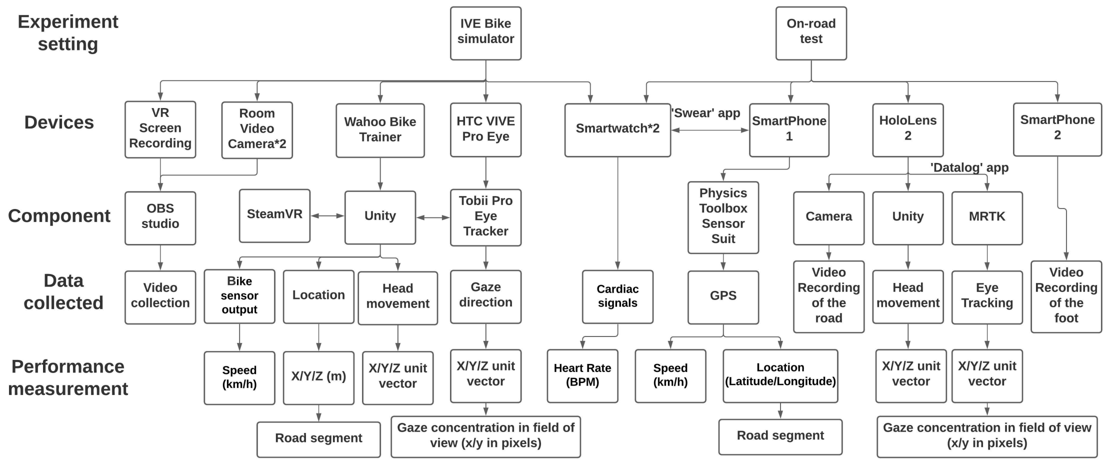


Figure 4.2: System Architecture and Data Framework

for both experiments. The Wahoo indoor bicycling training equipment was chosen to connect the physical bicycle to Unity, allowing us to collect the data necessary for this research study (real-time speed, instantaneous power, distance traveled) as well as provide haptic feedback to the participant. Additionally, two external video recording devices captured participants' movements during the IVE experiments. These recordings were used to monitor participants and understand their movements and reactions. What participants see in the IVE was also recorded with OBS studio software, which can integrate all videos with a uniform timestamp and frames per second. Furthermore, two android smartwatches equipped with the "SWEAR" app were utilized to track participants' physiological signals such as arm movement and heart rate (Boukhechba and Barnes 2020).

For the on-road test, there were two different sensors from the ones used in the IVE experiments: (1) instead of the HTC Vive headsets a Hololens 2 headset was used, and (2) an additional Android smartphone app - "Physics Toolbox Sensor Suite" was used to track participants' cycling performance. The Hololens 2 is a pair of mixed-reality smart glasses with eye-tracking features developed and manufactured by Microsoft. The authors developed a Hololens app based on MRTK SDK called 'datalog' to collect eye tracking and head movement data. Additionally, the Physics Toolbox Sensor Suite app can collect GPS, acceleration, sound, lighting, and many other environmental factors.

To achieve the identified research objectives, we conducted a pilot study in which we benchmarked participants cycling behaviors and performance in a real-life environment and the corresponding IVE setting. The benchmark study has a within-subjects design to control for variance between subjects.

4.2.3 Experiment procedure

As a pilot study, a total number of six participants (mean age = 30.3, SD = 3.3) were recruited for the experiment. All participants are 18 or older, without color blindness, and familiar with the chosen corridor.

For the IVE bike simulator experiment, after arriving at the lab space, the participants were asked to wear Android smartwatches and fill out a pre-experiment survey about demographic information. Before the formal experiment, the participants were placed into a training scenario to familiarize themselves with the virtual environment, navigation, as well as calibration for the eye tracker and bike simulator. The experimental task was cycling to the end of the experimental area, as indicated in Figure 3.2. After the experiment, the participants were asked to sign up for the on-road study to take place in a few weeks after. The IVE experiments took approximately half an hour to complete.

For on-road experiments, the same bicycle model was utilized, instrumented with sensors to collect a range of variables. All on-road tests were conducted on clear weather days, during peak traffic hours. The same steps as the IVE experiments were followed for the on-road study.

4.2.4 Performance Metrics and Analysis

To study the validity of the IVE bike simulator in different contextual settings, all the performance data are calculated separately for different road segments. As shown in Figure 3.2, road segment 1 is the area between intersections 1 and 2, it includes a 4% downhill grade with a wall on the right to block the parking lot. Road segment 2 is a level road, and the parking lot is visible to cyclists.

The following metrics are measured as indicators of participants' performance: speed of the cyclist (km/h), head movements (three-dimensional unit vector), gaze direction (gaze focus on the current field of view), and heart rate (beat per minute, BPM). All performance measurements include average and standard deviation (SD) in different road segments. Performance data from the on-road and simulator experiments were compared to assess the relative and absolute validity of the simulator. Validity refers to the ability of a simulator to accurately represent real-world driving (Wynne, Beanland, and Salmon 2019). There are two major forms of validity: absolute validity (direct value comparison of a simulator and on-road testing) and relative validity (the same patterns or effects are observed even if the study failed to establish absolute validity). In this study, absolute validity for on-road and simulator data was assessed using paired sample t-tests at a level of significance of 0.05. If absolute validity is failed to establish, Pearson's correlation between the two settings is used to verify the relative validity.

Table 4.1: Performance comparison of IVE simulator and on-road test

Performance	Road Segment 1		Road Segment 2	
	IVE mean (SD)	On-road mean (SD)	IVE mean (SD)	On-road mean (SD)
Speed(km/h)	13.87(1.16*)	18.44(3.97*)	16.31(0.49*)	20.10(3.20*)
Left/right Head movement	0.061(0.14)	0.089(0.084)	0.066(0.13)	0.066(0.041)
Up/down Head movement	-0.0018* (0.049)	-0.19* (0.050)	0.019(0.056)	-0.10(0.044)
Left/right gaze (pixel)	972.06(78.94)	937.38(99.65)	961.58(79.68)	1049.22(240.83)
Up/down gaze (pixel)	607.38(34.27*)	520.61(88.76*)	597.97* (40.74*)	536.17* (76.52*)
Mean Heart Rate(bpm)	86.19(2.06)	96.14(2.58)	91.63(1.83)	97.48(1.96)

4.3 Results and Discussion

Among the six participants, one participant was missing the eye tracking data and another participant was missing heart rate data for the IVE experiment. The corresponding data was excluded from reporting. Table 4.1 shows all validity results. Overall, the on-road experiment has a higher average speed and standard deviation than the IVE experiment, however, there is no significant difference in the average speed across both road segment 1 ($p=0.09$) and segment 2 ($p=0.23$); the difference in the standard deviation of speed is significant for both segments ($p1 = 0.04$, $p2 = 0.002$), and Pearson correlations are positive for both, while only segment 1 is significant (0.865, $p = 0.026$), which means relative validity is only achieved for road segment 1. After the downhill road segment, as the speed is higher, participants in real-road tend to have more control over the speed, resulting in a larger variance in the SD of their speed. This can be explained by the difficulty in modeling all real-world elements within the VR environment. The weight of the users is not considered in the IVE bike simulator either, which in the real world would impact their acceleration and resultant speed after the downhill section. Furthermore, road friction is another important factor that can affect the SD of the speed - in real road segment 2 most participants just keep freewheeling without pedaling, while in the IVE bike simulator, they keep pedaling to maintain the current speed.

For head movement, all pairs' t-test results are not significantly different except for the up and down head movement in road segment 1 ($p = 0.002$). Pearson correlation for the up/down head movement in road segment 1 is positive but not significant (0.689, $p = 0.20$). After checking the video recording from HoloLens 2, we found that in the real road, there are several manholes on the road (4 in segment 1 and 7 in segment 2) which are not modeled in the IVE. The presence of pedestrians was not modeled in the IVE which may also explain the significant difference between the vertical head movement. This finding indicates that on real roads, cyclists will lower their heads due to their concern with manholes or other roadway conditions and contextual settings.

For gaze direction, no significant differences are found across left and right gaze directions. Concerning up and down gaze direction, only the average of up/down gaze direction in road segment 1 is not significant. Participants have significantly more up-and-down scanning behaviors on-road than in IVE for both segments. The average gaze center is lower in on-road tests in road segment 2. Furthermore, according to Pearson correlations, relative validity is not established for other pairs either: mean up/down gaze direction in segment 2 (-0.276, p = 0.653), the SD of up/down gaze for both segment 1 (-0.660, p = 0.226) and 2 (0.013, p = 0.984). The trend in gaze direction is similar to head movement as participants will scan up and down more frequently due to the real road complexity.

For heart rate data from the smartwatches, there is no significant difference in all comparisons, the overall distribution of HR can be seen in Figure 4.1 B-1, and he distribution of IVE and real road are similar to each other.

4.4 Conclusion

The goal of this study is to evaluate the use of an IVE bike simulator in understanding cyclists' behavior with multiple low-cost human sensing devices. A pilot study was conducted both in an IVE and on a real road. Various sensors are applied to ensure that similar data output is obtained. Both absolute and relative validity is established across a range of cyclist performances. Results show that most of the performance measurements have absolute validity, but some of the features from eye tracking, most of which are in the vertical direction, could not establish either absolute or relative validity. This phenomenon may be caused by road geometry changes, the appearance of other road users, and hardware limitations (especially the headsets used in the study). Overall, the promising results indicate that the IVE bike simulator can be further utilized for understanding cyclists' behaviors. In addition to the difficulty in modeling all real-world elements as discussed above, the study is also limited in: (1) The small sample size of participants. After the pilot study, an experiment with a larger number will be conducted in the near future. (2) Another limitation is that participants could not change the bike gear within the current IVE bike simulator as the current version does not support gear change; therefore, the gear is consistently in the middle level for all experiments, which might potentially affect the speed change. (3) Lastly, as psychophysiological data are highly event-based, we did not annotate or collect specific events that took place within the real-road condition and in our future studies, we will consider collecting and integrating such information within the design of IVE. Further work can be done by including a wider range of age groups, validating more measurements from the sensors, and exploring different locations.

Chapter 5

Psycho-physiological Measures on a Bicycle Simulator in Immersive Virtual Environments: How Protected/Curbside Bike Lanes May Improve Perceived Safety

Bicycling, as a traditional means of mobility, has gained more popularity in recent years. Bicycle mode share is rising in response to issues in modern cities, such as increases in traffic jams, land use, energy consumption, air pollution, climate change, and physical inactivity (Flusche 2012; Rupi and Krizek 2019). This trend has continued during the COVID-19 pandemic, as it is reported that bicycling levels have significantly increased in many countries despite lockdowns and travel restrictions (Buehler and Pucher 2021). However, there has been an alarming increase in bicyclist fatalities over the last decade. The National Highway Traffic Safety Administration's report shows that in the United States, the number of bicyclist fatalities has increased by more than 35% since 2010 (NHTSA 2021a). One potential reason for this increase is due to the auto-centric nature of much of US roadway design, which often overlooks the safety and comfort needs of vulnerable road users like bicyclists (Schultheiss, Sanders, and Toole 2018). For instance, vehicles are now safer for the vehicle occupants with the development of active (e.g., electronic stability control, automatic emergency braking, head-up displays, and other systems that have been used to assist in the prevention of a crash) and passive safety features (e.g., seat belts, air-bags, laminated glass, and other settings to protect occupants during a crash) on a vehicle. However, to increase the comfort of the in-vehicle occupants, some adjustments on modern vehicles such as smaller windows, wider pillars, and larger headrests, may increase the risks to other vulnerable users, as indicated by a vehicle/bicycle crash analysis with 3350 motor vehicle/bicycle crashes in the U.S. (**lusk2015database**). In addition, these technologies don't help to avoid the 'Looked but failed to see' errors when the drivers saw the danger too late to avoid collision when they are overtaking another road user like bicyclists (**herslund2003looked**; **koustanai2008statistical**). Furthermore, current data sources have some limitations. First, the police crash reports mainly focus on vehicle-bicyclist crashes that result in fatalities. Second, hospital data recently drew more attention TO traffic safety especially for vulnerable road users as it can provide real-world fatality data with different accident types, such as single-bicycle crashes (**myhrmann2021factors**). The first study in North America with ambulance records showed that cycle tracks were safer compared to biking on the road. The cycle tracks had a 28% injury rate with 2.5 times as many bicyclists compared to parallel roads without bike provisions (**lusk2011risk**). Specifically, for bike lane designs, different treatments on the bike lane may result in different cycling behaviors. A study in the Netherlands with 62 participants from two experiments with different edge types has reported that for older bicyclists, the shoulder and edge strip treatments were related to more efficient path use and safer distances from the verge (**westerhuis2020cycling**). A comparison study of 70 young adults from five European cities (Oxford, London, Amsterdam, Houten, and Groningen) verified that physical segregation between cyclists and vehicles reduces the likelihood of stress (Teixeira et al. 2020). Even for the protected bike lane treatment, the different levels of separation can contribute to different risk levels in protected bike lanes (**cicchino2020not**). Although these studies shine A light on the importance of evaluating the impact of roadway design on cycling behaviors and emotional states, generally speaking, there are a very limited number of studies as well as sources of data for evaluations of roadway designs for vulnerable road users, including bicyclists.

The absence of applicable data has been recognized as a limiting factor for many transports and urban planning studies, especially for bicyclists (**willberg2021comparing**). Therefore, it is necessary to identify innovative approaches to understand how bicyclists' behavior, sense of safety, and comfort are affected under different roadway conditions during the design and planning phases. To achieve this, bicyclists' behavior and psycho-physiological state data should be explored and evaluated in different roadway settings.

Various methods of collecting bicycle safety, risk, comfort, and behavior data have been utilized in the past. These studies spanned surveys, observational studies, naturalistic studies, and simulation studies. For example, interviews and surveys ask participants about their behaviors and comfort in a certain design context either by imagination or after an actual bike ride. However, the subjective response does not always reflect what the participant will do in a real-road setting and can suffer from hypothetical bias (Fitch and S. L. Handy 2018). Observational studies can record realistic changes within the environment and bicyclists' responses in real-world conditions but are unable to track bicyclists' physiological changes (**chidester2001pedestrian**).

Naturalistic studies can further record bicyclists' responses in real-world conditions through different sensing modalities such as GPS, ECG, or mobile eye trackers (Rupi and Krizek 2019). However, these studies have potential risks for participants as they may be placed in dangerous roadway settings (e.g. distraction in high-traffic density areas) (Stelling-Konczak et al. 2018). Furthermore, in naturalistic studies, it is difficult to control many environmental factors that may impact the independent variables, especially for physiological and behavioral factors, which makes causal inferences especially difficult (Fitch, Sharpnack, and S. L. Handy 2020; Teixeira et al. 2020). Experimental studies conducted through Immersive Virtual Environments (IVE) are an emerging approach that minimizes the hypothetical bias of subjective surveys while offering a controlled, low-risk, and immersive environment to evaluate the responses of bicyclists to different roadway designs and conditions. One of the main challenges in previous IVE simulation studies was the integration of human sensing techniques into the experiment. In IVE-related literature, participants' physiological responses have been applied to evaluate different design alternatives for buildings (Francisco et al. 2018), hospitals (Chías et al. 2019), and other civil infrastructure systems (Awada et al. 2021). However, only a few recent studies have applied bicyclist physiological sensing in IVE simulators (Cobb, Jashami, and Hurwitz 2021), and a deeper understanding of bicyclists' psychological and physiological responses in different roadway designs and conditions is still needed. Overall, we still have very a limited understanding of bicyclists' physiological responses in different roadway environments, especially in IVE studies. Some naturalistic studies have been conducted to evaluate bicyclists' behavior and physiological responses in different contextual settings (**Guo2022**; McNeil, Monsere, and Dill 2015; Rupi and Krizek 2019; Teixeira et al. 2020). These preliminary studies revealed that psycho-physiological metrics (e.g., heart rate (HR), gaze variability, and skin conductance) are indicators of how participants' behaviors and perceptions may change in different contextual settings.

By integrating a bicycle simulator with an IVE, this research aims to overcome some of the limitations found in previous research by conducting repeated measures experiments to collect bicyclists' physiological responses (specifically, gaze variability and HR) in different urban roadway designs. With an IVE, we are able to control other roadway environmental factors, such as infrastructure design, vehicle traffic volume, traffic signal phase, vehicular speeds, and gaps, vehicle types, lighting, and weather conditions to better quantify the relationships between bicyclists' behavior (e.g., speed, lane position) and physiological responses (e.g., gaze variability, HR). These measurements can be used as surrogate data to better understand how different types of roadway conditions and infrastructure designs result in higher rates of physiological stimulation. In this paper, we leverage human sensing tools (e.g., wearable devices) together with a bicycle simulator in IVE to specifically evaluate and model the bicyclists' psycho-physiological and behavioral responses. We consider three roadway design scenarios (shared bike lane [sharrows], standard curbside bike lane, and protected bike lane with flexible delineators) and evaluate bicyclists' HR, gaze measures, and speed within each environment. After performing extensive feature extractions on HR and gaze data, we leverage linear mixed-effect models to compare bicyclists' responses across the simulated environments. With results from 50 participants (23 female/27 male; aged 18-68; student and faculty bicyclists) in the IVE simulator study, we investigated the following research questions: How can protected/curbside bike lanes improve bicyclists' cycling experience in terms of perceived safety, cycling behavior and psycho-physiological responses? The results reveal the subjective and objective preferences of different bicycle infrastructures while riding the simulator bike through the virtual environment and provide suggestions to policymakers on how to potentially increase the bike share in certain environments.

In this paper, we first provide detailed background on previous bicycle safety studies in naturalistic, simulated, and virtual reality (VR) environments. We then discuss the applicability of bicyclists' physiological responses and gaze measures in understanding their behaviors and physiological states, such as stress level and cognitive load. By providing the methodology of our experimental design, we dive into the details of the experiment. We then provide the results of bicyclists' HR and gaze variability together with bicycling performance within each environment with a linear mixed-effect modeling approach. We conclude with a discussion on comparing different physiological responses and the reasons behind the results.

5.1 Experiment Design

This research studies the effect of different roadway designs on bicyclists' physiological states. The independent variables are demographic information (i.e., age, gender, bicycling attitude, and VR experience), the subjective realism of the IVE, as well as the three categorical variables of different roadway designs in IVE with a bicycle simulator: (1) the as-built shared bike lane environment (sharrows), (2) separate bike lane, and, (3) protected bike lane with pylons. The dependent variables are different measurements of cycling performance (i.e., speed, lane position) and physiological responses (i.e., eye tracking and HR features) from integrated or mobile sensors.

5.2 Experiment Procedure

Figure 5.1 shows the experiment procedure. Once a prospective participant signed up, a researcher contacted the participant both via email and phone call a day before the experiment to confirm their reservation and health condition (due to the COVID-19 requirement). Upon arrival, each participant is asked to sign the consent form approved by the IRB office and put on two smartwatches on both wrists, before completing the pre-experiment survey. After finishing the pre-experiment survey, instructions are given on how to use the VR headset, controllers, and bike simulator. After the bike is adjusted to a comfortable position, the participant mounts the bike and puts on the headset. Next, the participant is guided through the eye tracker calibration. After the IVE system setup, the participant is placed into a familiarization scenario (without any vehicle traffic) to become accustomed to interacting with the IVE. In this environment, the participant can practice pedaling, steering, and braking and the practice procedure can be repeated until the participant feels comfortable. If the participant feels any motion sickness, they may stop the experiment at any point and still receive compensation for participation.

Once the participant is comfortable in the training environment, they experience the three design scenarios in random order, where each experiment trial lasts about two minutes, with a two-minute break between each scenario. Once the participant has completed all three scenarios, they are asked to complete the post-experiment survey. On average, each participant spends 30 minutes completing the experimental procedure.

