Prediction of Simulation Sickness induced by Virtual Reality Driving Simulation based on Individual, Behavioral, and Physiological Factors

DISSERTATION

zur Erlangung des akademischen Grades
Doktor der Informatik

Lehrstuhl für Multimodale Menschzentrierte Künstliche Intelligenz
Universität Augsburg

M.Sc. Stanislava Antonova Rangelova

2021

Datum der Disputation	Augsburg, 15.11.2021
Erstgutachter	Prof. Dr. Elisabeth André
Zweitgutachter	Prof. Dr.-Ing. Johannes Schilp

Abstract

This thesis investigates the prediction of simulation sickness induced by a virtual reality driving simulation. It assesses the effects of individual, behavioral, and physiological factors on the onset of simulation sickness and evaluates the prediction models that apply these factors. The users' sense of presence, as induced by virtual reality, is additionally explored as a contributory factor to simulation sickness.

Simulation sickness is a condition of physiological discomfort experienced during or after exposure to any virtual environment. It is a well-known phenomenon that is often compared to motion sickness. While virtual reality has and will continue to positively impact the development and testing of new automotive interior concepts, simulation sickness remains one of its most significant drawbacks. Despite advances in technology, the discomfort caused by using head-mounted displays has yet to be resolved.

To contribute to a solution, this thesis examines in experiments with a total of 94 participants the effects of a moving platform, gender, and types of driving, as well as evaluating different prediction models of simulation sickness. These features are considered a relevant combination of factors related to simulation sickness experienced in automotive interior development. As automated driving is considered the future of transportation, a driving simulation that imitates real-life automated driving could be a perfect testing or training tool. Nonetheless, virtual automated driving could potentially inherit one of the common problems of real-life automated driving, namely, motion sickness. In order to assemble a virtual reality driving system for user evaluations, a moving platform and physiological sensors were implemented. Three experiments, including static and dynamic driving, as well as automated and standard driving, were carried out. Furthermore, several prediction models based on individual factors (such as gender, motion sickness history, and previous experiences), behavioral factors (such as head movements and the speed of the virtual vehicle), and physiological factors (such as heart rate variability, skin conductance, and respiration) were evaluated.

The findings showed that simulation sickness induced by virtual reality driving simulation could be successfully predicted by individual, behavioral, and physiological factors. However, the individual factors showed a low variance in the explanation of the models. A further investigation showed that the features extracted from the cardiovascular signal could predict simulation sickness with similar accuracy to the combination model (behavioral and physiological factors). Moreover, simulation sickness was not significantly affected by the addition of motion cues or by changes in the driving conditions. Gender, as revealed by the significant main effect on simulation sickness, is a simulation sickness factor with a high susceptibility potential. Further work is required to

establish the importance of individual factors as well as physiological factors as predictors of simulation sickness.

Keywords: simulation sickness, virtual reality, driving simulation, physiological signals, machine learning, prediction models

Zusammenfassung

In der vorliegenden Doktorarbeit werden Möglichkeiten untersucht, das Auftreten der Simulationskrankheit vorherzusagen. Diese kann durch eine Fahrsimulation in virtueller Realität ausgelöst werden. Dazu werden die Einflüsse von individuellen, verhaltensbezogenen und physiologischen Faktoren sowie Vorhersagemodelle auf Grundlage dieser Faktoren untersucht. Zusätzlich wird das Präsenzgefühl von Individuen als beeinflussender Faktor der Simulationskrankheit in der virtuellen Realität erforscht.

Simulationskrankheit ist ein Zustand physiologischen Unbehagens, das während oder nach der Exposition gegenüber einer virtuellen Umgebung empfunden wird. Es ist ein bekanntes Phänomen, das ähnliche Symptome wie die Reisekrankheit aufweist. Während sich die virtuelle Realität positiv auf die Entwicklung und Erprobung neuer Konzepte für den Fahrzeuginnenraum ausgewirkt hat und weiterhin auswirken wird, stellt die Simulationskrankheit in ihren Auswirkungen auf ein Individuum einen erheblichen Nachteil dar. Trotz der Fortschritte in der Technologie ist das Unbehagen bei der Verwendung von Head-Mounted-Displays noch nicht behoben.

Im Rahmen dieser Arbeit wird bei insgesamt 94 Versuchspersonen der Einfluss einer bewegten Plattform, des Geschlechts und der Art des Fahrens sowie verschiedene Vorhersagemodelle der Simulationskrankheit untersucht. Die so erarbeiteten Faktoren sind besonders wichtig für die Entwicklung des automobilen Innenraums, da sie als relevante Faktoren in Bezug auf die Verringerung oder Behebung der Simulationskrankheit angesehen werden. Eine innovative Art der Mobilität, wie das autonome Fahren ohne Unterstützung durch den Fahrer, wird als die Zukunft der Mobilität angesehen. Nichtsdestotrotz könnte das virtuelle autonome Fahren möglicherweise zu einem der üblichen Probleme des realen autonomen Fahrens führen, nämlich der Reisekrankheit. Um ein Fahrsystem in der virtuellen Realität für Benutzerevaluierungen aufzubauen, wurden eine bewegliche Plattform und physiologische Sensoren eingeführt. Es wurden drei Experimentreihen durchgeführt, darunter Kombinationen von statischem und dynamischem Fahren sowie automatisiertem und selbstgesteuertem Fahren. Darüber hinaus wurden mehrere Vorhersagemodelle evaluiert, die individuelle Faktoren wie Geschlecht, eine eventuelle Vorgeschichte mit der Reisekrankheit, frühere Erfahrungen, Verhaltensfaktoren wie Kopfbewegungen und die Geschwindigkeit des virtuellen Fahrzeugs sowie physiologische Faktoren wie Herzfrequenzvariabilität, Hautleitwert und Atmung berücksichtigen.

Die Ergebnisse zeigen, dass die Simulationskrankheit, die durch die Fahrsimulation in der virtuellen Realität ausgelöst wird, durch individuelle, verhaltensbezogene und physiologische Faktoren erfolgreich vorhergesagt werden kann. Allerdings zeigen die einzelnen Faktoren eine geringe Varianz in der Erklärung der Modelle. Eine weitere Untersuchung zeigt, dass die aus dem kardiovaskulären Signal extrahierten Faktoren die Simulationskrankheit mit ähnlicher Genauigkeit

vorhersagen können wie das Kombinationsmodell (verhaltensbezogene und physiologische Faktoren). Allerdings wird die Simulationskrankheit nicht signifikant durch das Hinzufügen von Bewegungshinweisen oder die Veränderung der Fahrbedingungen beeinflusst. Das Geschlecht ist ein Krankheitsfaktor mit hohem Anfälligkeitspotential, wie der signifikante Effekt auf die Simulationskrankheit zeigt. Weitere Forschung ist erforderlich, um die Bedeutung individueller sowie physiologischer Faktoren als Vorhersagefaktoren für die Simulationskrankheit zu ermitteln.

Schlagwörter: Simulationskrankheit, Virtual Reality, Fahrsimulation, Physiologische Signale, Maschinelles Lernen, Vorhersagemodelle

На Максимилиан

(To Maximilian)

Acknowledgement

I would like to express my sincere gratitude to my supervisor Prof. Dr. Elisabeth André, for her invaluable help and support with this dissertation. She consistently allowed this project to be my own work but steered me in the right direction whenever she thought I needed it. She always provided me with valuable sources and reacted to my queries. Thank you very much!

I would like to thank my supervisor at BMW Group in Munich, Dr. Daniel Motus, for his support through each stage of this dissertation and for giving me the chance to work on such an exciting topic. I also wish to thank all employees but not limited to the Department of Interior Development for their support and participation in my studies. Many thanks to my students for their help in the preparation and conductance of the user studies.

Thanks should also go to all my colleagues at the Human-Centered Artificial Intelligence lab, University of Augsburg, for creating a friendly and inspirational working atmosphere. I would like to say a special thank you to Simon Flutura, Tobias Huber, and Katharina Weitz for their invaluable support, collaboration, and fruitful discussions.

In addition, I would like to express my gratitude to my family and friends, especially to my mother and my sister, for providing me with unfailing support and continuous encouragement throughout my years of study. I am incredibly grateful to my fiancé for his endless support. This accomplishment would not have been possible without them.

Abbreviations

BVP	Blood Volume Pulse
$CAVE$	Cave Automatic Virtual Environment
DOF	Degrees of Freedom
EDA	Electrodermal Activity
FMS	Fast Motion sickness Scale
FOV	Field-of-View
$FSSQ$	Follow-up Simulator Sickness Questionnaire
HMD	Head-Mounted Display
HR	Heart Rate
HRD	Heart Rate Difference
IPQ	iGroup Presence Questionnaire
$MSSQ$	Motion Sickness Susceptibility Questionnaire
PQ	Presence Questionnaire
RR	Respiration Rate
RRD	Respiration Rate Difference
SCL	Skin Conductance Level
$SCLD$	Skin Conductance Level Difference
SiS	Simulation Sickness
SSQ	Simulator Sickness Questionnaire
VR	Virtual Reality

Contents

List of Figures

List of Tables

1 Introduction

> Have you ever had a dream, Neo,
> that you were so sure was real?
> What if you were unable to wake
> from that dream, Neo? How would
> you know the difference between
> the dream world and the real world?
>
> *Morpheus*, "The Matrix"

1.1 Motivation

In recent years, virtual reality (VR) has become increasingly popular and is now accessible to the general public thanks to improvements in technology and cost reductions. Particularly in the automotive sector, VR offers new opportunities to evaluate interior and user interface concepts during development. The development of a vehicle's interior is a long and complicated process involving many experts such as designers, engineers, and researchers. Such a process sometimes demands quick changes while testing the interior, which can be a challenging task. To solve this problem, virtual environments have the potential to offer the advantage of rapidly changing the displayed environment. With VR systems, a safe, fully controlled, and yet highly immersive environment can be provided at a fraction of the costs of real driving studies. Thus, environments that are closely matched to realistic driving experiences can be utilised without risk of liability issues or high costs [225]. This is especially important in a fast-paced interior development area. The improvement of VR technology has been aided by the rapid advancement of computer technology. This technology is not only used to show stationary interior concepts, though; it can also be used to create a more interactive environment, such as a VR driving simulation.

Operationalized through head-mounted displays (HMD), VR offers a stereoscopic 3-dimensional (3D) environment with a wide field-of-view (FOV) and a low-latency fast-tracking system for better interaction in the virtual world. Thus, integrating an HMD into a driving simulation completely disconnects the user from the real world; in other words, the driver is tricked into believing that he or she is part of the virtual environment. One of the significant advantages of the phenomenon of presence is that it allows the user to feel sufficiently immersed in a virtual environment that realistic automated driving scenarios can be executed without the hazardous effects of a real-world study.

However, common human factors and usability issues remain and can impair the potential of this technology. The most common issue is the onset of simulation sickness (SiS), a condition of physiological discomfort experienced during or after exposure to any simulated environment. Individuals exposed to a simulated environment, such as a driving simulation, can experience symptoms such as nausea, dizziness, disorientation, or headache. The addition of VR technology by using an HMD as part of a driving simulation system could significantly contribute to a SiS outbreak [230]. In the relevant literature, discomfort induced by a virtual environment has been described using the terms "visually induced motion sickness", "VR sickness", and "cybersickness". Such discomfort is usually accompanied by visual symptoms: this is the main difference between the SiS induced by conventional simulators and the SiS caused by VR. Throughout this thesis, then, the term "simulation sickness" with the abbreviation "SiS" will refer to a condition that does not require a real physical motion. The condition is known to occur under visual motion simulation and develops using a wide FOV [18].

Symptoms such as *general discomfort, eye strain,* and *difficulty concentrating* are more likely to be experienced while a user is in a fully immersive virtual environment such as an HMD [97]. For example, *eye strain*, which in turn can cause *headache*, can result from the close distance between the eyes and the HMD's screen. These symptoms, also known as oculomotor symptoms, are the primary difference between SiS and motion sickness, in which nausea-like discomfort prevails [107, 218]. At first glance, the symptoms do not seem dangerous to participants during exposure to the simulated environment; however, it is possible that they could have a long-lasting effect after the simulation. This situation could be dangerous not only for the individual, but also for those surrounding them. Some sources note the potential danger posed to users who drive in the real world after prolonged exposure to a driving simulation. Fortunately, there have been no reports of automobile or motorcycle accidents as a result of SiS within twelve hours after exposure to a simulated environment [99]. Nevertheless, the potential risks of SiS should not be underestimated due to the unknown duration of the symptoms after exposure to the simulated environment. SiS should therefore be considered as serious as any other medical issue.

In a VR study, Kolasinski and Gilson [128] reported that 85% of participants experienced at least one SiS symptom after the virtual environment exposure. Another study reported that a large percentage of participants, between $60 - -70\%$, experienced SiS while using an HMD [196]. Furthermore, a more recent study by Clifton and Palmisano[41] found that 96% of the

participants reported SiS symptoms on at least one of the four experimental trials (steering - sitting and standing; teleportation - sitting and standing). These findings show that the percentage of users who experience SiS has not fallen over recent decades, which suggests that in spite of advances in VR technology, SiS remains a problem. Indeed, it can be a significant drawback in the usage of technology as people who have previously experienced discomfort are not willing to re-use the technology. In addition, as the vehicle's interior testing becomes more interactive and integrates driving simulation in the interior concepts, the risk of experiencing SiS increases exponentially.

According to cue conflict theory [172], the discrepancy between visual and motion cues is one of the assumed reasons for SiS while using driving simulators [42]. Based on this theory, reducing the mismatch between the visual and vestibular sensory systems should mitigate SiS. The addition of a motion platform to a static driving simulation is a commonly used technique to simulate the missing motion cues. However, the experimental data is somewhat controversial. For example, Aykent et al. [10] and Curry et al. [48] showed that the addition of motion cues reduces SiS, while Keshavarz et al. [108] observed no differences in SiS under motion condition compared to static motion condition. Nevertheless, the addition of motion cues to a static VR driving simulation is touted as one of the promising mitigation techniques.

Widespread use of automated vehicles for transportation is an anticipated future trend, so it is likely that the number of driving simulations using an automated driving type will increase. Automated driving is categorized into five levels, where level five relates to fully automated driving without any human intervention. Fully automated cars, therefore, have the highest level of automation. The driving system has full control over all driving tasks under all road conditions managed by a human driver at the lowest level [188]. It is known that drivers experience less motion sickness than passengers. Vehicle control, which separates drivers' and passengers' susceptibility to motion sickness, is significant in understanding SiS etiology [172]. Rolnick and Lubow [184] conducted a rotation experiment in which participants were divided into "drivers" and "passengers." The drivers performed a set of nauseogenic rotating movements, while the passengers merely watched and experienced the accompanying motion. The results showed that the drivers felt fewer motion sickness symptoms than the passengers due to their relative control over the rotations. This lack of control contributes significantly to motion sickness onset in individuals who are more susceptible to it due to their incapability to adequately predict the upcoming movement trajectory. The full driving automation would allow the driver to concentrate her or his attention on tasks other than driving. Engagement with non-driving tasks such as reading, watching videos, or communicating with other passengers face-to-face would be possible.

Furthermore, the driver's active role would change to the more passive role of passenger during the automated driving. These scenarios could significantly increase the possibility of motion sickness outbreak while using automated vehicles. Factors related to motion sickness in automated driving cars and possible scenarios were discussed previously in the literature as well as plausible mitigation methods [56]. One of these scenarios looked into the transition from a driver to a co-

driver or a passenger. An important aspect of this transition is the ability to anticipate upcoming motions. Thus, a design guiding principle was suggested to provide adequate visual information to the passengers about the automated vehicle path. Wada [223] discussed motion sickness in automated driving from a driver and a vehicle system perspective, as well countermeasures with an accent on a vehicle control based on a mathematical model of motion sickness. He also specifies the difference between drivers and passengers and the expansion of non-driving activities as plausible factors for increasing motion sickness susceptibility.

Similar to drivers of real-world vehicles, drivers of virtual vehicles have more extensive control over the vehicle than passengers. Dong, Yoshida, and Stoffregen [59] stated that these two types of vehicles are different regarding postural adjustments to motion. In virtual vehicles, the postural adjustment shows a tendency to reduce rather than to increase postural stability. It is considered that postural instability precedes motion sickness onset, but regarding SiS, discomfort can arise in the absence of postural instability. There is evidence-based research supporting the control of movements in virtual environments as a SiS related factor. A VR study by Stanney and Hash [205] compared three different levels of control: "active"—where the user had a full range of movements using a joystick; "passive"—where the user had no control; and "active-passive"—where the user had a limited range of movements. The results showed that the "active-passive" group felt significantly less SiS compared to the other two groups, and the "active" group felt less sickness than the "passive" group. It was implied that the full range of movements in the "active" group was more complicated and made the control more difficult. Thus, the adaptation to the virtual environment was slower than in the group with limited movements.

Moreover, in a similar manner to the rotating experiment of Rolnick and Lubow [184], Dong et al. [59] conducted an experiment to examine the effect of vehicle control within a virtual environment. One group of participants were assigned the role of drivers who could control the virtual vehicle during a driving simulator game displayed on a conventional flat-screen display, while the other group performed the role of passengers watching the recorded video of the drive. The results revealed that drivers were less likely to report SiS than passengers. These findings demonstrate that the more passive role of the passenger is more strongly associated with SiS onset in virtual vehicles as well as in physical vehicles. While some research has been carried out on the effect of vehicle control on SiS induced by virtual environments, no studies have been found which investigate this phenomenon in the context of fully automated modern HMD driving simulation that has been integrated with a moving platform.

It is a common assumption that women are more susceptible to SiS and experience more severe symptoms than men [72, 98, 127, 143, 212], although some research studies undermine this assertion [78, 128, 189, 229]. Possible reasons for this susceptibility could be women's hormone balance [40], their larger FOV [34], or their differences in motion perception [23]. Munafo et al. [152], for example, reported that the incidence of SiS was more common among female participants than male participants. However, the severity of SiS symptoms was not different between the genders. Moreover, Biocca [18] noted that men might tend to under-report their

symptoms, which could be a possible reason for the difference in the questionnaire scores measuring SiS across the two genders. Research on gender difference in SiS has not included a comparison between female and male participants in the context of automated VR driving simulation using modern HMD [72, 151]. Therefore, we investigated whether gender affects the incidence of SiS within an innovative VR framework.

In the literature, immersion often refers to a physiological state characterized by perceiving oneself to be enveloped by, included in, and interacting with a stimulating environment [233]. A sense of presence refers to experiencing the simulated environment rather than the actual (real-world) environment. A virtual environment that generates a higher level of immersion will create a greater sense of presence. The correlation between SiS in VR and the sense of presence appears to be complicated and indirect. A lower sense of presence may induce *disorientation*, which may increase the sickness onset [156]. A survey conducted by Schuemie and colleagues [192] has shown that the correlation between SiS and presence is controversial. However, another study found a positive correlation [136]. The study concluded that the relationship between SiS and presence might significantly change with different interactivity levels in the virtual environment. A recent review by Weech et al. [229] indicated that SiS is negatively related to the sense of presence. However, only three studies utilized a driving simulation environment. Two of them used a projection screen and the other a set of three conventional displays as a viewing system. Currently, there is no data on how or whether the sense of presence affects SiS in a VR driving simulation in the context of automated driving. The operationalization of a modern HMD with high graphics quality could increase the sense of being in the virtual world. Furthermore, a VR driving setup with an integrated moving platform could add a layer of immersion to the virtual environment as the body is physically moved. These two factors could contribute to an increased sense of presence, which in turn could lead to a reduction in SiS onset. Thus, we evaluated the sense of presence within an innovative VR driving setup and its relationship to SiS.

To assess SiS, not only subjective factors but also objective factors such as physiological factors are monitored. Physiological signals can serve as indicators of SiS occurrence [74, 174, 229]. Earlier research showed that SiS is significantly positively correlated with physiological signals such as heart signal [51, 120, 194] and sweating [51, 226], and breathing [120]. However, caution should be used while using physiological responses as an assessment tool. The physiological signals' response might be easily influenced by other factors such as excitement, anticipation, and boredom, all of which could be plausibly met in virtual environments. As an objective measurement, therefore, physiological factors should only be used in conjunction with subjective measures such as questionnaires in order to form a complete assessment of SiS. Thus, their integration into the driving setup and synchronization with the driving simulation is vital to the development and evaluation of the innovative VR driving setup.

Due to the considerable influence of SiS on an individual's physical and mental state during VR sessions, it is necessary to determine how to predict these symptoms. Prediction models of SiS were previously described in the literature [104, 126, 175]. Nevertheless, conventional statistical

methods on individual factors, primarily excluding physiological signals as predictors, were used for creating the models. Furthermore, the data utilized for building the models, was obtained through diverse virtual simulations other than virtual driving. Recent studies have started to use machine learning models for the prediction of SiS in virtual environments [86, 96, 118, 142]. From these studies, only one work used physiological data as a predictor [142]. The best results, though, were obtained through a combination of individual and task-related factors by Jin et al. [96]. Altogether these studies emphasize that prediction models which foresee the outbreak of SiS are still not thoroughly developed, particularly for HMD driving simulations. More focused data collected through an HMD driving simulation will contribute to more accurate SiS outbreak prediction induced by such driving setups. The limited number of recent studies on models based on objective data offers an opportunity to develop a prediction model to classify simulation sick users. Thus, we present separate models using individual factors as well as sophisticated machine learning methods on physiological and behavioral characteristics for SiS recognition in the context of VR driving simulation.

To summarise, in this thesis, we focus on the challenging problem of SiS induced by VR automated driving simulation in an urban driving scenario. The thesis concentrates on the prediction of SiS based on individual, behavioral, and physiological factors. Furthermore, an innovative driving setup, including a modern HMD, integrated moving platform, and physiological sensors, is developed for the evaluation of SiS onset. Additionally, a standard driving scenario is compared to an automated driving scenario using static and dynamic VR driving.

1.2 Research Objectives

A VR driving simulation with fewer discomfort outbreaks will encourage an uninterrupted evaluation process with user compliance. Some projects may require shorter sessions, while others may have a longer duration until the performed tasks are achieved. This will facilitate the exploration of the options provided by new products for comprehensive data collection and early error detection. The absence of SiS would enable users to enjoy VR without unwanted feelings of discomfort such as nausea, disorientation, and headaches. Another benefit would be that users could become more receptive to repeatedly using VR if there was no fear of the after-effects negatively affecting their daily tasks.

In order to improve user experience in the virtual world, the following research objective of this thesis should be fulfilled:

- **Integration of moving platform and physiological sensors into VR driving setup**
  A compact VR driving simulation setup contains an HMD, a computer, and a video controller. A more sophisticated setup contains a set of steering wheel and pedals instead of a video controller. In some cases, such a driving setup might include a driver seat. This

compact VR system is easy to set up and transport, so it is very often used in VR studies regarding SiS. However, these VR systems lack physical motion cues, which provide the necessary feedback to signal when the virtual car stops or turns. The integration of motion cues through the moving platform should be carried out accurately as a larger discrepancy between the visual and motion cues could provoke SiS onset. Previous studies operationalized driving setups, including one or more conventional displays, in order to present the virtual world. Nevertheless, with the comeback of HMDs in recent years and rapidly increasing usage, enhancing the driving setups viewing system is a step towards a more immersive driving experience. When the user turns on the modern HMD on, he or she is immersed in the virtual world and can see only what is happening there. In order to simulate a car's movement during the virtual drive, we integrated a motion platform into an HMD driving setup. The moving platform could move in three different directions synchronized to the movements of the virtual car.

Furthermore, we integrated several physiological sensors to record the user's bodily responses to the virtual drive. The bodily responses could be used for the detection of SiS during the driving simulation. To build a prediction model of SiS, we first had to create a dataset from the recorded physiological signals. The dataset should include the exact signals' values, and thus, the synchronization between the signals and the virtual drive is essential. In the literature, physiological signals are recorded and used as an objective measure for SiS. Some physiological factors, such as heart rate, breathing, and sweating, could be affected by the experienced discomfort. However, some studies failed to report a significant relationship between these signals and the onset of SiS. A combination of subjective and objective measures could provide more insights into SiS than only using one of these measures. Thus, integrating physiological signals into the VR setup ensured the setup was an effective evaluation tool for assessing SiS.

While similar driving simulation setups exist in previous work, this is the first time that a VR setup combines a modern HMD, a moving platform, and physiological sensors in one synchronized driving simulation. SiS can be readily assessed subjectively and objectively during the virtual drive, providing the researcher with the necessary data to evaluate the virtual experience.

- **Evaluation of SiS factors in standard and automated dynamic VR driving setup**
  The popularity of VR technology can contribute to using a compact VR driving setting that can be easily set up and transported for automotive user interaction research. The enhancement of such setups through the integration of a moving platform, as described in the previous section, could bring an additional layer of engagement and make the simulation more realistic. However, it is questionable whether these moving platforms actually reduce SiS onset or merely improve levels of immersion through feelings of interactivity and excitement. As automated driving is considered the future of transportation, a driving simulation that imitates real-life automated driving could be a perfect testing or training

tool. Nonetheless, virtual automated driving could potentially inherit one of the common problems of real-life automated driving, namely, motion sickness.

Although SiS is a primarily visually induced sickness, a similar discomfort to that felt by passengers of a real-life vehicle could occur. Little is known of the effect of automated driving on SiS, and the relevant body of literature speculates that this type of driving evokes more discomfort than standard driving in a VR environment due to the loss of control over the vehicle. Furthermore, the operationalization of an HMD could increase the chance of SiS outbreak as this type of viewing system is considered to induce more SiS than setups with conventional displays. Thus, we chose to evaluate SiS with a dynamic VR setup using fully automated driving simulation. The aggregation of automated driving in a highly immersive virtual world which simulates the vehicle's movements provides new insights into user well-being, both during and after the simulation.

In general, women have been reported to be more susceptible to SiS and to feel greater discomfort than men. However, some researchers found no evidence to support this assumption and concluded that there is no difference in symptoms between genders. Other individual factors, in combination with the user's gender, might contribute more to the SiS outbreak than gender on its own. This issue is further complicated by the fact that more male participants have been recruited than female participants to engage in previous test studies. This might lead to an unbalanced sample size, which excludes gender as a factor from the data analysis. The knowledge regarding gender and SiS in VR driving simulation is primarily based on previously conducted studies on motion sickness and SiS in conventional simulators. Susceptibility differences across genders have not been broadly studied in the framework of modern HMDs, and there are even fewer studies concerning driving simulations with HMDs. Additionally, the driver's changing role in automated driving could also affect SiS across genders. Therefore, we evaluated gender differences within automated virtual driving simulation on a dynamic HMD setup. In order to establish a comparison point with automated driving and investigate whether the automated type of driving evokes more discomfort, we compared standard and automated driving VR setup as well as dynamic versus static standard driving VR setup.

- **Relationships exploration between SiS-inducing factors and SiS**
  Some of these factors, however, are assumed to be related to SiS, and evidence of this relationship has been obtained from previous research on motion sickness or SiS induced by conventional display settings. With the operationalization of modern HMD as a viewing system in VR driving setups, the relationship between SiS factors and SiS might change as the hardware characteristics change. We explore the connection between SiS factors such as *motion sickness history, age, sense of presence, gender*, and *sleep deprivation* as well as physiological factors, and SiS to determine whether these factors are still relevant to the phenomenon in the context of VR driving simulations. Furthermore, a few factors, such as *physical activity, mood,* and *personal traits*, of which little is known regarding their

correlation to SiS in VR driving environments, are explored as well.

- **Prediction of SiS induced by VR driving simulation**
  The prediction of SiS is challenging as determining the exact moment of onset is a largely subjective exercise. Some individuals become aware of the discomfort when it is at its highest point of severity. Others under-report their symptoms as they fail to recognize or do not wish to admit that they felt discomfort. Another complicating factor is that SiS is mostly assessed after the VR simulation and less so during the simulation. Some studies developed statistical models to predict this phenomenon based on collected data through user evaluation studies. More recent work has started to include and build prediction models on physiological data. However, this study used numerous factors as predictors that could be difficult to implement in a more practical setting, such as a VR environment within the automotive industry. Moreover, the data used in these models is not derived from a virtual driving environment. In order to develop an accurate model for predicting SiS induced by HMD driving simulation, the data used for this model should be collected from such driving settings. That way, the model would be more accurate and applied to a similar setting by expecting a similar prediction rate. Although some prediction models have been developed, only one study has attempted to build an SiS prediction model based solely on physiological data. Thus, we develop models to predict SiS in the context of modern HMD driving simulation setup using separately individual, behavioral and physiological, and only physiological factors.

1.3 Outline of the Thesis

This thesis contains eleven chapters. The Theoretical Background section provides a fundamental introduction to the phenomenon of SiS, including a description of SiS-inducing factors. The Related Work Work section discusses the body of literature related to SiS induced by VR driving simulations. A description of the applied methodology, including the utilized measurements, is provided by the Methodology. The Technical Realization section addresses the challenge of integrating a moving platform and physiological sensors into the VR driving system. The experimental section of this thesis is reported across three chapters divided according to different types of driving. The Prediction Models present the developed models for predicting SiS based on individual and physiological factors. The Discussion reviews the reported results in the context of state-of-the-art research. Lastly, Contribution and Conclusion summarizes the research contributions of this thesis and suggests directions for future work.

Figure 1.1 illustrates a visual outline of the thesis.

This thesis is outlined as follows:

- **Chapter 2** lays out the theoretical background of SiS and briefly describes the human

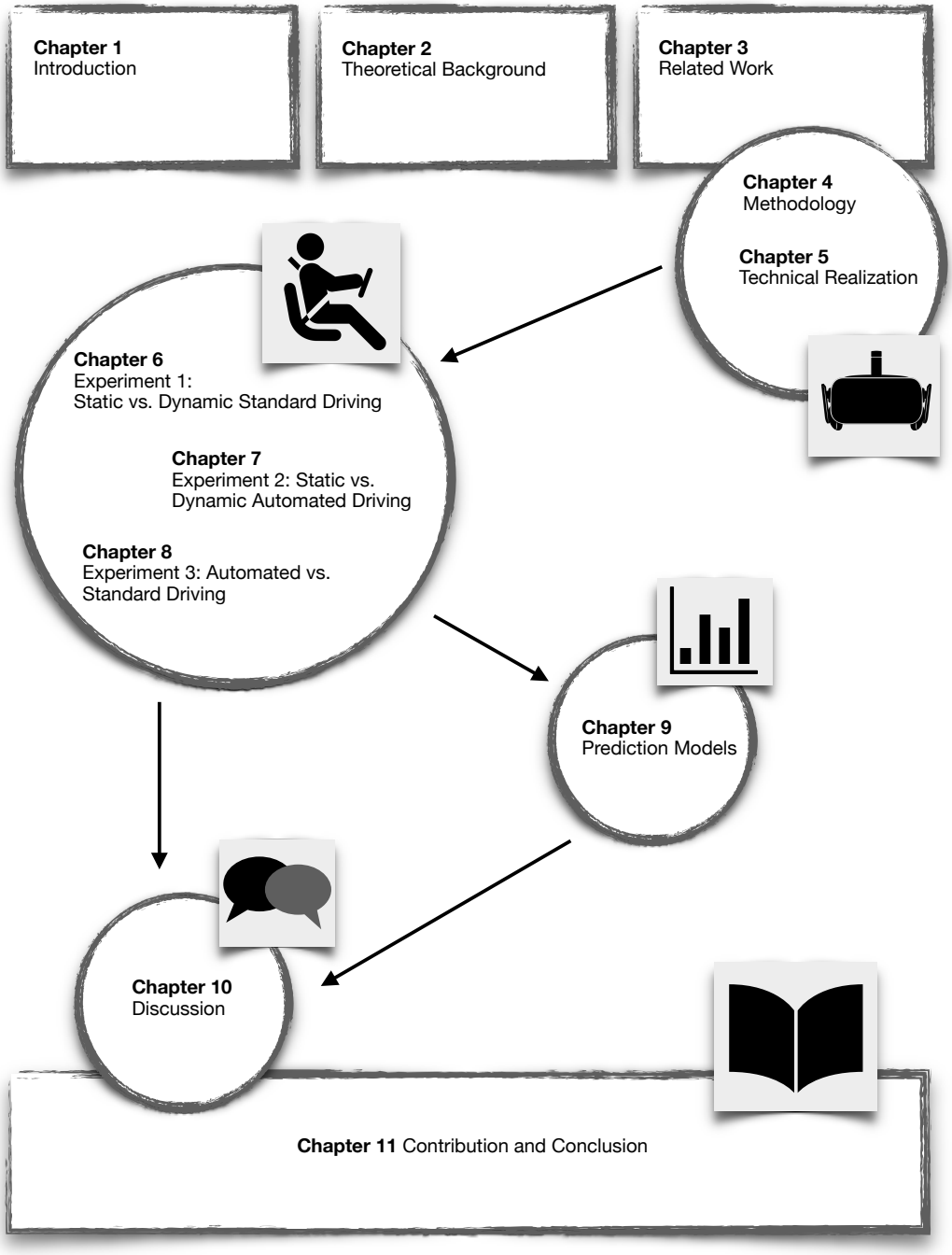


Figure 1.1: Thesis outline.

sensory systems and VR. For a better understanding of the physiological mechanisms underlying SiS onset, a basic knowledge of anatomy and how the visual, vestibular, and proprioceptive systems work is essential. The fundamental theory underlying our experience of SiS is described to outline the malaise's background and possible origin. Furthermore, to explain the possible relationship between SiS-inducing factors and SiS, an extended list of factors is presented. Each factor is discussed in relation to SiS being induced by a VR driving simulation.

- **Chapter 3** is concerned with previous works on SiS and, in particular, SiS induced by HMD driving simulations. The body of research on SiS is substantial, and it has been growing in recent years. However, the number of studies on SiS within a VR driving simulation with a modern HMD is still insufficient. Based on previously reported evidence, a list of mitigation techniques separated into three groups is described. Each technique is reviewed in terms of its applicability to HMD driving simulations. Furthermore, prediction models from previous works on SiS in virtual environments are presented. These findings guide the evaluation of SiS-related factors and the prediction of the phenomenon.

- **Chapter 4** presents the design of user evaluations, the used measurements, as well as the procedure of the experiments. In order to fulfill the research objectives, a factorial design for the experiments was chosen. Two user studies were conducted to test automated and standard VR driving simulations. Nevertheless, three experiments were carried out of these studies. The third experiment was used to evaluate the differences between the two types of driving environments. Additionally, three types of measurements were used before, during, and after the experiments. Questionnaires were chosen based on their previous validation and relevance to the objective of the thesis. Additional questions were included to more effectively tailor the questionnaires for better for the extraction of the required data. A primary aspect of choosing the right physiological sensors for the VR setup was the wireless connectivity between the sensors and the computer. In that way, the user's movement was not completely limited by the sensors' cables. Therefore, wireless sensors were used for physiological data collection. They were also easy to attach to the user's body and explain their purpose as part of the experimental procedure. An overview of the procedure is given to illustrate the order of the steps taken during the experiments.

- **Chapter 5** describes the integration of a moving platform and physiological sensors into the VR driving simulation setup. The moving platform was integrated into the setup by colleagues at BMW Group, while an external company conducted the software implementation. Therefore, only an overview of the integration and the main technical specifications are presented in the thesis. Then, the integration of the physiological sensors into the dynamic VR setup and their synchronization is described. Furthermore, two pilot studies were conducted to establish whether the setup is usable and technically ready for further user evaluations.

- **Chapter 6** evaluates the effect of *motion* and *gender* on SiS, physiological factors, and *sense of presence* in a standard VR driving simulation. We conducted an experiment to test how *gender* and two different driving settings, static and dynamic, affect SiS within a standard driving simulation in a virtual urban environment. As a relevant and vital topic in the automotive industry, we carried out a driving experiment operationalizing the innovative HMD driving setup described in the previous chapter. The current experiment allowed the participants to steer the virtual car following a pre-designed route. The SiS onset was assessed during the VR simulation, immediately after, and approximately one hour later. For a driving scenario, an urban low traffic scenario was used. Furthermore, we explored possible relationships between several individual factors such as *personality traits, age, gender, motion sickness history, sleep deprivation, sense of presence*, and *mood*, as well as physiological factors and SiS.

- **Chapter 7** investigates the effect of *motion* and *gender* on SiS, physiological factors, and *sense of presence* in an automated VR driving simulation. The same driving scenario and VR system as in the standard driving experiment were operationalized. The main difference between this and the previous experiment was the participant's involvement in the actual driving. In the current experiment, the vehicle was fully automated and no user interaction was required. The road was marked with arrows, which showed the path of the vehicle to the participants. A static automated driving condition was compared with a dynamic automated driving condition in order to test whether the addition of motion cues affects SiS or not. The SiS assessment and VR driving scenario were the same as in the previous experiment. Additionally, we explored possible relationships between the same individual factors listed in **Chapter 6** and SiS.

- **Chapter 8** compares the effect of the two different *types of driving*, explored in the previous two chapters, *motion* and *gender* on SiS, physiological factors, and *sense of presence* in a VR driving simulation. As one of the SiS-inducing factors, vehicle control was evaluated in the context of the innovative VR driving setup. We carried out a third experiment utilizing the data collected from the previous two experiments. The experiment included only the participants who took part in either automated or standard driving conditions. Thus, a possible effect of adaptation was limited. The SiS assessment and VR driving scenario was the same as in the two previous experiments. Additionally, we explored possible relationships between SiS and the same individual factors listed in **Chapter 6** and **Chapter 7**.

- **Chapter 9** presents prediction models developed to predict an individual's probability of experiencing SiS. The linear models predict the SiS, including only individual factors, and the machine learning models predict SiS, including behavioral and physiological factors.

- **Chapter 10** presents an overview of the results regarding hypothesis testing, relationships, and prediction models. A discussion of these results is presented concerning the previously

conducted study and the theoretical background of the SiS. Furthermore, the significant relationships between SiS and the other variables are discussed.

- **Chapter 11** concludes with an overview of the most significant contributions made by this thesis and outlines possible directions for future research.

2 Theoretical Background

> What is real?
> How do you define real?
> If real is what you can feel, smell, taste and see, then real is simply electrical signals interpreted by your brain.
>
> *Morpheus*, "The Matrix"

The main objective of this thesis is to evaluate several SiS-related factors and to develop a reliable model for predicting SiS onset in VR driving simulation. Before evaluating SiS, we should understand it as a phenomenon and the factors related to its onset. Therefore, this chapter lays out the theoretical background of the thesis. It provides an overview of theories, factors, and measurements of SiS. To better understand the possible origin of this malaise, a brief description of the most related human sensory systems, such as the visual, vestibular, and proprioceptive systems, is given. Next, the most common theories of SiS are introduced along with a list of possible causative factors that are divided into three main groups: individual, hardware, and simulation. Then measurements commonly used to assess SiS–including objective measures such as physiological signals and subjective measures such as questionnaires—-are outlined.

In the literature, depending on the setting, SiS induced by visual stimuli is variously referred to as cybersickness, virtual reality sickness, gaming sickness, and visually induced motion sickness. As SiS varies among individuals, the symptoms can vary as well. However, the symptoms can be narrowed down to three main groups as follows [101]. The first group is *Nausea* and includes symptoms such as stomach awareness, burping, and increased salivation. The second group, *Oculomotor*, includes symptoms such as eye strain, headache, difficulty focusing, and blurred vision. The last group is *Disorientation* and includes symptoms such as general discomfort, vertigo, and dizziness. Some of the symptoms are no longer considered relevant and therefore

new symptoms have been added to the list with each passing year as virtual reality becomes more commonplace and better understood. Nevertheless, the three main groups have remained broadly the same. SiS is considered a subset of motion sickness, and therefore many SiS symptoms are the same as those of motion sickness. However, there are also some differences. Motion sickness symptoms are mostly gastrointestinal (e.g., stomach awareness, nausea, and burping) whereas SiS symptoms are more related to oculomotor problems such as headaches, eye strain, and blurred vision [105]. Also, SiS symptoms very rarely escalate to emesis [103]. In the next subsection, we will look at a brief overview of human sensory systems to further understand this phenomenon.

2.1 Human Sensory Systems

We receive information about our surrounding environment through the sensory systems spread over our bodies. Our hearing provides 20% of this information, followed by smell (5%), touch (4%), and taste (1%) [85]. All these senses work together to not only bring realism to the virtual world but also to alleviate its side effects. When the sensitive balance between these senses is disrupted by VR, though, we can experience SiS. The primary sensory systems that contribute to the onset of SiS in VR are the visual, the vestibular, and the proprioceptive systems. Basic knowledge of these systems is required in order to better understand the background of SiS.

Visual System

The human visual system is one of the primary information input systems in the human body and provides the most information (70%) to the brain [85]. It includes the eyes and the pathways connecting to the visual cortex and other parts of the brain [79]. The human eye consists of three layers (i.e., outer, middle, and inner) and has a highly complex anatomical structure.

There are four main eye movements: saccades, vergence, pursuit, and vestibular ocular reflex (VOR) [166]. The saccades are rapid movements of the eye that instantly change the point of fixation. This ranges from small movements such as reading a sentence to much larger movements such as gazing around a room. The pursuit movements or smooth pursuit movements are much slower tracking movements of the eyes that keep a moving stimulus on the fovea [166]. Compared to the saccades, which can be made voluntarily and unconsciously, the pursuit movements are made mostly voluntarily. This is, a person can choose whether or not to track a moving stimulus. Vergence is eye movements that align each eye with targets placed at a different distance from the eyes. Vergence movements are disconjugate. In other words, they involve either convergence or disvergence of the line between an individual's eye and an object that is nearer or farther away. The VOR movements compensate for the head movements relative to the external world by stabilizing the eyes [166]. More about VOR is described in the subsection Vestibular System.

Vestibular System

The vestibular system is a sensory organ located in the labyrinth of the inner ear and which is responsible for the body's equilibrium and spatial orientation—both of which allow the body to coordinate movements and maintain balance (Fig. 2.1). The vestibular system consists of three semicircular canals and the otolith organs (the utricle and the saccule) placed inside the vestibule [154]. As described by Gray [79], the semicircular canals are located above and behind the vestibule, and they are three in number: anterior, posterior, and lateral. The semicircular canals are a series of differently lengthed looped tubes positioned at right angles to one another. They are responsible for rotary movements. Each canal presents an enlargement at one end, called ampulla. Endolymph flows through the canals and moves in a direction dependent on the head position. Within the vestibule, the membranous labyrinth does not entirely preserve the form of the bony cavity but rather consists of two membranous sacs, the utricle and the saccule, also called the otolith organs. The utricle, the larger of the two, is placed at the upper and back part of the vestibule. The saccule lies near the opening of the vestibular duct of the cochlea. These two organs are responsible for linear acceleration in the body [79].

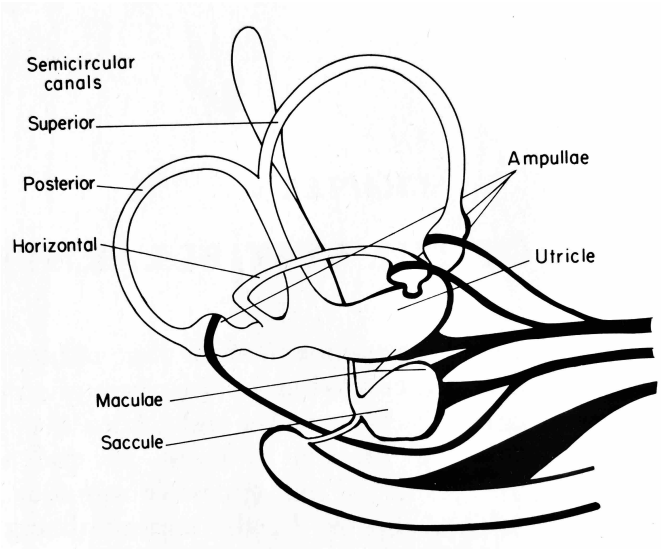


Figure 2.1: The vestibular system - semicircular canals and otolith organs (from Reason and Brand [172]).

The information from the vestibular system is transferred to the brain via the vestibular nerve [76]. The vestibular system is the only organ of balance, and therefore it is extremely sensitive to any kind of mismatch with the other sensory systems. A significant eye movement, coordinated with the vestibular system, which is directly related to the visual perceptions in VR, is the VOR.

The VOR is a reflex of the eye movement that stabilizes the retina's images when the head is moving. The reflex produces an eye movement in the direction opposite to the head movement in response to neural input from the vestibular system; thus, it maintains the image in the center of the visual field [65]. For example, when the head rotates rightward, the eyes move leftward. This

process applies for head movements such as right and left, up and down, and tilt to the right and left. All these movements provide information to maintain the ocular muscles' visual stability. The VOR helps to be created new neuron paths that assist the adaptation to a new environment. Moreover, a temporal modification of the VOR gain can be induced by virtual environments [55]. Draper [60] listed major virtual factors that might affect VOR and SiS, each of which should be studied in SiS research. The factors are time delays/update rate, tracking (position) inaccuracies, image quality/resolution, FOV, the weight of the HMD, and accommodation/vergence cues. Most of these are among the core list of SiS-related factors, which is presented later in this chapter.

Proprioceptive System

Proprioception is the perception of stimuli by individuals relating to their position, posture, equilibrium, or internal condition [29]. The receptors involved in this process provide information about mechanical forces arising from the body itself and are known as proprioceptors. The proprioceptors' purpose is to give continuous information about the position of the limbs and other body parts in space. Kinesthesia is another part of the proprioceptive system and is often used in regards to proprioception. The differentiation is that the kinesthesia is responsible for the muscle movements. Proprioception is cognitive, while kinesthesia is behavioral. For example, individuals with vestibular dysfunction can walk using only their sense of sight to maintain balance. But at the moment they close their eyes, they are unable to walk straight anymore.

Rupert and Kolev [186] found that tactile stimuli can alter motion perception. This alteration could lead to a reduction of SiS in static and motion simulators. Another study reported that noisy vestibular stimulation, when coupled with angular accelerations of the camera, reduced SiS onset in a VR application using two different setups [228]. Furthermore, it was found that proprioceptive vibrations affected SiS and reduced the feeling of discomfort by approximately 47% during a navigation task in VR [164]. Together these findings show that the stimulation of the proprioceptive system might reduce the onset of SiS in virtual environments. However, adverse results can be found in the literature as well. For example, it may be the case that tactile stimulation only minimizes the SiS due to its distraction effect [71].

2.2 Simulation Sickness Theories

Cue Conflict Theory

Cue conflict theory, also known as sensory conflict theory, is the most well-known theory of SiS. Originally developed to explain motion sickness occurrence [172], it was later discovered that it also applies to SiS. It describes how a so-called "mismatch" can arise between two or more human sensory systems, which then includes a perceived increase of discomfort as a response

from the body. For example, when an individual uses a static driving simulator, a mismatch between the motion perception of the human body through the vestibular cues and the visually perceived information through the visual cues, can arise. This mismatch can then result in SiS onset [215]. Nevertheless, SiS can also appear when using a motion driving simulator. A mismatch between the motion cues and visual cues might induce discomfort as the physical motion can occur with a slight delay from the visually perceived motion. Kolasinski [127] pointed out that vagueness in visual, vestibular, and proprioceptive cues may be created in VR to represent motion. The reason for this is that these sensory systems provide visual cues inconsistently with self-motion as opposed to matching vestibular cues. In VR vestibular cues and peripheral vision are especially significant for particular orientation and detection of self-motion. Furthermore, the conflict between the sensory systems is commonly used to clarify the connection between *vection*, the perception of the world moving away from an individual while one is still, and SiS. As an SiS-inducing factor, *vection* has been well researched and is described in more detail later in this chapter.

A variation of the cue conflict theory is the sensory rearrangement theory [172]. Reason [171] noted that this theory variation is developed based on two assumptions. The first assumption is that to provoke discomfort, a sensory rearrangement in some form between visual, vestibular, and proprioceptive cues should exist. Thus, it is expected that a discomfort should arise from the conflict between the current experience and that stored from previous experience. In other words, an individual can become sick due to inconsistency between the expected and the present experience. The second assumption is that regardless to what extent the other sensory systems are involved in the conflict, the vestibular system must be involved, either directly or indirectly. Considering the sensory rearrangement theory, repeated exposure to a given environment can help an individual to adapt to that environment. However, the duration of such a process can vary as it is strongly individual. For example, when an individual uses a driving simulator, she or he might experience SiS due to the difference between her previous driving experience and the present driving simulated experience. However, the body might adapt very fast to the new experience, and thus, the onset of SiS would not last long during the simulation. The rate of adaptation varies across individuals, though, and can range from a few minutes to a few sessions; or, in some cases, individuals never adapt to the simulated environment.

Although the cue conflict theory is widely used to explain the SiS onset, it suffers from some potential shortcomings. First, the fundamental physiological processes underlying the theory have not been fully identified. Second, fabrication of the cue conflict theory is challenging since this theory can explain almost every stimulus that induces motion sickness afterward, making for few testable predictions. Third, many researchers have noted the low predictive validity of the sensory mismatch idea. Although it can explain why a given stimulus can be problematic, it is difficult to predict the frequency and the severity of discomfort for a given stimulus regarding a single individual or a group. A widely accepted and unified theory on SiS should be able to make quantitative predictions regarding the average group of people when the given conditions

are known [113]. Given these shortcomings, some researchers oppose cue conflict theory and promote different theories of motion sickness and SiS.

Postural Instability Theory

One such theory is the postural instability theory put forward by Riccio and Stoffregen [182], which look at motion sickness and SiS from an ecological perspective. Also known as the ecological theory of motion sickness, the postural instability theory states that a disturbance in human postural balance leads to motion sickness onset. Postural stability is defined as "the state in which uncontrolled movements of the perception and action systems are minimized" [182]. Every new environment, such as a virtual environment, presents the individual with the challenge of maintaining balance within it [61]. Regarding VR, Reed-Jones et al. [176] stated that the postural balance is maintained by adjusting the body movements in response to simulated motion stimuli, and not in response to the gravitational body position. Consequently, body adjustment is incorrect and results in postural instability. The SiS symptoms arise due to the continuous adjustment of the body posture in the virtual environment.

The relationship between postural stability and SiS in virtual environments has been previously documented [1, 176, 199, 214, 221]. Furthermore, Arcioni et al. [7] reported that SiS could be predicted from individual differences in postural activity in an HMD simulation. In another study, the participants who performed more head movements during video game play experienced more SiS [147]. However, postural sway does not always precede SiS in VR. A recent study on SiS observed no connection between SiS and postural instability [54]. In other words, SiS can arise in the absence of postural instability. The participants who experienced more discomfort presented less variation in postural sway.

Postural instability theory suffers from shortcomings regarding its assumption, measures of postural sway, and co-founding factors. For example, according to Keshavarz et al. [113], one apparent inconsistency is that restraining a person's body should always reduce SiS onset. The evidence related to this in the literature, however, is inconsistent. A few studies reported that SiS is reduced by body movement restrictions [21, 37, 116], while other studies reported contrary results [227]. As shown above, there are several conventional theories of SiS. Despite the drawbacks of cue conflict theory, it is still one of the most commonly applied theories of SiS. The theory can be used under various conditions, which gives it the strength of a complete theoretical framework—and on that basis, it is used in this research. Thus, this thesis will primarily concentrate on SiS caused by conflicting cues.

2.3 Simulation Sickness Related Factors

SiS is associated with many factors that are involved in its onset (i.e., polygenic). Factors that can induce SiS can be grouped into three categories. The first category involves individual factors, which comprise gender, age, experience, and illness. The second category is simulator-related factors, which comprise lag, flicker, calibration, and motion platform. The last category is the task-related factors, which comprise factors control, duration, and scene complexity [127].

Rebenitsch [175] described numerous factors based on Kolasinski's [127] report on SiS in virtual environments (Table 2.1). Only 37 factors that are considered to have a relationship with SiS in VR driving simulations are described in this work. The following factors are excluded: eye dominance, stereoscopic visual ability, global visual flow, hats, mental rotation ability, screen distance to the eye, tracking method, ambient temperature, olfactory feedback, audio feedback, altitude above the terrain, rate of linear or rotational acceleration, orientation cues, and the ratio of virtual to the real world. The factors body mass index, history of headache/migraine, and state of health are combined into one factor—state of health. The factors motion platform and haptic feedback are combined into one factor—haptic feedback. Additionally, latency is included as a factor from the Hardware group, and personal traits are included as a separate factor.

Table 2.1: Potential factors related to SiS in virtual environments described in this thesis (adapted from Kolasinski [127] and Rebenitsch [175]).

Individual	**Hardware**	**Simulation**
Motion sickness history	Display type	Degree of control
State of health	Resolution/blur	Vection
Age	Field-of-view	Self-movement speed
Gender	Ergonomics	Contrast
Vision correction	Inter-pupillary distance	Luminance level
Personality traits	Update rate	Color
Emotions	Latency	
Experience with VR systems (Adaptation)	Calibration	Scene content/complexity
Experience with real-world task	Position tracking error	Independent visual background
Duration	Head movements	
Experience with video games	Haptic feedback	
Perceptual style		

The use of VR systems, like every simulated environment, is also associated with SiS onset. VR is not a new technology; therefore, some known aspects and effects of SiS related to the virtual environment have been previously researched. Many of the SiS factors (e.g., gender, motion

sickness history, and duration), which are listed for simulated environments, also apply to VR [52, 127, 131, 161]. Symptoms of general discomfort, eye strain, and difficulty concentrating, part of the Oculomotor discomfort group, are more likely to be experienced while an individual is using the fully immersive virtual environment [97]. For example, eye strain could be experienced due to the close distance between the eyes and the HMD's screen. Consequently, this could also induce a headache. Such Oculomotor symptoms are the primary difference between SiS and motion sickness (which is more inclined towards Nausea symptoms) [107, 218].

2.3.1 Individual Factors

Motion Sickness History

One of the SiS factors considered to be a predictor of SiS onset is the susceptibility to motion sickness. The individual's history of motion sickness is usually evaluated through a Motion Sickness Questionnaire (MSQ) [172], which measures the susceptibility level during childhood and adulthood. There is a shorter, updated version of the MSQ developed by Golding [77] called MSSQ-short, which has been used more commonly in recent years. The questionnaire is described in detail later in **Chapter 4.2.3**. Several studies have found a positive relationship between motion sickness history and SiS onset [26, 143, 173, 212]. If an individual had motion sickness episodes in the past, e.g., as a child during a boat trip, this could lead to SiS while using a simulator as an adult. It is assumed that a similar relationship exists between motion sickness history and SiS induced by an HMD driving system. However, no work on this relationship in the context of modern HMDs and urban driving scenarios was found. Thus, in this thesis, the association between the history of motion sickness and SiS induced by an HMD driving simulation will be evaluated.

State of Health

An individual's health condition is reported to be one of the factors that predicts a high probability of nausea while using a simulated environment. For example, if the participant has flu or other medical conditions that require medications that could restrict daily tasks, the overall state of health is not normal. This can raise the risk of SiS. Furthermore, a reduced sleeping condition named insomnia could be positively related to the onset of SiS. In a military study, it was reported that SiS symptoms increased with the users' reduced sleeping hours [98].

Another factor related to the users' health is the Body Mass Index (BMI), which is an indicator of a person's overall health level. It is calculated as dividing the weight by the squared height. BMI is measured in kg/m2 as a result of dividing the mass (kg) by height (m2). The BMI varies across individuals and is measured according to several categories. It is considered that a normal BMI is between 18.5 and 25 [70]. In an experimental study regarding human factors in the virtual

environment, the BMI appeared to have a significant negative relation to Oculomotor symptoms [212]. Moreover, those of the participants with higher BMI showed a lesser emetic response.

Physical activity is another factor that could be categorized under the term of state of health. A positive correlation was found between physical activity and user performance in VR [16]. Sharma [195] reported an association between physical activity and motion sickness. Among individuals whose professions involved great physical activities, the prevalence of motion sickness was none or lower than for the general population. It might be assumed that similar to motion sickness; there is a correlation between physical activity and SiS. It is a requirement to be in good health to participate in SiS studies, but it was not tested whether the level of physical activity could affect SiS. Therefore, in this work, the possible relationship between sleep deprivation, physical activity, and SiS is investigated.

Age

The connection between age and SiS onset is assumed to be the same as the relationship between age and motion sickness. Reason and Brand [172] reported that the highest level of motion sickness arises between the ages of 2 and 12. After this age, motion sickness gradually reduces through adulthood until the age of 50. Above this range of age, motion sickness is very rare. However, several studies have explicitly explored the relationship between age and SiS induced by driving simulators. It was observed that older participants are more prone to SiS symptoms during conventional driving simulator studies than the younger participants [30, 100, 108, 157].

Arns and Cerney [8] observed a pattern of SiS induced by a CAVE-like VR environment that seemed to contrast with the common understanding of the relationship between motion sickness and age; that is, older participants experienced more SiS than younger participants. The SiS onset was more likely to increase with age. Additionally, the average total sickness score was higher for participants 50 years and older than for younger participants. A few years earlier, Liu and colleagues [140] carried out an experiment on quantitative measures of driving performance, and participants reported qualitative measures including SiS across different age groups in a VR driving environment using VR glasses. The results showed that the amount and severity of SiS symptoms increased significantly with age. Nevertheless, symptoms reported by the older group ($56+$ years) were not greater than those reported by the middle-aged group ($36-55$ years). These findings point out that despite the similarities between the symptomatology of motion sickness and SiS, the two malaises may have a different relationship with the sickness-inducing factors. Thus, the direct application of motion sickness research data on SiS might sometimes be inaccurate.

Gender

In the literature, it is common to assume that women are more susceptible to SiS and experience more severe SiS symptoms than men [62, 72, 98, 127, 143, 212]. However, contrary research results exist as well [78, 121, 189, 229]. Possible reasons for this susceptibility could be women's hormone balance [40], their wider FOV [34], or their differences in motion perception [23]. Munafo et al. [152], for example, reported that the incidence of SiS was more common among female participants than male participants. However, the severity of SiS symptoms was not different between the genders. Moreover, Biocca [18] noted that men might tend to under-report their symptoms, which could be a possible reason for the difference in the questionnaire scores measuring SiS across the two genders. In psychological terms, women and men have different perceptions about driving a vehicle [75].

Stanney et al. [212] reported that female participants experienced more Oculomotor and Disorientation symptoms than male participants. These results could be related to the slightly wider FOV in women [34]. The wider FOV might provoke more discomfort in women as previous research showed that simulators with a wider FOV induce more SiS than simulators with a narrow FOV due to the more extensive perception of visual signals [107]. The FOV is listed as one of the SiS-related hardware factors later in this chapter. Furthermore, Czerwinski et al. [50] pointed out that women benefit more from displays with a wider FOV as they tend to navigate through the virtual world using landmarks. A display with a wider FOV can show more landmarks at once. The women's navigation preference might be linked to the wider FOV they exhibit [34].

Boyd [23] suggested that the gender differences are mainly in the perception of depth. She stated that women biologically rely more on shape from shading perception, and men rely on motion parallax perception regarding depth perception. Motion parallax is the visual effect when an individual watches moving objects from a range of different distances: the objects that are closer to the viewpoint seem to move faster than the objects that are farther away. However, in reality, all the objects move at the same speed. Shape from shading is a technique used for visual perception when shading gives the illusion of depth, e.g., a shading gives information to the brain about the distance between the object and the eyes [167]. Boyd [23] explained that it is easier to create a depth perception through motion parallax in virtual worlds than through shape from shading technique. As the creation of sufficiently realistic shading within the virtual environment is more computationally demanding, the other technique, motion parallax, is preferred for recreating depth virtually. Therefore, unintentionally, virtual environments are more welcoming to male users than female users in terms of felt discomfort. The choice of which depth technique to use, however, might change with the fast-developing hardware (e.g., HMDs, CPUs, GPUs) and matching software (e.g., game engines using a global illumination).

Another HMD specification is discovered to have a different effect on female and male participants regarding SiS. Stanney et al. [213] found that the inter-pupillary distance (later explained in a subsection) adjustment, while using an HMD, can significantly impact SiS onset. Female

participants experienced more SiS when the inter-pupillary distance of the HMD did not fit well to their eyes. However, the motion sickness susceptibility was another factor contributing to the higher SiS levels in female participants. Nevertheless, the study showed that the consumer HMD hardware did not have an appropriate display adjustment for a considerable proportion of female participants.

A recent study, conducted by Nickkar et al. [157], showed that female participants experienced more headache and nausea symptoms than male participants. The correlation between gender and SiS is an interesting window into how the physiology of the users affects the VR experience.

Vision Correction

Vision correction can result from wearing glasses or contact lenses. Individuals who wear glasses while they are wearing an HMD can experience ergonomic discomfort. Moreover, the glasses could contribute to SiS symptoms such as blurred vision and difficulty focusing as result of the possible small offset between the eyes and the HMD's lenses. This could lead to blurred vision and difficulty focusing. Rebenitsch and Owen [173] found a statistically significant positive correlation between vision correction and SiS. 45% of participants needed a correction of their vision, with 25% wearing contact lenses, 10% glasses, and 10% using both. In the same work, the authors built an SiS prediction model using the vision correction that explained only 17.9% of the variance. Furthermore, when this factor was combined with motion sickness history score from carnival rides, as much as 58% of the variance was explained. This shows that vision correction could contribute significantly to predicting SiS onset. Therefore, the relationship between vision correction and SiS is investigated in this thesis.

Experience with VR System (Adaptation)

Previous experience with a VR system refers to how frequently, if at all, a participant has used VR prior to being involved in a test study. This factor is directly related to repeated exposure to the virtual environment, also called adaptation or habituation. Adaptation is considered one of the solutions against SiS, and has been studied [88, 145, 178]. To become fully adapted to virtual environments, it is essential that users use VR regularly. Stanney and Kennedy [206] suggest a two-to-five days time gap between the VR sessions, although some adjustments to this time scale may be required depending on the individual user.

Howarth and Hodder [91] conducted an experiment on adaptation using a PC racing game and 7 different groups of participants. Ten participants were exposed every day, another ten participants were exposed every two days, and so on up to a week. The results from their experiment showed that SiS symptoms decreased drastically with adaptation between one and seven days apart. Having said that, the lasting effect of adaptation has not been well researched.

Furthermore, adaptation only works in a specific way; that is to say, just because a user has adapted to one virtual system does not mean they will be habituated to another virtual system. This shows the non-transferable nature of habituation. For example, the user could adapt to a specific driving simulator. However, when the HMD of this simulator is replaced with a similar but not identical HMD, the user's body will react differently as the system is a new virtual environment. It might be that adaptation to one system makes it easier to adapt to another, but adaptation to a new system would still require at least a few sessions.

Experience with Real-world Task

Research into an army flight simulator revealed a positive relationship with SiS [26]. The pilots with more flying hours experienced the SiS onset earlier than the pilots with fewer hours. Several other papers showed evidence of the same relationship [127, 145, 234]. Thus, it is clear that user's past experience is an essential factor is SiS due to the conflict between what is seen and what is expected to be seen. This supports the cue conflict theory, which states that the more experience the user has, the more significant their neural store is, which in turn increases the chance of mismatching patterns with the current simulator and also leads to a higher possibility of SiS onset.

Duration

In terms of duration, the longer the period spent in the simulated environment, the higher the chance to experiencing SiS [144]. Stanney et al. [210] reported that the most individuals (37%) dropped-out between 11 and 20 minutes of VR exposure. Thus, they suggested that VR exposure should perhaps be limited to less than 15 minutes. Furthermore, Fisher [67] noted that the test time frame should be between 5 and 25 minutes, and, if necessary, the user should take a 10-minute break. There is no specific rule regarding how long the users should be exposed to VR, but it is recommended that the exposure not exceed two hours [98]. Moreover, the session's duration should be reduced to the minimum time needed for the task to be completed.

Video Games Experience

Video game experience is assumed to be one of the predictive factors of SiS onset. It is stated that users who play daily video games (e.g., console or PC games) are more resilient to the sickness. The virtual world is a computer-generated world not unlike the 3D gaming world. Whilst there are differences in the immersion level and surrounding environment, conventional video games do to a certain extent prepare users for some of the potentially disorienting features of VR immersion, such as quick changes in the player's point of view. Thus, players who have no experience in this style of gaming may be at greater risk of experiencing eye strain and mild disorientation. Further

research is needed, though, to investigate the relationship between previous gaming experience and SiS

Perceptual Style

Perceptual style is the way in which users perceive motion during a virtual simulation. In the section regarding factor gender, it was explained that female and male users perceive depth in two different ways. The current HMDs (e.g., Oculus Rift, HTC Vive) use the parallax motion to simulate depth during motion in a virtual environment. This perceptual style matches the way that male users perceive depth in real life. Therefore, it is assumed that, unintentionally, the HMDs are made primarily for male users [23]. Moreover, visual discomfort is a subjective perception and can be expressed through a headache, blurred vision, double vision, and eye strain. One suggestion is that this discomfort might be related to a viewer's perception of motion [93]. However, in this thesis, a consumer version of a modern conventional HMD is used in the experiments. An investigation of the different perceptual styles is therefore outside the scope of this thesis.

Emotions

Valence describes the positive or negative feeling caused by a situation or an object. For example, anger and anxiety represent a negative valence, and joy represents a positive valence (Fig. 2.2). Arousal describes the physiological and psychological state of a person being awoken. Excitation represents positive arousal, and boredom portrays negative arousal (Fig. 2.2).

Reason and Brand [172] noted that anxiety might accelerate the onset of symptoms during motion exposure. More anxious participants may experience greater discomfort than participants who are less anxious. A previous study revealed a correlation between anxiety and SiS symptoms such as *general discomfort, nausea*, and *stomach awareness* [137]. However, whilst anxiety might be a contributory factor, it is not a primary inducing factor. Kim et al. [119] presented a comparison study between desktop, HMD, and CAVE during a stressful task. The results showed that HMD-elicited negative emotional changes are comparable to the emotional changes induced by the other two systems. Contrary to this, another study showed that using an HMD led to significantly higher valence and arousal scores [81], particularly in terms of positive emotions. Walch et al. [225], meanwhile, reported no significant difference between HMD and standard flat screen driving simulation on valence and arousal scores.

Personality Traits

Neuroticism is the tendency to experience negative mood states, frustration, and anxiety [15]. This personality trait has been related to increased levels of motion sickness [44]. Nonetheless, the

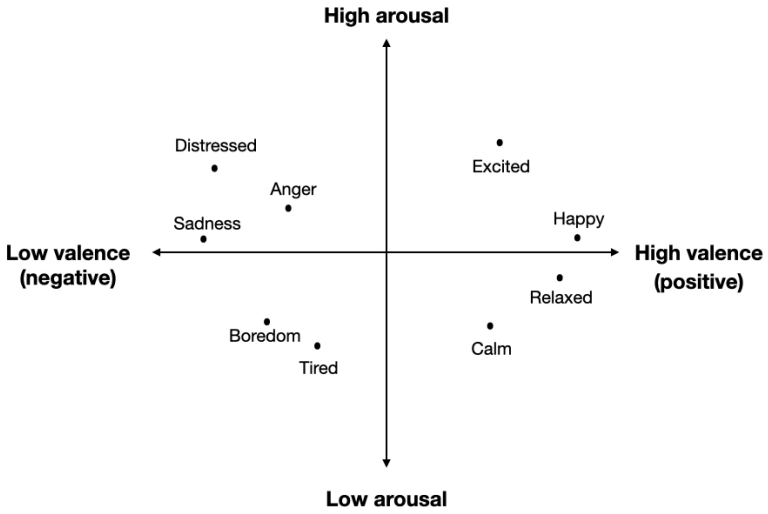


Figure 2.2: Two-dimensional emotion arousal-valence model adapted by Russell [187].

reason why *neuroticism* is related to motion sickness is still under discussion. Wilding and Meddis [232] reported a significant positive correlation between *neuroticism* and motion sickness. In their study, the more neurotic individuals reported a higher level of motion sickness, but whether those individuals actually experienced more motion sickness is unclear. Nonetheless, Reason and Brand [172] noted that a correlation between *neuroticism* and motion sickness might exist to some extent but that a causal link between them seems unlikely. Furthermore, Owen, Leadbetter and Yardley [158] found that anxiety, which is among the traits associated with *neuroticism*, was correlated to motion sickness within a VR environment. Regarding the *sense of presence*, Riccelli et al. [181] revealed that participants who reported higher levels of *neuroticism* experienced greater *sense of presence* during the VR rollercoaster ride than the participants with lower levels. No research, though, has been found on personality traits and SiS in VR HMD driving simulations, and hence, their relationship within this context can only be assumed.

Extroversion is described by sociability, talkativeness, and outgoingness [5]. Reason and Brand [172] discussed the relationship between the extroversion–introversion spectrum and motion sickness. They revealed that although no studies directly confirmed the correlation between *extroversion* and motion sickness, many noticed that extroverts are more resistant to sickness than introverts. Collins and Lentz [44] found that individuals who scored higher on *extroversion* scale experienced less motion sickness than those with a lower score. A focused evaluation on extroversion–introversion and SiS is seldom in the literature. Thus, it can be assumed that the relationship between these two factors is the same as the relationship between *extroversion* and motion sickness. In the current thesis, we investigate the possible correlation between personality

traits (*neuroticism* and *extroversion*) on SiS in the context of a VR driving simulation.

2.3.2 Simulation Factors

Degree of Control

Degree of control describes the extent to which users feel they have control over the movement initiated in the virtual environment. Stanney and Hash [205] investigated three different degrees of control (passive, active, and active-passive) in order to determine which one induced the least SiS. The results showed that the active condition (i.e., complete control) reduced the experienced symptoms as compared to the passive condition (i.e., no control). However, compared to the active-passive condition (i.e., coupled control), the latter condition induced fewer sickness symptoms. Thus, the study showed that user-initiated control could be manipulated in order to reduce SiS onset [205]. In another study, Stanney et al. [211] found a significant positive correlation between user's control over movements and Nausea symptoms. Moreover, when the users had complete control, their performance was enhanced in stationary tasks and in tasks, which required only head movements.