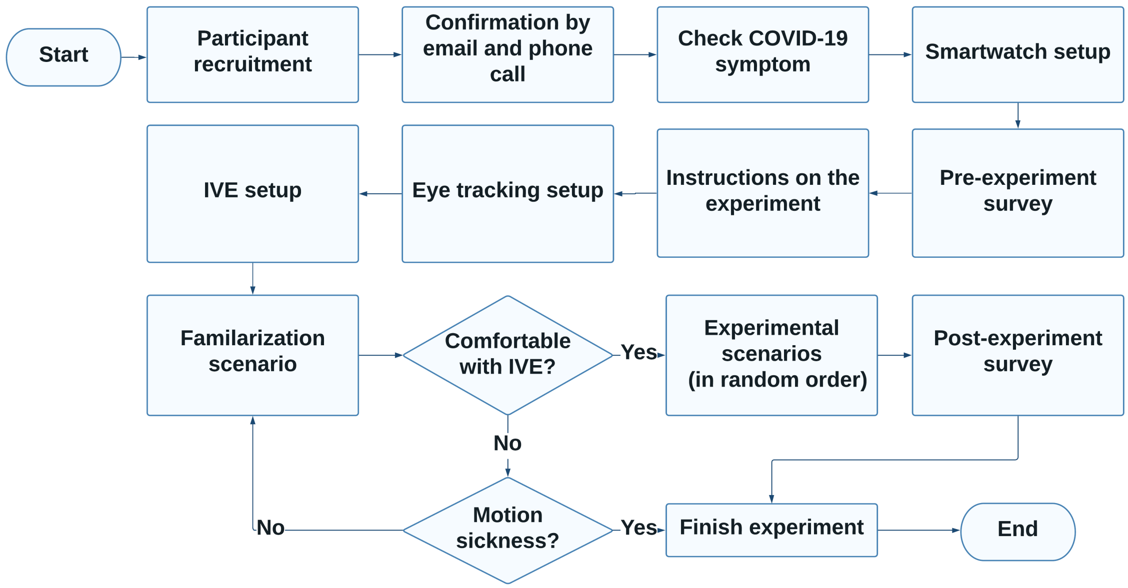


Figure 5.1: Experiment procedure

5.3 Participants

51 participants were recruited for the experiment. Most of the participants are local bicyclists, university students, and faculty members who are familiar with the study corridor. All participants are 18 or older and without color blindness. During the study, one participant could not finish the experiment due to motion sickness. For the remaining 50 participants (23 female and 27 male), the mean age is 34.1 with a standard deviation of 12.9 (1 participant did not reveal his/her age information); the age distribution is shown in Figure 5.2.

5.4 Statistical Modeling

a Linear Mixed Effects Model (LMM) was chosen to model the different response variables across participants. LMM facilitates the analysis as it has the ability to (1)

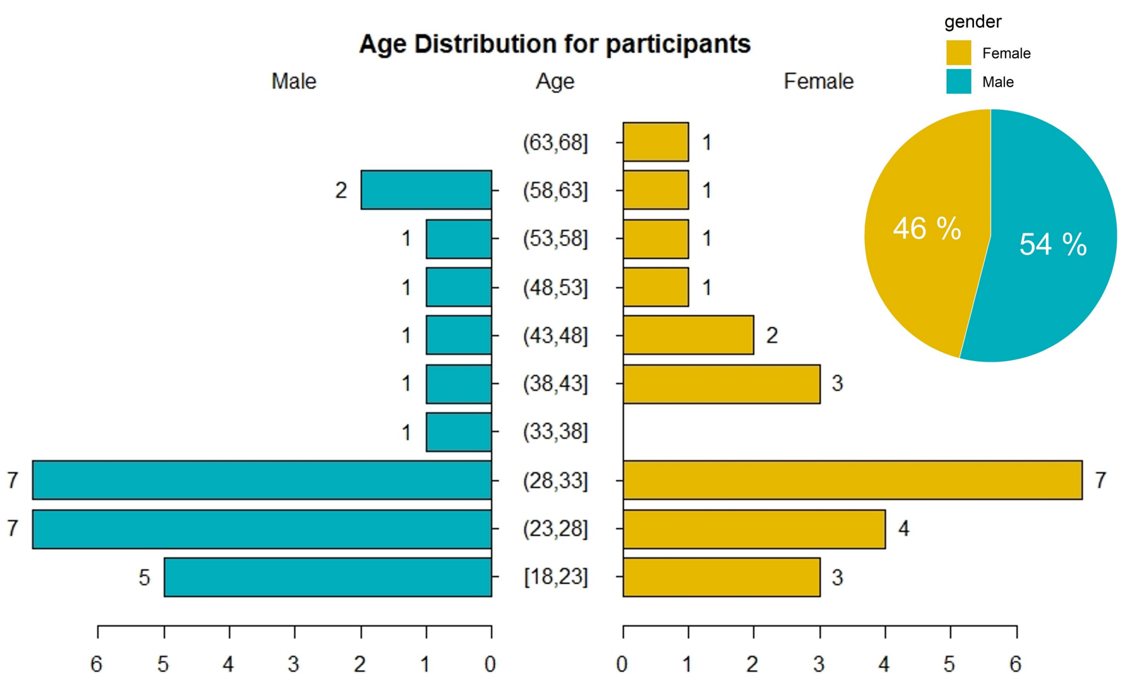


Figure 5.2: Age and gender distribution of participants based on the demographic data. Overall 23(46%) are females and 27(54%) are males. One male participant didn't reveal his age information

characterize group and individual behavior patterns in a formal way, (2) acknowledge both group and individual differences, and (3) incorporate additional covariates (Krueger and Tian 2004).

To evaluate participants' physiological responses in different road environments, as shown in Table 5.1, different behavioral and physiological responses are treated as dependent variables in each LMM model, including cycling performance (speed and lateral lane position), eye tracking metrics (SGE, GTE, PRC, and mean fixation length), and HR metrics (mean HR and number of HR change points). Independent variables include the type of bicycle infrastructure (as-built, separate bike lane, and protected bike lane), age (older or younger than 30), attitude towards cycling (based on pre-survey response), prior VR experience, and participants' sense of realism for bike speed and braking in the IVE. Additionally, each participant is treated as a random effect in the model. If statistically significant effects are revealed in any LMM models for the scenario variable, post hoc contrasts will be performed for multiple comparisons using Fisher's Least Significant Difference (LSD) L. J. Williams and Abdi 2010. All statistical analyses were performed at a 95% confidence level ($\alpha = 0.05$).

5.5 Results

This section reports the results of the experiment. The following subsections describe the summary statistics (from the pre-and post-experiment surveys), the bicyclists' physical behavior (cycling speed and lateral lane position), and physiological responses (eye tracking and HR) in different roadway designs.

Table 5.1: Summary of independent and dependent variables in LMM models

Variable Type	Variable Name	Categories	Data Source
Independent (Categorical)	Scenario	3	IVE Design
	Age	2	Pre-survey
	Gender	2	Pre-survey
	Type of bicyclist	4	Pre-survey
	VR experience	4	Pre-survey
	Realism of bike steering	5	Post-survey
	Realism of bike speed	5	Post-survey
Dependent (Continuous)	Cycling speed	km/h	Unity
	Lateral lane position	m	Unity
	Horizontal gaze variability	pixel	Eye Tracking
	Percentage of road center gaze	percentage	Eye Tracking
	Mean fixation length	second	Eye Tracking
	Stationary gaze entropy	bits	Preprocessing
	Gaze transition entropy	bits	Preprocessing
	HR	bpm	Smartwatch
	Number of HR change points	count	Preprocessing

5.5.1 Survey Response

All participants indicated that they have some level of prior knowledge of VR, although only one participant owns VR equipment and uses it regularly, as shown in Figure 5.3-a. The majority of the participants have a positive attitude toward cycling, as shown in Figure 5.3-b, with only two participants expressing hesitancy of cycling under any condition.

In the post-experiment survey, the majority of participants indicated that the virtual environment was immersive, with 94% of participants choosing a 4 or 5 on the 5-point Likert scale (mean=4.42), with 4 and 5 indicating ”immersed” and ”very immersed”, respectively. Most participants also found that the virtual environment was to scale (94% chose 4 or 5, mean=4.54). The participant’s feelings of speed and steering realism were both above average with 50% indicating a 4 or 5 level of realism (mean

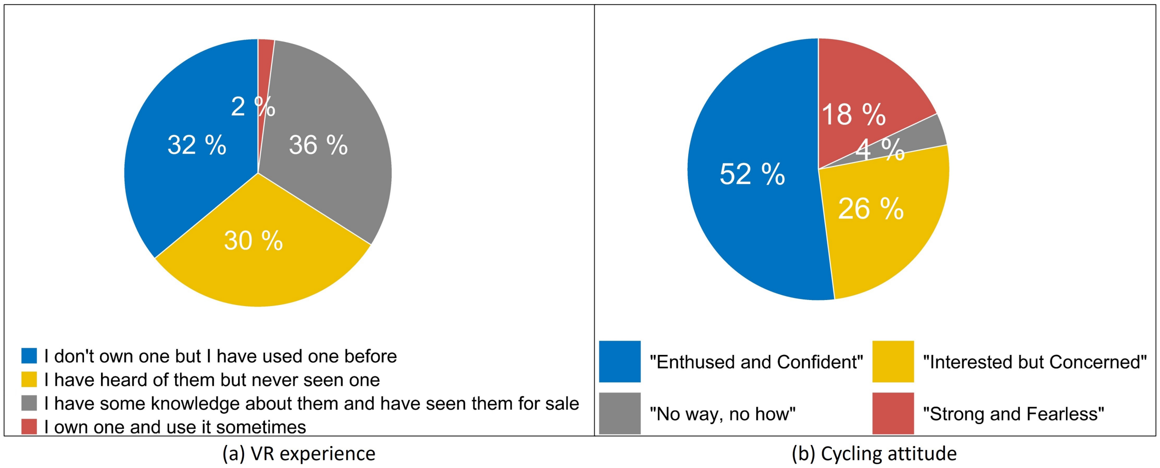


Figure 5.3: Summary of some survey responses. (a) Prior knowledge on VR, (b) Type of bicyclist

= 3.56), and 54% indicating a 4 or 5 for steering realism (mean = 3.60). Figure 5.4 shows the results of participants' scenario preferences. Participants indicated an overwhelming preference towards the protected bike lane (69% rate it as the safest), followed by separate bike lane (22% rate it as the safest), and the as-built scenario rated the least preferred safe environment (10% rate it as the safest).

5.5.2 Cycling Performance

Two LMMs are built individually for average speed and lateral lane position to estimate the relationship between the independent variables and participants' cycling performance. For the mean speed LMM, there is a significant difference between the as-built and protected bike lane scenarios ($\beta = -1.209, SE = 0.383, p < 0.01$). Bicyclists' mean speed in the protected bike lane with pylons scenario (13.88 km/h) is significantly lower compared to the as-built scenario (15.09 km/h). No significant differences are found between the separate bike lane scenario (14.94 km/h) and the as-built scenario, as shown in Figure 5.5 - a. Similarly, there is no signif-

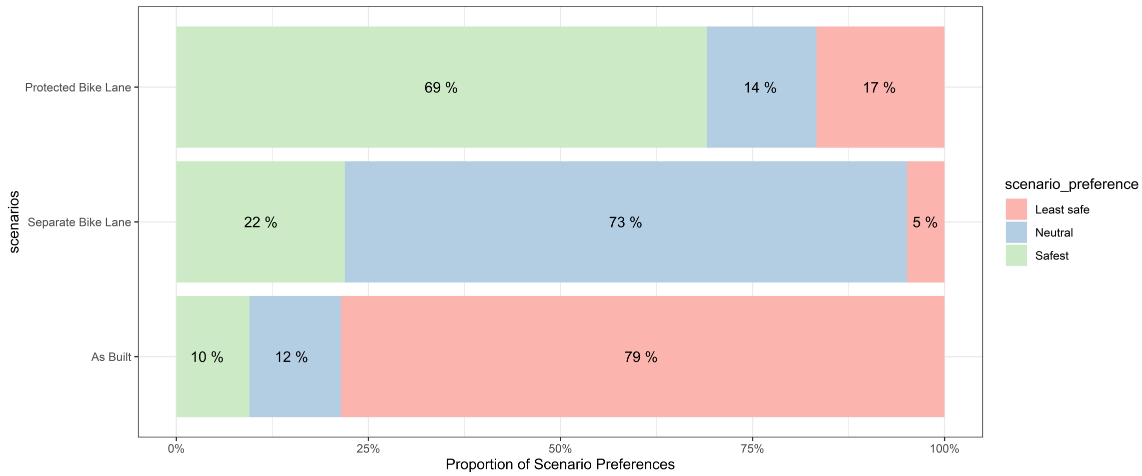


Figure 5.4: Scenario preference across participants based on the post-experiment survey. Note that 94% of participants choose a 4 or 5 on the 5-point Likert scale (mean=4.42), with 5 indicating a full immersion

icant difference in participants' speed between the bike lane scenario and protected bike lane with pylons. The random effect for the mean speed model is significant ($\beta = -1.209, SE = 0.383, p < 0.001$), suggesting that it is necessary to treat the participant as a random factor in the model. This is also indicative of individual differences across participants in their cycling performance.

For the mean lateral lane position, no significant differences are found across the three scenarios, although the difference between as-built and protected bike lanes with pylons scenarios are marginally significant ($\beta = -0.129, SE = 0.070, p = 0.068$). As shown in Figure 5.5 - b, the average distance to the roadside curb for the three scenarios (as-built, separate bike lane, and protected bike lane with pylons) are 0.97 m, 0.88 m and 0.84 m respectively. The greater the average distance to the curb the smaller the lateral distance between the bicycle and the vehicle. Therefore, there is a trend of participants moving closer to the curb to stay away from vehicles with the presence of separate bike lanes or protected bike lanes with pylons. The random effects for this model are significant ($\beta = 0.971, SE = 0.067, p < 0.001$) as well.

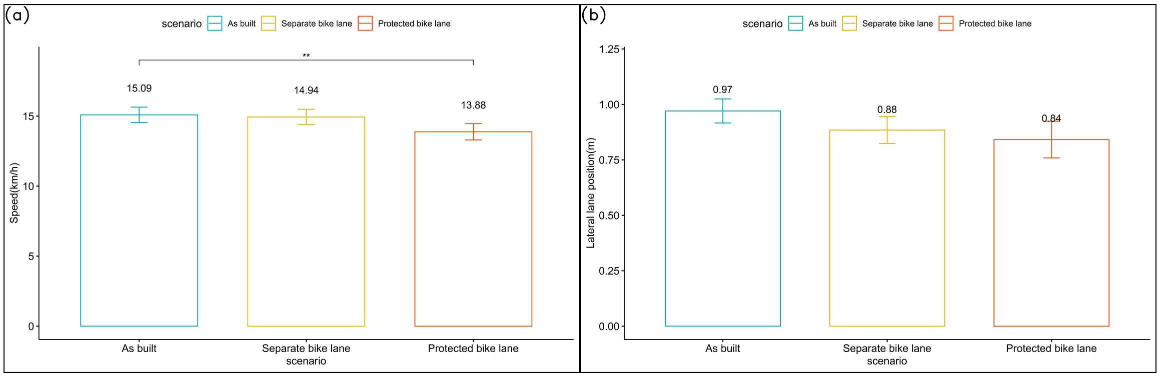


Figure 5.5: Cycling performance measured through speed (a) and lateral position (b) across different scenarios. Note that there is a trend for participants to move closer to the road curb to stay away from vehicles with the presence of a separate bike lane or protected bike lane with pylons.

5.5.3 Eye Tracking

Five LMM are built individually for each eye-tracking dependent variable (Table 5.1) to estimate the relationship between the independent variables and participants' eye-tracking metrics (horizontal gaze variability, PRC, mean fixation length, SGE, and GTE). We first plot the eye-tracking heat map in the field of view to get an overview of the gaze distribution. As shown in Figure 5.6, visual observations from the gaze heat map indicate that the as-built scenario has a more dispersed distribution than the other two scenarios. The separate bike lane scenario appears to have a higher concentration in the center of the gaze area, followed by the protected bike lane with pylons scenario.

As illustrated above, an LMM is built for evaluating the relationship between participants' horizontal gaze variability and the independent variables. In Figure 5.7, the result of the horizontal gaze variability model shows the random effects were significant ($\beta = 86.257, SE = 32.990, p < 0.05$), which suggests that it was necessary to treat the participant as a random factor in the model. Both the sepa-

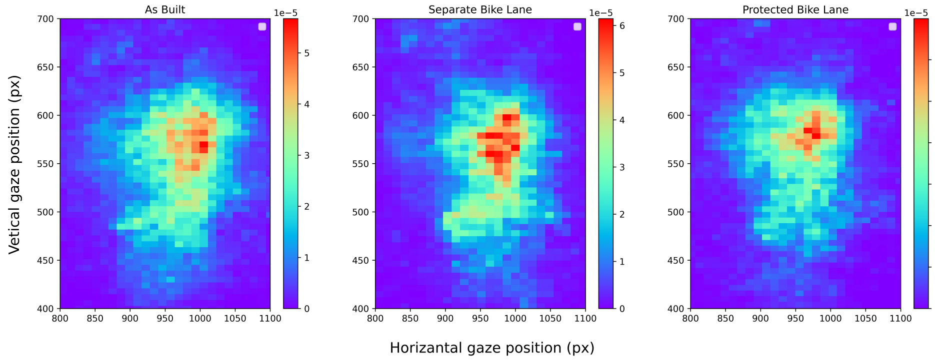


Figure 5.6: Gaze density heat map for different scenarios. Note that visual observations from the gaze heat map indicate that the as-built scenario has a more dispersed distribution than the other two scenarios

rate bike lane and protected bike lane scenarios are statistically significant predictors for the horizontal gaze variability ($\beta = -16.349, SE = 4.288, p < 0.001$ and $\beta = -12.645, SE = 4.278, p < 0.01$, respectively). As shown in Figure 5.7 - a, a significantly lower horizontal gaze variability is observed both in the separate bike lane and protected bike lane, which indicates that participants are more focused directly ahead rather than laterally looking around the road environment. Another significant factor revealed by the model is the realism score of the bike speed from the post-experiment survey ($\beta = -13.991, SE = 6.287, p < 0.05$). Generally speaking, the higher realism of bike speed the participants indicate, the lower horizontal gaze variability they show during the experiment (Figure 5.7 - b, except for the small group who selected 2). No significant results are found in terms of the steering realism score.

A similar LMM is built for the percentage of road center fixation. As shown in Figure 5.8, a similar result is presented by the LMM for the horizontal gaze variability; the random effects are also significant ($\beta = 91.993, SE = 6.045, p < 0.001$). For the independent variables, both the separate bike lane ($\beta = 4.083, SE = 0.947, p < 0.001$)

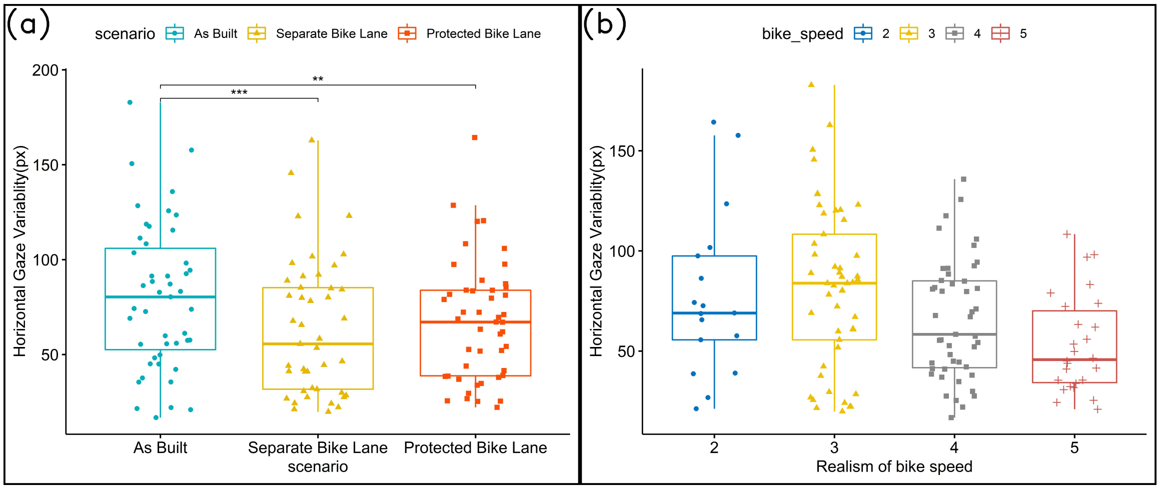


Figure 5.7: Horizontal gaze variability within different scenarios (a) as well as within different ratings of the realism of the bike speed (b). Note that a significantly lower horizontal gaze variability is achieved both in the separate bike lane and protected bike lane. Additionally, the higher realism of bike speed the participants indicate, the lower horizontal gaze variability they show during the experiment

and protected bike lane ($\beta = 2.558, SE = 0.938, p < 0.01$) scenarios are statistically significant. The percentage of road center fixation in the separate bike lane is slightly higher than in the protected bike lane, which aligns with visual observation of the gaze heat map (Figure 5.6). This result indicates participants focus their gaze most on the road center in the separate bike lane. The realism score of the bike speed from the post-experiment survey is also significant ($\beta = 2.892, SE = 1.218, p < 0.05$).