Within the more specific context of a driving simulator, the degree of control can be categorized into two levels—active and passive. The active level is when the user is the driver and has full control over the vehicle. The passive level is when the user is the passenger and has no control over it, such as in a fully automated car.

According to previous research, passengers experienced more SiS than drivers in a driving simulation [36, 59, 184]. For example, Rolnick and Lubow [184] conducted a rotation experiment in which participants were divided into "drivers" and "passengers." The drivers performed a set of nauseogenic rotating movements, while the passengers merely watched and experienced the accompanying motion. The results showed that the drivers felt fewer motion sickness symptoms than the passengers due to their relative control over the rotations. This lack of control contributes significantly to motion sickness onset in individuals who are more susceptible to it due to their incapability to adequately predict the upcoming movement trajectory.

Furthermore, Dong et al. [59] evaluated the influence of vehicle control on motion sickness while playing a console video game from a driver and a passenger perspective, in a similar manner to the rotating experiment of Rolnick and Lubow [184]. The experiment had a coupled design whereby the participants were randomly divided into pairs. Each pair consisted of an assigned driver and an assigned passenger. A video of the driver's driving was recorded and shown to the passenger. That way, the participants were exposed to the same video motion cues from a different vehicle control perspective. Their results confirm that the control over the simulated vehicle reduced motion sickness and that postural instability precedes motion sickness. Furthermore, before the onset of subjective symptoms of motion sickness, the movement varied between the participants who reported sickness and those who did not [59].

A recent study conducted by Curry et al. [47] presented the opposing conclusion, however. That is, no effect of vehicle control on SiS was reported. They carried out the same coupled design as Dong et al. [59]. The difference was that they paired only the same gender participants in the study, and they utilized a modern HMD. The results showed that no significant difference was observed between the drivers and the passengers as well as between female and male participants regarding SiS. Interestingly, female drivers discontinued significantly earlier than male drivers. These findings contradict those from the earlier studies and show that utilizing a different viewing device has an impact on the induced discomfort. The results point out that gender might have a more substantial effect on SiS than vehicle control in the context of VR driving. Nevertheless, the study used a stool instead of a car seat, leading to reduced back support which could result in discomfort to some extent. Also, the car seat contributes to making the virtual experience feel more closely matched to a real-world experience.

McGill et al. [146] investigated the utilization of an HMD within an in-car driving setting. They tested three different visual in-motion conditions (conveyed rotations only, conveyed no motion, and peripherally conveyed rotations and accelerations) regarding sickness. At the same time, the participants had the role of passengers. The results showed no significant differences between the conditions. Nevertheless, the participants reported different preferences across the conditions. McGill et al. [146] pointed out that the wrong visual in-motion condition might induce more discomfort than the preferred condition. These findings show that discomfort during real-world automated driving while using an HMD might be strongly related to individual preferences.

Without control over the vehicle, the driver's role is transformed into the role of a passenger. It is assumed that the same effect applies to any simulated automated driving as well. Therefore, removing the users' control in a virtual vehicle could lead to a high drop rate due to SiS onset. Furthermore, there was a suggestion of designing this automated driving environment with consideration of the SiS impact [56]. Automated driving could offer the opportunity to choose a smoother and less sudden maneuver to reduce SiS. However, that kind of setting would require personalization of the vehicle, and thus the vehicle would not be operationalized as a shared means of transportation.

Vection

Vection is the perception of illusory self-motion while the body is in a static position [87]. A classic example of vection is when a person is sitting on a stationary train and sees another train move in parallel. The person in the stationary train perceives the motion as his or her own motion and so experiences illusory self-motion. Vection can be experienced along all six degrees of freedom (DOF) in body motion [113]. Body motion is described with respect to the x-axis (longitudinal) or the fore-aft axis, the y-axis (lateral) or the left-right axis, and the z-axis (spinal) or the up-down axis [89] (see Fig. 2.3).

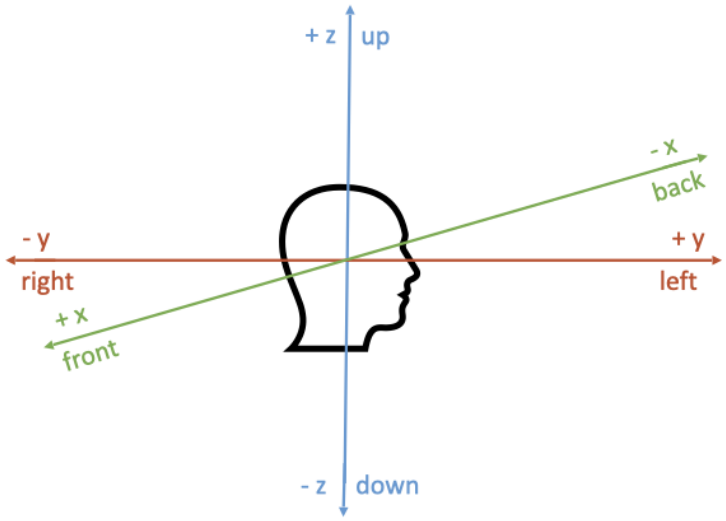


Figure 2.3: Cardinal head axes including x, y, z-axes, representing the fore-aft (front-back) axis, left-right axis, and up-down axis, respectively.

Illusory translation along one of the axes is called linear vection, and illusory rotation around one of the axes is called circular vection. Rotation around the x, y, and z-axes are known as pitch, yaw, and roll, respectively. In virtual environments such as driving simulators, the linear vection (particularly along the fore-aft axis) is more common than the circular vection [113]. The relationship between vection and SiS has been well researched, although the results are controversial. Some studies reported a positive correlation between vection and SiS [57, 132], while other reported a negative correlation was also reported [20, 159]. Furthermore, one study found no significant correlation [114]. Altogether these findings showed out that the relationship between these two phenomena appears sophisticated and affected by other factors during the VR simulation.

Self-movement Speed

In a virtual environment, the users receive information from different visual cues to help them estimate their self-movement speed. A higher speed movement should result in blurred vision as the human visual system is used to such as experience in the real world. Although there are no studies specifically on how self-movement speed affects SiS, there are studies that have evaluated the effect of rotational speed. So, Lo and Ho [203] investigated the effect of navigation speed on vection and SiS in an HMD virtual environment. There were eight-speed levels in the fore-and-aft axis. The duration of exposure was 30 minutes. The results showed that vection and SiS symptoms increased when the speed rose from 3 m/s to 10 m/s root mean square. Beyond 10 m/s, the navigation speed was steady until it reached the speed of 59 m/s. Another notable finding was that the vection outraged a bit earlier than nausea, which supports the idea that SiS

induced by virtual environments is predominantly a visual-induced motion sickness.

Keshavarz and Hecht [110] investigated the role of combined roll, pitch, and yaw motion on SiS while using a virtual rollercoaster. The results showed that the lowest sickness scores were obtained in the pitch condition and a significantly higher score was reported in a dual- and triple-axes combination. The authors concluded that the complexity of visual motion did not increase SiS onset linearly and proposed that the sickness reached a steady position in the dual-axis condition. The addition of a third axis did not exceed the reached severity of sickness. However, the results should be interpreted cautiously because the experimental setup used a projection screen, not an HMD.

Contrast

Visual contrast has shown no relationship with SiS in VR environments. However, contrast is related to the level of luminance as the difference between the highest and the lowest level of luminosity. The high contrast could lead to eye strain.

Luminance Level

Screen luminance is how much luminance is collected by the lenses of the VR device and then projected to the user's eyes, e.g., final illuminance.

$$Illuminance(inlux)I = L * Sin(theta/2)^2 * PI \tag{2.1}$$

where L is the screen brightness, and theta is the displayed FOV. Usually, the virtual world simulation is brighter than the real world but not brighter than a good lid with the artificial light place. The luminance level is related to flicker. A higher luminance level could lead to a greater flicker [155].

Color

A study, carried out by Bonato, Bubka, and Alfieri [19], compared three optokinetic drum conditions with black and white, gray shade, and chromatic stripes (including white, red, yellow, black, green, blue colors) regarding motion sickness. The result showed that color affected the motion sickness onset and, more specifically, increased the likelihood of headaches. So and Yuen [201], however, did report that changing the color did not affect SiS in a comparison study of four similar VR setups with two changing colors. Thus, we assume that color might have an effect on SiS within modern HMD environments; but the investigation of this factor is out of the scope of the current thesis.

Scene Content/Complexity

The virtual scene content might contribute to induced some of SiS symptoms. For example, if a user is exposed to a driving simulation, which includes many complicated maneuvers, this could lead to dizziness, vertigo, and nausea. Therefore, the simulation scenario should be created in terms of a user-friendly environment to reduce potential discomfort without compromising the simulation's tasks. Lo and So [141] reported that the increased scene velocity (i.e., scene complexity) along with pitch, yaw, and roll axes in VR significantly increased the SiS onset. However, in a standard driving simulation, the level of scene velocity is stable during the whole session. Moreover, in some cases, the addition of visual assets to the scene complexity of an HMD driving simulator can reduce SiS [95]. The VR simulation should display an appropriate level of content. In the current thesis, we chose a low-traffic urban driving simulation, including other vehicles. We excluded pedestrians or bicycles as they might contribute to crowding the scene, lower the simulation performance, and overwhelm the participants.

Independent Visual Background

Prothero [165] presented an "independent visual background" (IVB) as a mitigation technique against SiS. An IVB is a visual background (e.g., grid-like background) displayed behind the visual content of interest. Prothero also conducted an experiment to test the effect of an IVB on SiS [165]. The results showed that the addition of an IVB significantly reduced the discomfort compared to non-IVB conditions. Nevertheless, it was remarked that the IVB technique could only reduce SiS induced by rest frame conflict.

Furthermore, Duh et al. [61] investigated the effect of an IVB on SiS in two different display conditions (stereo and non-stereo) while using a driving simulator. They found that an IVB reduced SiS in both conditions. These findings are consistent with those of Prothero [165].

2.3.3 Hardware Factors

Display Type

Different display types (e.g., desktop monitor, HMD, and projection display) might induce different levels of SiS due to their different technical specifications. Sharples et al. [196] compared four different displays (HMD, desktop display, projection screen, and reality theatre) regarding SiS in VR factory simulation. The results showed that the HMD induced more Nausea symptoms than the desktop display and the reality theatre. Comparing the difference between the pre- and post-exposure SiS, the HMD showed significantly higher severity across all display types. Moreover, the HMD provoked more Disorientation symptoms compared to the desktop display.

One of the described earlier types of eye movements is vergence (see Chapter 2.1). Hoffman et al. [90] displayed in their work the vergence and the focal distance with real stimuli and stimuli presented on conventional 3D displays. They explained that the user's gaze is fixed on a pivot in the real world and it is fixed on a simulated pivot on a computer display screen. The vergence distance is the same as in the real world, but the focal distance, which is now the distance to the display, is shorter. A vergence-accommodation conflict arises from the differences between focal distance and vergence distance in a simulated environment. This conflict causes fatigue and discomfort in users [90].

The prolonged use of 3D displays, including HMDs, might lead to increased fatigue. Fatigue is one of the SiS symptoms that might be induced by the vergence-accommodation conflict. Therefore, objects in the VR should be arranged in such a way as to reduce the chance of fatigue or to at least postpone its onset.

Park et al. [160] proposed a solution based on the vergence-accommodation conflict, namely an oculomotor exercise performed immediately before viewing the VR content while using an HMD. This exercise combined the following oculomotor exercises: range, saccadic, pursuit, and vergence. Although the sample size was small (n = 8), the results showed significant differences in favor of the exercise paradigm in the three different sickness groups: Nausea, Disorientation, and Oculomotor. This simple for integration method of reducing cybersickness might be a key to minimizing the discomfort and caused by prolonged use of HMDs. Still, more tests set in different environments are needed and not only with images but also with actual simulation in high-end HMDs.

Resolution/Blur

Resolution refers to the graphics quality of the display. The higher the resolution, the higher is the quality of the visual scene in VR. In a study by Mon-Williams et al. [149], it was reported that resolution quality might lead to undesired discomfort. Moreover, the resolution is related to flicker, which could increase the SiS onset. However, modern HMDs have a good resolution specifications of approximately $1080x1200$ pixels per eye. Therefore, the graphic quality is suitable for most applications. Further investigation is needed, though, to establish a connection between resolution and SiS onset in VR driving simulations. But the exploration of this factor is outside the scope of this thesis.

Field-of-View

FOV is the extent of the observable world that a user can see on a display at any given moment. Simulators with a wider FOV induce more SiS than simulators with a narrow FOV due to the larger perception of visual signals and vection [107].

Moreover, the wider FOV and relatively larger display could lead to a higher possibility of flicker, which is another factor of SiS [127]. The wider the FOV, the greater the likelihood of SiS onset because a user receives more visual cues than is necessary. Furthermore, a study conducted by Seay et al. [193] explored the links between FOV size, SiS and *sense of presence*, showed that a higher level of nausea was related to wider FOVs. However, users enjoyed a superior level of realism with the high FOV. This shows that the FOV and the level of immersion are controversial factors. On one hand, the user feels more immersed in the simulated environment with a wider FOV. On the other hand, a wider FOV could lead to symptoms of discomfort, such as headache and disorientation.

Ergonomics

Ergonomics is the design of a system or device that is used by a user to enter a simulated environment. This is strongly related to the HMDs, and the manufacturers have been working hard in recent years to overcome this problem. With HMD usage still in its infancy, the design of a lightweight and well face-fitting HMD is necessary to ensure that usage becomes more common. The consumer versions presented on the market by companies like Oculus and HTC Vive have reached a comfort level that permits users to operate within a virtual environment for a longer time. SiS symptoms such as a headache and general discomfort come along with the lousy ergonomics due to the low comfort level of the system [52]. Thus, it can be assumed that the weight of the HMD is not a sickness induced factor in HMD driving applications.

Inter-pupillary Distance

Inter-pupillary distance is the distance between the centers of the pupils of the eyes. This distance is critical for binocular viewing systems where the pupils' centers should be positioned within the pupils' exits of the viewing system. Kolasinski and Gilson [128] found a significant negative correlation between *eye strain*, an Oculomotor symptom, and inter-ocular distance. On the other hand, Howarth [92] pointed out that the mismatch between the system inter-ocular distance and the inter-screen distance, theoretically, is more likely to induce discomfort than the mismatch between inter-ocular distance and inter-pupillary distance. Therefore, it can be assumed that the inter-ocular distance configuration of the HMD should not differ significantly from the user's inter-ocular distance.

Update Rate

The update rate is the frequency of the monitor display updates the cycles displayed on. If the update rate is low, then a visible fading between the cycles, called flicker, occurs. The flicker could distract the user during the simulation due to the inconsistent display of the images and

might result in eye strain and headache. In general, users have a different level of sensitivity to the flicker due to physical differences [127]. To avoid the flicker, the update rate of the HMD should not be below $75Hz$.

Latency

In terms of lag, also called latency, there must be no significant delay between action and reaction while using the simulator or any other simulated environment for that matter [52]. The human brain is a fine-tuned sensitive organ, which can detect the slightest delay between the user's actions and the displayed outcome. If the lag takes too much time, the brain is confused, and this results in SiS [218]. Therefore, the latency should be less than 20 milliseconds; otherwise, flicker could occur [12].

Calibration

Calibration is the setting of a VR system in order to provide usability comfort. Davis [52] noted that a poorly calibrated VR system can increase SiS symptoms Therefore, calibration is an essential part of the preparation for any user evaluation. The calibration is directly related to the user's preferences and could induce or reduce SiS based on how well it is carried out.

Position Tracking Error

A position tracker is a vital part of every HMD. It can be integrated into the standalone HMD, or it can be an external part. The role of the tracker is to map the position of the HMD in any given moment in the virtual world. Therefore, the tracker's accuracy must be high. The tracker accuracy represents the difference between the object's actual 3D position and that reported by the tracker measurements [33]. The occurrence of tracking errors, such as misplace of the user's position, might induce vertigo and disorientation. Therefore, this error should be reduced to a minimum before any user evaluation.

Head Movements

As one of the most common virtual environment exploration methods, moving the head in the direction of interest is used nearly every VR application. Even in VR applications, where the user is only an observer, the head movements are essential to navigation. Previous research found that participants who performed more head movements while playing video games using an HMD experienced more SiS [147]. A more recent study carried out by Palmisano, Mursic and Kim [159] showed that head movements could contribute to SiS when a mismatch exists between the perceived and physical head movements. Some participants might reduce their head movements

in order to minimize the felt discomfort during the VR. This coping mechanism is a collective body reaction to sickness-induced stimuli. For example, in a VR driving simulation, individuals could reduce their head movements to just the most essential head movements. That, however, could lead to unrealistic behavior, which could compromise the user study itself. Furthermore, Keshavarz et al. [116] reported that restraining the upper body against the backrest of a driving seat reduced SiS during a video driving game. The findings point out that the connection between head movements and SiS is questionable, and that other factors might be involved. Nevertheless, the studies with more recent HMD models showed that SiS could be reduced through lessening the head movements. Thus, in this thesis, head movements are used as a predictor of SiS as categorised under behavioral factors.

Haptic Feedback

Haptic feedback is a combination of tactile and kinesthetic feedback, which can bring realism to the VR simulation through the addition of vibration and a motion platform. It is assumed that vibration received as haptic feedback from the driving seat helps to reduce cybersickness during virtual driving. However, in a study on airflow and seat vibration in VR, D'Amour et al. [62] reported that the presence of a seat vibration did not have an impact on SiS. These results should be interpreted with caution because the virtual environment used in the experiment was a video of a bicycle ride shot from a first-person view.

Another way for the user to receive haptic feedback is through a motion platform. The addition of motion cues could reduce the discrepancy between the visual and vestibular systems. Several studies show that SiS is less likely to occur in a dynamic driving simulator than in a static driving simulator [9, 10, 48]. But other studies did not support this statement [108, 123]. For example, Klüver et al. [122] compared two motion-based and three static driving simulators without finding a difference in the onset of SiS. However, it has to be mentioned that in this study, the driving simulators also differed in terms of other potentially confounding variables, such as FOV, resolution, and motion delay between the visual input and physical motion response.

Keshavarz et al. [108] examined the mitigating effect of physical motion cues while participants drove actively in a VR driving simulator. Contrary to their expectations, the results indicated only a small impact of physical motion cues on SiS. Nevertheless, the authors emphasized the more significant impact that physical motion cues could have if they are produced more congruent to the visual cues of the driving simulation than in their conducted study. These results and recommendations are in accordance with the findings of Klüver et al. [122]. Furthermore, Keshavarz et al. [108] emphasize how challenging it is to achieve this kind of precision when self-driving, meaning the participants can choose any driving maneuver at any time. They propose it is easier to achieve precision when participants drive fully automated, and the driving maneuvers are predefined for every participant.

Many factors, described in the literature, are related to SiS onset. With the up-rise of the modern HMDs, some factors are no longer relevant, such as resolution. Other factors, such as gender and motion platform, are still establishing their relationship to SiS, despite their long-time presence in the SiS research. Potentially new factors such as personality traits, which were previously less investigated, might bring new insights. Moreover, the degree of control as a factor will become an essential aspect in future VR studies due to automated driving and interaction within this innovative vehicle environment. Therefore, this thesis will focus mainly on individual factors while using a different type of driving and motion in a VR driving simulation.

2.4 Simulation Sickness Measurements

2.4.1 Subjective Measurements

The majority of research studies on SiS have used some form of questionnaire as an SiS assessment tool. Sometimes the questionnaires are used in combination with other measurement tools such as physiological or behavior measurements. One of the main reasons to use questionnaires is that SiS is a largely subjective phenomenon that differs widely between individuals. Furthermore, this type of measurements are easy to administrate and to analyze.

Simulator Sickness Questionnaire (SSQ) [106] is the most widely used questionnaire for evaluating SiS onset. The SSQ gives insights into how the user felt after the experiment. The questionnaire consists of 16 questions, each of which has four possible answers, e.g., none, slight, average, and severe. Each answer reflects the degree to which a participant experiences a symptom. The scoring system of the SSQ uses unit weights (Table 2.2) based on the following scores: 0 - none, 1 - slight, 2 - moderate, and 3 - severe. Each of these pertains to one or more of three clusters, e.g., Nausea (N), Oculomotor (O), and Disorientation (D). The general discomfort, difficulty focusing, difficulty concentrating, and blurred vision symptoms are part of two clusters. Each symptom is given a score, which is multiplied by the relative weight of one of the clusters and then summed alongside other symptoms' scores at the end of the same cluster (e.g., nausea-related [1], oculomotor-related [2] and disorientation-related [3]).

Table 2.2: SSQ scoring system [106].

Item	Nausea	Oculomotor	Disorientation
General discomfort	1	1	0
Fatigue	0	1	0
Headache	0	1	0
Eye strain	0	1	0
Difficulty focusing	0	1	1
Increased salivation	1	0	0
Sweating	1	0	0
Nausea	1	0	1
Difficulty concentrating	1	1	0
Fullness of head	0	0	1
Blurred vision	0	1	1
Dizzy eye open	0	0	1
Dizzy eye closed	0	0	1
Vertigo	0	0	1
Stomach awareness	1	0	0
Burping	1	0	0

The summed score of all symptoms with weight = 1 was multiplied by a specific coefficient. The following equation shows the total cluster score:

$$N = [1] * 9.54, \; where \; N \; stands \; for \; Nausea \; cluster \tag{2.2}$$

$$O = [2] * 7.58, \; where \; O \; stands \; for \; Oculomotor \; cluster \tag{2.3}$$

$$D = [3] * 13.92, \; where \; D \; stands \; for \; Disorientation \; cluster \tag{2.4}$$

The total score (TS) of all symptoms is the sum of the cluster scores before they are multiplied with the coefficient:

$$TS = ([1] + [2] + [3]) * 3.74 \tag{2.5}$$

One of the limitations of the SSQ is that it is administrated after the simulation is over. Therefore, it can not assess precise SiS at a given moment and can not follow the development of the malaise over time. Under-reporting the symptoms is another limitation as a person could unintentionally fail to report the severity of the actual symptoms.

Virtual Reality Symptom Questionnaire (VRSQ) is a VR questionnaire based on SSQ, where specific factors showed more weight than the other regarding VR sickness [4]. The final version of the questionnaire consists of 13 symptoms categorised into two groups: general body symptoms, and eye-related symptoms. Although this tool was developed specifically for the VR environment, it lacks validation. Further recommendations for using a more significant subject group and diverse age group were made in the research. A recent study by Del Sid et al. [53] validated the VRSQ with a total number of 100 participants.

Although the SSQ was developed many years ago, it is still the most commonly used tool for any sickness onset in simulated environments. Twenty years later, after the first introduction of SSQ, a validation was conducted with 530 participants. It was confirmed that the SSQ is still a useful tool for assessing SiS onset [13].

Fast Motion Sickness Scale (FMS) is a verbal rating scale ranging from 0 (no sickness at all) to 20 (frank sickness) that assesses sickness severity at any given moment [111]. The higher scores indicate a severer SiS onset. The FMS focuses on symptoms such as general discomfort, nausea, and stomach awareness. Moreover, Keshavarz and Hecht [111] cross-validated this verbal rating scale with the SSQ with a total number of 126 participants. The FMS validation demonstrates that SiS can be assessed reasonably well using a single-item measure, which is easily administrated. Also, the scale allows researchers to capture the development of SiS over time. Despite the benefits of using the FMS, there are still some drawbacks. The scale was created to focus on general well-being, and therefore it lacks an assessment of visual-related symptoms. The cross-validation results with the SSQ support the lack of visual-related symptoms with the lowest correlation among the SSQ scores.

2.4.2 Objective Measurements

It has been demonstrated that SiS onset was significantly positively correlated to heart rate (HR) [51], heart rate variability (HRV) [120], high frequency range of HRV [194], electrodermal activity (EDA) [51, 226], and respiration rate (RR) [120], while low frequency/high frequency ratio of HRV [194] showed a negative correlation. Furthermore, previous research shows that participants with severer SiS symptoms showed higher levels of these physiological signals than participants who did not experience SiS [42]. However, it must be noted that some research results show that these physiological changes can individually differ regarding their direction [98].

Electrocardiography (ECG) is the measurement of the heart signals that the heart sends every second while pumping blood out into the arteries. On the electrocardiogram, several different waves are displayed. The signal recorded for every given second, and from that, the HR and HRV are analyzed. The peak of the highest wave in the upper part of the electrocardiogram is called R-peak. The HR is calculated based on the distance between two R-peaks, also called the RR interval.

$$HR = 60 seconds/interbeatinterval \tag{2.6}$$

Prior studies demonstrated that HR had a positive relationship to SiS induced by VR environment [153]. HR was higher at the beginning when the users entered the virtual world and later, it reverted to normal values. Earlier in this chapter, it was shown that duration in the virtual world is one of the SiS inducing factors. The prolonged stay in VR could provoke SiS onset [211]. However, a change in HR is not necessary to be caused solely by SiS. Other factors can influence HR such as excitement, which users may experience when using VR for the first time.

Another study reported that HRV as an objective factor could help to predict SiS [32]. The connection is unclear, but the results showed that users who experienced more SiS also had higher HRV. This could be connected to the stress level and the situation when entering a new and highly stimulating virtual environment. Kim et al. [120] reported a positive significant correlation ($r = 0.373$) between HRV and SiS. In the same study, the relationship between SSQ clusters and SiS was calculated, and among the clusters the strongest correlation was between the Oculomotor cluster ($r = 0.426$) and SiS. Furthermore, these findings were supported in a recent study by Garcia-Agundez et al. [73] where a strong positive correlation between ECG parameters and the Oculomotor or the Disorientation cluster was identified.

Blood Volume Pulse (BVP) records changes in the volume of blood in the blood vessels. A non-invasive sensor can measure the cardiovascular dynamics through light absorption of the skin and tissues and their level of illumination [2]. When the heart pumps blood, the arteries become denser, and the light passing through is reduced. Usually, the sensor is attached to the non-dominant hand on the tip of the ring finger. The HR is calculated from the BVP signal by estimating the time interval between the heartbeats, named interbeat interval, measured in seconds [163]. Furthermore, similar to ECG, most of the factors related to the heart signal can be extracted from the BVP signal.

Electrodermal Activity (EDA) is the continuous change in the electrical properties of the skin [25]. It is measured with sensors that detect the changes in the passing through the skin a neglected amount of currency, measured in micro-Siemens (μS). Two main parts of the EDA can be used to measure ongoing electrodermal changes [25]. The first one is the tonic level, which is

related to the signal's background characteristics and slower acting changes. The second one is the phasic level, which is associated with faster changes in the signal.

EDA can be measured through the galvanic skin response (GSR), the electrodermal response (EDR), the psychogalvanic reflex (PGR), the sympathetic skin response (SSR), and the skin conductance level (SCL). Throughout this thesis, we use the SCL as a measure of the EDA. Typically, the EDA changes when people are under stress, and and manifests through perspiration on the inner side of the hand, the palms, the fingers, and the forehead. Previous research reported an increased SCL while using a VR environment [63, 148]. For example, it was found that SCL increased during a VR navigation [120]. The SCL was significantly increased in the last minute of the VR navigation when compared to the baseline value. It is unclear, though, exactly why this is, as it can be argued that the skin response changes not only because of the felt discomfort but because of the stressful situation in the simulation.

Respiration is the process of breathing in (inhalation) and breathing out (exhalation) [6]. The respiratory effort is measured by the number of breaths taken in a certain amount of time, often referred to as respiration rate (RR). Typically, RR is measured in breaths per minute.

Bruck and Watters [32] theorized that an increase in *arousal* level assists changes in RR, which leads to reduced carbon dioxide levels in central blood flow. The reduction of carbon dioxide levels results in symptoms such as *fullness of head, fatigue, difficulty concentrating*, and *dizziness*. They also suggested that RR changes might occur in response to the increased *arousal* level resulting from stimuli of high movement VR. Furthermore, Kim et al. [120] reported a significant positive correlation between the SSQ total score ($r = 0.392$), the SSQ Nausea ($r = 0.342$), the SSQ Oculomotor ($r = 0.386$), the SSQ Disorientation ($r = 0.382$), and RR.

2.5 Summary

In summary, SiS is a complex malaise that involves a multisensory integration process. Typically, different sensory systems produce a well-tuned information input that is received by the human body. However, the usual perception of self-motion is disrupted by VR experience in which artificially simulated motion occurs, such as in a driving simulation. This can lead to experiencing vection and in some cases SiS.

In this chapter, we have presented an overview of the human sensory systems related to SiS and outlined the theoretical background of the current thesis. We have also described the factors related to SiS and their possible relationship to the discomfort induced by VR driving. In the following chapter, we will discuss related work and prediction models of SiS regarding HMD driving simulations.

3 Related Work

I'm trying to free your mind, Neo.
But I can only show you the door.
You're the one that has to walk
through it.

Morpheus, "The Matrix"

Although a large and growing body of literature has investigated SiS, not so many studies have specifically investigated SiS induced by VR driving applications. Even less common are the studies using an HMD in their setups. Thus, before we continue to the methodological part of the thesis, we will consider the relevant setups and studies. The outcome of this review is a list of mitigation techniques against SiS. The techniques are separated into three groups aligned with the groups used to categorize SiS factors in the previous **Chapter 2**. Each mitigation technique is reviewed in terms of its potential applicability to HMD driving simulations. Even though the utilization of techniques for reducing SiS can significantly improve user experience, the techniques are sometimes also related to technical or simulation changes in the VR setup. Furthermore, the goal of the techniques is to treat potentially sick individuals, not to prevent them from becoming sick in the first place. Thus, predicting SiS is essential for improving the user experience and the related usability of VR simulations. In order to build a basic knowledge of the utilized classification methods, we provide a brief overview of these methods. A discussion on the previous prediction models of SiS in the literature and their shortcomings is presented in the last part of this chapter.

The chapter is based on the following publication:

- Rangelova, S., & André, E. (2019). A Survey on Simulation Sickness in Driving Applications with Virtual Reality Head Mounted Displays. MIT Press. Presence, Vol. 27, No. 1, Winter 2019, 1–17 [168].

3.1 HMD Driving Setups Overview

There are several different types of VR driving system, ranging from a set of screens to CAVE-like systems. The compact driving setups usually employ a screen or three-screen set or an HMD as a viewing system. With the revival of HMDs as a highly immersive VR medium, the number of setups using a modern HMD has increased. A typical compact VR setup contains a PC, an HMD (including positional trackers), a steering wheel, pedals, a gear changer, and a car seat. Some setups might include a moving platform in order to replicate motion feedback from a moving vehicle. An overview of VR setups with a modern HMD is shown in Table 3.1. As modern HMDs are focused on here, studies from earlier than 2017 are not presented.

Table 3.1: An overview of VR driving setups with a modern HMD, where CPU stands for central processing unit, GPU - graphic processing unit, ECG - electrocardiography, EDA - electrodermal activity, EMG - electromyographic.

HMD	PC (CPU/GPU)	Steering wheel	Pedals	Car seat	Moving platform	Bio sensors	Reference
HTC Vive	-	Yes	Yes	Yes	No	No	Walch et al. [225]
Oculus Rift CV1	- / NVIDIA GeForce GTX 970	Yes	Yes	No	No	ECG, EMG, EDA	Eudave and Valencia [63]
HTC Vive	Intel Core i7-6700K 4 GHz, 32 GB / 2 x NVIDIA GeForce 1080	Yes	Yes	Yes	6 DOF	No	Ropelato et al. [185]
Oculus Rift	Intel Core i7-6700K 4 GHz, 16 GB / NVIDIA GeForce GTX 980 12 GB	Yes	Yes	Yes	No	No	Suwarno et al. [216]
HTC Vive	Intel Xeon 64 GB / NVIDIA GTX 1080Ti	Yes	Yes	Yes	No	No	Venkatakrishnan et al. [219]
Oculus Rift	Intel Core i7-6700K 4 GHz, 16 GB / NVIDIA GeForce GTX 980 12 GB	-	Yes	Yes	No	HR	Setiowati et al. [194]
Oculus Rift CV1	-	Yes	Yes	Stool	No	No	Curry et al. [47]

3.2 Simulation Sickness Associated to HMD Driving Simulation

Research into VR driving simulations and, more specifically, driving simulations with HMDs, has experienced growth in recent years. A few studies are focusing on VR driving simulator evaluation with modern HMD. SiS in VR driving applications should be considered more complex than other comparable sicknesses in other VR applications, namely because of visual and physical motion cues, which introduce the possibility of motion sickness onset. This means that discomfort can be induced from a multitude of sources, such as visual-vestibular mismatch or visual-proprioceptive mismatch.

Table 3.2 summarizes a review of SiS associated with VR HMD driving applications since 2012. This period has been chosen due to the lower hardware parameters of HMDs before 2012, such as FOV, refresh rate, and visual graphic quality. For comparison, the Oculus Rift Development Kit 1 (DK1), released in 2013, has a resolution of $640x800$ pixels per eye, a refresh rate of 60 Hz, a latency (end-to-end) of $50 - 60$ ms, a FOV of $110°$, and weights 380g [31]. The consumer version of the Oculus Rift (CV1), released in 2016, has a resolution of $1080x1200$ pixels per eye, a refresh rate of 90 Hz, a latency (end-to-end) of approximately 25 ms, a FOV of $110°$, and weights 360g. Parameters related to the SiS outbreak, such as latency, are reduced by almost 50% with the newer HMD model, Oculus Rift (CV1).

In the literature, immersion often refers to a physiological state characterized by perceiving oneself to be enveloped by, included in, and interacting with a stimulating environment [233]. A sense of presence refers to experiencing the simulated environment rather than the actual (real-world) environment. In other words, presence is the illusion of "being there" [84]. Furthermore, presence can be defined by three ways: personal presence, social presence, and environmental presence. Personal presence measures the extent to which the user feels like he or she is in the virtual world. Social presence identifies the extent to which the user is aware of and can interact with other participants in the virtual world. Environmental presence identifies the extent to which the environment notices the user's existence and reacts to it [84].

It is assumed that a virtual environment that creates a higher immersion level would create a greater sense of presence. Nevertheless, many researchers have pointed out a distinction between the feeling of "being there" and immersion. A person can be disconnected from the real world by the VR experience (immersion) but still not feel present there [14, 156]. The sense of presence can be measured by the subjective seven-point scale questionnaire called the Presence Questionnaire (PQ). The PQ's original version includes 32 items divided into four major factor categories: control, sensory, distraction, and realism factors [233]. Some of the items appear in two or three categories of factors. For example, the item "How completely were you able to actively survey or search the environment using vision?" appears in the control, sensory, and realism factor categories.

The correlation between SiS in VR and the sense of presence appears to be complicated and

indirect. A lower sense of presence may induce disorientation, which may increase sickness onset [156]. A survey such as that conducted by Schuemie and colleagues [192] has shown that the correlation between SiS and presence is controversial. Several researchers reported a negative correlation between SiS and presence [156, 211, 233]. However, other studies showed a positive relationship between sickness and a *sense of presence* [138, 139]. Lin et al. [136] found that the relationship between SiS and presence is positive and may significantly change with different interactivity levels in the virtual environment. Nevertheless, a recent review reported that there is more evidence of a negative correlation between presence and SiS than a positive correlation [229].

An essential inclusion criterion for the review was not only the year of publication but also the occurrence of SiS. It should be noted that most of the studies have a small sample size, and therefore, they lack statistical power to generalized their conclusion. Nevertheless, these studies provide valuable insights into the influence of VR driving applications on SiS and how it is measured.

Table 3.2: Overview of SiS evaluation in VR HMD driving applications

Objective	HMD	N (f/m)	Measurement	Findings	Reference
Investigation of the sense of presence and physiological response induced by an immersive virtual environment.	Oculus Rift CV1	5 (all male) M = 31.2 SD = 4.6	- HR - EDA - A customized presence and cybersickness questionnaire	HMD increased the sense of presence. None of the participants reported cybersickness symptoms. The emergency maneuvering increased the response of HR and EDA.	Eudave & Valencia [63]
Comparison of VR and non-VR driving simulations influence on physiological responses, SiS, and driving performance.	Oculus Rift Development Kit 2 (DK2)	94 (24/70) M = 24.8 SD = 4.7	SSQ	HMD induced significantly more discomfort than the stereoscopic 3D simulation.	Weidner et al. [230]

Continue on the next page

Table 3.2: Overview of SiS evaluation in VR HMD driving applications (cont.).

Objective	HMD	N (f/m)	Measurement	Findings	Reference
Evaluation of VR driving simulation in relation to user's immersion in low-cost setup.	HTC Vive	18 (3/14; 1-not reported gender) M = 25.1 SD = 3.2	SSQ	No significant difference between HMD and flat-screen condition regarding SiS was reported.	Walch et al. [225]
Evaluation of additional visual assets on SiS outbreak in HMD driving simulation.	Oculus Rift DK2	72 (18/54) M = 25.1 SD = 3.2	- SSQ - A customized questionnaire	Significantly reduced SiS onset in a city VR environment with additional assets (e.g pedestrians and other cars.) No relation found between *motion sickness history* and SiS.	Ihemedu-Steinke et al. [95]
HMDs for HCI validation while performing a lane change task in an immersive environment.	Oculus Rift DK1	20 (3/17) M = 28.3 SD = 3.6	- SSQ - PIT - NASA-TLX	SiS and spatial presence were significantly higher with HMD setup than with PC setup. Users reported awkwardness when they did not see their own hands.	Reich & Stark [180]

Continue on the next page

Table 3.2: Overview of SiS evaluation in VR HMD driving applications (cont.).

Objective	HMD	N (f/m)	Measurement	Findings	Reference
Motion sickness comparison between a CAVE environment and an HMD.	Oculus Rift DK2	24 (6/18) M = 36 SD = 9	SSQ	SiS might increase with increasing of yaw acceleration.	Colombet, Kemeny & George [45]
An investigation of galvanic cutaneous stimulation and auditory stimulation in mitigation SiS.	VR-based visual system	15 (6/9) M = 23.2	- SSQ - Head sway	SiS onset was 47% less with galvanic cutaneous stimulation and audio stimulation in a static driving simulator.	Gálvez-Garcia [82]
Development and evaluation of VR driving simulator with HMD.	Oculus Rift DK2	25 (5/20) M = 37.6	PQ	Four out of five females felt sick in the first five minutes of the test. Almost all participants with no prior HMD experience got simulation sick.	Ihemedu-Steinke et al. [94]
Comparison between a static HMD and a medium range FOV driving simulators regarding SiS.	Oculus Rift DK1	14 (2/12) M = 24.4 SD = 2.3	- Customized questionnaire - Vehicle acceleration - Head (vestibular) acceleration	SiS increased with the HMD driving simulator. Users of the HMD driving simulator felt more discomfort such as *nausea*, *dizziness* and *eye strain*. The level of immersion was higher with the HMD driving simulator and it could deliver a better experience despite SiS occurance.	Aykent et al. [11]

Continue on the next page

Table 3.2: Overview of SiS evaluation in VR HMD driving applications (cont.).

Objective	HMD	N (f/m)	Measurement	Findings	Reference
Assess the extent to which level the vehicle in a loop elicits realistic driving responses.	NVIS ST50	48 (24/24) M = 27.3	SSQ	Disorientation cluster showed higher scores, followed by the Nausea and Oculomotor clusters. The most severe symptoms were *fullness of head, eye strain, difficulty concentrating*.	Sieber et al. [198]
Development of VR driving simulator with an HMD for training users' driving skills.	HTC Vive	17 (5/12) M = 29.5 SD = 8.3	- SSQ - PQ	SiS was higher after the virtual driving than before the simulation. Nausea cluster showed the most increase of symptoms. However, the Oculomotor symptoms were slightly reduced after the VR simulation. 23.5% stopped the experiment due to severe SiS symptoms.	Ropelato et al. [185]
Comparison between an HMD and a three-monitor display VR driving simulator regarding SiS.	Oculus Rift	20 (10/10)	SSQ	The HMD driving simulator induced more discomfort than the three-monitor display setup. Participants reported more Disorientation and Nausea symptoms compared to the standard display condition.	Suwarno et al. [216]

Continue on the next page

Table 3.2: Overview of SiS evaluation in VR HMD driving applications (cont.).

Objective	HMD	N (f/m)	Measurement	Findings	Reference
Proof of concept to study SiS in a highly realistic immersive VR driving simulator.	HTC Vive	15	- SSQ - MSSQ - PQ	Based on the participants' feedback, the VR setup can be used as a reliable platform to study SiS. The results indicated that participants experienced more SiS after exposure to the VR driving simulator.	Venkatakrishnan et al. [219]

The HMD model, number (N), gender (f – female and m – male), and age (M – mean and SD – standard deviation) of the subjects as well as the measurements are mentioned. HR - heart rate; EDA - electrodermal activity; SSQ - Simulation Sickness Questionnaire; PQ - Presence Questionnaire; MSSQ - Motion Sickness Susceptibility Questionnaire; PIT - Presence and Immersive Tendency questionnaire (see [180]); NASA-TLX - NASA Task Load Index.

3.3 Mitigation Techniques Against Simulation Sickness

Despite technological advances, SiS has still not been completely eradicated. At first glance, the specific symptoms related to this sickness do not seem inherently dangerous for VR users during the actual immersive experience. However, upon further examination, SiS is not only a highly unpleasant experience among users but can also can have a long-lasting and potentially dangerous effect well after the end of the simulation [102, 144, 207, 209]. Some sources warn about specific dangers for users who drive after extended exposure to a virtual environment. At least until 2005, there were no reported car accidents due to SiS within twelve hours after a simulation's conclusion [98].

To reduce SiS onset in virtual environments, some mitigation techniques can be applied. Not every technique is suitable or has been tested with VR driving applications. Table 3.3 is adapted from a review study on behavioral techniques by Keshavarz [109], where feasible solutions for reducing SiS were discussed. Additional techniques are included for which a positive contribution to the minimization of SiS has been reported. The table excludes the technique of galvanic cutaneous stimulation due to an unpleasant feeling which can occur. This technique stimulates a large diameter of the skin surface, which affects the skin's nerve fibers, with electric current with values below the motor threshold [82].

Table 3.3: Overview of SiS mitigation techniques and associated SiS induced factors divided into three major groups

Factors	Mitigation technique	Description	N (f/m)	Application type	Findings	Reference
INDIVIDUAL	*Hyoscine*	An anti-motion sickness medication which is used for the prevention of travel sickness.	39	Exploring coridors with several rooms.	Hyoscine had a significant effect on reducing the nauseous effect of SiS in an immersive VR environment.	Regan [177]
	Ginger	A plant which root is widely used as a spice and a home remedy.	18 (8/10)	*nausea* assesment induced by circular vection	Ginger sufficiently reduced the severity of *nausea*, delayed the *nausea* outbreak and shortened the recovery time.	Lien et al. [134]
	Adaptation	The process of gradually present a new environment to the user using scheduled repeated sessions.	30	Exploring corridor with several rooms	Significantly decreased *nausea* from the first immersion to the fourth. No statistically significant differences	Regan [177]

Continue on the next page

Table 3.3: Overview of SiS mitigation techniques and associated SiS induced factors divided into three major groups (cont.)

Factors	Mitigation technique	Description	N (f/m)	Application type	Findings	Reference
			95 (56/39)	Driving simulation	A single brief exposure followed by a day with no exposure reduced SiS symptoms.	Domeyer et al. [58]
	Passive restraint	Users' upper body is fixed to the backrest of the driving seat using elastic straps.	37 (25/12)	Driving simulation	Passive restraint reduced SiS during video driving game.	Keshavarz et al. [116]
SYSTEM	*Pleasant music*	Music which is perceived as pleasant by the user.	93 (50/43)	Watching video of a bicycle ride	The pleasant music alleviated SiS. In the study relaxing music was played as pleasant music.	Keshavarz & Hecht [112]
	Pleasant odor	Odor which is perceived as agreeable by the user.	62 (47/15)	Watching video of a bicycle ride	The participants reported less SiS during the experiment when the pleasant odor was noticed. The smell of roses was used as a pleasant odor.	Keshavarz et al. [115]

Continue on the next page

Table 3.3: Overview of SiS mitigation techniques and associated SiS induced factors divided into three major groups (cont.)

Factors	Mitigation technique	Description	N (f/m)	Application type	Findings	Reference
	Fresh air	Continuous airflow.	82 (43/39)	Watching video of a bicycle ride	The SSQ Nausea cluster, the SSQ Oculomotor cluster, and the SSQ total score were reduced.	D'Amour et al. [62]
	Motion cues	Motion cues provided by a moving platform which moves the user's body according to the displayed visual stimulation.	117 (58/59)	Driving simulation	SiS was significantly lower with the dynamic driving simulator.	Curry et al. [48]
			20 (2/18)	Driving simulation	The SiS symptoms *nausea*, *dizziness*, and *eye strain* were significantly reduced with the dynamic driving simulator.	Aykent et al. [10]

Continue on the next page

Table 3.3: Overview of SiS mitigation techniques and associated SiS induced factors divided into three major groups (cont.)

Factors	Mitigation technique	Description	N (f/m)	Application type	Findings	Reference
SIMULATION	*Dynamic FOV*	Dynamically change of user's FOV without noticeable effect by the user during the exploration of the virtual environment.	30 (11/19)	Villa exploration	The dynamic FOV modification reduced the experienced discomfort and extended the time spent in the virtual environment.	Fernandes & Feiner [64]
	Visual assets	Addition of visual assets such as artificial intelligent vehicles and pedestrians.	72 (18/54)	Driving simulation	The addition of visual assets significantly reduced SiS induced by VR driving simulation.	Ihemedu-Steinke et al. [95]
	Virtual nose	An integrated im*Age* of human nose displayed in the HMD.	41	Roller-coaster and villa exploration	Users explored the "Tuscany Villa" application around 90 seconds longer and played a roller-coaster game two seconds longer with the addition of the virtual nose.	Whittinghill & Case [231]

Continue on the next page

Table 3.3: Overview of SiS mitigation techniques and associated SiS induced factors divided into three major groups (cont.)

Factors	Mitigation technique	Description	N (f/m)	Application type	Findings	Reference
	Independent visual background	A visual scene component which provides a reference point of visual motion cues.	11 (7/4)	Driving simulation	The independent visual background reduced SiS in a driving simulator.	Duh et al. [61]
	Virtual hand-eye co-ordination task	A virtual peg-in-hole game where the participant should insert the pegs into the straw-like holes from back to front. This technique is used after the user gets sick.	21 (9/12)	Custom made an obstacle course	The results reported that doing a coordination task after the user felt a discomfort significantly reduced the experienced discomfort.	Curtis et al. [49]

The HMD model, number (N), gender (f – female and m – male), and age (M – mean and SD – standard deviation) of the subjects as well as the measurements are mentioned. HR - heart rate; EDA - electrodermal activity.

Individual-related mitigation techniques

Regan [177] found that hyoscine alleviated not only *nausea*, but also sickness symptoms, such as *eye strain, stomach awareness, disorientation*, and *headache* induced by an HMD. For this evaluation, a malaise scale and the SSQ were used to measure the discomfort. The malaise scale measures the level of malaise on a scale from 1 (no symptoms) to 6 (being sick). None of the participants had been previously exposed to VR. The hyoscine has a quicker mitigation effect than adaptation [177]. Besides the relieving effect, other side effects such as *drowsiness, dizziness, blurred vision*, and *dry mouth*, which in many cases are very similar to motion sickness symptoms, may occur [204]. The study was conducted with the PROVISION 200 immersion VR system and attached flight helmet with a resolution of $360x240$. Compared to current VR technology standards, this VR system has become outdated, and therefore the effect of hyoscine with the current HMDs is unknown.

Lien et al. [134] showed that ginger effectively reduced *nausea* severity and even reduced the recovery period after the experiment. A cutaneous electrogastrography was used as a measurement tool for vasopressin infusion and plasma vasopressin determination. The findings confirmed that ginger could be successfully used against motion sickness and aligned with previous research on motion sickness [80]. However, ginger only alleviates symptoms of *nausea*, which are primarily related to motion sickness, other Oculomotor cluster symptoms, such as *disorientation* or *dizziness*, can still affect the user. The experimental setup implemented circular vection to induce *nausea* and did not use any VR technology. Therefore, it can be assumed that ginger might alleviate SiS in virtual environments. Further research is needed to assess the effect of ginger on SiS in VR applications.

Regan [177] reported that adaptation is a possible technique against SiS induced by HMDs. The adaptation process consisted of four immersive sessions. The time between the first, the second, and the third session was approximately four months. The time between the third and the last session was only one week. Each session followed the same instructions and used the same virtual environment as the previously conducted experiment on the frequency of occurrence and severity of side-effects of immersion in VR [177]. The measurement tools were the malaise scale and the SSQ. The results demonstrated that all participants experienced symptoms after the first session.

In contrast, 57% of the participants reported no symptoms after the last session. Although there was a slight increase in the Disorientation and Oculomotor cluster between the second and the third session, the adaptation was successful. It must be considered that all of the sessions were conducted in the same virtual environment, and therefore it can be debated whether adaptation is possible only when the environment does not change. The virtual environment should consist of the same or very similar elements so that the user's body would not notice any differences.

Domeyer et al. [58] investigated adaptation within a short period (one day) between the sessions

as a mitigation technique against SiS onset in a driving simulator. A short familiarization session preceded each testing session with the simulator. Nevertheless, some of the test sessions were conducted on the same day as the familiarization sessions and some two days later due to the counterbalanced study design. A Revised Simulator Sickness Questionnaire (RSSQ) by Kim, Parker, and Park [117] was used to evaluate SiS. The RSSQ is a modified version of the SSQ that extends the SSQ by adding a cluster (strain/confusion), eight items (*drowsiness, visual flashbacks, stomach awareness, confusion, vomiting, pallor, difficulty equilibrating, muscle stiffness for strain*), and changes the rating scale to a broader range from 1 to 10. It transpired that the adaptation process, in a sequence of a short session followed by a day of no exposure before the test session, decreased the SiS symptoms induced by a driving simulator. Thus, it could be a promising option to give users sufficient time to adapt to a particular simulated environment. However, this approach is relatively a time- and resource-consuming.

Keshavarz et al. [116] evaluated the effect of passive restraint of the upper body movements on SiS onset in relation to *age* during a simulated driving task. The participants were located in two groups related to their age. The FMS, the SSQ, and a postural sway test (with eyes closed) were used in the experiment. Additionally, a questionnaire measured presence, simulation realism, and vection. Presence was assessed using a scale from 0 (not at all) to 10 (very strong). Realism was assessed using a scale from 0 (very unrealistic) to 10 (very realistic). Vection was assessed using a scale from 0% (never) to 100% (constantly). After each test session, the participants were asked whether they feel sick or not. Following their answer (yes or no), the participants were divided into a sick and well group.

Furthermore, Keshavarz et al. [116] reported a higher level of realism and a stronger sense of presence during the restraint condition. No effect on vection was found in older participants in the unrestrained driving condition. SiS was significantly reduced in the restraint condition among older participants who were assigned to the sick group and experienced sickness in the unrestrained condition. Besides the positive effects, the upper body restraint approach was not evaluated with an HMD, which was reported to induce more SiS compared to projection displays [196]. Furthermore, the virtual environment's exploration is constrained due to limited movement capability, which could result in discomfort or willingness to terminate the session. Thus, further research would be needed to determine exactly how this approach affects SiS induced by HMDs.

System-related mitigation techniques

Keshavarz and Hecht [112] reported a significant alleviation of SiS when pleasant music was played during a bicycle ride video projection. The FMS and the SSQ were used to measure sickness onset. Additionally, the participants were asked to rate the music's pleasantness on a 7-point Likert scale (1 – very pleasant, and 7 – very unpleasant). The study evaluated three different types of music: relaxing instrumental music, neutral mainstream pop music, and stressful electronic music. The results of the study indicated that relaxing music likely reduced SiS.

Its subjective pleasantness grouped the music into pleasant and unpleasant music. The participants who perceived the played music as pleasant, regardless of the type, reported less SiS. Keshavarz and Hecht [112] have drawn attention to the assumption that sounds, such as engine sounds or traffic noise, do not have any effect on SiS. The background music should be part of a more sophisticated musical piece such as instrumental or pop music. Nevertheless, the participants' head movements were limited by a chinrest. Thus, it can be argued that the head restraint may contribute to reducing SiS onset due to the reported positive effect of the passive body restraint on SiS [116].

Regarding VR driving simulations, the use of music might be a promising approach for mitigating SiS. Music can be integrated as radio music within the car entertainment system where the user can choose between radio stations with different music types. However, the effect of music and the type of music on SiS induced by HMDs is unknown.

Keshavarz et al. [115] investigated the effect of different odors on SiS. The FMS and the SSQ assessed SiS severity. The participants were divided into three groups: pleasant odor (rose), unpleasant odor (leather), and no odor. The categorization of the odors was made according to a previous survey on odors and their level of pleasantness. The experimental setup was identical to a previously conducted experiment investigating different types of music and SiS [112]. It has been found out that only half of the participants exposed to the scents noticed them. Thus, these data was re-assigned to a fourth group (odor not noticed). The results showed that the SiS onset was significantly reduced among the pleasant odor group compared to the group that did not detect any odor. These findings demonstrate that not every user will be affected by a pleasant odor and therefore some users will experience less discomfort due to highly subjective scent preferences.

Furthermore, it can be argued that a head restraint, just as in the previous study [112], may contribute to reducing SiS onset as it was reported that there was a positive effect of the passive body restraint on SiS [116]. The effect of pleasant odor on SiS induced by an HMD during driving applications is unknown. Nevertheless, the use of a pleasant odor as a subtle air refresher might be a potential method for minimizing SiS.

D'Amour et al. [62] found that a continuous stream of fresh air directed to the user significantly reduced SiS. The FMS and the SSQ measured the sickness outbreak. The experiment setup was similar to previous studies [112, 115] except that the restraint of the head movements was removed, two fans were used to stream airflow to the participant, and vibration was added to the participant's seat. The participants were assigned to one of the following groups: control (no airflow and no vibration), airflow, vibration, and airflow and vibration. Only the fresh air approach is included as a mitigation technique in Table 3.3 because the other approachs showed no significant results regarding SiS. The study showed that the airflow can be quickly provided by fans that stream cool fresh air to the user. This technique is appropriate for almost every VR driving application. Moreover, this method can help sustain the body temperature at a comfortable

level during the simulation in order to reduce discomfort.

Curry et al. [48] compared static and dynamic driving simulators regarding the severity of SiS in virtual environments. Sickness outbreak was measured by the SSQ, which, in contrast with the studies above, was administrated verbally. It has been found that the dynamic driving simulator significantly reduced SiS compared to the static one. The study included an acclimation time of approximately five to ten minutes. It can be argued that the addition of motion cues did not on its own reduce SiS but that rather the combination of acclimation time and motion cues helped to reduce symptoms. Nonetheless, it is a standard procedure of driving simulation studies to include an acclimation time before the actual driving evaluation. The addition of a motion platform to a static VR driving simulation has the potential to reduce SiS onset.

Aykent et al. [10] found that SiS was significantly reduced with a dynamic driving simulator compared to a static one, which is in line with Curry et al. [48]. The measurement method was a Motion Sickness Dose Value (MSDV) and a questionnaire on perception due to psychophysics. The MSDV is a method for objectively rating motion sickness. The coefficients used in this method are calculated by the frequency and direction of vibration to which the body is exposed (ISO 2631-1:1997). In this study, Aykent and colleagues [10] applied an illness rating method originated from the MSDV where the scores ranged from 0 ("I felt good.") to 3 and greater than 3 ("I felt absolutely terrible."). The longitudinal, lateral, and vertical acceleration of the heads of the participants were used to determine the illness rating. The questionnaire on perception contained 12 items measured from 1 (very little) to 10 (very strong). The purpose of the questionnaire was to assess not only the perceived physical discomfort, like in the SSQ, but also to assess the participants' impressions (psychophysics), such as mental pressure, fear, and anxiety. The results showed that *nausea*, *dizziness*, and *eye strain* symptoms had lower values compared to the static driving simulator.

Moreover, it was observed that longitudinal head movements induced discomfort with the static driving simulator, and vertical head movements induced discomfort with the dynamic driving simulator. The addition of physical motion cues is a promising approach for alleviating SiS, and can be applied to VR driving simulations. Having said that, the sensation of a moving vehicle might cause nausea, such as motion sickness, and therefore physical motion cues should be applied with caution.

Simulation-related mitigation techniques

Fernandes and Feiner [64] were able to show that a subtle dynamic change of the FOV reduced SiS. The FOV was changing dynamically until it reached $80°$ or $90°$ soft-edged cutout FOV displayed on the HMD. The measurement tools were the SSQ, the PQ, and a discomfort score obtained by a question regarding the current state of discomfort graded from 0 ("how you felt coming in") to 10 ("want to stop"). Additionally, a post-questionnaire was used to assess whether

the participants noticed the FOV constraints and if they did whether the constraints restricted their experience. This technique allowed participants to immerse themselves in the virtual environment and experience less discomfort while exploring the virtual world. The findings suggested that the change in FOV was subtle and did not disrupt the level of presence. However, a reduced FOV decreases the level of presence [46]. The dynamically reduced FOV might not be entirely applicable to VR driving applications due to the restriction of the virtual environment's exploration and immersion. Nonetheless, further research focusing on this approach regarding SiS in HMD driving applications is necessary before any concrete conclusions can be made.

Ihemedu-Steinke et al. [95] evaluated the effect of the addition of visual assets (e.g., artificial intelligent vehicles and pedestrians) on SiS in a VR driving simulation. For the evaluation, the SSQ and a questionnaire to assess the virtual experience and enjoyment based on the PQ and a measurement tool by Lin et al. [136] were used. The results showed that the participants experienced significantly less sickness when the driving simulation included additional visual assets.

The study did not report the previous VR experience of the participants. Therefore, it is unclear whether they were using an HMD for the first time or had previous experience. The excitement of using VR technology for the first time could have suppressed the experienced discomfort, and therefore the participants may have stayed immersed in the virtual world for longer despite the felt discomfort.

As another study indicated that users may still enjoy the VR experience regardless of SiS [222]. The group with additional visual assets drove a little more than one minute longer than the other group. Thus, the mitigation technique of including visual assets could potentially be a solution against SiS in a short VR driving sessions (5-10 min). The approach of adding visual assets could enhance the VR experience not only by reducing the SiS onset but also by improving realism.

Whittinghill et al. [231] reported SiS reduction by adding a virtual nose to the center of the FOV of the HMD. The time duration was recorded as well as the electrodermal activity. The results showed a time delay of a few seconds before the sickness was acknowledged. This time delay is not sufficient for VR driving applications where a delay of a few seconds would not make a notable difference for the users. Their findings might have been more persuasive if more information about the research was given. The approach of adding a virtual nose is not entirely applicable to VR driving applications due to extensive user evaluation sessions.

Duh et al. [61] investigated the effect of an independent visual background on SiS in a driving simulator. The SSQ was used to assess SiS symptoms. The E2i questionnaire was used for the assessment of "sense of presence" and "enjoyment." This questionnaire was developed to assess engagement, enjoyment, and immersion in a virtual environment [135, 136]. The participants reported significantly fewer symptoms with the independent visual background than the no-visual background condition. A possible flaw of this approach is that the sense of presence might be interrupted. Contrary to that, the results showed that the sense of presence and enjoyment were

slightly higher in the independent visual background condition. However, the small sample size failed to provide sufficient evidence for a conclusive decision. The addition of an independent visual background has the potential to be applied against SiS in HMD driving applications.

Curtis et al. [49] proposed a mitigation technique of a virtual hand-eye coordination task against SiS. The SSQ assessed the induced sickness, the mitigation effectiveness, and the sickness outbreak. The study reported a significant decrease in SiS symptoms. Moreover, the authors noted that the participants carried out the task while still immersed in the virtual environment. This approach is not entirely suitable for VR driving simulations due to a possible distracting effect on the driver. Nonetheless, this approach may possibly be helpful for VR fully autonomous driving simulations to keep the users immersed for longer in the virtual environment without interrupting the immersion.

Some of the mitigation techniques have already been tested with VR driving applications and yielded positive results. Others have been successfully tested with SiS stimuli, which suggests that they could potentially be applied with HMD driving applications. Additionally, this work has shown that the mitigation techniques could unintentionally incorporate more than one technique. For example, a combination of two techniques such as pleasant music and head restraint [112], pleasant odor and head restraint [115], and acclimation time and motion cues [48]. Depending on the user evaluation's objective, specific factors related to SiS should be considered during the development of the VR driving simulation. The improvement of the hardware (e.g., *central processing unit, graphics processing unit*, and *HMD performance*) and also the right setup could reduce some of the factors (e.g., *refresh rate, FOV*, and *flicker*).

3.4 Predicting Simulation Sickness

Predicting SiS onset will significantly benefit not only the research areas, which are often using VR environments but also areas such as entertainment, automotive, aviation, and education. However, predicting when an individual will become sick is still a challenge due to the complex nature of SiS and the suddenness with which it occurs. Before we continue to review previous attempts to predict SiS, we should briefly overview the prediction models' algorithms.

3.4.1 Linear Regression

The goal of linear regression, also known as a linear model, is to predict an output value $y \in R$ from input samples x (see equation 3.1).

$$y_i = (b_0 + b_1 x_i) + \varepsilon_i \tag{3.1}$$

An outcome variable (y_i), also known as dependent variable, can be predicted from a model, and some error associated with that prediction can be calculated [66]. The error (ε_i) is the difference between the predicted and observed value of y for the i-th sample. The y_i is predicted from a predictor (x_i), also known as independent variable, and a parameter, b_1, related to the predicted variable that quantifies the relationship it has with the outcome variable. The parameter b_0 is a parameter that tells the outcome variable's value when the predictor is zero, also known as intercept. For evaluation of the linear model performance, is used an *R-Squared* (R^2), also known as a *coefficient of determination*. The R^2 shows the proportion of the variance in the dependent variable explained by the variance in the independent variable (predictor); it is a *Pearson correlation* coefficient squared [66]. The R^2 is a number between 0 and 1, which, when represented as percentage, 0 represents 0% and 1 represents 100%. The higher the R^2 is, the most effectively the predictors can predict the outcome. The R^2 is calculated from the *Total Sum of Squares* (TSS) and the *Residual Sum of Squares* (RSS), where $\hat{y}$ is predicted value of y_i, and $\overline{y}$ is the overall mean.

$$R^2 = 1 - \frac{RSS}{TSS} \qquad RSS = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad TSS = \sum_{i=1}^{n}(y_i - \overline{y})^2 \tag{3.2}$$

In many cases, more than one variable can be associated with the dependent variable. A regression with two or more independent variables is known as multiple regression. In order to add a predictor to the model, a coefficient b should be given for each added independent variable. The formula below pertains to multiple regression linear model, where b_1 is the coefficient of the first predictor (x_1), b_2 is the coefficient of the second predictor (x_2), and b_n is the coefficient of the nth predictor (x_n) [66].

$$y_i = (b_0 + b_1 x_{1i} + b_2 x_{2i} ... b_n x_{ni}) + \varepsilon_i \tag{3.3}$$

3.4.2 Machine Learning Methods

Classification is a supervised learning method that classifies new data based on the learned knowledge from a given data sample. Models built using this method are called classification models. Classification models can be binary or multi-class model as well as multi-label model. This section presents a brief overview of some machine learning algorithms used for classification in the current thesis. To better understand the results of the reported prediction models in **Chapter 9**, we will offer an overview of the algorithms and methods for evaluations.

Logistic Regression is a multiple regression, where the outcome variable is a discrete variable (e.g., gender - male and female) and the independent variable is continuous or categorical variable

[66]. Using a sigmoid function, this type of regression identifies the probability of predictor (x), also known as independent variable, producing a discrete outcome variable (y), also known as dependent variable. The equation for logistic regression is shown in Equation 3.4, where x is the independent variable which will be transformed and e is *Euler*'s constant equal to 2.718.

$$y = \frac{1}{1 + e^{-x}} \tag{3.4}$$

The sigmoid function creates an S-shaped curve that can transform any numeric value and map it into a value between 0 and 1 but without ever reaching the exact limits [217]. Applying the formula, the sigmoid function converts the independent variable into probability between 0 and 1 concerning the occurrence of the dependent variable. A value of 0 means no chance of occurrence, and 1 represents a specific chance of happening. For each data point, a different probability is calculated, and based on that, this data point is assigned to one of the classes. Logistic regression can be used for a multi-class model called a multinational logistic regression.

Support Vector Machine (SVM) is a classification algorithm, similar to logistic regression, that automatically adjusts the capacity of the classification function by maximizing the margin between the training data points and the class boundaries [22]. The margin helps with additionally supporting the new data points, which may violate the logistic regression data assumptions. Also, SVM helps to reduce the impact of outlying data points. Compared to the logistic regression, which tries to fit all data points, including the outliers, the SVM is less perceptive to such data points.

Linear Discriminant Analysis (LDA) is another classification algorithm that is commonly used with two classes. LDA is related to Fisher's linear discriminant [68], but the latter does not support the normal distribution assumption. LDA tries to find the combination of predictors that maximize the variance between the data groups while minimizing the variance within the data groups. For an input variable, the mean and the variance for each class are calculated. When the variables are more than one, the same statistical parameters for each variable are calculated over the multivariate Gaussian. Similar to other linear classification techniques, the LDA can only be used with normally distributed data.

Naive Bayes (NB) is a non-linear classification technique based on the Bayes Theorem of probabilities [183]. The NB assumes that there is an independence between the features, also known as independent variables. A feature vector is built from each data point based on its set of features. NB assigns a given data point to a most likely class based on its feature vector. One of the NB advantages is that it is straightforward to implement and it is especially useful for datasets with many features (multidimensional). Gaussian NB is a particular case of the NB algorithm,

which assumes that the features have a normal (gaussian) distribution and they are continuous variables. For the models evaluated in this thesis, a Gaussian NB will be used, abbreviated only with NB.

K-Nearest Neighbors (KNN) is a supervised algorithm that classifies input data points based on their location [69]. A new data point will be set to a class based on its distance to K number of nearest data points (neighbors). To predict the values of the new data points, the algorithm uses feature similarities to the already known data points. KNN is a simple and straightforward technique to implement. However, with the increase of the features number, its performance can get slower.

Decision Tree is another supervised algorithm used for classification and regression [28]. Decision tree classifier is a tree-like structure that starts with a single primary node and then splits into two descendent nodes after a decision is applied to it. Each of those internal nodes is a branch from the primary node and leads to additional nodes, which split into further nodes. In the end, the tree stops growing when there are no more splitting nodes. The last nodes in the tree are also called terminal nodes. Each terminal node is assigned to a class. There may be more than one terminal node assign to a specific class. In that case, the classification outcome is based on placing all terminal nodes per class together. In this thesis, the decision tree will be used as a binary classifier.

Random Forest (RFC) is a classification algorithm consisting of a collection of decision tree classifiers where for the kth tree, an independent identically distributed random vector is calculated, and each tree votes for the most popular class [27]. One of RFC advantages is that it is considered a highly robust method due to the usage of many decision trees. The model takes an average from all made predictions, and therefore the chance of experiencing overfitting is very low. However, the model is relatively slow in making predictions as it uses many decision trees. Moreover, it is difficult to interpret and thus not as easily understandable as the Decision Tree model because of the use of many trees.

Evaluation

Cross-validation is an evaluation method for measuring how well the model performed on unseen data [125]. Holdout cross-validation is an uncomplicated validation method where a random sample from the dataset is selected. This random sample is commonly called a test set. The rest of the dataset generally called a training set, is used for training the model. After the model is trained, an evaluation of the model is performed on the test set. The evaluation result is presented as a performance metric such as accuracy, precision, and recall. The results from

the holdout method should be used with caution as the method has a notable drawback; it is heavily dependent on the data allocated to the test set. Thus, different iterations of this method can produce significantly different results.

In order to overcome this drawback, K-fold cross-validation is generally used. In the K-fold cross-validation method, the holdout method is used for k-number of iterations instead of one iteration [125]. Every time a sample of data is used as a test set, the other $k-1$ data is used for the training set. For each k-iteration, a performance metric is calculated. In the end, the result of the evaluation is presented as an averaged performance metric across all iterations. In that way, the method is more robust to potential variance in the dataset as every data point is only once in the test set and $k-1$ times in the training set. However, this cross-validation method requires more computational power to run k-number of iterations. In turn, that could cost much more time when evaluating big datasets compared to the holdout validation.

A more exhaustive cross-validation method is Leave-one-out (LOO) cross-validation, where the number of iterations is equal to the number of data points in the dataset [125]. Each iteration has only one data point selected from the original dataset as a test set. All other data points are allocated in the training set. Thus, the LOO method could be even more computational demanding than the K-fold validation. For small datasets, that could be a plausible evaluation method, but this method may require substantial computer power and time for large datasets.

Confusion matrix is a form of table that displays the model performance. The confusion matrix consists of four different combinations between actual and predicted values. The actual values are described as true and false, and the predicted values are described as positive and negative. The four combinations are True Positive, True Negative, False Positive, and False Negative (Fig. 3.1).