The LMM model for the mean fixation duration shows the random effects are significant ($\beta = 0.242, SE = 0.051, p < 0.001$), and both the separate bike lane and protected bike lane scenarios are statistically significant predictors of mean fixation duration ($\beta = 0.015, SE = 0.007, p < 0.05$ and $\beta = 0.014, SE = 0.007, p < 0.05$, respectively). As shown in Figure 5.9, significantly higher fixation duration is observed both in the separate bike lane and protected bike lane scenarios compared to the as-built scenario. No significant results are found for other independent variables.

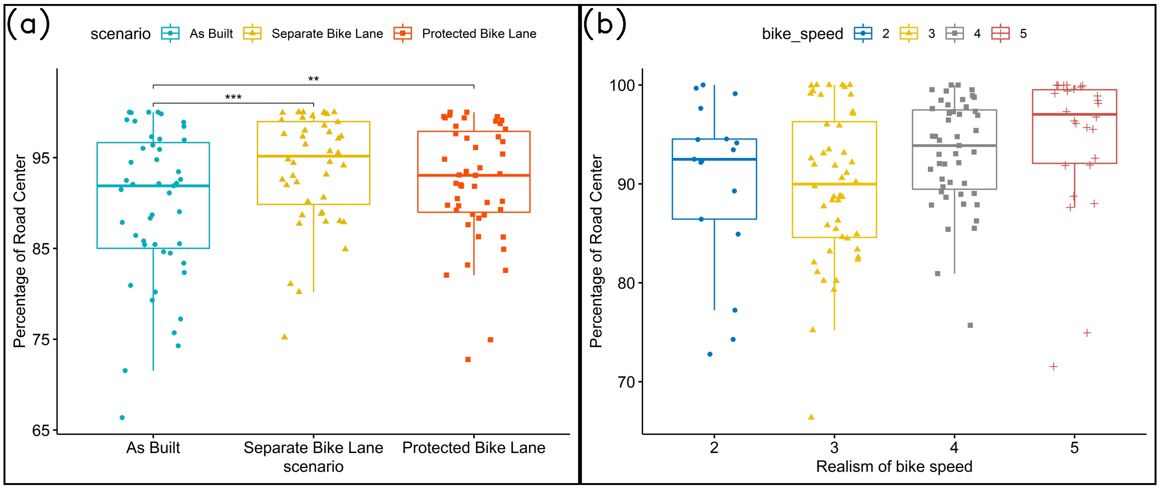


Figure 5.8: PRC and mean fixation length within different scenarios. Note that the percentage of road center fixation in a separate bike lane is slightly higher than in the protected bike lane, which aligns with the visual observation in the gaze heat map

Two LMMs are built for SGE and GTE. In both models, the random effects are significant ($\beta = 1.951, SE = 0.413, p < 0.001$ and $\beta = 0.791, SE = 0.306, p < 0.05$, respectively). Other than the random effects, only the separate bike lane in the SGE model is a significant predictor ($\beta = -0.224, SE = 0.079, p < 0.01$). As shown in Figure 5.10 - a, the SGE in the separate bike lane environment is significantly lower than in the as-built environment. No significant results are observed in the GTE model.

5.5.4 HR

An LMM is built for mean HR during each experiment to compare the overall HR levels in different infrastructure designs. The random effects are significant ($\beta = 92.892, SE = 29.191, p < 0.01$). No significant results are found between different road designs (Figure 5.11 - a). A significant result is shown in the type of bicyclist attitude based on the survey result ($\beta = -8.410, SE = 4.082, p < 0.05$). As shown in

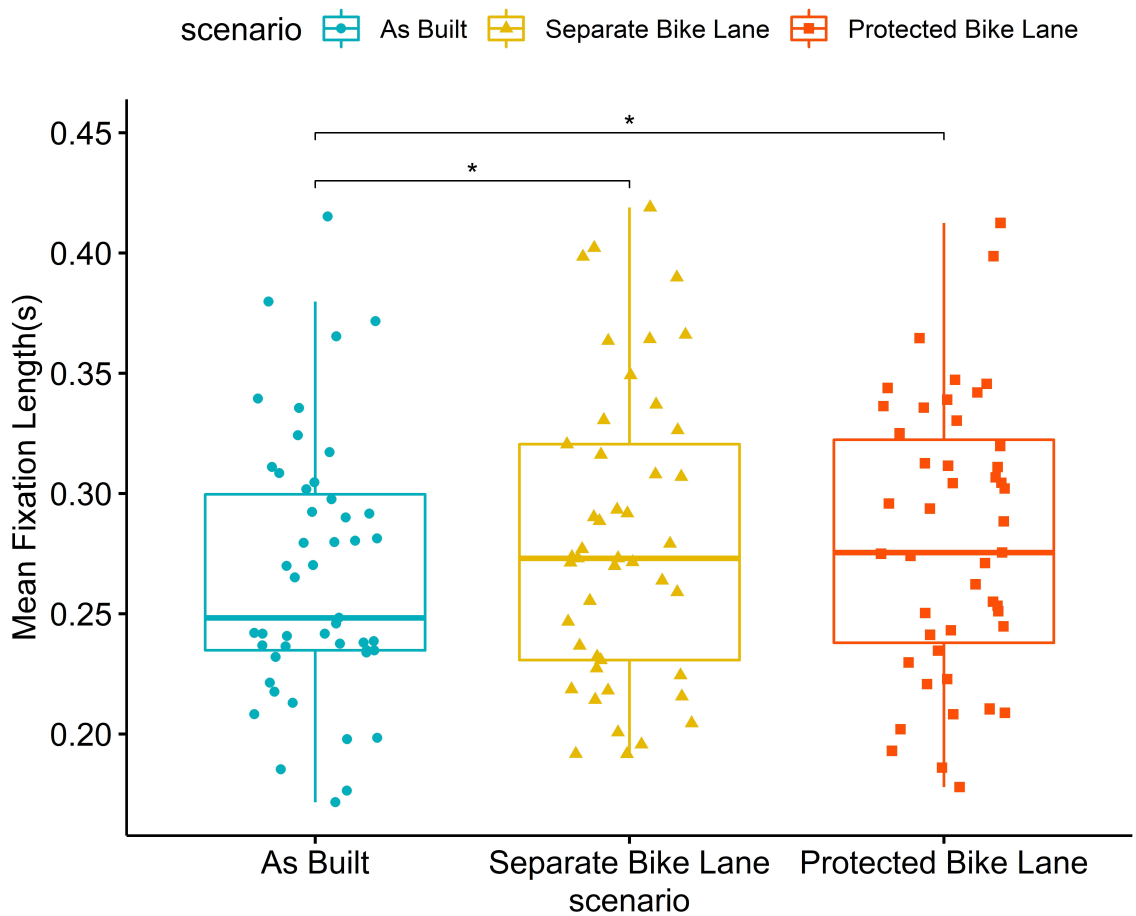


Figure 5.9: Mean fixation duration within different scenarios. A significantly higher fixation duration is observed both in the separate bike lane and protected bike lane scenarios compared to the as-built scenario.

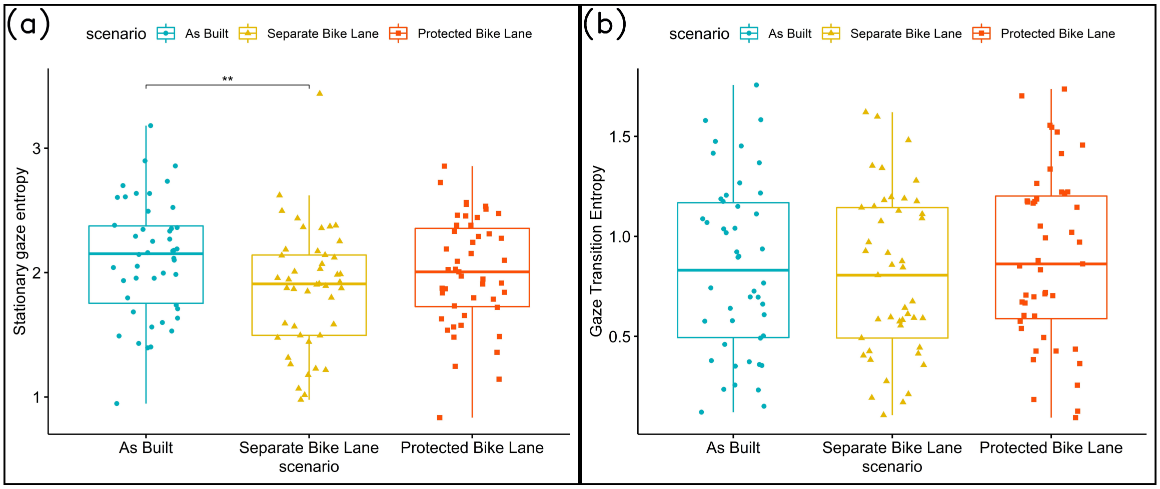


Figure 5.10: Stationary gaze entropy (a), and gaze transition entropy (b) within different scenarios. Note that the SGE in the separate bike lane environment is significantly lower than in the as-built environment.

Figure 5.11 - b, the more positive attitude participants have on biking, the lower mean HR they have during the experiment. Note that only two participants responded to the bicycling attitude question with 'No way, no how' so only six data points for this category are available.

In addition to the overall HR level, we are also interested in the abrupt changes in participants' HR. By utilizing the BCP method, we can extract the abrupt HR changes for each scenario. Figure 5.12 illustrates the average frequency of HR change points in different scenarios. The LMM model shows that both the separate bike lane ($\beta = -0.393, SE = 0.145, p < 0.01$) and protected bike lane ($\beta = -0.360, SE = 0.145, p < 0.05$) have a significantly lower frequency of HR change point than the as-built scenario. The frequency of HR change points in the as-built design is almost twice that of the separate bike lane and protected bike lane. The distribution of HR change points is shown in Figure 5.13. There are three peaks in Figure 5.13 - a, where all take place before the participant arrives at an intersection. Among the three intersections, the peak of the HR change point in the third intersection, which

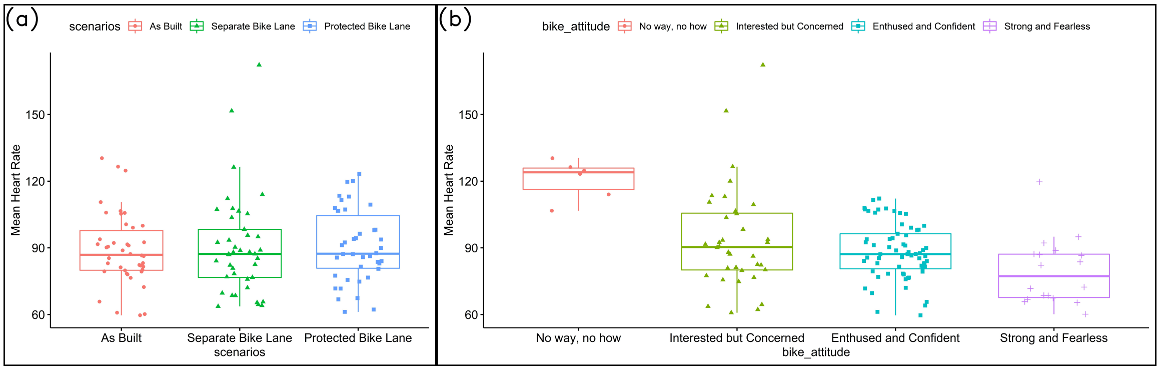


Figure 5.11: Mean HR within different scenarios (a) as well as attitude towards biking (b). Note that the more positive the attitude participants have on biking, the lower the mean HR they have during the experiment.

has a traffic signal, happens earlier than the other two intersections. Figure 5.13 - b is the density plot of HR change points for different scenarios. Each scenario appears to have two peaks, with the as-built design having higher peaks in the first intersection and the third intersection. The density plots of the separate bike lane and protected bike lane scenarios are smoother than the as-built design. This indicates that the HR change points in the as-built scenario are more subjective to roadway environmental changes. In other words, the separate bike lane and protected bike lane may reduce the effect of environmental changes on the HR changes.

5.6 Discussion

5.6.1 Cycling Performance

The results show that the roadway design can affect cycling performance. Among the three roadway designs, the speed in the protected bike lane is significantly lower than the other two designs. On average, participants cycled at a lower speed when they are separated from the vehicle lane. However, this contradicts a similar IVE

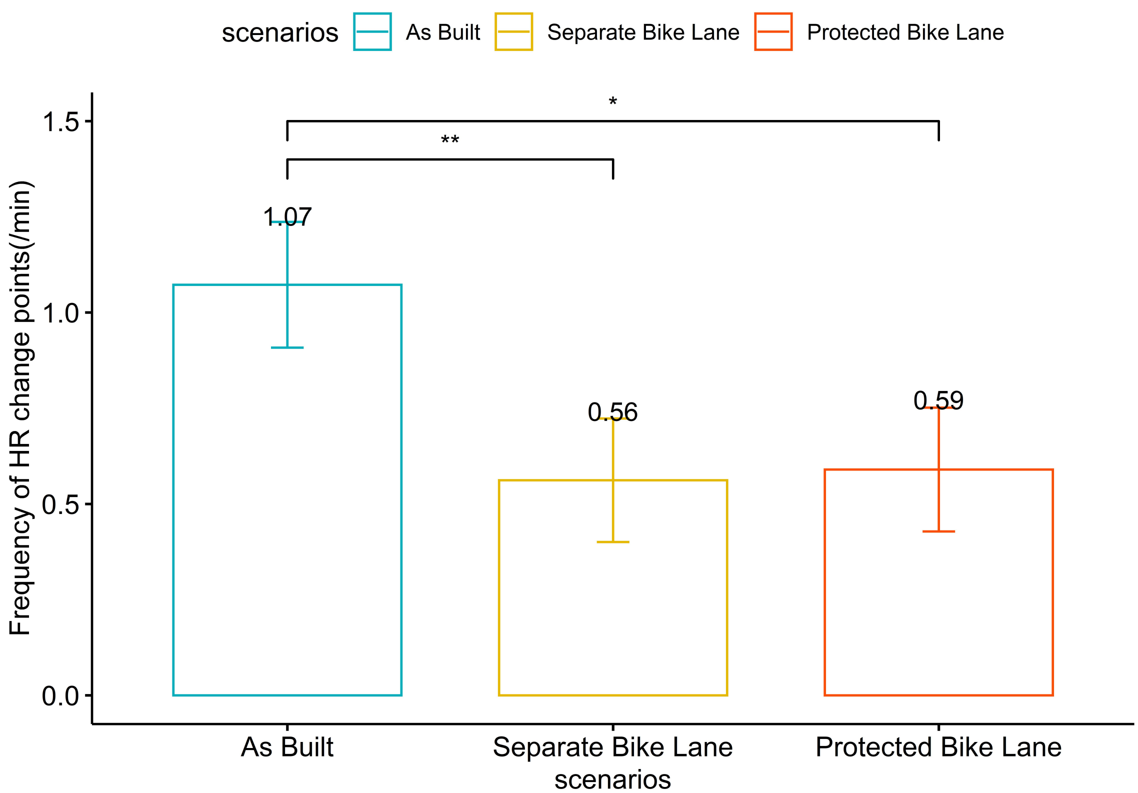


Figure 5.12: Frequency of HR change points within different scenarios. Note that both the separate bike lane and protected bike lane have a significantly lower frequency of HR change points than the as-built scenario.

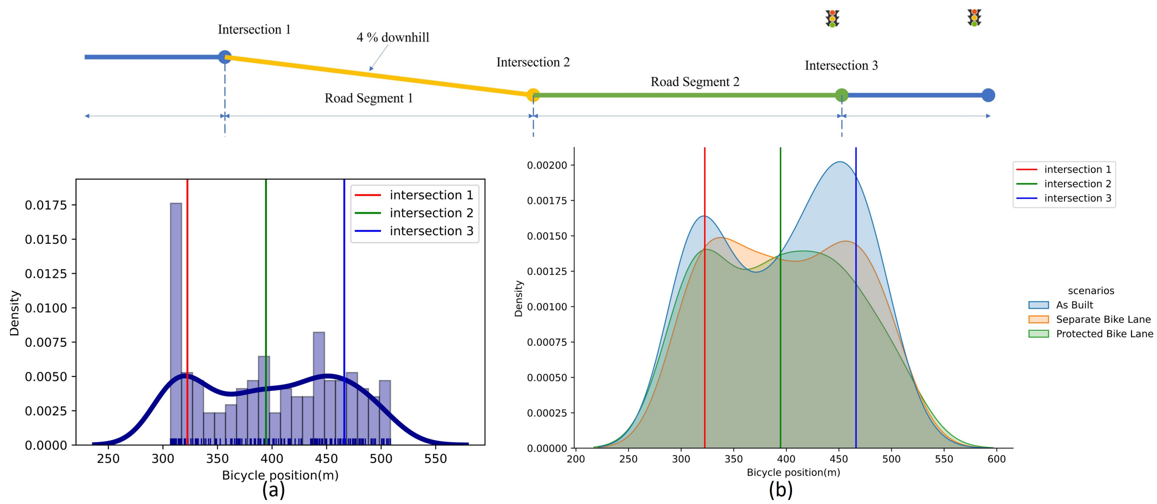


Figure 5.13: HR change points distribution (a) overall distribution on the road; (b) distribution over each scenario. Note that each scenario appears to have two peaks, with the as-built design having higher peaks in the first intersection and the third intersection. The density plots of the separate bike lane and protected bike lane scenarios are smoother than the as-built design, which indicates the HR change points in the as-built scenario are more subjective to roadway environmental changes.

study (with 50 participants) where bicyclists ride at lower speeds in the no bike lane condition versus the bike lane condition (Cobb, Jashami, and Hurwitz 2021). The differences may be due to (1) the different IVE settings. In (Cobb, Jashami, and Hurwitz 2021), A screen provides the forward view while in this study, a head-mounted display provides the forward view, this implies that different IVE settings will affect bicyclists' responses, as discussed by previous studies (Bogacz, Hess, Calastri, Choudhury, Erath, et al. 2020; Guo, Angulo, et al. 2022). (2) the different road environments. Our IVE is modeled from a real road and all the participants are local bicyclists, they are more familiar with the as-built designs. (3) The vehicle settings. The vehicles are randomly generated based on the empirically observed distribution of vehicle arrivals from the start point with a fixed routine in the vehicle lane. For some participants, when they are cycling in the shared vehicle lane in the as-built scenario, the approaching vehicles from behind will slow down and follow them until the vehicle lane is cleared. Therefore, in those cases, the participants will only see an open street ahead of them without any cars passing by. In the post-experiment survey, some participants mentioned that they were motivated to ride faster when no vehicles were passed by them. Based on our statistics of how many vehicles have passed the participants during the experiment, the average number of the passed vehicle during the experiment for as-built design (0.77) is less than the separate bike lane (0.88) and protected bike lane (1.03). (4) The traffic volume and speed. The traffic volume is relatively low in our experiment. On average, there is 0.9 vehicle passing the participant, which is lower than in related studies. Water Street is an urban road with a speed limit of 25 mph and vehicles within the IVE were designed to travel constantly at this speed limit, which may indicate that the effect of different roadway designs on bicyclists' speed is subjective to traffic volume and vehicle speed. In addition, some participants report that they feel the protected bike lane is narrower than expected, they want to avoid hitting the pylons during the experiment, which may potentially lower their speed and decrease the lateral distance to the curb. For the lateral lane position, the LMM model result is marginally significant, bicyclists tend to stay away from the vehicle lane when there is a bike lane, which will lead to a larger lateral distance when a vehicle is trying to pass them. This can help to increase bicyclists' comfort level and safety, which has been shown by many previous studies (McNeil, Monsere, and Dill 2015; Nazemi, Eggermond, Erath, Schaffner, et al. 2021).

5.6.2 Gaze Behavior

The gaze behaviors also vary in different roadway designs. Generally, compared with a separate bike lane and protected bike lane scenarios, participants in the as-built scenario have a wider horizontal distribution of fixations, indicated by a larger horizontal gaze variability and a lower PRC, which can be a sign of active searching strategy in the environment. The higher PRC in the separate bike lane and protected bike lane scenarios might be an indicator of higher cognitive workload (Engström, Johansson, and Östlund 2005), but this does not necessarily suggest a sub-optimal state for the gaze strategy. According to the Yerkes-Dodson law (Yerkes, Dodson, et al. 1908), there is an optimal range of cognitive load; if the current cognitive load is lower than the optimal range, increased cognitive workload with higher arousal level can help to improve the performance. The addition of a separate bike lane may have impacted the cognitive workload of the participants to focus on keeping the bicycle in the center of the bike lane. The increase in cognitive workload was lower for a protected bike lane, where bicyclists tend to keep closer to the curbside instead of being in the center of the bike lane.

In the as-built scenario, a shorter fixation duration is observed, which as discussed by previous research is related to a higher hazard estimation of cyclists (Stülpnagel 2020). It is further verified by our post-experiment survey in which most of the participants rate the as-built scenario as the least safe scenario. In terms of the two alternative designs, the separate bike lane scenario seems to have a more focused gaze behavior than the protected bike lane. The phenomenon is also identified by the gaze entropy results. Only the separate bike lane scenario has a significantly lower stationary gaze entropy, which quantifies the overall spatial dispersion of gaze. To our knowledge, the effect of roadway design on the gaze entropy of bicyclists has not previously been examined. An increase in SGE indicates a change in the spatial areas that information is being sampled from, which is illustrated by the less populated and more dispersed depiction of fixation density, as discussed in a driver-related study (B. A. Shiferaw et al. 2018). Meanwhile, for the GTE, no significant differences are found between the three scenarios. The increase in GTE reflects a more random pattern of transitions between fixations. Variation in GTE is related to scene complexity and task demand (B. Shiferaw, Downey, and Crewther 2019). One possible reason for the result may be when designing a separate bike or protected bike lane, participants may feel obligated to maintain a lateral bike lane position (especially in IVE), which offsets the effect of new roadway designs with less scene complexity for necessary visual information retrieval.

5.6.3 Heart Rate Variation

For the HR response, we did not find any significant difference in mean HR between the three scenarios. The correlations between HR/HRV and subjective safety ratings were also found to be weak in previous naturalistic studies (Doorley et al. 2015; Fitch, Sharpnack, and S. L. Handy 2020). In our controlled experiment, the association is even less conclusive. However, when considering the abrupt changes in HR, after extracting the HR change point from the raw data, it is found that both the separate bike lane and protected bike lane scenarios have a significantly lower frequency of HR change points than the as-built scenario. We further explore the spatial distribution of the HR change point and find that the peaks of the HR change point occur more frequently prior to reaching the intersections. Based on our results, in our low-task requirement scenario, the intersection is more associated with a higher number of change points, possibly showing a higher stress level than other sites.

The position of the HR change point depends on the types of intersections. The first intersection has more HR change points than the other two intersections. There are two possible reasons. First, as the first intersection is at the beginning of the experiment, participants may still need some time to get used to the IVE, even if they have practiced before. Second, as indicated by a previous study, spatially open locations increase the level of perceived risk (Stülpnagel 2020), and the road segment after the first intersection is a downhill road. After entering the first intersection, bicyclists will have a wider field of view, which can lead to an increased level of perceived risk. Our study further verified this finding through abrupt changes in HR. Moreover, participants seem to have HR change points earlier in more complex intersections (specifically, the third intersection with traffic lights and stop lines). This can be explained by the fact that more complex intersections require more time to prepare for the crossing. This aligns with another naturalistic study which shows bicyclists' first fixation on the traffic light occurs earlier in a no-bike-lane road (Rupi and Krizek 2019). While not in biking studies, similar results were achieved in other transportation studies with respect to other road users such as drivers' stress levels and emotions when getting closer to the intersections. For instance, recent studies both through subjective measures (Bustos et al. 2021), self-reports (Dittrich 2021), and increases in HR (Tavakoli, Kumar, Guo, et al. 2021; Tavakoli, Boukhechba, and Heydarian 2021) have all shown that drivers experience higher stress level as they arrive at an intersection. However, we note that our findings need to be further verified by future studies due to the limited number and type of intersections in this study. Moreover, the distribution of HR change points for the separate bike lane and protected bike lane scenario is smoother than the as-built scenario. The separate bike lane has more delayed peaks than the other two designs. The reasons behind it should be further explored in future studies.

One important point with respect to HR in our study is the short duration of each scenario as well as the overall experiment. Because the HR was sampled at a relatively lower frequency (1 Hz), the number of data points per scenario becomes significantly smaller as compared to other data sources such as gaze measures (120 Hz). The low frequency might be another reason for the insignificant results in the comparison of mean HR across the three scenarios. In the future work of this study, we are planning to use the raw PPG readings from the watches to enhance the depth of the HR modeling within different scenarios. While HR is sampled at 1 Hz, PPG is sampled at 100 Hz but with the caveat of being affected by the motion artifacts. However, we should note that even with a lower number of data points, a change point detector, when applied to the overall data of a participant, can learn the proper distribution and find the moments of abrupt increases, which are spatially intuitive as well (e.g., being close to intersections).

5.6.4 Demographics and Survey

Although a majority of the 50 participants rate the protected bike lane with pylons scenario as the safest design, the post hoc comparison does not reveal too many differences between the separate bike lane and protected bike lane scenarios. The protected bike lane scenario has a lower average speed. The average lane position is closer to the road curbside and the separate bike lane design has a slightly higher gaze concentration. Other than these findings, there is little evidence showing significant differences between these two alternative designs in terms of cycling behavior and physiological responses. These results indicate there exist some differences between participants' subjective ratings and objective behavioral responses.

No gender or age differences are found to be statistically significant in this study. It is widely accepted that female bicyclists have lower cycling participation rates (Mitra and Nash 2019), as well as stronger preferences than males for greater separation from motor traffic (Aldred et al. 2017). While some studies report minor gender differences (Cobb, Jashami, and Hurwitz 2021), it is argued that female bicyclists using the lanes had significantly more positive associations with the protected lanes than males. In other words, the protected bike lanes can somehow help close the gender gap in cycling (Dill et al. 2014; Aldred et al. 2017). It is worth noting that most of the participants in this study are regular bicyclists, so they are not representative of the entire population. Evidence of stronger preferences among older people is also limited (Aldred et al. 2017). Another notable finding from this study is that the realism of speed is more related to bicyclists' physiological changes than the realism of steering. This can be task-dependent, as in our experiment, most of the road is straight and only a few steering maneuvers are required around the second intersection.

5.7 Limitations and Future Work

The limitations of this study are as follows. First, the duration of the experiment is short, and some findings in this experiment need further study. Building a longer road segment in the IVE with more street blocks can be a solution. However, we should note that a longer duration is more likely to cause motion sickness or fatigue, which can lead to performance degradation. Thus, future work needs to find an optimal duration for the study accounting for the trade-off between avoiding motion sickness and retrieving longer time series of data. Second, although a practice scenario is introduced in the beginning and the order of the scenarios is randomized, a learning effect can exist. This means that the participant might become more familiar with the experiments as they progress through the scenarios, which can affect HR, gaze, and speed. Third, more benchmark studies are needed to verify the findings in real-world environments. We have conducted a pilot study of six participants both in IVE and real road, and preliminary findings show that most of the physiological responses in IVE are representative of the real world (Guo, Robartes, et al. 2021). Nonetheless, a benchmark study in the real world is needed for the same participant groups to validate the findings. Future work will be focused on benchmarking the IVE setup in more diverse scenarios and with a higher number of participants.

While in this study we focused on the design, it can be the case that different events within each design can impact the participant's psycho-physiology. For instance, when considering the effect of a separate bike lane, it could be the case that increases in traffic density, speed of vehicles, and other environmental factors can affect how participant's psycho-physiology changes within each scenario. In an attempt to isolate the effect of the roadway design, these variables had little or no variation in this experimental design. Future work will be focused on simulating more detailed events within each alternative design to better illustrate the interaction between environmental properties (e.g., the presence of bike lane) and events (e.g., vehicles passing at high speed).

While we focused on HR and gaze, we note that human psycho-physiology is a complicated matter which is not bound by only two measures. The addition of wearable devices with more detailed sensors (e.g., skin conductance, skin temperature, and breathing patterns) may provide additional insight on the bicyclists' psycho-physiology.

5.8 Conclusion

This study explores bicyclists' physiological and behavioral changes in different urban roadway designs. In an immersive virtual environment, a bicycle simulator with integrated mobile sensing devices is used in the experiment to record bicyclists' behavioral and physiological responses on the same road with three different roadway designs: shared bike lane (as-built), separate bike lane, and protected bike lane with pylons. Results from 50 participants indicate that (1) the protected bike lane design has the highest subjective safety rating; (2) participants in the protected bike lane scenario have the lowest cycling speed and highest lateral distance to the vehicle lane, indicating the potential for safer bicycling behavior with lower speeds and increased separation from vehicles; (3) bicyclists focus their gaze on the cycling task more in the separate and protected bike lane scenarios, indicating the potential for decreased distractions when cycling in a separate or protected bike lane compared to shared bike lane; (4) creating separate zones for bicyclists (whether separate bike lane or protected bike lane) has the potential to reduce the stress level, as indicated by decreased HR changes compared to the shared bike lane; and (5) the immersive virtual environment can be an efficient and safe tool to evaluate bicyclists' behavioral and physiological responses in different alternative roadway designs.

Chapter 6

The Effect of Cognitive Distraction on Cyclist Behavior

6.1 Introduction

Bicycle users are expanding as an increasing number of cities in the United States are encouraging Low-carbon transportation by investing in infrastructure to accommodate bicyclists (Pucher, Dill, and S. Handy 2010). This increasing trend hasn't been slowed down during the COVID-19 pandemic, it is reported that bicycling levels have significantly increased in many countries even with lockdowns and travel restrictions (Buehler and Pucher 2021).

However, bicyclist fatalities are increasing as more bicyclists are using the road over the last decade. The report from the National Highway Traffic Safety Administration shows that in the United States, the number of bicyclist fatalities has increased by more than 35% since 2010 (NHTSA 2021a).

Distraction has been identified as one of the main reasons for traffic accidents. In the US, nine percent of fatal crashes, 15 percent of injury crashes, and 15 percent of all police-reported motor vehicle traffic crashes in 2019 were reported as distraction-affected crashes. There were 566 vulnerable road users including pedestrians, pedal cyclists, and others killed in distraction-affected crashes (NHTSA 2021b). The data for the fatalities and injuries are underestimated as only vehicle-related crashes are reported. For now, very limited information about cyclists' distraction on fatality is available. Not only the limited data sources but also the number of studies that have been published on distracted biking is small. As a result, our knowledge about the effect of distraction on cycling is insufficient. Previous studies have reported that distractions have a major prevalence among bike users and that they play a significant role in the prediction of the traffic crash rates of cyclists, through the mediation of risky behaviors (Useche et al. 2018). Studying cyclists behaviors under the influence of distraction can provide evidence for interventions to address safety-related issues. Similar to the diving distraction, the cycling distraction can be categorized into three main types: Visual (taking the eyes off the road), Manual(taking the hands off the handlebar), and Cognitive (taking the mind off cycling). This study will focus on cognitive distraction as it is related to the most frequently reported secondary task during cycling, such as listening to music or talking in the earphones (Mwakalonge, White, and Siuhi 2014; Wolfe et al. 2016).

The current state of knowledge on cyclist distraction is mostly retrieved from surveys or observational studies. For example, an observational study in New York City shows that headphone use is the most prevalent distraction among local cyclists (Ethan et al. 2016). However, observational studies are unable to track cyclists' physiological changes and get the details of secondary tasks (e.g., headphone use can be either music listening or talking on the phone). Surveys from different areas around the world have been collected to study cyclists' distracted behavior, listening to music or talking with earphones have been identified as the most prevalent distractions (Terzano 2013; Wolfe et al. 2016; Young et al. 2020).

The most frequent secondary tasks, both listening to music and talking with earphones can be categorized as a cognitive distraction. One of the main challenges in the quantitative analysis of cognitive distraction is the difficulty in measuring the workload needed for certain tasks. To understand the mechanism of distraction, a standardized secondary task with different levels of workload is required in the experimental study. To our knowledge, no prior studies have applied such methods for cyclist distraction. In other research fields, several standardized secondary tasks have been developed to simulate different levels of workload. For instance, to simulate the phone conversation, an alternative mock cellphone task was used in a driving-related study as a cognitive distraction (Ebadi et al. 2020). The mock cellphone task was designed to simulate cognitive load when talking on the phone, and the impact of this type of task was reported to be similar to a hands-free cellphone conversation in a prior study (Muttart et al. 2007).

Physiological responses such as physical measures (e.g., eye movement) and biological measures (e.g., EDA) are reported to be effective measurements to detect distractions in driving, however, these conclusions haven't been verified in cyclist studies (Yusoff et al. 2017).

The goal of this experiment is to study the effect of cognitive distraction on cyclist behavior, especially, we are interested in applying the standardized secondary task in the IVE bicycle simulator to simulate different levels of cognitive workload, and explore cyclists' physiological responses in different situations. The research hypothesis is 1. Listening to high-tempo music results in a higher biking speed, gaze/head movement, and a higher rate of SCR peaks. 2. Talking on the phone result in lower speed, lower gaze/head movement variation, and lower rate of SCR peaks. 3. Standardized secondary tasks can be used to simulate different cognitive distractions in the IVE.

6.2 Methodology

6.2.1 Experiment Design

This research studies the effect of cognitive distraction on cyclist behavior in the proposed IVE bicycle simulator framework. The cognitive distraction will be triggered by both the standardized secondary task (Mock phone conversation task) and the actual task (music listening). Each participant will experience 3 different conditions (baseline, music listening, and mock phone conversation) in random order.

As an alternative to cellphone conversations, a mock cellphone task was used in this study as a cognitive distraction, particularly because typical conversations are much more difficult to experimentally control. While performing the distraction task, participants were instructed to listen to a series of generic English language sentences synonymous with the previously validated grammatical reasoning task and respond aloud to the subject, object, and whether the sentence was plausible or not. The experimenter would remotely initiate the task with a button press at the beginning of each scenario and similarly, terminate it a few seconds before the end of every scenario. The participants listening to each sentence were then asked to reply aloud: the subject of the sentence, the object of the sentence, and whether or not the sentence was plausible. For example, for the sentence, "A child jumped a rope," the correct answer is: "Child, Rope, and Yes." Similarly, an implausible sentence would be, "A cat baked the television," and the correct answer would be: "Cat, Television, and No." The mock cellphone task was designed in such a way that the driver would be cognitively loaded, and as found in prior studies, the impact of this type of task would be similar to that of a hands-free cellphone conversation (Ebadi et al. 2020).

In addition to the standardized secondary tasks, an actual secondary task, music listening is considered in this experiment as well. In the music listening condition, the participants will be asked to listen to a popular song of the year. The song will be played during cycling automatically by the experimenter, the participants can hear the song from the earphone of the VR headset, similar to listening to music with earphones in the real world.

6.2.2 Road Environment in IVE

The IVE for this study is developed from the IVE of our last study in Chapter 5. We further develop the IVE with the real-world measurement of the Water Street corridor in the city of Charlottesville, Virginia. The IVE road environment has been extended to 8 street blocks as shown in Figure 6.1, the IVE road starts from the intersection of West Main Street and Ridge Street, and ends at the intersection of East Water Street and 9th Street NE (at the Belmont Bridge). Bike lanes are designed for the road with a standard bike lane width of 4 feet (1.2m).

6.2.3 Experiment Procedure

A similar experiment procedure will be followed as the study in Chapter 5. Upon arrival, each participant is asked to sign the consent form approved by the IRB office (Appendix A) and put on two smartwatches on both wrists, before completing the pre-experiment survey.

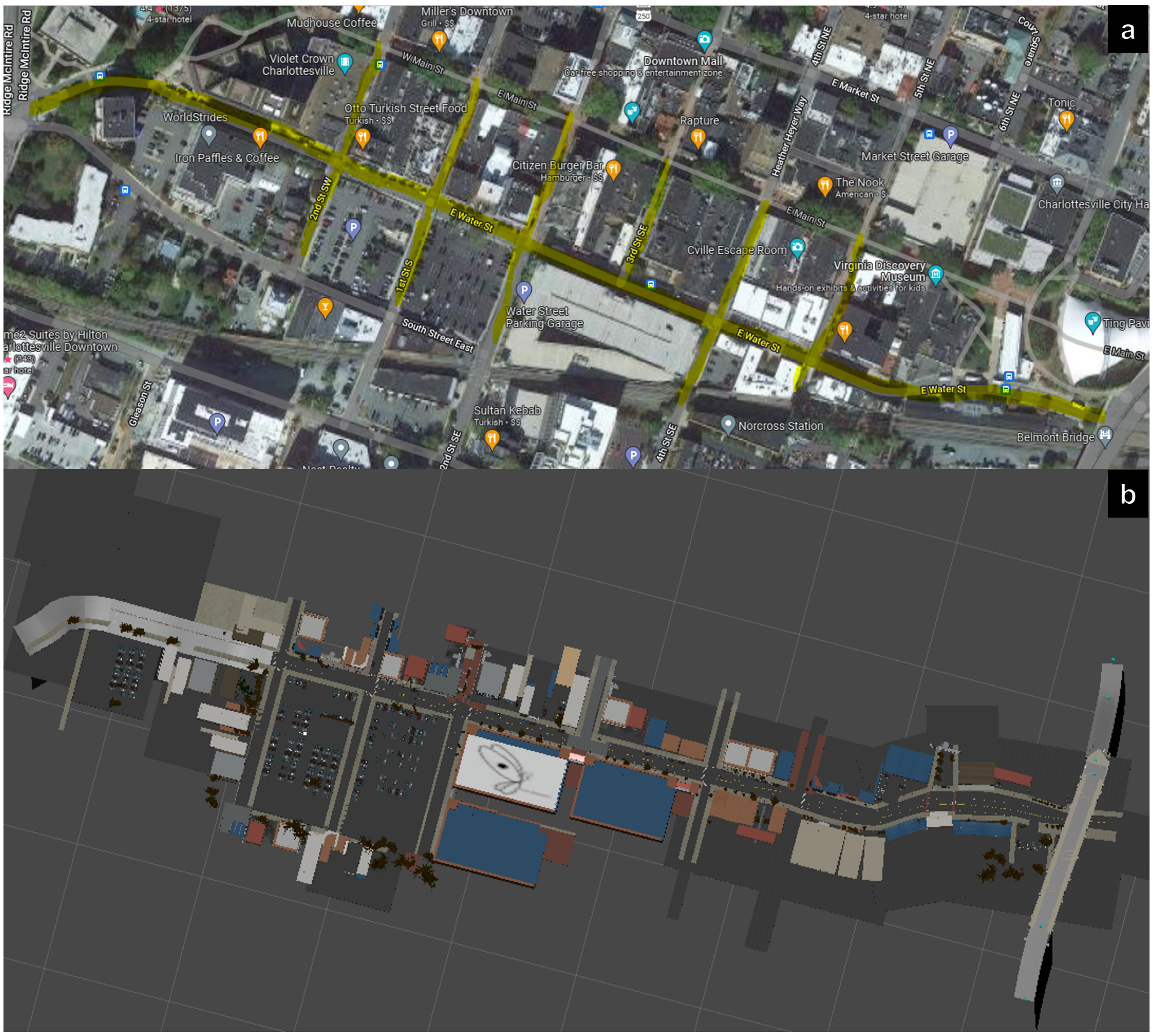


Figure 6.1: Illustration of extended IVE road environment. (a) Bird view from the map of real road. The built road segments are highlighted in yellow. (b) Bird view of the modeled streets within the IVE in Unity software

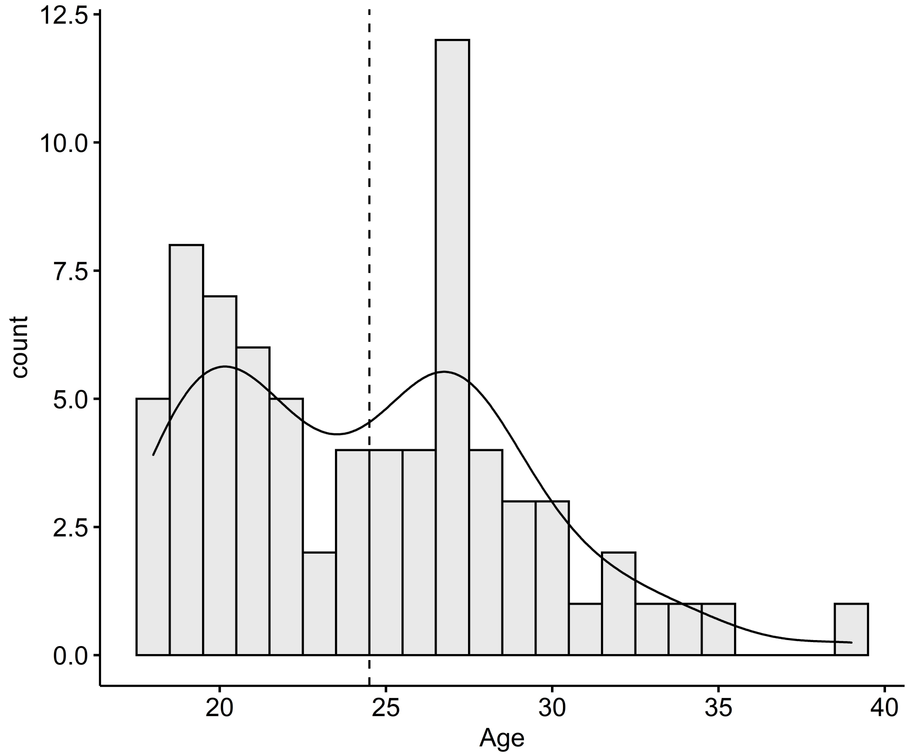


Figure 6.2: Age distribution of 75 participants. The mean age is 24.5 years old, cyclists who are younger than this age are categorized as a younger group

6.2.4 Participants

75 participants were recruited for the experiment. Among them 40 are female, 33 are male, 1 participant identified as other and 1 participant didn't provide gender information. Most of the participants are local bicyclists, students, and faculty members from the University of Virginia. All participants are 18 or older and without color blindness. The mean age is 24.5 with a standard deviation of 4.7, and the median age is 24.5 years old as well (one participant didn't provide the age information). The age distribution is shown in Figure 6.2.