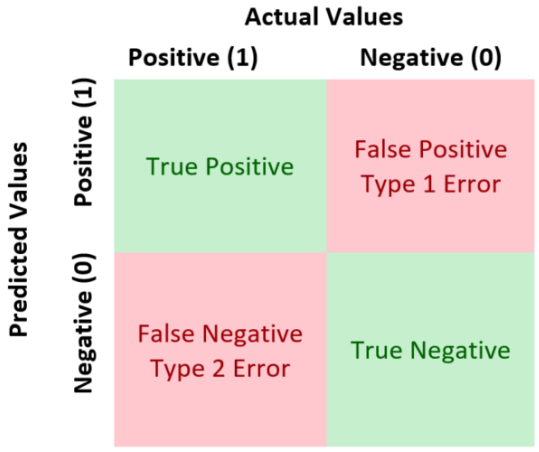


Figure 3.1: A confusion matrix showing the different combinations between the actual and predicted values.

True Positive (TP) is when the actual value is positive, and the predicted value is positive. **True Negative (TN)** is when the actual value is negative, and the predicted value is negative. **False Positive (FP)**, also known as a Type I Error, is when the actual value is negative, but the predicted value is positive. **False Negative (FN)**, also known as a Type II Error, is when the actual value is positive, but the predicted value is negative.

Performance measurements, such as accuracy, recall, precision, and F1 score, can be calculated from the confusion matrix.

Accuracy is calculated from the summation between the correct positive and negative predictions over the total values, and it presents overall how accurate the classifier is.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.5}$$

Precision is calculated from the correct positive predictions over all actual values, and it presents the classification model's ability to return only correct instances.

$$Precision = \frac{TP}{TP + FP} \tag{3.6}$$

Recall is calculated from the correct positive predictions over all predicted values, and it presents the classification model's ability to identify all correct instances.

$$Recall = \frac{TP}{TP + FN} \tag{3.7}$$

F1 score is a performance metric that combines recall and precision.

$$F1\ score = \frac{2 * Recall * Precision}{Recall + Precision} \tag{3.8}$$

Receiver Operating Characteristic (ROC) curves is another evaluation tool for model performance. The ROC curves show how well the model distinguishes between the classes compared to a specific threshold. The ROC plot is based on the calculated sensitivity (the TP value) and specificity (the FP value) for each instance in the dataset, and then the ROC curve displays the sensitivity against 1 - specificity [3]. The further the plotted curve is from the diagonal of the plot, which represents the prediction by chance, the better the model is identifying the classes. Area Under the Curve (AUC) is a model's overall performance metric, and it can be calculated from the ROC curve plot. The bigger the area under the ROC curve, the higher the AUC score is. The AUC can also be calculated for each class. In that case, the score shows how well the model identifies a particular class. The higher the score, between 0 and 1, the better the performance is.

3.4.3 Previous Prediction Models

A straightforward administration method of SiS prediction is the MSSQ-short [77]. Based on an individual's previous experience with motion sickness in childhood and adulthood, an estimation of her or his susceptibility to SiS is made. The questionnaire showed good predictability of motion sickness and, more specifically, predicting which individual will be motion-sensitive. However, the MSSQ-short is less accurate in identifying notion-resistant individuals [77]. Furthermore, the MSSQ-short is not optimized to assess past VR experiences, and thus, it might lack predictability of SiS in VR, although it is widely used in SiS research. In this thesis, the MSSQ-short will be used to evaluate the relationship between motion sickness history and SiS onset in VR.

Several previous studies have attempted to predict SiS. The most effective of them used correlation analysis and reported the relationships between the variables and the SiS rather than creating prediction models. Despite Kennedy, Dunlap, and Fowlkes [104] not creating a prediction model for SiS, they did perform an extensive meta-analysis on possible factors that could aid the development of models. Their analysis was based on more than 2000 articles on SiS and other forms of motion sickness, and they concluded that operational measures, which measure sickness in the actual environment, account for a large part of variance (67%). Another factor that also showed a high percentage of variance (38%) was a provocative test used to deliberately induce sickness. *Motion sickness history* also showed a good percentage of variance (34%), and it appeared in large numbers of studies. As much as 26% of variance came from psychological measures. This estimate was higher than reviews in the literature imply, and thus the contribution of personality might be underestimated. Therefore, the current thesis investigates the relationship between the users' personality and SiS onset in a VR driving simulation. Kennedy et al. [104] concluded that it was likely that the best prediction is based on a combination of operational measures, *motion sickness history*, psychological variables (personality and perceptual style), laboratory provocative tests, and physiological measures (autonomic and sensory function).

Kolasinski [126] proposed a linear model that explained 34.3% of the variance. The model included four variables that showed significance only concerning each other. Separately, the variables did not show a significant correlation to the prediction of SiS. The variables were *age*, *gender*, *mental rotation ability*, and *pre-exposure prototype value*.

Another prediction of SiS was attempted through Cybersickness Dose Value (CSDV), which was based on the MSDV used for predicting seasickness [200]. The CSDV quantifies the changes of scene movements along the vertical, horizontal, and radial axes through a metric called "spacial velocity" and multiplies the metric by the display, task, and subject-related scaling factors. An enhanced version of CSDV was presented, which included a weighted "spacial velocity," which presents navigation velocity and scene complexity [202]. However, the authors noted that other factors might influence SiS, and said that further research is needed to ensure better predictions. The display, task, and subject-related scaling factors were taken out from the metric. Chen, Ho, Lor, and So [39] reported that CSDV, including exposure duration, display's FOV, scene

complexity of virtual environments, and the navigation velocity as factors, could explain 82% of the variance for Nausea and 68% for SSQ total score.

Rebenitsch and Owen [173] built models to predict cybersickness that might explain close to half of the variance including as most influencing factors *vision correction* (17.9%) *motion sickness history* (10.8%). The best model explained 58% of the variance, including the score only from the carnival rides from the MSSQ and the three vision correction type variation (i.e., contact lenses, glasses, both).

Furthermore, Rebenitsch [175] proposed an individual model which can explain 37% of adjusted variance, including *motion sickness history*, *headache*, and *video game play* as factors. A model including only *motion sickness history* as an individual factor can explain 31% of the variance, which showed a high contribution rate in the multivariate model. A cross configuration model, which was developed from the literature, can explain an adjusted variance of 55%. However, the model included not only individual factors but also simulation and hardware factors such as *type of movement, type of controller, the realism of the simulation, sitting or standing position, FOV*, and *duration of the simulation*.

Recent studies have started to use machine learning models for the prediction of SiS in virtual environments. Hell and Argyriou [86], for example, described a novel framework for automated ratings on SiS using neural networks. Machine learning architectures based on deep neural networks were trained using data collected through VR rollercoaster applications to predict SiS levels. Two different datasets of features were built. The rollercoaster data points set consisted of absolute position, relative position from previous roller coaster point, current speed, and gravitational forces exerted at this point; including vertical and horizontal axes parameters: maximal positive force, maximal negative force, average positive force (only sampled at points with force greater than zero), average negative force, percentage of the ride where a positive force is exerted, and percentage of the ride where a negative force is exerted. The custom input parameters set consisted of maximum speed, average speed, total length, maximal downwards angle, maximal upwards angle, and type of rollercoaster. From these data, Hell and Argyriou [86] developed three different neural network models. As an output, each model had 20 output neurons, 5 outputs each for nausea, fun, intensity, and price levels. Each output neuron represents one out of 5 possible stars—with 5 being the most nauseous experience. The paper focused on the nausea level as an output. The best model contained the custom input parameters as features and achieved 50% of correct answers and *Mean Error* of 0.75. Despite the relatively small dataset, the approach showed promising results on SiS prediction in VR.

Kim et al. [118] proposed a framework called VR sickness predictor using the interaction model between the user's motion and the vestibular system. The framework extracts two types of features: perceptual motion feature through a visual-vestibular interaction model and statistical content feature that affects user motion perception. The results showed that the correlation between the method and the subjective SiS score was as much as 72%.

In another study, Jin et al. [96] presented a machine learning approach to predict the level of SiS for five VR games automatically. They developed a novel ranking-rating score to measure the ground-truth annotations for SiS. Three machine learning methods were compared: Convolutional Neural Network (CNN), Long-Short Term Memory Recurrent Neural Networks (LSTM-RNN), and Support Vector Regression (SVR). Two different datasets were used for the different models as follows. For the CNN and the LSTM-RNN: head movement, motion intensity, contrast, smoothness, and entropy. For the SVR: video game experience, VR experience, MSSQ, head movement, hue, saturation, brightness motion intensity, texture. The results showed that the LSTM-RNN performed the best with $R^2 = 0.87$ and *Mean Square Error* = 0.009.

Although more research has been conducted in recent years on predicting SiS, a few studies include physiological factors as a possible predictor in the context of HMD virtual environments. Martin et al. [142] presented machine learning models for SiS detection in a virtual environment based on physiological data. The data was collected from three VR games using a modern HMD, with a total duration of 30 minutes. The features were extracted from cardiac and skin conductance measures. Three machine learning regression methods were tested for the model evaluation: RFC, Gradient Boosting, and SVM. The best results came from the Gradient Boosting model with $R^2 = 0.48$ and *Mean Square Error* = 1.05.

However, several shortcomings of the work by Martin et al. [142] have been highlighted. Firstly, the data was obtained from 27 participants, of whom only five were female. The low ratio of female to male participants can increase the model bias towards male participants. A more gender-equal dataset is needed to overcome this flaw in the model as earlier research showed differences in some cardiac measures across genders [124]. Secondly, the experimental description did not include information concerning the operationalized VR setup, such as technical details regarding the system's computing and video graphic power. The VR game description and the participants' interactions within the VR environments are sparse as well. Lastly, the frequency domain's extracted cardiac features might provide inaccurate information as the sampling frequency was low for that type of data extraction. The frequency of recording in the study was set to 128 Hz when a sampling frequency of 250 Hz or more would provide excellent results for cardiac signal analysis in time and frequency domain [130]. When only the time domain is used, a frequency of 100 Hz is also acceptable. Like the cardiac signal, the EDA signal was recorded with the same sampling frequency of 128 Hz - a low sampling rate for this type of signal. The low sampling frequency is a source of concern about how accurate the model is when the extracted features might be inaccurate.

Altogether these findings show that prediction models of SiS remain an ongoing research effort, especially in the context of VR driving simulations that are modelled on real-life traffic situations. This provides an opportunity to develop a model based on objective data to differentiate simulation sick from non-sick VR users within a modern HMD virtual driving environment. Thus, in this thesis, we investigate if such a model can be built and to what extent this model will fit the collected data from the conducted later VR driving experiments.

3.5 Summary

In this chapter, we have reviewed previous work on SiS related to HMD driving simulations. Although efforts to solve VR-induced SiS have been ongoing, a concrete solution to the problem is still yet to be found. The survey conducted on SiS in modern HMD driving simulations demonstrated that a few studies explicitly investigated SiS within driving applications. The overall findings of the studies showed that modern HMD driving setups induced more discomfort than conventional display setups. The level of immersion and the sense of presence were higher within VR driving simulations. Nevertheless, few of the surveyed works included more than 30 participants. The small sample size can produce issues with the generalization of the results on the population. Furthermore, we observed from the reviewed works that a few driving studies utilizing HMDs measured the felt discomfort through questionnaires and physiological sensors. That shows that results rely strongly on the participants' subjective responses. Another observation is that gender difference was reported in fewer studies without being on focus. However, this aspect was evaluated without the sample size being balanced, which might lead to inaccurate conclusions regarding SiS and gender. This suggests that there is a lack of gender-based research in the context of modern HMD driving simulations.

With the rapid expansion of vehicle automation, automated driving has the potential to become an integral part of the automotive industry's future. That would increase VR driving application utilization in the early development stage of the cars' interior development. Thus, we have discussed previously evaluated mitigation techniques to present possible methods that apply to SiS in VR driving applications. Some of the mitigation techniques have already been tested with VR driving applications and yielded positive results. Others have been successfully tested with SiS stimuli, which implies their potential to be applied with HMD driving applications. Additionally, the findings from Section 3.3 have shown that the mitigation techniques could unintentionally incorporate more than one method. Such combinations of two techniques were pleasant music and head restraint, pleasant odor and head restraint, or acclimation time and motion cues. Depending on the user evaluation's objective, specific factors related to SiS should be considered during the development of the VR driving simulation. The improvement of the hardware (e.g., central processing unit, graphics processing unit, and HMD performance) and the right setup could reduce some of the factors (e.g., refresh rate, FOV, ergonomics, and latency).

We have also described classification machine learning methods that we will use later for predicting SiS, and we have discussed previous prediction models. The earlier prediction models ranged from conventional statistical models to more sophisticated machine learning models. Only one of the studies proposed a machine learning model based solely on physiological data collected partially through a VR racing game. The study showed similarities with our work. However, we chose to use a classification approach to predict SiS using solely physiological data rather than a regression approach, as proposed in the study. Previous attempts to predict SiS provide valuable insights on already tested models and factors related to SiS. With this knowledge, and

considering the technical limitations of the current work, new models are proposed, which are described and evaluated later in the thesis (**Chapter 9**).

4 Methodology

> That which is measured improves.
> That which is measured and
> reported improves exponentially.
>
> *Karl Pearson*

In order to evaluate some of the SiS factors and to collect data for building prediction models, we carried out three experiments. This chapter gives an overview of the methodology that was followed in every experiment. The VR system architecture outlines the operationalized hardware components during the user evaluations. As a central part of the VR system, the virtual environment itself is described, including a visual presentation from the actual driving VR application. The same measurements, subjective and objective, were used in every user evaluation to fulfill the goal of this thesis, which complied with the evaluation design. One of the questionnaires is considered a "standard" in the current body of SiS research, and thus, it has been used very often in the SiS studies. Other tools, such as a questionnaire for personality determination, are not so often used in VR experiments. Although objective measures such as physiological signals have been previously used in the SiS research, the results are often controversial. Therefore, including this measure could contribute to the SiS prediction and bring valuable new insights regarding their relationship to SiS. Furthermore, as a significant part of this thesis, collecting physiological data was necessary for SiS prediction modeling. The chapter finishes with a brief description of the procedure, which shows the conducted steps during the experiments.

4.1 Experimental Design Overview

In order to evaluate the VR driving setup and investigate the SiS outbreak, two studies were conducted. The collected data from the studies was used for three different experiments described

in the next chapters. The only difference between the studies was the type of driving (e.g., automated and standard driving). The first two experiments had a 2 x 2 factorial design, and the last experiment had a 2 x 2 x 2 factorial design. The first experiment evaluated a VR automated driving simulation including two factors, motion (with motion, without motion) and gender (male, female). In total, 62 participants took part in the experiment. The second experiment evaluated a VR standard driving simulation (e.g., the vehicle has a driver), including the same factors as the first experiment. The sample consisted of 63 participants. The third experiment investigated whether the different types of driving (automated, standard), the addition of motion (with motion, without motion), and gender (male, female) affect the SiS. The total number of participants was 66. The data was acquired from the participants of the previous two experiments who took part only in one driving condition, e.g., either in automated driving or standard driving condition.

4.1.1 System architecture

The VR driving simulation system consisted of an HMD, including a tracking system, a 3 DOF moving platform with four pneumatic actuators, a PC for rendering the driving scene, and physiological sensors (see Figure 4.1 and Figure 4.2).

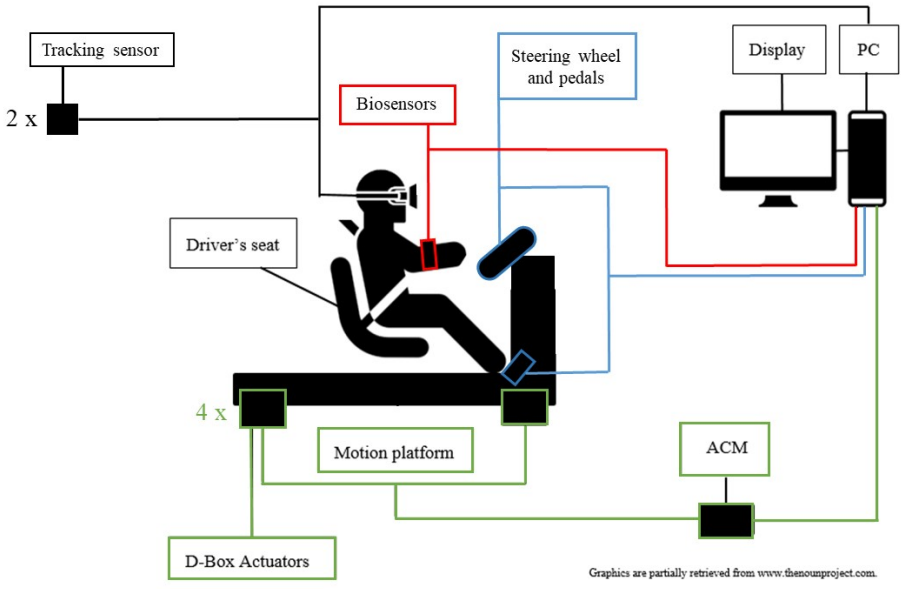


Figure 4.1: Concept of the VR driving simulation with an HMD operationalized in the experiments.

As an HMD, the HTC Vive Pro[1] with a FOV of 110°, uses a binocular view and a resolution of 1440 x 1600 pixels per eye was used. The two tracking sensors were SteamVR Base Station 2.0. They were placed diagonally opposite to each other to track the head movements of the participant for correctly updating the VR environment. In that way, the whole setup was covered by the

[1] https://www.vive.com/us/

Figure 4.2: Setup of the VR driving simulation including the HMD (HTV Vive Pro), motion platform(D-Box), driving seat, steering wheel, and pedals (Fanatec).

tracking sensors, which provided almost no latency. The HTC Vive Pro tracking is outside-in. The Base Stations are sending light signals to track the position of the HMD.

A 3 DOF moving platform manufactured by D-Box[2] was chosen as hardware. Four actuators, vertically attached to each corner of the moving platform, can simulate pitch, roll, and heave movements. The pitch movements simulate breaking and acceleration; the roll movements simulate turns while driving; the heave movements simulate the road surface. Each actuator has a maximum payload of 227 kg, a maximum velocity of $\pm$ 100 mm/sec, a maximum frequency range of $0 - 100$ Hz, a maximum acceleration of 1 *g-force*, a maximum angle of $15°$, and a stroke of 152.4 mm. Every two actuators had one master box. The master box translated the signals

[2]http://tech.d-box.com/training-and-simulation/

from the Actuator Control Module and controlled the actuators. The module was connected to a high-end PC, where a motion code was executed. The motion code is a software product that uses motion algorithms to respond to user commands. The code was required for the moving platform to behave as closely as possible to the virtual environment's actions. Motion signals are sent to the motion controller, decrypting the signal and sending the queue to the motion actuators. The development and the implementation of the motion code, which is compatible with Unreal Engine 4, were handled by an external company. In order to create an immersive environment, a car seat was added to the platform for a more realistic experience.

On the same PC, a $3D$ scene of the driving simulation is displayed. The virtual scene was powered with a 3.6 GHz Intel Core $i7$ processor with 128 GB of memory and an NVIDIA GeForce RTX 2080Ti graphic card running on Windows 10. The movement induced by the platform was aligned to the driving maneuvers shown in the VR driving simulation. To create an immersive environment, a car seat was added to the platform for a more realistic experience.

For the experiment with the standard driving, pedals and a steering wheel, manufactured by Fanatec[3], were added to the setup. The pedals setup was ClubSport Pedals V3 with a vibration feedback on the throttle, and brake The steering wheel setup included a ClubSport Wheel Base V2.5 and a ClubSport Steering Wheel BMW GT2. As the gearbox was set to automatic, only two out of three pedals were used in the experiments.

Compared with the VR driving setups presented earlier in this thesis (see Chapter 3), our VR setup has a moving platform, and physiological sensors integrated (Table 4.1). That way, the researcher can collect not only subjective data through questionnaires but also objective data through the sensors. The moving platform provides physical motion feedback while the virtual vehicle is driven. Our VR driving setup is technically innovative and can be utilized as a helpful SiS evaluation tool.

[3]https://www.fanatec.com/eu-en/

Table 4.1: A comparison table between the VR driving setups previously described in **Chapter 3** and our VR setup. CPU stands for central processing unit, GPU - graphic processing unit, ECG - electrocardiography, EDA - electrodermal activity, EMG - electromyographic

HMD	PC (CPU/GPU)	Steering wheel	Pedals	Car seat	Moving platform	Bio sensors	Reference
HTC Vive	-	Yes	Yes	Yes	No	No	Walch et al. [225]
Oculus Rift CV1	- / NVIDIA GeForce GTX 970	Yes	Yes	No	No	ECG, EMG, EDA	Eudave and Valencia [63]
HTC Vive	Intel Core i7-6700K 4 GHz, 32 GB / 2 x NVIDIA GeForce 1080	Yes	Yes	Yes	6 DOF	No	Ropelato et al. [185]
Oculus Rift	Intel Core i7-6700K 4 GHz, 16 GB / NVIDIA GeForce GTX 980 12 GB	Yes	Yes	Yes	No	No	Suwarno et al. [216]
HTC Vive	Intel Xeon 64 GB / NVIDIA GTX 1080Ti	Yes	Yes	Yes	No	No	Venkatakrishnan et al. [219]
Oculus Rift	Intel Core i7-6700K 4 GHz, 16 GB / NVIDIA GeForce GTX 980 12 GB	-	Yes	Yes	No	HR	Setiowati et al. [194]
Oculus Rift CV1	-	Yes	Yes	Stool	No	No	Curry et al. [47]
HTC Vive Pro	**Intel Core i7 3.6 GHz, 128 GB / NVIDIA GeForce RTX 2080Ti**	**Yes**	**Yes**	**Real car seat**	**3 DOF**	**ECG, EDA respiratory effort**	**Our VR setup**

4.1.2 Virtual Environment

The virtual environment was developed with the game engine software Unreal Engine 4.20[4]. Game developers widely use this game engine because it can deliver high-quality graphics that can be optimized for VR applications. Its support of physically based shaders together with increasingly powerful graphic processors enable almost photorealistic, high-definition real-time graphics. High-quality graphics are desirable elements in the development of VR driving simulations [150]. The BMW Group uses the Unreal Engine 4 in the development and evaluation of interior concepts at the Department for Interior Development.

A high-quality BMW car model and high-quality environmental assets, such as trees, roads, traffic signs, buildings, and vehicles, were provided by the Department of Interior Development at BMW Group. In addition, this included a traffic system and city environment, which was used as a driving scenario. This environment was based on the typical traffic system in a German city and the futuristic look of a Singapore city. This scenario was chosen because the VR environment was planned to be used for future interior concepts, and therefore, the city presented a possible future environment. The virtual environment included regular traffic without pedestrians or cyclists. The participants had to follow the regular traffic rules and the traffic signs. They had to drive or be driven through a roundabout, a sharp turn, and traffic lights (Fig. 4.3). The total driving time was set to 24 min or until the participants decided to stop, whatever came first. In the standard driving study, the participants had to pay attention and follow the blue arrows marked on the road, as they showed the correct route. However, this road attention was not necessary for the automated driving study as the participants were driven.

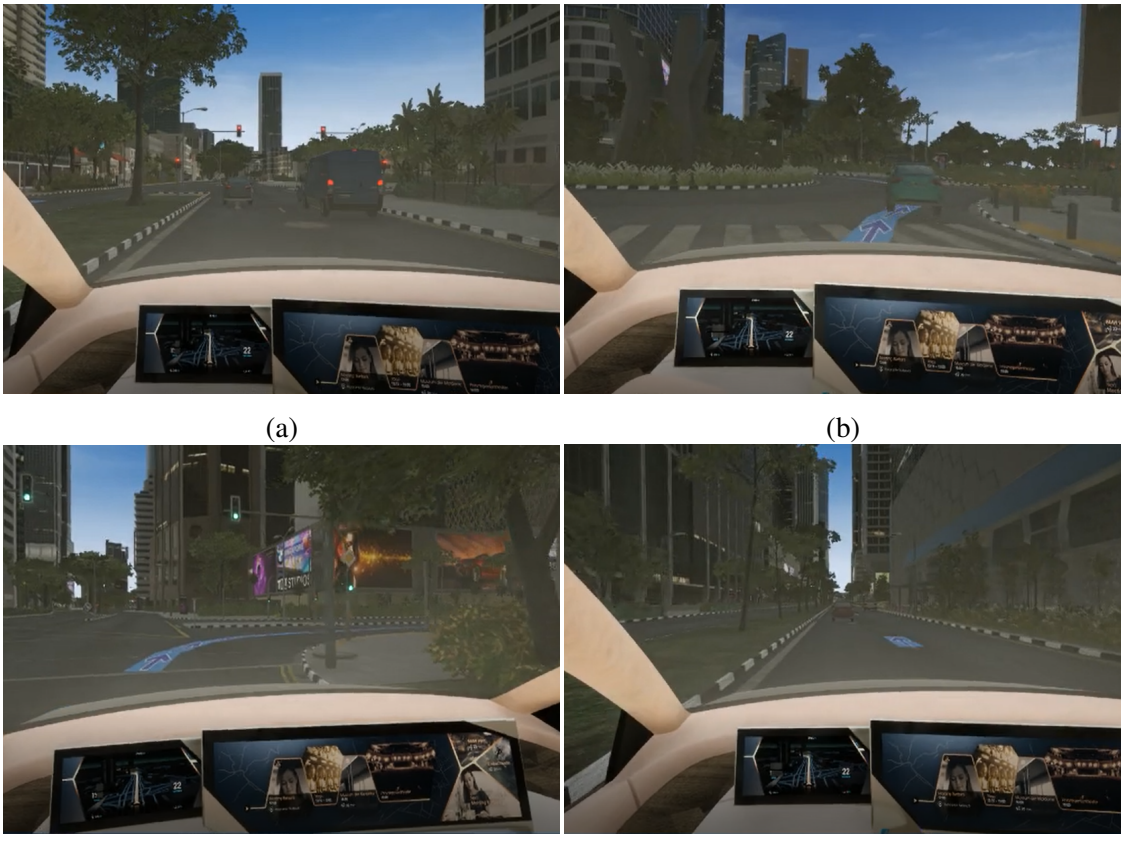
4.2 Measurements

4.2.1 Behavioral Measurements

The variability in the behavior during VR driving can strongly individual. However, previous research pointed out that some participants tend to change their head movement behavior within the VR environment (see **Chapter 2**). Thus, the location and the rotation of the head and virtual vehicle were recorded during the VR driving to investigate a possible relationship and predictive power to SiS. Furthermore, the vehicle's acceleration and speed might be SiS predictors contributing significantly to the SiS outbreak.

It should be noted that the axes of the HMD's coordination system do not match straight the cardinal head axes. As shown below, the x-axis is the left-right axis, the y-axis is the up-down axis, and the z-axis is the front-back axis in the HMD (Fig. 4.4). However, in the cardinal

[4]https://www.unrealengine.com/en-US/



(a) (b)

(c) (d)

Figure 4.3: Screenshots from the virtual environment during the standard driving experiment. The images show the diversity of turns in the city driving scenario, including traffic lights (a), roundabout (b), turns (c), and straight road segments (d).

coordination system, the x-axis is the front-back axis, the y-axis is the left-right axis, and the z-axis is the up-down axis (see **Chapter 2**)

Location

Location of the vehicle and the HMD was recorded. Both variables were used to plot the location during the driving simulation on a graph when required. The vehicle location was acquired in order for the current vehicle velocity as well as the average velocity to be calculated. The data was recorded with x, y, and z values. According to the physics formula (4.1), the velocity can be calculated from a distance divided by the time. To calculate the distance, the difference between each recorded sample was calculated for each axis x,y, and z and recorded as a new value. The new value was up to the power of two and recorded under calculated velocity value for each x,y, and z-axis (e.g., V_x, V_y, V_z). Then, the Pythagorean theorem was used to calculate the distance. The results were then converted to km/h values using the reference that one Unreal unit is equal

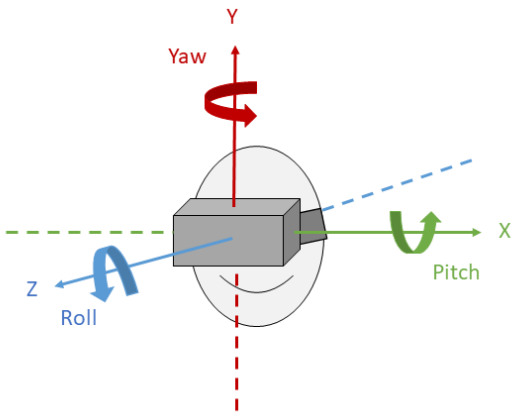


Figure 4.4: A coordinate system of 6 DOF. The position is listed as a set of x,y,z coordinates, where x is for left-right, y is for up-down, and z is for forward-back. Orientation is listed as a quaternion where the pitch is leaning forward or backward (rotation x-axis), yaw is rotating left, or right (rotation y-axis), and roll is bending left or right (rotation z-axis).

to one meter.

$$Velocity = Distance \ / \ Time \tag{4.1}$$

$$Distance = \sqrt{V_x^2 \ + V_y^2 \ + V_z^2} \tag{4.2}$$

Rotation

Rotation of the vehicle and the HMD was recorded. The vehicle rotation was acquired in order for the current rotation of the vehicle to be calculated. Additionally, the HMD rotation was recorded to get insights on participants' head rotation. The data was recorded as a quaternion, which is a specialized representing 3D orientation reference as a 4-dimensional number (w, x, y, z) in order to correct for strange behaviors encountered when rotating 3D objects. For calculating the correct head rotation, the quaternion acquired from the head rotation was multiplied by the inverted quaternion acquired from car rotation 4.3.

$$New\ Quaternion_{head} = Quaternion_{head} \ * \ Quaternion_{car}^{-1} \tag{4.3}$$

In order to convert the quaternion to Euler angles, the following formula was used. The roll, pitch, and yaw rotations are represented from ϕ, θ, and ψ, respectively (4.5, 4.6, 4.7). The results from

the conversion were recorded in radians and then converted to degrees.

$$q = [q_w\ q_x\ q_y\ q_z]^T = [q_0\ q_1\ q_2\ q_3]^T \tag{4.4}$$

$$\phi = \arctan(\frac{2\ (q_0q_1 + q_2q_3)}{1 - 2\ (q_1^2 + q_2^2)}) \tag{4.5}$$

$$\theta = \arcsin 2\ (q_0q_2 - q_3q_1) \tag{4.6}$$

$$\psi = \arctan(\frac{2\ (q_0q_3 + q_1q_2)}{1 - 2\ (q_2^2 + q_3^2)}) \tag{4.7}$$

4.2.2 Physiological Measurements

Physiological sensors, such as BVP, ECG, EDA, and respiration, from Plux Biosignals[5], were used for data acquisition (Fig. 4.5). The four sensors were connected to an 8-channel wireless hub, including eight generic inputs and one ground. The resolution of the hub is up to 16-bit (per channel) with a sampling rate of up to 3000 Hz per channel. The communication was through Bluetooth Class II with a range of up to 10m. All sensors were operated at a sample rate of 1kHz (i.e., 1000 samples were recorded for every second). The physiological signals HR, SCL, and RR were collected using the sensors. For the HR, a 1-lead ECG attached to the right side of the participants was used. The same signal was collected through the BVP sensor attached to the index finger of the hand. The double record of the HR signal was used as a measure against data loss. The SCL was assessed through electrodermal activity sensors attached to the third and fourth fingers of the non-dominant hand. The RR was collected using a piezoelectric respiration sensor attached to the abdomen via an elastic strapped belt. For building the SiS prediction model, from the physiological signals, additional features were extracted. The description of these features is provided in **Chapter 10**.

4.2.3 Questionnaires

Pre-questionnaire

A pre-questionnaire via an online platform LimeSurvey [6] was given to assess social-demographic information (e.g., *age, gender, education*) and *previous experience with video gaming, VR*, and

[5]https://www.biosignalsplux.com
[6]https://www.limesurvey.org

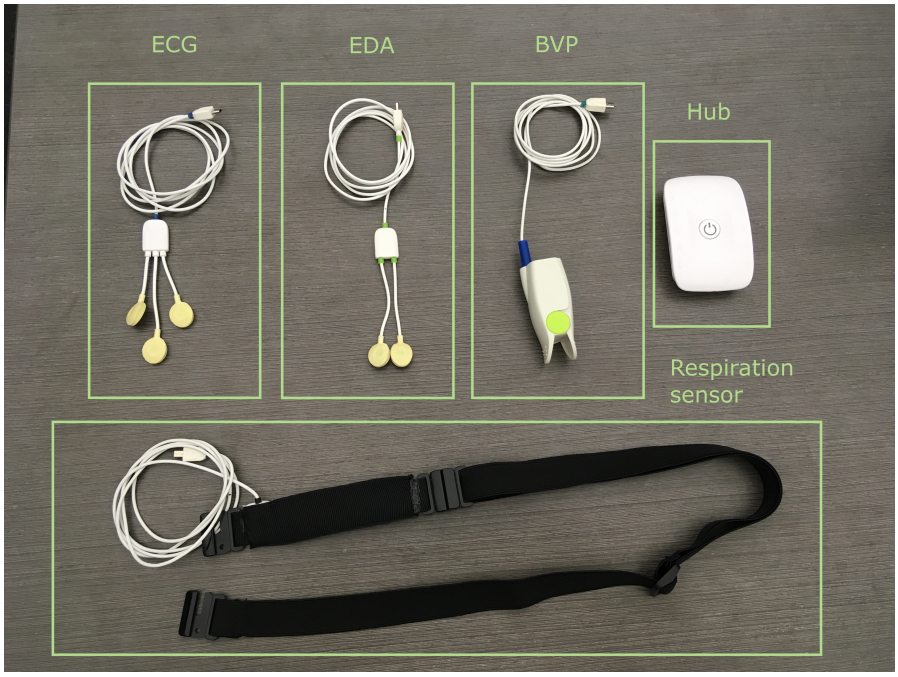


Figure 4.5: Physiological sensors used in the experiments.

driving simulators (see Appendix B). The participants' driving habits (*driving frequency* and kilometers(km) per year) were assessed with two items, *"How often did you drive a car in the past 12 months?"* and *"What is your annual mileage in kilometers (including vacation and business trips)?"*. The first time was measured on a scale from 1 (Never) to 5 (Daily), and the second item was measured on a scale from 1 (No driving) to 7 (More than 40 000km per year). A higher score indicates higher *driving frequency* for the first item and higher mileage per year for the second item. Furthermore, the participants needed to rate statements regarding their video gaming habits and experience with VR as well as driving simulators on a five-point Likert-Scale, ranging from strongly disagree to strongly agree. Higher scores indicated more experience with video games, VR, and driving simulators. For example, *"I am playing video games daily (e.g., PC or console games)."* Questions regarding the *physical activity* of the participants were added to the pre-questionnaire in order to explore the possibility of a connection between *physical activity* and SiS. The questions were the following: *"I am physically active in my daily life (e.g., climbing stairs, walking)."* and *"I am often doing Sport during the week."* The five-point Likert scale, ranging from strongly disagree to strongly agree, was used as a measurement scale.

Questions regarding personality traits were added to the pre-questionnaire. The questions were adopted from NEO-Five-factor Inventory short version for the German population [129]. The questions represented as 12 items were divided into two factors: *neuroticism* and *extroversion*. For example, an item from the *neuroticism* factor is *"I feel often tense and nervous."* An example from the *extroversion* factor is *"I have many people around me gladly."*

Additionally, the participant's history regarding motion sickness was assessed with the MSSQ-

Short [77]. The questionnaire was integrated into the pre-questionnaire. The MSSQ-Short is divided into two parts, A and B. Part A should be rated regarding their childhood and part B regarding the last ten years on a 5-point scale (never traveled, never felt sick, rarely felt sick, sometimes felt sick, frequently felt sick). Both parts are asking the participants to rate how often they felt sick during the ride in nine different types of transportation (e.g., cars, boats, buses, and airplanes). The minimum score is 0, and the maximum score is 54. A higher score indicates a higher susceptibility to motion sickness.

On-spot Questionnaire

At the beginning of the VR session, an on-spot questionnaire was given to the participants. The questionnaire included items regarding *caffeine consumption, sleep deprivation, current mood*, current well-being state, presence, and SiS.

The participants were asked to answer the item regarding the caffeine intake, *"I consumed in the last* 4 *hours caffeine (e.g., Coffee, Coca-Cola, Red Bull)."*, on a 5-point Likert scale, no caffeine at all - very much caffeine (more than 4 cups). Regrading the *sleep deprivation*, the participants were asked to answer three items regarding their sleep in the last seven days, in the last night, and if they feel currently tired. The scale was a 5-point Likert scale, totally disagree - totally agree.

The participants' current mood was evaluated before and after the VR driving simulation with a Self-assessment manikin (SAM) [24]. SAM is a questionnaire using pictures as items for measuring the affective dimensions of *valence, arousal*, and *dominance*. In this work, only two dimensions were used to evaluate participants' *arousal* and *valence* level (Fig. 4.6). The questionnaire was administrated as a picture questionnaire where each picture had value from 1 (Negative) to 5 (Positive) for *valence* and from 1 (calm and relaxed) to 5 (restless and exited) for *arousal*.

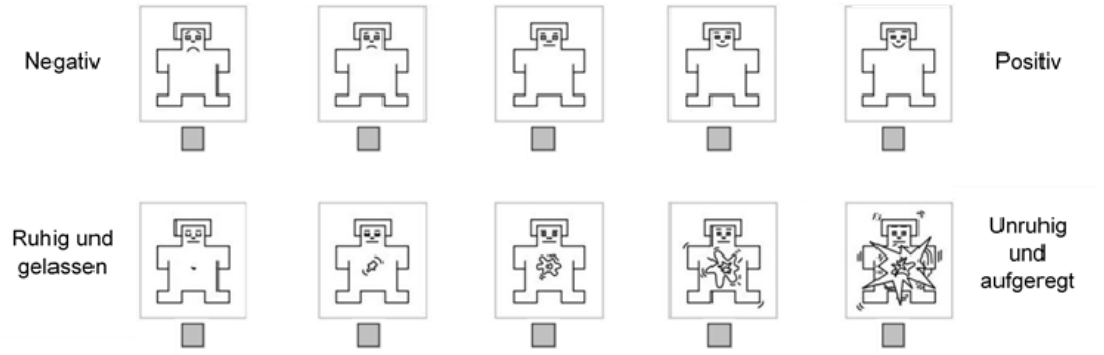


Figure 4.6: An example of SAM which was given to the participants before and after the virtual driving.

During the driving simulation, the experienced SiS level was assessed through the FMS [111]. Every 3 min, the participants gave a score on the FMS scale and reported their most severe

symptoms verbally. The higher scores indicate a severer SiS. The SiS was immediately assessed after the VR driving by the SSQ [106].

The last part of the on-spot questionnaire was to assess the presence during the simulation through the iGroup Presence Questionnaire (IPQ) [191]. It consists of 15 questions where each question could be answered once through a 5-point Likert scale (strongly disagree - strongly agree). The higher scores indicating a higher experienced sense of presence. Four questions regarding enjoyment [136] were included in the IPQ questionnaire to measure the users' emotional reaction to the VR simulation.

Follow-up Questionnaire

This questionnaire was administrated one hour after the VR session via the online platform LimeSurvey. However, some of the participants filled it out later than one hour. The questionnaire was the SSQ, including all 16 items. In this work, to differentiate the on-spot SSQ and the follow-up SSQ, the latter is abbreviated as FSSQ.

4.3 Procedure

An overview of the procedure is shown in Figure 4.7. The same procedure was applied for all three experiments, including filling in a consent form, receiving a short information about the driving simulation, and filling in questionnaires. A further description of the design and procedure for each experiment is presented in separate chapters regarding the experiments.

An online version of the pre-questionnaire described in this chapter earlier (see Appendix B) was sent to the participants before the experiment's start via online survey platform LimeSurvey. On the day of the experiment, the participants signed a consent form (see Appendix E) and filled in the first part of the on-spot questionnaire (see Appendix C). After that, the user was accompanied to the VR setup, and the physiological sensors were attached to the upper body of the participant. Then a physiological data for 5 min was recorded, and it was used as a baseline in the data analysis. During that time, the driving simulation was not started yet; thus, the participant's body had the chance to reduce the stress due to entering a novel environment. Another 2 min allowed the participant to look around and got accommodated to the HMD. During that time, a calibration of the HMD was performed. The driving simulation started after the introduction period was over.

The participants had only one task during simulated driving, which was to examine the interior. Before starting the driving session, the FMS and how it works were explained. The participants were asked at the beginning, every 3 minute, and at the end of the simulation to rate their current well-being state on a scale ranging from 0 (no sickness at all) to 20 (frank sickness) and

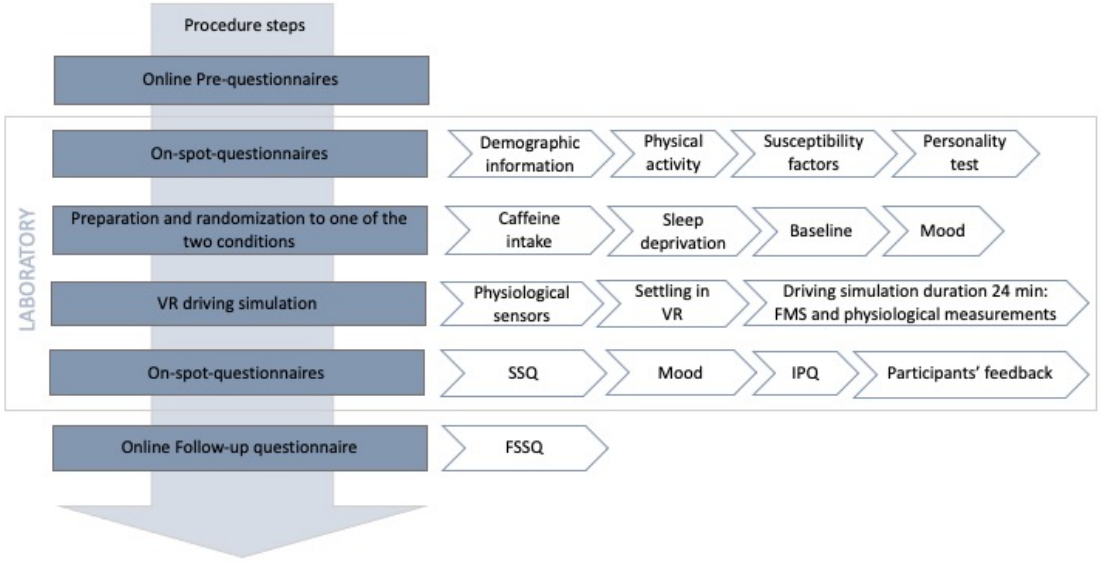


Figure 4.7: The procedure including all steps, questionnaires, and physiological measurements.

report their symptoms, if any verbally. The researcher asked how the participants were currently feeling, including the FMS range, three to four times at the VR driving's begging. After that, the researcher asked only shortly how the participants felt without mentioning the FMS as the participants were already acquainted with the scale. If participants stated a rating above 15, they were offered the option to terminate the VR driving simulation early in order to avoid frank sickness. Of course, the participants could end at any time the simulation as well.

Additionally, head location and rotation, and vehicle location and rotation in the virtual world were recorded. These behavioral data were recorded continuously during the experiment, as well as the physiological data. After completing the driving simulation, the participants were released from the sensors, and they left the setup. In the end, the second part of the on-spot questionnaire was given. The last element of the on-spot questionnaire was a text field left for the participants to write down their comments. They also had the chance to give their feedback directly to the researcher. The follow-up questionnaire via LimeSurvey was given one hour after the experiment to follow the SiS development over time (see Appendix D). Each participant drove for 24 min or until she or he wanted to stop, whatever occurred first. All questionnaires were administrated in the German language.

4.4 Summary

We have summarized the methodology of the experiments which we will present in the next three chapters, namely, **Chapter 6, Chapter 7**, and **Chapter 8**. The methodology overview

included the design of all experiments and the cross-reference design between them as well as the description of the VR system and the virtual urban environment. The chapter also presented the measurements used in the experiments to assess the SiS and factors related to SiS. As we now have an overview and knowledge of the methodology, we will continue to the moving platform and physiological signals implementation in the next chapter.

5 Technical Realization

> Any sufficiently advanced technology is indistinguishable from magic.
>
> *Arthur C. Clarke*

A significant part of the user evaluation setup preparation is the implementation of the motion platform and physiological sensors (described in the previous chapter). Before we describe the implementation of these elements, we will discuss VR driving setups utilized in earlier studies. The hardware implementation of the motion platform was carried out by an external partner company and supervised by employees at the Department of Interior Development at BMW Group. Employees conducted the software implementation and the mapping of the platform's motion into the virtual world. Therefore, no detailed description of the motion platform implementation is provided in this thesis.

Another essential implementation for the VR setup is the implementation of the physiological sensors. The sensors should continuously provide signals from an individual's body, which responds to the virtual environment during the virtual driving with minimal limitation of one's natural movements. Thus, the sensors are connected via wireless transmission to the PC. In order to complete the user evaluation preparation, two pilot studies were conducted to test the feasibility and readability of the VR system. Furthermore, two of the main SiS factors that were evaluated in the experiments were assessed in the pilot studies, namely, motion platform and type of driving.

The physiological signals implementation was part of a master thesis, supervised by the author, and written by Stefan Büttner [35], during a collaboration between the BMW Group and the University of Augsburg, Germany.

The first sub-chapter is based on the following publication:

- Rangelova, S., Decker, D., Eckel, M., & André, E. (2018, July). Simulation Sickness Evaluation While Using a Fully automated Car in a Head Mounted Display Virtual Environment. In International Conference on Virtual, Augmented and Mixed Reality (pp. 155-167). Springer, Cham [169].

5.1 Moving Platform Implementation

A moving platform was integrated into a static VR driving simulation setup in order to add missing motion cues to the VR user experience. Furthermore, based on the previous studies described in **Chapter 3**, the moving platform should not merely be added, but should also be synchronized to the visual output of the driving simulation. This way, SiS onset should be minimized, and a more pleasant VR experience should be delivered.

The VR driving system consisted of an HMD with a tracking system, a moving platform with four pneumatic actuators, and a computer for rendering the driving scene. The Oculus Rift CV1 with a FOV of 110° and a resolution of 2160 x 1200 pixels per eye was the used HMD. The moving platform was provided externally by the BMW Group.

5.1.1 Evaluation

A fully automated driving study with a within-subjects design was conducted. The participants were exposed to a scenario that did not require any intervention or monitoring from the user. Two conditions were presented: a scenario without and a scenario with an additional moving platform. Each participant took part in both scenarios with a 24-hours gap between the sessions. Approximately 80% of the participants started with the static condition due to minor organizational issues. The reminder started with a dynamic scenario. The total duration of the driving simulation was approximately 11 min. The participants were instructed to sit comfortably on the driver's seat and freely explore the car's interior and surroundings.

The simulation contained a traffic environment with no other movable visual assets (i.e., no traffic situations with vehicles or pedestrians). The driving simulation started with a right turn to a terminal of an airport. There the car stopped for 5 seconds and afterward continued towards a highway. Before reaching the highway, the participants experienced driving down under a bridge. During the highway driving, the car changed lane from left to right side. After that, the participants experienced driving up and down on a hill, followed by a right turn to a country road, including two left and two right turns. The simulation stopped when the car reached a specific point where the participants could see a city in the distance.

Participants

Nine participants aged between 30 and 53 years ($M = 35.67, SD = 7.04$) took part in the study. Only one female participated, and therefore gender as a variable was unequally distributed and could not be evaluated. All participants were recruited via direct request from the Department of Interior Development by the BMW Group. Five of the participants described themselves as frequent drivers who had driven more than 10.000km in the past year. Regarding the previous experience with driving simulations, 67% responded positively. Only one of the participants reported discomfort after being exposed for over 30 min to a driving simulator. Two participants reported that they play video games daily. In respect to previous experience with HMDs, 67% had used an HMD before this trial, one participant did not respond, and one experienced SiS during their previous VR session. They had spent between 5 and 300 min ($M = 73.3, SD = 111.79$) in VR environments.

Measurements

For this pilot study, objective physiological measurements, such as an ECG and EDA, and subjective measurements, such as questionnaires, were used. To measure the physiological signals, 3-channel-ECG and SCL data were recorded with medical sensors by g.Tec Medical Engineering GmbH[1] with a sampling frequency of 512 Hz. It should be noted that the sensors from g.Tec were used only in this pilot study because they did not provide a wireless connection to the computer. For the subsequent pilot study and the complete experiments, another physiological sensor set was utilized. A baseline signal was recorded for 2 min in a resting position in the VR environment before the driving scenario.

Two subjective questionnaires were issued to participants before each simulation and two after each simulation. Prior to the trial, the participants were given a questionnaire about primary demographic and biographical data. The MSSQ-short was also used before the trial [77]. The post-questionnaires were the SSQ [106] and the IPQ. A few questions regarding enjoyment [136] were included in the IPQ questionnaire to measure the users' emotional reaction to the VR simulation. Additionally, interviews were developed to collect participants' responses. These interviews aimed to collect additional information about the trial.

The study was designed as a pilot study to assess the feasibility of a procedure that could be used for future studies. Due to the small sample size, the data were analyzed with the help of descriptive statistics. HR was obtained from ECG signal and SCL was obtained from EDA signals and processed with the Ledalab software [17].

[1] https://www.gtec.at/

5.1.2 Results and Discussion

The results of the subjective data collected through the SSQ indicate that no clear trend can be observed between the motion and static conditions regarding SiS. Only light to non-existent symptoms were reported in both conditions. The overall score indicates that the *general discomfort* and *stomach awareness* score is higher in the static condition. However, *difficulty focusing* and *fullness of the head* shows a trend towards higher scores in the moving platform condition. Figure 5.1 shows the calculated clusters' scores based on the SSQ's weight system. The symptoms of the Nausea cluster, which showed changes in the responses, are *sweating, increased salivation,* and *stomach awareness*. In the static condition, one participant experienced slight *sweating*, and three participants had slight *stomach awareness*. In the moving platform condition, *sweating* and *stomach awareness* were not reported, and one participant felt *increased salivation*.

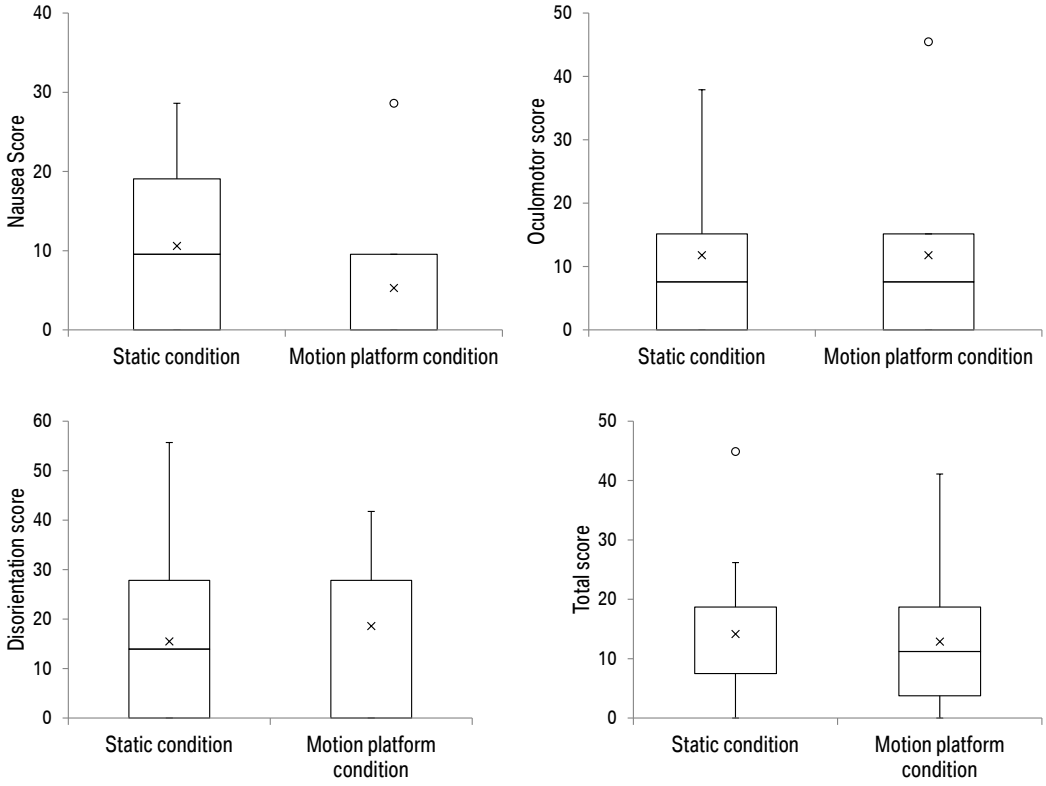


Figure 5.1: SSQ score divided into clusters for static and moving platform conditions.

The results of the MSSQ-short questionnaire showed that two users had a high level, two had a moderate level, and five had a low level of susceptibility to motion sickness. Among the users with a high level, one showed a lower SSQ score in the static condition. One of the users with a moderate level demonstrated a lower SSQ score in the static condition, while the third showed a lower SSQ score in the moving platform condition.

Regarding the sense of presence, the answers to the question "I was involved in the virtual reality experience" from IPQ indicated a tendency towards the moving platform condition. The question "I experienced a delay between my actions and expected outcome" suggested that the static condition provoked a higher experienced delay compared with the moving platform condition. The last four questions regarding enjoyment showed mostly positive answers.

The exploratory data analysis did not show any noticeable difference in physiological signals across the conditions. The mean HR in both conditions was approximately 70 bpm and the mean SCL was approximately 7μS. These results support the self-reported discomfort in both conditions. More detailed information for a particular participant is shown in Figure 5.2. The figure shows that the HR and SCL were higher in the static condition. The last event, in which the car drives up, shows a lower HR level than the previous event, in which the car drives down, in the motion platform condition. This differs from the static condition where the effect was reversed across the two events.

The participants were asked to describe their overall experience with the VR automated driving in both conditions: with and without motion. The experience was described as interesting and exciting, as one participant said: "It was very exciting because it was my first time to try a driving simulation with virtual reality and at the same time it was a bit strange." Another participant responded to the question regarding when their greatest degree of discomfort was felt by saying, "Breaking and acceleration in both conditions made you feel strange. The motion felt mostly like vibration." Other participants said to the same question that the turns were the most unpleasant part of the simulation. One participant stated that "the simulation felt very artificial like nothing is happening there, which got boring at some moments."

Another remarked, "It was much better than using only a screen." One of the participants stated that they felt a light discomfort shortly after the trial, which increased with time. One hour later, the participant got nauseous. According to the experiment's notes, a second participant experienced similar discomfort, but after a short duration of time this returned to the average level. The overall response to the VR automated driving experience (with and without motion) was positive. Participants expressed willingness to participate again.

5.2 Physiological Signals Implementation

In the previous section, we described the moving platform pilot study with physiological sensors. These sensors, however, were not integrated into the VR driving setup, and the data collected through them was synchronized manually to the VR simulation. Thus, in this section, we will describe the integration of physiological signals into the VR driving simulation in order to improve the data synchronization. Furthermore, signals from the virtual environment were recorded for complete SiS evaluation. To help identify the characteristics of the virtual environment that

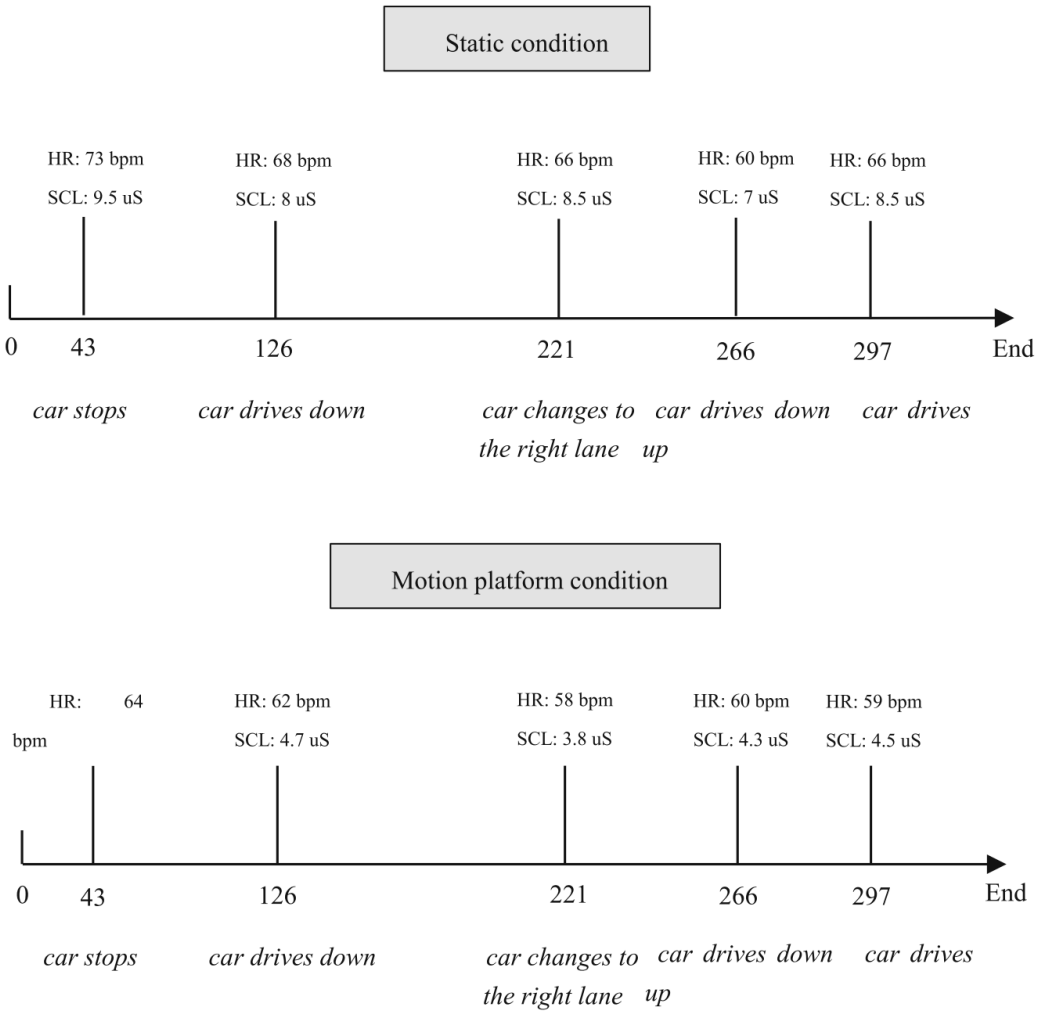


Figure 5.2: HR and SCL over time (in seconds) in static and motion platform condition during the following events: car stops, car drives down, car changes to the right lane, car drives down, and car drives up.

triggered a physiological response, characteristics are expressed by the simulation itself rather than annotating a video recording afterward.

All the data of the pilot study is recorded by SSI [224], which stored the recorded data in its own simple, text-based data format. This way, the data of all sessions can later be further analyzed using different tools such as data analysis and machine learning features of SSI, Python, or Matlab.

The SSI framework offers tools to record, analyze, and recognize human behavior in real-time, such as gestures, mimics, head nods, and emotional speech [224]. It supports streaming from multiple sensors and includes mechanisms for their synchronization. To allow for integration with other applications, SSI features a set of network plugins such as Transmission Control Protocol (TCP), User Datagram Protocol (UDP), User Datagram Protocol, OSC, or Websockets. Beneath

continuous streams, sporadic events can be recorded synchronously. Particularly, SSI supports the machine learning pipeline in its full length (pre-processing, feature extraction, and online classification and fusion) and offers a graphical interface that assists a user to collect own training corpora and to obtain personalized models. It also suits the fusion of multimodal information at different stages, including early and late fusion. SSI is written in C++, and source code is available under LGPL.

Unreal Engine SSI Plugin

In order to connect the VR environment to the SSI pipeline, a plugin[2] for the Unreal Engine was developed, which sends the data via OSC packages to SSI. Therefore, an existing OSC plugin[3] for the Unreal Engine was adapted to SSI's data paradigm. The plugin enables the delivery of different aspects of an environment to SSI or the Unreal Engine's debug output for debugging and development purposes. Configuring a stream includes choosing a sending target, a stream name, a sample rate, and a chunk size.

Physiological Sensors

The physiological signals are recorded through the sensors provided by the Biosensors Plux[4]. Then they are sent to the data collection computer via Bluetooth and recorded by SSI. The signals from the VR environment are sent via the OSC protocol to SSI, where its OSC module records them (Fig. 5.3).

Both the physiological signals and the signals from the VR environment can be categorized into two basic categories: streams and events. Streams are recorded continuously from the beginning to the end of the simulation without interruption with a constant, user-defined sample rate. Events are single points in time and mainly consist of a timestamp and an event name. In the following experiments, only streams were used due to time constraints.

Unique consideration had be taken of the information streams in order to synchronize accurately on the grounds that SSI expected them at a constant example rate. In this way, sampling in the game loop would sample the signal of interest with a constant sample rate. SSI would interpret such an irregularly sampled signal to be sampled with a constant sample frequency. This results in inaccurately calculated timestamps and causes a time discrepancy of the signal, as shown in Figure 5.4. Therefore, SSI's new plugin for the Unreal Engine queries the requested information at a constant rate in a separate thread and saves it into a buffer. If the buffer is full, its content is sent off to SSI. This way, the requested information, such as the car's position, is sampled at a constant rate in the simulation and sent to SSI, enabling correct data synchronization.

[2]https://github.com/stefanbuettner/UE4-SSI

[3]https://github.com/monsieurgustav/UE4-OSC

[4]https://www.biosignalsplux.com

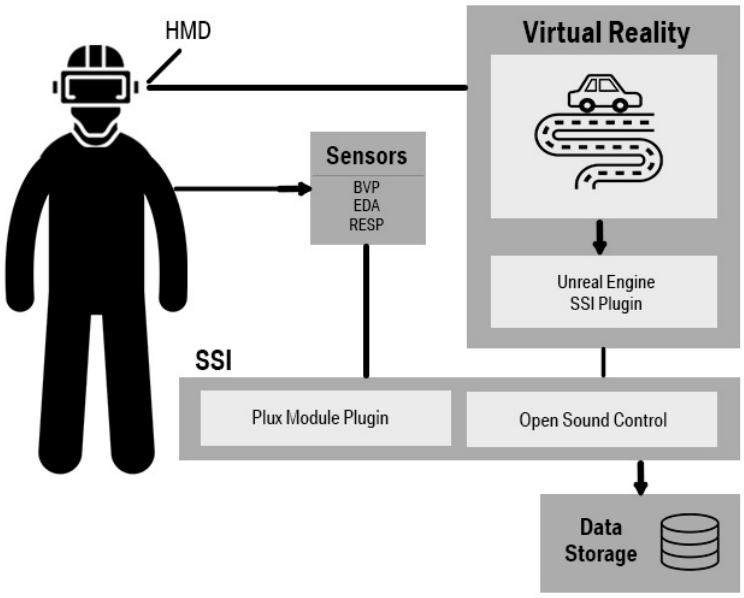


Figure 5.3: Architecture of the proposed approach showing the connection between the sensors and the virtual environment. RESP stands for respiratory sensor.

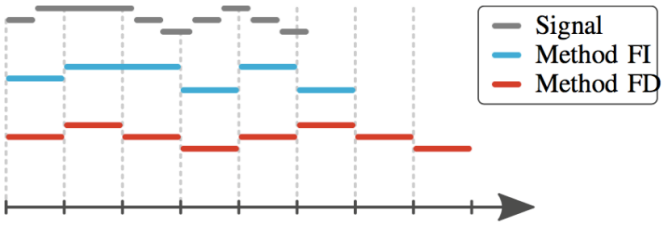


Figure 5.4: This figure shows the reconstructions of two different sampling methods in a game. The frame-dependent method FI is diverging in time whereas the frame-independent method FI is at most $\frac{1}{f}$ seconds off, where f is the sampling frequency. From Büttner [35].

5.2.1 Evaluation

To determine whether the integration of the sensors was successful, a small pilot study was conducted with ten participants. All of the participants were employees of the University of Augsburg, Germany. The study compared two conditions: automated and standard driving on a highway. In addition to the recorded physiological signals, a pre-questionnaire was used to collect standard demographic data. The participants were divided into two groups randomly. Each group of participants took part under only one condition.

The system setup consisted of an HMD with a tracking system, sensors, and a computer for rendering the driving scene. The HMD was the Oculus Rift CV1 with a FOV of 110° and a resolution of 2160 x 1200 pixels per eye. Plux sensors recorded the HR, SCL, and RR signals. All three signals were recorded with a frequency of 1 kHz. The sampling rate of the car location

and head position in the VR was 20 Hz and 10 Hz, respectively.

The VR environment was developed with the engine software Unreal Engine 4.20. A highway, based on highways in Germany, was the developed scenario in the virtual environment since it represented a common driving scenario.

Highway Automated Driving

Four male and one female participants aged between 25 and 37 years ($M = 30.60, SD = 4.72$) tested the system setup. After the introduction, the participants filled in the questionnaire, including questions about demographic data (e.g., age, gender, education), previous driving simulation experience, and VR experience. For example, "I have experience with driving simulations," "I am playing daily video games (e.g., PC- or console games)" and "I have already used a head-mounted display (e.g., Oculus Rift, HTC Vive)." The questionnaire used a 5-Likert scale which scorees from 1 (totally disagree) to 5 (totally agree). Three of the participants reported previous experience with an HMD. Two participants stated "neutral" on their previous experience with a driving simulation, and the others had no prior experience regardless of the driving simulation's type. One of the participants plays video games regularly.

Before starting the VR session, the participants were informed of the procedure and the aim of the study. A consent form was issued to each participant, followed by the pre-questionnaire. It was then explained that the driving simulation was a fully automated driving simulation on a highway road and that the participants were not required to do anything during the driving experience as they did not have control of the vehicle. Then the sensors were attached to the body. The driving simulation started with a pre-spawned car on the highway. After joining the running traffic, the vehicle followed the road. The total duration of the session was approximately 3 minutes.

Highway Standard Driving

Four male and one female participants aged between 24 and 36 years ($M = 29.60, SD = 4.39$) tested the second setup. After the introduction, the participants filled in the same demographic questionnaire from the first setup. Four of the participants did not have previous experience with a driving simulation regardless of the type. Two of the participants had used an HMD before the evaluation. All but one of the participants reported that they play video games relatively often. The procedure in the second condition was the same as the first condition except for the control over the vehicle. Under this condition, the participants used an Xbox 360 controller to operate the vehicle.

5.2.2 Results and Discussion

A software toolbox based on Python 3, named Biosppy[5], was used to analyze the data. For every participant, the HR, SCL, and RR were recorded via the BVP, EDA, and respiratory effort sensors, respectively.

The results from the automated driving condition showed that the mean HR of all participants was 80 bpm (beats per minute), where the min HR was 62 bpm, and the max HR was 154 bpm. Due to technical problems, BVP data for one participant was missing. The average resting HR of an adult is between 60 bpm and 100 bpm [197]. This reveals that the participants had typical HR values during virtual driving and experienced no stress or arousal. The collected data from the EDA signal was converted to skin conductance amplitude. The mean SCL among the participants was 0.28 μS, where the max was 0.40 μS.

The mean RR in the automated driving condition was 0.23Hz, which means that on average the participants managed 14 breaths per minute. The RR for an adult ranges between 12 and 20 breaths per minute or, when calculated in Hz, between 0.20Hz and 0.33Hz. Thus, at an average of 14 breaths per minute, the participants were within the normal range. Two signals, car location, and head position were acquired as streams from the virtual environment. Each of these signals was recorded within the x, y, z coordination system. The results from the standard driving condition were slightly different from the automated driving condition, as it was expected. The mean HR of all participants was 88 bpm, where the min HR was 59 bpm, and the max HR was 149 bpm. Due to technical problems, BVP data for one participant was missing. The mean HR was 8 bpm higher than the automated driving condition, and the min and max HR were slightly lower. This could be explained by the fact that the participants were more excited to drive themselves than sit and ride in the vehicle.

Regarding the EDA, the mean SCL among the participants was 0.38 μS, where the max was 0.48 μS. Due to technical problems, EDA data for one participant was missing. It is interesting to report that in comparison with the mean SCL from the automated driving condition, the mean SCL in the standard driving condition was 0.10 μS higher. The higher mean SCL was in line with the higher mean HR. Previously, Cobb et al. [42] reported an increased HR with prolonged stay in the virtual environment. However, in the pilot study, the exposure did not exceed 5 min, which means the effects of a prolonged VR session were not considered. Kim et al. [120] found that SCL increased during a virtual-navigation task. Driving in a virtual environment can be considered a navigation task in a VR simulation as an individual moves herself or himself through the virtual world using a steering wheel. Thus, our results align with the previous research.

The last recorded physiological signal was respiration. The mean RR was 0.22Hz; this means that on average, the participants took 13 breaths per minute. These results differ with one breath from the automated driving condition, where the average RR was 14 breaths per minute. Compare to

[5]https://github.com/PIA-Group/BioSPPy

the mean HR and mean SCL, which demonstrated increased values, the mean RR was slightly decreased.

5.3 Summary

In this chapter, we have described the implementation of two different appliances: a moving platform and physiological sensors. We have integrated these two appliances into a standard setup in order to create an innovative VR driving simulation setup for subsequent user experiments. The implementation results showed that the moving platform addition is a feasible VR setup, and a reduction in SiS symptoms was seen. However, the pilot study was used only to test the concept of a moving platform and an HMD in terms of driving simulation. The implementation of physiological sensors was the logical next step as the sensors used in the pilot study were not automatically synchronized with the driving simulation. The small evaluation showed that the sensors and the data acquisition were working as expected, and they could be used for user experiments. The experiments reported in the upcoming chapters were all conducted using the same driving simulation setup. All acquired data was necessary to build a complete profile of SiS induced by VR simulation in two different driving conditions and two different moving conditions.

6 Experiment 1: Static vs. Dynamic Standard Driving

> An approximate answer to the right problem is worth a good deal more than an exact answer to an approximate problem.
>
> *John Tukey*

This chapter describes the evaluation of *motion* and *gender* influence on SiS in a standard VR driving simulation. The standard driving simulation utilized a virtual vehicle with an automatic gearbox. The experiment had the same setup described in the methodology **Chapter 5**, including an HMD, a moving platform, and physiological sensors. The SiS onset was assessed during the VR simulation, immediately after and around one hour later. The driving scenario was an urban driving scenario with a length of 24 min.

We hypothesized that there would be a significant difference in the discomfort onset between the motion and without motion condition. According to previous research and the cue conflict theory, the addition of aligned motion cues to a stationary driving simulation might alleviate the SiS [10, 48]. Thus, the group using a moving platform should report visibly lower SiS scores. A significantly positively correlation between SiS and physiological signals such as heart signal [51, 120, 194], sweating [51, 226], breathing [120] was previously described. Therefore, it is expected that the difference across conditions would be visible in the physiological signals. Previous studies reported that the SiS and *sense of presence* are negatively related, and therefore, the increased *sense of presence* is correlated to a lower SiS onset [229]. Therefore, it is assumed that the *sense of presence* will increase with the addition of motion feedback of the VR driving.

Regarding *gender*, we assumed that female participants would experience higher discomfort than male participants in the VR driving simulation. Earlier research showed that women felt sicker than men [72, 98, 127, 143, 212]. The reason for that is unclear but possible explanations could be

the hormone balance [40], wider FOV in women [34], or difference in the depth perception [23]. Furthermore, we assumed that the differences across genders would be visible in the physiological signals. As *sense of presence* is related to the VR user experience and SiS, it is hypothesized that female and male participants experience the VR simulation differently. Moreover, the interaction effect between *gender* and moving platform has been not well reported in the literature. Thus, in this thesis, it is hypothesized that such an effect exists, and it is significant. This interaction might affect the SiS, physiological signals, and *sense of presence*.

Before reporting the results from the experiments, we have briefly described the experimental design. Furthermore, descriptive statistics regarding the participants displayed the *Standard Deviation (SD)* and the *Mean (M)* of variables related to socio-demographics and previous experiences as well as motion sickness history. Next, the chapter continues with reporting the results. In this chapter, two aspects of the data analysis were in focus: the hypotheses testing and the correlation analysis. These two types of analysis revealed insights regarding SiS in VR driving simulations. The chapter concludes with a summary of the reported results.

6.1 Study Design

A 2 x 2 factorial design was chosen to investigate whether the addition of *motion* and *gender* affect the onset of SiS or not in a standard VR driving simulation. Each factor had two levels: *motion* (with motion, without motion) and *gender* (male, female). A comparison between the control and the experimental groups was conducted as well as correlation analysis. The dependent variables were SiS, *sense of presence*, HR, SCL, and RR. The control variables were *age*, *vision correction*, *motion sickness history*, *experience with VR*, *experience with driving simulation*, *gaming*, and *driving habits*. The *driving habits* represented the questions regarding the *driving frequency* and the *driving km per year*. *Driving frequency* measured how often the participants drove a car in the last 12 months on a scale from 1 (Never) to 5 (Daily). The variable *driving km per year* measured what was the participants' annual mileage in km on a scale from 1 (No driving) to 7 (More than 40 000km per year). *Gaming experience* measured the daily gaming frequency on a 5-Likert scale from 1 (Totally disagree) to 5 (Totally agree). On the same scale, the participants' previous experience with VR was measured. The previous *driving simulation experience* was also measured on the 5-Likert scale. The *MSSQ-short* score is calculated according to the questionnaire calculations guidance, where the minimum value is 0, and the maximum value is 54. Correlation analysis on human factors (i.e., *gender*, *motion sickness history*, *previous experience*, *vision correction*, *emotions*, and *physical activity*) was conducted to calculate the appearance and the strength of the relationship.

The participants were randomly assigned into two groups, with a moving platform and without moving platform. The exclusion criteria for participation were a dysfunctional vestibular or visual system, heavy medication taken, over caffeine intake, alcohol intake in less than 24h, sickness

(flu or cold), pregnancy, any problems with the vestibular system. The data was collected through questionnaires (pre-questionnaire, on-spot-questionnaire, follow-up questionnaire), physiological sensors (ECG, EDA, and respiratory effort), behavioral data (vehicle location in the virtual world, head location and rotation). For the data analysis, the IBM SPSS Statistics version 23.0 and the Python-based library Biosppy [1] were used. The average vehicle's speed was 27 km/h and the maximum speed was 96 km/h.

6.2 Participants

The study sample consisted of 63 participants (Fig. 6.1), between 18 and 61 years old ($M = 34.51, SD = 11.78$), from which 40 were male. All participants were recruited via internal e-mail system. The eligibility criteria for the participants were: being able to understand the German language, having a driver's license class B and having normal or corrected to normal vision. The participants were randomly assigned to one of the two conditions with a total number of 32 (female - 11, male - 21) for without motion and 31 (female - 12, male - 19) for motion group. Thirty participants wore some type of vision correction (e.g., glasses, lenses) and 33 did not. The considered study sample regarding the physiological measurements was smaller than 63 participants due to technical issues with the used sensors.

The motion condition group did not differ significantly regarding the following susceptibility factors: *age*, *driving kilometer per year*, *gaming*, *VR experience*, *simulator experience*, and *motion sickness history* (Table 6.1). The two conditions did differ significantly regarding the driving frequency. The motion condition participants drove significantly more frequently than the without motion condition participants.

The female and male groups differed only significantly regarding *age*, *video game usage*, and *motion sickness history*. Male participants were older, scored higher on *video game usage*, and lower on *motion sickness history* than females (Table 6.2.)

6.3 Results

The results showed a significant difference regarding SiS onset between the two *motion* conditions. Surprisingly, the results showed no significant differences regarding *gender*. All graphs related to the hypotheses testing are displayed in Appendix G.

A frequency regarding the answers of the self-report scales concerning SiS was calculated in order to get an overview of the magnitude of SiS occurrence in the study. The overall occurrence of discomfort, verbally reported by the participants via FMS during the VR driving, was divided

[1] https://github.com/PIA-Group/BioSPPy

Table 6.1: Descriptive statistics of the susceptibility factors differences: age, driving frequency, driving km per year, gaming experience, VR experience, driving simulator experience, and motion sickness history, between the two conditions - motion and without motion. One case from the motion condition is missing.

	Without motion			Motion		
	M	SD	n	M	SD	n
Age	35.59	12.16	32	32.70	10.98	31
Driving frequency	3.69	0.86	32	3.83	0.82	31
Driving km per year	3.53	1.37	32	3.60	1.1	30
Gaming experience	1.88	1.34	32	1.47	.90	30
VR experience	2.69	1.47	32	2.50	1.23	30
Driving simulation experience	2.34	1.43	32	2.27	1.39	30
MSSQ-short	9.25	9.27	31	7.57	6.52	30

Table 6.2: Descriptive statistics of the susceptibility factors differences: age, driving frequency, driving km per year, gaming experience, VR experience, driving simulator experience, and motion sickness history, between genders.

	Female			Male		
	M	SD	n	M	SD	n
Age	28.26	6.51	23	37.69	12.56	39
Driving frequency	3.78	.85	23	3.74	.85	39
Driving km per year	3.39	.89	23	3.67	1.4	39
Gaming experience	1.74	1.36	23	1.64	1.04	39
VR experience	2.61	1.31	23	2.59	1.39	39
Driving simulation experience	2.39	1.44	23	2.26	1.39	39
MSSQ-short	10.87	8.89	23	7.00	7.23	39

Figure 6.1: A participants using the standard VR driving setup.

into five clusters (Fig. 6.2). As many as 30% of the participants experienced moderate discomfort. None of the participants experienced severe or frank sickness. Similar results were reported from the SSQ, where the severity level was low as slight (46%) sickness severity. As expected, the severity level dropped after one hour and more (Fig. 6.4). The results from the FSSQ showed that as much as 89% of the participants reported slight sickness severity. From them 23% reported no symptoms at all. Compare to the results of the Experiment 1 from the same questionnaire, here the results showed that the participants had more slight symptoms. Only one participant reported severe sickness in the FSSQ. The same participants reported very slight sickness symptoms immediately after the VR simulation. Furthermore, the participant discontinues the driving simulation between the 9th and the 12th min. The participant was a female and took part in the without motion condition.

The most self-reported symptoms during the VR motion condition were *nausea*, *unease*, and *feeling warm* (Fig. 6.3). For the other condition, the most reported symptoms were *dizziness*, *nausea*, and *unease*. It should be noted that the participants were allowed to report more than one symptom at a time.

Moreover, 19 participants (11 male) of the considered sample size had to terminate the driving

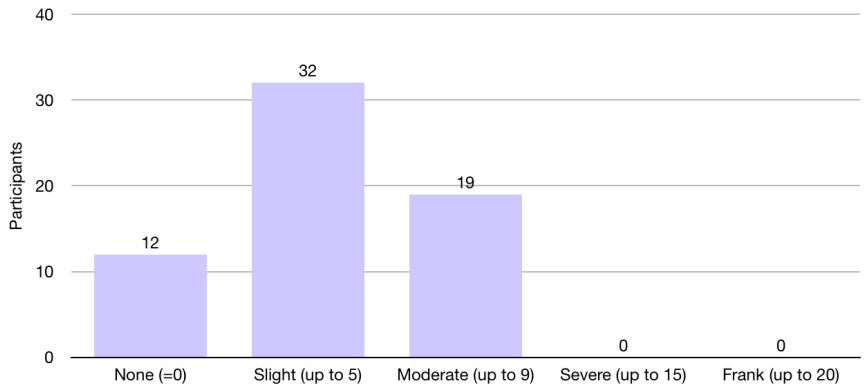


Figure 6.2: Mean FMS score frequency separated into five clusters: none, slight, moderate, severe, and frank sickness.

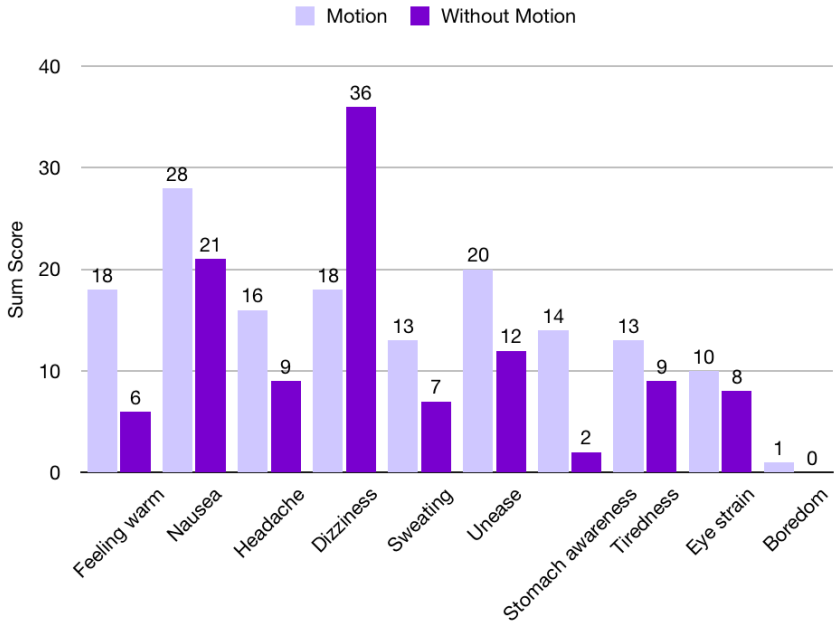


Figure 6.3: Symptoms frequency during the VR simulation for motion and without motion conditions.

simulation early due to severe feelings of discomfort. These participants were not excluded from the data analysis because the mean value of the FMS score was taken for the statistical analysis. For the analysis of the rest of the data, the earlier withdraw was not relevant as the aim of the experiment was to evaluate the sickness onset. A considerable number of those participants (n = 9), quit before the 12th min of the driving simulation. Furthermore, most of them (32%) stopped the test between the 6th and the 9th min, followed by 21% between the 12th and 15th min. Overall, most participants (78%) stopped before the 15 min mark. However, only 26% reported a moderate severity level on the SSQ total score. The same amount of participants reported severity level equal to none. The sickness symptoms followed the scheme Nausea > Disorientation >

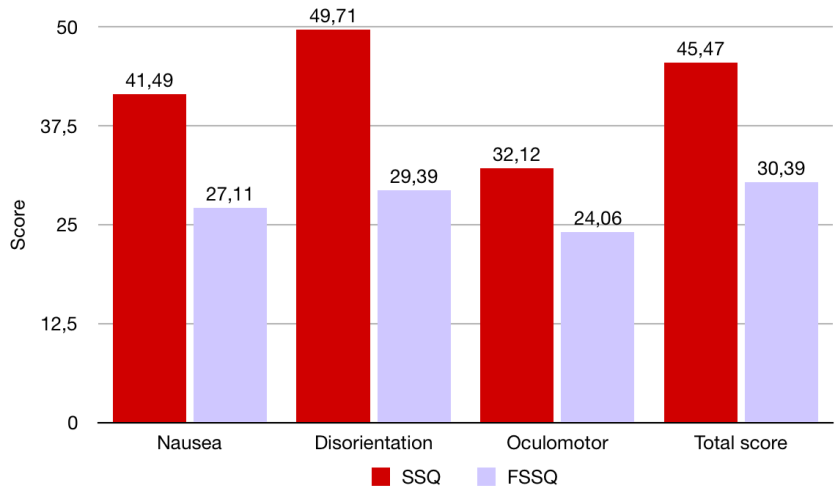


Figure 6.4: Symptoms frequency (mean score) immediately after and one hour later after the VR simulation.

Oculomotor.

Most of the participants (68%), who stopped the virtual driving prematurely, were part of the group in the without motion condition. On the question, if the participants feel tired today, 32% answered positively, and 44% answered negatively. Furthermore, all of the participants defined themselves as *"happy"*, and 68% defined themselves as *"calm"* before the start of the VR driving simulation.

6.3.1 Hypotheses Testing

FMS and SSQ Score

A two-way ANOVA was conducted on the influence of two independent variables (*gender*, *motion*) on the SiS onset during, immediately after, and one hour later, after the driving simulation. *Gender* included two levels (female, male) and *motion* included two levels (with motion, without motion). The dependent variables were FMS score, SSQ Nausea score, SSQ Disorientation score, SSQ Oculomotor score, SSQ total score, FSSQ Nausea score, FSSQ Disorientation score, FSSQ Oculomotor score, and FSSQ total score. The effect size, Eta squared (η^2), for each variable that showed significance is reported. As guidance to which degree the variable is presented in the population, we used Cohen's guidance, where the small effect limit is 0.10, the average effect limit is 0.25, and the large effect limit is 0.40 [43]. Table 6.3 presents an overview of the Means and Standard Deviations for the ANOVA results of the independent variables.

Table 6.3: Overview of Means and Standard Deviations for the two-way ANOVA results using the FMS score, SSQ total score, SSQ Nausea score, SSQ Disorientation score, FSSQ Nausea score, and FSSQ Oculomotor score as criterion.

Measure	Motion		Without Motion		Male		Female	
	M	SD	M	SD	M	SD	M	SD
FMS score	2.94	2.92	4.51	2.98	3.40	2.92	4.33	3.20
SSQ total score	42.59	32.47	48.27	37.74	48.34	35.60	40.49	34.36
SSQ Nausea score	31.08	33.14	51.58	40.54	36.97	38.54	49.36	37.13
SSQ Disorientation score	39.51	46.88	59.60	48.64	43.85	44.97	59.92	53.49
FSSQ Nausea score	21.12	28.01	37.21	36.08	22.69	26.06	41.34	40.96
FSSQ Oculomotor score	20.57	18.78	31.33	31.39	20.69	20.11	35.73	33.31

For the FMS score (Table 6.4), only the effect of *motion* was statistically significant at the $.05$ significance level with small effect size of 0.09.

Table 6.4: Two-way ANOVA results using FMS score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	49.22	1	49.22	5.743	.020*	0.09
Gender	15.10	1	15.10	1.762	.190	
Motion * Gender	10.33	1	10.33	1.206	.277	
Residuals	505.65	59	8.57			

*Note. *p < .05, **p < .01, ***p < .001.*

For the SSQ total score (Table 6.5), all effects were not statistically significant at the $.05$ significance level.

For the SSQ Nausea score (Table 6.6), only the effect of *motion* was statistically significant at the $.05$ significance level with small effect size of 0.01.

For the SSQ Disorientation score (Table 6.7), all effects were not statistically significant at the $.05$ significance level.

Due to a normality violation, the ANOVA was not a suitable analysis for the SSQ Oculomotor score. A Mann-Whitney test indicated that the SSQ Oculomotor score was greater for participants in the without motion condition ($Mdn = 38.16$) than for participants in the motion condition ($Mdn = 25.65$), $U = 299.0, p = .006, r = -.35$. A Mann-Whitney test indicated no significant difference between female ($Mdn = 22.74$) and male ($Mdn = 26.53$) participants regarding

Table 6.5: Two-way ANOVA results using SSQ total score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	1085.66	1	1085.66	0.880	.352
Gender	816.67	1	816.74	0.662	.419
Motion * Gender	2156.60	1	2156.60	1.748	.191
Residuals	72787.5	59	1233.69		

*Note. *p < .05, **p < .01, ***p < .001.*

Table 6.6: Two-way ANOVA results using SSQ Nausea score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	7462.59	1	7462.59	5.450	.023*	0.01
Gender	2628.53	1	2628.53	1.920	.171	
Motion * Gender	485.73	1	485.73	0.355	.554	
Residuals	80787.8	59	1369.28			

*Note. *p < .05, **p < .01, ***p < .001.*

Table 6.7: Two-way ANOVA results using SSQ Disorientation score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	6337.49	1	6337.49	2.769	.101
Gender	4232.63	1	4232.63	1.850	.179
Motion * Gender	0.25	1	0.25	0.000	.992
Residuals	135020	59	2288.47		

*Note. *p < .05, **p < .01, ***p < .001.*

SSQ Oculomotor score, $U = 382.0, p = .261, r = -.14$.

FSSQ Score

Due to a normality violation, the ANOVA was not a suitable analysis for the FSSQ total score. A Mann-Whitney test indicated that the FSSQ total score was greater for participants in the without motion condition ($Mdn = 33.66$) than for participants in the motion condition ($Mdn = 14.96$), $U = 292.0, p = .046, r = -.27$. The same test indicated a significant difference between female ($Mdn = 37.40$) and male ($Mdn = 14.96$) participants regarding FSSQ total score, $U = 258.0, p = .034, r = -.28$.

For the FSSQ Nausea score (Table 6.8), the two main effects were statistically significant at the $.05$ significance level with small effect size.

Table 6.8: Two-way ANOVA results using FSSQ Nausea score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	4199.66	1	4199.66	4.218	.045*	0.07
Gender	4441.02	1	4441.02	4.460	.039*	0.07
Motion * Gender	576.02	1	576.02	0.579	.450	
Residuals	53768.36	54	995.71			

*Note. *p < .05, **p < .01, ***p < .001.*

Due to normality violation, the ANOVA was not suitable analysis for the FSSQ Disorientation score. A Mann-Whitney test indicated that the FSSQ Disorientation score was greater for participants in the without motion condition ($Mdn = 27.84$) than for participants in the motion condition ($Mdn = 13.92$), $U = 293.0, p = .041, r = -.27$ (Fig. G.4). The same test indicated a significant difference between female ($Mdn = 27.84$) and male ($Mdn = 13.92$) participants regarding FSSQ Disorientation score, $U = 258.5, p = .030, r = -.29$.

For the FSSQ Oculomotor score (Table 6.9), the effect of *gender* was statistically significant at the $.05$ significance level and a effect size of 0.07.

Presence Scores

A two-way ANOVA was conducted on the influence of two independent variables (*gender*, *motion*) on the *sense of presence*. *Gender* included two levels (female, male) and *motion* included two levels (with motion, without motion). For the *presence score* (Table 6.10), all the effects were not statistically significant at the $.05$ significance level.

Table 6.9: Two-way ANOVA results using FSSQ Oculomotor score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	1965.01	1	1965.01	3.056	.086	
Gender	2895.69	1	2895.69	4.504	.038*	0.07
Motion * Gender	397.27	1	397.27	0.618	.435	
Residuals	34719.40	54	642.95			

*Note. *p < .05, **p < .01, ***p < .001.*

Table 6.10: Two-way ANOVA results using presence score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	2.45	1	2.45	0.102	.750
Gender	0.12	1	0.12	0.005	.944
Motion * Gender	7.68	1	7.68	0.321	.573
Residuals	1413.65	59	23.96		

*Note. *p < .05, **p < .01, ***p < .001.*

Physiological Signals

This study pursues the conception that an increase of the measured physiological signals (ECG, EDA, and respiratory effort) indicates severer SiS occurrence. A five-minute baseline measurement of the signals was recorded in order to have reference values of the participants' average physiological signals. During the driving simulation, the physiological signals were continuously assessed. After the data acquisition, the signals were averaged for every participant in order to be able to analyze them further. The signals of the baseline and driving simulation were separately averaged over time for every participant. The difference between these two scores was calculated (Driving Simulation minus Baseline) for every physiological signal, HRD, SCLD, RRD, respectively. The higher values indicating a severer SiS level. It should be taken into account that only the physiological data which had both records, during the simulation and the baseline, was included in the calculation of the variable differences.

A two-way ANOVA was conducted on the influence of two independent variables (*gender*, *motion*) on the physiological signals response during the driving simulation. *Gender* included two levels (female, male) and *motion* included two levels (with motion, without motion). Table 6.11 presents an overview of the Means and Standard Deviations for the ANOVA results of the independent variables.

For the HRD (Table 6.12), all effects were not statistically significant at the .05 significance level.

Table 6.11: Overview of Means and Standard Deviations for the two-way ANOVA results using the presence score and the physiological signals as criterion.

Measure	Motion		Without Motion		Male		Female	
	M	SD	M	SD	M	SD	M	SD
Presence	44.00	4.61	43.78	5.03	43.85	4.39	43.96	5.52
HRD	2.65	9.70	1.41	15.10	2.86	12.63	0.34	12.35
HR	77.86	9.31	78.81	10.99	76.73	9.42	81.26	11.00
SCLD	0.92	0.68	0.63	0.40	0.75	0.64	0.87	0.52
SCL	1.11	0.63	0.88	0.40	0.99	0.55	1.02	0.54
RRD	0.01	0.04	-0.02	0.03	-0.004	0.04	-0.14	0.02
RR	0.24	0.02	0.26	0.12	0.26	0.11	0.23	0.23

Table 6.12: Two-way ANOVA results using HRD as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	51.10	1	51.10	0.314	.578
Gender	93.77	1	93.77	0.576	.452
Motion * Gender	25.47	1	25.47	0.156	.694
Residuals	7492.71	46	162.89		

*Note. *p < .05, **p < .01, ***p < .001.*

For the HR during the VR driving simulation (Table 6.13), all effects were not statistically significant at the .05 significance level.

Table 6.13: Two-way ANOVA results using HR as criterion.

	Sum of Squares	*df*	*Mean Square*	*F*	*p*
Motion	0.95	1	0.95	0.010	.923
Gender	205.09	1	205.09	2.067	.157
Motion * Gender	163.21	1	163.21	1.645	.206
Residuals	4862.33	49	99.23		

*Note. *p < .05, **p < .01, ***p < .001.*

For the SCLD (Table 6.14), all effects were not statistically significant at the .05 significance level.

Table 6.14: Two-way ANOVA results using SCLD as criterion.

	Sum of Squares	*df*	*Mean Square*	*F*	*p*
Motion	0.55	1	0.55	1.514	.233
Gender	0.05	1	0.05	0.125	.727
Motion * Gender	0.08	1	0.08	0.208	.653
Residuals	7.20	20	0.36		

*Note. *p < .05, **p < .01, ***p < .001.*

For the SCL during the VR driving simulation (Table 6.15), all effects were not statistically significant at the .05 significance level.

Table 6.15: Two-way ANOVA results using SCL as criterion.

	Sum of Squares	*df*	*Mean Square*	*F*	*p*
Motion	0.39	1	0.39	1.272	.270
Gender	0.02	1	0.02	0.054	.819
Motion * Gender	0.01	1	0.01	0.035	.853
Residuals	8.05	26	0.31		

*Note. *p < .05, **p < .01, ***p < .001.*

For the RRD (Table 6.16), all effects were not statistically significant at the .05 significance level.

Table 6.16: Two-way ANOVA results using RRD as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	0.00	1	0.00	1.609	.224
Gender	0.00	1	0.00	0.419	.527
Motion * Gender	0.00	1	0.00	2.333	.148
Residuals	0.02	15	0.00		

*Note. *p < .05, **p < .01, ***p < .001.*

For the RR during the VR driving simulation (Table 6.17), all effects were not statistically significant at the $.05$ significance level.

Table 6.17: Two-way ANOVA results using RR as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	0.00	1	0.00	0.208	.651
Gender	0.01	1	0.01	0.649	.427
Motion * Gender	0.00	1	0.00	0.265	.611
Residuals	0.24	29	0.01		

*Note. *p < .05, **p < .01, ***p < .001.*

6.3.2 Correlations

All data using different scales (e.g., physiological data, emotions, personality, MSSQ-short) were standardized for the statistical analysis. A Spearman's correlation showed no significant relationship between FMS, SSQ, and FSSQ scores, and *age* (see appendix F). Another correlation analysis showed a significant relationship between MSSQ-short and sickness scores as follows. For SSQ total score, $r_s = .221, p = .085$; SSQ Nausea, $r_s = .255, p = .046$; SSQ Disorientation, $r_s = .259, p = .042$; SSQ Oculomotor, $r_s = .357, p = .004$; FMS mean score, $r_s = .256, p = .045$ (see Appendix F).

Another Spearman's correlation showed a significant relationship between SSQ Disorientation, SSQ Oculomotor, and *neuroticism* (Table 6.18). A Spearman's correlation showed a significant relationship between SSQ Oculomotor and *extroversion* (Table 6.18).

No relationship between FMS, SSQ, FSSQ scores, and *presence* was found through Spearman's correlation (see Appendix F).

Table 6.18: Correlations between FMS mean, SSQ, and FSSQ scores, neuroticism and extroversion.

	Neuroticism	n	Extroversion	n
FMS mean score	.103	62	-.139	62
SSQ total score	.055	62	-.170	62
SSQ Nausea	.210	62	-.231	62
SSQ Disorientation	.301*	62	-.168	62
SSQ Oculomotor	.394**	62	-.283*	62
FSSQ total score	.110	58	-.004	58
FSSQ Nausea	.050	58	-.040	58
FSSQ Disorientation	.107	58	.125	58
FSSQ Oculomotor	.161	58	-.011	58

*Note. *p < .05, **p < .01, ***p < .001.*

A point-biserial correlation was run to determine the relationship between FMS, SSQ, FSSQ scores, and *gender* as well as *vision correction*. There was a negative correlation between FSSQ scores and *gender*, which was statistically significant as follows (Fig. 6.5). FSSQ total score, $r = -.288, p = .028$; FSSQ Nausea, $r = -.273, p = .038$; FSSQ Disorientation, $r = -.268, p = .042$; FSSQ Oculomotor, $r = -.276, p = .036$. No statistically significant relationships were found for *vision correction* (see Appendix F).

Furthermore, no significant relationship between FMS, SSQ, FSSQ scores, and *sleep deprivation* as well as *physical activity* was found (see Appendix F). A Spearman's correlation showed a significant relationship between FMS, SSQ Disorientation, SSQ Oculomotor, and *arousal* before the driving simulation (Table 6.19). The same test showed no significant relationship between FMS, SSQ, FSSQ scores, and *valence* before the driving simulation (Table 6.19).

Table 6.19: Correlations between FMS mean, SSQ, FSSQ total score, arousal, and valence.

	Arousal	n	Valence	n
FMS mean score	.389**	63	-.136	63
SSQ total score	.139	63	-.151	63
SSQ Nausea	.212	63	-.128	63
SSQ Disorientation	.252*	63	-.045	63
SSQ Oculomotor	.296*	63	-.009	63
FSSQ total score	0.66	58	-.194	58
FSSQ Nausea	-.044	58	-.145	58
FSSQ Disorientation	.041	58	-.124	58
FSSQ Oculomotor	.128	58	-.182	58

*Note. *p < .05, **p < .01, ***p < .001.*

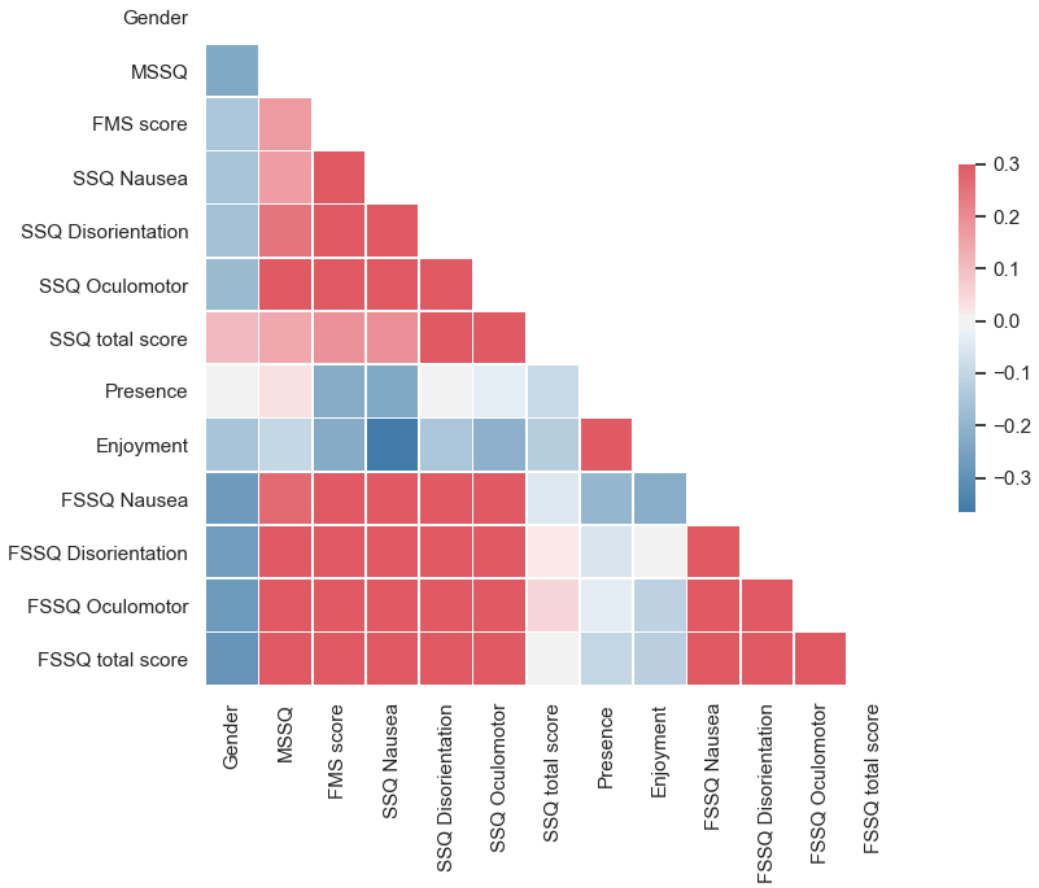


Figure 6.5: A Pearson's correlation matrix showing the relationship between MSSQ-short, FMS, SSQ, presence, FSSQ scores, and gender. For the complete results including the p-values, please see Appendix F.

A Pearson's correlation was run to investigate the relationship between FMS, SSQ, FSSQ scores, and *driving frequency*, *video games play*, *driving simulation experience*, *VR experience*. Only the significant results are reported. A significant positive relationship between the SSQ Nausea score and self-reported *driving frequency* as "except driving once per month" was found, $r = .223, p(1 - tailed) = .040$. Other significant positive relationships between SSQ Disorientation, FMS score and not *playing video games daily* was reported, $r = .239, p(1 - tailed) = .031$, and $r = .212, p(1 - tailed) = .049$, respectively. Furthermore, a positive correlation between FMS score and *driving simulation experience* as "neutral" answered was showed, $r = .240, p(1-tailed) = .030$. A significant negative relationship between SSQ Nausea and previous *VR experience* as positive answered was reported, $r = -.301, p(1-tailed) = .009$.

Regarding HRD, a Spearman's correlation showed significant relationship between SSQ total score and the calculated difference of HR, $r_s = .281, p = .048$. The same statistical test showed no significant relationship between FMS, SSQ Nausea, SSQ Disorientation, SSQ Oculomotor, FSSQ scores, and HRD (see Appendix F). Furthermore, a Spearman's correlation showed no

significant relationship between FMS, SSQ, FSSQ scores, and HR recorded during the driving simulation F).

Furthermore, no significant relationship between FMS, SSQ, FSSQ scores, and SCL data was found (see Appendix F). A Spearman's correlation showed a significant relationship between SSQ Nausea, and RR recorded during the driving simulation as well as FSSQ Nausea and RR record. For SSQ Nausea, $r_s = -.513, p = .002$, and for FSSQ Nausea, $r_s = -.441, p = .013$. The same statistical test showed no significant relationship FMS, SSQ Disorientation, SSQ Oculomotor, SSQ total score, FSSQ Disorientation, FSSQ Oculomotor, and FSSQ total score (see Appendix F). No relationship between FMS, SSQ, FSSQ scores, and RRD was found (see Appendix F).

6.4 Discussion

A significant difference across the *motion* conditions supported the assumed hypotheses regarding the influence of *motion* on SiS. Participants felt significantly more SiS during the static virtual driving than the dynamic driving. Furthermore, participants in the static condition experienced more symptoms from Nausea (e.g., *stomach awareness, nausea, sweating*) and Oculomotor (e.g., *headache, eye strain, fatigue*) clusters. These findings aligned with earlier research on SiS induced by driving simulations [10, 48]. Furthermore, the questionnaire given one hour later (FSSQ) showed significant results across the two conditions. Participants in the static condition felt longer sick than the participants in the dynamic condition. A possible explanation is the higher level of SiS, which participants in the static condition experienced during the VR driving. Thus, this group of participants needed a longer time to regain a normal state of well-being.

Surprisingly, *gender* showed no effect on SiS onset. The groups had no significantly different results regarding sickness based on the participants' *gender*. However, a trend was observed that women experienced more discomfort in the absence of motion cues than men. In two of the symptoms clusters, Oculomotor, and Total score, male participants felt slightly more unwell than the female participants. Even though the self-report data indicate no severe SiS occurrence, the early dropout rate of 30.0% emphasizes the need to find solutions in order to mitigate SiS in VR driving simulations. Nevertheless, the results from the experiment aligned with Klosterhalfen et al. [121] that the female participant did not experience higher discomfort than male participants. A possible reason for that could be that female participants did not differ significantly than the male participants regarding previous *VR experience, gaming experience*, and *driving simulation experience*, as well as *driving frequency*. The two individual factors that differed were *age* and *motion sickness history*. However, these factors did not contribute to the felt discomfort in the female participants. Furthermore, women recovered slower from the felt discomfort than men, and they experienced longer the side effects of the standard VR driving. These results contradict the findings from the FMS and SSQ. There was no significant difference between the genders, but one hour later or more, the women reported higher SiS severity than men. That points out

to unstable nature of the malaise. It could outbreak even when an hour has passed from the VR simulation.

A statistically significant negative relationship was reported between *gender* and FSSQ scores. Men tend to feel more discomfort one hour after the VR driving simulation then women. This correlation contradicts the reported above hypothesis that female participants reported higher FSSQ scores than male participants. Furthermore, the side effects after exposure to a virtual environment could be substantial [207]. The authors stated that the felt discomfort does not quickly disappear upon post VR exposure but lingers on for a significant period. That points out that SiS could be a severe problem even when the exposure to the VR simulation was a while ago.

It was observed that the female participants experienced more discomfort in the absence of motion cues. The responses to the questionnaire one hour after the VR simulation showed the same results. Another trend was observed in all sickness measures in the experiment. The female participants felt more discomfort during the standard driving simulation. In the condition of standard driving, the participant's head should move fast according to the situation on the road, like in the real world. Therefore, some visual blur from these movements could occur. Due to the wider FOV [34], the women could be more affected from this effect than men.

Sense of presence were not affected by the addition of a moving platform or *gender*. The results showed that the participants felt almost the same *sense of presence* in the static and dynamic conditions. It was surprising to found out that the addition of motion cues to the VR driving simulation did not contribute to the better feeling of *presence* in the dynamic driving condition. Despite the previous research on *presence* and SiS, in which the report showed controversial results [229], it is assumed that stimulating two of the primary sensory systems, vestibular and visual systems, the *sense of presence*, would be higher.

Regarding the physiological data, no significant difference was found across the *motion* conditions or *gender*. The difference between the recorded signals of HR, RR, and SCL, and the baseline values, as well as the signals recorded during the VR simulation, were almost the same for the female and male participants and the static and dynamic conditions. These findings are surprising as it was expected that physiological responses would follow the results from the subjective measurements. Partially the results regarding *gender* during and immediately after the simulation were supported. However, no significant difference was observed across the conditions that oppose the FMS and SSQ measurements. It could be argued that the participants did not assess their state of well-being accurately, and they reported higher values on the scales. That points out that both types of measurements should be used to assess all aspects of SiS accordingly.

Furthermore, a significant positive correlation was reported between SSQ total score and the difference between the baseline and the recorded HR. When the overall discomfort increases, the HR differentiates significantly from the HR recorded in a normal state of well-being, and vice versa. These results aligned with previous research where participants with severer SiS symptoms showed higher levels of HR than participants who did not experience SiS [42]. However, it has to

be noted that some research results show that these physiological changes can individually differ regarding their direction [98].

A strong negative relationship was found between the SSQ Nausea score and RR recorded during the driving simulation. When the RR increases, the severity of the Nausea cluster symptoms such as *nausea, sweating*, and *stomach awareness* decreases, and vice versa. A similar strong negative relationship was reported between FSSQ Nausea and RR. This negative relationship between these SiS symptoms and RR records pointed to that the participants slow down breathing when they experienced more discomfort. Similar results were reported in previous research [120], where an RR decrease during the VR simulation was reported. A possible explanation could be that this action is taken as an automatic response from the human body to reduce the aroused nervous system. Nevertheless, the results should be interpreted with caution because the RR sample size significantly smaller than the HR sample size.

An average positive correlation was reported between MSSQ-short total score and sickness scores. Participants who indicate a higher motion sickness susceptibility most likely might experience more discomfort. These findings were expected as the same relationship was previously presented [143, 173, 212]. Furthermore, in the literature, it is widely assumed that the more susceptible to motion sickness individuals would be more susceptible to SiS. The findings in this thesis provide an empirical evidence to support this assumption. Previous history with motion sickness might be one factor related to gender differences in the conducted experiments. Women reported higher susceptibility to motion sickness, and they experienced more discomfort than men.

Another positive relationship was found, which was statistically significant, between *presence* and *extroversion*. Individuals who perceived themselves as extroverts feel more present in the standard VR driving simulation. Sas and OHare[190] reported that individuals who defined themselves as introverts experienced much greater *sense of presence* than individuals who defined themselves as extroverts. They concluded that the sensitive type, feeling type, more introvert, or more judging type of people experienced a higher sense of themselves at the moment of *presence*. Thus, it is interesting that our findings showed an opposing statement that extroverts feel more present in the VR driving system.

Consistent with the literature, this study found that participants who scored higher on the *neuroticism* scale experienced an increased level of SiS [44, 158, 232]. In particular, increased susceptibility to SiS symptoms such as *difficulty focusing, dizziness, blurred vision, fatigue, headache*, and *eye strain*. A possible reason could be that anxiety, which is part of neuroticism's personality dimension, directly influences VR's visual and physical perception and thus increases SiS susceptibility.

Interestingly, we found a significant positive correlation between the *arousal* before the VR driving and SiS. Participants who are more excited before the VR driving are more susceptible to higher discomfort during the driving and symptoms such as *difficulty focusing, dizziness, blurred vision, fatigue, headache*, and *eye strain*. Gugenheimer et al. [81] reported that using an HMD

led to significantly higher valence and arousal scores, particularly in terms of positive emotions. Nonetheless, they measured the emotions post-simulation to compare the level of emotion to the other non-HMD condition. The VR experience itself directly influenced the measured emotion level. In contrast, in our experiment, participants report their *arousal* level before the start of the VR simulation in order to evaluate this SiS factor as a potential predictor. Possibly, they were already excited to some extent that they would try an HMD. Nonetheless, individuals' emotional state could be influenced by other factors before the experiment on which the researcher had no control in the current investigation.

Contrary to the reported positive relationship between SiS and *sleep deprivation* [98], no correlation was reported in the experiment. Participants' sleeping aspects did not correlate with the felt discomfort. However, the *sleep deprivation* was not fully controlled, and it was not investigated over a prolonged time, and thus, the results should be interpreted with caution. Another possible factor of SiS inducing factors was evaluated for a potential relationship. However, no correlation between the *physical activity* and SiS was found. This result did not support the assumption that less physically active individuals would suffer more from SiS induced by VR simulation. In contrast to earlier findings, no evidence of relationship between *vision correction* and SiS was detected [173]. In other words, individuals who wear eyeglasses or contact lenses are not prone to feel more discomfort than individuals with normal vision.

Moreover, a positive relationship between FMS score and SSQ Disorientation and video game experience was reported. Individuals who play video games daily experience less discomfort, including symptoms such as *difficulty focusing, blurred vision*, and *dizziness* during the standard VR driving simulation. That supports the assumption that gamers feel less sick within a VR environment than non-gamers. Gamers are used to the interaction within the virtual world. Thus, their VR driving actions induced less discomfort for them as the interaction is similar to video game interactions. Besides, a significant negative relationship between SSQ Nausea scores and previous VR experience was found. Participants who had more previous experience with VR felt less nauseous and vice versa. Experience with the VR system as a SiS factor is directly related to repeated exposure to the virtual environment, also called adaptation or habituation. Adaptation is considered one of the solutions against SiS, which was well studied [88, 145, 178]. Our results show that individuals who have prior interaction with VR are less prone to symptoms such as *nausea, and stomach awareness*. However, there were no data on how often the individuals were exposed to VR and for how long. Furthermore, these findings aligned with the literature on the mismatch between current and expected experiences [171]. Individuals who already have been in the VR world have stored that experience, and in the next interaction with such an environment, they would feel less sick due to the reduced mismatch.

6.5 Summary

In this chapter, we have illustrated the assessment of *motion* and *gender* regarding SiS in the context of a standard HMD driving simulation. Standard driving with an automatic gearbox was utilized. This evaluation was the first of three user evaluations, which are presented separately in consecutive chapters. The gathered data was used for the development of the SiS prediction models. We carried out an experiment with 63 participants separated into two moving conditions: static and dynamic. The driving scenario was the same as the previous experiment: standard urban traffic, including other vehicles. The results revealed that as many as 30% of the participants experienced moderate discomfort. The most self-reported symptoms during the dynamic condition were *nausea unease*, and *feeling warm*. For the static condition, the most reported symptoms were *dizziness, nausea*, and *unease*.

This evaluation showed that participants felt significantly more SiS during the static virtual driving than the dynamic driving. Furthermore, participants in the static condition experienced more symptoms from Nausea (e.g., *stomach awareness, nausea, sweating*) and Oculomotor (e.g., *headache, eye strain, fatigue*) clusters. Surprisingly, *gender* showed no effect on SiS onset. The groups had no significantly different results regarding sickness based on the participants' gender. Further, we observed a trend that women experienced more discomfort in the absence of motion cues than men. This effect was observed as well in response to the questionnaire one hour after the VR simulation. Regrading the physiological data, no significant difference was found across the *motion* conditions or *gender*. Additionally, we investigated associations between individual factors and SiS. Several factors such as *gender, HRD, RR, motion sickness history, neuroticism, arousal, video game experience*, and *VR experience* showed statistically significant relationships to SiS. Further, we discussed the results in the context of prior research and related literature.

7 Experiment 2: Static vs. Dynamic Automated Driving

Every adventure requires a first step.

Lewis Carroll, "Alice's Adventures in Wonderland"

This chapter describes the evaluation of *motion* and *gender* influence on SiS in an automated VR driving simulation. In order to test whether the addition of moving platform and gender affects the SiS or not, we carried out a user evaluation with the VR setup described in **Chapter 5**. The SiS onset was assessed during the VR simulation, immediately after, and around one hour later. The driving scenario was an urban driving scenario with a length of 24 min.

In this experiment, fully automated driving was utilized. Automated driving is categorized into five levels, where level five relates to fully automated driving without any human intervention. The driving system has full control over all driving tasks under all road conditions managed by a human driver at the lowest level [188]. Earlier research showed that the addition of motion cues to a stationary driving simulation might alleviate the SiS [10, 48]. Analogically, we assumed that adding motion cues to an automated driving simulation would reduce the SiS occurrence. In the literature, the focus is more on the changing active (drives) to passive (passenger) role rather than on the static-dynamic aspect of automated driving. Thus, our hypotheses are based on previous studies utilized on standard driving simulators regarding SiS.

Previous studies showed that women felt more discomfort than men [72, 98, 127, 143, 212]. Thus, analogous to the previous experiment's hypotheses in **Chapter 6**, we assumed that female participants would experience higher discomfort than male participants in the VR driving simulation. Additionally, the same hypotheses regarding the interaction between *motion* and *gender*, are assumed for this experiment. Nevertheless, the driving type was different, and therefore, the driving experience differentiated from conventional driving. With the standard driving, the participants were in control of the virtual vehicle, and they were responsible for their behavior

on the road. Similar to the previous experiment, we hypothesized that the physiological signals and *sense of presence* would be affected by the addition of *motion* or *gender*. An increase in physiological signals, such as heart signal [51, 120, 194], sweating [51, 226], breathing [120], is significantly positively correlated to SiS. Therefore, it is expected that the difference across conditions would be visible in the physiological signals.

Before reporting the results from the experiments, we present a brief description of the experimental design. Moreover, participants' descriptive statistics show the *Standard Deviation (SD)* and the *Mean (M)* of variables related to socio-demographics and previous experiences as well as motion sickness history. Then the chapter continues with reporting the results. As in the previous experiment, two aspects of the data analysis focused on the hypotheses testing and the correlation analysis. These two types of analysis revealed insights regarding SiS in VR driving simulations. The chapter finishes with a summary of the reported results.

The user evaluation of the moving platform in an urban automated VR driving simulation was part of a master thesis, supervised by the author, and written by Karolin Rehm [179], during a collaboration between BMW Group and Ludwig-Maximilians-University Munich, Germany.

The chapter is based on the following publication:

- Rangelova, S., Rehm, K., Diefenbach, S., Motus, D., & André, E. (2020, July). Gender differences in simulation sickness in static vs. moving platform VR automated driving simulation. In International Conference on Human-Computer Interaction (pp. 146-165). Springer, Cham. [170]

7.1 Study Design

A 2 x 2 factorial design was chosen to investigate whether the addition of *motion* and *gender* affect the onset of SiS or not in an automated VR driving simulation. Each factor had two levels: *motion* (with motion, without motion) and *gender* (male, female). Furthermore, an investigation of human factors and how they are related to SiS was performed.

A comparison between the control and the experimental groups was conducted as well as exploratory analysis within the experimental group. The dependent variables were SiS, *sense of presence*, HR, SCL, and RR. The control variables are *age*, *vision correction*, *motion sickness history*, *experience with VR*, *experience with driving simulation*, *gaming*, and *driving habits*. Correlation analysis on human factors (i.e., *gender*, *motion sickness history*, previous experience, *vision correction*, emotions, and *physical activity*) was conducted to calculate the appearance and the strength of the correlation.

The participants were randomly assigned into two groups, driving with a moving platform and without moving platform. The exclusion criteria of the participants were the same as the

criteria in the previous experiment (Chapter 6). The data was obtained via questionnaires (Pre-questionnaire, On-Spot questionnaire, Follow-up questionnaire), physiological sensors (ECG, EDA, and respiratory effort), behavioral data (vehicle location, and rotation in the virtual world and head location and rotation). For the data analysis, the IBM SPSS Statistics version 23.0 and the Python-based library Biosppy[1] were used. The average vehicle's speed in the virtual world was 30 km/h, and the maximum speed was 59 km/h.

7.2 Participants

The study sample consisted of 62 participants (Fig. 7.1), between 19 and 61 years old ($M = 31.94, SD = 10.56$), from which 31 were female. All participants were recruited via internal e-mail system. The eligibility criteria for the participants were: being able to understand the German language, having a driver's license class B and having normal or corrected to normal vision. The participants were randomly assigned to one of the two conditions with a total number of 32 (female - 17, male - 15) for without motion and 30 (female - 14, male - 16) for motion group. The considered study sample regarding the physiological measurements is smaller than 62 participants due to technical issues with the used sensors.

The motion condition groups did not differ significantly regarding the following susceptibility factors: *age*, *driving kilometer per year*, *gaming*, *VR experience*, *simulator experience*, and *motion sickness history*. The two conditions did differ significantly regarding the *driving frequency*. The motion condition participants drove significantly more frequently than without motion condition (Table 7.1).

The female and male groups differed only significantly regarding *age*, *video game usage*, and *motion sickness history*. Male participants were older, scored higher on *video game usage*, and lower on *motion sickness history* than female participants (Table 7.2).

7.3 Results

The results showed no significant difference regarding SiS onset, *sense of presence* and physiological signals response between the *motion* and *without motion* conditions. As expected, the results showed a significant differences across genders. That opposed the results from Experiment 1 and it pointed out that in two different driving environments, the effect of the same variables (*motion* and *gender*) on SiS onset could change. Further, the results showed significant differences regarding SiS onset and some physiological signals response between female and male participants. Furthermore, the *arousal* of participants might be a possible predictor of SiS susceptibility. All graphs related to the hypotheses testing are displayed in Appendix G.

[1] https://github.com/PIA-Group/BioSPPy

Table 7.1: Descriptive statistics of the susceptibility factors differences: age, driving frequency, driving km per year, gaming experience, VR experience, driving simulator experience, and motion sickness history, between the two conditions - motion and without motion.

	Without motion			Motion		
	M	SD	N	M	SD	n
Age	32.22	11.18	32	31.63	10.04	30
Driving frequency	3.59	0.98	32	4.07	0.74	30
Driving km per year	3.53	1.30	32	3.93	1.08	30
Gaming experience	1.66	1.13	32	1.76	1.06	29
VR experience	2.84	1.32	32	2.69	1.51	29
Driving simulation experience	2.38	1.45	32	2.41	1.55	29
MSSQ-short	9.31	8.92	31	8.18	6.91	29

Table 7.2: Descriptive statistics of the susceptibility factors differences: age, driving frequency, driving km per year, gaming experience, VR experience, driving simulator experience, and motion sickness history, between female and male participants.

	Female			Male		
	M	SD	N	M	SD	n
Age	27.39	4.86	31	36.48	12.66	31
Driving frequency	3.71	0.86	31	3.94	.93	31
Driving km per year	3.45	0.10	31	4	1.34	31
Gaming experience	1.19	0.48	31	2.23	1.28	30
VR experience	2.65	1.36	31	2.33	1.52	30
Driving simulation experience	2.45	1.50	31	2.33	1.52	30
MSSQ-short	10.78	8.98	31	6.61	6.16	29

Figure 7.1: A participants using the automated VR driving setup.

A frequency regarding the answers of the self-report scales concerning SiS was calculated in order to get an overview of the magnitude of SiS occurrence in this study. The overall occurrence of discomfort, verbally reported by the participants via FMS during the VR driving, was divided into five clusters (Fig. 7.2). As many as 36% of the participants experienced moderate discomfort. Similar results were reported from the SSQ, where the severity level was low as slight (68%) and moderate (21%) sickness severity. As expected, the severity level dropped drastically after one hour and more (Fig. 7.4). The results from the FSSQ showed that all participants reported slight sickness and 16% reported no symptoms at all.

Furthermore, the most self-reported symptoms during the VR session in the motion condition were *nausea*, *dizziness*, *unease*, and *tiredness*. For the *without motion* condition, the most reported symptoms were *nausea*, *dizziness*, and *headache* (Fig. 7.3).

Moreover, 11 participants (9 female) of the considered sample size had to terminate the driving simulation early due to severe feelings of discomfort. These participants were not excluded from the data analysis because the mean value of the FMS score was taken for the statistical analysis. For the analysis of the rest of the data, the earlier withdraw was not relevant as the aim of the

experiment was to evaluate the sickness onset. A considerable number of those participants (n = 7), all female, quit between 3 min and 12 min of the driving simulation. Furthermore, most participants (36%) stopped the test between 9 min and 12 min.

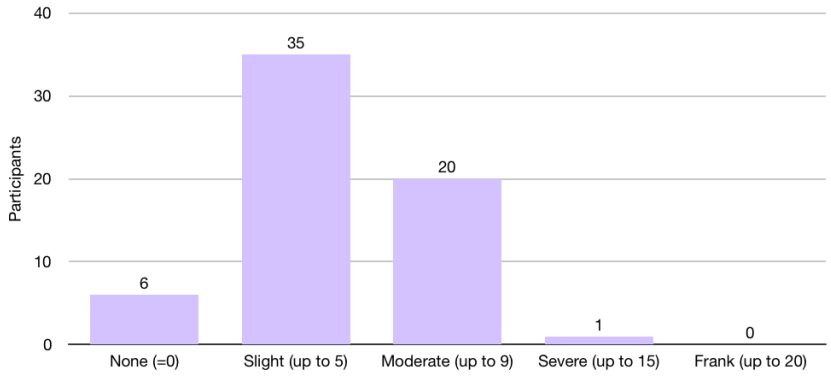


Figure 7.2: Mean FMS score frequency separated into five clusters: none, slight, moderate, severe, and frank sickness.

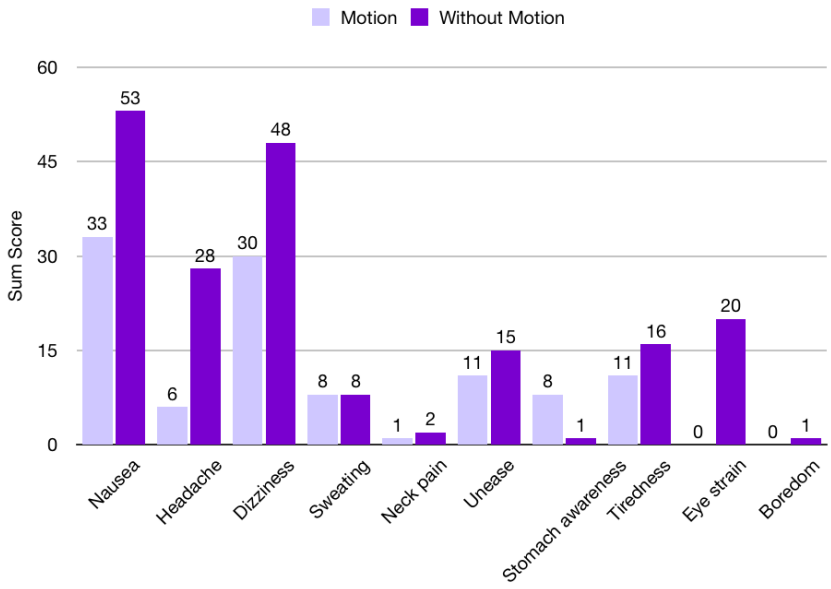


Figure 7.3: Symptoms frequency during the VR simulation for motion and without motion conditions.

7.3.1 Hypotheses Testing

FMS and SSQ Scores

A two-way Analysis of Variance (ANOVA) was conducted on the influence of two independent variables (*gender*, *motion*) on the SiS onset during, immediately after, and one hour later after the

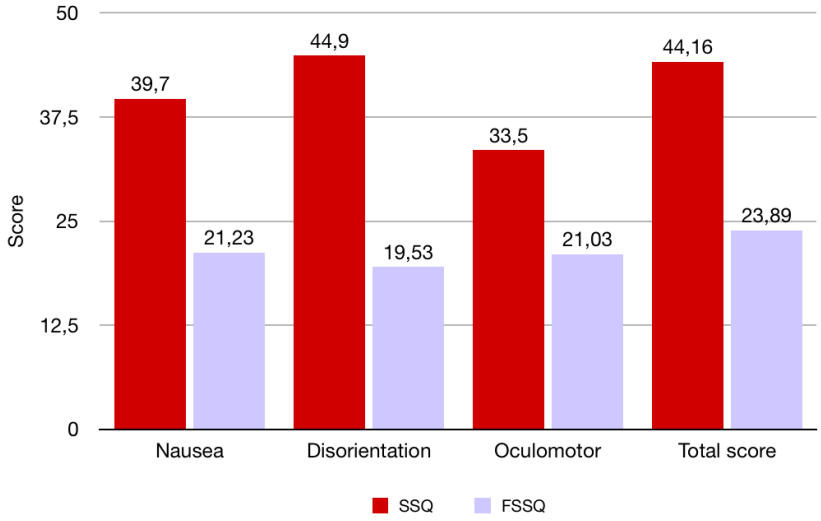


Figure 7.4: Symptoms frequency (mean score) immediately after and one hour later after the VR simulation.

driving simulation. *Gender* included two levels (female, male), and *motion* included two levels (with motion, without motion). The dependent variables were FMS score, SSQ Nausea score, SSQ Disorientation score, SSQ Oculomotor score, SSQ total score, FSSQ Nausea score, FSSQ Disorientation score, FSSQ Oculomotor score, and FSSQ total score. The effect size, Eta squared (η^2), for each variable that showed significance is reported. As guidance to which degree the variable is presented in the population, we used Cohen's guidance, where the small effect limit is 0.10, the average effect limit is 0.25, and the large effect limit is 0.40 [43]. Table 7.3 presents an overview of the Means and Standard Deviations for the ANOVA results of the independent variables.

For the FMS score (Table 7.4), only one effect (*Gender*) was statistically significant at the $.05$ significance level with average effect size of 0.17.

For the SSQ total score (Table 7.5), only one effect was statistically significant at the $.05$ significance level and an effect size of 0.17.

For the SSQ Nausea score (Table 7.6), only one effect was statistically significant at the $.05$ significance level with average effect size of 0.16.

For the SSQ Disorientation score (Table 7.7), only one effect was statistically significant at the $.05$ significance level and an average effect size of 0.16.

For the SSQ Oculomotor score (Table 7.8), only one effect was statistically significant at the $.05$ significance level with average effect size of 0.10.

Table 7.3: Overview of Means and Standard Deviations for the two-way ANOVA results using the FMS score, SSQ total score, SSQ Nausea score, SSQ Disorientation score, SSQ Oculomotor score, FSSQ total score, FSSQ Nausea score, FSSQ Disorientation score, and FSSQ Oculomotor score as criterion.

Measure	Motion		Without Motion		Male		Female	
	M	SD	M	SD	M	SD	M	SD
FMS score	3.62	2.95	3.93	2.97	2.58	2.57	4.98	2.82
SSQ total score	42.89	39.77	45.35	35.43	29.20	39.77	59.12	37.03
SSQ Nausea score	40.07	42.40	39.35	39.58	24.00	30.56	55.39	43.75
SSQ Disorientation score	44.08	45.08	45.68	43.62	27.39	34.56	62.42	45.87
SSQ Oculomotor score	31.08	26.88	35.77	27.24	25.19	26.58	41.81	25.05
FSSQ total score	31.58	36.03	21.80	18.41	19.99	22.38	33.52	31.81
FSSQ Nausea score	28.97	38.11	18.42	19.24	16.12	20.73	31.45	36.36
FSSQ Disorientation score	25.78	37.17	17.76	22.29	16.32	25.25	27.32	34.64
FSSQ Oculomotor score	26.95	26.12	19.87	16.76	18.82	19.18	28.07	23.86

Table 7.4: Two-way ANOVA results using FMS score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	0.38	1	0.38	0.050	.824	
Gender	88.46	1	88.46	11.779	.001**	0.17
Motion * Gender	0.15	1	0.15	0.020	.887	
Residuals	435.58	58	7.51			

*Note. *p < .05, **p < .01, ***p < .001.*

Table 7.5: Two-way ANOVA results using SSQ total score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	3.73	1	3.73	0.003	.956	
Gender	14061.1	1	14061.1	11.723	.001**	0.17
Motion * Gender	1422.84	1	1422.84	1.186	.281	
Residuals	69569.3	58	1199.47			

*Note. *p < .05, **p < .01, ***p < .001.*

Table 7.6: Two-way ANOVA results using SSQ Nausea score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	120.28	1	120.28	0.083	.774	
Gender	15660.1	1	15660.1	10.816	.002**	0.16
Motion * Gender	1325.66	1	1325.66	0.916	.343	
Residuals	83979.1	58	1447.92			

*Note. *p < .05, **p < .01, ***p < .001.*

Table 7.7: Two-way ANOVA results using SSQ Disorientation score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	7.86	1	7.86	0.005	.946	
Gender	19335.3	1	19335.3	11.534	.001**	0.16
Motion * Gender	1724.01	1	1724.01	1.028	.315	
Residuals	97227.4	58	1676.34			

*Note. *p < .05, **p < .01, ***p < .001.*

Table 7.8: Two-way ANOVA results using SSQ Oculomotor score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	200.37	1	200.37	0.297	.588	
Gender	4248.46	1	4248.46	6.287	.015*	0.10
Motion * Gender	620.54	1	620.54	0.918	.342	
Residuals	39195.4	58	675.78			

*Note. *p < .05, **p < .01, ***p < .001.*

FSSQ Scores

For the FSSQ total score (Table 7.9), there was a statistically significant interaction at the .05 significance level between the effect of *motion* and *gender*, $F(1, 52) = 5.571, p = .022, \eta^2 = 0.09$.

Table 7.9: Two-way ANOVA results using FSSQ total score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	1843.47	1	1843.47	2.628	.111	
Gender	3129.56	1	3129.56	4.462	.039*	0.07
Motion * Gender	3907.78	1	3907.78	5.571	.022*	0.09
Residuals	36473.1	52	701.41			

*Note. *p < .05, **p < .01, ***p < .001.*

For the FSSQ Nausea score (Table 7.10), there was a statistically significant interaction at the .05 significance level between the effect of *motion* and *gender*, $F(1, 52) = 9.582, p = .003, \eta^2 = 0.14$.

Table 7.10: Two-way ANOVA results using FSSQ Nausea score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	2222.45	1	2222.45	3.076	.085	
Gender	4061.74	1	4061.74	5.622	.021*	0.08
Motion * Gender	6921.96	1	6921.96	9.582	.003**	0.14
Residuals	37565.7	52	722.42			

*Note. *p < .05, **p < .01, ***p < .001.*

For the FSSQ Disorientation score (Table 7.11), all effects were not statistically significant at the .05 significance level.

For the FSSQ Oculomotor score (Table 7.12), all effects were not statistically significant at the .05 significance level.

Presence Scores

A two-way ANOVA was conducted on the influence of two independent variables (*gender*, *motion*) on the *sense of presence*. *Gender* included two levels (female, male), and *motion* included two levels (with motion, without motion). For the *presence* score (Table 7.13), all the effects were not statistically significant at the .05 significance level.

Table 7.11: Two-way ANOVA results using FSSQ Disorientation score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	1243.13	1	1243.13	1.435	.236
Gender	2082.25	1	2082.25	2.403	.127
Motion * Gender	2892.54	1	2892.54	3.338	.073
Residuals	45056.7	52	866.48		

*Note. *p < .05, **p < .01, ***p < .001.*

Table 7.12: Two-way ANOVA results using FSSQ Oculomotor score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	921.89	1	921.89	2.060	.157
Gender	1435.99	1	1435.99	3.208	.079
Motion * Gender	979.95	1	979.95	2.189	.145
Residuals	23274.6	52	447.59		

*Note. *p < .05, **p < .01, ***p < .001.*

Table 7.13: Two-way ANOVA results using presence score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	0.00	1	0.00	0.022	.884
Gender	0.08	1	0.08	0.417	.521
Motion * Gender	0.00	1	0.00	0.004	.947
Residuals	11.20	58	0.19		

*Note. *p < .05, **p < .01, ***p < .001.*

Physiological Signals

This study pursues the same conception as described in Chapter 6, that an increase of the measured physiological signals (ECG, EDA, and respiratory effort) indicates severer SiS occurrence. The procedure of recording the signals was precisely the same as in the previous chapter. A five-minute baseline measurement of these signals was recorded in order to have reference values of the participant's average physiological signals. During the driving simulation, these signals were continuously assessed. After the data acquisition, the signals were averaged for every participant in order to be able to analyze them further. The signals of the baseline and driving simulation were separately averaged over time for every participant. The difference between these two scores was calculated (Driving The difference between these two scores was calculated (Driving Simulation minus Baseline) for every physiological signal, HRD, SCLD, RRD, respectively. The higher values indicating a severer SiS level. It should be taken into account that only the physiological data which had both records, during the simulation and the baseline, was included in the calculation of the variable differences.

A two-way ANOVA was conducted on the influence of two independent variables (*gender*, *motion*) on the physiological signals response during the driving simulation. *Gender* included two levels (female, male), and *motion* included two levels (with motion, without motion). Table 7.14 presents an overview of the Mean and Standard Deviation for the ANOVA results of the independent variables.

Table 7.14: Overview of Means and Standard Deviations for the two-way ANOVA results using the presence score and the physiological signals as criterion.

Measure	Motion		Without Motion		Male		Female	
	M	SD	M	SD	M	SD	M	SD
Presence	41.07	6.53	40.75	6.48	41.45	5.80	40.35	7.10
HRD	4.77	12.61	0.07	14.88	-2.86	14.88	7.93	10.39
HR	79.39	14.33	74.18	11.24	72.84	12.24	80.83	12.68
SCLD	0.86	0.97	0.62	0.69	0.47	0.44	1.22	1.13
SCL	0.91	0.94	0.85	0.81	0.73	0.52	1.08	1.16
RRD	0.003	0.02	-0.01	0.04	0.003	0.03	-0.01	0.03
RR	0.24	0.04	0.24	0.03	0.25	0.03	0.23	0.04

For the HRD (Table 7.15), one effect was statistically significant at the $.05$ significance level with a small effect size 0f 0.06.

For the HR record (Table 7.16), one effect was statistically significant at the $.05$ with small effect size of 0.10.

For the SCLD (Table 7.17), one effect was statistically significant at the $.01$ significance level due

Table 7.15: Two-way ANOVA results using HRD as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	398.31	1	398.31	2.406	.127	
Gender	1615.82	1	1615.82	9.759	.003**	0.06
Motion * Gender	2.58	1	2.58	0.016	.901	
Residuals	7947.17	48	165.57			

*Note. *p < .05, **p < .01, ***p < .001.*

Table 7.16: Two-way ANOVA results using HR as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	455.33	1	455.33	3.027	.088	
Gender	992.80	1	992.80	6.600	.013*	0.10
Motion * Gender	121.97	1	121.97	0.811	.375	
Residuals	8122.51	54	150.42			

*Note. *p < .05, **p < .01, ***p < .001.*

to the violation of the homogeneity of variances. The effect size for *gender* was average - 0.24

Table 7.17: Two-way ANOVA results using SCLD as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Motion	1.54	1	1.54	2.908	.097	
Gender	5.98	1	5.98	11.285	.002*	0.24
Motion * Gender	0.40	1	0.40	0.757	.390	
Residuals	18.56	35	0.53			

*Note. *p < .01, **p < .001.*

For the SCL record (Table 7.18), all effects were not statistically significant at the $.01$ significance level due to the violation of the homogeneity of variances.

For the RRD (Table 7.19), all effects were not statistically significant at the $.05$ significance level.

For the RR record (Table 7.20), all effects were not statistically significant at the $.05$ significance level.

Table 7.18: Two-way ANOVA results using SCL as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	0.07	1	0.07	0.095	.759
Gender	1.32	1	1.32	1.757	.192
Motion * Gender	0.18	1	0.18	0.235	.630
Residuals	33.83	45	0.75		

*Note. *p < .01, **p < .001.*

Table 7.19: Two-way ANOVA results using RRD as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	0.00	1	0.00	2.405	.135
Gender	0.00	1	0.00	2.538	.125
Motion * Gender	0.00	1	0.00	0.221	.643
Residuals	0.02	22	0.00		

*Note. *p < .05, **p < .01, ***p < .001.*

Table 7.20: Two-way ANOVA results using RR as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Motion	.000	1	.000	.028	.867
Gender	.003	1	.003	2.516	.119
Motion * Gender	.002	1	.002	1.439	.236
Residuals	.052	47	.001		

*Note. *p < .05, **p < .01, ***p < .001.*

7.3.2 Correlations

The relationship of SiS, physiological signals, and presence was analyzed in order to get more insight regarding their reliability. All data using different scales (e.g., physiological data, emotions, personality, MSSQ-short) were standardized for the statistical analysis. The significant level was set to .05 with 95% Confidence Interval.

A Spearman's correlation showed a significant negative relationship between MSSQ-short score and *age*, $r_s = -.327, p = .011$ (see Appendix F). The same correlation analysis showed a significant positive relationship between SSQ Oculomotor and MSSQ-short score, $r_s = .323, p = .012$ (see Appendix F).

Another correlation analysis presented a significant positive relationship between SSQ Oculomotor and *neuroticism*, $r_s = .291, p = .024$. The other dependent variables were not associated with *neuroticism* (see Appendix F). Additionally, no correlation between FMS, SSQ, FSSQ, *presence*, and *extroversion* was found (see Appendix F).