6.3 Results

This section reports the results of the experiment. The following subsections describe the summary statistics (from the pre-and post-experiment surveys), the bicyclists' physical behavior (cycling speed, input power, and lateral lane position), and physiological responses in different roadway designs.

6.3.1 Survey Response

Pre-experiment Survey

Participants' attitude towards cycling is collected in the pre-experiment survey. 74 participants answered the question about what types of bicyclists they are. The majority of the participants have a positive attitude towards cycling, as shown in Figure 6.3, 6 participants (8.11%) indicated their attitude towards cycling as "No way, no how" - I do not ride a bike, 25 participants (37.84%) identified themselves as "Interested but Concerned" - I like the idea of riding but have concerns. The rest of the participants had a higher preference for cycling as 28 (33.78%) indicated themselves as "Enthused and Confident" -I like to ride and will do so with dedicated infrastructure and the remaining 15 (20.27%) chose "Strong and Fearless" - I will ride anywhere, no matter the facilities provided.

The engagement of secondary tasks both in daily life and during cycling is also collected in the pre-experiment survey. We asked the participants to estimate how many hours they spend in music listening and phone usage, as well as the frequency of music listening, and phone talking when they ride a bike. The distribution of daily hours spend on music listening and phone are displayed in Figure 6.4. The average music

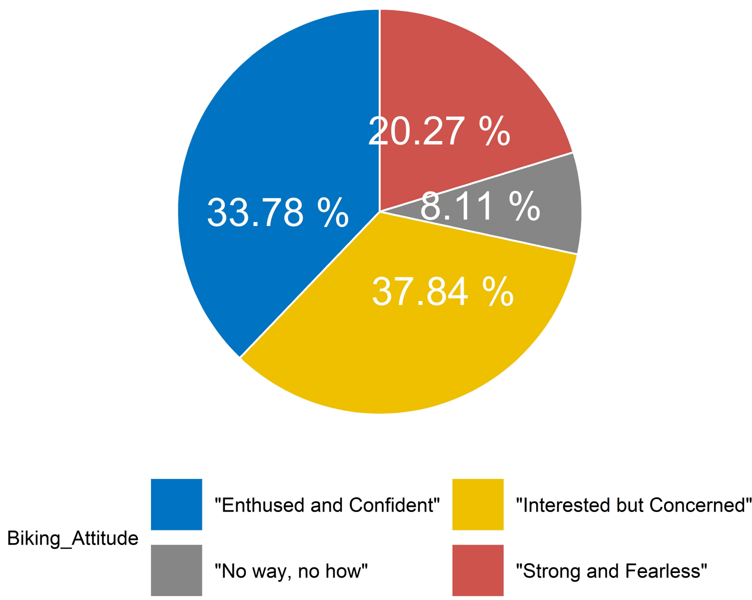


Figure 6.3: Participants' attitude towards cycling

listening hours are 2.82 hours (sd = 2.88 hours). And unsurprisingly, the average hours spent on the phone is higher (mean = 4.30 hours, sd = 2.07 hours). Participants' music listening and phone talking frequency while riding a bike is self-reported with 5-point Likert scales, with 5 options of "Never (<10%)", "Seldom (about 25%)", "Sometimes (about 50%)", "Often (about 75%)", and "Always (>90%)". Participants were asked to choose an option that is closest to them. Table 6.1 summarizes the results for these two questions. The participants have a higher frequency of music listening than talking on the phone while biking. More than half of the participants admitted that they have music-listening behavior while biking, and only about 25% of the participants reported that they had the experience talking on the phone while biking before.

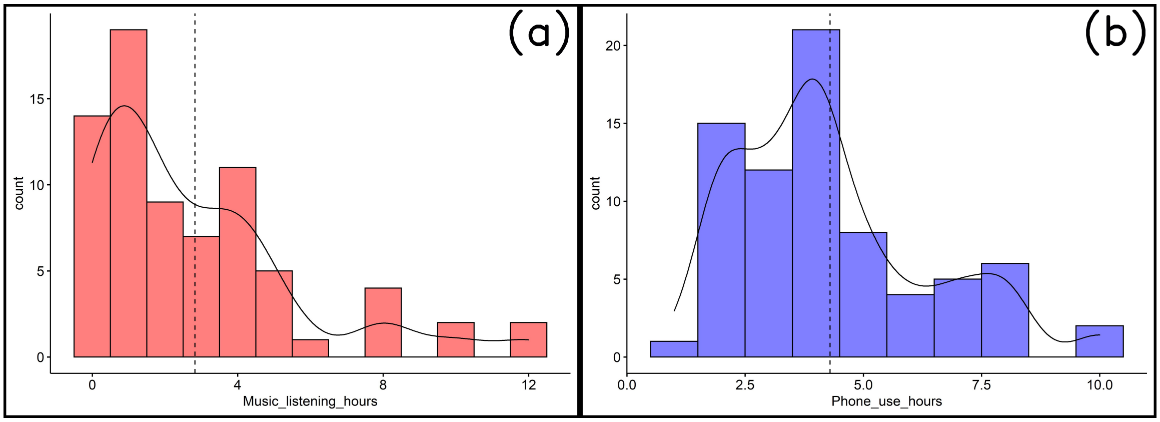


Figure 6.4: Participants' daily hours spend in music listening (a) and phone usage (b)

Table 6.1: Summary of participants' music listening and phone talking frequency while riding a bike

Secondary Task Type	Frequency	Number	Percentage (%)
Music Listening while Biking	Never (<10%)	36	48
	Seldom (about 25%)	9	12
	Sometimes (about 50%)	7	9.3
	Often (about 75%)	15	20
	Always (>90%)	7	9.3
	NA	1	1.3
Talking on the Phone while Biking	Never (<10%)	55	73.3
	Seldom (about 25%)	18	24
	Sometimes (about 50%)	1	1.3
	Often (about 75%)	0	0
	Always (>90%)	0	0
	NA	1	1.3

Post-experiment Survey

In the post-experiment survey, participants' stated preferences over the three scenarios are collected in three aspects: safety, comfort and distraction. For each question, the answer is to choose their subjective ratings from a 5-point Likert scale.

For subjective safety rating, the Baseline scenario is rated as the safest scenario with an average score of 4.31/5.0, followed by the Music Listening (3.93/5.0), then the

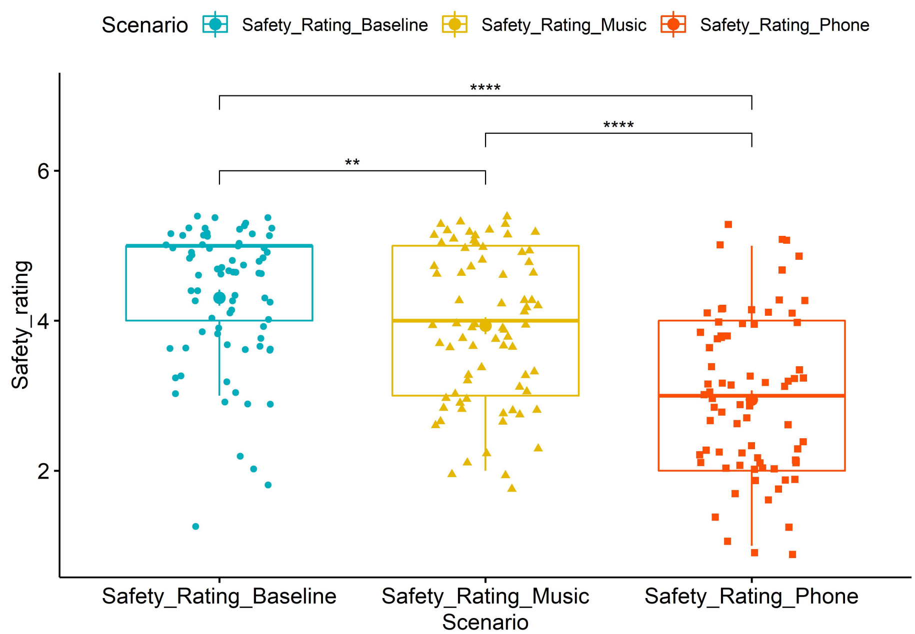


Figure 6.5: Subjective safety rating of different scenarios (jitter points to avoid over-plotting)

Mock phone conversation scenario (2.95/5.0), the differences between all the three scenarios are significant, as shown in Figure 6.5 (Baseline v.s. Music listening, $p = 0.00297$; Baseline v.s. Mock phone conversation, $p < 0.0001$; Music listening v.s. Mock phone conversation, $p < 0.0001$).

For subjective comfort rating, the scores of the Baseline (4.35/5.0) and Music listening (4.32/5.0) scenarios are close to each other, and both are significantly higher than the Mock phone conversation scenario (2.88/5.0), with both p values smaller than 0.0001, as shown in Figure 6.6.

For subjective distraction rating, the Mock phone conversation scenario is rated as the most distracting scenario with an average score of 3.74/5.0, followed by the Music Listening (2.42/5.0), then the Mock phone conversation scenario (1.64/5.0), the differences between all the three scenarios are significant, as shown in Figure 6.7 (all

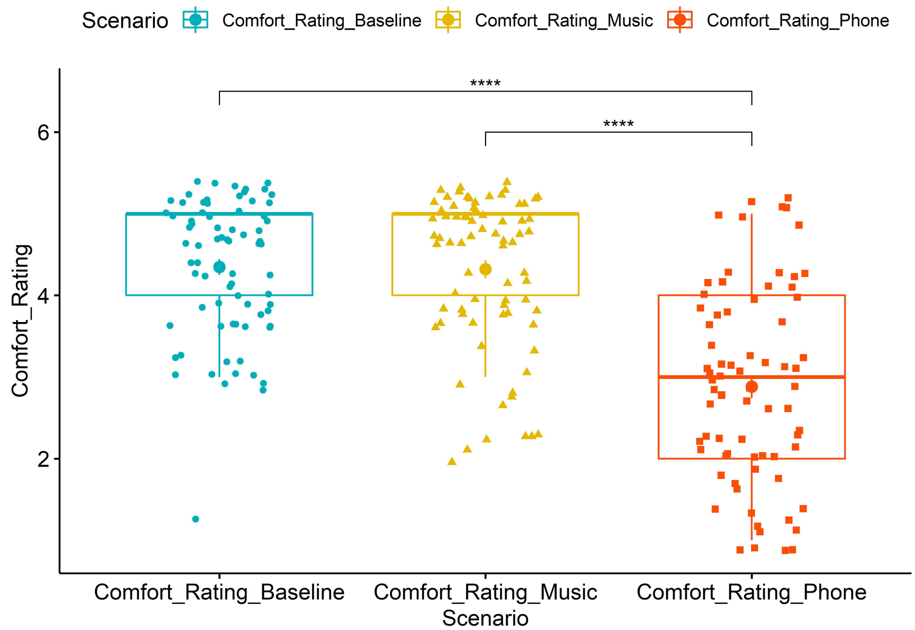


Figure 6.6: Subjective comfort rating of different scenarios (jitter points to avoid overplotting)

the p values are smaller than 0.0001).

6.3.2 Cycling Performance

Speed

For the mean speed, as indicated in Figure 6.8-a, there is a significant difference between the Baseline and the Mock phone conversation scenarios ($\beta = -1.262, SE = 0.435, p = 0.0428$) and between the Music listening and the Mock phone conversation scenarios ($\beta = 1.178, SE = 0.312, p = 0.0007$). Bicyclists' mean speed in the three scenarios (Baseline, Music listening, and Mock phone conversation) are 18.6 km/h, 19.1 km/h, and 17.9km/h, respectively. Age group differences are found only in the Mock phone conversation scenario, where the younger group (19.3 km/h) has a

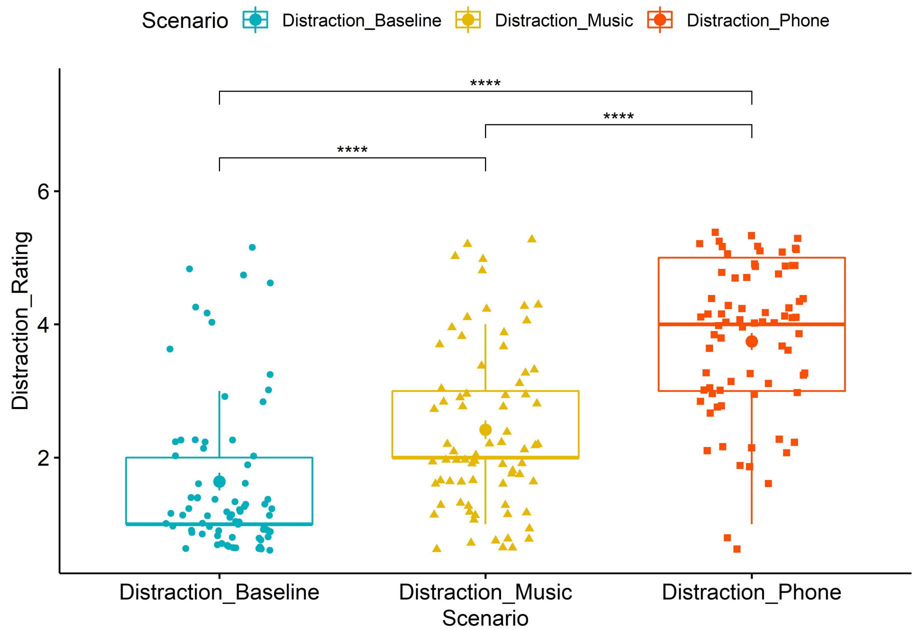


Figure 6.7: Subjective distraction rating of different scenarios (jitter points to avoid overplotting)

significantly higher cycling speed ($\beta = 2.903, SE = 1.058, p = 0.00783$) than the older group (16.7 km/h), as indicated in Figure 6.8-b.

For the standard deviation of speed, the results show there is a significant difference between the Baseline and the Music listening scenarios ($\beta = -0.30669, SE = 0.14391, p = 0.0348$), as shown in Figure 6.9-a. Bicyclists' standard deviation of speed in the three scenarios (Baseline, Music listening, and Mock phone conversation) are 1.92 km/h, 1.74 km/h, and 1.83 km/h, respectively. For the Music listening scenario, it is found that participants who listen to music a lot (>4 hours daily) have a lower standard deviation of speed ($\beta = -0.572, SE = 0.226, p = 0.0142$), as shown in Figure 6.9-b.

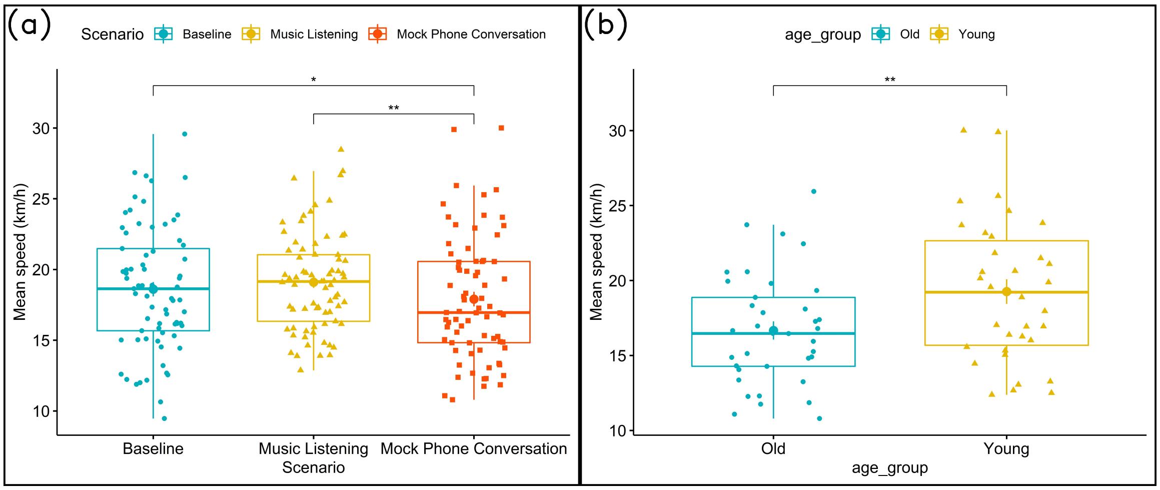


Figure 6.8: Mean speed of (a) different scenarios, (b) different age groups in the Mock phone conversation scenario

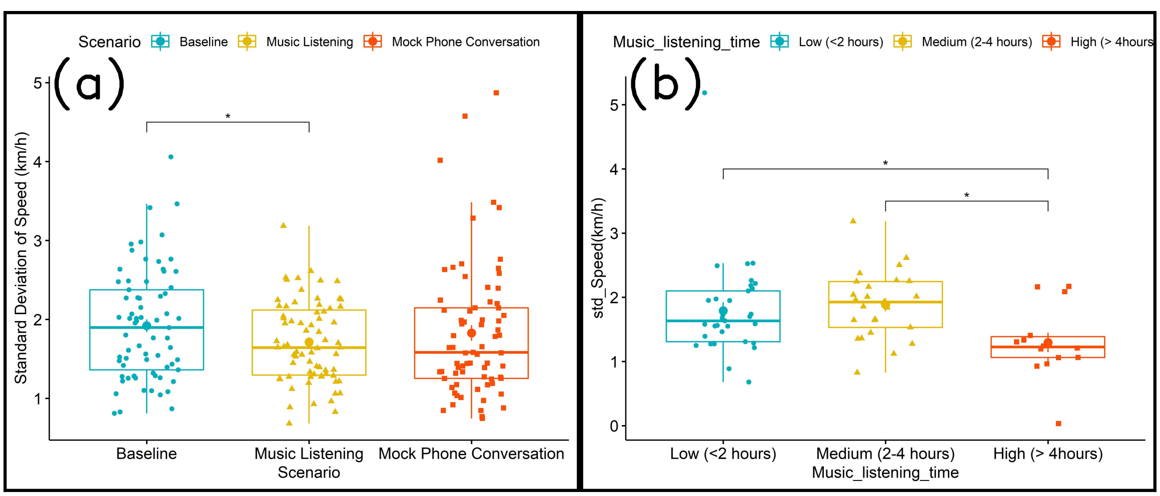


Figure 6.9: Standard deviation of the speed of (a) different scenarios, (b) participants with different music listening time in the Music listening scenario

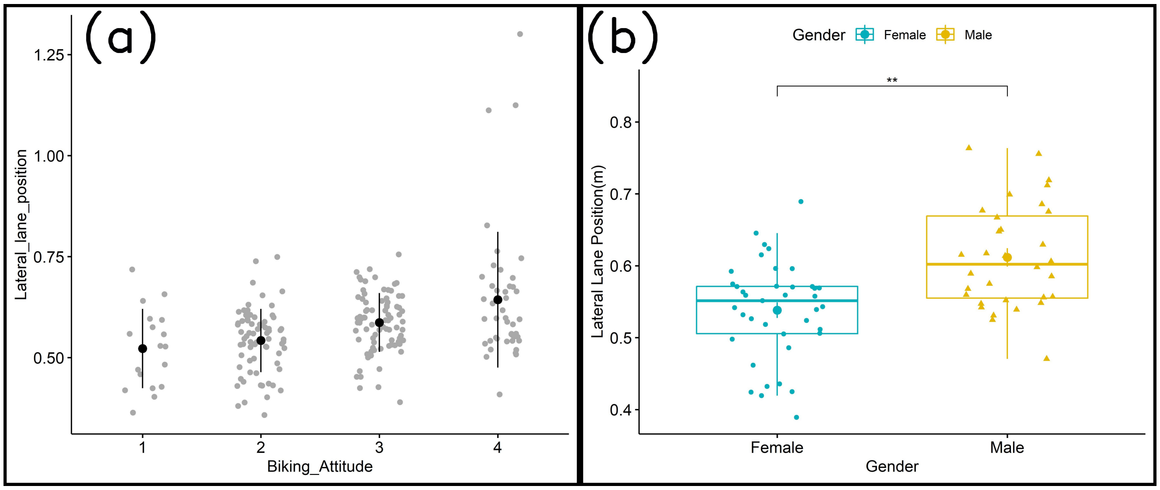


Figure 6.10: Mean lateral lane position of (a) different attitudes toward cycling, a higher score means a more positive attitude towards cycling. 1-4 indicates 'No way, no how', 'Interested but Concerned', 'Enthused and Confident', and 'Strong and Fearless', respectively. (b) different genders in the Music listening scenario

Lateral Lane Position

For the lateral lane position, no significant differences are found between different scenarios. The average lateral lane position for the three scenarios is Baseline - 0.575m, Music listening - 0.569m, and Mock phone conversation - 0.592m. However, significant differences are found between the participants with different attitudes toward cycling. As can be seen from Figure 6.10-a, participants who hold more positive attitudes toward cycling will go more on the left (closer to the vehicle lane).