Pearson correlation was calculated between FMS, SSQ, FSSQ scores, and *sense of presence*. The results pointed out that there was no significant association between self-reported *presence* and the SiS scores (see Appendix F).

A point-biserial correlation was run to determine the relationship between FMS, SSQ, FSSQ scores, and *gender*. There was a correlation positive between FMS, SSQ scores and *gender*, which was statistically significant (Fig. 7.5). Furthermore, the same analysis showed no correlation between FMS, SSQ, FSSQ, *presence*, and *vision correction* (see Appendix F).

Regarding the *sleep deprivation* no significant associations were found (see Appendix F). A Spearman's correlation pointed out a significant positive relationship between SSQ Oculomotor and *arousal* before the driving simulation, $r_s = .270, p = .034$. The other variables showed no relationship to *arousal* (see Appendix F). However, no relationship between the SiS variables and *valence* before the driving simulation was found (see Appendix F).

A Spearman's correlation showed a significant positive relationship between the HRD and the SSQ total score ($r_s = .291, p = .037$), the SSQ Nausea ($r_s = .308, p = .026$), and the SSQ Disorientation ($r_s = .321, p = .021$) (Table 7.21). A Kendall's Tau correlation was conducted and there was a significant relationship between HR record and SSQ Nausea score ($r_\tau = .186, p = .047$) (Table 7.21).

The same correlation analysis, Spearman's correlation, showed no significant relationship between the tested variables and the SCLD, as well as the tested variables and the RRD (see Appendix F). There was a significant positive relationship between SSQ Nausea score and SCL record ($r_\tau = .218, p = .036$) (see Appendix F). Furthermore, there was a significant negative correlation between RR record and FMS score ($r_\tau = -.216, p = .034$), SSQ total score ($r_\tau = -.240, p = .019$), SSQ Nausea score ($r_\tau = -.300, p = .004$), and FSSQ Nausea ($r_\tau = -.240, p = .033$) (Table 7.22).

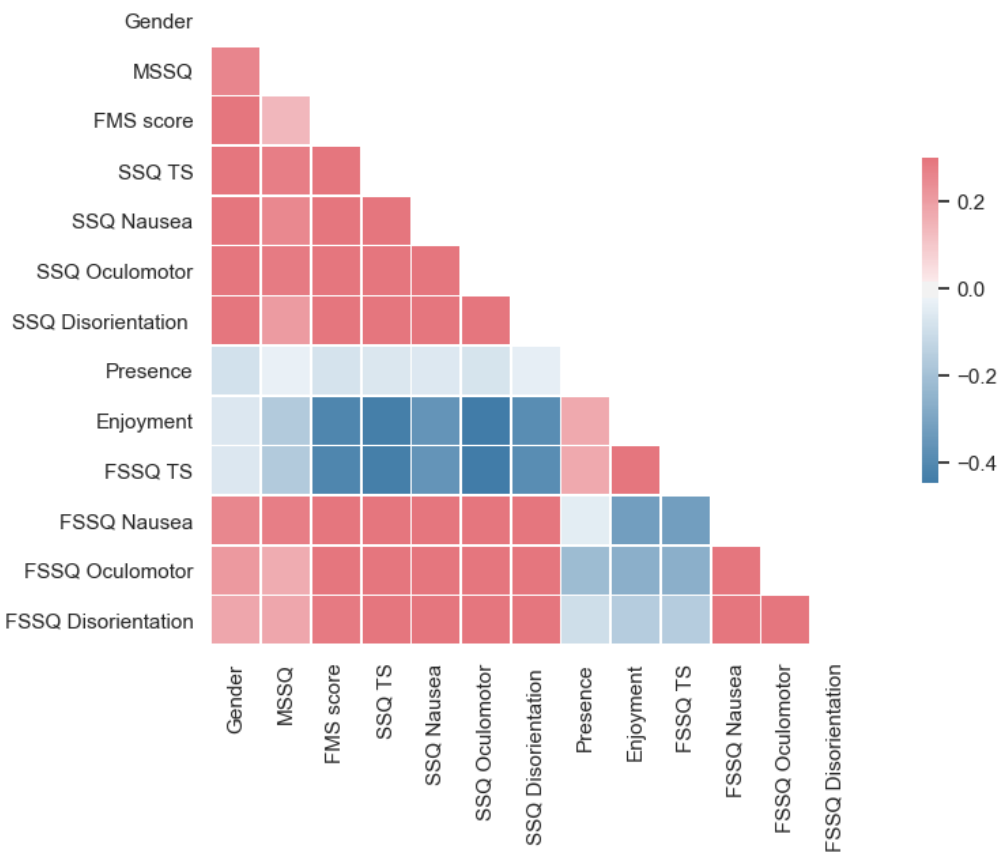


Figure 7.5: A Pearson's correlation matrix showing the relationship between MSSQ-short, FMS, SSQ, presence, FSSQ scores, and gender. For the complete results including the p-values, please see Appendix F.

Table 7.21: Correlations between MSSQ-short, FMS, SSQ, FSSQ scores, HR difference (HRD), and HR record.

	HRD (Spearman)	n	HR record (Kendall)	n
FMS score	.124	52	.094	58
SSQ total score	.291*	52	.163	58
SSQ Nausea	.308*	52	.186*	58
SSQ Disorientation	.321*	52	.149	58
SSQ Oculomotor	.196	52	.112	58
FSSQ total score	.017	46	.046	52
FSSQ Nausea	.119	46	.119	52
FSSQ Disorientation	.092	46	.064	52
FSSQ Oculomotor	-.007	46	.014	52

Note. *$p < .05$, **$p < .01$, ***$p < .001$.

Table 7.22: Correlations between MSSQ-short, FMS, SSQ, FSSQ scores, RR difference (RRD), and RR record.

	RRD (Spearman)	n	RR record (Kendall)	n
FMS score	-.251	26	-.216*	51
SSQ total score	-.215	26	-.240*	51
SSQ Nausea	-.317	26	-.300**	51
SSQ Disorientation	-.052	26	-.193	51
SSQ Oculomotor	-.038	26	-.185	51
FSSQ total score	.202	25	-.115	47
FSSQ Nausea	.064	25	-.240*	47
FSSQ Disorientation	.151	25	-.115	47
FSSQ Oculomotor	.261	25	-.114	47

*Note. *p < .05, **p < .01, ***p < .001.*

7.4 Discussion

None of the hypotheses regarding the effect of *motion* were supported by the results, which showed no significant difference across the conditions regarding SiS onset. Contrary to the expectations, the addition of motion cues did not contribute to reducing the SiS symptoms significantly. The mitigating effect of physical motion cues regarding SiS in VR driving simulations could not be replicated [10, 48]. However, a previous study showed that the addition of motion cues did not affect the sickness outbreak [108]. The results were explained with the physical limits of the moving platform or with slight time delays between driver inputs and the motion response. The authors suggested that the addition of a moving platform could have a more significant effect when the vehicle's maneuvers are intentionally limited to current movements, which the participants can anticipate. In the current study, however, this assumption was not supported. Thus, the response time of the moving platform is a possible reason for the not significant effect of the moving platform. Despite the DOF of the platform, some of the maneuvers could be difficult to replicate accurately, and therefore, a mismatch between the sensory systems might arise. Furthermore, the questionnaire given one hour later (FSSQ) did not report any significant results across the *motion* conditions. The participants recovered similarly in both conditions regardless of the physical cues.

As expected, most of the hypotheses regarding the effect of *gender* were supported, which showed a significant difference between female and male participants. These results match those observed in earlier studies [98, 127, 212]. A possible reason could be the wider FOV [34], and the more influence of the visual cues [220] which could lead to more significance of the visual input and therefore, to contribute to the higher mismatch between the visual and motion cues.

The SiS severity in the experiments appears to be slight to moderate since the frequencies of the self-reported SiS severity mean scores range from the lower to the middle section of the used scales. Even though the SiS occurrence was not severe, the early dropout rate of 17.7% emphasizes the importance of finding mitigation against SiS in VR driving simulations. Moreover, no interaction was reported between *motion* and *gender* on SiS onset during the VR driving.

A trend was observed in all subjective measurements that female participants felt more discomfort as a result of adding motion to the VR driving simulation. We observed the same effect in the response of the follow-up questionnaire (FSSQ). In particular, the interaction between *motion* and *gender* regarding the FSSQ total score and FSSQ Nausea score was statistically significant. Overall, women recovered slower from the felt discomfort than men and they experienced longer the side effect of the VR driving including symptoms such as *nausea, sweating*, and *general discomfort*.

An average positive correlation was reported between SiS and *gender*. Women tend to feel more discomfort during and immediately after the automated VR driving simulation. Numerous studies have documented the relationship between *gender* and SiS, and more specifically, women are more susceptible to SiS than men [62, 98, 127, 131, 174, 212]. Stanney et al. [212], for example, found a linear relationship between SSQ Disorientation, SSQ Oculomotor, and SSQ Total score and *gender*. The results from the experiment showed a statistically significant positive relationship between sickness scores (e.g., SSQ and FSSQ scores) and *gender*. Our findings are in accord with previously reported *gender* differences in VR simulations. However, the exact reason for this relationship is still not apparent. A possible explanation might be the wider FOV [34], different motion perception in virtual environments [23], or insufficient IPD calibration [213].

We found that the addition of the moving platform or *gender* did not have a significant effect on the *sense of presence*. These findings match those reported in the previous chapter regarding the standard VR driving experiment (see Chapter 6.3). The addition of motion cues influenced the feeling of *presence* neither in the standard nor in the automated driving experiment. The participants felt almost the same in terms of *presence* regardless of the motion conditions or their *gender*. The results are surprising, as we assumed that stimulating two of the primary sensory systems, vestibular and visual systems, would contribute to experiencing a higher *sense of presence*.

Regarding the physiological data, the recorded HR and the measured difference between the baseline and the recorded HR (HRD) and skin conductance (SCLD) showed a significant difference across *genders*. The female participants had a significantly higher level of HRD, and HR record than the male participants. These findings aligned with previous research on physiological signals and SiS [42, 51]. However, no significant interaction was reported between *motion* and *gender* regarding any of the measured physiological signals. These results pointed out that HR is one of the reliable objective measurements regarding SiS onset. The objectively recorded data supported the self-reported data. These findings are in line with those of previous studies that women get

more simulation sick than men [72, 143, 212].

A significant positive correlation was reported between SSQ Nausea, SSQ Disorientation, SSQ total score, and the difference between the baseline and the recorded HR (HRD). When the discomfort increases, the recorded HR differentiates significantly from baseline and vice versa. Our results are consistent with those of Cobb et al. [42] who reported that participants with severer SiS symptoms showed increased HR than participants who did not experience SiS. There are similarities between the correlation expressed by HR in this study and described individual differences by Johnson [98]. Nevertheless, the author remarked that the direction of these physiological changes could individually differ. Furthermore, another significant positive correlation was found between the HR record and SSQ Nausea. In other words, when the HR during the VR driving increases, then the susceptibility to symptoms such as *nausea* and *stomach awareness* increases and vice versa. These findings match those observed in earlier studies [51, 153]. Although the HR showed a correlation to some sickness symptoms, these results should be interpreted with caution because of the low correlation value.

Additionally, the SSQ Nausea score showed a significant positive relationship with the SCL record. Individuals are sweating more when the severity of SiS symptoms such as *nausea* and *stomach awareness* increases. These results are consistent with data obtained in earlier work [51, 226]. Further, a positive relationship was found between skin temperature and SiS [153]. A higher skin temperature leads to sweating and shows that the body is under stress. A possible explanation might be that many participants reported that they feel warm during the VR simulation, which is an indicator of increased body temperature. The found positive correlation between SiS and SCL further supports the assumption that an increase in physiological signals is related to discomfort outbreak. Nevertheless, caution must be applied with the small sample size of SCL, as the findings might not significantly impact a general population.

Similar to the previous experiment, the standard VR driving (see Chapter 6.3), a significant negative relationship was found between FMS, SSQ total score, SSQ Nausea score, FSSQ Nausea score, and RR record. The participants slow down their breathing when they experienced more discomfort and vice versa. These results seem to be consistent with other research, which found a decrease in the RR during a VR simulation [120]. However, the same study reported a positive relationship between SiS and RR, which contradicts this thesis's current findings. As discussed in the previous chapter, the findings could be explained with the human body's automatic response to the aroused nervous system and the actions to calm the nervous system. Nonetheless, with the small sample size of RR, caution must be applied, as the findings might not significantly impact a general population.

We found a positive correlation between MSSQ-short total score and SSQ Oculomotor score. Individuals who indicated a higher susceptibility to motion sickness tend to experience SiS symptoms such as *fatigue, headache*, and *eye strain*. These findings are in line with those of earlier studies [143, 173, 212] and previous standard VR driving experiment (see Chapter 6.3).

Interestingly, we found that the measured level of *arousal* before the VR simulation showed a significantly positive correlation to SiS symptoms from the Oculomotor cluster. Participants who are more excited before the VR driving are more susceptible to symptoms such as *headaches, fatigue*, and *eye strain*. These results are similar to the findings reported in the previous chapter regarding the standard VR driving experiment (see Chapter 6.3). However, here only the Oculomotor symptoms showed significance. As discussed in the previous chapter, a possible explanation for the positive relationship could be the preceding excitement of using an HMD for the first time. However, other factors, on which the research had no control in the current study, could have influence individuals' emotional state.

Regarding *neuroticism*, our results match those reported from the previous chapter's standard driving experiment (see Chapter 6.3). The results align with those observed in earlier studies that participants who scored higher on the *neuroticism* scale experienced an increased discomfort [44, 158, 232]. For example, SiS symptoms such as *fatigue, headache*, and *eye strain*. A possible explanation could be that SiS susceptibility increases due to anxiety, which is part of neuroticism's personality dimension and might directly influence VR's visual and physical perception.

7.5 Summary

We have described, in this chapter, the evaluation of *motion* and *gender* regarding SiS induced by an automated HMD driving simulation. This evaluation was the second of the three user evaluations, which are presented separately in this thesis. The collected data was used for the development of the SiS prediction models. We carried out an experiment with 62 participants separated into two moving conditions: static and dynamic. The driving scenario was standard urban traffic, including other vehicles. The results showed that as many as 36% of the participants experienced moderate discomfort. The most self-reported symptoms during the dynamic condition were *nausea*, *dizziness*, *unease*, and *tiredness*. For the static condition, the most reported symptoms were *nausea*, *dizziness*, and *headache*.

This investigation revealed that none of the hypotheses regarding the effect of *motion* were supported by the findings, which showed no significant difference across the motion conditions regarding SiS onset. Contrary to the expectations, motion cues did not contribute to reducing the SiS symptoms significantly. As expected, most of the hypotheses regarding *gender* were supported, which displayed a significant difference between female and male participants. We observed a trend in all subjective measurements that female participants felt more discomfort due to the addition of motion. Furthermore, the same effect was seen in the questionnaire's responses one hour after the VR simulation. Regarding the physiological data, the results showed that female participants had a significantly higher level of HRD and HR record than male participants. Further, we investigated relationships between individual factors and SiS. A few factors such as *gender, motion sickness history, arousal, neuroticism, HRD, HR, SCL*, and *RR* showed a

statistically significant relationship to SiS. Furthermore, we discussed the findings in the setting of relevant research and earlier work.

8 Experiment 3: Standard vs. Automated Driving

"And what is the use of a book," thought Alice, "without pictures or conversations?"

Lewis Carroll, "Alice's Adventures in Wonderland"

This chapter describes the evaluation of *motion*, *gender*, and *type of driving* influence on SiS in a VR driving simulation. The experiment had the same setup described in the methodology **Chapter 5**, including an HMD, a moving platform, and physiological sensors. The SiS onset was assessed during the VR simulation, immediately after, and around one hour later. The driving scenario was an urban driving scenario with a length of 24 min.

The previous two chapters (**Chapter 6** and **Chapter 7**) reported results from the experiments with standard and automated driving separately. The focus was on the SiS factors within the VR driving simulation operationalized by a particular driving type. This chapter continues reporting the results regarding SiS factors, but a new independent factor, *type of driving*, is added. We investigated whether the type of driving affects SiS or not. Vehicle control is one of the SiS inducing factors, and it is especially relevant within virtual driving environments. Earlier studies showed that passengers are more likely to experience SiS than drivers in a virtual environment [59]. The more passive role of the passenger can increase the tendency to SiS onset in virtual vehicles. As the role of the driver changes within the automated driving, it is assumed that the driver in this type of driving will feel more discomfort than the driver in the standard driving condition.

Furthermore, as sensory rearrangement theory states in **Chapter 2**, the current experience is continuously compared with the experiences collected in the neural store. Although the driving experience in the automated condition is very similar to riding a car as a passenger, the sitting

individual on the driving seat and not controlling the vehicle could be perceived as a new experience for some individuals. Thus, it could lead to SiS outbreak, higher than in the standard driving condition. We hypothesized that there would be a difference between the standard and automated driving regarding SiS. It is expected that automated driving would evoke a higher level of SiS. Moreover, analogically to the previous experiments, we assumed that **motion** and **gender** would have an effect on SiS. Motion condition would induce less discomfort than the static condition, and female participants would suffer more from SiS than male participants.

Regarding the physiological signals, we hypothesized that there would be differences between the participants in the automated and the standard driving conditions. The physiological response will be stronger in automated driving than on standard driving, which would align with the hypothesis stated above regarding the SiS outbreak. Regarding *sense of presence*, we assumed that individuals would feel more present while driving the virtual vehicle themselves. As the standard driving condition involves more a participant in the VR simulation than the automated driving, it is expected that the participant experiences higher *sense of presence*. Furthermore, it is expected that there is a significant interaction between *motion*, *gender*, and *type of driving* on SiS, physiological signals, and *sense of presence*.

Before reporting the results from the experiments, a brief description of the experimental design is presented. Furthermore, we have reported descriptive statistics of participants' socio-demographic and previous experiences, as well as motion sickness history. Next, the chapter continues with reporting the results. As in the previous experiment, two aspects of the data analysis were in focus: the hypotheses testing and the correlation analysis. These two types of analysis revealed insights regarding SiS in VR driving simulations. The chapter concludes with a summary of the reported results.

8.1 Study Design

A 2 x 2 x 2 factorial design was chosen to investigate whether one of the following factors affects the onset of SiS or not. Each factor had two levels: *type of driving* (standard, automated), *motion* (with motion, without motion), and *gender* (male, female). Additionally, a correlation analysis on human factors (i.e., *gender*, *motion sickness history*, *previous experience*, *vision correction*, *emotions*, and *physical activity*) was conducted to calculate the appearance and the strength of the relationship. The dependent variables were SiS, *sense of presence*, HR, SCL, and RR. The control variables were *age*, *vision correction*, *motion sickness history*, *experience with VR*, *experience with driving simulation*, *gaming*, and *driving habits*.

The data for this experiment were collected from the two previous experiments described in Chapter 6 and Chapter 7. Only the participants who participated in one of the driving conditions, standard or automated, were included in the sample. For the data analysis, the IBM SPSS

Statistics version 23.0 and the Python-based library Biosppy[1] were used.

8.2 Participants

The study sample consisted of 66 participants, between 18 and 60 years old ($M = 32.79, SD = 10.59$). All participants were recruited via internal e-mail system. The sample size for the factor *driving* was almost equally distributed, 35 (female - 10, male - 25) for standard and 31 (female - 16, male - 15) for automated driving. The sample size for the factor *motion* was with ratio 1.5 between the groups, 28 (female - 9, male - 19) for the with motion and 38 (female - 17, male - 21) for the without motion. The sample size for the factor *gender* was similarly distributed, with ratio 1.5 between the groups, 26 females and 40 males. The standard and automated driving groups differed slightly regarding *age*, *VR experience*, *driving simulation experience*, and *motion sickness history*. The participants in the standard driving group were older, scored lower on *VR experience*, *driving simulation experience*, and *motion sickness history* than the participants in the automated driving group (Table 8.1). The motion condition groups did not differ significantly regarding the following susceptibility factors: *age*, *driving kilometer per year*, *gaming*, *VR experience*, *simulator experience*, and *motion sickness history* (Table 8.2). The female and male groups differed only significantly regarding *age* and *motion sickness history*. Male participants were older and scored lower on *motion sickness history* than females (Table 8.3).

Table 8.1: Descriptive statistics of the susceptibility factors differences: age, driving frequency, driving km per year, gaming experience, VR experience, driving simulator experience, and motion sickness history, between the two conditions - standard and automated driving.

	Standard			Automated		
	M	SD	n	M	SD	n
Age	34.14	11.85	35	31.26	8.90	31
Driving frequency	3.60	4.00	35	3.87	0.81	31
Driving km per year	3.40	1.24	35	3.81	1.28	31
Gaming experience	2.06	1.33	35	1.81	1.20	31
VR experience	2.51	1.40	35	3.00	1.32	31
Driving simulation experience	2.14	1.27	35	2.90	1.60	31
MSSQ-short	9.00	8.60	35	10.45	9.62	31

[1] https://github.com/PIA-Group/BioSPPy

Table 8.2: Descriptive statistics of the susceptibility factors differences: age, driving frequency, driving km per year, gaming experience, VR experience, driving simulator experience, and motion sickness history, between the two conditions - with motion and without motion.

	Without motion			Motion		
	M	SD	n	M	SD	n
Age	31.68	10.75	38	34.29	10.36	28
Driving frequency	3.63	0.85	38	3.86	0.76	28
Driving km per year	3.42	1.29	38	3.82	1.22	28
Gaming experience	1.97	1.37	38	1.89	1.13	28
VR experience	2.68	1.44	38	2.82	1.31	28
Driving simulation experience	2.53	1.48	38	2.46	1.48	28
MSSQ-short	10.45	9.87	38	8.64	7.86	28

Table 8.3: Descriptive statistics of the susceptibility factors differences: age, driving frequency, driving km per year, gaming experience, VR experience, driving simulator experience, and motion sickness history, between female and male participants

	Female			Male		
	M	SD	n	M	SD	n
Age	28.50	5.16	26	35.58	12.23	40
Driving frequency	3.81	0.80	26	3.68	0.83	40
Driving km per year	3.38	0.804	26	3.73	1.49	40
Gaming experience	1.85	1.46	26	2.00	1.13	40
VR experience	2.81	1.23	26	2.70	1.47	40
Driving simulation experience	2.62	1.47	26	2.43	1.48	40
MSSQ-short	13.31	10.06	26	7.33	7.56	40

8.3 Results

The most self-reported symptoms after the VR motion session were *general discomfort*, *fatigue*, and *nausea* (Fig. 8.1). All graphs related to the hypotheses testing are displayed in Appendix G.

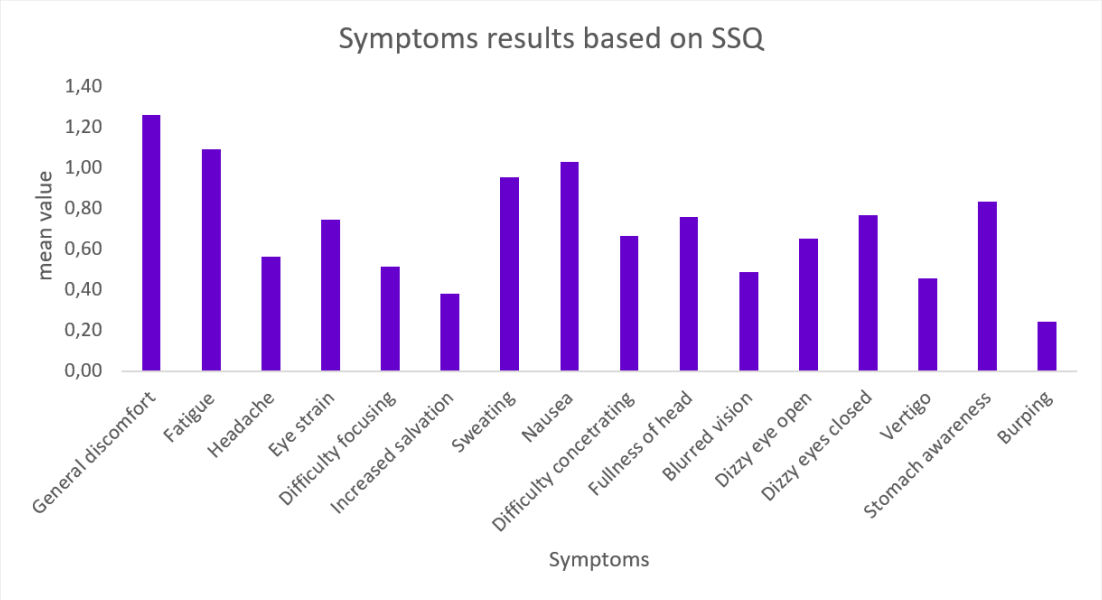


Figure 8.1: Symptoms reported immediately after the VR simulation for automated and standard driving conditions.

8.3.1 Hypotheses Testing

FMS and SSQ Scores

A three-way ANOVA was conducted on the influence of three independent variables (*drive*, *motion*, *gender*) on the SiS during, immediately after, and one hour later, after the driving simulation. *Drive* included two levels (standard, automated), the *motion* included two levels (with motion, without motion), and *gender* included two levels (female, male). The dependent variables were FMS score, SSQ Nausea score, SSQ Disorientation score, SSQ Oculomotor score, SSQ total score, FSSQ Nausea score, FSSQ Disorientation score, FSSQ Oculomotor score, and FSSQ total score. The effect size, Eta squared (η^2), for each variable that showed significance is reported. As guidance to which degree the variable is presented in the population, we used Cohen's guidance, where the small effect limit is 0.10, the average effect limit is 0.25, and the large effect limit is 0.40 [43]. Table 8.4 presents an overview of the Means and Standard Deviations for the ANOVA results of the independent variables.

Table 8.4: Overview of Means and Standard Deviations for the two-way ANOVA results using the FMS score, SSQ total score, SSQ Nausea score, SSQ Disorientation score, SSQ Oculomotor score, FSSQ total score, FSSQ Nausea score, FSSQ Disorientation score, and FSSQ Oculomotor score as criterion.

Measure	Motion		Without Motion		Male		Female		Standard drive		Automated drive	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
FMS score	2.41	2.50	3.78	2.43	2.73	2.60	3.92	2.60	3.35	2.46	3.03	2.65
SSQ total score	48.09	43.87	62.79	37.11	43.76	36.36	76.24	39.10	58.66	41.61	54.17	39.67
SSQ Nausea score	45.32	45.98	58.24	43.11	38.40	38.98	74.85	43.98	56.97	43.69	48.01	45.59
SSQ Disorientation score	51.21	56.45	65.20	48.66	46.98	47.04	78.17	54.85	64.03	57.51	53.88	45.71
SSQ Oculomotor score	33.57	27.86	45.48	26.02	31.59	25.64	52.48	25.63	38.77	28.23	42.30	26.46
FSSQ total score	30.93	33.56	42.35	47.72	21.76	23.70	61.17	52.72	39.70	49.00	34.08	32.16
FSSQ Nausea score	29.03	37.35	36.06	40.20	18.30	22.14	55.61	47.86	34.63	41.04	30.74	36.50
FSSQ Disorientation score	24.28	34.48	40.51	61.95	16.68	26.03	58.75	69.16	37.94	62.66	27.32	34.20
FSSQ Oculomotor score	26.39	22.74	35.13	35.74	20.47	19.74	47.81	36.95	32.28	36.25	29.76	22.64

For the FMS score (Table 8.5), only the effect of *gender* was statistically significant at the .05 significance level with small effect size of 0.07.

Table 8.5: Three-way ANOVA results using FMS score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Drive	6.91	1	6.91	1.144	.289	
Motion	10.75	1	10.75	1.781	.187	
Gender	27.47	1	27.47	4.552	.037*	0.07
Drive * Motion	1.86	1	1.86	0.308	.581	
Drive * Gender	3.49	1	3.49	0.579	.450	
Motion * Gender	14.88	1	14.88	2.465	.122	
Drive * Motion * Gender	0.816	1	0.816	0.135	.714	
Residuals	350.08	58	6.04			

*Note. *p < .05, **p < .01, ***p < .001.*

For the SSQ total score (Table 8.6), only the effect of *gender* was statistically significant at the .05 significance level and an average effect size of 0.15.

Table 8.6: Three-way ANOVA results using SSQ total score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Drive	1099.67	1	1099.67	0.763	.386	
Motion	1259.56	1	1259.56	0.874	.354	
Gender	15775.5	1	15775.5	10.951	.002**	0.15
Drive * Motion	1935.26	1	1935.26	1.343	.251	
Drive * Gender	262.17	1	262.17	0.182	.671	
Motion * Gender	2.21	1	2.21	0.002	.969	
Drive * Motion * Gender	56.39	1	56.39	0.039	.844	
Residuals	83552.4	58	1440.56			

*Note. *p < .05, **p < .01, ***p < .001.*

For the SSQ Nausea score (Table 8.7), only the effect of *gender* was statistically significant at the .05 significance level and an average effect size of 0.16.

For the SSQ Disorientation score (Table 8.8), only the effect of *gender* was statistically significant at the .05 significance level with small effect size of 0.10.

For the SSQ Oculomotor score (Table 8.9), only the effect of *gender* was statistically significant at the .05 significance level with small effect size of 0.10.

Table 8.7: Three-way ANOVA results using SSQ Nausea score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Drive	2790.45	1	2790.45	1.633	.206	
Motion	756.15	1	756.15	0.442	.509	
Gender	21129.8	1	21129.8	12.365	.001**	0.16
Drive * Motion	2217.09	1	2217.09	1.297	.259	
Drive * Gender	417.95	1	417.95	0.245	.623	
Motion * Gender	1.85	1	1.85	0.001	.974	
Drive * Motion * Gender	214.79	1	214.79	0.126	.724	
Residuals	99111.6	58	1708.82			

*Note. *p < .05, **p < .01, ***p < .001.*

Table 8.8: Three-way ANOVA results using SSQ Disorientation score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Drive	3234.99	1	3234.99	1.228	.272	
Motion	369.38	1	369.38	0.243	.624	
Gender	17043.9	1	17043.9	6.470	.014*	0.10
Drive * Motion	1169.01	1	1169.01	0.444	.508	
Drive * Gender	340.13	1	340.13	0.129	.721	
Motion * Gender	255.47	1	255.47	0.097	.757	
Drive * Motion * Gender	18.70	1	18.70	0.007	.933	
Residuals	152784	58	2634.20			

*Note. *p < .05, **p < .01, ***p < .001.*

Table 8.9: Three-way ANOVA results using SSQ Oculomotor score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Drive	32.88	1	32.88	0.048	.827	
Motion	1318.53	1	1318.53	1.936	.169	
Gender	4620.18	1	4620.18	6.785	.012*	0.10
Drive * Motion	1097.53	1	1097.53	1.612	.209	
Drive * Gender	42.64	1	42.64	0.063	.803	
Motion * Gender	21.28	1	21.28	0.031	.860	
Drive * Motion * Gender	35.16	1	35.16	0.052	.812	
Residuals	39494.8	58	680.95			

*Note. *p < .05, **p < .01, ***p < .001.*

FSSQ Score

For the FSSQ total score (Table 8.10), only the main effect of *gender* was statistically significant at the .01 significance level. The statistical level was reduced due to a significant result of Levene's test and unequal sample size. The effect size for *gender* was average - 0.18.

Table 8.10: Three-way ANOVA results using FSSQ total score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η
Drive	1301.77	1	1301.77	1.004	.321	
Motion	579.84	1	579.84	0.447	.507	
Gender	18281.22	1	18281.22	14.099	<.001**	0.18
Drive * Motion	5755.84	1	5755.84	4.439	.040	
Drive * Gender	1628.66	1	1628.66	1.256	.268	
Motion * Gender	372.43	1	372.43	0.287	.594	
Drive * Motion * Gender	2912.83	1	2912.83	2.246	.140	
Residuals	66128.92	51	1296.65			

*Note. *p < .01, **p < .001.*

For the FSSQ Nausea (Table 8.11), only the main effect of *gender* was statistically significant at the .01 significance level. The statistical level was reduced due to a significant result of Levene's test and unequal sample size. The effect size of *gender* was average - 0.20.

For the FSSQ Disorientation score (Table 8.12), only the main effect of *gender* was statistically significant at the .01 significance level. The statistical level was reduced due to a significant result

Table 8.11: Three-way ANOVA results using FSSQ Nausea score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Drive	554.89	1	554.89	0.484	.490	
Motion	56.37	1	56.37	0.049	.825	
Gender	17359.35	1	17359.35	15.153	<.001**	0.20
Drive * Motion	3986.13	1	3986.13	3.480	.068	
Drive * Gender	162.64	1	162.64	0.142	.708	
Motion * Gender	0.27	1	0.27	0.000	.988	
Drive * Motion * Gender	3862.30	1	3862.30	3.372	.072	
Residuals	58424.15	51	1145.57			

*Note. *p < .01, **p < .001.*

of Levene's test and unequal sample size. The effect size of *gender* was slightly above the small effect size value - 0.13.

Table 8.12: Three-way ANOVA results using FSSQ Disorientation score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Drive	3081.01	1	3081.01	1.531	.222	
Motion	1961.11	1	1961.11	0.975	.328	
Gender	19687.79	1	19687.79	9.785	.003*	0.13
Drive * Motion	10035.51	1	10035.51	4.988	.030	
Drive * Gender	3571.45	1	3571.45	1.775	.189	
Motion * Gender	1893.52	1	1893.52	0.941	.337	
Drive * Motion * Gender	3643.90	1	3643.90	1.811	.184	
Residuals	102613.04	51	2012.02			

*Note. *p < .01, **p < .001.*

For the FSSQ Oculomotor score (Table 8.13), only the main effect of *gender* was statistically significant at the $.01$ significance level. The statistical level was reduced due to a significant result of Levene's test and unequal sample size. *Gender* showed an average effect size of 0.16.

Presence Scores

A three-way ANOVA was conducted on the influence of three independent variables (*drive*, *motion*, *gender*) on the *sense of presence*. *Drive* included two levels (standard, automated), *motion* included two levels (with motion, without motion), and *gender* included two levels

Table 8.13: Three-way ANOVA results using FSSQ Oculomotor score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Drive	584.54	1	584.54	0.806	.373	
Motion	350.03	1	350.03	0.483	.490	
Gender	8637.26	1	8637.26	11.913	.001*	0.16
Drive * Motion	2404.45	1	2404.45	3.316	.074	
Drive * Gender	1529.94	1	1529.94	2.110	.152	
Motion * Gender	250.28	1	250.28	0.345	.559	
Drive * Motion * Gender	735.59	1	735.59	1.015	.319	
Residuals	36976.5	51	725.03			

*Note. *p < .01, **p < .001.*

(female, male). The dependent variable was *sense of presence*. Table 8.15 presents an overview of the Mean and Standard Deviation for the ANOVA results of the independent variables.

For the *presence* (Table 8.14), only the effect of the *drive* was statistically significant at the $.05$ significance level and an effect size of 0.09.

Table 8.14: Three-way ANOVA results using presence score as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Drive	272.5	1	272.5	6.241	.015**	0.09
Motion	1.05	1	1.05	0.024	.877	
Gender	10.50	1	10.50	0.241	.626	
Drive * Motion	0.21	1	0.21	0.005	.945	
Drive * Gender	8.77	1	8.77	0.201	.656	
Motion * Gender	2.49	1	2.49	0.057	.812	
Drive * Motion * Gender	75.43	1	75.43	1.728	.194	
Residuals	2532.48	58	43.66			

*Note. *p < .05, **p < .01, ***p < .001.*

Physiological Signals

A three-way ANOVA was conducted on the influence of three independent variables (*drive*, *motion*, *gender*) on the physiological signals response to the VR driving simulation. *Drive* included two levels (standard, automated), *motion* included two levels (with motion, without motion), and *gender* included two levels (female, male). The physiological signals of the

baseline and driving simulation were separately averaged over time for every participant. The difference between these two scores was calculated (Driving Simulation minus Baseline) for every physiological signal; HR difference (HRD), SCL difference (SCLD), RR difference (RRD). Thus, the dependent variables were HRD, SCLD, RRD, HR, SCL, and RR. Table 8.15 presents an overview of the Means and Standard Deviations for the ANOVA results of the independent variables.

Table 8.15: Overview of Means and Standard Deviations for the two-way ANOVA results using the FMS score, SSQ total score, SSQ Nausea score, SSQ Disorientation score, SSQ Oculomotor score, FSSQ total score, FSSQ Nausea score, FSSQ Disorientation score, and FSSQ Oculomotor score as criterion.

Measure	Motion		Without Motion		Male		Female		Standard drive		Automated drive	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Presence	41.46	6.76	42.32	6.78	42.73	5.29	40.77	8.48	43.97	5.14	39.68	7.64
HRD	6.26	10.81	0.71	14.98	1.17	15.03	6.30	10.02	3.17	14.44	2.78	12.98
HR	78.56	11.62	74.93	12.05	75.00	10.24	79.19	12.05	77.60	8.46	75.16	13.19
SCLD	0.86	0.78	0.69	0.61	0.52	0.53	1.30	0.69	0.90	0.70	0.68	0.67
SCL	1.06	0.76	0.91	0.66	0.80	0.49	1.24	0.89	0.89	0.57	1.03	0.79
RRD	0.02	0.03	-0.01	0.04	0.01	0.03	-0.01	0.04	0.02	0.04	-0.01	0.03
RR	0.24	0.03	0.24	0.02	0.24	0.22	0.24	0.02	0.24	0.02	0.24	0.03

For the HRD (Table 8.16), all effects were not statistically significant at the .05 significance level.

Table 8.16: Three-way ANOVA results using HRD as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Drive	1.63	1	1.63	0.008	.929
Motion	171.07	1	171.07	0.927	.341
Gender	252.55	1	252.55	1.368	.248
Drive * Motion	215.35	1	215.35	1.167	.286
Drive * Gender	87.03	1	87.03	0.471	.496
Motion * Gender	285.15	1	285.15	1.545	.220
Drive * Motion * Gender	54.90	1	54.90	0.297	.588
Residuals	8491.68	46	184.60		

*Note. *p < .05, **p < .01, ***p < .001.*

For the HR recorded during the VR session (Table 8.17), only the interaction effect between *drive* and *motion* was statistically significant at the .05 significance level. The interaction effect between *drive* and *motion*, $F(1, 48) = 5.805, p = .020, \eta^2 = 0.09$ was significant. Simple main effects analysis showed that there was a trend to significance in HR record between the *motion* conditions in the automated driving ($p = .059$) than the standard driving condition ($p = .536$). The motion condition induced a higher HR in the participants than without motion during the automated driving.

Table 8.17: Three-way ANOVA results using HR as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p	η^2
Drive	6.95	1	6.95	0.065	.800	
Motion	182.93	1	182.93	1.715	.197	
Gender	338.85	1	338.85	3.177	.081	
Drive * Motion	619.17	1	619.17	5.805	.020*	0.09
Drive * Gender	257.97	1	257.97	2.419	.126	
Motion * Gender	0.02	1	0.02	0.000	.990	
Drive * Motion * Gender	74.28	1	74.28	0.696	.408	
Residuals	5119.75	48	106.66			

*Note. *p < .05, **p < .01, ***p < .001.*

For the SCLD (Table 8.18), all effects were not statistically significant at the .01 significance level. The statistical level was reduced due to a significant result of Levene's test and unequal sample size.

Table 8.18: Three-way ANOVA results using SCLD as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Drive	0.12	1	0.12	0.310	.585
Motion	0.16	1	0.16	0.388	.541
Gender	2.93	1	2.93	7.340	.014
Drive * Motion	0.01	1	0.01	0.018	.894
Drive * Gender	0.00	1	0.00	0.011	.918
Motion * Gender	0.11	1	0.11	0.285	.600
Drive * Motion * Gender	0.30	1	0.30	0.760	.395
Residuals	7.19	18	0.40		

*Note. *p < .01, **p < .001.*

For the SCL recorded during the VR sessions (Table 8.19), all effects were not statistically significant at the .01 significance level. The statistical level was reduced due to a significant result of Levene's test and unequal sample size.

Table 8.19: Three-way ANOVA results using SCL as criterion.

	Sum of Squares	*df*	*Mean Square*	F	p
Drive	0.00	1	0.00	0.000	.982
Motion	0.43	1	0.43	0.882	.354
Gender	1.57	1	1.57	3.204	.082
Drive * Motion	0.15	1	0.15	0.300	.587
Drive * Gender	0.02	1	0.02	0.037	.850
Motion * Gender	0.02	1	0.02	0.042	.839
Drive * Motion * Gender	0.00	1	0.00	0.005	.944
Residuals	17.68	36	0.49		

*Note. *p < .01, **p < .001.*

For the RRD (Table 8.20), all effects were not statistically significant at the .05 significance level.

For the RR recorded during the VR sessions (Table 8.21), all the effects were not statistically significant at the .05 significance level.

Table 8.20: Three-way ANOVA results using RRD as criterion.

	Sum of Squares	*df*	*Mean Square*	*F*	*p*
Drive	0.00	1	0.00	1.965	.183
Motion	0.01	1	0.01	4.313	.057
Gender	0.00	1	0.00	2.591	.130
Drive * Motion	0.00	1	0.00	0.055	.819
Drive * Gender	0.00	1	0.00	0.006	.939
Motion * Gender	0.00	1	0.00	0.421	.527
Drive * Motion * Gender	0.00	1	0.00	0.000	1.00
Residuals	0.15	14	0.00		

*Note. *p < .05, **p < .01, ***p < .001.*

Table 8.21: Three-way ANOVA results using RR as criterion.

	Sum of Squares	*df*	*Mean Square*	*F*	*p*
Drive	0.00	1	0.00	1.269	.268
Motion	0.00	1	0.00	0.295	.590
Gender	0.00	1	0.00	1.783	.190
Drive * Motion	0.00	1	0.00	0.453	.505
Drive * Gender	0.00	1	0.00	0.029	.867
Motion * Gender	0.00	1	0.00	0.566	.457
Drive * Motion * Gender	0.00	1	0.00	0.388	.538
Residuals	0.18	35	0.00		

*Note. *p < .05, **p < .01, ***p < .001.*

8.3.2 Correlations

All data using different scales (e.g., physiological data, emotions, personality, MSSQ-short) were standardized for the statistical analysis. A Spearman's correlation showed only significance between SSQ Oculomotor score and *age*, $r_s = -.278, p = .024$ (Table 8.22). Nevertheless, a trend to negative significant relationship between MSSQ total score and *age*, $r_s = -.233, p = .059$, and SSQ Disorientation score and *age*, $r_s = -.215, p = .084$, was found. The same correlation analysis showed a positive significant relationship between SSQ, FSSQ and MSSQ-short score (Table 8.22).

Table 8.22: Correlations between FMS, SSQ, FSSQ, age, and MSSQ-short score.

	Age	n	MSSQ-short	n
FMS Score	-.127	66	.077	66
MSSQ-short Score	-.233	66	-	-
SSQ total score	-.147	66	.346**	66
SSQ Nausea	-.017	66	.382**	66
SSQ Disorientation	.046	66	.269*	66
SSQ Oculomotor	-.278*	66	.306*	66
FSSQ total score	.021	66	.358**	66
FSSQ Nausea	.080	63	.363**	66
FSSQ Disorientation	.097	63	.276*	66
FSSQ Oculomotor	-.089	63	.390**	66

*Note. *p < .05, **p < .01, ***p < .001.*

A Spearman's correlation was conducted between SSQ, FSSQ scores, and *neuroticism* as well as *extroversion*. The results showed a significant relationship between SSQ total score, SSQ Disorientation, SSQ Oculomotor, FSSQ total score, FSSQ Oculomotor, and *neuroticism*, and no significant relationship between the listed variables and *extroversion* (Table 8.23).

A point-biserial correlation was run to determine the relationship between FMS, SSQ, FSSQ scores, and *gender*. There was a positive correlation between SSQ scores, FSSQ scores, and *gender*, which was statistically significant (Fig. 8.2). The same correlation analysis was run on *vision correction*. The results showed no significant relationship between FMS, SSQ, FSSQ scores, and *vision correction* (see Appendix F).

A Spearman's correlation showed no significant relationship between FMS, SSQ, FSSQ scores, and *sleep deprivation* as well as *physical activity* (see Appendix F). Furthermore, there was no significant relationship between FMS, SSQ, FSSQ scores and *valence* before the driving simulation. However, a trend of significant positive relationship between *presence* and *valence* before the driving simulation was found, $r_s = .240, p = .052, N = 66$. No significant correlation

Table 8.23: Correlations between FMS, SSQ, FSSQ, neuroticism, and extroversion.

	Neuroticism	n	Extroversion	n
FMS Score	.201	66	.016	66
SSQ total score	.284*	66	-.066	66
SSQ Nausea	.131	66	-.089	66
SSQ Disorientation	.265*	66	-.053	66
SSQ Oculomotor	.400**	66	-.043	66
FSSQ total score	.265*	59	.052	59
FSSQ Nausea	.237	59	.027	59
FSSQ Disorientation	.219	59	.088	59
FSSQ Oculomotor	.285*	59	.057	59

Note. *$p < .05$, **$p < .01$, ***$p < .001$.

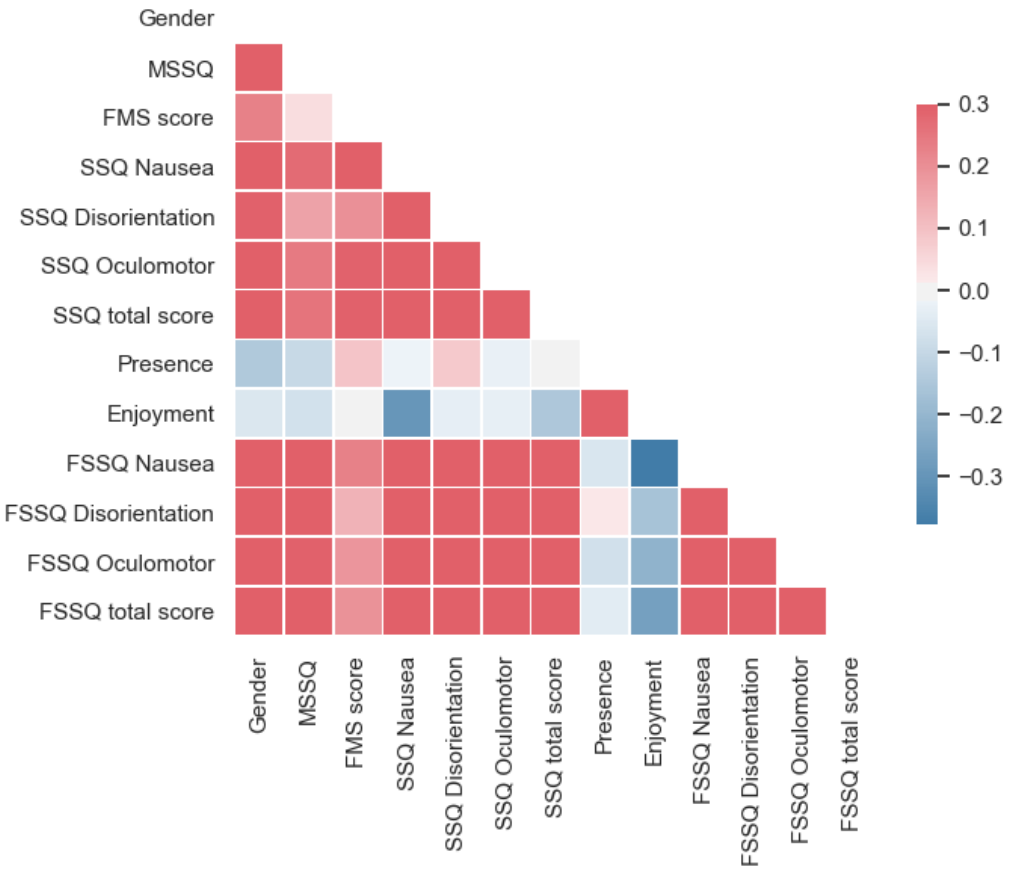


Figure 8.2: A Pearson's correlation matrix showing the relationship between MSSQ-short, FMS, SSQ, presence, enjoyment, FSSQ scores, and gender. For the complete results including the p-values, please see Appendix F.

between FMS, SSQ, FSSQ, and *arousal* before the driving simulation was found (see Appendix F).

A Spearman's correlation showed no significant relationship between the FMS, SSQ, FSSQ scores, and HR recorded during the VR simulation. Nevertheless, a positive relationship between HRD and SSQ total score, $r_s = .274, p = .045, N = 54$; SSQ Nausea, $r_S = .291, p = .033, N = 54$; and SSQ Disorientation, $r_s = .300, p = .028, N = 54$, was found (Table 8.24).

Table 8.24: Correlations between FMS, SSQ, FSSQ scores, HR difference (HRD), and HR record.

	HRD	n	HR record	n
FMS Score	-.099	54	-.074	56
SSQ total score	.274*	54	.259	56
SSQ Nausea	.291*	54	.220	59
SSQ Disorientation	.300*	54	.236	56
SSQ Oculomotor	.153	54	.219	56
FSSQ total score	.117	47	.246	49
FSSQ Nausea	.182	47	.216	49
FSSQ Disorientation	.152	47	.233	49
FSSQ Oculomotor	.047	47	.232	49
Presence	-.185	54	-.075	56
Enjoyment	-.195	49	-.297*	51

*Note. *p < .05, **p < .01, ***p < .001.*

No correlation between SCLD as well as SCL recorded during the VR driving simulation, and sickness score was found (see Appendix F). A Spearman's correlation indicated a negative relationship between RR record and SSQ Nausea score, $r_s = -.349, p = .022, N = 43$. The same correlation analysis showed no relationship between FMS, SSQ, FSSQ scores, and RRD (see Appendix F).

8.4 Discussion

Similarly to the results from the automated VR driving experiment (see Chapter 7.3), none of the hypotheses regarding *motion* were supported by the results, which showed no significant difference across the *motion* conditions regarding SiS onset measured through FMS, SSQ, and FSSQ scores. Contrary to the assumptions, motion cues did not influence the felt discomfort. The SiS reducing effect of physical motion cues in VR driving simulations could not be replicated [10, 48]. A possible reason for this might be the used sample. The findings from the standard VR driving evaluation showed that the dynamic condition evoked less SiS. In contrast, the findings from the automated VR driving showed that the moving platform did not affect SiS.

For this experiment, we selected only the participants who took part in one of the previous two experiments. However, it might be that in the sample from the standard VR driving experiment, most participants, who did not report significant sickness scores, were included. Therefore, the current study's sample size most likely contained participants with low sickness scores, among which no difference could be observed. Surprisingly, no differences were found across the driving conditions regarding SiS. Automated driving did not induce more discomfort than standard driving in the context of HMD driving simulation. This outcome is contrary to that of Dong et al. [59] who found that drivers were less likely to report SiS than passengers but align with that of Curry et al. [47] who revealed no discomfort differences between drivers and passengers. The passenger's passive role during our automated VR driving did not increase the tendency to SiS onset. These findings are likely to be related to the marked driving path. Earlier work suggested that providing adequate visual information to the passengers about the automated vehicle path might reduce motion sickness in real-life driving [56]. It can be assumed that the labeled road provided sufficient visual information to the participants in the automated VR driving condition, and thus, the SiS onset did not differ from the standard VR driving condition. Nonetheless, it should be noted that in the Dong et al. [59] experiment, the participants were exposed to the same driving simulation. In the current study, the driving simulation might slightly differ due to personal driving preferences. The standard driving participants drove the same marked road, but it depended on the individual driving style to take turns or stop at a traffic light. Thus, the participants in both conditions might not be exposed to the same visual and motion cues. Furthermore, a trend was observed that the dynamic condition induced more sickness during the VR simulation in the standard then automated driving condition, which could be explained by the participants' driving preferences.

All hypotheses regarding *gender* were supported, which showed a significant difference between female and male participants. Previous studies reported similar findings [98, 127, 212]. Similarly to the *motion* conditions, a possible reason for these results might be the used sample. It might be that the female participants, who were taken as a sample in the current experiment, were the participants who reported higher sickness scores in the previous two experiments. Moreover, a trend was observed that the standard driving condition induced more sickness in women than the automated driving condition, which could be explained by the participants' driving preferences. Another observation showed that women felt more SiS in the static driving condition than in the dynamic one. A possible reason for that could be the wider FOV [34] which could lead to more significance of the visual input and therefore, to contribute to the higher mismatch between the visual and motion cues in the static driving condition.

We found an average positive correlation between SiS and *gender*. Women tend to feel more discomfort immediately after and around one hour after the VR driving simulation. Similar results we reported from the automated VR driving simulation experiment. The relationship between *gender* and SiS, and more specifically that women experienced higher discomfort than men has been reported by several studies [62, 98, 127, 131, 174, 212]. Stanney et al. [212], for

example, found a linear relationship between SSQ Disorientation, SSQ Oculomotor, and SSQ Total score and *gender*. Our findings are in line with previously reported *gender* differences in VR simulations. However, the precise cause of this relationship is still unclear. It seems possible that these results might be due to the women wider FOV [34], different depth perception in virtual environments [23], or insufficient IPD calibration [213].

As expected, the *type of driving* had an effect on *sense of presence*, the participants in the standard VR driving condition experienced more *presence* than the participants in automated VR driving. This finding is consistent with that of Seay et al. [193] who reported that the drivers experienced a higher level of *presence* than the passengers. A more interactive VR environment, such as the standard driving condition, made the participants feel more present than the automated driving controlled entirely by the vehicle itself.

Regrading the physiological data, only HR recorded during the VR driving simulation showed a significant result on the interaction between the *type of driving* and *motion*. The HR during the automated VR driving was higher in the motion condition than in standard VR driving. It may be that these participants in the automated condition felt more excited, driving the dynamic condition as the moving platform adds a physical motion. Furthermore, a trend of increased HR during the automated VR driving for women was observed. In contrast, male participants had a higher HR during the standard VR driving condition. These results are likely to be related to the participants' driving preferences, and in particular, that women have a different perception about driving than men [75]. Another observation showed that female participants had a higher HR during the dynamic driving condition.

Consistent with the previous two experiments' results, a significant positive correlation was reported between SiS and the difference between the baseline and the recorded HR. When symptoms such as *nausea, stomach awareness, general discomfort, difficulty focusing, blurred vision*, and *dizziness* increase, the HR differentiates significantly from the HR recorded in a normal state of well-being, and vice versa. These results aligned with previous research where participants with severer SiS symptoms showed higher levels of HR [42]. However, Johnson [98] noted that some research findings show that the direction of these physiological changes can individually differ.

Additionally, we found a significant negative relationship between the SSQ Nausea score and RR recorded during the driving simulation. When individuals slow down their breathing, the severity of the SiS symptoms such as *nausea, sweating*, and *stomach awareness* increase, and vice versa. The same correlation was reported in the results from the previous two experiments, the standard VR driving (see Chapter 6.3) and the automated VR driving (see Chapter 7.3). These findings are in line with those of Kim et al. [120] who reported an RR decrease during the VR simulation. As we discussed previously, the observed correlation between the SSQ Nausea symptoms and RR record might be explained in this way: the nervous system arouses because of the felt discomfort in the body, and the breathing slows down as an automatic response. The felt discomfort is

not so severe that the individual decides to stop the VR simulation, but it is severe enough to feel discomfort to which the body responds. Nonetheless, the RR sample size was significantly smaller than the HR sample size, and thus, the results should be interpreted with caution.

A negative relationship was found between *age* and SSQ Oculomotor score. When the *age* increases, SiS symptoms such as *headache, fatigue*, and *eye train* decrease, and vice versa. These findings are contrary to previous studies that have suggested that SiS induced by VR environments increased with increasing the *age* [8, 140]. Nevertheless, the average *age* of the current study participants was around 35 years with a deviation of 12 years, which could not accurately represent an elderly aged group above 50 years.

Not surprisingly, we found a positive correlation was between MSSQ-short total score and sickness scores. The participants who indicated a higher susceptibility to motion sickness most likely experience severer SiS. These findings are consistent with those of previous studies [143, 173, 212] and the two previous driving experiments reported in Chapter 6 and Chapter 7.

Furthermore, a positive, statistically significant relationship was reported between SSQ total score, SSQ Disorientation, SSQ Oculomotor, FSSQ total score, FSSQ Oculomotor, and *neuroticism*. When *neuroticism* score increases, the sickness score increases as well, and vice versa. In accordance with the present results, the previous two experiments (see Chapter 6.3 and Chapter 7.3) and previous studies have demonstrated that participants who scored higher on the *neuroticism* scale experienced an increased level of SiS [44, 158, 232]. For example, SiS symptoms such as *fatigue, headache*, and *eye strain*. A possible explanation could be that increased discomfort is affected by anxiety. Some individuals experienced for the first time an HMD driving simulation, and their excitement prior to the VR driving could increase their anxiety, which is part of neuroticism's personality dimension. Furthermore, our findings showed that the participants who score higher on the *neuroticism* scale felt more prolonged discomfort (e.g., Oculomotor symptoms). Individuals' personality combined with the excitement of using a VR simulation might be a possible factor contributing to experiencing discomfort immediately after and an hour or more later.

8.5 Summary

In this chapter, we have described the assessment of *motion*, *gender*, and *type of driving* regarding SiS induced by an HMD driving simulation. This evaluation was the third and last of the three user evaluations. Data from the two previous experiments was not only used to test whether there are difference across *motion* conditions and *gender* but also across two different *types of driving*, namely, standard and automated driving. We carried out an experiment with 66 participants. The driving scenario was the same as the prior two experiments, a standard urban traffic including other vehicles. Furthermore, we used the collected data for the development of the SiS prediction

models. The most self-reported symptoms after the VR driving session were *general discomfort, fatigue*, and *nausea*.

This assessment revealed that contrary to the expectations, motion cues did not contribute to reducing the SiS symptoms significantly. These results are similar to the results from the automated VR driving experiment. All hypotheses regarding *gender* were supported, which displayed a significant difference between female and male participants. Surprisingly, the automated driving did not induce more discomfort than the standard driving in the context of HMD driving simulation. The passenger's passive role during the automated VR driving did not increase the tendency to SiS onset. Regarding the physiological data, only HR recorded during the VR driving simulation showed a significant result on the interaction between the type of driving and motion. The HR during the automated VR driving was higher in the motion condition than in standard VR driving. Furthermore, we investigated correlations between individual factors and SiS. Several factors such as *gender, age, motion sickness history, neuroticism, HRD*, and *RR* showed statistically significant relationships to SiS. Further, we discussed the results in the setting of relevant literature and previous work.

9 Prediction Models

All models are wrong, but some are useful.

George E. P. Box

The previous three chapters reported the results from the hypotheses testing and correlation analysis carried out on the VR driving simulation experiments. In this chapter, we focus on the prediction of SiS and which models are most suitable. As one of the objectives in this thesis, the prediction of SiS induced by VR driving simulation was made using individual, behavioral, and physiological factors. We have used regression analysis, and machine learning techniques to investigate whether or not the chosen factors can predict the discomfort outbreak. Nevertheless, the individual factors were included only in the linear models as the data for those factors was collected before the VR session. The SiS severity was extracted from the self-reported scores during the VR simulation, immediately after, and one hour later. Although the SiS might differ between these time points, the prediction models could help identify the most critical predictors of discomfort.

The behavioral and physiological factors were for building a few different models to predict whether an individual is a simulation sick or not. We assumed that combining the two groups of factors will contribute to the better prediction of SiS. Thus, some of the models used the two groups of factors together. Furthermore, different machine learning techniques were carried out to compare the performance and accuracy among them. In that way, the models can be compared to the previous prediction models reported in the literature. The chapter concludes with a summary of the reported prediction models.

9.1 Individual Factors

A linear regression analysis was performed in order to explore which human factors can be used as predictors of SiS in VR driving simulation. The predictors were *gender, age, motion sickness history, vision correction, drive frequency, previous experience with driving simulations, previous experience with VR, previous experience with an HMD, video games play frequency, caffeine consumption, sleep deprivation, physical activity, valence*, and *arousal* before the VR session. The outcome variables were FMS, SSQ total score, SSQ Nausea, SSQ Disorientation, SSQ Oculomotor, and FSSQ total score. A few factors, from the all conducted data analysis, showed significant influence in the models as well as the significance of the model itself. Therefore, the following section presents only the regression analysis models that can explain 10% or higher variation. The variables *valence* and *arousal* after the VR driving simulation were excluded from the analysis due to their dependency on experienced discomfort during the simulation.

9.1.1 Simulation Sickness during VR Simulation

From the data derived from Experiment 1, described in **Chapter 6**, a simple linear regression was carried out to predict SiS measured by the FMS score based on the participant's previous experience with an HMD. The R^2 value was 0.171. Therefore, 17.1% of the variation in the FMS score can be explained by the model containing only the *previous HMD experience*. A significant regression equation was found ($F(4, 57) = 2.935, p = .028$). The complete model includes the four possible answers on the question regarding previous experience with an HMD, namely, "totally disagree", "disagree", "neutral", and "agree". Due to the transformation of the independent variable, the last possible answer, "totally agree", was taken as a constant. While having no previous experience with an HMD contributed significantly to the model ($B = 20.600, p = .007$), being neutral ($B = -17.200, p = .224$), having previous experience ($B = .2.685, p = .691$) as well as having no experience at all did not ($B = .535, p = .932$).

$$\begin{aligned} FMS\ score = 21.700 &+ (0.535 * HMD\ Experience \\ &Totally\ Disagree) \\ &+ (20.600 * HMD\ Experience\ Disagree) \\ &+ (-17.200 * HMD\ Experience\ Neutral) \\ &+ (2.685 * HMD\ Experience\ Agree) \end{aligned} \tag{9.1}$$

Another simple regression was carried out to investigate whether *gender* could significantly predict participants' FMS score. The data for this regression analysis was derived from Experiment 2 described in **Chapter 7**. The FMS score indicates the level of SiS during the VR driving simulation. The results of the regression indicated that the model explained 17.1% of the variance and that the model was a significant predictor of SiS, $F(1, 60) = 12.370, p = .001$. The factor

gender contributed significantly to the model ($B = 21.676, p = .001$). The final predictive model is:

$$FMS\ score = 23.171 + (21.676 * Gender) \tag{9.2}$$

9.1.2 Simulation Sickness after VR Simulation

Overall Simulation Sickness

From the data derived from Experiment 1, described in **Chapter 6**, a simple linear regression was carried out to predict SiS measured by the SSQ total score based on the participant's previous experience with an HMD. The SSQ total score was calculated based on the weight system of the questionnaire (see Chapter 2). The R^2 value was 0.198. Therefore 19.8% of the variation in the SSQ total score can be explained by the model containing the *previous HMD experience*. A significant regression equation was found ($F(4, 57) = 3.513, p = .012$). The complete model includes the four possible answers on the question regarding previous experience with an HMD, namely, "totally disagree", "disagree", "neutral", and "agree". Due to the transformation of the independent variable, the last possible answer, "totally agree", was taken as a constant. While having no previous experience, a self-reported disagreement with the statement of the question regarding the HMD experience, contributed significantly to the model ($B = 422.522, p = .002$), being neutral ($B = -309.717, p = .227$), having previous experience ($B = 163.385, p = .185$) as well as having no experience at all did not ($B = 63.277, p = .577$).

$$\begin{aligned} SSQ\ total\ score = 355.907 + (63.277 * HMD\ Experience \\ Totally\ Disagree) + (422.522 * HMD\ Experience\ Disagree) \\ + (-309.717 * HMD\ Experience\ Neutral) \\ + (163.385 * HMD\ Experience\ Agree) \end{aligned} \tag{9.3}$$

Additionally, a linear regression was executed based on the data from Experiment 1, described in **Chapter 6**. The regression analysis was carried out to predict SiS measured by the SSQ total score based on the *motion sickness history*. The R^2 value was 0.136. Therefore, 13.6% of the variation in the SSQ total score can be explained by the model containing MSSQ-short score. A significant regression equation was found ($F(3, 57) = 3.001, p = .038$). None of the factors contributed significantly to the model: motion sickness in childhood ($B = -676.461, p = .064$), motion sickness in adulthood ($B = -659.941, p = .071$), and motion sickness total score ($B = 683.209, p = .062$).

$$\begin{aligned} SSQ\ total\ score = 340.522 + (-676.461 * Childhood\ Motion\ Sickness) \\ + (-659.941 * Adulthood\ Motion\ Sickness) \\ + (683.209 * Total\ score\ Motion\ Sickness) \end{aligned} \tag{9.4}$$

A simple regression was carried out to investigate whether *gender* could significantly predict participants' discomfort measured through the SSQ total score. The data for this regression was derived from Experiment 2 described in **Chapter 7**. SSQ total score indicates the level of SiS immediately after the VR driving simulation. The results of the regression showed that the model explained only 16.3% of the variance and that the model was a significant predictor of SiS, $F(1, 60) = 11.727, p = .001$. The factor *gender* contributed significantly to the model ($B = 29.920, p = .001$). The final predictive model was:

$$SSQ\ total\ score = 29.196 + (29.920 * Gender) \tag{9.5}$$

Another simple regression was carried out to investigate whether *gender* could significantly predicts participants' discomfort measured through the SSQ total score. The data for this regression was derived from Experiment 3 described in **Chapter 8**. The results of the regression analysis indicated that the model explained 21.4% of the variance and that the model was a significant predictor of SiS, $F(1, 64) = 17.43, p < .001$. The factor *gender* contributed significantly to the model ($B = 50.310, p < .001$). The final predictive model is:

$$SSQ\ total\ score = 71.528 + (50.310 * Gender) \tag{9.6}$$

Furthermore, a multiple regression was conducted based on the data from Experiment 3, described in **Chapter 8**. The multiple regression was carried out to investigate whether *arousal* before the VR driving simulation, measured by the SAM, and *gender*, could significantly predict SiS immediately after the simulation measured by the SSQ total score. The results of the regression analysis pointed out that the model explained as much as 23.2% of the variance and that the model was a significant predictor of the SSQ total score, $F(2, 63) = 9.517, p < .001$. The self-reported *arousal* level contributed significantly to the model ($B = 14.954, p = .015$) as well as *gender* ($B = 35.299, p < .001$).

$$SSQ\ total\ score = 11.608 + (14.954 * Arousal) + (35.299 * Gender) \tag{9.7}$$

Nausea Symptoms

From the data derived from Experiment 1, described in **Chapter 6**, a simple linear regression was run to predict SiS measured by the SSQ Nausea score based on the participant's previous experience with an HMD. SSQ Nausea score indicates the level of the symptoms from the Nausea cluster such as *nausea, stomach awareness*, and *sweating* immediately after the VR driving simulation. The R^2 value was 0.258. Therefore, 25.8% of the variation in the Nausea symptoms such as *nausea, stomach awareness*, and *sweating* can be explained by the model containing the *previous HMD experience*. A significant regression equation was found ($F(4, 57) = 4.948, p = .002$). The complete model includes the four possible answers on the question regarding previous

experience with an HMD, namely, "totally disagree", "disagree", "neutral", and "agree". Due to the transformation of the independent variable, the last possible answer, "totally agree", was taken as a constant. While having no previous experience, a self-reported disagreement with the statement of the question regarding the HMD experience, contributed significantly to the model ($B = 54.855, p < .001$), being neutral ($B = -22.419, p = .382$), having previous experience ($B = 12.439, p = .313$) as well as having no experience at all did not ($B = 13.777, p = .228$).

$$\begin{aligned} SSQ\ Nausea\ score = 27.189 &+ (13.777 * HMD\ Experience\ Totally\ Disagree) \\ &+ (54.855 * HMD\ Experience\ Disagree) \\ &+ (-22.419 * HMD\ Experience\ Neutral) \\ &+ (12.439 * HMD\ Experience\ Agree) \end{aligned} \tag{9.8}$$

A simple regression was carried out to investigate whether *gender* could significantly predict participants' discomfort measured through the SSQ Nausea score. The data for this regression was derived from Experiment 2 described in **Chapter 7**. The results of the regression analysis indicated that the model explained only 15.2% of the variance and that the model was a significant predictor of the SiS symptoms such as *nausea, stomach awareness*, and *sweating*, $F(1, 60) = 10.727, p = .002$. The factor *gender* contributed significantly to the model ($B = 31.390, p = .002$). The final predictive model is:

$$SSQ\ Nausea\ score = 24.004 + (31.390 * Gender) \tag{9.9}$$

Another simple regression was carried out to investigate whether *gender* could significantly predict participants' discomfort measured through SSQ Nausea score. The data for this regression was derived from Experiment 3 described in **Chapter 8**. The results of the regression indicated that the model explained as much as 23.8% of the variance and that the model was a significant predictor of the SiS symptoms such as *nausea, stomach awareness*, and *sweating*, $F(1, 64) = 20.039, p < .001$. The factor *gender* contributed significantly to the model ($B = 52.507, p < .001$). The final predictive model is:

$$SSQ\ Nausea\ score = 63.441 + (52.507 * Gender) \tag{9.10}$$

Disorientation Symptoms

A simple linear regression was carried out to predict the score from the Disorientation cluster measured by the SSQ Disorientation score based on the participants' gender. The data for the regression model was derived from Experiment 1, described in **Chapter 6**.