Input Power

For the mean input power, the average input power in music listening (50.6 Wattage) is 16% higher than mock phone conversation (43.8 Wattage), and 7.5% higher than baseline (47.1 Wattage). There is a significant difference between the Baseline and the Mock phone conversation scenarios ($\beta = -6.658, SE = 2.443, p = 0.00725$) and be-

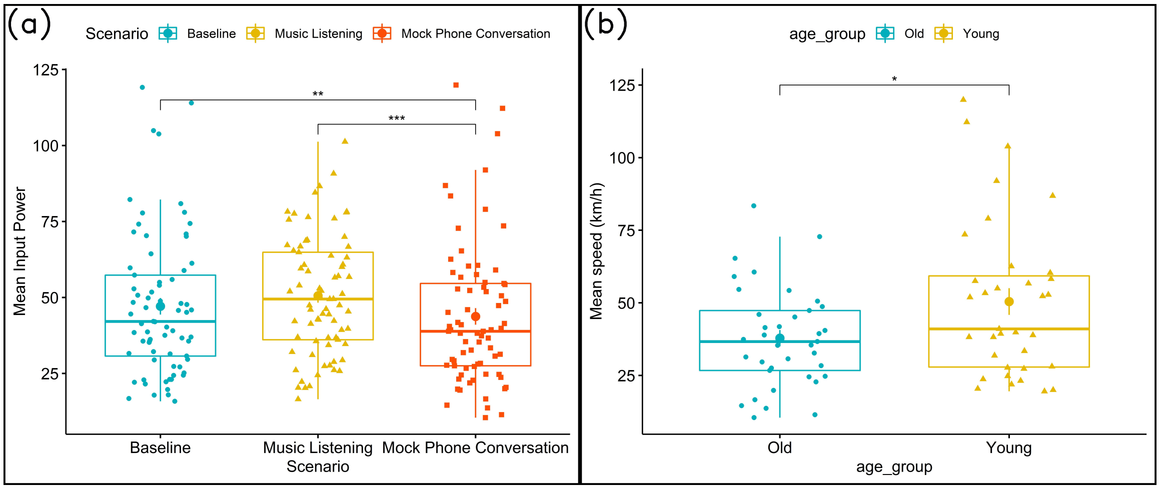


Figure 6.11: Mean input power of (a) different scenarios, (b) different age groups in the Mock phone conversation scenario

tween the Music listening and the Mock phone conversation scenarios ($\beta = 6.72, SE = 1.75, p = 0.0006$), as shown in Figure 6.11-a. Age group differences are found only in the Mock phone conversation scenario, where the younger group (50.4 Wattage) has a significantly higher cycling input power ($\beta = 11.864, SE = 5.712, p = 0.0418$) than the older group (37.8 Wattage), as indicated in Figure 6.11-b.

Head Movement

The head movement result shows a significant difference between the Baseline and the Mock phone conversation scenario with $\beta = -0.00636, SE = 0.00220, p = 0.00435$, as shown on Figure 6.12-a, participants had a lower variation of head movement direction in the Mock phone conversation scenario than the Baseline. Additionally, in the music listening scenario, male participants have a higher head movement variation than female participants with $\beta = 0.0122, SE = 0.00538, p = 0.0266$ (Figure 6.12-b).

Figure 6.12: Head movement variation (a) different scenarios; (b) gender differences in music listening scenario

6.3.3 Physiological Response

Heart Rate

The mean heart rate result indicates that there are no significant differences between the three scenarios a 95% confidence level. The mean HR (beat per minute) of the Baseline, Music listening, and Mock phone conversation are 92.89, 92.07, and 90.66, respectively. No other factors are found to significantly affect the mean HR.

When calculating the frequency of increasing HR change points, the result in Figure 5.12 illustrates the average frequency of HR change points in different scenarios. The LMM model shows that both the Baseline ($\beta = 0.653, SE = 0.226, p = 0.0131$) and the Music listening scenarios ($\beta = -0.546, SE = 0.226, p = 0.0171$) have a significantly higher frequency of increasing HR change points than the Mock phone conversation scenario. Bicyclists' average frequency of increasing HR change points in the three scenarios (Baseline, Music listening, and Mock phone conversation) are 1.98 counts/min, 2.09 counts/min, and 1.44 counts/min, respectively (Figure 5.12-a).

Figure 6.13: Frequency of increasing HR change points (a) different scenarios; (b) gender differences

Male participants have a significantly lower frequency of increasing HR change points than female participants (1.54 vs. 2.10), with $\beta = -0.612, SE = 0.291, p = 0.0403$, as shown in Figure 6.13-b.

The distribution of HR change points is illustrated in Figure 6.14. With this figure, it is easier to find which road segment causes more HR change points. The peak is in the big curve after intersection 6. For other flat or uphill road segments, there will be more HR change points as the bicyclists are approaching an intersection; While for the downhill road segment, there will be fewer HR change points as the bicyclists are approaching an intersection.

Skin Temperature

The mean skin temperature (°C) for the Baseline, Music listening, and Mock phone conversation scenarios are 32.63°C, 32.68°C, and 32.71°C, although the Mock phone conversation scenario has a slightly higher, we didn't find any significant differences between the three scenarios a 95% confidence level (Figure 6.15-a). The age factor

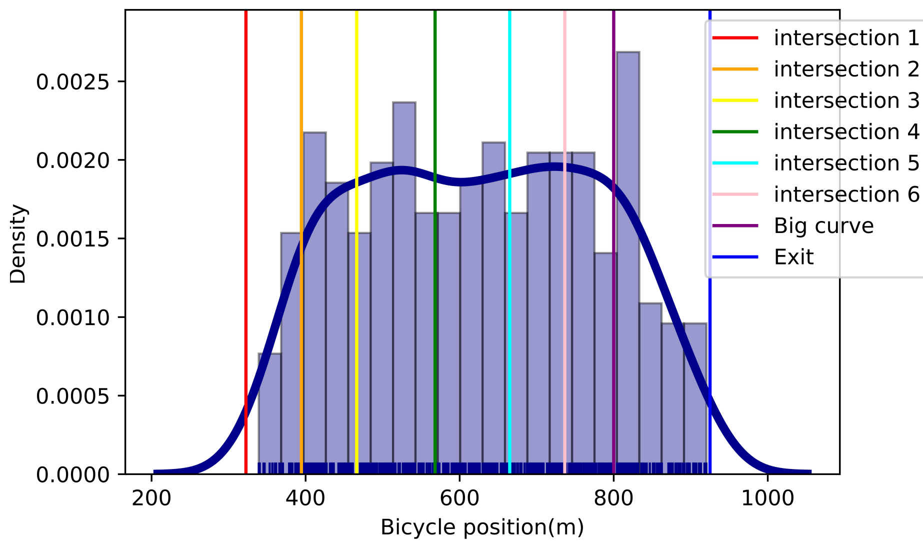


Figure 6.14: Spatial distribution of HR change points

is found to be marginal significant ($\beta = 0.682, SE = 0.347, p = 0.0544$), as shown in Figure 6.15-b. No other factors are found to significantly affect the mean Skin Temperature.

EDA

The numbers of SCR peaks for the Baseline, Music listening, and Mock phone conversation scenarios are 28.0, 29.3, and 25.8. Although the Mock phone conversation scenario has a slightly lower number of SCR peaks, we didn't find any significant differences between the three scenarios with a 95% confidence level (Figure 6.16-a). No other factors are found to significantly affect the number of SCR peaks.

The mean amplitude of SCR peaks for the Baseline, Music listening and Mock phone conversation scenarios are 0.151 μS, 0.171μS, and 0.205 μS (Figure 6.16-b). The Mock phone conversation scenario appears to have a sightly higher number of SCR

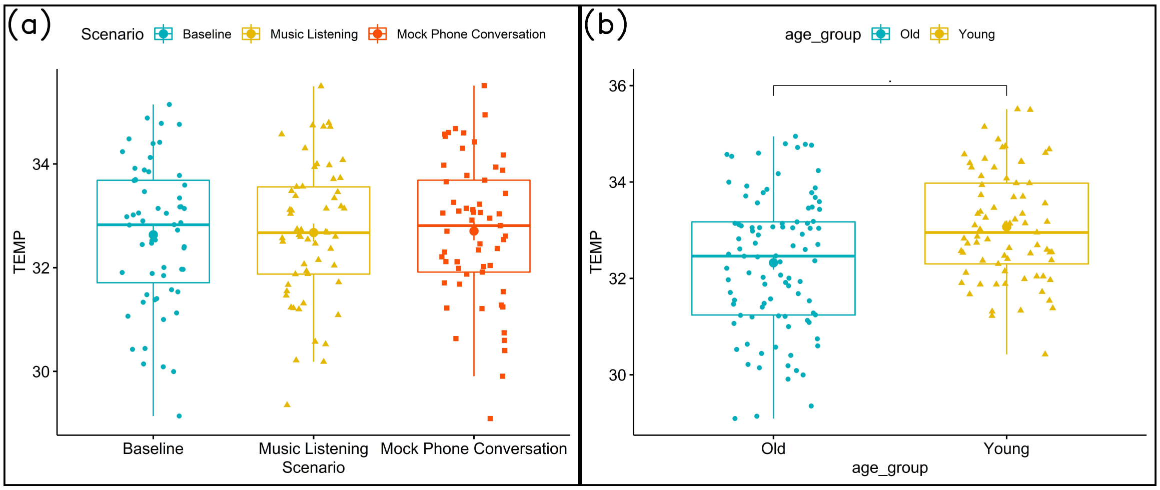


Figure 6.15: Skin Temperature of (a) different scenarios; (b) different age groups

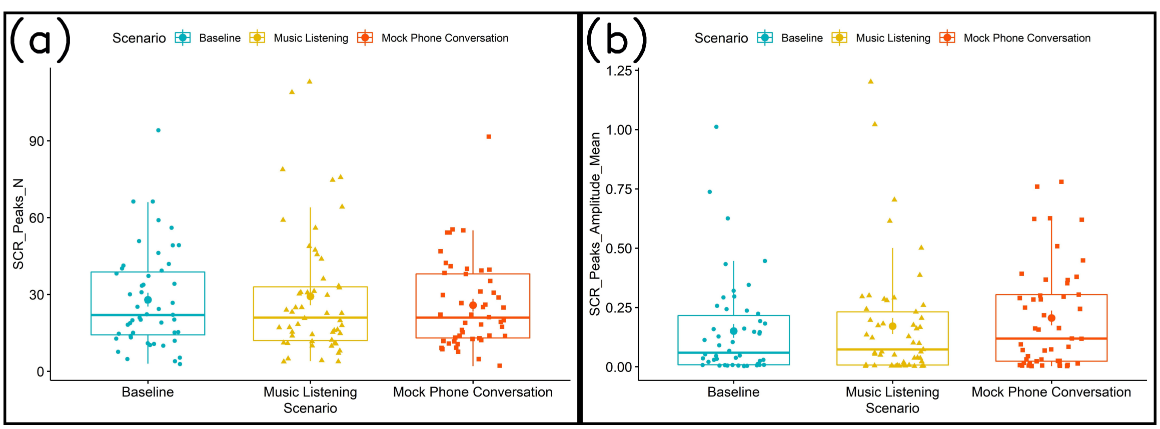


Figure 6.16: EDA data results (a) number of SCR peaks different scenarios; (b) mean amplitude of SCR peaks

peaks than the Baseline, but the differences are not significant ($\beta = 0.05279, SE = 0.03218, p = 0.1042$). In the Mock phone conversation scenario, the gender factor is found to be significant (male = 0.125 μS, female = 0.267 μS, $\beta = -0.152, SE = 0.0717, p = 0.0394$). No other significant differences are found in a 95% confidence level.

6.4 Discussion

We measured three subjective ratings from the post-experiment survey: safety, comfort, and distraction. For the two types of secondary tasks, not surprisingly, the Mock phone conversation is rated as the most distracting scenario, as it requires both listening to the audio (input) and speaking out the response (output). And for music listening, the cyclists only need to listen to the audio (input). The safety rating is correlated with the distraction rating, lower distraction rating scenarios have higher safety ratings. In terms of comfort rating, no significant results are found between the Baseline and Music listening scenarios, both scenarios have a higher comfort rating than the Mock phone conversation scenario.

Different levels of cognitive distraction have different effects on cycling behavior and physiological response. For cycling performance, the Music listening scenario has a significantly higher average speed and input power than the Mock phone conversation, as cyclists have a lower subjective rating on the distraction of music listening scenario, they are more confident to keep a higher speed in the IVE with more input power, although the safety rating of music listening is lower than the Baseline.

In a previous virtual reality-based distracted cycling study, it was found that those in a low perceptual load (visual distraction) VR cycled at a higher intensity despite greater pain (Wender et al. 2022). An earlier real road study also reported that telephoning coincided with reduced speed, reduced peripheral vision performance, and increased risk and mental effort ratings (De Waard et al. 2010). That study created different levels of perceptual load by displaying different items in the VR for the item detection task. In our study, we generated different levels of cognitive distraction, and with low cognitive load (music listening), we observe a similar effect as a visual distraction. With high cognitive load (mock phone conversation), the adaptive cycling performance includes lowering speed, less input power, and less head movement, indicating a degraded perception ability of the surrounding environment, which is aligned with previous research findings with drivers made fewer saccades, spent more time looking centrally and spent less time looking to the right periphery (Harbluk, Noy, and Eizenman 2002).

However, in terms of lateral lane position, our findings of cycling performance under the influence of cognitive distraction is different from driving. With the findings from the last experiment in Chapter 5, we design bike lanes for the whole road in this experiment, which is different from the real road of the shared bike lane with vehicles. The introduction of bike lanes in the last experiment was found to help the bicyclist to keep closer to the road curbside. In this study, a similar effect is found as there are no significant differences between different scenarios in lateral lane position. While in driving-related studies, cognitive load led to a diminished standard deviation of lateral position, implying a better lane-keeping performance. However, a systematic comparison of time-to-line crossing calculations suggested a degraded safety margin of lane keeping (P. Li et al. 2018).

Music listening has been found to be related to emotional arousal, which has the potential to affect cycling performance. For example, listening to preferred music showed no ergogenic benefit during repeated anaerobic cycling sprints when compared to non-preferred music. However, preferred music increased motivation to exercise and decreased perceived exertion (Ballmann et al. 2019). The cycling task in this experiment is low intensity, listening to preferred music, as indicated by the survey results with higher familiarity and preference of the song played in the Music listening scenario, which may help to explain the increased speed and input power. Cyclists' engagement in the music also leads to a decreased standard deviation of speed, therefore, they will keep a high speed while avoiding any additional speed changes during the whole cycling process.

For the physiological response, the HR results are more related to the speed results. No significant results are found in the mean HR, but the HR change points results showed that there are fewer heart rate change points in the Mock phone conversation scenario than in the Baseline and Music listening scenarios. The EDA data has no significant difference across different scenarios, but the trend of SCR is worth-noting. The Mock phone conversation scenario has a lower number of SCR peaks but a higher mean amplitude of SCR peaks. This may indicate that with a higher level of cognitive load, cyclists will have a lower frequency but higher intensity in their skin conductance response. This result requires further validation in the future.

Demographic differences are found in several aspects. Generally speaking, participants who hold more positive attitudes toward cycling will go more on the left (closer to the vehicle lane), as they may be more confident about their ability to control the bike. The HR change points data reveals the gender difference as male participants have a significantly lower frequency of increasing HR than female participants. The differences vary in different levels of cognitive load. In the Music listening scenario, the younger group has a significantly higher cycling speed and input power as they are more used to the selected music. In the Mock phone conversation scenario, participants who listen to music a lot (>4 hours daily) have a lower standard deviation of speed, indicating they are more engaged in the music. The physiological response also shows that female cyclists are more affected by the cognitive load in the Mock phone conversation scenario, as (1) male participants have a higher head movement variation than female participants, and (2) Female participants have a higher mean amplitude of SCR peaks. These results highlight the groups of people that require more attention when studying cyclist cognitive distraction (young people who listen to music a lot in their daily life and female cyclists under the effect of higher cognitive distraction such as talking on the phone).

6.5 Conclusion

This research explores the effect of different cognitive distractions on bicyclists' physiological and behavioral changes. In an immersive virtual environment, a bicycle simulator with multiple physiological sensing devices is utilized to collect bicyclists' behavioral and physiological responses on the same road design with bike lanes. Data collection includes demographic information (age, gender, biking attitude), engagement in a secondary task such as music listening and talking on the phone during their daily life or cycling, cycling performance in the simulator (speed, lane position, input power, head movement) and physiological responses (heart rate, skin temperature). Results from 75 participants who rode on a bicycle simulator through a virtual environment indicate that (1) Cyclists would have a significantly higher speed, a lower standard deviation of speed, and higher input power in the music listening scenario. (2) When talking on the phone, cyclists will try a lower speed with less input power and less head movement variation. (3) When listening to music, cyclists who had a strong habit of daily music listening (> 4 hours/day) had a higher engagement in the music, with a significantly lower sd of speed. Male cyclists stayed closer to the vehicle lane and had a higher head movement variation. (4) Lane position is not affected by the scenario, this may be the effect of introducing bike lanes in the environment.

Chapter 7

Conclusion and Future Directions

In this dissertation, we address the lack of pre-accident data issues for vulnerable road users, especially for cyclists, who are facing increasing risks from the road environment nowadays. To understand cyclists' safety, behavior, and comfort levels under different design contexts, a low-cost, risk-free, more controllable solution with the ability of multimodal data collection is needed for the research community. This dissertation develops an Immersive Virtual Environment (IVE) bicycle simulator and tests it with three different experiments for validation in different situations. The key conclusions and the answers to research objectives (RO) and research questions (RQ) at the end of Chapter 1 are listed below:

We start the work with the development of Omni-Reality and Cognition Lab Simulator (ORCLSim) in Chapter 3. Previous research highlights the advantages of using an Immersive Virtual Environment (IVE) in conducting bicyclist and pedestrian studies. These environments do not put participants at risk of injury, are low-cost compared to on-road or naturalistic studies, and allow researchers to fully control variables of interest. In this study, we propose ORCLSim to support human sensing techniques within IVE to evaluate bicyclist and pedestrian physiological and behavioural changes in different contextual settings. To showcase this framework, we present two case studies where pilot data from five participants' physiological and behavioral responses in an IVE are collected and analyzed, representing real-world roadway segments and traffic conditions. Results from these case studies indicate that physiological data is sensitive to road environment changes and real-time events in the IVE, especially changes in heart rate and gaze behavior. By analyzing these changes, future studies can identify how participants' stress level and cognitive load is impacted by the surrounding environment. The ORCLSim system architecture is a prototype that can be customized for future studies in understanding users' behavioral and physiological responses in virtual reality settings. with this study, we are able to answer RQ 1 of RO 1: *What set of information can we capture in the Immersive Virtual Environment (IVE) bike simulator?* We are able to collect cycling performance (speed, steering, braking, acceleration, and lane position), eye Tracking (gaze direction, fixation), physiological responses (Heart rate, head movement, hand acceleration), video recording and stated preference surveys (subjective rating).