The linear regression was carried out to predict SiS measured by the SSQ Disorientation score based on the *motion sickness history*. The R^2 value was 0.152. Therefore 15.2% of the variation in the Disorientation symptoms such as disorientation *blurred vision, dizziness*, and *vertigo* can be

explained by the model containing MSSQ-short score. A significant regression equation was found ($F(3, 57) = 3.409, p = .023$). While the motion sickness in childhood ($B = -96.392, p = .044$) and the motion sickness total score ($B = 96.391, p = .044$) contributed significantly to the model, the motion sickness in adulthood ($B = -92.266, p = .053$) did not.

$$\begin{aligned} SSQ\ Disorientation\ score = 34.561 &+ (-96.392 * Childhood\ Motion\ Sickness) \\ &+ (-92.266 * Adulthood\ Motion\ Sickness) \\ &+ (96.391 * Total\ score\ Motion\ Sickness) \end{aligned} \tag{9.11}$$

Additionally, a linear regression was executed based on the data from Experiment 2, described in **Chapter 7**. The R^2 value was 0.161. Therefore 16.1% of the variation in the Disorientation symptoms such as disorientation *blurred vision, dizziness*, and *vertigo* can be explained by the model containing only factor *gender*. A significant regression equation was found ($F(1, 60) = 11.529, p = .010$). Participant's discomfort related to symptoms from the Disorientation cluster increased 35.025 units for each score of *gender*.

$$SSQ\ Disorientation\ score = 27.391 + (35.025 * Gender) \tag{9.12}$$

Another simple linear regression was executed to predict the score from the Disorientation cluster measured by the SSQ Disorientation score based on the *gender*. The data for this regression model was derived from Experiment 3, described in **Chapter 8**. The R^2 value was 0.157. Thus, 15.7% of the variation in the Disorientation symptoms such as disorientation *blurred vision, dizziness*, and *vertigo* can be explained by the model containing only factor *gender*. A significant regression equation was found ($F(1, 64) = 11.924, p < .001$). The Gender contributed significantly to the model ($B = 51.076, p = .001$). Participant's discomfort related to symptoms from the Disorientation cluster increased 77.952 units for each score of *gender*.

$$SSQ\ Disorientation\ score = 77.952 + (51.076 * Gender) \tag{9.13}$$

Oculomotor Symptoms

Additionally, a simple linear regression was carried out to predict the score from the Oculomotor cluster measured by the SSQ Oculomotor score based on *gender*. The data for the regression model was derived from Experiment 1, described in **Chapter 6**. The linear regression was carried out to predict SiS measured by the SSQ Oculomotor score based on the *motion sickness history*. The R^2 value was 0.135. Therefore 13.5% of the variation in the Oculomotor symptoms such as *fatigue, headache*, and *eye strain* can be explained by the model containing MSSQ-short scores. A significant regression equation was found ($F(3, 57) = 2.967, p = .039$). None of the factors contributed significantly to the model: motion sickness in childhood ($B = -8.551, p = .714$),

motion sickness in adulthood ($B = -7.359, p = .753$), and motion sickness total score ($B = 9.146, p = .695$).

$$\begin{aligned} SSQ\ Oculomotor\ score = 23.396 &+ (-8.551 * Childhood\ Motion\ Sickness) \\ &+ (-7.359 * Adulthood\ Motion\ Sickness) \\ &+ (9.146 * Total\ score\ Motion\ Sickness) \end{aligned} \tag{9.14}$$

A linear regression was executed based on the data from Experiment 3, described in **Chapter 8**. The R^2 value was 0.166. Thus 16.6% of the variation in the Oculomotor symptoms such as *fatigue, headache*, and *eye strain* can be explained by the model containing only *gender*. A significant regression equation was found ($F(1, 64) = 12.771, p < .001$). The Gender contributed significantly to the model ($B = 32.434, p = .001$). Participant's symptoms from the Oculomotor cluster increased 32.434 units for each score of *gender*.

$$SSQ\ Oculomotor\ score = 52.113 + (32.434 * Gender) \tag{9.15}$$

Another simple regression was carried out to investigate whether personal trait *neuroticism* could significantly predict participants' discomfort measured through SSQ Oculomotor score. The data for this regression was derived from Experiment 3 described in **Chapter 8**. The results of the regression indicated that the model explained only 13.1% of the variance and that the model was a significant predictor of the SiS symptoms such as *fatigue, headache*, and *eye strain*, $F(1, 64) = 9.622, p = .003$. The *neuroticism* contributed significantly to the model ($B = 2.156, p = .003$). The final predictive model is:

$$SSQ\ Oculomotor\ score = 17.726 + (2.156 * Neuroticism) \tag{9.16}$$

9.1.3 Simulation Sickness One Hour after VR Simulation

A simple linear regression was carried out to predict SiS one hour after the simulation measured by FSSQ total score based on *gender*. The data for the regression model was derived from Experiment 1, described in **Chapter 6**. The R^2 value was 0.170. Therefore 17.0% of the variation in the FSSQ total score can be explained by the model containing only factor *gender*. A significant regression equation was found ($F(1, 64) = 13.125, p = .001$). Participant's sickness one hour after the VR simulation decreased with 34.533 units for each score of *gender*.

$$FSSQ\ total\ score = 19.583 + (34.533 * Gender) \tag{9.17}$$

Additionally, a linear regression was executed based on the data from Experiment 1, described in **Chapter 6**. The linear regression was carried out to predict SiS one hour after the simulation measured by the FSSQ total score based on the *motion sickness history*. The R^2 value was 0.158 Thus, 15.8% of the variation in the FSSQ total score can be explained by the model containing

MSSQ-short scores. A significant regression equation was found ($F(3, 53) = 3.311, p = .027$). None of the factors contributed significantly to the model: motion sickness in childhood ($B = 89.680, p = .808$), motion sickness in adulthood ($B = 122.926, p = .739$), and motion sickness total score ($B = -85.977, p = .816$).

$$\begin{aligned} FSSQ\ total\ score = 181.086 + (89.680 * Childhood\ Motion \\ Sickness) \\ + (122.926 * Adulthood\ Motion\ Sickness) \\ + (-85.977 * Total\ score\ Motion\ Sickness) \end{aligned} \tag{9.18}$$

Another linear regression was executed based on the data from Experiment 3, described in **Chapter 8**. The linear regression was carried out to predict SiS one hour after the simulation measured by the FSSQ total score based on the *motion sickness history*. The R^2 value was 0.181. Thus, 18.1% of the variation in the FSSQ total score can be explained by the model containing MSSQ-short scores. A significant regression equation was found ($F(3, 62) = 4.579, p = .006$). None of the factors contributed significantly to the model: motion sickness in childhood ($B = 35.968, p = .494$), motion sickness in adulthood ($B = 41.263, p = .433$), and motion sickness total score ($B = -35.846, p = .496$).

$$\begin{aligned} FSSQ\ total\ score = 43.259 + (35.968 * Childhood\ Motion\ Sickness) \\ + (41.263 * Adulthood\ Motion\ Sickness) \\ + (-35.846 * Total\ score\ Motion\ Sickness) \end{aligned} \tag{9.19}$$

Furthermore, from the data derived from Experiment 3, described in **Chapter 8**, a simple linear regression was carried out to predict SiS one hour after the simulation measured by the FSSQ total score based on the *gender*. The R^2 value was 0.162. Therefore, 16.2% of the variation in the FSSQ total score can be explained by the model containing only *gender*. A significant regression equation was found ($F(1, 64) = 12.40, p < .001$). The factor *gender* contributed significantly to the model ($B = 45.753, p = .001$). Participant's sickness increased 45.763 units for each score of *gender*.

$$FSSQ\ total\ score = 45.763 + (45.753 * Gender) \tag{9.20}$$

Another linear regression was executed based on the data from Experiment 3, described in **Chapter 8**. The linear regression was carried out to predict SiS one hour after the simulation measured by the FSSQ total score based on the participants' *caffeine consumption*. The *caffeine consumption* was divided into three levels based on the answers on the 5-Likert scale. The levels were low, neutral, and high. The R^2 value was 0.120. Thus, 12% of the variation in the FSSQ total score can be explained by the model containing levels of *caffeine consumption*. A significant regression equation was found ($F(3, 62) = 2.828, p = .046$). While the neutral consumption

($B = 37.179, p = .028$) and high consumption ($B = 59.005, p = .014$) contributed significantly to the model, the low consumption ($B = 20.892, p = .238$) did not.

$$\begin{aligned} FSSQ\ total\ score = 40.318 + (20.892 * Low\ Consumption) \\ + 37.179 * Neutral\ Consumption) \\ + (59.005 * High\ Consumption) \end{aligned} \tag{9.21}$$

9.2 Behavioral and Physiological Factors

9.2.1 Dataset

In order to build a model for predicting SiS, datasets were created from the collected data through the first two experiments (**Chapter 6** and **Chapter 7**). In total, three datasets were created, each including 20 features and a class label. There were two class labels: ***Sick*** and ***Not sick***. The data from each participant were manually annotated based on the FMS score. The range of the FMS scale was from 0 to 20. However, above 15 was considered that the participants would experience emesis, and therefore the scale was reduced till 15 for the annotation purposes. The label *Not sick* contained the answers with a score between 0 and 6, and the label *Sick* - the answers with the score between 7 and 15. The FMS score was reported every three minutes, and during this time window, every data point was labeled with one of the labels.

However, these data points were not enough to build a machine learning model, and therefore, a resampling method was used for creating more data. The data extraction was based on a time window of one minute and an overlapping window with the previous data point of one minute. That way, from one participant's session, 24 data samples instead of eight was extracted. Nevertheless, not all participants finished the full session, and therefore, their data was less than 24 samples.

Due to some technical problems, an interruption of the signal between the sensors and the computer occurred at some moments. A pre-processing data method was used to check the raw data for zeros and replace them with the previous value in order not to disrupt the timeline of the data. Furthermore, all data samples with features which were containing an out of normal range data were removed. For example, one of the SCL means was over $15muS$, which is obviously out of the SCL range. Therefore, this sample was removed. By removing these outlying data, the bias towards outlier points was minimized as well.

Three datasets were created, namely, standard driving dataset (SD), automated driving dataset (AD), and all dataset (ALL) combining the sample data from the standard and the automated driving datasets. The SD dataset included 904 data samples from 52 participants, and the AD dataset included 968 data samples from 48 participants. ALL dataset included 1872 data samples from the combined previous two datasets.

The total number of features of each dataset was 20, from which seven were extracted from ECG data, eight from EDA and RR data, and five from virtual world data. For the features extraction from the raw physiological signals, the Python library Biosppy[1] was used. However, the library was used only to extract the SCL, RR, HR, and R peaks from the raw signals. Before extracting the ECG signal features, a Finite Impulse Response with bandpass frequency between 3 and 45 Hz was applied to filter the signal. Furthermore, R peaks were calculated using the Hamilton segmenter [83]. A Butterworth filter with a lowpass band of 5 Hz was applied to the EDA raw signal. Additionally, the signal was smoothed using a Boxzen kernel with a kernel size of 75% of the sampling rate. For the RR raw signal, a Butterworth filter with a bandpass frequency between 0.10 and 0.35 Hz was utilized. Furthermore, in order to smooth the signal, a Boxcar kernel with kernel size 3 was applied.

The following features were extracted from the ECG signal:

Time-domain features

Time-domain indices of HRV quantify the amount of variability in measurements of the interbeat interval, which is the time period between successive heartbeats.

- **HR mean** - the mean value of the HR
- **RMSSD** - the root mean square of successive differences. The square root of the mean of the squares of the successive differences between adjacent beat to beat or RR. RR is the interval between successive R peaks. Instead RR, the term "NN" is used to highlight that the processed beats are "normal" beats.
- **SDNN** - the standard deviation of NN intervals
- **NN50** - the number of pairs of successive NNs that differ by more than 50 ms

Frequency-domain features

The frequency aspect of the HRV is measured by counting the number of NN intervals that match different frequency bands.

- **Low frequency (LF)** - band from 0.04 to 0.15 Hz
- **High frequency (HF)** - band from 0.15 to 0.40 Hz
- **LF/HF ratio** - the ratio between LF to HF power measured in percentage

[1] https://github.com/PIA-Group/BioSPPy

The following features were extracted from the EDA signal:

- **SCL Mean** - the mean value of the SCL
- **SCL Max** - the maximum value of the SCL
- **SCL Min** - the minimum value of the SCL
- **SCL SD** - the standard deviation of the SCL

The following features were extracted from the RR signal:

- **RR Mean** - the mean value of the RR
- **RR Max** - the maximum value of the RR
- **RR Min** - the minimum value of the RR
- **RR SD** - the standard deviation of the RR

The following features were extracted from the virtual world:

- **Speed** - it was calculated based on the data from the virtual vehicle's location and converted to km per hour
- **Acceleration** - it was calculated from the speed and measured in meters over squared seconds
- **Pitch**, **Yaw**, and **Roll** - they were calculated from the position of the HMD

All datasets had a problem with unbalanced classes between *Not sick* and *Sick* classes. The *Not sick* class was almost four times bigger than the *Sick* class. To correct this problem, we performed an oversampling of the undersampled class using the Python library Imbalanced-learn [133]. The library uses a Synthetic Minority Over-sampling Technique (SMOTE) [38] to do the oversampling. The SMOTE algorithm crates "synthetic" data samples to oversample the minority class. For each minority class sample, a "synthetic" sample is calculated along the line segments joining any or all of the minority class K-nearest neighbors where the number of neighbors is randomly chosen. Using the SMOTE algorithm, the balance between the classes was returned.

Furthermore, another method, called under-sampling, for reducing the imbalance between the classes was applied. The method was performed on the datasets using the same Python library as for the over-sampling [133]. The under-sampling method uses an algorithms which selects samples from the original dataset S. Thus, S' is defined such as $|S'| < |S|$ and $S' \not\subset S$. The

class *RandomUnderSampler* with *RandomState* equal to zero was used for performing a random under-sampling in the three datasets.

In the thesis, two different cross-validation methods were used, 10 K-fold and Leave-one-out cross-validation.

9.2.2 Results

Seven different classification algorithms were used for solving a classification problem of who is simulation sick and who is not. For performing all machine learning algorithms and their optimizations, the Python library called SciKitLearn was used [162]. The algorithms were Logistic Regression, LDA, KNN, Decision Trees (CART), NB, SVM, and RFC. The data from the standard (SD) and automated (AD) experiments were evaluated separately as well as together dataset (ALL). For each dataset, the two techniques for fighting the unbalance in the dataset (over- and under-sampling) were used and separately evaluated. For each model, the three different datasets (AD, SD, and ALL) were used, each evaluated with both cross-validation methods (K-fold and LOO) described earlier in the chapter. The complete table of results can be seen in Appendix H.

The models which showed the highest unweighted accuracy were using the RFC, CART, LDA, and KNN methods (see Table 9.1). The RFC gave better results regarding the prediction of the simulation sick participants compare to the other algorithms. Furthermore, the over-sampling technique showed a bit better results compared to the other technique of under-sampling. As we can see from the tables below, the most promising results were delivered from the SD.

Table 9.1: Overview of the best results measured through unweighted accuracy using over- and under-sampling algorithms.

Dataset	Classifier	Under-sampling	Over-sampling
AD dataset	RFC	0.68	0.68
	CART	0.66	0.61
	LDA	0.62	0.61
SD dataset	RFC	0.72	0.81
	SVM	0.72	-
	KNN	-	0.71
	LDA	0.74	0.72
ALL dataset	RFC	0.65	0.68
	CART	0.60	0.62
	KNN	0.62	0.68

Note. RFC - Random Forest Classifier, LDA - Linear Discriminant Analysis, KNN - K-nearest neighbors classifier, CART - Decision Trees Classifier.

Furthermore, an AUC score was calculated to compare the skills between the models. The graphs of the ROC curves, including the AUC scores of the evaluated models for each dataset can be found in Appendix H. The presented graphs were derived from datasets using the under-sampling technique. Nevertheless, the results were very similar to the over-sampling technique. The highest AUC scores were calculated for the RFC, CART, KNN, and LDA algorithms. That supports the results from Table 9.1.

However, the results are not reliable predictors of SiS. Most of the models showed results that were close to a random choice. In order to increase the accuracy of the model, hyperparameter optimization was applied to find the most suitable set of parameters. For this purpose, two search algorithms were deployed, namely Grid search and Randomized search. Grid search builds for every combination of hyperparameters a different model, which is then evaluated. The hyperparameters that reach the best evaluation score are considered the best set of hyperparameters for the tested model. Randomized search combines the hyperparameters randomly and finds the best solution. Like the Grid search, the model's hyperparameters with the best evaluation score are considered as the most suitable set of hyperparameters.

The hyperparameters of the RFC model which were optimized were *max_depth* and *max_features*, which sets the maximum depth of the tree, and sets the number of features to consider when looking for the best split, respectively. For the CART model, the same hyperparameters were optimized. For the KNN model, the tuned hyperparameters were *n_neighbors, weights*, and *algorithm*, which sets the number of neighbors to use by the queries, sets the weight function used in prediction and sets the algorithm used to compute the nearest neighbors, respectively. For the LDA model, the optimized hyperparameters were *solver* and *tol*, which sets which solver will be used, and sets the threshold used for rank estimation in singular value decomposition solver, respectively.

Both search algorithms were used, and the model's setting, which returned better results were reported. Table 9.2 presents the results from the optimization of the hyperparameters over the different datasets. For the RFC model, the hyperparameters were the following: *max_depth* = 100 and *max_features* = 5. The same hyperparameters setting was used for All datasets. For the CART model, the hyperparameters were: *max_depth* = 50, and *max_features* = 15 for the AD and the All datasets. For the LDA model, the hyperparameters were: *solver* = *'svd'*, and *tol* = 0.0001 for the AD and the SD datasets. These were also the setting of the model by default. For the KNN model, the hyperparameters were: *n_neighbors* = 15, *weights* = *'distance'*, and *algorithm* = *'auto'* for the SD dataset, and *n_neighbors* = 5, *weights* = *'distance'*, and *algorithm* = *'auto'* for the AD dataset.

The best accuracy (0.87) was the result of the RFC model trained and tested on the SD over-sampled dataset. Therefore, we applied the built-in function *feature_importances* from the same machine learning library, ScikitLearn, mentioned earlier in the chapter [162], to understand better which features contributed the most to the model. The function calculates the feature importance

Table 9.2: Overview of the results of the optimized models including all 20 features.

Dataset	Classifier	Accuracy	Class	Precision	Recall	F1 score
AD Dataset	RFC	0.82	Not sick	0.87	0.88	0.87
			Sick	0.63	0.50	0.56
	CART	0.70	Not sick	0.85	0.78	0.81
			Sick	0.41	0.52	0.46
	LDA	0.62*	Not sick	0.86	0.61	0.71
			Sick	0.33	0.66	0.44
SD Dataset	RFC	0.87	Not sick	0.92	0.90	0.91
			Sick	0.65	0.69	0.67
	CART	0.74	Not sick	0.93	0.68	0.79
			Sick	0.37	0.77	0.50
	LDA	0.74*	Not sick	0.89	0.77	0.83
			Sick	0.39	0,60	0.47
ALL Dataset	RFC	0.82	Not sick	0.88	0.87	0.88
			Sick	0.56	0.53	0.54
	CART	0.72	Not sick	0.87	0.76	0.81
			Sick	0.37	0.55	0.44
	KNN	0.67	Not sick	0.87	0.68	0.76
			Sick	0.33	0.64	0.44

*Note. * Used the undersampled dataset.*

RFC - Random Forest Classifier, LDA - Linear Discriminant Analysis, KNN - K-nearest neighbors classifier, CART - Decision Trees Classifier.

based on the node impurity, also known as *Mean Decrease of Impurity* or *Gini importance* [27]. To calculate the *Gini importance*, the node impurity total decrease is normalized and then averaged over all trees in the model. For weighting the node impurity, a probability of reaching the node is estimated by the number of samples which reach that node.

The top five features are the LF/HF ratio, the yaw movement of the head, the LF, the RR Minimum, and the SCL mean. The features are a combination of the physiological and behavioral data, but the predominant features are from the ECG domain. A model with only those five features, trained and tested on the SD dataset, showed an accuracy of 0.82. The results for the *Sick* class were the following, *Precision* - 0.62, *Recall* - 0.71, and *F1 score* - 0.67. These findings pointed out that with four times fewer features, the accuracy of the model stayed almost the same. However, with different iterations of the model, the fifth feature has changed, but the other four stayed the same.

Additionally, the same machine learning algorithm was used with only physiological features, which were in total 15 features (Table 9.3). It can be noticed that the accuracy of the model did not differ significantly from the model with all features. Nevertheless, the best performance was using the SD dataset. The feature reducing algorithm was further applied on the SD dataset and the results indicated an accuracy of 0.83. The results for the *Sick* class were *Precision* - 0.56, *Recall* - 0.69, and *F1 score* - 0.62. The features were reduced to seven, and there were the following ordered by significance: *LF/HF ratio, HR mean, LF, SDNN, RMSSD, NN50*, and *SCL mean*.

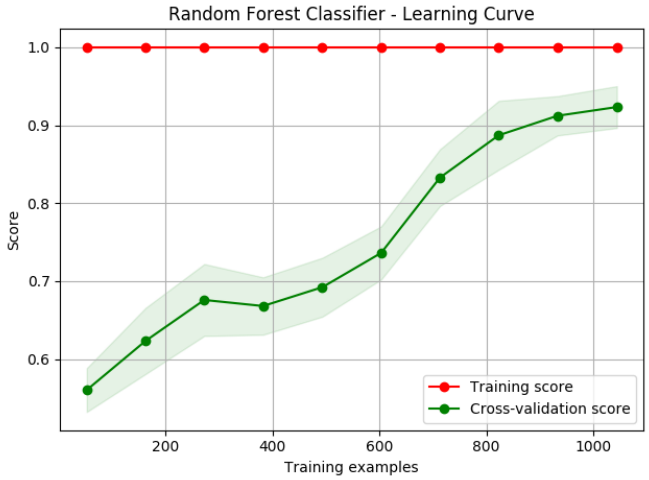
Most of the features were from the ECG data, and therefore, another model was build using the SD dataset and only the features from the same physiological signal, namely the ECG signal. The results showed that the model accuracy was the same (0.83) as the previous featured reduced model. The *Precision* was 0.55, the *Recall* was 0.60, and the *F1 score* was 0.58. These findings pointed out that the model with only ECG features performed almost as good as the model, which included an SCL feature. However, this model did not identify so well the relevant instances of SiS. Using the SCL mean as a feature contributes to better identification of the SiS. Nevertheless, the results showed that the classification of SiS in a VR driving simulation could be done using only physiological data, and in particular, using only ECG data.

Furthermore, to complete the prediction model evaluation, a learning curve of the model, including all features and including only the features extracted from the ECG signal were plotted. With increasing the training examples, the first model (Fig. 9.1 (a)) pointed to a better score above 0.90 level. However, the second model (Fig. 9.1 (b)) showed also a good score almost 0.90. That indicates that the second model is similar in performance as the first model, but it has far fewer features. From the graphs, it can be observed that both models predicted perfectly while using the training set but not so well when they were tested on the validation set. Therefore, the models were overfitting. The graphs also show that with the growing number of training samples, the models learn fast and predict better. Therefore, a bigger dataset with more samples could overcome the problem of overfitting.

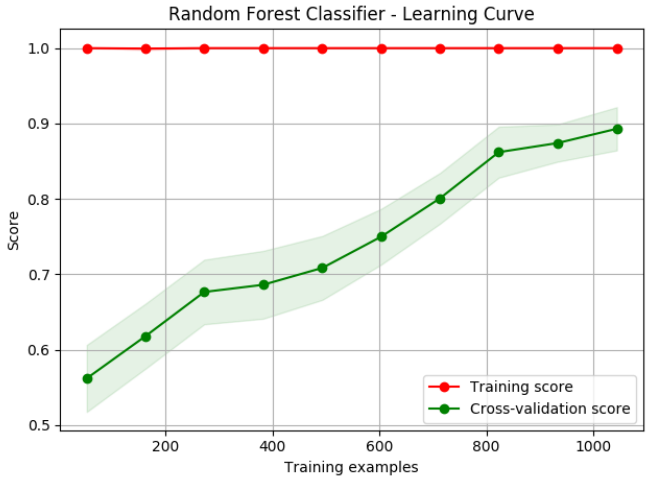
Table 9.3: Overview of the optimized models' results, including only 15 physiological features.

Dataset	Classifier	Accuracy	Class	Precision	Recall	F1 score
AD Dataset	RFC	0.81	Not sick	0.86	0.91	0.88
			Sick	0.61	0.50	0.55
SD Dataset	RFC	0.85	Not sick	0.92	0.88	0.90
			Sick	0.59	0.69	0.63
ALL Dataset	RFC	0.80	Not sick	0.88	0.87	0.88
			Sick	0.51	0.53	0.52

RFC - Random Forest Classifier.



(a)



(b)

Figure 9.1: Learning curves of the model including all 20 features (a) and the model including 7 ECG features (b), showing the accuracy of the training and the validation sets. Both models are trained on the SD dataset.

9.3 Discussion

The most prominent single factor for predicting SiS onset was *gender*. It was a part of each model, sometimes in two models, for each outcome. However, except in the second model for SSQ total score and for the SSQ Nausea, the single *gender* factor models showed that not more than 17.1% of the variation in the sickness outbreak could be explained. *Gender* can explain 21.4% and 23.8% of the variation in the SSQ total score and SSQ Nausea models, respectively. In previous research, *gender* was one of the essential factors in the model suggested by Kolasinski [126]. The model included as well factors such as *age, mental rotation ability*, and *pre-exposure value*, and it could explain 34% of the variance.

Not surprisingly, one of the predictors, which showed significant results, was *motion sickness history*. Contrary to the previously suggested model by Rebenitsch [175], the *motion sickness history* was not able to explain more than 19% of the variation in the current SiS models. Nevertheless, that result is close to the reported results of a model build on individual factors [173]. The findings could be explained with the low level of previously experienced motion sickness by the participants. A screening method can be used to choose only individuals with high sensitivity to motion sickness for future studies. However, this type of screening could raise some ethical questions in the context of HCI research in automotive industry.

Another factor that showed significant results was the *previous experience with HMDs*. This factor combined the answers from a question regarding the *HMD experience*. The most significant answer which might influence the sickness outbreak was the lack of previous experience. However, the top results were that 25.8% of the variation in the SSQ Nausea model could be explained by containing the *HMD experience* as a single predictor. The results can be connected with one of the SiS related factors - *adaptation* (see **Chapter 2**. Without previous experience with the VR world, the individual is exposed to a novice environment. *Adaptation* is one of the well-known and used solution against SiS [88, 145, 178]. Therefore, there is no surprise that this factor can explain a good percentage of variance in the prediction model.

Surprisingly, personality traits *arousal* before the VR driving simulation and *neuroticism* showed promising results. In combination, these two factors can explain 11.5% of the variation in the model regarding SiS during VR simulation. These findings support previous research that related higher *arousal* [81] with SiS in virtual environments and *neuroticism* with motion sickness [44]. Based on the similarities shared between motion sickness and SiS, it can be assumed that *neuroticism* is likely associated with SiS onset. Nevertheless, the *arousal* predictor, in combination with *gender*, can explain as much as 25.2% of the variation in the SSQ total score model. A possible explanation for that could be that women were more excited to try the VR driving simulation than men. Alone the predictor *neuroticism* can explain 13.1% of the variation in the SSQ Oculomotor model. All these findings show that personality traits could be a possible SiS predictor, and further, more focused studies are needed.

Furthermore, *caffeine consumption* on the day of the VR session showed small but significant results. This factor was extracted from a question regarding whether the participants consumed in the last 4 hours a caffeine drink. The results were divided into three groups based on the answers: low, neutral, and high. Then each group represented a different variable in the model. The factor *caffeine consumption* can explain 12% of the variation in the FSSQ total score model. The results point to that *caffeine consumption* is a possible important factor for recovering after the VR session. Therefore, individuals who planned to use HMDs in the next four hours are advised to stop consuming high caffeine drinks.

From all evaluated machine learning models, the best results for predicting SiS was delivered by an RFC algorithm model (87%) trained on the SD dataset. The model included physiological and behavioral features. With the AD dataset, the same model showed an accuracy of 0.82, which was the same as the results with ALL dataset. However, the more important measure was the *F1 score*. For the *Sick* class, in the RFC model with the standard dataset, the score was 0.67, and for the *Not sick* class, the score was 0.91. The *precision* (0.65) and the *recall* was 0.69, which was also better than the other models. Another model was build using only features extracted from the physiological data. The model used the RFC algorithm as the most prominent one from the previous model comparison. The results showed an accuracy of 0.85 using the SD dataset. Furthermore, we have applied a feature reducing algorithm to reduce the features to seven in number. The accuracy of the reduced features model was 0.83.

Another model using the same RFC algorithm was built with only features from the ECG signal. The accuracy of the model with only ECG features was very similar to the model with all features (83%). These findings point out the possibility of using only one physiological signal to predict the discomfort outbreak. A comparison with the earlier reported models (Chapter 3.4.3) is not applicable because each model had different performance metrics based on the model's task (classification or regression). For example, Hell and Argyriou [86] reported the *Mean Error*, Kim et al. [118] reported the *Pearson correlation coefficient*, and Martin et al. [142] reported the *coefficient of determination*. A common metric to measure the model's performance should be used to carry out an appropriate comparison across the prediction models. Nonetheless, similarities with the previous works exist based on the applications used for collecting the data for building the datasets. Martin et al. [142], and Jin et al. [96] utilized similar VR applications based on their interactivity and environmental elements; all VR applications were VR games. Furthermore, these two studies utilized the same evaluation metrics, and thus, they can be directly compared to each other based on the models' performance.

A note of caution is due here since the *Sick* class's performance results were close to random choice (*F1 score* - 0.67). However, the model predicts with very high certainty (0.91) which individual is not simulation sick. One of the weaknesses of the model is that it does not return the relevant instances correctly. A possible explanation might be the low level of the severity of SiS during the VR driving simulation. Also, an additional contribution to that could come from the unequal distribution of data between the classes. Despite the application of the over-sampling

algorithm, the data in the *Sick* class was not sufficiently realistic. Another weakness is that the proposed model uses the RFC algorithm, which is a sophisticated algorithm and sometimes could run slower compared to other algorithms. For the current thesis, that was not an issue due to the size of the used dataset. The performance speed of the model could be of high importance if the model is used for real-time recognition applications.

Interestingly, the top five features in the all features model were derived from the ECG, respiration signal, and the yaw movement of the head. The movement of the head can be easily related to driving, as it is probably one of the most basic head movements while driving. The driver should turn his head very often, for example, to check whether the road ahead on a crossroad is clear to proceed or not. Therefore, it is not a surprise that participants often turn their heads in the virtual drive when they have control over the vehicle as in the standard VR driving simulation. The surprising finding is that the head movement contributes significantly to the SiS classification model. Furthermore, the top four features from the ECG features model were *LF/HF ratio, HR mean, LF, SDNN, RMSSD*, and *NN50*.

9.4 Summary

We have described the development of the SiS prediction models. Two different approaches were presented. The first approach included the prediction of SiS based on individual factors collected through the questionnaires. The second approach used machine learning methods to predict SiS based on behavioral and physiological data collected through the VR driving and the physiological sensors. Overall, we have developed 11 simple, 11 multiple linear, and 13 machine learning models. The most prominent individual factor was *gender*, and it took part in half the linear models. The individuals' gender can explain 21.4% and 23.8% of the SSQ total score and SSQ Nausea models' variation, respectively. The best linear model predicted Nausea symptoms, using the previous HMD experience as a predictor. The lack of previous experience might influence the sickness symptoms outbreak such as *nausea stomach awareness* significantly, and *sweating*. Moreover, the findings showed that the classification of SiS in a modern HMD driving simulation could be done using physiological data. More notably, that model was built with only features extracted from the ECG signal. However, the model built on the RFC method, including behavioral and physiological features, performed slightly better, and reached an accuracy of 87%. Further, we discussed the developed model with reference to prior work.

10 Discussion

> They're a million worlds. All different...All similar. Constants and variables.
>
> *Elisabeth*, Bioshock: Infinite

This thesis aimed to assess and predict the phenomenon of SiS in the context of an automated VR driving simulation utilizing an innovative HMD driving setup. The setup was developed explicitly throughout this thesis for the evaluation of SiS. Furthermore, a comparison across two different driving types, automated and standard, was carried out using static and dynamic VR driving simulations. This chapter includes a discussion on major findings as associated with the literature on SiS induced by VR driving applications and relationships between individual factors and SiS. It also considers the theoretical and practical implications of the conducted experiments. The chapter concludes by acknowledging the limitations of the conducted experiments, which helps to give a more balanced impression of the reported results.

The objectives of the current thesis were as follows: to integrate a moving platform and physiological signals into an innovative HMD driving system; to evaluate SiS factors in a standard and automated dynamic VR driving setup; to explore individual factors; and to predict SiS using individual, behavioral, and physiological factors. In order to meet the objectives, we have conducted three experiments that tested the effects of a moving platform and two types of driving. One unanticipated finding was that the experiments revealed inconsistent results regarding the effect of *motion* and *gender* on SiS. Two out of the three experiments showed that motion does not affect SiS induced by a VR driving simulation. The same two experiments also showed that *gender* does affect SiS. Therefore, caution should be applied when generalizing the findings of all experiments. In the current thesis, we determine that the more frequently reported findings will be considered as generalized findings for the overall discussion and conclusion. For a detailed

disputation of each experiment's results, the reader can refer to the Discussion section of **Chapter 6**, **Chapter 7**, and **Chapter 8**.

Although the severity of SiS was reported at only a low level, its appearance, which sometimes resulted in participants discontinuing the virtual drive, is nevertheless an essential indication that the SiS should not be overlooked. The most frequently experienced symptoms across all experiments were from the Disorientation cluster. This result aligned with previous studies [208] where the VR systems showed more Disorientation than Nausea and Oculomotor symptoms. This suggests that individuals exposed to VR driving simulations are more likely to experience *difficulty focusing, blurred vision, dizziness*, and *vertigo*, than symptoms of discomfort such as *fatigue, headache, eye strain*, and *difficulty concentrating*. The profile of SiS follows the scheme D > N > O, where D is Disorientation, N is Nausea, and O is Oculomotor.

10.1 General Discussion

In order to assess SiS induced by a VR driving simulation, we developed a VR driving setup that could be used later as an HCI evaluation tool of automobile interior design concepts. Thus, there were requirements for the setup that had to be fulfilled. Some of the requirements were hardware related, such as the size of the setup and its mobility. Others were related to the software, such as the implementation compatibility with the already used 3D software. The successful integration of the moving platform and the physiological sensors into a modern HMD driving setup fulfilled one of the thesis objectives. The VR driving setup was utilized during all user experiments described in this thesis. Prior works operationalized similar driving systems; however, none of them combined a modern HMD viewing system with a dynamic 3 DOF moving platform and physiological sensors, including a realistic virtual urban environment [63, 185, 216, 219, 225]. Continuously recording synchronized physiological signals can contribute to the enhancement of SiS assessment during a virtual drive. The results of the objective measures can be considered in conjunction with the results of the self-reported measures to form a more comprehensive overall evaluation. Therefore, the developed HMD driving system will aid the HCI research in the context of automotive interior development.

Evaluation of SiS factors in standard and automated dynamic VR driving setup

In the standard driving experiment, the SiS was more severe in the static condition than in the dynamic condition. These results are in line with previous studies that show that physical motion cues have an impact on the induced discomfort [10, 48]. Although the moving platform had some movement constraints due to the DOF, the produced motion feedback was sufficient for the participants to feel better than without any motion feedback.

A moving platform as an additional layer to the VR driving setup contributes to a more realistic driving experience. It should be noted that this might led to more cautious driving than in the static condition, where the driving had no physical consciences. For example, one participant stopped at the last minute at a traffic light. In the static condition, nothing would happen, but in the dynamic condition, the participant would feel this sudden stop and the move of the platform on the x-axis (front-back axis), which could make her or him nauseous. Thus, the participants in the dynamic condition might have driven more carefully than the participants in the static condition. As a result, the SiS scores reported in the dynamic condition were lower than those in the static condition. Furthermore, the *without motion* group participants had a slightly higher motion sickness susceptibility based on their self-reported motion sickness history. It is assumed that individuals with higher susceptibility to motion sickness are more prone to experience SiS than individuals with less susceptibility. Therefore, in our experiment's results, the participants in the static condition might experience more discomfort because of their slightly increased susceptibility to motion sickness than those in the dynamic condition. Nonetheless, the MSSQ-short score reported across the groups was not sufficient to substantially impact the SiS onset.

In the VR automated driving experiment the moving platform had no significant effect on SiS. The participants experienced almost the same discomfort in both static and dynamic conditions. However, we found that SiS symptoms were slightly higher in the static condition than in the dynamic condition. The added motion showed a very slight trend towards reducing the induced discomfort, but it was not strong enough to be significant. A possible reason for this might be the low vehicle speed during the virtual ride. The average speed was 30 km/h, which could be too low to produce sufficiently strong physical motion feedback in order to reduce SiS.

Interestingly, the standard and the automated VR driving experiments reported controversial results regarding the moving platform. The only difference between these two experiments was the control over the virtual vehicle. In the direct comparison between the two driving conditions in the third experiment, no significant difference between the condition was found as well as no interaction between motion and type of drive on SiS. This inconsistency may be due to the used sample. As we discussed in **Chapter 8.4**, the group of participants chosen for the last experiment might be those who did not report high discomfort in the previous two experiments. Thus, the results were not sufficiently different to show significance between the conditions.

The level of sickness in two out of three experiments did not differ across the *motion* conditions. Therefore, overall, the addition of motion cues did not influence SiS induced by a modern HMD driving simulation. These findings are consistent with those obtained from previous studies, which have suggested that a dynamic driving simulation did not evoke less discomfort than a static driving simulation [108, 123]. Nevertheless, it should be noted that the evoked SiS was generally in the low range of symptoms severity. This result may be explained by the fact that the vehicle speed was low during both virtual drives. As discussed earlier, the low speed might be insufficient to give the required physical motion feedback in order to reduce SiS. Nevertheless,

the VR drive was an urban environment simulation that involved numerous stops at the traffic lights, which could have led to the overall reduced vehicle speed.

According to the cue conflict theory, a mismatch could arise between the visual and vestibular sensory systems, which could incite a perceived increase of discomfort as a response from the body [172]. Thus, static driving simulators might induce a higher mismatch than dynamic driving simulators, which can result in SiS outbreak [215]. However, our findings did not provide any confirmation of this postulation. A plausible explanation might be that the *motion* evoked by the moving platform was mild, and the participants did not actually feel a substantial difference between the conditions. The simulated acceleration might be perceived as a rotating movement around the center of the platform instead of the lateral axis due to the motion platform limitations. Ropelato et al. [185] stressed that the illusion of linear acceleration could be interrupted by the inadequate response of the moving platform during sudden and swift vehicle deceleration. Instead of a change in velocity, an individual can perceive this movement as a rotation around the platform's center. Platforms with 6 DOF can overcome that, maintaining this rotation by an existent linear movement along the corresponding axis. However, the VR driving system developed in this thesis utilized a 3 DOF moving platform, and thus the linear movements on the lateral and longitudinal axes were not supported. Additionally, the motion condition induced a higher SCL than the static condition. This finding illustrates that participants felt more stressed during the dynamic VR driving simulation. A possible reason for this might be the exposure to a novel and highly immersive virtual environment that includes physical motion feedback. Another reason might be the physical limitations of the moving platform. As mentioned above, a platform with 3 DOF has more restricted movements than one with 6 DOF, and thus, some of the vehicle movements might not be correctly replicated.

Furthermore, participants' level of discomfort was the same across the standard and automated driving conditions. The standard driving did not induce less discomfort than the automated driving in the context of HMD driving simulation. This outcome is inconsistent with that of Dong et al. [59], who found that drivers were less likely to report SiS than passengers, but it supports that of Curry et al. [47], who revealed no discomfort differences between drivers and passengers. The passenger's passive role during our automated VR driving did not increase the tendency to experience discomfort. Nonetheless, fewer participants discontinued the automated VR driving simulation. This suggests that the motion cues in the experiments were felt differently. This result contradicts the suggestion that the addition of a moving platform could have a more significant effect on SiS when the vehicle's maneuvers are intentionally limited to current movements similar to the automated driving, which the participants can anticipate [108]. A possible explanation might be that the standard driving experiment required every participant to drive, and thus participants' driving preferences could influence the VR driving experience. In contrast, during the automated driving experiment, the motion was the same for every participant and did not depend on the participants' driving preferences. Together these findings suggest that the addition of motion cues does not affect SiS onset per se. These results reflect those of Klüver et al. [123],

who concluded that a moving platform does not necessarily decrease the incidence of SiS, but it is more important how well a motion system simulates the motion cues adequately does have an impact. Therefore, the physical limits of the moving platform utilized in the current VR setup might be the cause of the insignificant influence of the motion cues on SiS within the HMD driving simulation.

Overall, the findings of the current thesis revealed that female participants experienced more SiS than male participants in the context of modern HMD driving simulations. These results match those observed in earlier studies that women are more susceptible to SiS than men [62, 72, 127, 143, 212]. Interestingly, the male participants reported more discomfort in the standard VR driving than in the automated driving. The SiS incidence was still higher for the female participants in both VR simulations, but due to the fluctuation of the male sickness scores, the automated VR driving showed significant difference across genders and the standard VR driving did not. It should be noted that women reported more motion sickness susceptibility than men across all experiments. Individuals who have a motion sickness history are more prone to SiS onset during driving simulations [143]. Thus, the effect of *gender* might be influenced by the motion sickness history aspect of the experiment's sample. Nonetheless, these findings should be interpreted with caution due to potential sample bias.

The physiological records revealed that women had a higher HR and skin response than men during the VR driving across the experiments than men. These results are in line with those obtained from the questionnaires. The higher HR might result from increased excitement about driving a virtual vehicle in a novel environment or result from the increased severity of the SiS symptoms. A higher SCL is related to a strong body response to stress, and sweating is one of the SiS symptoms. These observations match those from earlier studies that showed that SiS is significantly positively correlated with physiological signals such as heart signal [51, 120, 194] and sweating [51, 226], and breathing [120].

Moreover, prolonged VR driving could be an additional contributor to the elevated HR [42]. While these studies investigated the bodily response to SiS, they did not explicitly evaluate this response across genders. The combination of the increased pulse and sweating indicates that the female participants indeed experienced more discomfort than the male participants. Nevertheless, it should be noted that our findings displayed only a tendency and not a statistical significance of differences between the physiological signals across genders.

We observed a trend that the female participants suffered more during the standard VR driving condition than during the automated condition. Women might find the automated VR driving simulation more pleasant because of the visual control flow during the ride. Due to their wider FOV [34], women could be more affected by the visual blur than men. In the standard driving condition, the participant's head may move fast from one direction to another, according to the situation on the road, like in the real world, and therefore some visual blur from these movements might be induced.

Another observation showed that women experienced more SiS in the static driving condition than in the dynamic driving condition. Possible reasons for this might be the wider FOV [34] which could lead to more significance of the visual input and therefore, to contribute to the higher mismatch between the visual and motion cues in the static driving condition.

Surprisingly, neither the addition of motion cues nor *gender* affected *sense of presence* within an HMD driving simulation. The feeling of "being there" was not increased during the dynamic driving condition. A recent review by Weech et al. [228] showed that SiS is negatively related to *sense of presence*. Nevertheless, our findings revealed no difference between the *motion* conditions or *gender*, despite there being reported differences regarding the felt discomfort. A possible explanation for these results might be the moderate level of presence experienced across all experiments. As expected, the *type of driving* had an effect on *sense of presence*, with the participants in the standard VR driving condition feeling more present than in the participants in automated VR driving. These results match those observed by Seay et al. [193], who reported that the drivers experienced a higher level of *presence* than the passengers. A more interactive VR environment, such as the standard driving, made the participants feel more present than the fully controlled automated driving condition. These findings show that interaction with the virtual environment during the automated VR driving simulation could increase the *sense of presence*. The conducted experiments in this thesis were explicitly designed to include only an observational task in order to reduce possible co-founding factors of SiS during the virtual drive. In a more practical setting, interactions with the virtual environment naturally occur from daily tasks conducted within this environment, such as a quality check of the vehicle interior. *Sense of presence* is an excellent enhancement of the VR simulation, but it is not a factor that impairs daily work such as SiS. Moreover, a reduction in symptoms of discomfort outbreak might positively influence the *sense of presence* [229].

Relationships exploration between SiS-inducing factors and SiS

Establishing which individual factors are possibly related to SiS onset in the context of an HMD driving simulation is a step towards SiS prevention. In total, 17 individual factors were investigated for a possible association to SiS. An overview of all tested correlations is displayed in Table 10.1, where, for each factor, the correlation coefficient and its significance per experiment are presented. The SiS is evaluated using either the FMS score or one of the SSQ scores (TS—total score, N—Nausea, O—Oculomotor, D—Disorientation.) The source of the SiS value is shown in brackets after the coefficient for each correlation. The highest coefficient among the sickness values is presented. The last column of the table shows whether the factor is related to SiS and describes the direction of their relationship. Only four factors showed significant, directionally consistent relationships across all experiments. Two factors showed similar relationships in two out of three experiments; thus, this section also discusses these factors.

Table 10.1: Overview of significant correlations with SiS reported during or immediately after the VR simulation across the experiments.

Individual factor	Experiment 1	Experiment 2	Experiment 3	Overall correlation
Age	-.216 (O)	-.101 (O)	-.278* (O)	No correlation
Gender	-.184 (O)	.413** (FMS)	.404** (N)	Positive
Motion sickness history	.357** (O)	.323** (O)	.382** (N)	Positive
Presence	-.239 (N)	-.077 (FMS)	.116 (FMS)	No correlation
Vision correction	.076 (N)	.139 (D)	.138 (N)	No correlation
Sleep deprivation	-.223 (N)	.238 (O)	.094 (N)	No correlation
Physical activity	-.193 (N)	.242 (O)	.161 (O)	No correlation
Arousal	.389** (FMS)	.270** (O)	.201 (D)	Positive
Valence	-.151 (TS)	-.104 (D)	-.034 (O)	No correlation
Neuroticism	.394** (O)	.291** (O)	.400** (O)	Positive
Extroversion	-.283* (O)	-.102 (N)	-.089 (N)	No correlation
HRD	.281* (TS)	.321* (D)	.300* (D)	Positive
HR	-.135 (FMS)	.186* (N)	.259 (TS)	No correlation
SCLD	-.381 (TS)	.269 (D)	.292 (O)	No correlation
SCL	-.341 (TS)	.218* (N)	.220 (O)	No correlation
RRD	-.324 (N)	-.317 (N)	-.304 (N)	No correlation
RR	-.513** (N)	-.300** (N)	-.349* (N)	Negative

*Note. *p < .05, **p < .01.*

Gender as a factor related to SiS, and more specifically that women are more susceptible to SiS than men, has been previously reported in several studies [62, 72, 98, 127, 143, 212]. Stanney et al. [212], for example, reported a linear relationship between SSQ Disorientation, SSQ Oculomotor, and SSQ Total score and *gender*. We found that *gender* showed a significant positive relationship with SiS in two of the user evaluations. Women tend to feel more discomfort than men when using a modern HMD driving simulation. This finding supports the results from the hypotheses testing and aligns with previously reported *gender* difference in virtual simulations. Plausible explanations of why women experienced more SiS within VR driving applications were presented earlier in the chapter.

Another individual factor that showed a significant positive association to SiS in two experiments is *arousal* before the VR driving simulation. Participants who are more excited before the standard or automated VR driving are more susceptible to symptoms such as *general discomfort, fatigue, headache, eye strain, difficulty focusing, difficulty concentrating*, and *blurred vision*. Gugenheimer et al. [81] reported that HMDs evoked increased positive emotions. However, the

authors measured the emotions after the simulation to compare the emotional level to the other condition of the non-HMD group. These findings were directly related to the HMD experience itself.

In contrast, in our experiments, participants reported their *arousal* level before the start of the VR simulation. By doing this, *arousal* as an SiS factor can be evaluated as a potential predictor. It is possible that some participants were already excited to some extent about trying an HMD as for many of them it was their first time using such a device. This finding's practical implication can be an early warning regarding individuals visibly excited before using the HMD driving system. In that way, the susceptibility to SiS might be reduced, and the VR session can be more productive. Furthermore, Bruck and Watters [32] identified four factors of SiS during a VR exposure: general cybersickness, vision, arousal, and fatigue. The factor *arousal* consisted of the following variables *respiration, eye focusing, concentration, blurred vision, vertigo*, and *stomach awareness*. They even suggested that the changes in RR might occur in response to the increased *arousal* level resulting from stimuli of high movement VR. In turn, the increased level of *arousal* may result in SiS symptoms such as *fatigue, difficulty concentrating*, and *dizziness*. Thus, our findings are consistent with those of Bruck and Watters [32].

However, the results from the experiments revealed no significant difference regarding RR during the VR driving simulation. We even found a negative significant association between RR and SiS induced by a modern HMD driving simulation. The participants slowed down their breathing when they experienced more discomfort and vice versa. These results contradict the assumption that increased RR leads to increased *arousal* level resulting in SiS symptoms [32]. Interestingly, in the same two experiments, RR is negatively related to SiS, and the *arousal* is positively related. A possible explanation for this might be that other factors contribute to the increased *arousal* in our studies rather than just RR on its own. Additionally, the results seem to be consistent with another research study, which found a decrease in RR during a VR simulation [120]. Nevertheless, the same study reported a positive relationship between SiS and RR, which contradicts this thesis's current findings. A possible explanation could be that this action is taken as an automatic response by the human body to reduce the arousal levels of the nervous system. Nonetheless, with the small sample size of RR, caution must be applied, as the findings might not significantly impact the general population.

Another physiological signal, the HRD, showed a significant correlation with SiS. The HR differentiates significantly from HR recorded in a normal state of well-being when SiS increases. In other words, individuals who had a higher HR during the VR simulation felt greater discomfort than individuals who had not. This result supports evidence from previous observations reported by Cobb et al. [42], where participants with more severe SiS symptoms showed higher levels of HR than participants who did not experience SiS. Furthermore, there are resemblances between the described individual differences by Johnson [98] and the relationship expressed by HRD in this study. Nonetheless, our findings are contrary to the previously reported lack of association between HR and SiS induced by HMD simulations [74]. Nonetheless, our findings are contrary

to the previously reported lack of difference between baseline recording and condition recording of the HR regarding SiS induced by HMDs [74]. Moreover, we found that the highest correlation was between the HRD and the Disorientation cluster symptoms. This relationship is in accord with the self-reported SiS severity across the experiments and assumed profile of SiS induced by VR environments [208]. The findings might be explained by the responsive nature of the HR signal to the SiS onset [51].

A positive correlation was reported between motion sickness history and SiS scores. Participants who indicate a higher motion sickness susceptibility most likely feel more discomfort in the HMD driving simulation. This finding was assumed as the same relationship was previously presented between motion sickness history and SiS induced by virtual environments [26, 143, 173, 212]. In our experiments, the self-reported motion sickness history might be a factor associated with *gender* differences. Women reported higher susceptibility to motion sickness, and they felt more discomfort than men. These results are in agreement with those obtained by Stanney et al. [212]. It was reported in their work that the female participants had a 22% higher motion sickness susceptibility score than the male participants. They presumed that motion sickness history might be a stronger predictor of SiS induced by virtual environments than *gender*. This conclusion aligned with a prior study by Graeber and Stanney [78], which stated that *gender* differences within a virtual environment might not be due to *gender* itself but to motion sickness susceptibility.

Moreover, a significant positive relationship was found between *neuroticism* and SiS in the context of an HMD driving simulation. Participants who score higher on the *neuroticism* personality dimension are more susceptible to SiS symptoms such as *general discomfort, fatigue, headache, eye strain, difficulty focusing, difficulty concentrating*, and *blurred vision*. This finding agrees with earlier research that has related *neuroticism* to an increased level of discomfort [44, 158, 232]. Nonetheless, the reason why *neuroticism* is associated with SiS is still under discussion. Anxiety, which is a trait associated with neuroticism's personality dimension, was associated with increased motion sickness within a VR environment [158]. Furthermore, a previous study revealed a correlation between anxiety and SiS symptoms such as *general discomfort, nausea* and *stomach awareness* [137]. As a strongly arousing emotion, *anxiety* is associated with the individual factors *arousal* and *neuroticism*. Both factors showed a positive correlation to SiS onset in the correct thesis. Thus, it might be assumed that a more anxious and nervous individual is more susceptible to SiS induced by a modern HMD driving simulation.

10.2 Limitations

The current thesis has several limitations that need to be considered when interpreting the present findings and that need to be addressed in future research. One limitation was the considered sample size for some of the physiological data of the hypotheses testing. This leads to small

statistical power, meaning differences and effects could not have been identified even if they existed. Some of the physiological data, for example, was mostly presented through the recorded signal during the VR simulation and less through the baseline. Therefore, the sample size of the difference was small, as it was difficult to calculate the difference between the two variables.

A second limitation was the driving simulation's performance. The urban environment required numerous virtual assets to be included in a realistic traffic representation. Furthermore, the number of included vehicles also contributed also to the usage of more computer and graphics power, which sometimes led to a drop in performance.

A third limitation of the current work was that the researcher did not entirely control the follow-up questionnaire. The results should therefore be interpreted with caution. A link to the FSSQ was sent one hour after the VR session was over through an email client. However, some of the participants filled out the questionnaire on the next day. Therefore, the results might not be entirely accurate because the participants filled out the FSSQ based on their recollections of the experiment rather than based on their impressions immediately after the driving simulation.

A fourth limitation was the physical limitation of the moving platform and the potentially inefficient platform power. According to the results, participants did not report differences across motion conditions, which may be due to the set platform power. Moreover, due to the restricted degrees of freedom, some vehicle movements might be inaccurately replicated. An updated and fine-tuned moving platform setting could improve platform performance.

A fifth limitation was that the current study was a between-subject design. The findings of the current study may be influenced by individual differences, especially concerning SiS during the VR driving. Due to potential adaptation effects and time constraints, a between-subject design was used.

A final limitation was the severity of the SiS, which was mostly light to moderate. Based on these severity levels, the prediction models did not easily differentiate between the sick and not sick individuals. Since the work aimed to investigate SiS onset in a working environment, the pre-screening of the participants was not planned, and thus it was not conducted.

11 Contribution and Conclusion

> If a conclusion is not poetically balanced, it cannot be scientifically true.
>
> *Isaac Asimov*

This thesis started with the idea of evaluating SiS-related factors in a VR driving environment and analyzing their relationship to SiS. For that purpose, we developed an innovative driving system combining a modern HMD, a moving platform, and physiological sensors. We expanded this idea during the research phase, and built prediction models of SiS based on the evaluated factors as well as on more objective factors such as physiological signals. Therefore, the thesis ultimately combines the evaluation, exploration, and prediction of SiS into one project that contributes to the current body of research.

11.1 Contribution

More in-depth insights into SiS induced by a VR driving simulation have been provided by answering the research questions that were included in the thesis:

1. One of the significant contributions of this thesis is the development of a technically innovative VR driving simulation setup. VR driving setups utilized in earlier studies consisted of one or multiple screens as a viewing system, car seat, steering wheel, and pedals. These VR systems were, however, mostly static. In order to enhance the virtual driving experience and create a usable SiS evaluation tool, we integrated a modern HMD, a moving platform, and wireless physiological sensors into a VR driving setup. Additionally, the VR setup included a state-of-the-art computer that is powerful enough to support highly

demanding realistic VR graphics. Physiological signals such as ECG, SCL, respiratory effort, and signals from the virtual world such as HMD position and vehicle position were synchronized and continuously recorded. Furthermore, the VR system is compact, which allows for simple transportation and setup. Two pilot studies were conducted to test the feasibility of the setup before the start of the main experiments. The VR setup was developed in a manner that allowed it to be incorporated in the daily work of the Department of Interior Development at the BMW Group. Considering the results from the main experiments, the current ability to replicate vehicle movements should be improved in order to simulate these movements in the virtual world more accurately. The dynamic VR driving system is intended to be used for evaluating new interior concepts shortly after the improvements are carried out. Thus, interior concepts and interaction within these aforementioned interiors can be safely assessed in a fully controlled virtual environment.

2. Another major contribution is the utilization of various subjective, and objective measures in the context of VR automated driving simulation. In general, to assess SiS onset, subjective measures such as questionnaires are used. It must be noted, however, that the data from these questionnaires can at times be inaccurate as, unintentionally, individuals might understate their symptoms. A number of studies have started to utilize physiological sensors as SiS measures, in addition to the standard questionnaires. Similarly, we combined the subjective measures of discomfort with more objective measures such as ECG, SCL, and respiratory effort. In order to obtain more insight into SiS, we utilized a self-reporting sickness scale during the VR driving, questionnaires immediately after the driving simulation as well as one hour later, and physiological signals. Furthermore, behavioral data, as another subjective measure, can be used in the SiS evaluation. Individual head rotations and vehicle speed and acceleration can be calculated from the online recorded HMD and vehicle movements in the virtual world. In the current thesis, we used this behavioral data exclusively to build prediction models due to time constraints. The combination of questionnaires and physiological and behavioral measurements creates an excellent assessment tool for SiS in the context of a modern HMD driving simulation. The VR system will be utilized for HCI research as an evaluation tool within the BMW Group.

3. Furthermore, the gathered insights of SiS induced by an HMD driving simulation is an additional contribution of this thesis. Particularly, the findings from the automated VR driving experiments extended the current knowledge on SiS and automated virtual driving in an urban scenario. Our findings strengthen the assumption that SiS induced by HMDs is primarily a visually-induced sickness and is more influenced by the visual motion cues than the physical motion cues. In support of previous studies, we found out that women are more susceptible than men during an automated VR driving simulation. The findings enhance the current literature by providing empirical evidence that gender has an effect on the novel virtual driving environment. Moreover, women should be more cautious when using such a VR system, especially if the environment includes any moving platforms.

Other findings regarding virtual vehicle control contribute to the current body of research. The automated driving did not induce more discomfort than the standard driving in the context of modern a HMD driving simulation. Contrary to the assumption that less control over the vehicle would evoke more discomfort, we found that there were no differences between the two types of driving.

Additionally, only six individual factors out of 17 showed a relationship to SiS in the context of modern HMD driving simulations. As assumed susceptibility to motion sickness, physiological responses and *gender* are related to SiS susceptibility. However, two of the factors, namely *arousal and neuroticism*, revealed an interesting association to the felt discomfort. A possible underlying connection between these two factors might be the reason for their relationship to SiS. These findings show that the role of personality traits is often underestimated, and thus there could be more to offer in the SiS research. Nevertheless, personality traits vary significantly among individuals, and various external factors could affect them before the VR session.

4. The last main contribution is the demonstration that physiological signals can predict SiS onset within a VR driving simulation environment. Previous research showed that bodily responses could be utilized as SiS predictors. Nevertheless, that concept has never been evaluated before within a modern HMD driving urban environment. Furthermore, the prediction model, which used only the cordial signal, performed similarly to the model using all physiological signals. A few factors extracted from the cordial signal contributed significantly to the SiS prediction model. Thus, these insights are valuable and contribute to the current body of SiS research. Moreover, the reported results hint at the potential role played by bodily responses, so these too could be utilized as an online predictor.

11.2 Conclusion

Although not every hypothesis was supported, the primary research questions were answered. As evaluated in two different driving types, the moving platform had neither a significant effect on SiS onset nor on the sense of presence in the current study. Gender, as revealed by the significant main effect on SiS, is an SiS factor with a high susceptibility potential. As presented by standard and automated driving, the type of driving did not significantly affect SiS onset. However, the standard type of driving affected presence. The combination of physiological and behavioral factors can predict SiS induced by a VR driving simulation. A further investigation showed that the cordial signal could alone predict SiS with an accuracy very close to the combination model. The models built solely on individual factors were not particularly successful when excluded from complex models with too many factors. This demonstrates that SiS can be predicted more effectively when using objective measures during the VR driving simulation rather than the VR sessions' factors. Nonetheless, the individual factors could be used as recommendations to the

users and developers and could also identify which users are more susceptible to SiS. Early prevention and knowledge is necessary to find a proper solution.

The main implication of the current dissertation is the several insights offered on how to reduce SiS in an HMD driving simulation. First, the exposure time should be limited to 12 minutes. Twelve minutes was the time limit at which most of the participants reported greater level of sickness or discontinued the simulation. Second, before the first long working VR session, an individual should have a short session or two with the VR system: the less experience a user has with an HMD, the greater the possibility of feeling discomfort. The human body needs time to adapt to the new environment and become acquainted with the new rules of perception. Third, the caffeine beverage consumption should be reduced approximately 4 hours before the VR session. The higher caffeine consumption is, the stronger the sickness symptoms are one hour later or more after the VR exposure. Fourth, the measured arousal before the VR driving simulation can be used as an early indicator of SiS. This is especially important for women as their level of excitement could additionally contribute to an onset of sickness. Fifth, the measured personality traits can be used as a premature indicator for SiS. Individuals who have neurotic tendencies experience more discomfort in the VR driving simulation. Sixth, when using an HMD, the users should reduce their head movements. Users should be instructed first to adapt to the virtual environment with slow head or body movements and then, during the VR driving simulation, to limit their head movements to only the most essential ones. Finally, before using an HMD, female users should be aware of their high susceptibility to SiS. Suitable refreshments, as well as short breaks, should be provided in an easily accessible area. However, all these factors are still related to using questionnaires, and therefore the data might be strongly subjective.

Another implication is that even though a modern HMD was used, the SiS was still present. This shows that SiS is more related to perception cues than to the use of technology. The most common symptoms demonstrate that the discomfort comes primarily from the HMD and visual motion perception rather than from the missing physical motion cues. Lastly, but in no way less important, the current thesis contributes to existing knowledge of SiS by providing a piece of empirical evidence that physiological signals can be used as an indicator of SiS onset. This establishes a direction for future research in detecting and preventing SiS symptoms within a VR driving simulation.

11.3 Future Work

Further work is required to establish the viability of individual and physiological factors as predictors of SiS. Considering physiological signals hint at a promising prediction effect, additional invasive physiological signals could be explored. Using more data could help to improve the prediction model. Furthermore, using different sensors such as those for measuring gastric or brain activity could extend the use of physiological signals. A threshold dependency of the

physiological values before the VR simulation could be developed. Using a within-subject design can benefit from investigating in depth the individuals' differences in SiS onset. Future application development for online automatic SiS recognition is a possible next step in SiS research. Individuals who could be informed that they should stop the VR simulation due to a high chance of becoming sick would be able to enjoy a longer and more pleasant time in the virtual world. Such an application could possibly be a SmartWatch device that reads the heart or skin conductance data while the application is used.

Considering the findings, a future study might include a different driving scenario, such as a highway or country road. Induced boredom during automated driving can be reduced by assigning a task such as the objective of driving or in-car interaction. This task might contribute additionally to reducing SiS. As a predominantly visually induced discomfort, the findings show that individuals mostly suffered from disorientation-related symptoms. Nonetheless, according to the SSQ, one of these symptoms is the symptom of *nausea*. The same symptom is actually part of another symptom cluster, namely the Nausea cluster. A more precise subjective measure is needed, which is designed specifically for the sickness induced by HMDs. Such a questionnaire has been attempted, but it suffers from a lack of validity. Subjective measures can be partially unreliable, as some participants did not answer honestly, whether intentionally or not. Thus, objective measures such as physiological signals can contribute to a more precise measure of SiS.

The lack of significant difference between the motion conditions in most experiments points to the operationalization of a different moving platform performance configuration. Different levels of motion could establish a motion threshold for the HMD VR systems, which are built using consumer electronics. Even though gender demonstrated affects on SiS in the current thesis, future research should further explore the relationship between gender and factors such as motion sickness history, vision correction, adaptation, personality traits, and global visual flow. Several SiS factors have been described in the current work in an attempt to aid the investigation of possible inducing factors in the context of an HMD driving simulation. Nonetheless, further research should be undertaken to study how other simulation factors (such as different focus areas and independent visual backgrounds) as well as hardware factors (such as resolution and position tracking errors) affect SiS.

A Physiological sensors specifications

ECG

The ECG sensor has a gain of 1000 with a range of $\pm 1.5mV$ (with VCC = $3V$), bandwidth between 0.5 and $100Hz$, and energy consumption of ~$1mA$. The transfer function for the ECG signal is shown below (see Equation 5.8). Where VCC is the operating voltage – $3V$, G_{ECG} is the sensor gain – 1000, ADC (Analog-to-Digital Converter) is the value sampled from the channel, and n is a number of bits of the channel – by default n = 16. The transfer function of the ECG signal is shown in the equation below. The first equation shows the ECG value in Volt (V) and the second, the value in milli Volts (mV). The ECG sensor used in the experiments and a sample ECG signal is shown in Figure A.1.

$$ECG(V) = \frac{(\frac{ADC}{2^n} - \frac{1}{2})VCC}{G_{ECG}} \tag{A.1}$$

$$ECG(mV) = ECG(V)1000 \tag{A.2}$$

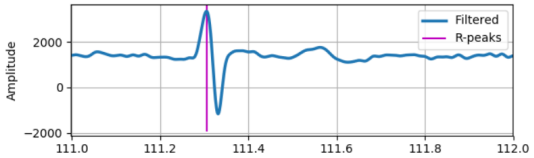


Figure A.1: The image shows a sample signal from an ECG sensor.

BVP

The BVP sensor used in the experiments and a sample BVP signal is shown in Figure A.2. The BVP sensors were used together with the ECG sensor to record the heart signal. Both sensors were used at the same time as a precocious measure of sensor failure.

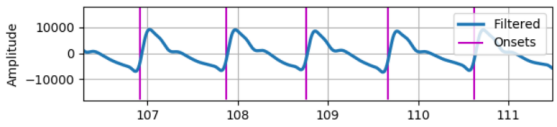


Figure A.2: The image shows a sample signal from a BVP sensor.

EDA

The EDA sensor has a range of $0 - 25\mu$S, bandwidth of $0 - 3Hz$, and energy consumption of $0.72mA$. The transfer function for the EDA signal is shown in Equation 2. The EDA sensor used in the experiments and a sample EDA signal is shown in Figure A.3.

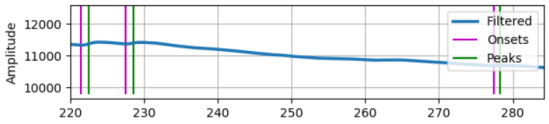


Figure A.3: The image shows a sample signal from an EDA sensor.

RESP

The respiration sensor uses piezoelectric film technology. The transfer function for the respiration signal can range from - 50% to 50%. The respiration sensor used in the experiments and a sample respiration signal is shown in Figure A.4.

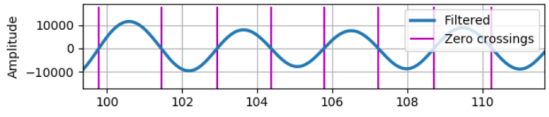


Figure A.4: The image shows a sample signal from a respiration sensor.

B

Pre-questionnaire

Vielen Dank, dass Sie sich Zeit nehmen, an dieser Studie teilzunehmen.

Bitte lesen Sie sich vor Beginn der Studie folgende Teilnehmerinformationen sorgfältig durch.

Ihre Daten sind selbstverständlich vertraulich und werden nur in anonymisierter Form gespeichert, wissenschaftlich ausgewertet und ggf. publiziert. Ausgehend von diesen Daten können keine Rückschlüsse auf Ihre Person gezogen werden. Ihre Teilnahme an dieser Studie ist freiwillig. Es steht Ihnen zu jedem Zeitpunkt frei, Ihre Teilnahme abzubrechen, ohne dass Ihnen daraus Nachteile entstehen.

Im Folgenden werden Sie kurz zu Ihren demographischen Daten, Gewohnheiten hinsichtlich Autofahren und Videospielen, physischer Aktivität und Erfahrung mit VR Fahrsimulatoren befragt.

Falls Sie noch Fragen zu dieser Studie haben sollten, wenden Sie sich bitte an: stanislava.rangelova@bmw.de

Diese Befragung wird maximal 5 Minuten in Anspruch nehmen.

Bitte klicken Sie nun auf "Weiter", um den Teilnahmebedingungen zuzustimmen und die Umfrage zu starten.

Teil A: Soziodemographische Daten

Bitte füllen Sie die vorhanden Textfelder aus und wählen Sie die auf Sie zutreffendste Antwort aus.

A1. Bitte generieren Sie hier zunächst Ihren Versuchspersonen-Code. Er setzt sich wie folgt zusammen:

Erste 2 Buchstaben des Vornamens Ihrer Mutter, Letzte 2 Buchstaben des Vornamens Ihres Vaters, Tag Ihres eigenen Geburtstags.

z.B.: Mutter: Anna; Vater: Peter; eigener Geb. 05.11.1983 = ANER05

Der Versuchspersonen-Code dient zur Anonymisierung Ihrer Antworten. Bitte achten Sie auf Richtigkeit Ihrer Angabe.

Zu Beginn des Versuchs im VR Fahrsimulator werden Sie gebeten, den selben Code erneut zu generieren.

A2. Wählen Sie bitte das heutige Datum aus.

A3. Alter

A4. Geschlecht

- ☐ weiblich
- ☐ männlich

A5. Haben Sie eine Sehschwäche?

- ☐ Ja
- ☐ Nein

A6. Was ist Ihr höchster Bildungsabschluss?

- ☐ Hauptschule/Realschule
- ☐ Gymnasium oder vergleichbar
- ☐ Bachelor oder vergleichbar
- ☐ Master oder vergleichbar
- ☐ Doktor/Professor

Teil B: Autofahren

Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

B1. Wie oft sind Sie in den letzten 12 Monaten im Durchschnitt Auto gefahren?

- ☐ Nie
- ☐ Weniger als 1 mal pro Monat
- ☐ Mehrmals pro Monat
- ☐ Mehrmals pro Woche
- ☐ Täglich

B2. Wie hoch ist Ihre jährliche Fahrleistung in Kilometern (inklusive Urlaubsfahrten und geschäftlichen Fahrten)?