We then run a benchmarking study in Chapter 4 to answer RQ 2 of RO1: *Will the cyclists' cycling/Heart Rate/eye tracking behavior in the IVE be representative of the real world?* Immersive virtual environments (IVE) have been shown to provide a realistic representation of real-world conditions, allowing researchers, designers, and engineers to study the impact of design features on end users. However, these tools have not been evaluated and validated for vulnerable road users, such as cyclists. The purpose of this study is to assess the use of an IVE bike simulator to study the impact of design and environmental conditions on cyclists' perceived safety and behavioral changes. By benchmarking cyclists' behaviors and perceived safety in real-life settings compared to their representative IVE bike simulations, we can validate whether these IVE simulators are realistic representations of real-world conditions. A pilot study was conducted both in IVE and on the real road. Various sensors are applied to ensure that similar data output is obtained. Both absolute and relative validity is established across a range of cyclist performances. Results show that most of the performance measurements have absolute validity, but some of the features from eye tracking, most of which are in the vertical direction, could not establish either absolute or relative validity. This phenomenon may be caused by road geometry changes, the appearance of other road users, and hardware limitations (especially the headsets used in the study). Overall, the promising results indicate that the IVE bike simulator can be further utilized for understanding cyclists' behaviors.

Next, we run two experimental studies in the IVE with larger numbers of participants. In Chapter 5, we evaluate different design alternatives in the IVE. Participants bike in a simulated virtual environment modeled to scale from a real-world street with a shared bike lane (sharrows) to evaluate how the introduction of a curbside bike lane and a protected bike lane with flexible delineators may impact perceptions of safety, as well as behavioral and psycho-physiological responses. Results from 50 participants (representing both genders and across a wide age range) show that the protected bike lane design received the highest perceived safety rating and exhibited the lowest average cycling speed. Furthermore, both the curbside bike lane and the protected bike lane scenarios show a less dispersed gaze distribution than the as-built sharrows scenario, reflecting a higher gaze focus among bicyclists on the biking task in the curbside bike lane and protected bike lane scenarios, compared to when bicyclists share right of way with vehicles. Additionally, heart rate change point results from the study suggest that creating dedicated zones for bicyclists (curbside bike lanes or protected bike lanes) has the potential to reduce bicyclists' stress levels. Female participants show a higher preference for the protected bike lane design and a lower perceived safety rating on the sharrows. With this study, we are able to answer Yes to both RQ 3 and RQ 4 of RO 2: *Can the proposed framework support IVE* *bike simulator to study cyclists' behaviors in different roadway designs?* and *Are the cyclists' physiological response different from their stated preferences?*.

In Chapter 6, we are more interested in the cyclists' internal states, especially, the effect of cognitive distraction on cycling behavior and psychophysiological responses. This research explores the effect of different cognitive distractions on bicyclists' physiological and behavioral changes. In an immersive virtual environment, a bicycle simulator with multiple physiological sensing devices is utilized to collect bicyclists' behavioral and physiological responses on the same road design with bike lanes. Data collection includes demographic information (age, gender, biking attitude), engagement in a secondary task such as music listening and talking on the phone during their daily life or cycling, cycling performance in the simulator (speed, lane position, input power, head movement) and physiological responses (heart rate, skin temperature). Results from 75 participants indicate that (1) Cyclists would have a significantly higher speed, lower sd of speed, and higher input power in the music listening scenario. (2) When talking on the phone, cyclists will try a lower speed with less input power and less head movement variation. (3) When listening to music, cyclists who had a strong habit of daily music listening ($>$ 4 hours/day) had a higher engagement in the music, with a significantly lower sd of speed. Male cyclists stayed closer to the vehicle lane and had a higher head movement variation. (4) Lane position is not affected by the scenario, this may be the effect of introducing bike lanes in the environment. These results provide answers to RQ 5 and RQ 6 of RO 3: *Can we use a standardized secondary task to simulate cognitive distraction during cycling?* and *What's the effect of different types of cognitive distraction on cycling behavior?*. For RQ 5, based on our example of the mock phone conversation, it can create proper levels of distraction during cycling with great experiment control ability. For RQ 6, we find that depending on the levels of cognitive distraction tasks, different adaptive behaviors are observed. For lower levels of distraction like music listening, cyclists tend to have a higher engagement in the secondary task with higher speed, less adjustment in speed, higher input power, and more heart rate change points. With higher levels of cognitive distraction like phone talking, cyclists keep a lower speed with less input power and they will have less observation on the surrounding environment with lower head movement variation.

7.1 Takeaways and Practical Outcomes

The key takeaways and practical outcomes from the dissertation are listed as followed.

The IVE-based framework can be utilized for studying the behavioral changes of cyclists in different contextual settings. The lack of pre-accident data is one of the main challenges in finding the reasons behind traffic accidents, especially for vulnerable road users. The IVE-based framework is proved to be an effective tool with relatively lower cost and risk for understanding cyclists' responses to different infrastructures.

For designers, when evaluating new designs with user tests, it is important to conduct experiments with proper experiment design and include the users' physiological responses in addition to subjective measurements. As the users' objective measurements can be different from their subjective preferences, the designs should be evaluated in a more comprehensive way. The results in Chapter 5 show that we may not need protected bike lanes in all locations on the same road and we can more strategically add a hybrid protected and not protected bike lanes. For instance, through our gaze tracking analyses, we see that participants' stress level changes as they are approaching an intersection or a traffic light. Similarly, going downhill/uphill has an impact on their focus and awareness of the surrounding environments. By optimizing the locations where we have protected bike lanes, we can 1) reduce the cost of protected bike lanes since we won't have them everywhere, but also provide more sense of safety to cyclists.

For urban planners, it's crucial to consider contextual roadway settings such as traffic patterns, speed of vehicles, presence of other roadway users as well as the target populations that will be biking on this road (based on the surrounding neighborhood and community information). It is also important to study why certain users do not bike. For instance, across our participants, we identified that females were more inclined toward safer bike lanes such as the protected bike lanes. meanwhile, in their subjective responses, they indicated that safety is one of the main reasons they might not choose biking as the main mode of transportation. So as planners, we can potentially increase the usability of bike lanes by a larger number of members of the surrounding community if we can increase their sense of safety when they are biking on crowded streets. In other words, understanding the user and impacted community members' needs can inform the different options we can choose. Chapter 5 and Chapter 6 provide more details about these findings.

7.2 Future Directions

We discuss the limitations of our proposed work and possible future research directions during the development and use of the VR simulator in this dissertation, several limitations, and directions have been identified as areas of future research.

- **Development of simulators for other road users in IVE**. In addition to bicyclists and pedestrians, we haven't tested the ORCLSim framework for other road users such as drivers yet. We are planning to integrate the driving simulator in the ORCL lab into our system in the near future. In addition, the emergence of new road users like e-scooters and e-bikers has brought up new conflicts between the existing road users and the fatalities of these new road users have been increasing. The development of simulators for these road users will help researchers to better understand the new challenges.

- **Distributed multi-agent virtual reality simulator**. One shortcoming of existing virtual VRU simulators and driving simulators is that you are limited to integrating one user (agent) within the virtual space. Once the system framework can support the simulation of different road users, we can propose integrating multiple users, who could represent drivers (through a driving simulator), pedestrians, and bicyclists (through virtual simulators) within the same virtual environment. Furthermore, it would be possible to remotely connect multiple simulators across different sites. By connecting these distributed simulators together within a single IVE, we will allow multiple participants realistically interact with each other and the surrounding virtual environment. Through this system, we can evaluate each participant's behavioral and physiological responses and create a realistic dataset that can better inform future design decision-making.

- **Integration of other physiological sensing methods in the simulator**. In addition to the physiological sensing used in the dissertation, more physiological sensing could be included in the framework. For example, Electromyography (EMG) is a diagnostic procedure to assess the health of muscles and the nerve cells that control them (motor neurons). Cycling is a physical activity that requires a lot of manual input, the EMG data can provide more information about the physical activity. An electroencephalogram (EEG) is a test that measures electrical activity in the brain using small, metal discs (electrodes) attached to the scalp. EEG is more sensitive in users' internal states such as cognitive states.

- **Extension of current road network in IVE**. Our IVE is built from the Water Street Corridor in the city of Charlottesville, the initial IVE in Chapter 3 only has 4 street blocks, and in Chapter 6 we have extended the road to 8 street blocks. Our future vision is to extend the current road network to the whole city of Charlottesville.

- **Further improvement on the bicycle simulator components**. The current IVE-based is highly immersive and realistic both when rated by our participants and when compared to other existing bike simulators, however, it is not perfect. With the feedback from participants of different experiments, we have learned that the realism of steering can be further improved. The current steering is controlled by the controller, which requires steering calibration at the beginning of each scenario. The latest bike trainer has integrated the handlebar with the steering data input/output, we can make use of the new technology to upgrade the current simulator. Furthermore, the bike simulator is fix-based, and more degrees of freedom can be considered such as moving up and down according to the road environment.

- **Long-term benchmarking study both in the real world and in the IVE**. Although a benchmarking study is conducted in Chapter 4, a long-term benchmarking study is required. The study in Chapter 4 has a relatively small participant number and only tests in the daytime hours during working days with sunny weather. As we further extend the road network in IVE, a benchmarking study is needed both in the real world and in the IVE with different environment settings such as time of the day, weather conditions, and traffic conditions.

Bibliography

Yerkes, Robert M, John D Dodson, et al. (1908). "The relation of strength of stimulus to rapidity of habit-formation". In: *Punishment: Issues and experiments*, pp. 27–41.

Shannon, Claude Elwood (1948). "A mathematical theory of communication". In: *The Bell system technical journal* 27.3, pp. 379–423.

May, James G et al. (1990). "Eye movement indices of mental workload". In: *Acta psychologica* 75.1, pp. 75–89.

Coyle, EF et al. (1991). "Physiological and biomechanical factors associated with elite endurance cycling performance." In: *Medicine and science in sports and exercise* 23.1, pp. 93–107.

Barry, Daniel and John A Hartigan (1993). "A Bayesian analysis for change point problems". In: *Journal of the American Statistical Association* 88.421, pp. 309–319.

Van Veen, Hendrik AHC et al. (1998). "Navigating through a virtual city: Using virtual reality technology to study human action and perception". In: *Future Generation Computer Systems* 14.3-4, pp. 231–242.

Kwon, Dong-Soo et al. (2001). "KAIST interactive bicycle simulator". In: *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*. Vol. 3. IEEE, pp. 2313–2318.

Purves, Dale et al. (2001). "Types of eye movements and their functions". In: *Neuroscience*, pp. 361–390.

Harbluk, Joanne L, Y Ian Noy, and Moshe Eizenman (2002). *The impact of cognitive distraction on driver visual behaviour and vehicle control*. Tech. rep.

Krueger, Charlene and Lili Tian (2004). "A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points". In: *Biological research for nursing* 6.2, pp. 151–157.

Engström, Johan, Emma Johansson, and Joakim Östlund (2005). "Effects of visual and cognitive load in real and simulated motorway driving". In: *Transportation research part F: traffic psychology and behaviour* 8.2, pp. 97–120.

Erdman, Chandra and John W Emerson (2007). "bcp: an R package for performing a Bayesian analysis of change point problems". In: *Journal of Statistical Software* 23.1, pp. 1–13.

Muttart, Jeffrey W et al. (2007). "Driving without a clue: Evaluation of driver simulator performance during hands-free cell phone operation in a work zone". In: *Transportation Research Record* 2018.1, pp. 9–14.

Parkin, John, Mark Wardman, and Matthew Page (2007). "Models of perceived cycling risk and route acceptability". In: *Accident Analysis & Prevention* 39.2, pp. 364–371.

Caird, Jeff K et al. (2008). "A meta-analysis of the effects of cell phones on driver performance". In: *Accident Analysis & Prevention* 40.4, pp. 1282–1293.

Chapman, Andrew R et al. (2008). "Patterns of leg muscle recruitment vary between novice and highly trained cyclists". In: *Journal of Electromyography and Kinesiology* 18.3, pp. 359–371.

Daniels, Stijn, Erik Nuyts, and Geert Wets (2008). "The effects of roundabouts on traffic safety for bicyclists: an observational study". In: *Accident Analysis & Prevention* 40.2, pp. 518–526.

Drews, Frank A et al. (2009). "Text messaging during simulated driving". In: *Human factors* 51.5, pp. 762–770.

Sener, Ipek N, Naveen Eluru, and Chandra R Bhat (2009). "An analysis of bicycle route choice preferences in Texas, US". In: *Transportation* 36.5, pp. 511–539.

De Waard, Dick et al. (2010). "Mobile phone use while cycling: Incidence and effects on behaviour and safety". In: *Ergonomics* 53.1, pp. 30–42.

Pucher, John, Jennifer Dill, and Susan Handy (2010). "Infrastructure, programs, and policies to increase bicycling: an international review". In: *Preventive medicine* 50, S106–S125.

Wang, Ying, Bruce Mehler, et al. (2010). "The validity of driving simulation for assessing differences between in-vehicle informational interfaces: A comparison with field testing". In: *Ergonomics* 53.3, pp. 404–420.

Williams, Lynne J and Herve Abdi (2010). "Fisher's least significant difference (LSD) test". In: *Encyclopedia of research design* 218, pp. 840–853.

Haufe, Stefan et al. (2011). "EEG potentials predict upcoming emergency brakings during simulated driving". In: *Journal of neural engineering* 8.5, p. 056001.

Holmqvist, Kenneth et al. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Zahabi, Seyed Amir H et al. (2011). "Estimating potential effect of speed limits, built environment, and other factors on severity of pedestrian and cyclist injuries in crashes". In: *Transportation research record* 2247.1, pp. 81–90.

Flusche, Darren (2012). "Bicycling means business: The economic benefits of bicycle infrastructure". In.

Poulos, Roslyn G et al. (2012). "Exposure-based cycling crash, near miss and injury rates: The Safer Cycling Prospective Cohort Study protocol". In: *Injury prevention* 18.1, e1–e1.

Van Hoof, Wannes et al. (2012). “Comparing lower lumbar kinematics in cyclists with low back pain (flexion pattern) versus asymptomatic controls–field study using a wireless posture monitoring system”. In: *Manual therapy* 17.4, pp. 312–317.

Chaurand, Nadine and Patricia Delhomme (2013). “Cyclists and drivers in road interactions: A comparison of perceived crash risk”. In: *Accident Analysis & Prevention* 50, pp. 1176–1184.

Terzano, Kathryn (2013). “Bicycling safety and distracted behavior in The Hague, the Netherlands”. In: *Accident Analysis & Prevention* 57, pp. 87–90.

Thompson, Leah L et al. (2013). “Impact of social and technological distraction on pedestrian crossing behaviour: an observational study”. In: *Injury prevention* 19.4, pp. 232–237.

Dill, Jennifer et al. (2014). “Can protected bike lanes help close the gender gap in cycling? Lessons from five cities”. In: *Urban Studies and Planning Faculty Publications and Presentations*.

Mwakalonge, Judith L, Jamario White, and Saidi Siuhi (2014). “Distracted biking: a review of the current state-of-knowledge”. In: *International Journal of Traffic and Transportation Engineering* 3.2, pp. 42–51.

Wang, Ying, Bryan Reimer, et al. (2014). “The sensitivity of different methodologies for characterizing drivers’ gaze concentration under increased cognitive demand”. In: *Transportation research part F: traffic psychology and behaviour* 26, pp. 227–237.

Ayres, Thomas J et al. (2015). “Bicyclist behavior at stop signs”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 59. 1. SAGE Publications Sage CA: Los Angeles, CA, pp. 1616–1620.

Doorley, Ronan et al. (2015). "Analysis of heart rate variability amongst cyclists under perceived variations of risk exposure". In: *Transportation research part F: traffic psychology and behaviour* 28, pp. 40–54.

Heydarian, Arsalan, Joao P Carneiro, et al. (2015). "Immersive virtual environments versus physical built environments: A benchmarking study for building design and user-built environment explorations". In: *Automation in Construction* 54, pp. 116–126.

McNeil, Nathan, Christopher M Monsere, and Jennifer Dill (2015). "Influence of bike lane buffer types on perceived comfort and safety of bicyclists and potential bicyclists". In: *Transportation research record* 2520.1, pp. 132–142.

Stoker, Philip et al. (2015). "Pedestrian safety and the built environment: a review of the risk factors". In: *Journal of Planning Literature* 30.4, pp. 377–392.

Chen, Peng and Qing Shen (2016). "Built environment effects on cyclist injury severity in automobile-involved bicycle crashes". In: *Accident Analysis & Prevention* 86, pp. 239–246.

Correa, Alejandro et al. (2016). "Indoor pedestrian tracking by on-body multiple receivers". In: *IEEE Sensors Journal* 16.8, pp. 2545–2553.

Ethan, Danna et al. (2016). "An analysis of technology-related distracted biking behaviors and helmet use among cyclists in New York City". In: *Journal of community health* 41.1, pp. 138–145.

Nikolas, Molly A et al. (2016). "Risky bicycling behavior among youth with and without attention-deficit hyperactivity disorder". In: *Journal of child psychology and psychiatry* 57.2, pp. 141–148.

Standardization, International Organization for (2016). *Road vehicles–Transport information and control systems–Detection Response Task (DRT) for assessing attentional effects of cognitive load in driving.*

Tapiro, Hagai et al. (2016). “Where do older pedestrians glance before deciding to cross a simulated two-lane road? A pedestrian simulator paradigm”. In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 60. 1. SAGE Publications Sage CA: Los Angeles, CA, pp. 11–15.

Wang, Ying, Xiang Guo, et al. (2016). “Color Block Task: A New Surrogate Secondary Task to Measure the Impact of Drivers’ Incrementally Increased Workload”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 60. 1. SAGE Publications Sage CA: Los Angeles, CA, pp. 1889–1893.

Wolfe, Elizabeth Suzanne et al. (2016). “Distracted biking: an observational study”. In: *Journal of trauma nursing: the official journal of the Society of Trauma Nurses* 23.2, p. 65.

Zeile, Peter et al. (2016). “Urban Emotions and Cycling Experience–enriching traffic planning for cyclists with human sensor data”. In: *GI_Forum* 1.2013, pp. 204–216.

Akbar, Izzat A et al. (2017). “Three drowsiness categories assessment by electroencephalogram in driving simulator environment”. In: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, pp. 2904–2907.

Aldred, Rachel et al. (2017). “Cycling provision separated from motor traffic: a systematic review exploring whether stated preferences vary by gender and age”. In: *Transport reviews* 37.1, pp. 29–55.

Cao, Zhe et al. (2017). “Realtime multi-person 2d pose estimation using part affinity fields”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7291–7299.

Cloutier, Marie-Soleil et al. (2017). "“Outta my way!” Individual and environmental correlates of interactions between pedestrians and vehicles during street crossings". In: *Accident Analysis & Prevention* 104, pp. 36–45.

Coulacoglou, C and DH Saklofske (2017). "Executive function, theory of mind, and adaptive behavior". In: *Psychomterics and Psychological assessment. Principles and applications*, pp. 91–119.

Dawson, Michael E, Anne M Schell, and Diane L Filion (2017). "The electrodermal system." In.

Eudave, Luis and Miguel Valencia (2017). "Physiological response while driving in an immersive virtual environment". In: *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, pp. 145–148.

Heydarian, Arsalan and Burcin Becerik-Gerber (2017). "Use of immersive virtual environments for occupant behaviour monitoring and data collection". In: *Journal of Building Performance Simulation* 10.5-6, pp. 484–498.

Kwigizile, Valerian et al. (2017). *Real Time Bicycle Simulation Study of Bicyclists' Behaviors and their Implication on Safety*. Tech. rep. Western Michigan University. Transportation Research Center for Livable …

Lee, Oliver et al. (2017). "Description of a model based bicycle simulator". In: *Proceedings of the 6th Annual International Cycling Safety Conference*.

O'Hern, Steve, Jennie Oxley, and Mark Stevenson (2017). "Validation of a bicycle simulator for road safety research". In: *Accident Analysis & Prevention* 100, pp. 53–58.

Stroh, Oliver (2017). "The Design of an Electro-Mechanical Bicycle for an Immersive Virtual Environment". In.

Xu, Jing, Yingzi Lin, and David Schmidt (2017). “Exploring the Influence of Simulated Road Environments on Cyclist Behavior”. In: *International Journal of Virtual Reality* 17.3, pp. 15–26.

Yusoff, Norhasliza M et al. (2017). “Selection of measurement method for detection of driver visual cognitive distraction: A review”. In: *IEEE Access* 5, pp. 22844–22854.

Zou, Hao, Nan Li, and Lijun Cao (2017). “Emotional response–based approach for assessing the sense of presence of subjects in virtual building evacuation studies”. In: *Journal of Computing in Civil Engineering* 31.5, p. 04017028.

Abadi, Masoud Ghodrat and David S Hurwitz (2018). “Bicyclist’s perceived level of comfort in dense urban environments: How do ambient traffic, engineering treatments, and bicyclist characteristics relate?” In: *Sustainable cities and society* 40, pp. 101–109.

Barnard, Megan Patricia and Peter Chapman (2018). “The effects of instruction and environmental demand on state anxiety, driving performance and autonomic activity: Are ego-threatening manipulations effective?” In: *Transportation research part F: traffic psychology and behaviour* 55, pp. 123–135.

Fitch, Dillon T and Susan L Handy (2018). “The relationship between experienced and imagined bicycling comfort and safety”. In: *Transportation research record* 2672.36, pp. 116–124.

Francisco, Abigail et al. (2018). “Occupant perceptions of building information model-based energy visualizations in eco-feedback systems”. In: *Applied Energy* 221, pp. 220–228.

Keler, Andreas et al. (2018). “A bicycle simulator for experiencing microscopic traffic flow simulation in urban environments”. In: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 3020–3023.

Kim, Hye-Geum et al. (2018). “Stress and heart rate variability: A meta-analysis and review of the literature”. In: *Psychiatry investigation* 15.3, p. 235.

Li, Penghui et al. (2018). “Does cognitive distraction improve or degrade lane keeping performance? Analysis of time-to-line crossing safety margins”. In: *Transportation research part F: traffic psychology and behaviour* 57, pp. 48–58.

Napoli, Nicholas J et al. (2018). “Uncertainty in heart rate complexity metrics caused by R-peak perturbations”. In: *Computers in biology and medicine* 103, pp. 198–207.

Nazemi, Mohsen, Michael AB van Eggermond, Alexander Erath, and Kay W Axhausen (2018). “Studying cyclists’ behavior in a non-naturalistic experiment utilizing cycling simulator with immersive virtual reality”. In: *Arbeitsberichte Verkehrs- und Raumplanung* 1383.

Nazemi, Mohsen, Michael AB van Eggermond, Alexander Erath, Dorothea Schaffner, et al. (2018). “Speed and space perception in virtual reality for bicycle research”. In: *7th International Cycling Safety Conference (ICSC 2018)*.

Organization, World Health (2018). *Global status report on road safety 2018: summary*. Tech. rep. World Health Organization.

Ridel, Daniela et al. (2018). “A literature review on the prediction of pedestrian behavior in urban scenarios”. In: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 3105–3112.

Robartes, Erin and T Donna Chen (2018). “Crash histories, safety perceptions, and attitudes among Virginia bicyclists”. In: *Journal of safety research* 67, pp. 189–196.

Schultheiss, William, Rebecca L Sanders, and Jennifer Toole (2018). “A historical perspective on the AASHTO guide for the development of bicycle facilities and the impact of the vehicular cycling movement". In: *Transportation research record* 2672.13, pp. 38–49.

Shaaban, Khaled, Deepti Muley, and Abdulla Mohammed (2018). "Analysis of illegal pedestrian crossing behavior on a major divided arterial road". In: *Transportation research part F: traffic psychology and behaviour* 54, pp. 124–137.

Shiferaw, Brook A et al. (2018). "Stationary gaze entropy predicts lane departure events in sleep-deprived drivers". In: *Scientific reports* 8.1, pp. 1–10.

Stelling-Konczak, A et al. (2018). "A study in real traffic examining glance behaviour of teenage cyclists when listening to music: results and ethical considerations". In: *Transportation research part F: traffic psychology and behaviour* 55, pp. 47–57.

Sun, Carlos and Zhu Qing (2018). "Design and construction of a virtual bicycle simulator for evaluating sustainable facilities design". In: *Advances in Civil Engineering* 2018.

Trefzger, Mathias et al. (2018). "A visual comparison of gaze behavior from pedestrians and cyclists". In: *Proceedings of the 2018 ACM symposium on eye tracking research & applications*, pp. 1–5.

Useche, Sergio A et al. (2018). "Distraction of cyclists: how does it influence their risky behaviors and traffic crashes?" In: *PeerJ* 6, e5616.

Uttley, Jim, James Simpson, and Hussain Qasem (2018). "Eye-tracking in the real world: Insights about the urban environment". In: *Handbook of Research on Perception-Driven Approaches to Urban Assessment and Design*. IGI Global, pp. 368–396.

Abadi, Masoud Ghodrat, David S Hurwitz, et al. (2019). "Factors impacting bicyclist lateral position and velocity in proximity to commercial vehicle loading zones: Application of a bicycling simulator". In: *Accident Analysis & Prevention* 125, pp. 29–39.

Ballmann, Christopher G et al. (2019). "Effects of listening to preferred versus non-preferred music on repeated wingate anaerobic test performance". In: *Sports* 7.8, p. 185.

Chías, Pilar et al. (2019). "3D MODELLING AND VIRTUAL REALITY APPLIED TO COMPLEX ARCHITECTURES: AN APPLICATION TO HOSPITALS'DESIGN." In: *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*.

Ergan, Semiha et al. (2019). "Quantifying human experience in architectural spaces with integrated virtual reality and body sensor networks". In: *Journal of Computing in Civil Engineering* 33.2, p. 04018062.

Guo, Xiang, Lian Cui, et al. (2019). "How Will Humans Cut Through Automated Vehicle Platoons in Mixed Traffic Environments? A Simulation Study of Drivers' Gaze Behaviors Based on the Dynamic Areas of Interest". In: *Systems Engineering in Context*. Springer, pp. 691–701.

Huemer, Anja Katharina, Selvi Gercek, and Mark Vollrath (2019). "Secondary task engagement in German cyclists–An observational study". In: *Safety Science* 120, pp. 290–298.

Kim, Jinwoo, Megha Yadav, Changbum R Ahn, et al. (2019). "Saliency detection analysis of pedestrians' physiological responses to assess adverse built environment features". In: *Computing in Civil Engineering 2019: Smart Cities, Sustainability, and Resilience*. American Society of Civil Engineers Reston, VA, pp. 90–97.

Lohani, Monika, Brennan R Payne, and David L Strayer (2019). "A review of psychophysiological measures to assess cognitive states in real-world driving". In: *Frontiers in human neuroscience* 13, p. 57.

Mitra, Raktim and Sean Nash (2019). "Can the built environment explain gender gap in cycling? An exploration of university students' travel behavior in Toronto, Canada". In: *International journal of sustainable transportation* 13.2, pp. 138–147.

Paridon, Kjell N van et al. (2019). "Visual search behaviour in young cyclists: A naturalistic experiment". In: *Transportation research part F: traffic psychology and behaviour* 67, pp. 217–229.

Rupi, Federico and Kevin J Krizek (2019). "Visual eye gaze while cycling: Analyzing eye tracking at signalized intersections in urban conditions". In: *Sustainability* 11.21, p. 6089.

Shiferaw, Brook, Luke Downey, and David Crewther (2019). "A review of gaze entropy as a measure of visual scanning efficiency". In: *Neuroscience & Biobehavioral Reviews* 96, pp. 353–366.

Wynne, Rachael A, Vanessa Beanland, and Paul M Salmon (2019). "Systematic review of driving simulator validation studies". In: *Safety science* 117, pp. 138–151.

Yang, Hongtai et al. (2019). "Underreporting, crash severity and fault assignment of minor crashes in China–a study based on self-reported surveys". In: *International journal of injury control and safety promotion* 26.1, pp. 30–36.

Zhao, Yongxiang et al. (2019). "Quantitative measurement of social repulsive force in pedestrian movements based on physiological responses". In: *Transportation research part B: methodological* 130, pp. 1–20.

Administration, National Highway Traffic Safety et al. (2020). "Overview of motor vehicle crashes in 2019". In: *US Department of Transportation: Washington, DC, USA*.

Bogacz, Martyna, Stephane Hess, Chiara Calastri, Charisma F Choudhury, Alexander Erath, et al. (2020). "Comparison of cycling behavior between keyboard-controlled and instrumented bicycle experiments in virtual reality". In: *Transportation research record* 2674.7, pp. 244–257.

Boukhechba, Mehdi and Laura E Barnes (2020). "Swear: Sensing using wearables. Generalized human crowdsensing on smartwatches". In: *International Conference on Applied Human Factors and Ergonomics*. Springer, pp. 510–516.

Ebadi, Yalda et al. (2020). "Impact of cognitive distractions on drivers' hazardous event anticipation and mitigation behavior in vehicle–bicycle conflict situations". In: *Transportation research record* 2674.7, pp. 504–513.

Fitch, Dillon T, James Sharpnack, and Susan L Handy (2020). "Psychological stress of bicycling with traffic: examining heart rate variability of bicyclists in natural urban environments". In: *Transportation research part F: traffic psychology and behaviour* 70, pp. 81–97.

Gao, Yuyan et al. (2020). "Unsafe bicycling behavior in Changsha, China: A video-based observational study". In: *International journal of environmental research and public health* 17.9, p. 3256.

Guo, Xiang, Yichen Jiang, and Inki Kim (2020). "Interacting with Autonomous Platoons: Human Driver's Adaptive Behaviors in Planned Lane Changes". In: *2020 Systems and Information Engineering Design Symposium (SIEDS)*. IEEE, pp. 1–5.

Kim, Jinwoo, Megha Yadav, Theodora Chaspari, et al. (2020a). "Environmental Distress and Physiological Signals: Examination of the Saliency Detection Method". In: *Journal of Computing in Civil Engineering* 34.6, p. 04020046.

— (2020b). "Saliency detection analysis of collective physiological responses of pedestrians to evaluate neighborhood built environments". In: *Advanced Engineering Informatics* 43, p. 101035.

Lee, Gaang et al. (2020). "Wearable biosensor and collective sensing–based approach for detecting older adults' environmental barriers". In: *Journal of Computing in Civil Engineering* 34.2, p. 04020002.

National Highway Traffic Safety Administration, NHTSA (2020). "Fatality analysisi reporting system encyclopedia". In: *http://www-fars. nhtsa. dot. gov/Main/index.aspx*.

Noghabaei, Mojtaba and Kevin Han (2020). "Hazard recognition in an immersive virtual environment: Framework for the simultaneous analysis of visual search and EEG patterns". In: *Construction Research Congress 2020: Computer Applications*. American Society of Civil Engineers Reston, VA, pp. 934–943.

Rybarczyk, Greg et al. (2020). "Physiological responses to urban design during bicycling: A naturalistic investigation". In: *Transportation research part F: traffic psychology and behaviour* 68, pp. 79–93.

Shoman, M and H Imine (2020). "Modeling and simulation of bicycle dynamics". In: *Proc. TRA*, pp. 1–10.

Soares, Sónia, Sara Ferreira, and António Couto (2020). "Driving simulator experiments to study drowsiness: a systematic review". In: *Traffic injury prevention* 21.1, pp. 29–37.

Stülpnagel, Rul von (2020). "Gaze behavior during urban cycling: Effects of subjective risk perception and vista space properties". In: *Transportation research part F: traffic psychology and behaviour* 75, pp. 222–238.

Tavakoli, Arash, Mehdi Boukhechba, and Arsalan Heydarian (2020). "Personalized driver state profiles: A naturalistic data-driven study". In: *International Conference on Applied Human Factors and Ergonomics*. Springer, pp. 32–39.

Teixeira, Inaian Pignatti et al. (2020). “Does cycling infrastructure reduce stress biomarkers in commuting cyclists? A comparison of five European cities”. In: *Journal of Transport Geography* 88, p. 102830.

Young, Kristie L et al. (2020). “Australian cyclists’ engagement in secondary tasks”. In: *Journal of Transport & Health* 16, p. 100793.

Adami, Pooya et al. (2021). “Effectiveness of VR-based training on improving construction workers’ knowledge, skills, and safety behavior in robotic teleoperation”. In: *Advanced Engineering Informatics* 50, p. 101431.

Awada, Mohamad et al. (2021). “An integrated emotional and physiological assessment for VR-based active shooter incident experiments”. In: *Advanced Engineering Informatics* 47, p. 101227.

Baee, Sonia et al. (2021). “MEDIRL: Predicting the visual attention of drivers via maximum entropy deep inverse reinforcement learning”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 13178–13188.

Bogacz, Martyna, Stephane Hess, Chiara Calastri, Charisma F Choudhury, Faisal Mushtaq, et al. (2021). “Modelling risk perception using a dynamic hybrid choice model and brain-imaging data: Application to virtual reality cycling”. In: *Transportation Research Part C: Emerging Technologies* 133, p. 103435.

Brand, Christian et al. (2021). “The climate change mitigation effects of daily active travel in cities”. In: *Transportation Research Part D: Transport and Environment* 93, p. 102764.

Buehler, Ralph and John Pucher (2021). *COVID-19 Impacts on Cycling, 2019–2020.*

Bustos, Cristina et al. (2021). “Predicting Driver Self-Reported Stress by Analyzing the Road Scene”. In: *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE Computer Society, pp. 1–8.

Cobb, Douglas P, Hisham Jashami, and David S Hurwitz (2021). "Bicyclists' behavioral and physiological responses to varying roadway conditions and bicycle infrastructure". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 80, pp. 172–188.

Dittrich, Monique (2021). "Why Drivers Feel the Way they Do: An On-the-Road Study Using Self-Reports and Geo-Tagging". In: *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 116–125.

Guo, Xiang (2021). *ORCL VR EyeTracking Set Up.* URL: https://git.io/JK3hq.

Guo, Xiang, Erin Robartes, et al. (2021). "Benchmarking the use of immersive virtual bike simulators for understanding cyclist behaviors". In: *Computing in Civil Engineering 2021*, pp. 1319–1326.

Nazemi, Mohsen, Michael AB van Eggermond, Alexander Erath, Dorothea Schaffner, et al. (2021). "Studying bicyclists' perceived level of safety using a bicycle simulator combined with immersive virtual reality". In: *Accident Analysis & Prevention* 151, p. 105943.

NHTSA (2021a). *Fatality analysis reporting system encyclopedia.*

— (2021b). "Traffic Safety Facts Research Note: Distracted Driving 2019". In: *Report No. DOT HS* 813, p. 111.

Noghabaei, Mojtaba, Kevin Han, and Alex Albert (2021). "Feasibility Study to Identify Brain Activity and Eye-Tracking Features for Assessing Hazard Recognition Using Consumer-Grade Wearables in an Immersive Virtual Environment". In: *Journal of Construction Engineering and Management* 147.9, p. 04021104.

Rodriguez-Valencia, Alvaro et al. (2021). "Towards an enriched framework of service evaluation for pedestrian and bicyclist infrastructure: acknowledging the power of users' perceptions". In: *Transportation*, pp. 1–24.

Sharif, Marissa A and Daniel M Oppenheimer (2021). "The effect of categories on relative encoding biases in memory-based judgments". In: *Organizational Behavior and Human Decision Processes* 162, pp. 1–8.

Tavakoli, Arash, Mehdi Boukhechba, and Arsalan Heydarian (2021). "Leveraging Ubiquitous Computing for Empathetic Routing: A Naturalistic Data-driven Approach". In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–6.

Tavakoli, Arash, Shashwat Kumar, Mehdi Boukhechba, et al. (2021). "Driver state and behavior detection through smart wearables". In: *2021 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 559–565.

Tavakoli, Arash, Shashwat Kumar, Xiang Guo, et al. (2021). "HARMONY: A Human-Centered Multimodal Driving Study in the Wild". In: *IEEE Access* 9, pp. 23956–23978.

Zou, Xin et al. (2021). "On-road virtual reality autonomous vehicle (VRAV) simulator: An empirical study on user experience". In: *Transportation Research Part C: Emerging Technologies* 126, p. 103090.

Guo, Xiang, Austin Angulo, et al. (Aug. 2022). "ORCLSim: A System Architecture for Studying Bicyclist and Pedestrian Physiological Behavior through Immersive Virtual Environments". In: *Journal of Advanced Transportation* 2022, p. 2750369. issn: 0197-6729. doi: 10.1155/2022/2750369. url: https://doi.org/10.1155/2022/2750369.

Wender, Carly LA et al. (2022). "Virtual reality-based distraction on pain, performance, and anxiety during and after moderate-vigorous intensity cycling". In: *Physiology & Behavior* 250, p. 113779.

Appendices

Appendix A

Informed Consent Agreement

Please read this consent agreement carefully before you decide to participate in the study.

Purpose of the research study: The purpose of this research is to test the effectiveness of Virtual Reality (VR) as tool to replicate realistic environmental settings at a low cost while reducing risk to the user during experimentation. In this experiment, we aim to increase understanding of perceived safety and technological acceptance as it relates to bicyclists and the road environment. This information can be used by planners and engineers to better design technology and infrastructure for bicyclists.

With VR, we can study human behaviors in settings/scenarios that (1) we have limited or no access to (e.g., design of a new intersection that has not been built yet) or (2) are considered high-risk environments for collecting real-life data (e.g., bicyclist safety or crash rates at an intersection).

What you will do in the study: The goal of the Bicyclist Study is to place bicyclists in an environment in which they can naturally interact with vehicles. The participant will be seated on a stationary bike and will be wearing a VR headset and physiological sensing. The instrumented bicycle will allow their actions to be replicated in the virtual environment (speeding up, slowing down, steering). Specifically, this research aims to study how bicyclists behave in scenarios where they will be involved with different secondary tasks. You will be given a short questionnaire after each test in which you will respond to your thoughts and feelings regarding your experience.

Time required: The study will require about 1 hour of your time.

Risks: The physical components of these tasks are not stressful, and include head and body turning, moving, and pointing. Light and sound intensities are well within normal ranges. The only foreseeable physical risks are slight eye strain, dizziness, and mild nausea. There are no known mental risks. You will be asked to remove the head mounted display if they experience any eye strain, dizziness, or nausea during the sessions. They will be given rest breaks in between the sessions. Upon request, you will also be allowed to stop and leave the experiment if you feel uncomfortable or cannot continue the experiment. A loss of confidentiality would not put you at risk, and the researchers will use caution in handling the data.

Benefits: There are no direct benefits associated with the participation in this study. The proposed experiments are straightforward tests of performance and visual comfort using standard virtual environments displays and trackers.

Confidentiality: The information that you give in the study will be handled confidentially. Your information will be assigned a code number. The list connecting your email to this code will be kept in a locked file. When the study is completed and the data have been analyzed, this list will be deleted. Your name will not be used in any report. Once any data is deleted from a request, the changes will propagate correspondingly to the backup drives.

Voluntary participation: Your participation in the study is completely voluntary. Deciding not to participate will have no effect on your education at the University of

Virginia.

Right to withdraw from the study: You have the right to withdraw from the study at any time without penalty.

How to withdraw from the study: If you want to withdraw from the study, please contact the ORCL lab at xg3kb@virginia.edu indicating that you would like to withdraw from the study. There is no penalty for withdrawing. You may request that your archived data to be destroyed upon withdrawing from the study.

Payment: You will receive a $20 gift card as payment for participating in the study.

If you have questions about the study, contact:

Donna Chen

Engineering Systems and Environment

151 Engineer's Way, Room 101G

University of Virginia, Charlottesville, VA 22904

Telephone: (434) 924-6224

Email address: tdchen@virginia.edu

Arsalan Heydarian

Engineering Systems and Environment

151 Engineer's Way, Room 281

University of Virginia, Charlottesville, VA 22904

Telephone: (434) 924-1014

Email address: ah6rx@virginia.edu

Research Assistants

Xiang Guo

Engineering Systems and Environment

Thornton Hall, Room D101

University of Virginia, Charlottesville, VA 22904

Email address: xg3kb@virginia.edu

To obtain more information about the study, ask questions about the research procedures, express concerns about your participation, or report illness, injury or other problems, please contact:

Tonya R. Moon, Ph.D.

Chair, Institutional Review Board for the Social and Behavioral Sciences

One Morton Dr Suite 500

University of Virginia, P.O. Box 800392

Charlottesville, VA 22908-0392

Telephone: (434) 924-5999

Email: irbsbshelp@virginia.edu

Website: www.virginia.edu/vpr/irb/sbs

Agreement:

I agree to participate in the research study described above.

Print Name:________________________

Signature: ____________________

Date: ____________________

You will receive a copy of this form for your records.