- ☐ 0 km pro Jahr (keine Fahrt)
- ☐ 0 - 5.000 km pro Jahr
- ☐ 5.001 - 10.000 km pro Jahr
- ☐ 10.001 - 20.000 km pro Jahr
- ☐ 20.001 - 30.000 km pro Jahr
- ☐ 30.001 - 40.000 km pro Jahr
- ☐ mehr als 40.000 km pro Jahr

Teil C: Videospielen

Bitte wählen Sie die auf Sie zutreffendste Antwort aus und füllen Sie die vorhanden Textfelder aus.

C1. *Bitte wählen Sie die auf Sie zutreffendste Antwort aus.*

	Trifft gar nicht zu				Trifft völlig zu
Ich spiele täglich Videospiele (z.B. PC oder Konsole).	☐	☐	☐	☐	☐
Ich habe während oder nach dem Videospielen Unannehmlichkeiten (z.B. Desorientierung, Kopfschmerzen, Übelkeit oder Schwindel) verspürt.	☐	☐	☐	☐	☐

C2. Welche Unannehmlichkeiten haben Sie verspürt?

Falls Sie keine Videospiele spielen, können Sie diese Frage überspringen.

- ☐ Keine

Desorientierung ☐

Kopfschmerzen ☐

Übelkeit ☐

Schwindel ☐

Sonstiges ☐

Sonstiges

Teil D: Physische Aktivität

Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

D1. Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

	Trifft gar nicht zu				Trifft völlig zu
Ich bin in meinem Alltag körperlich aktiv (z.B. Treppen steigen, Spaziergänge etc.).	☐	☐	☐	☐	☐
Ich mache mehrmals die Woche Sport.	☐	☐	☐	☐	☐

Teil E: Erfahrung mit Fahrsimulatoren

Bitte füllen Sie die vorhanden Textfelder aus und wählen Sie die auf Sie zutreffendste Antwort aus.

E1. Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

	Trifft gar nicht zu				Trifft völlig zu
Ich habe Erfahrung mit Fahrsimulatoren.	☐	☐	☐	☐	☐
Ich habe während oder nach der Fahrsimulation Unannehmlichkeiten (z.B. Desorientierung, Kopfschmerzen, Übelkeit oder Schwindel) verspürt.	☐	☐	☐	☐	☐
Ich musste eine vorherige Fahrt in einem Fahrsimulator beenden, weil ich mich unwohl fühlte.	☐	☐	☐	☐	☐

E2. Welche Unannehmlichkeiten haben Sie verspürt? Falls Sie keine Erfahrung mit Fahrsimulatoren haben, können Sie diese und die nächste Frage überspringen.

Keine ☐

Desorientierung ☐

Kopfschmerzen ☐

Übelkeit ☐

Schwindel ☐

Sonstiges ☐

Sonstiges

E3. Wie lange waren Sie damals in der Lage in einem Fahrsimulator zu fahren?

Teil F: Erfahrung mit Virtual Reality

Bitte füllen Sie die vorhanden Textfelder aus und wählen Sie die auf Sie zutreffendste Antwort aus.

F1. Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

	Trifft gar nicht zu				Trifft völlig zu
Ich habe Erfahrung mit Virtual Reality.	☐	☐	☐	☐	☐
Ich habe schon mal ein Head-Mounted-Display (HMD), klassische VR Brille, verwendet.	☐	☐	☐	☐	☐
Haben Sie während oder nach der Verwendung Unannehmlichkeiten verspürt?	☐	☐	☐	☐	☐

F2. Welche Unannehmlichkeiten haben Sie verspürt? Falls Sie keine Erfahrung mit VR haben, können Sie diese und die nächste Frage überspringen.

Keine ☐

Desorientierung ☐

Kopfschmerzen ☐

Übelkeit ☐

Schwindel ☐

Sonstiges

Sonstiges

F3. Wie lange haben Sie pro Sitzung mit der virtuellen Welt interagiert?

Teil G: Erfahrung mit Bewegungskrankheit

Dieser Teil der Fragen betrifft Ihre Erfahrungen, die Sie mit Fortbewegung und Bewegungskrankheit in Ihrer Kindheit (bis 12 Jahre) gemacht haben.

Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

G1. Als Kind (bevor Sie 12 Jahre alt waren), wie oft fühlten Sie sich krank oder übel während der Fahrt in…

	Nie verwendet	Nie krank	Selten krank	Manchmal krank	Häufig krank
Autos	☐	☐	☐	☐	☐
Busse oder Mini-Busse	☐	☐	☐	☐	☐
Züge	☐	☐	☐	☐	☐
Flugzeuge	☐	☐	☐	☐	☐
Kleine Boote	☐	☐	☐	☐	☐
Schiffe / Fähren	☐	☐	☐	☐	☐
Schaukeln	☐	☐	☐	☐	☐
Karussells auf Spielplätzen	☐	☐	☐	☐	☐
Achterbahnen und ähnliche Fahrgeschäfte auf Jahrmärkten oder in Freizeitparks	☐	☐	☐	☐	☐

Teil H: Erfahrung mit Bewegungskrankheit

Dieser Teil der Fragen betrifft Ihre Erfahrungen, die Sie mit Fortbewegung und Bewegungskrankheit in den letzten 10 Jahren gemacht haben.

Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

H1. In den letzten 10 Jahren, wie oft fühlten Sie sich krank oder übel während der Fahrt in…

	Nie verwendet	Nie krank	Selten krank	Manchmal krank	Häufig krank
Autos	☐	☐	☐	☐	☐
Busse oder Mini-Busse	☐	☐	☐	☐	☐
Züge	☐	☐	☐	☐	☐
Flugzeuge	☐	☐	☐	☐	☐
Kleine Boote	☐	☐	☐	☐	☐
Schiffe / Fähren	☐	☐	☐	☐	☐
Schaukeln	☐	☐	☐	☐	☐
Karussells auf Spielplätzen	☐	☐	☐	☐	☐
Achterbahnen und ähnliche Fahrgeschäfte auf Jahrmärkten oder in Freizeitparks	☐	☐	☐	☐	☐

Teil I: Persönlichkeit

I1.

	Trifft gar nicht zu				Trifft völlig zu
Ich fühle mich anderen oft unterlegen.	☐	☐	☐	☐	☐
Wenn ich unter starkem Stress stehe, fühle ich mich manchmal, als ob ich zusammenbräche.	☐	☐	☐	☐	☐
Ich fühle mich oft angespannt und nervös.	☐	☐	☐	☐	☐
Manchmal fühle ich mich völlig wertlos.	☐	☐	☐	☐	☐
Zu häufig bin ich entmutigt und will aufgeben, wenn etwas schief geht.	☐	☐	☐	☐	☐
Ich fühle mich oft hilflos und wünsche mir eine Person, die meine Probleme löst.	☐	☐	☐	☐	☐

I2.

	Trifft gar nicht zu				Trifft völlig zu
Ich habe gern viele Leute um mich herum.	☐	☐	☐	☐	☐
Ich bin leicht zum Lachen zu bringen.	☐	☐	☐	☐	☐
Ich bin gerne im Zentrum des Geschehens.	☐	☐	☐	☐	☐
Ich habe oft das Gefühl, vor Energie überzuschäumen.	☐	☐	☐	☐	☐
Ich bin ein fröhlicher, gutgelaunter Mensch.	☐	☐	☐	☐	☐
Ich bin ein sehr aktiver Mensch.	☐	☐	☐	☐	☐

Vielen Dank für Ihre Teilnahme! Wir freuen uns Sie schon bald in unserem VR Fahrsimulator begrüßen zu dürfen.

C On-spot Questionnaire

Vielen Dank, dass Sie sich Zeit nehmen, an dieser Studie teilzunehmen.

Bitte lesen Sie sich vor Beginn der Studie die Ihnen vorliegende Teilnehmerinformationen sorgfältig durch. Falls Sie noch Fragen zu dieser Studie haben sollten, wenden Sie sich bitte an den Versuchsleiter. Mit Ihrer Unterschrift auf dem vor Ihnen liegenden Formular geben Sie Ihr Einverständnis zur Teilnahme an dieser Studie.

Diese Studie wird maximal 50 Minuten in Anspruch nehmen.

Wir wünschen Ihnen viel Spaß :-).

Bitte klicken Sie nun auf "Weiter", um einige Fragen zu Ihrem Wohlbefinden zu beantworten.

Teil A: Probandencode

A1. **Bitte generieren Sie hier zunächst Ihren Versuchspersonen-Code. Er setzt sich wie folgt zusammen:**

Erste 2 Buchstaben des Vornamens Ihrer Mutter, Letzte 2 Buchstaben des Vornamens Ihres Vaters, Tag Ihres eigenen Geburtstags.

z.B.: Mutter: Anna; Vater: Peter; eigener Geb. 05.11.1983 = ANER05

Teil B: Koffeinkonsum

Bitte wählen Sie die auf Sie zutreffendste Antwort aus und füllen Sie das vorhande Textfeld aus.

B1. Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

	Gar kein Koffein				Sehr viel Koffein (mehr als 4 Tassen)
Ich habe in den letzten 4 Stunden Koffein (z.B. Kaffee, Cola, Red Bull etc.) zu mir genommen.	☐	☐	☐	☐	☐

Teil C: Schlaf

Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

C1. Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

	Trifft gar nicht zu				Trifft völlig zu
Ich habe in den letzten 7 Nächten ausreichend geschlafen.	☐	☐	☐	☐	☐
Ich habe in der gestrigen Nacht ausreichend geschlafen.	☐	☐	☐	☐	☐
Ich fühle mich heute müde.	☐	☐	☐	☐	☐

Teil D: Derzeitige Stimmung

Wie fühlen Sie sich jetzt gerade?

D1. Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

Skala 1

	Negativ				Positiv
Skala 1	☐	☐	☐	☐	☐

D2. Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

Skala 2

	Ruhig und gelassen				Unruhig und aufgeregt
Skala 2	☐	☐	☐	☐	☐

Teil E: Wenden sie sich bitte an den Versuchsleiter.

Teil F: Fast Motion Sickness Scale (Ausfüllung durch Versuchsleiter)

(Ausfüllung durch Versuchsleiter)

F1.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Intensität 0 min

F2.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Symptom 0 min

F3.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Intensität 3 min

F4.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Symptom 3 min

F5.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Intensität 6 min

F6.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Symptom 6 min

F7.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Intensität 9 min

F8.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Symptom 9 min

F9.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Intensität 12 min

F10.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Symptom 12 min

F11.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Intensität 15 min

F12.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Symptom 15 min

F13.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Intensität 18 min

F14.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Symptom 18 min

F15.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Intesität 21 min

F16.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Symptom 21 min

F17.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Intesität 24 min

F18.

Alle drei Minuten: Nennen Sie bitte auf einer Skala von 0 - 20 wie unwohl Sie sich derzeit fühlen.

Welches Symptom überwiegt?

Symptom 24 min

Teil G: Wenden sie sich bitte an den Versuchsleiter.

Teil H: Wohlbefinden

Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

H1. Ich verspüre in diesem Moment…

**Gleichgewichtsstörungen wird als Orientierungsverlust in Bezug auf das aufrechte Stehen erlebt*

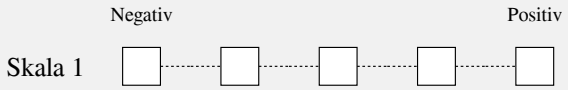
***Magen macht sich bemerkbar beschriebt gewöhnlich ein Gefühl von Unbehagen, das kurz vor der Übelkeit ist*

	Nicht	Leicht	Moderat	Schwer
Allgemeines Unwohlsein	☐	☐	☐	☐
Ermüdung	☐	☐	☐	☐
Kopfschmerzen	☐	☐	☐	☐
Angestrengte Augen	☐	☐	☐	☐
Schwierigkeiten, scharf zu sehen	☐	☐	☐	☐
Erhöhter Speichelfluss	☐	☐	☐	☐
Schwitzen	☐	☐	☐	☐
Übelkeit	☐	☐	☐	☐
Konzentrationsschwierigkeiten	☐	☐	☐	☐
Kopfdruck	☐	☐	☐	☐
Verschwommene Sicht	☐	☐	☐	☐
Schwindel (bei offenen Augen)	☐	☐	☐	☐
Schwindel (bei geschlossenen Augen)	☐	☐	☐	☐
Gleichgewichtsstörungen*	☐	☐	☐	☐
Magen macht sich bemerkbar**	☐	☐	☐	☐
Aufstoßen	☐	☐	☐	☐

Teil I: Derzeitige Stimmung

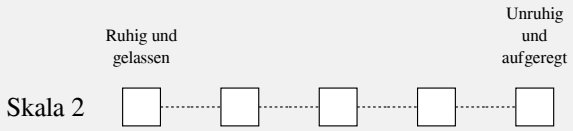
I1. Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

Skala 1



I2. Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

Skala 2



Teil J: Fahrvergnügen

Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

J1.

	Trifft gar nicht zu				Trifft völlig zu
In der computererzeugten Welt hatte ich den Eindruck, “dort gewesen zu sein".	☐	☐	☐	☐	☐
Ich hatte das Gefühl, dass die virtuelle Umgebung mich umgibt.	☐	☐	☐	☐	☐
Ich hatte das Gefühl, nur Bilder wahrzunehmen.	☐	☐	☐	☐	☐
Ich hatte nicht das Gefühl, im virtuellen Raum zu sein.	☐	☐	☐	☐	☐
Ich hatte das Gefühl, in dem virtuellen Raum zu handeln anstatt etwas von außen zu bedienen.	☐	☐	☐	☐	☐
Ich fühlte mich im virtuellen Raum anwesend.	☐	☐	☐	☐	☐
Mir war die reale Welt um mich herum bewusst, während ich mich durch die virtuelle Welt bewegte (z.B. durch Geräusche, Raumtemperatur, andere Personen etc.).	☐	☐	☐	☐	☐
Meine reale Umgebung war mir nicht mehr bewusst.	☐	☐	☐	☐	☐
Ich achtete noch auf die reale Umgebung.	☐	☐	☐	☐	☐
Meine Aufmerksamkeit war von der virtuellen Welt völlig in Bann gezogen.	☐	☐	☐	☐	☐
Meine Erfahrung in der virtuellen Umgebung glich der Erfahrung meiner realen Umgebung.	☐	☐	☐	☐	☐
Die virtuelle Welt erschien mir realistischer als die reale Welt.	☐	☐	☐	☐	☐
Die Aktionen, die ich initiiert (oder durchgeführt) habe, haben auf die Umwelt reagiert.	☐	☐	☐	☐	☐
Ich war an der Erfahrung der virtuellen Realität beteiligt.	☐	☐	☐	☐	☐
Ich habe eine Verzögerung zwischen meinen Handlungen und den erwarteten Ergebnissen erlebt.	☐	☐	☐	☐	☐
Mir hat die Simulation Spaß gemacht.	☐	☐	☐	☐	☐
Ich fühlte mich traurig, als die Simulation vorbei war.	☐	☐	☐	☐	☐
Ich möchte die Erfahrung wiederholen, die ich gerade gemacht habe.	☐	☐	☐	☐	☐
Mein Fahrerlebnis war interessant.	☐	☐	☐	☐	☐

Teil K: Anmerkungen

K1. Zum Abschluss würde uns noch Ihre persönliche Meinung (Anregungen) zur virtuellen Fahrt interessieren.

Diese Angabe ist freiwillig.

Vielen Dank für Ihre Teilnahme!

D Follow-up Questionnaire

Vielen Dank, dass Sie sich Zeit nehmen, an dieser Studie teilzunehmen.

Dieser Fragebogen wird ca. 5 Minuten in Anspruch nehmen.

Falls Sie noch Fragen zu dieser Studie haben sollten, wenden Sie sich bitte an: stanislava.rangelova@bmw.de.

Bitte klicken Sie nun auf "Weiter", um einige Fragen zu Ihrem Wohlbefinden zu beantworten.

Teil A: Probandencode

A1. **Bitte generieren Sie hier zunächst Ihren Versuchspersonen-Code. Er setzt sich wie folgt zusammen:**

Erste 2 Buchstaben des Vornamens Ihrer Mutter, Letzte 2 Buchstaben des Vornamens Ihres Vaters, Tag Ihres eigenen Geburtstags.

z.B.: Mutter: Anna; Vater: Peter; eigener Geb. 05.11.1983 = ANER05

Teil B: Wohlbefinden

Bitte wählen Sie die auf Sie zutreffendste Antwort aus.

B1. Ich verspüre in diesem Moment…

**Gleichgewichtsstörungen wird als Orientierungsverlust in Bezug auf das aufrechte Stehen erlebt*

***Magen macht sich bemerkbar beschriebt gewöhnlich ein Gefühl von Unbehagen, das kurz vor der Übelkeit ist*

	Nicht	Leicht	Moderat	Stark
Allgemeines Unwohlsein	☐	☐	☐	☐
Ermüdung	☐	☐	☐	☐
Kopfschmerzen	☐	☐	☐	☐
Angestrengte Augen	☐	☐	☐	☐
Schwierigkeiten, scharf zu sehen	☐	☐	☐	☐
Erhöhter Speichelfluss	☐	☐	☐	☐
Schwitzen	☐	☐	☐	☐
Übelkeit	☐	☐	☐	☐
Konzentrationsschwierigkeiten	☐	☐	☐	☐
Kopfdruck	☐	☐	☐	☐
Verschwommene Sicht	☐	☐	☐	☐
Schwindel (bei offenen Augen)	☐	☐	☐	☐
Schwindel (bei geschlossenen Augen)	☐	☐	☐	☐
Gleichgewichtsstörungen*	☐	☐	☐	☐
Magen macht sich bemerkbar**	☐	☐	☐	☐
Aufstoßen	☐	☐	☐	☐

Vielen Dank für Ihre Teilnahme!

E

Consent Form

Einwilligungserklärung

Vorname, Nachname _____

Datum/Uhrzeit _____ / _____

Einwilligungserklärung (nicht Zutreffendes streichen)

Ich wurde mündlich und schriftlich ausreichend über diese Studie informiert. Ich nehme freiwillig an der Testfahrt teil und kann jederzeit von der Fahrt zurücktreten bzw. diese abbrechen.

Hiermit gebe ich außerdem bekannt, dass ich während des Versuchs nicht unter dem Einfluss von Alkohol, Drogen oder Medikamenten stehe.

Ich verpflichte mich, mit dem Fahrsimulator sorgfältig umzugehen und den Anweisungen des/der Versuchsleiter/in Folge zu leisten. Während der Versuchsfahrt im Simulator sind die realen Verkehrsregeln (Straßenverkehrsordnung) einzuhalten.

Personen, bei denen eine Schwangerschaft vorliegt oder die an Herzkrankheiten, Innenohrerkrankungen, Grippe, akuten Augen- oder Ohrinfektionen, Schwindelgefühl, Epilepsie, Klaustrophobie, Kreislaufbeschwerden oder vergleichbaren Krankheiten leiden, sind aufgrund erhöhtem Risiko von der Fahrt ausgeschlossen. Mit meiner Unterschrift bestätige ich, dass diese Punkte nicht auf mich zutreffen.

Ich bin darüber informiert, dass ich während der Fahrt gefilmt/fotografiert werde und bin damit einverstanden.

Ich wurde darüber informiert, die erhobenen Daten in BMW IT Systemen anonymisiert gespeichert und während der Studie auf folgenden Wegen dokumentiert und verarbeitet werden:

☐ **Fragebogendaten**
☐ **Psychophysiologische Messdaten** (EKG, BDV, Atemfrequenz, EDA, Eye Tracker)
☐ **Videodaten**
☐ **Fotos**

Ich bin damit einverstanden, dass die im Rahmen dieser Studie an mir erhobenen Daten aufgezeichnet, gespeichert und die Ergebnisse in wissenschaftlichen Veröffentlichungen (ggf. auch Präsentationen der Videos) in anonymisierter Form verwendet werden können.

Ich habe die Einwilligungserklärung gelesen und verstanden.

_____

Ort, Datum Unterschrift

F

Correlations

Table F.1: Experiment 1 results: correlations between SSQ, FMS, and FSSQ scores, age, and MSSQ score.

	Age	n	MSSQ	n
MSSQ score	-.425**	62	-	-
FMS score	.137	63	.256*	62
SSQ total score	-.020	63	.221	62
SSQ Nausea	.134	63	.255*	62
SSQ Disorientation	-.081	63	.259*	62
SSQ Oculomotor	-.216	63	.357**	62
FSSQ total score	.014	58	.213	58
FSSQ Nausea	.095	58	.161	58
FSSQ Disorientation	-.072	58	.145	58
FSSQ Oculomotor	-.031	58	.211	58

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.2: Experiment 1 results: correlations between FMS, SSQ, FSSQ scores, gender, and vision correction.

	Gender	n	Vision correction	n
Gender	-	-	-.003	63
FMS score	-.150	63	-.008	63
SSQ total score	-.109	63	-.040	63
SSQ Nausea	-.157	63	.076	63
SSQ Disorientation	-.161	63	.026	63
SSQ Oculomotor	-.184	63	.029	63
FSSQ total score	-.288*	58	-.039	58
FSSQ Nausea	-.273*	58	.055	58
FSSQ Disorientation	-.268*	58	-.045	58
FSSQ Oculomotor	-.276*	58	-.120	58

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.3: Experiment 1 results: correlations between FMS, SSQ, FSSQ scores, sleep deprivation, physical activity, and sense of presence

	Sleep deprivation	n	Physical activity	n	Presence	n
FMS score	-.173	63	-.133	62	-.065	63
SSQ total score	-.040	63	.091	62	-.076	63
SSQ Nausea	-.223	63	-.193	62	-.239	63
SSQ Disorientation	-.211	63	-.031	62	-.105	63
SSQ Oculomotor	-.187	63	.057	62	-.023	63
FSSQ total score	-.100	58	-.119	58	-.225	58
FSSQ Nausea	-.139	58	-.172	58	-.254	58
FSSQ Disorientation	-.060	58	-.051	58	-.145	58
FSSQ Oculomotor	-.057	58	-.066	58	-.157	58

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.4: Experiment 1 results: correlations between FMS mean, SSQ, FSSQ total score, HR difference (HRD), and HR record.

	HR difference	n	HR record	n
FMS score	-.061	50	-.135	53
SSQ total score	.281*	50	.001	53
SSQ Nausea	.090	50	.106	53
SSQ Disorientation	.133	50	.067	53
SSQ Oculomotor	.206	50	.087	53
FSSQ total score	-.127	46	.176	49
FSSQ Nausea	-.075	46	.161	49
FSSQ Disorientation	-.168	46	.163	49
FSSQ Oculomotor	-.121	46	.170	49

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.5: Experiment 1 results: correlations between FMS, SSQ, FSSQ, RR difference (RRD), and RR record.

Pearson's correlation				
	RR Difference	n	RR record	n
SSQ total score	-.055	19	-	-
SSQ Nausea	-.324	19	-	-
FSSQ total score	.024	18	-	-
FSSQ Oculomotor	.210	18	-	-
Spearman's correlation				
	RR difference	n	RR record	n
FMS score	-.213	19	-.211	33
SSQ total score	-	-	.113	33
SSQ Nausea	-	-	-.513**	33
SSQ Disorientation	-.271	19	-.227	33
SSQ Oculomotor	-.014	19	-.119	33
FSSQ total score	-	-	-.298	31
FSSQ Nausea	-.265	18	-.441*	31
FSSQ Disorientation	-.263	18	-.259	31
FSSQ Oculomotor	-	-	-.098	31

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.6: Experiment 1 results: correlations between FMS, SSQ, FSSQ, SCL difference (SCLD), and SCL record.

Pearson's correlation				
	SCL Difference	n	SCL record	n
SSQ total score	-.381	24	-	-
FSSQ total score	-.110	19	-	-
Spearman's correlation				
	SCL difference	n	SCL record	n
FMS score	-.293	24	-.279	30
SSQ total score	-	-	-.341	30
SSQ Nausea	-.159	24	-.077	30
SSQ Disorientation	-.342	24	-.104	30
SSQ Oculomotor	-.296	24	-.107	30
FSSQ total score	-	-	.029	25
FSSQ Nausea	.004	19	.106	25
FSSQ Disorientation	-.098	19	-.027	25
FSSQ Oculomotor	-.019	19	.032	25

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.7: Experiment 2 results: correlations between FMS, SSQ, FSSQ, age, and MSSQ score.

	Age	n	MSSQ	n
Age	-	-	-.327*	60
FMS score	.060	62	.123	60
SSQ total score	-.022	62	.252	60
SSQ Nausea	.089	62	.211	60
SSQ Disorientation	-.035	62	.163	60
SSQ Oculomotor	-.101	62	.323*	60
FSSQ total score	-.052	56	.251	54
FSSQ Nausea	.112	56	.221	54
FSSQ Disorientation	.048	56	.234	54
FSSQ Oculomotor	.010	56	.243	54

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.8: Experiment 2 results: correlations between FMS, SSQ, FSSQ, neuroticism, and extroversion.

	Neuroticism	n	Extroversion	n
FMS score	.201	60	.015	60
SSQ total score	.246	60	-.062	60
SSQ Nausea	.156	60	-.102	60
SSQ Disorientation	.255	60	.011	60
SSQ Oculomotor	.291*	60	-.055	60
FSSQ total score	.222	54	.040	54
FSSQ Nausea	.175	54	-.062	54
FSSQ Disorientation	.146	54	.046	54
FSSQ Oculomotor	.250	54	.090	54

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.9: Experiment 2 results: correlations between FMS, SSQ, FSSQ, gender, and vision correction.

	Gender	n	Vision correction	n
FMS score	.413**	62	.183	62
SSQ total score	.404**	62	.081	62
SSQ Nausea	.389**	62	.139	62
SSQ Disorientation	.401**	62	.139	62
SSQ Oculomotor	.311*	62	-.026	62
FSSQ total score	.240	56	.033	56
FSSQ Nausea	.257	56	.031	56
FSSQ Disorientation	.183	56	.089	56
FSSQ Oculomotor	.213	56	.022	56

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.10: Experiment 2 results: correlations between FMS, SSQ, FSSQ, sleep deprivation, physical activity, and sense of presence.

	Sleep deprivation	n	Physical activity	n	Presence	n
FMS score	.209	62	.045	61	-.077	62
SSQ total score	.192	62	.138	61	-.063	62
SSQ Nausea	.101	62	.024	61	-.056	62
SSQ Disorientation	.195	62	.173	61	-.038	62
SSQ Oculomotor	.238	62	.242	61	-.076	62
FSSQ total score	.156	56	.110	55	-.128	56
FSSQ Nausea	.161	56	.046	55	-.044	56
FSSQ Disorientation	.032	56	.135	55	-.095	56
FSSQ Oculomotor	.174	56	.128	55	-.218	56

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.11: Experiment 2 results: correlations between FMS, SSQ, FSSQ, valence and arousal before driving simulation.

	Valence	n	Arousal	n
FMS score	.078	62	.235	62
SSQ total score	-.044	62	.207	62
SSQ Nausea	.-.011	62	.172	62
SSQ Disorientation	-.104	62	.155	62
SSQ Oculomotor	-.070	62	.270*	62
FSSQ total score	-.097	56	.123	56
FSSQ Nausea	-.085	56	.058	56
FSSQ Disorientation	.014	56	.044	56
FSSQ Oculomotor	-.161	56	.142	56

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.12: Experiment 2 results: correlations between MSSQ, FMS, SSQ, FSSQ scores, SCL difference (SCLD), and SCL record.

	SCL difference (Spearman)	n	SCL record (Kendall)	n
FMS score	.108	39	.060	49
SSQ total score	.223	39	.191	49
SSQ Nausea	.227	39	.218*	49
SSQ Disorientation	.269	39	.197	49
SSQ Oculomotor	.099	39	.110	49
FSSQ total score	.040	36	.117	45
FSSQ Nausea	.096	36	.129	45
FSSQ Disorientation	.115	36	.111	45
FSSQ Oculomotor	.087	36	.135	45

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.13: Experiment 3 results: correlations between FMS, SSQ, FSSQ scores, gender, and vision correction.

	Gender	n	Vision correction	n
FMS Score	.230	63	-.021	66
SSQ total score	.395**	66	.103	66
SSQ Nausea	.404**	66	.138	66
SSQ Disorientation	.294*	66	.092	66
SSQ Oculomotor	.359**	66	.035	66
FSSQ total score	.463**	59	-.107	59
FSSQ Nausea	.474**	59	-.072	59
FSSQ Disorientation	.402**	59	-.099	59
FSSQ Oculomotor	.440**	59	-.133	59

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.14: Experiment 3 results: correlation between FMS, SSQ, FSSQ scores, sleep deprivation, and physical activity.

	Sleep deprivation	n	Physical activity	n
FMS score	.000	66	-.010	66
SSQ total score	.060.	66	0.54	66
SSQ Nausea	.094	66	-.055	66
SSQ Disorientation	.038	66	.098	66
SSQ Oculomotor	.048	66	.161	66
FSSQ total score	.155	59	0.41	59
FSSQ Nausea	.182	59	0.60	59
FSSQ Disorientation	.109	59	.153	59
FSSQ Oculomotor	.156	59	-.007	59

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.15: Experiment 3 results: correlations between FMS, SSQ, FSSQ scores, arousal, and valence before the driving simulation.

	Arousal	n	Valence	n
FMS Score	.147	66	.006	66
SSQ total score	.187	66	.003	66
SSQ Nausea	.188	66	.029	66
SSQ Disorientation	.201	66	-.031	66
SSQ Oculomotor	.155	66	-.034	66
FSSQ total score	.047	59	-.148	59
FSSQ Nausea	.033	59	-.120	59
FSSQ Disorientation	.095	59	-.102	59
FSSQ Oculomotor	.063	59	-.218	59

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.16: Experiment 3 results: correlations between FMS, SSQ, FSSQ scores, SCL difference (SCLD), and SCL record.

	SCL difference	n	SCL record	n
FMS Score	.174	26	.088	44
SSQ total score	.069	26	.153	44
SSQ Nausea	-.132	26	.121	44
SSQ Disorientation	.038	26	.133	44
SSQ Oculomotor	.292	26	.220	44
FSSQ total score	.214	24	.179	39
FSSQ Nausea	.138	24	.188	39
FSSQ Disorientation	.175	24	.190	39
FSSQ Oculomotor	.298	24	.139	39

*Note. *p< .05, **p< .01, ***p< .001.*

Table F.17: Experiment 3 results: correlations between FMS, SSQ, FSSQ, RR difference (RRD), RR record, and sense of presence.

Pearson's correlation

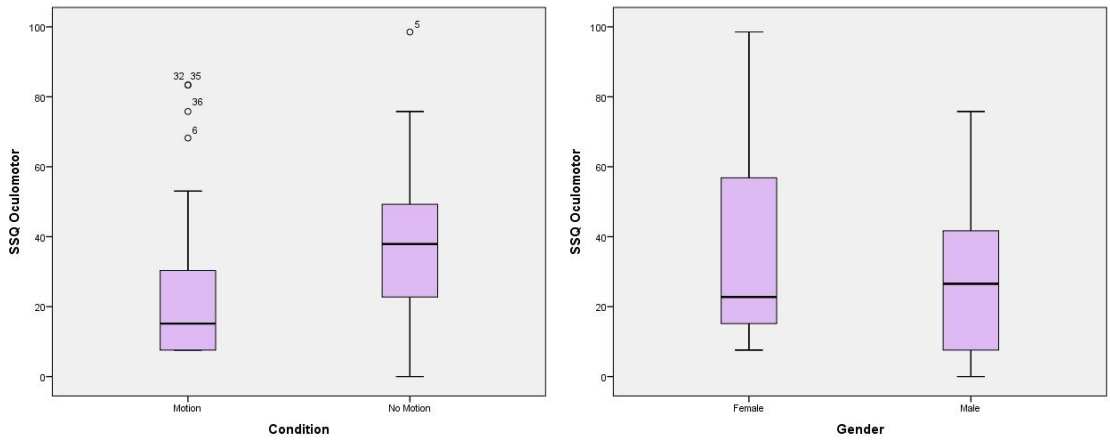
	RR Difference	n	RR record	n	Presence	n
FMS score	-.097	22	-	-	-	-
SSQ total score	-.083	22	-	-	.014	66
SSQ Oculomotor	.183	22	-	-	-	-
SSQ Nausea	-	-	-	-	.018	66

Spearman's correlation

	RR Difference	n	RR record	n	Presence	n
FMS score	-	-	-.167	43	.116	66
SSQ total score	-	-	-.223	43	-	-
SSQ Nausea	-.304	22	-.349*	43	-	-
SSQ Disorientation	-.044	22	-.173	43	.053	66
SSQ Oculomotor	-	-	-.044	43	-.035	66
FSSQ total score	.217	19	-.132	39	-.067	59
FSSQ Nausea	.148	19	-.269	39	-.049	59
FSSQ Disorientation	.176	19	-.181	39	-.058	59
FSSQ Oculomotor	.124	19	.004	39	-.127	59

*Note. *p< .05, **p< .01, ***p< .001.*

G Experiments Graphs



(a) (b)

Figure G.1: Experiment 1: Box plots of SSQ Oculomotor score for motion (a) and gender (b).

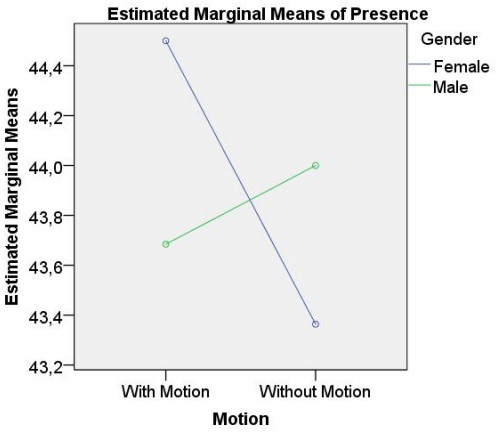


Figure G.2: Experiment 1: Plots of presence score over motion and gender.

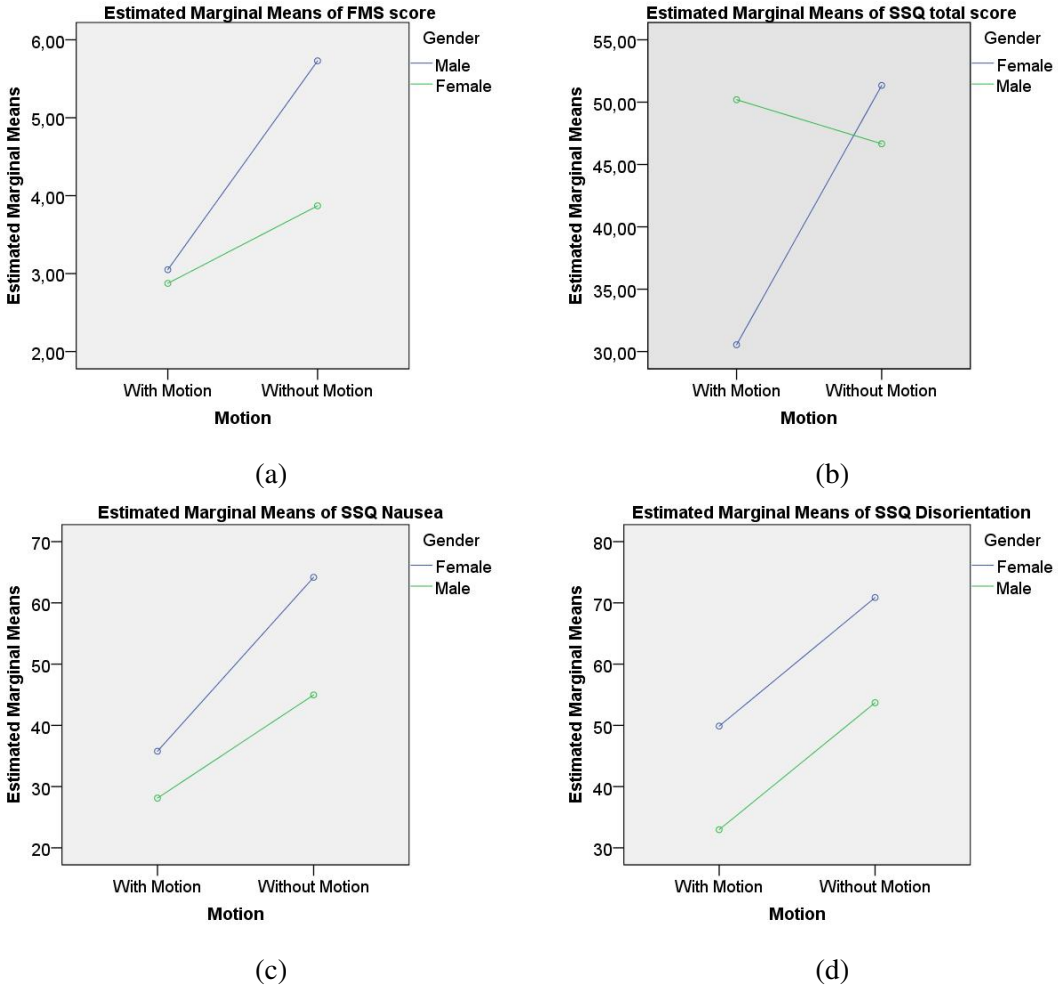


Figure G.3: Experiment 1: Plots of FMS (a), SSQ total score (b), SSQ Nausea score (c), SSQ Disorientation (d) over motion and gender.

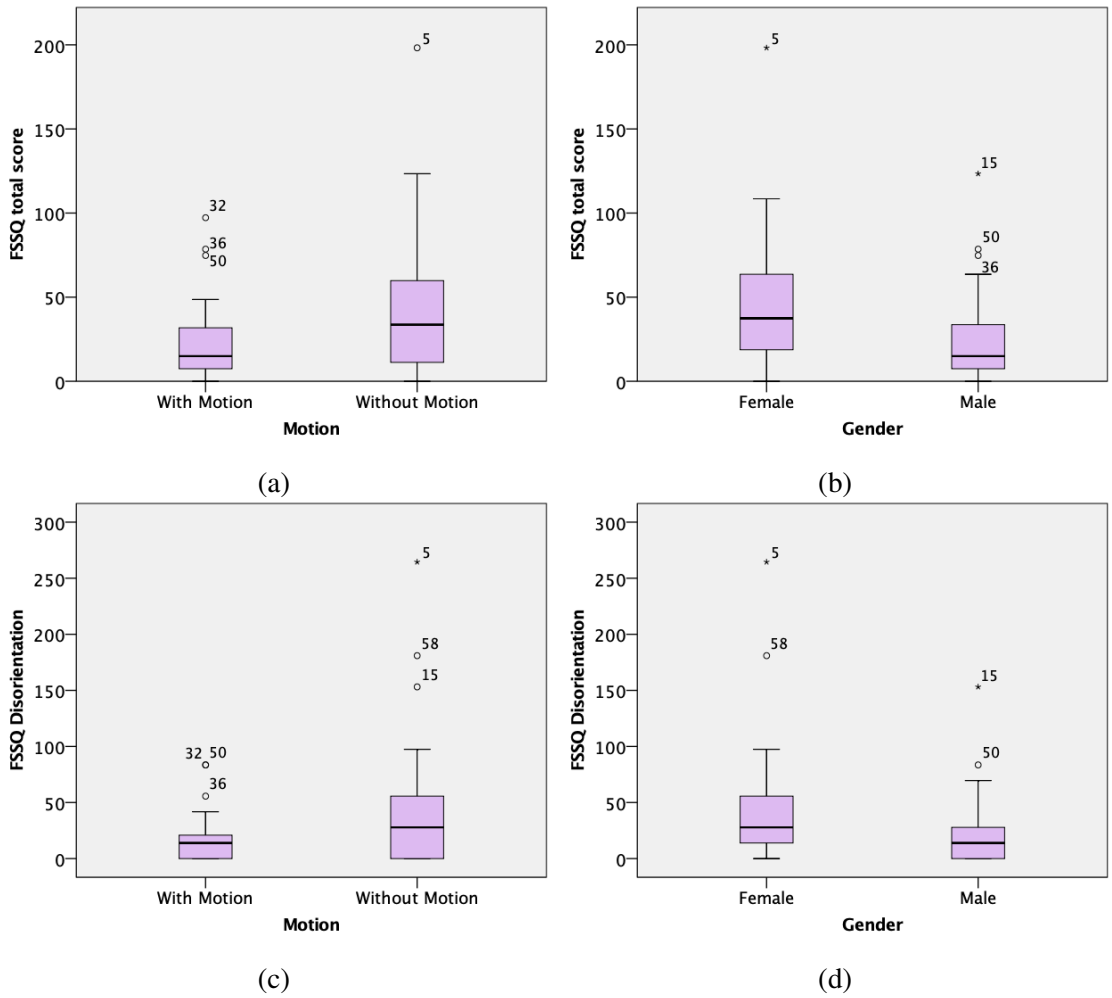


Figure G.4: Experiment 1: Plots of FSSQ total score for motion (a) and for gender (b), FSSQ Disorientation score for motion (c) and for gender (d).

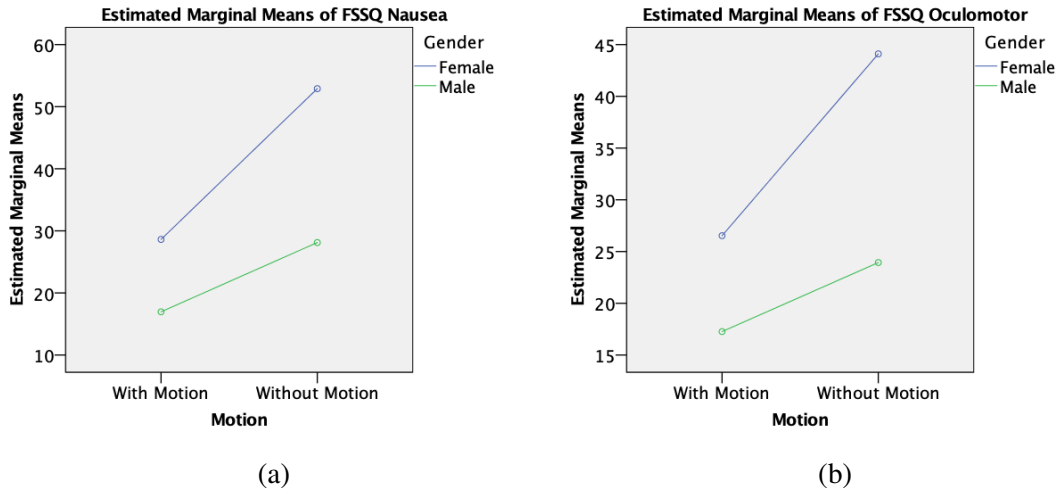
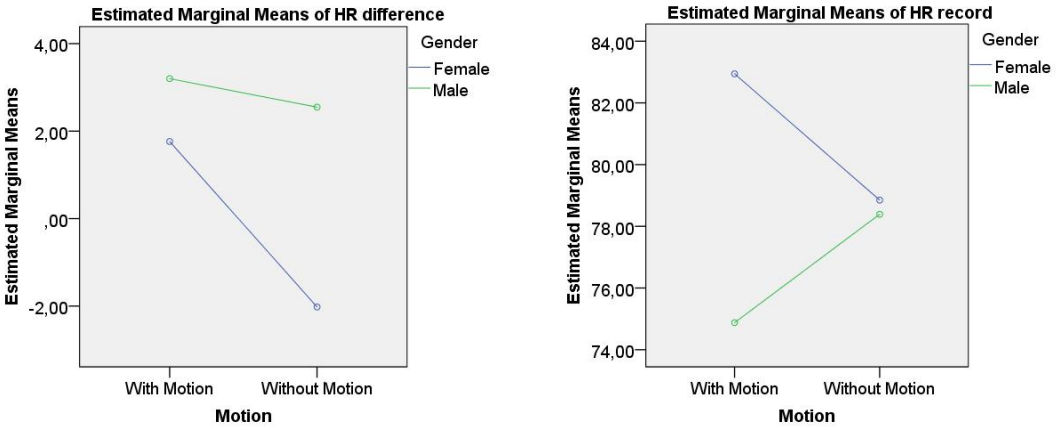
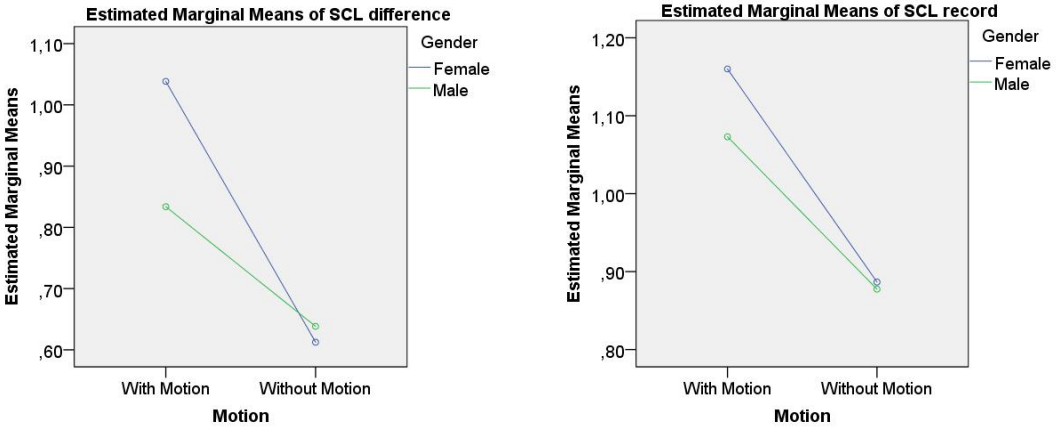


Figure G.5: Experiment 1: Plots of FSSQ Nausea score (a) and FSSQ Oculomotor score (b) over motion and gender.



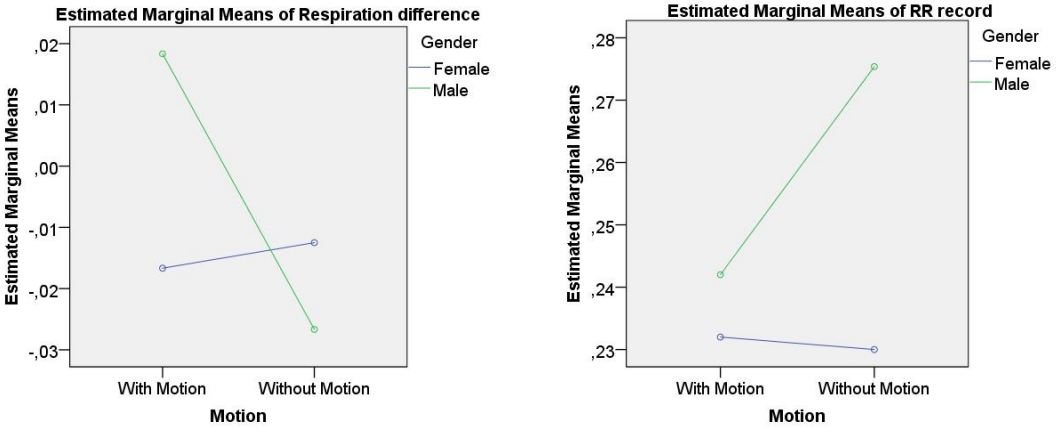
(a)

(b)



(c)

(d)



(e)

(f)

Figure G.6: Experiment 1: Plots of HR difference (a), HR record (b), SCL difference (c), SCL record (d), RR difference (e), and RR record (f) for factors motion and gender.

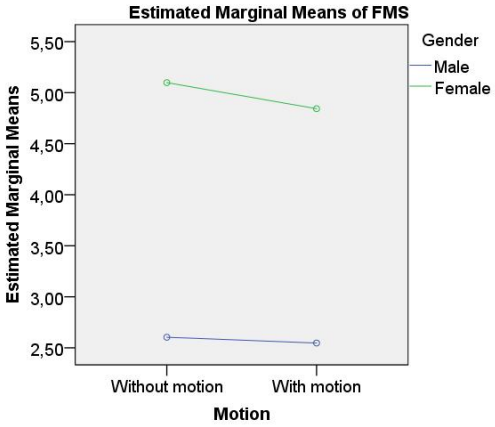


Figure G.7: Experiment 2: Plot of estimated marginal means of FMS mean score over motion and gender.

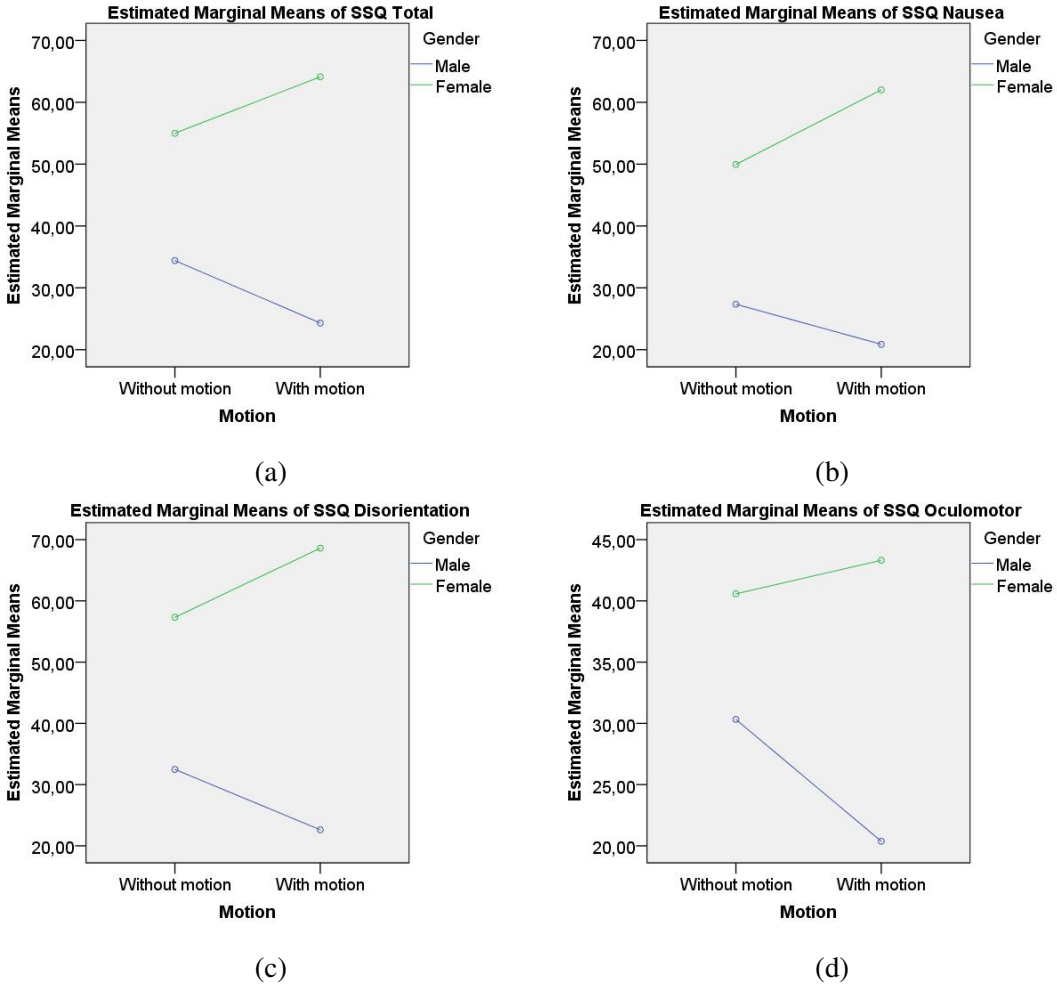


Figure G.8: Experiment 2: Plots of SSQ total score (a), SSQ Nausea (b), SSQ Disorientation (c), and SSQ Oculomotor (d) score over motion and gender.

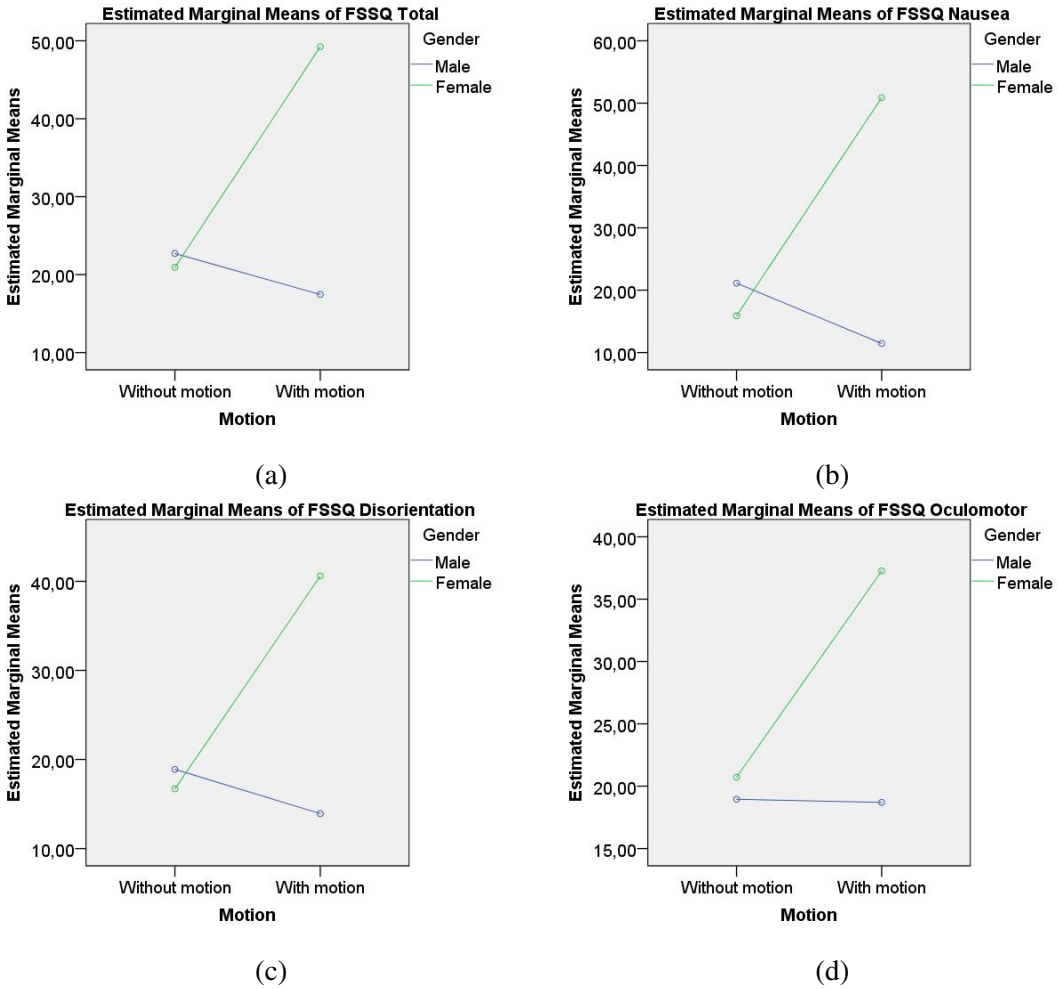


Figure G.9: Experiment 2: Plot of estimated marginal means of FSSQ total score (a), FSSQ Nausea (b), FSSQ Disorientation (c), and FSSQ Oculomotor (d) score over motion and gender.

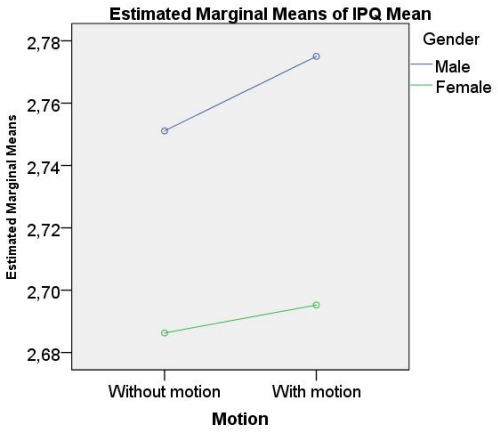


Figure G.10: Experiment 2: Plot of presence score over motion and gender.

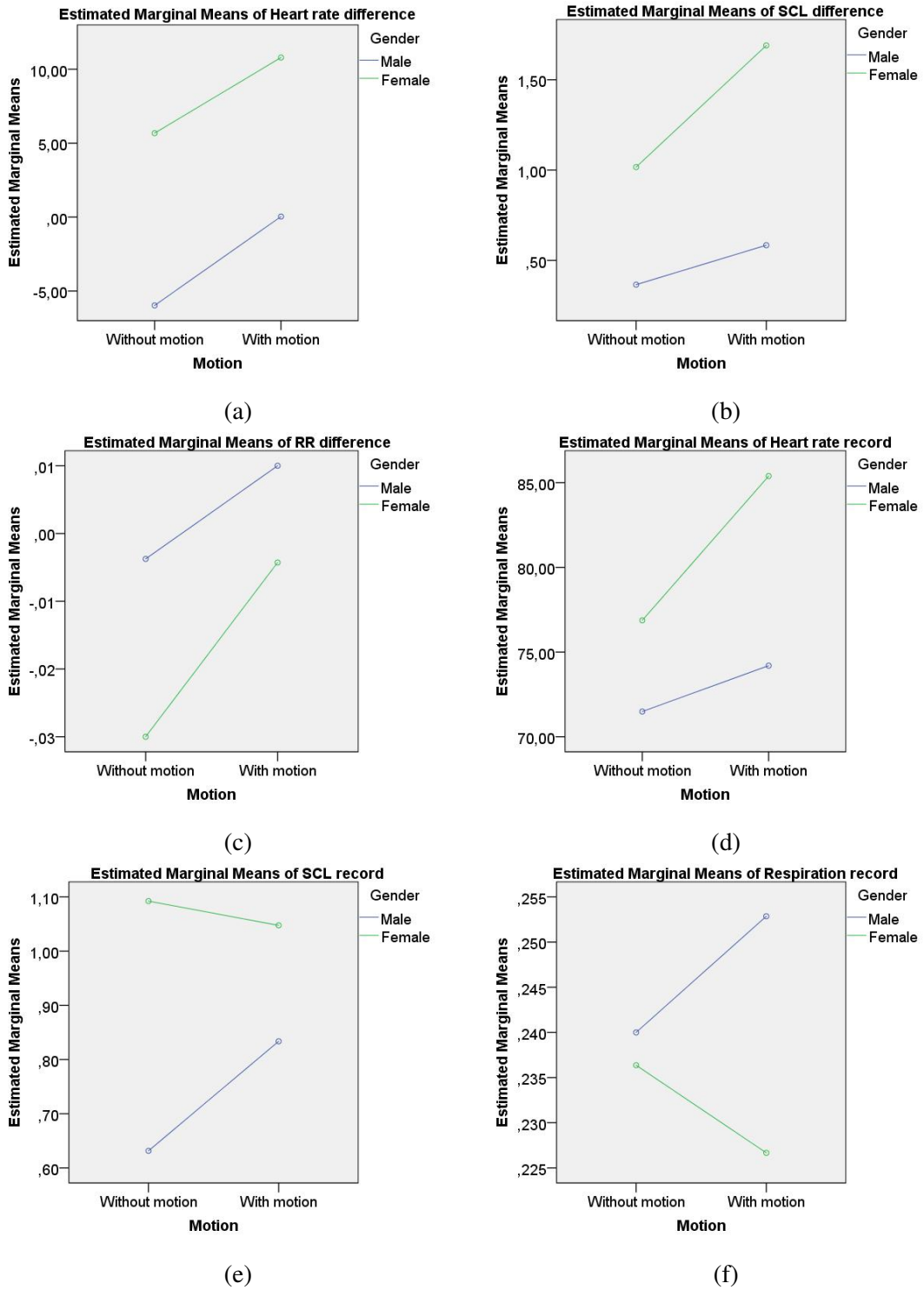


Figure G.11: Experiment 2: Plot of estimated marginal means of HRD (a), SCLD (b), RRD (c), HR record (d), SCL record (e), and RR record (f) over motion and gender.

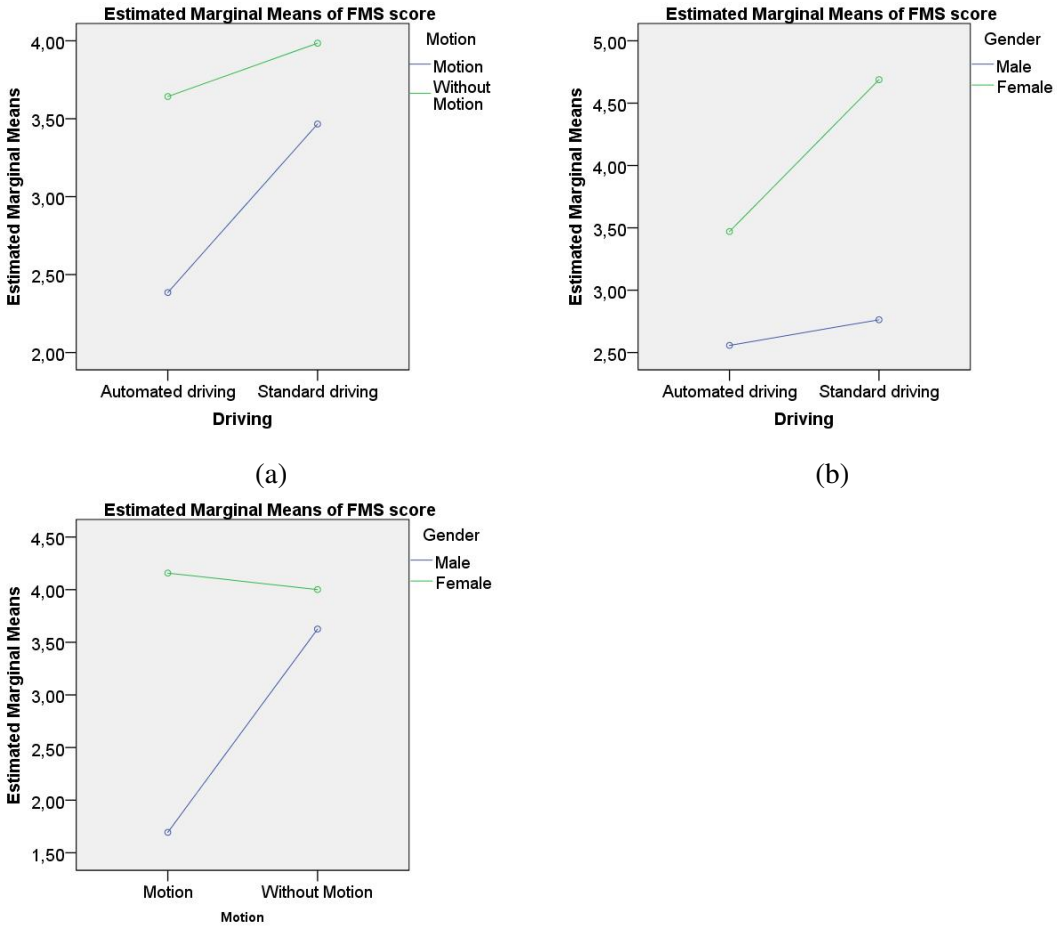


Figure G.12: Experiment 3: Plots of FMS score over drive and motion (a), drive and gender (b), motion and gender (c).

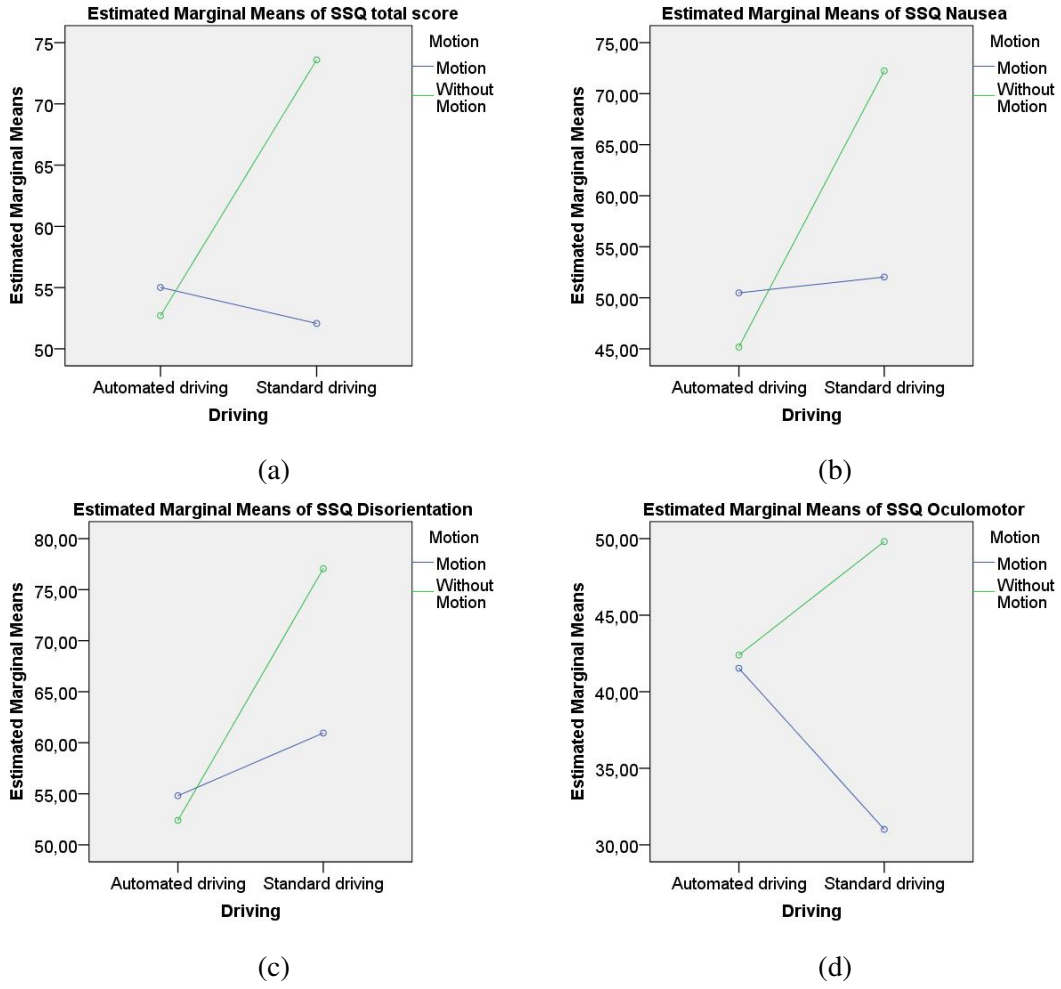


Figure G.13: Experiment 3: Plots of SSQ total score (a), SSQ Nausea score (b), SSQ Disorientation (c), and SSQ Oculomotor score (d) over drive and motion.

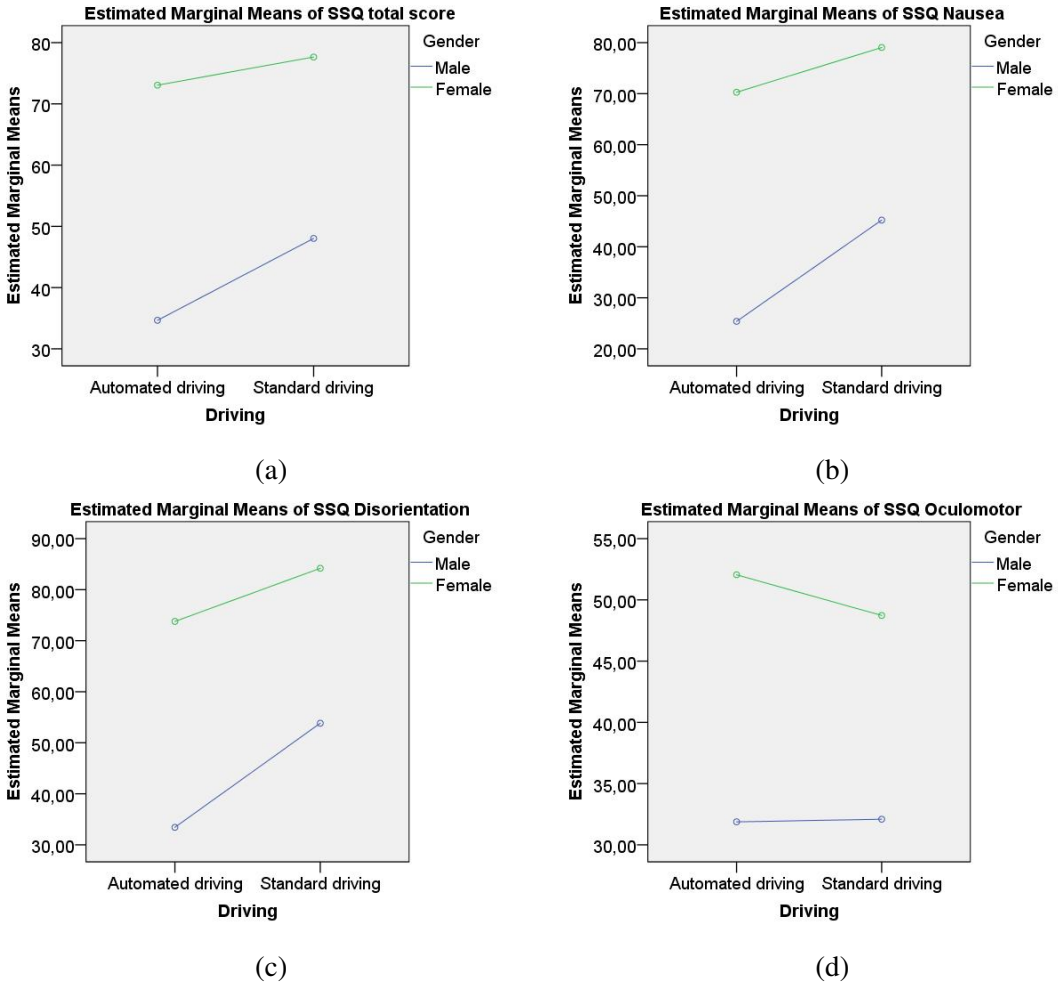


Figure G.14: Experiment 3: Plots of SSQ total score (a), SSQ Nausea score (b), SSQ Disorientation (c), and SSQ Oculomotor score (d) over drive and gender.

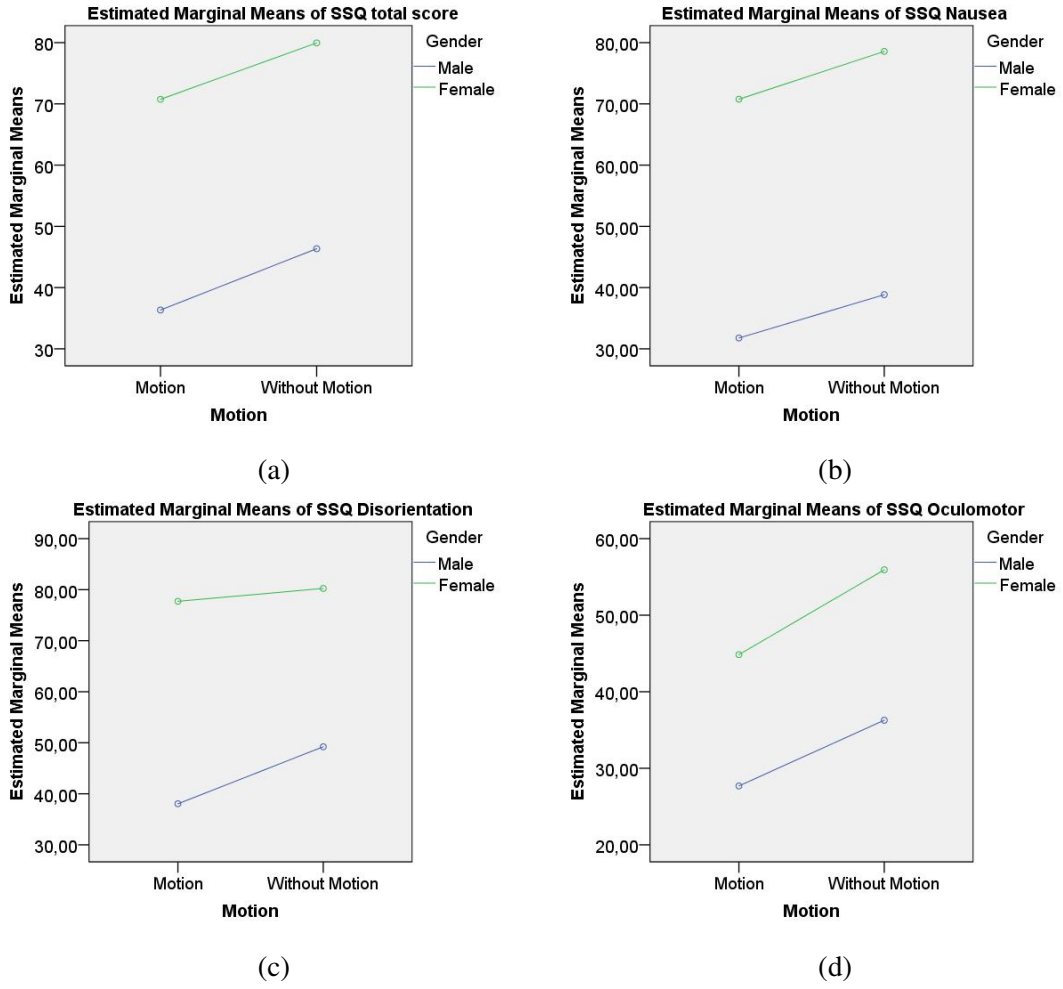


Figure G.15: Experiment 3: Plots of SSQ total score (a), SSQ Nausea score (b), SSQ Disorientation (c), and SSQ Oculomotor score (d) over motion and gender.

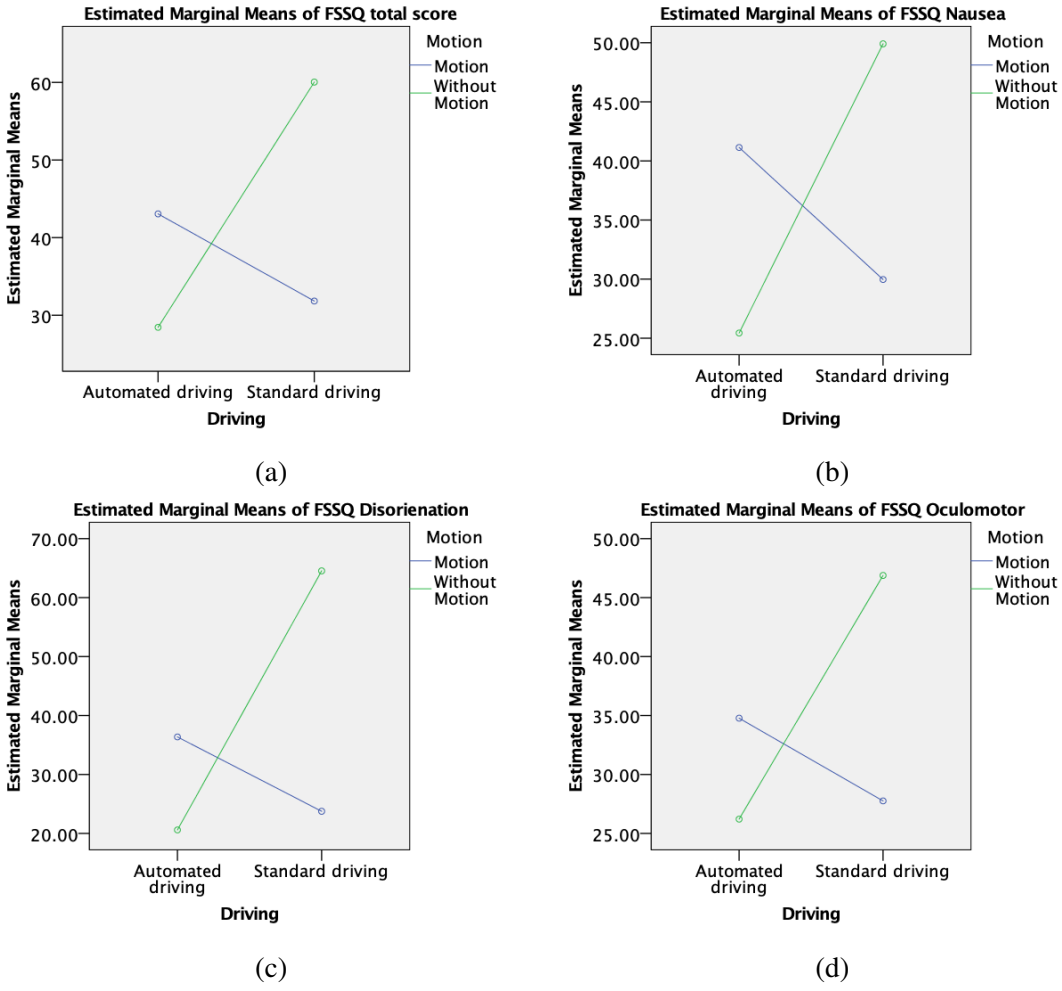


Figure G.16: Experiment 3: Plots of FSSQ total score (a), FSSQ Nausea score (b), FSSQ Disorientation (c), and FSSQ Oculomotor score (d) over drive and motion.

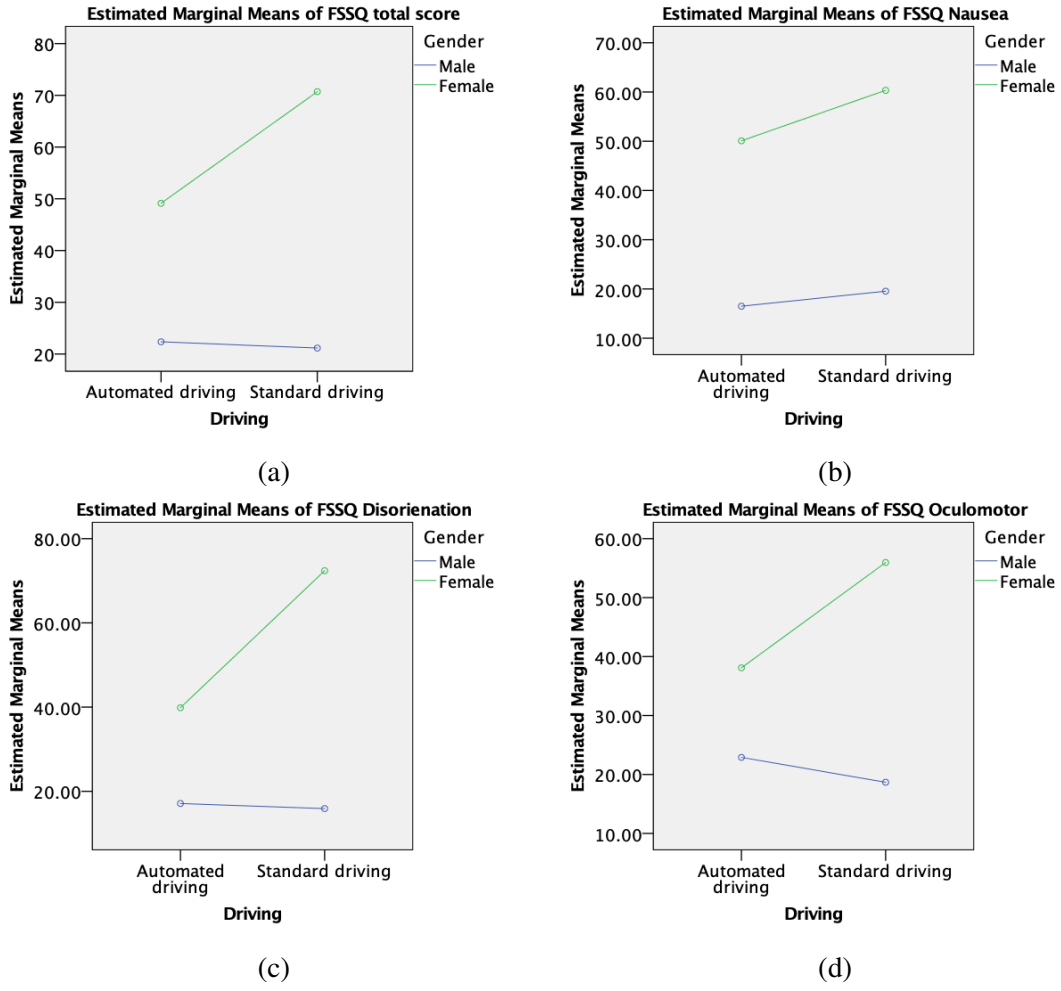


Figure G.17: Experiment 3: Plots of FSSQ total score (a), FSSQ Nausea score (b), FSSQ Disorientation (c), and FSSQ Oculomotor score (d) over drive and gender.

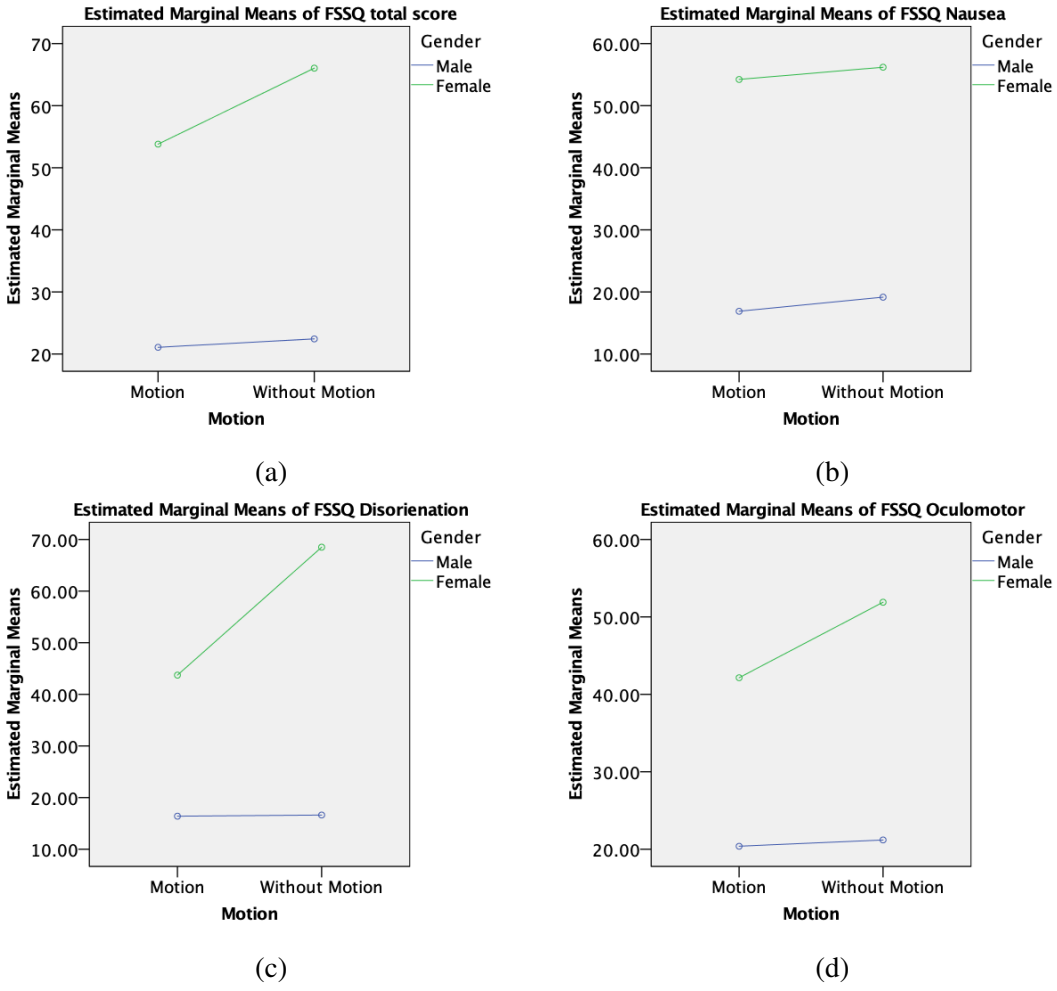


Figure G.18: Experiment 3: Plots of FSSQ total score (a), FSSQ Nausea score (b), FSSQ Disorientation (c), and FSSQ Oculomotor score (d) over motion and gender.

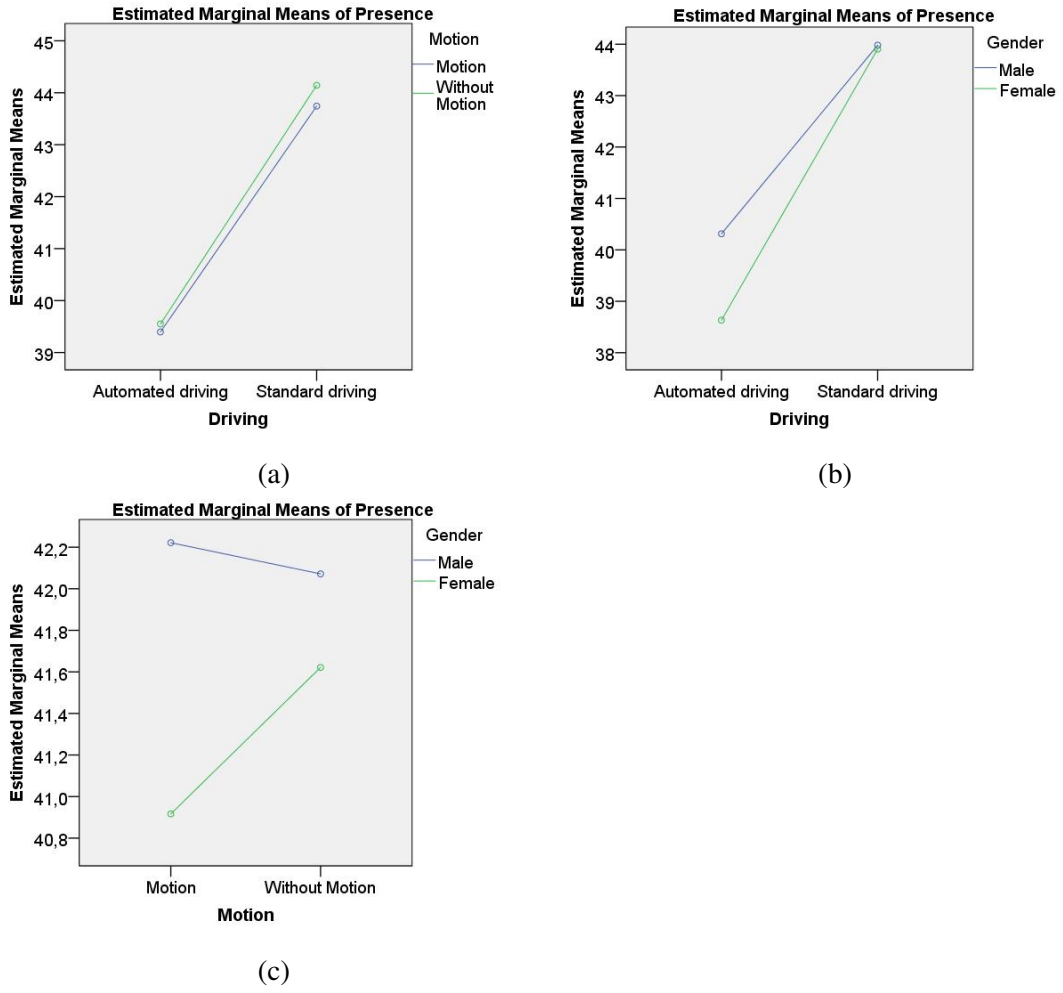


Figure G.19: Experiment 3: Presence score over drive and motion (a), drive and gender (b), and motion and gender (c).

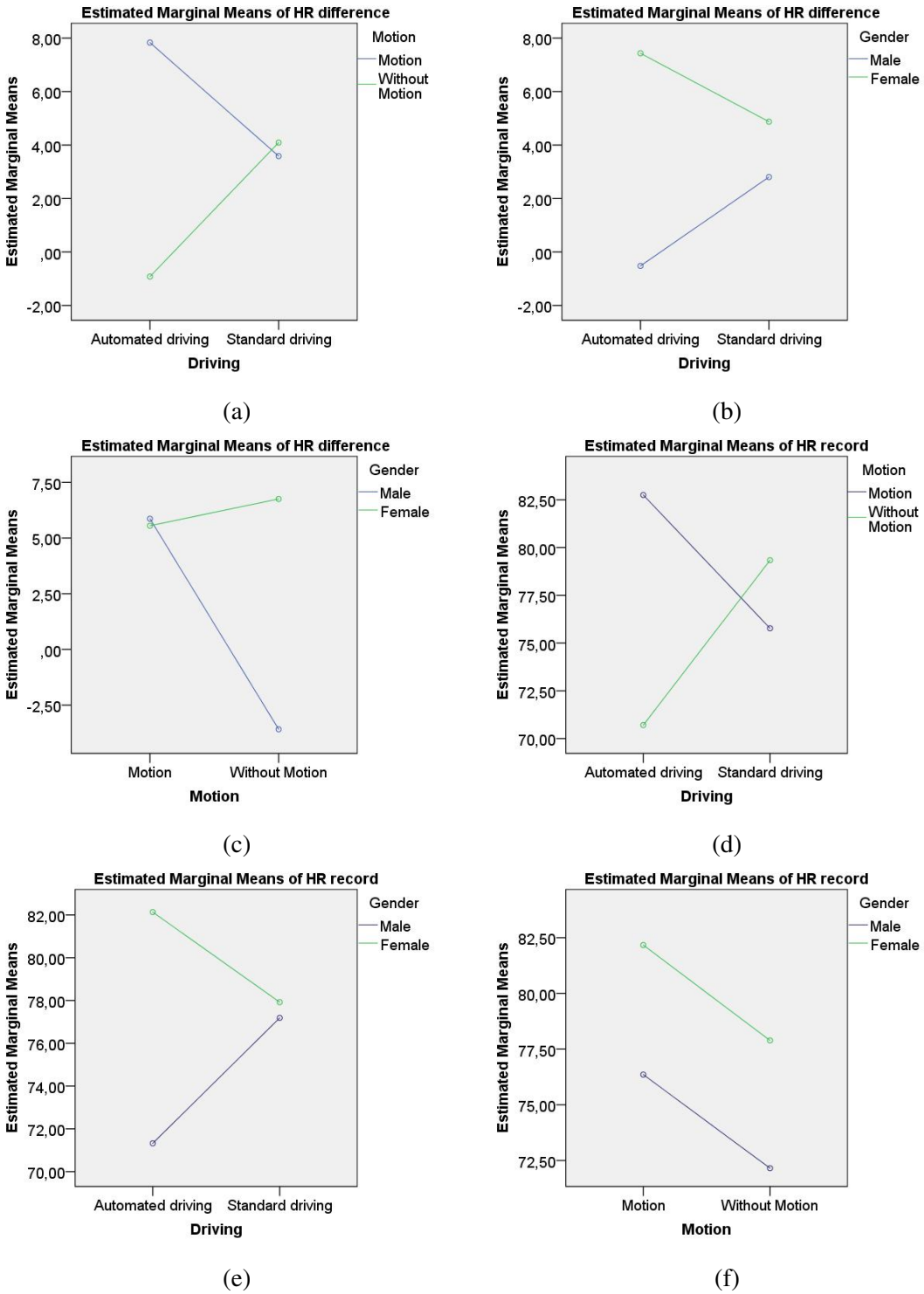
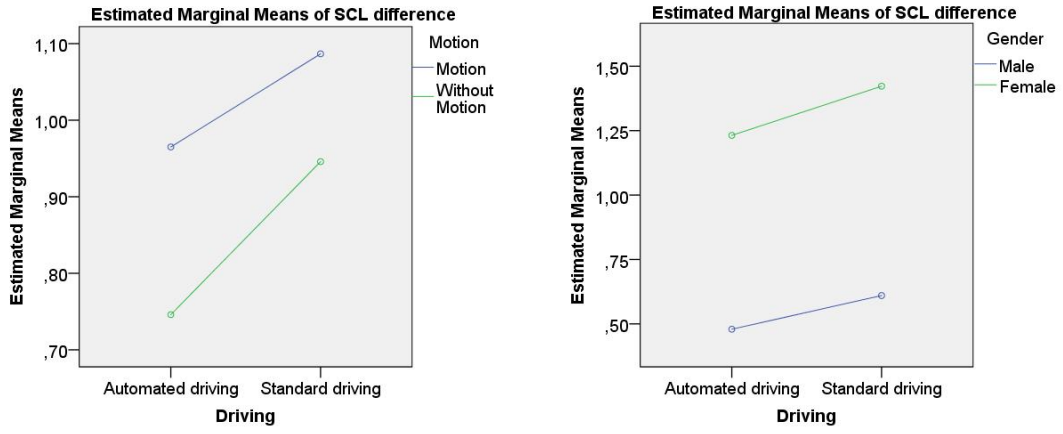
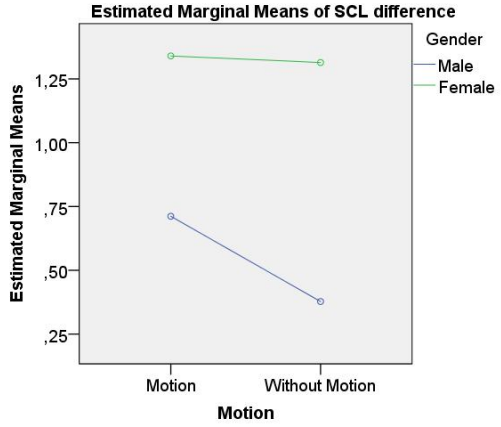


Figure G.20: Experiment 3: Plots of HR difference over drive and motion (a), drive and gender (b), motion and gender (c); and plots of HR record over drive and motion (d), drive and gender (e), motion and gender (f).

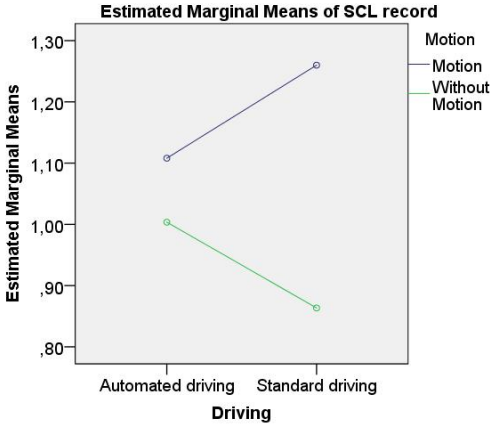


(a)

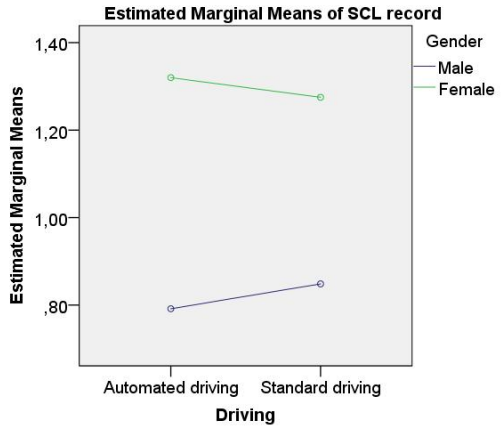
(b)



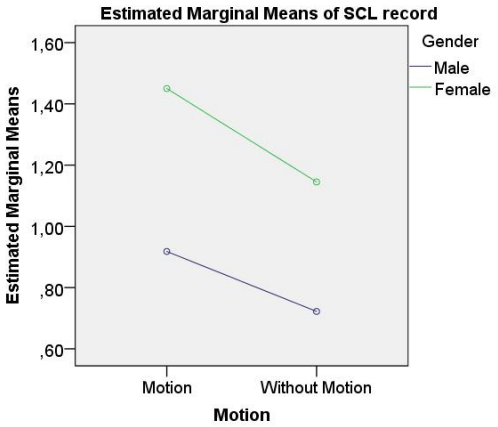
(c)



(d)

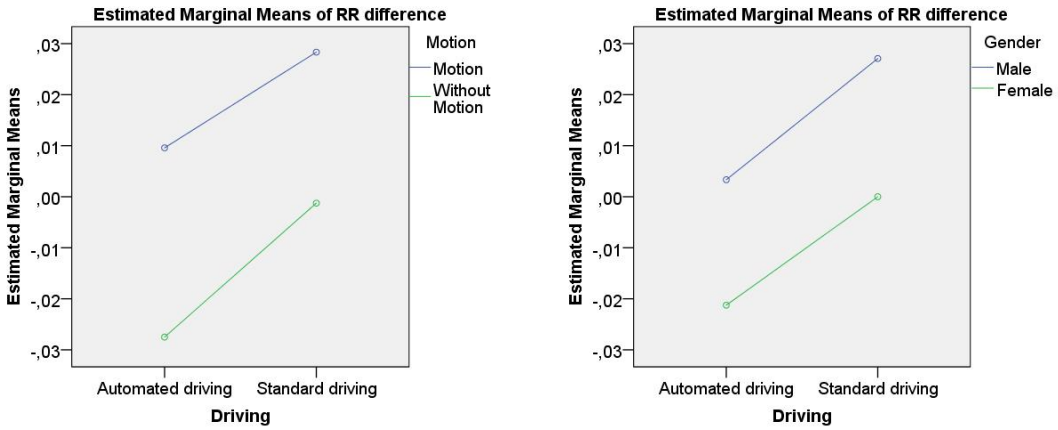


(e)



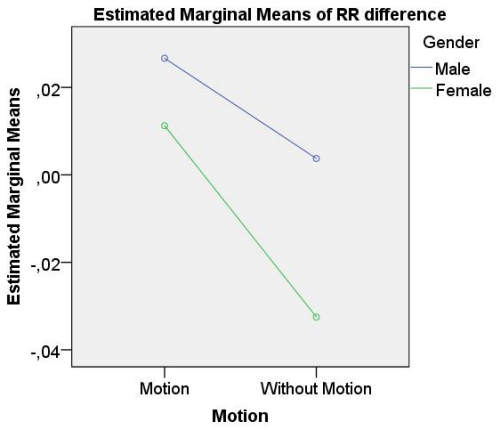
(f)

Figure G.21: Experiment 3: Plots of SCL difference over drive and motion (a), drive and gender (b), motion and gender (c); and plots of SCL record over drive and motion (d), drive and gender (e), motion and gender (f).

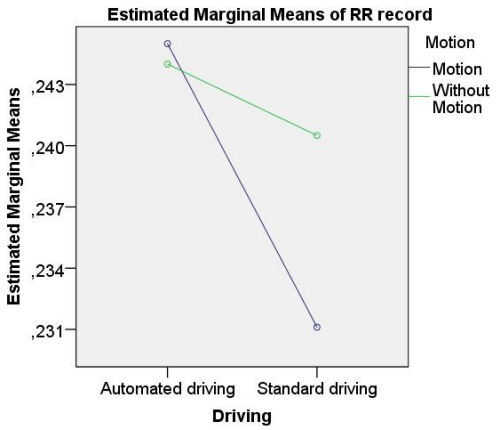


(a)

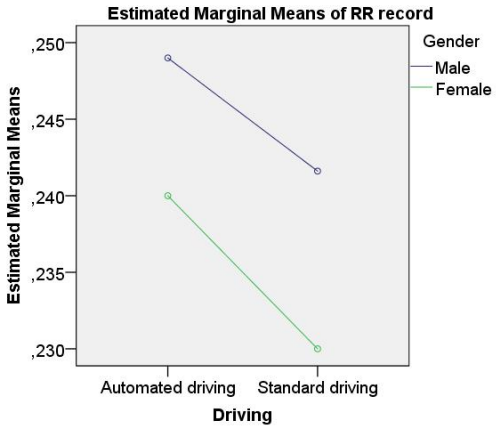
(b)



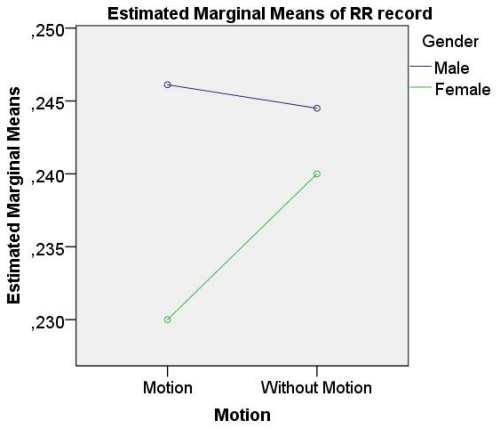
(c)



(d)



(e)



(f)

Figure G.22: Experiment 3: Plots of RR difference over drive and motion (a), drive and gender (b), motion and gender (c); and plots of RR record over drive and motion (d), drive and gender (e), motion and gender (f).

H Prediction Models

Table H.1: Results with under- and over sampling techniques on the Automated driving dataset measured in unweighted models' accuracy.

Classifier	Under sampling	Over sampling
LR	0.58	0.60
LDA	0.62	0.65
KNN	0.61	0.56
CART	0.66	0.61
NB	0.57	0.57
SVM	0.60	0.54
RFC	0.68	0.68

*Note. * Used the undersampled dataset.*
LR - Logistic Regression, LDA - Linear Discriminant Analysis,
KNN - K-nearest neighbors classifier, CART - Decision Trees Classifier,
NB - Naive Bayes, SVM - Support Vector Machine,
RFC - Random Forest Classifier.

Table H.2: Results with under- and over sampling techniques on the Standard driving dataset measured in unweighted models' accuracy.

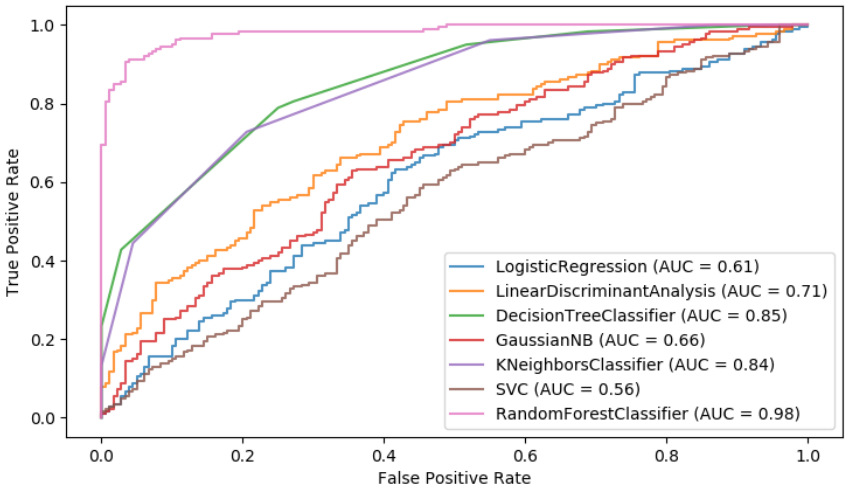
Classifier	Under sampling	Over sampling
LR	0.63	0.64
LDA	0.74	0.72
KNN	0.67	0.71
CART	0.65	0.69
NB	0.70	0.70
SVM	0.72	0.70
RFC	0.72	0.81

*Note. * Used the undersampled dataset.*
LR - Logistic Regression, LDA - Linear Discriminant Analysis,
KNN - K-nearest neighbors classifier, CART - Decision Trees Classifier,
NB - Naive Bayes, SVM - Support Vector Machine,
RFC - Random Forest Classifier.

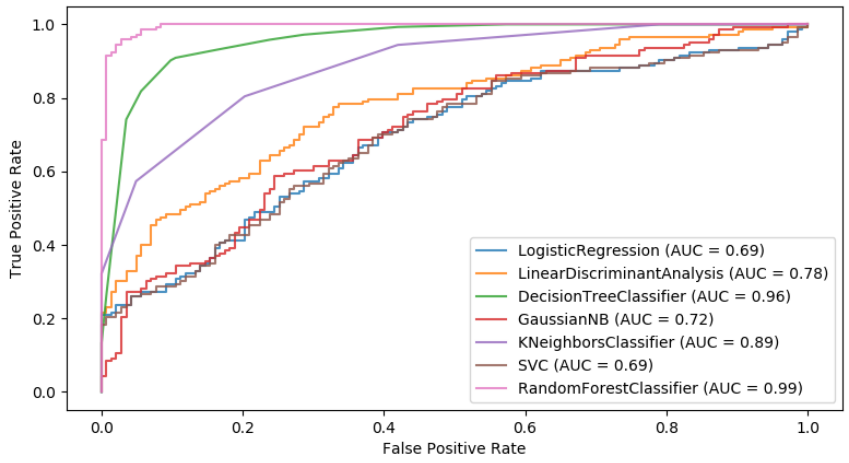
Table H.3: Results with under- and over sampling techniques on the All driving dataset measured in unweighted models' accuracy.

Classifier	Under sampling	Over sampling
LR	0.52	0.52
LDA	0.56	0.58
KNN	0.62	0.68
CART	0.60	0.62
NB	0.57	0.54
SVM	0.54	0.53
RFC	0.65	0.68

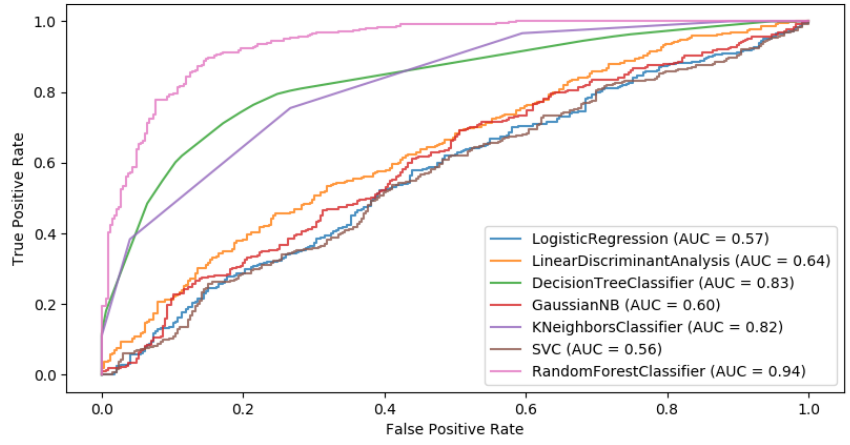
*Note. * Used the undersampled dataset.*
LR - Logistic Regression, LDA - Linear Discriminant Analysis,
KNN - K-nearest neighbors classifier, CART - Decision Trees Classifier,
NB - Naive Bayes, SVM - Support Vector Machine,
RFC - Random Forest Classifier.



(a)



(b)



(c)

Figure H.1: ROC curves including AUC scores for automated dataset (a), standard dataset (b), and all dataset (c).

Bibliography

[1] H. Akiduki, S. Nishiike, H. Watanabe, K. Matsuoka, T. Kubo, and N. Takeda, "Visual-vestibular conflict induced by virtual reality in humans," *Neuroscience letters*, vol. 340, no. 3, pp. 197–200, 2003.

[2] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological measurement*, vol. 28, no. 3, p. R1, 2007.

[3] D. G. Altman and J. M. Bland, "Statistics notes: Diagnostic tests 3: receiver operating characteristic plots," *BMJ*, vol. 309, no. 6948, p. 188, 1994.

[4] S. L. Ames, J. S. Wolffsohn, and N. A. McBrien, "The development of a symptom questionnaire for assessing virtual reality viewing using a head-mounted display," *Optometry and Vision Science*, vol. 82, no. 7, pp. 571–572, 2005.

[5] APA Dictionary of Psychology Online, "Extraversion," https://dictionary.apa.org/extraversion, 2020, Accessed: 6 April 2021.

[6] APA Dictionary of Psychology Online, "Respiration," https://dictionary.apa.org/respiration, 2020, Accessed: 6 April 2021.

[7] B. Arcioni, S. Palmisano, D. Apthorp, and J. Kim, "Postural stability predicts the likelihood of cybersickness in active HMD-based virtual reality," *Displays*, vol. 58, pp. 3–11, 2019.

[8] L. L. Arns and M. M. Cerney, "The relationship between age and incidence of cybersickness among immersive environment users," in *IEEE Proceedings. VR 2005. Virtual Reality, 2005.* IEEE, 2005, pp. 267–268.

[9] B. Aykent, F. Merienne, D. Paillot, and A. Kemeny, "Influence of inertial stimulus on visuo-vestibular cues conflict for lateral dynamics at driving simulators," *Journal of Ergonomics*, vol. 3, pp. 1–8, 2013.

[10] B. Aykent, F. Merienne, C. Guillet, D. Paillot, and A. Kemeny, "Motion sickness evaluation and comparison for a static driving simulator and a dynamic driving simulator," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 228, no. 7, pp. 818–829, 2014.

[11] B. Aykent, Y. Zhao, F. Merienne, and A. Kemeny, "Simulation sickness comparison between a limited field of view virtual reality head mounted display (Oculus) and a medium range field of view static ecological driving simulator (Eco2)," in *Proceedings Driving Simulation Conference Europe 2014*, 2014, pp. 65–71.

[12] R. E. Bailey, J. J. Arthur III, and S. P. Williams, "Latency requirements for head-worn display S/EVS applications," in *Enhanced and Synthetic Vision 2004*. International Society for Optics and Photonics, 2004, pp. 98–109.

[13] S. A. Balk, M. A. Bertola, and V. W. Inman, "Simulator sickness questionnaire: Twenty years later," in *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 2013, pp. 257–263.

[14] W. Barfield, D. Zeltzer, T. Sheridan, and M. Slater, "Presence and performance within virtual environments," *Virtual environments and advanced interface design*, pp. 473–513, 1995.

[15] D. H. Barlow, K. K. Ellard, S. Sauer-Zavala, J. R. Bullis, and J. R. Carl, "The origins of neuroticism," *Perspectives on Psychological Science*, vol. 9, no. 5, pp. 481–496, 2014.

[16] A. Basu, C. Ball, B. Manning, and K. Johnsen, "Effects of user physical fitness on performance in virtual reality," in *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 2016, pp. 233–234.

[17] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *Journal of neuroscience methods*, vol. 190, no. 1, pp. 80–91, 2010.

[18] F. Biocca, "Will simulation sickness slow down the diffusion of virtual environment technology?" *Presence: Teleoperators & Virtual Environments*, vol. 1, no. 3, pp. 334–343, 1992.

[19] F. Bonato, A. Bubka, and L. Alfieri, "Display color affects motion sickness symptoms in an optokinetic drum," *Aviation, space, and environmental medicine*, vol. 75, no. 4, pp. 306–311, 2004.

[20] F. Bonato, A. Bubka, S. Palmisano, D. Phillip, and G. Moreno, "Vection change exacerbates simulator sickness in virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 17, no. 3, pp. 283–292, 2008.

[21] C. T. Bonnet, E. Faugloire, M. A. Riley, B. G. Bardy, and T. A. Stoffregen, "Self-induced motion sickness and body movement during passive restraint," *Ecological Psychology*, vol. 20, no. 2, pp. 121–145, 2008.

[22] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*, 1992, pp. 144–152.

[23] D. Boyd, "Is the oculus rift sexist," *Quartz*, 2014.

[24] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[25] J. J. Braithwaite, D. G. Watson, R. Jones, and M. Rowe, "A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments," *Psychophysiology*, vol. 49, no. 1, pp. 1017–1034, 2013.

[26] M. Braithwaite and B. Braithwaite, "Simulator sickness in an army simulator," *Occupational Medicine*, vol. 40, no. 3, pp. 105–110, 1990.

[27] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[28] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*. CRC press, 1984, pp. 20–23.

[29] Britanica, "Proprioception," https://www.britannica.com/science/proprioception, 2018, Accessed 26 June 2019.

[30] J. O. Brooks, R. R. Goodenough, M. C. Crisler, N. D. Klein, R. L. Alley, B. L. Koon, W. C. Logan Jr, J. H. Ogle, R. A. Tyrrell, and R. F. Wills, "Simulator sickness during driving simulation studies," *Accident Analysis & Prevention*, vol. 42, no. 3, pp. 788–796, 2010.

[31] R. Brown, "Oculus Rift DK1: Full Specification," https://www.vr-compare.com/headset/oculusriftdk1, 2020, Accessed 10 July 2020.

[32] S. Bruck and P. A. Watters, "The factor structure of cybersickness," *Displays*, vol. 32, no. 4, pp. 153–158, 2011.

[33] G. C. Burdea and P. Coiffet, *Virtual reality technology*. John Wiley & Sons, 2003, pp. 65–66.

[34] A. Burg, "Lateral visual field as related to age and sex," *Journal of applied psychology*, vol. 52, no. 1p1, p. 10, 1968.

[35] S. Büttner, "Development of a physiological-signal-acquiring VR driving simulation with environmental events for automated signal annotation," Master's thesis, University of Augsburg, Germany, 2019.

[36] J. G. Casali and L. H. Frank, "Perceptual distortion and its consequences in vehicular simulation: Basic theory and incidence of simulator sickness," *Transportation Research Record*, vol. 1059, pp. 57–65, 1986.

[37] C. H. Chang, W. W. Pan, F. C. Chen, and T. A. Stoffregen, "Console video games, postural activity, and motion sickness during passive restraint," *Experimental Brain Research*, vol. 229, no. 2, pp. 235–242, 2013.

[38] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[39] R. Chen, A. Ho, F. Lor, and R. So, "Enhancing the predictive power of cybersickness dose value (CSDV) to include effects of field-of-view and binocular views," in *Proceedings of the 7th International Conference on Working With Computing Systems*, 2004.

[40] S. A. Clemes and P. A. Howarth, "The menstrual cycle and susceptibility to virtual simulation sickness," *Journal of biological rhythms*, vol. 20, no. 1, pp. 71–82, 2005.

[41] J. Clifton and S. Palmisano, "Effects of steering locomotion and teleporting on cybersickness and presence in HMD-based virtual reality," *Virtual Reality*, vol. 24, no. 3, pp. 453–468, 2020.

[42] S. V. Cobb, S. Nichols, A. Ramsey, and J. R. Wilson, "Virtual reality-induced symptoms and effects (VRISE)," *Presence: Teleoperators & Virtual Environments*, vol. 8, no. 2, pp. 169–186, 1999.

[43] J. Cohen, *Statistical power analysis for the behavioral sciences. 2nd ed.* Lawrence Earlbam Associates, 1988.

[44] W. E. Collins and J. M. Lentz, "Some psychological correlates of motion sickness susceptibility," *Aviat Space Environ Med.*, vol. 48, pp. 587–594, 1977.

[45] F. Colombet, A. Kemeny, and P. George, "Motion sickness comparison between a CAVE environment and a HMD," in *Driving Simulation Conference Europe*, 2016, pp. 201–208.

[46] J. J. Cummings and J. N. Bailenson, "How immersive is enough? A meta-analysis of the effect of immersive technology on user presence," *Media Psychology*, vol. 19, no. 2, pp. 272–309, 2016.

[47] C. Curry, R. Li, N. Peterson, and T. A. Stoffregen, "Cybersickness in virtual reality head-mounted displays: examining the influence of sex differences and vehicle control," *International Journal of Human–Computer Interaction*, vol. 36, no. 12, pp. 1161–1167, 2020.

[48] R. Curry, B. Artz, L. Cathey, P. Grant, and J. Greenberg, "Kennedy SSQ results: fixed-vs motion-based FORD simulators," in *Proceedings of DSC*, 2002, pp. 289–300.

[49] M. K. Curtis, K. Dawson, K. Jackson, L. Litwin, C. Meusel, M. C. Dorneich, S. B. Gilbert, J. Kelly, R. Stone, and E. Winer, "Mitigating visually induced motion sickness: A virtual hand-eye coordination task," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, 2015, pp. 1839–1843.

[50] M. Czerwinski, D. S. Tan, and G. G. Robertson, "Women take a wider view," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2002, pp. 195–202.

[51] J. Dahlman, A. Sjörs, T. Ledin, and T. Falkmer, "Could sound be used as a strategy for reducing symptoms of perceived motion sickness?" *Journal of neuroengineering and rehabilitation*, vol. 5, no. 1, p. 35, 2008.

[52] S. Davis, K. Nesbitt, and E. Nalivaiko, "A systematic review of cybersickness," in *Proceedings of the 2014 Conference on Interactive Entertainment*. ACM, 2014, pp. 1–9.

[53] D. A. Del Cid, D. Larranaga, M. Leitao, R. L. Mosher, S. R. Berzenski, V. Gandhi, and S. A. Drew, "Exploratory factor analysis and validity of the virtual reality symptom questionnaire and computer use survey," *Ergonomics*, vol. 64, no. 1, pp. 69–77, 2021.

[54] M. S. Dennison and M. D'Zmura, "Cybersickness without the wobble: Experimental results speak against postural instability theory," *Applied ergonomics*, vol. 58, pp. 215–223, 2017.

[55] S. Di Girolamo, P. Picciotti, B. Sergi, W. Di Nardo, G. Paludetti, and F. Ottaviani, "Vestibulo-ocular reflex modification after virtual environment exposure," *Acta otolaryngologica*, vol. 121, no. 2, pp. 211–215, 2001.

[56] C. Diels and J. E. Bos, "Self-driving carsickness," *Applied ergonomics*, vol. 53, pp. 374–382, 2016.

[57] C. Diels, K. Ukai, and P. A. Howarth, "Visually induced motion sickness with radial displays: effects of gaze angle and fixation," *Aviation, space, and environmental medicine*, vol. 78, no. 7, pp. 659–665, 2007.

[58] J. E. Domeyer, N. D. Cassavaugh, and R. W. Backs, "The use of adaptation to reduce simulator sickness in driving assessment and research," *Accident Analysis & Prevention*, vol. 53, pp. 127–132, 2013.

[59] X. Dong, K. Yoshida, and T. A. Stoffregen, "Control of a virtual vehicle influences postural activity and motion sickness," *Journal of Experimental Psychology: Applied*, vol. 17, no. 2, p. 128, 2011.

[60] M. H. Draper, "Can your eyes make you sick? Investigating the relationship between the vestibulo-ocular reflex and virtual reality," *Human Interface Technology Laboratory Technical Report*, 1996.

[61] H. B. Duh, D. E. Parker, and T. A. Furness, "An independent visual background reduced simulator sickness in a driving simulator," *Presence: Teleoperators & Virtual Environments*, vol. 13, no. 5, pp. 578–588, 2004.

[62] S. D'Amour, J. E. Bos, and B. Keshavarz, "The efficacy of airflow and seat vibration on reducing visually induced motion sickness," *Experimental brain research*, vol. 235, no. 9, pp. 2811–2820, 2017.

[63] L. Eudave and M. Valencia, "Physiological response while driving in an immersive virtual environment," in *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, May 2017, pp. 145–148.

[64] A. S. Fernandes and S. K. Feiner, "Combating VR sickness through subtle dynamic field-of-view modification," in *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 2016, pp. 201–210.

[65] M. Fetter, "Vestibulo-ocular reflex," in *Neuro-Ophthalmology*. Karger Publishers, 2007, vol. 40, pp. 35–51.

[66] A. Field, *Discovering statistics using IBM SPSS statistics*. SAGE Publications Ltd., 2013.

[67] D. L. Fisher, M. Rizzo, J. Caird, and J. D. Lee, *Handbook of driving simulation for engineering, medicine, and psychology*. CRC Press, 2011, pp. 14–17.

[68] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of eugenics*, vol. 7, no. 2, pp. 179–188, 1936.

[69] E. Fix and J. L. Hodges, "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties," *International Statistical Review / Revue Internationale de Statistique*, vol. 57, no. 3, pp. 238–247, 1989.

[70] K. M. Flegal, B. K. Kit, and B. I. Graubard, "Body mass index categories in observational studies of weight and risk of death," *American journal of epidemiology*, vol. 180, no. 3, pp. 288–296, 2014.

[71] G. Gálvez-García, J. Albayay, L. Rehbein, and F. Tornay, "Mitigating simulator adaptation syndrome by means of tactile stimulation," *Applied ergonomics*, vol. 58, pp. 13–17, 2017.

[72] A. Garcia, C. Baldwin, and M. Dworsky, "Gender differences in simulator sickness in fixed-versus rotating-base driving simulator," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, 2010, pp. 1551–1555.

[73] A. Garcia-Agundez, C. Reuter, P. Caserman, R. Konrad, and S. Göbel, "Identifying cybersickness through heart rate variability alterations," *International Journal of Virtual Reality*, vol. 19, no. 1, pp. 1–10, 2019.

[74] A. M. Gavgani, D. M. Hodgson, and E. Nalivaiko, "Effects of visual flow direction on signs and symptoms of cybersickness," *PloS one*, vol. 12, no. 8, p. e0182790, 2017.

[75] A. I. Glendon, L. Dorn, D. R. Davies, G. Matthews, and R. G. Taylor, "Age and gender differences in perceived accident likelihood and driver competences," *Risk Analysis*, vol. 16, no. 6, pp. 755–762, 1996.

[76] J. M. Goldberg, V. J. Wilson, D. E. Angelaki, and K. E. Cullen, *The vestibular system: a sixth sense*. Oxford University Press, 2012.

[77] J. F. Golding, "Predicting individual differences in motion sickness susceptibility by questionnaire," *Personality and Individual Differences*, vol. 41, no. 2, pp. 237–248, 2006.

[78] D. A. Graeber and K. M. Stanney, "Gender differences in visually induced motion sickness," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, 2002, pp. 2109–2113.

[79] H. Gray, *Gray's anatomy. 15th Edition. Drawings by HV Carter*. Barnes & Noble, 2010.

[80] A. Grøntved, T. Brask, J. Kambskard, and E. Hentzer, "Ginger root against seasickness: A conctrolled trial on the open sea," *Acta oto-laryngologica*, vol. 105, no. 1-2, pp. 45–49, 1988.

[81] J. Gugenheimer, E. Stemasov, J. Frommel, and E. Rukzio, "ShareVR: Enabling co-located experiences for virtual reality between HMD and non-HMD users," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 4021–4033.

[82] G. Gálvez-García, "A comparison of techniques to mitigate simulator adaptation syndrome," *Ergonomics*, vol. 58, no. 8, pp. 1365–1371, 2015, pMID: 25653095.

[83] P. Hamilton, "Open source ECG analysis," in *Computers in cardiology*. IEEE, 2002, pp. 101–104.

[84] C. Heeter, "Being there: The subjective experience of presence," *Presence: Teleoperators & Virtual Environments*, vol. 1, no. 2, pp. 262–271, 1992.

[85] M. L. Heilig, "El cine del futuro: The cinema of the future," *Presence: Teleoperators & Virtual Environments*, vol. 1, no. 3, pp. 279–294, 1992.

[86] S. Hell and V. Argyriou, "Machine learning architectures to predict motion sickness using a virtual reality rollercoaster simulation tool," in *2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE, 2018, pp. 153–156.

[87] L. J. Hettinger, K. S. Berbaum, R. S. Kennedy, W. P. Dunlap, and M. D. Nolan, "Vection and simulator sickness," *Military psychology*, vol. 2, no. 3, pp. 171–181, 1990.

[88] K. Hill and P. A. Howarth, "Habituation to the side effects of immersion in a virtual environment," *Displays*, vol. 21, no. 1, pp. 25–30, 2000.

[89] W. C. Hixson, J. I. Niven, and M. J. Correia, *Kinematics nomenclature for physiological accelerations: With special reference to vestibular applications*, vol. 14. Naval Aerospace Medical Institute, 1966.

[90] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of vision*, vol. 8, no. 3, p. 33, 2008.

[91] P. A. Howarth and S. G. Hodder, "Characteristics of habituation to motion in a virtual environment," *Displays*, vol. 29, no. 2, pp. 117–123, 2008.

[92] P. A. Howarth, "Oculomotor changes within virtual environments," *Applied Ergonomics*, vol. 30, no. 1, pp. 59–67, 1999.

[93] A. D. Hwang and E. Peli, "Instability of the perceived world while watching 3D stereoscopic imagery: a likely source of motion sickness symptoms," *i-Perception*, vol. 5, no. 6, pp. 515–535, 2014.

[94] Q. Ihemedu-Steinke, D. Sirim, R. Erbach, P. Halady, and G. Meixner, "Development and evaluation of a virtual reality driving simulator," in *Mensch und Computer 2015–Workshopband*, 2015, pp. 491–500.

[95] Q. C. Ihemedu-Steinke, S. Rangelova, M. Weber, R. Erbach, G. Meixner, and N. Marsden, "Simulation sickness related to virtual reality driving simulation," in *International Conference on Virtual, Augmented and Mixed Reality*, 2017, pp. 521–532.

[96] W. Jin, J. Fan, D. Gromala, and P. Pasquier, "Automatic prediction of cybersickness for virtual reality games," in *2018 IEEE Games, Entertainment, Media Conference (GEM)*. IEEE, 2018, pp. 1–9.

[97] C. Jinjakam and K. Hamamoto, "Simulator sickness in immersive virtual environment," in *The 5th 2012 Biomedical Engineering International Conference*, 2012, pp. 1–4.

[98] D. M. Johnson, "Introduction to and review of simulator sickness research," Army Research Inst Field Unit Fort Rucker AL, Technical Report, 2005.

[99] D. M. Johnson, "Simulator sickness during emergency procedures training in a helicopter simulator: Age, flight experience, and amount learned," Army Research Inst for the behavioral and social sciences Fort Rucker AL, Technical Report, 2007.

[100] N. Kawano, K. Iwamoto, K. Ebe, B. Aleksic, A. Noda, H. Umegaki, M. Kuzuya, T. Iidaka, and N. Ozaki, "Slower adaptation to driving simulator and simulator sickness in older

adults aging clinical and experimental research," *Aging clinical and experimental research*, vol. 24, no. 3, pp. 285–289, 2012.

[101] R. S. Kennedy and J. E. Fowlkes, "Simulator sickness is polygenic and polysymptomatic: Implications for research," *The International Journal of Aviation Psychology*, vol. 2, no. 1, pp. 23–38, 1992.

[102] R. S. Kennedy and M. G. Lilienthal, "Implications of balance disturbances following exposure to virtual reality systems," in *Proceedings Virtual Reality Annual International Symposium'95*. IEEE, 1995, pp. 35–39.

[103] R. S. Kennedy, M. G. Lilienthal, K. S. Berbaum, D. Baltzley, and M. McCauley, "Simulator sickness in US Navy flight simulators," *Aviation, space, and environmental medicine*, vol. 60, no. 1, pp. 10–16, 1989.

[104] R. S. Kennedy, W. P. Dunlap, and J. E. Fowlkes, "Prediction of motion sickness susceptibility," *Motion and space sickness*, pp. 179–216, 1990.

[105] R. S. Kennedy, N. E. Lane, M. G. Lilienthal, K. S. Berbaum, and L. J. Hettinger, "Profile analysis of simulator sickness symptoms: Application to virtual environment systems," *Presence: Teleoperators & Virtual Environments*, vol. 1, no. 3, pp. 295–301, 1992.

[106] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *The international journal of aviation psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[107] R. Kennedy, K. Berbaum, M. Lilienthal, W. Dunlap, and B. Mulligan, "Guidelines for alleviation of simulator sickness symptomatology," Naval Training Center Orlando FL, Technical Report, 1987.

[108] B. Keshavarz, R. Ramkhalawansingh, B. Haycock, S. Shahab, and J. Campos, "Comparing simulator sickness in younger and older adults during simulated driving under different multisensory conditions," *Transportation research part F: traffic psychology and behaviour*, vol. 54, pp. 47–62, 2018.

[109] B. Keshavarz, "Exploring behavioral methods to reduce visually induced motion sickness in virtual environments," in *International Conference on Virtual, Augmented and Mixed Reality*. Springer, 2016, pp. 147–155.

[110] B. Keshavarz and H. Hecht, "Axis rotation and visually induced motion sickness: the role of combined roll, pitch, and yaw motion," *Aviation, space, and environmental medicine*, vol. 82, no. 11, pp. 1023–1029, 2011.

[111] B. Keshavarz and H. Hecht, "Validating an efficient method to quantify motion sickness," *Human factors*, vol. 53, no. 4, pp. 415–426, 2011.

[112] B. Keshavarz and H. Hecht, "Pleasant music as a countermeasure against visually induced motion sickness," *Applied ergonomics*, vol. 45, no. 3, pp. 521–527, 2014.

[113] B. Keshavarz, H. Hecht, and B. D. Lawson, "Visually induced motion sickness: causes, characteristics, and countermeasures," in *Handbook of virtual environments*. CRC Press, 2014, pp. 652–703.

[114] B. Keshavarz, B. E. Riecke, L. J. Hettinger, and J. L. Campos, "Vection and visually induced motion sickness: how are they related?" *Frontiers in psychology*, vol. 6, p. 472, 2015.

[115] B. Keshavarz, D. Stelzmann, A. Paillard, and H. Hecht, "Visually induced motion sickness can be alleviated by pleasant odors," *Experimental brain research*, vol. 233, no. 5, pp. 1353–1364, 2015.

[116] B. Keshavarz, A. C. Novak, L. J. Hettinger, T. A. Stoffregen, and J. L. Campos, "Passive restraint reduces visually induced motion sickness in older adults," *Journal of experimental psychology: applied*, vol. 23, no. 1, p. 85, 2017.

[117] D. H. Kim, D. E. Parker, and M. Y. Park, "A new procedure for measuring simulator sickness - the RSSQ," University of Washington Human Interface Technology Laboratory, Seattle, WA, Technical Report, 2004.

[118] J. Kim, W. Kim, S. Ahn, J. Kim, and S. Lee, "Virtual reality sickness predictor: Analysis of visual-vestibular conflict and VR contents," *10th International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–6, 2018.

[119] K. Kim, M. Z. Rosenthal, D. Zielinski, and R. Brady, "Comparison of desktop, head mounted display, and six wall fully immersive systems using a stressful task," in *2012 IEEE Virtual Reality Workshops (VRW)*. IEEE, 2012, pp. 143–144.

[120] Y. Y. Kim, H. J. Kim, E. N. Kim, H. D. Ko, and H. T. Kim, "Characteristic changes in the physiological components of cybersickness," *Psychophysiology*, vol. 42, no. 5, pp. 616–625, 2005.

[121] S. Klosterhalfen, S. Kellermann, F. Pan, U. Stockhorst, G. Hall, and P. Enck, "Effects of ethnicity and gender on motion sickness susceptibility," *Aviation, space, and environmental medicine*, vol. 76, no. 11, pp. 1051–1057, 2005.

[122] M. Klüver, C. Herrigel, S. Preuß, H.-P. Schöner, and H. Hecht, "Comparing the incidence of simulator sickness in five different driving simulators," in *Proceedings of Driving Simulation Conference*, 2015.

[123] M. Klüver, C. Herrigel, C. Heinrich, H.-P. Schöner, and H. Hecht, "The behavioral validity of dual-task driving performance in fixed and moving base driving simulators,"

Transportation research part F: traffic psychology and behaviour, vol. 37, pp. 78–96, 2016.

[124] J. Koenig and J. F. Thayer, "Sex differences in healthy human heart rate variability: a meta-analysis," *Neuroscience & Biobehavioral Reviews*, vol. 64, pp. 288–310, 2016.

[125] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the 14th international joint conference on Artificial intelligence*, 1995, pp. 1137–1143.

[126] E. M. Kolasinski, "Prediction of simulator sickness in a virtual environment," Ph.D. dissertation, University of Central Florida, 1996.

[127] E. M. Kolasinski, "Simulator sickness in virtual environments," Army Research Inst for the Behavioral and Social Sciences Alexandria VA, Technical Report, 1995.

[128] E. M. Kolasinski and R. D. Gilson, "Simulator sickness and related findings in a virtual environment," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, 1998, pp. 1511–1515.

[129] A. Körner, M. Geyer, M. Roth, M. Drapeau, G. Schmutzer, C. Albani, S. Schumann, and E. Brähler, "Personality assessment with the NEO-five-factor inventory: the 30-item-short-version (NEO-FFI-30)," *Psychotherapie, Psychosomatik, Medizinische Psychologie*, vol. 58, no. 6, pp. 238–245, 2008.

[130] O. Kwon, J. Jeong, H. B. Kim, I. H. Kwon, S. Y. Park, J. E. Kim, and Y. Choi, "Electrocardiogram sampling frequency range acceptable for heart rate variability analysis," *Healthcare informatics research*, vol. 24, no. 3, pp. 198–206, 2018.

[131] J. J. LaViola, "A discussion of cybersickness in virtual environments," *ACM SIGCHI Bulletin*, vol. 32, no. 1, pp. 47–56, 2000.

[132] G. C. Lee, Y. Yoo, and S. Jones, "Investigation of driving performance, vection, postural sway, and simulator sickness in a fixed-based driving simulator," *Computers & industrial engineering*, vol. 33, no. 3-4, pp. 533–536, 1997.

[133] G. Lemaître, F. Nogueira, and C. K. Aridas, "Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning," *Journal of Machine Learning Research*, vol. 18, no. 17, pp. 1–5, 2017.

[134] H.-C. Lien, W. M. Sun, Y.-H. Chen, H. Kim, W. Hasler, and C. Owyang, "Effects of ginger on motion sickness and gastric slow-wave dysrhythmias induced by circular vection," *American Journal of Physiology-Gastrointestinal and Liver Physiology*, vol. 284, no. 3, pp. 481–489, 2003.

[135] J. J.-W. Lin, H. Abi-Rached, D.-H. Kim, D. E. Parker, and T. A. Furness, "A "natural" independent visual background reduced simulator sickness," in *Proceedings of the human factors and ergonomics society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA, 2002, pp. 2124–2128.

[136] J.-W. Lin, H. B.-L. Duh, D. E. Parker, H. Abi-Rached, and T. A. Furness, "Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment," in *Proceedings IEEE Virtual Reality 2002*. IEEE, 2002, pp. 164–171.

[137] Y. Ling, W.-P. Brinkman, H. T. Nefs, C. Qu, and I. Heynderickx, "Cybersickness and anxiety in virtual environments," in *2011 Joint Virtual Reality Conference. Nottingham UK: VTT Symposium*, 2011, pp. 80–82.

[138] Y. Ling, H. T. Nefs, W.-P. Brinkman, C. Qu, and I. Heynderickx, "The relationship between individual characteristics and experienced presence," *Computers in Human Behavior*, vol. 29, no. 4, pp. 1519–1530, 2013.

[139] C.-L. Liu and S.-T. Uang, "Effects of presence on causing cybersickness in the elderly within a 3D virtual store," in *International Conference on Human-Computer Interaction*. Springer, 2011, pp. 490–499.

[140] L. Liu, B. Watson, and M. Miyazaki, "VR for the elderly: Quantitative and qualitative differences in performance with a driving simulator," *CyberPsychology & Behavior*, vol. 2, no. 6, pp. 567–576, 1999.

[141] W. Lo and R. H. So, "Cybersickness in the presence of scene rotational movements along different axes," *Applied ergonomics*, vol. 32, no. 1, pp. 1–14, 2001.

[142] N. Martin, N. Mathieu, N. Pallamin, M. Ragot, and J.-M. Diverrez, "Automatic recognition of virtual reality sickness based on physiological signals," in *IBC*, 2018.

[143] N. A. Matas, T. Nettelbeck, and N. R. Burns, "Dropout during a driving simulator study: a survival analysis," *Journal of safety research*, vol. 55, pp. 159–169, 2015.

[144] M. E. McCauley, *Research issues in simulator sickness: Proceedings of a workshop*. National Academies, 1984.

[145] M. E. McCauley and T. J. Sharkey, "Cybersickness: Perception of self-motion in virtual environments," *Presence: Teleoperators & Virtual Environments*, vol. 1, no. 3, pp. 311–318, 1992.

[146] M. McGill, A. Ng, and S. Brewster, "I am the passenger: how visual motion cues can influence sickness for in-car VR," in *Proceedings of the 2017 CHI conference on human factors in computing systems*, 2017, pp. 5655–5668.

[147] O. Merhi, E. Faugloire, M. Flanagan, and T. A. Stoffregen, "Motion sickness, console video games, and head-mounted displays," *Human Factors*, vol. 49, no. 5, pp. 920–934, 2007.

[148] B.-C. Min, S.-C. Chung, Y.-K. Min, and K. Sakamoto, "Psychophysiological evaluation of simulator sickness evoked by a graphic simulator," *Applied ergonomics*, vol. 35, no. 6, pp. 549–556, 2004.

[149] M. Mon-Williams, J. Wann, and S. Rushton, "Design factors in stereoscopic virtual-reality displays," *Journal of the Society for information Display*, vol. 3, no. 4, pp. 207–210, 1995.

[150] R. R. Mourant and M. T. Schultheis, "A HMD-based virtual reality driving simulator," in *First International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 2001, pp. 300–304.

[151] R. R. Mourant and T. R. Thattacherry, "Simulator sickness in a virtual environments driving simulator," in *Proceedings of the human factors and ergonomics society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA, 2000, pp. 534–537.

[152] J. Munafo, M. Diedrick, and T. A. Stoffregen, "The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects," *Experimental brain research*, vol. 235, no. 3, pp. 889–901, 2017.

[153] E. Nalivaiko, S. L. Davis, K. L. Blackmore, A. Vakulin, and K. V. Nesbitt, "Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time," *Physiology & behavior*, vol. 151, pp. 583–590, 2015.

[154] R. Naunton, *The vestibular system*. Academic Press, 1975.

[155] S. Nichols and H. Patel, "Health and safety implications of virtual reality: a review of empirical evidence," *Applied ergonomics*, vol. 33, no. 3, pp. 251–271, 2002.

[156] S. Nichols, C. Haldane, and J. R. Wilson, "Measurement of presence and its consequences in virtual environments," *International Journal of Human-Computer Studies*, vol. 52, no. 3, pp. 471 – 491, 2000.

[157] A. Nickkar, M. Jeihani, and S. Sahebi, "Analysis of driving simulator sickness symptoms: zero-inflated ordered probit approach," *Transportation research record*, vol. 2673, no. 4, pp. 988–1000, 2019.

[158] N. Owen, A. G. Leadbetter, and L. Yardley, "Relationship between postural control and motion sickness in healthy subjects," *Brain research bulletin*, vol. 47, no. 5, pp. 471–474, 1998.

[159] S. Palmisano, R. Mursic, and J. Kim, "Vection and cybersickness generated by head-and-display motion in the Oculus Rift," *Displays*, vol. 46, pp. 1–8, 2017.

[160] W. D. Park, S. W. Jang, Y. H. Kim, G. A. Kim, W. Son, and Y. S. Kim, "A study on cyber sickness reduction by oculo-motor exercise performed immediately prior to viewing virtual reality (VR) content on head mounted display (HMD)," *Vibroengineering PROCEDIA*, vol. 14, pp. 260–264, 2017.

[161] R. Pausch, T. Crea, and M. Conway, "A literature survey for virtual environments: Military flight simulator visual systems and simulator sickness," *Presence: Teleoperators & Virtual Environments*, vol. 1, no. 3, pp. 344–363, 1992.

[162] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[163] E. Peper, R. Harvey, I.-M. Lin, H. Tylova, and D. Moss, "Is there more to blood volume pulse than heart rate variability, respiratory sinus arrhythmia, and cardiorespiratory synchrony?" *Biofeedback*, vol. 35, no. 2, 2007.

[164] J. Plouzeau, D. Paillot, J.-R. Chardonnet, and F. Merienne, "Effect of proprioceptive vibrations on simulator sickness during navigation task in virtual environment," in *International Conference on Artificial Reality and Telexistence Eurographics Symposium on Virtual Environments*, 2015.

[165] J. D. Prothero, "The role of rest frames in vection, presence and motion sickness," Ph.D. dissertation, University of Washington, 1998.

[166] D. Purves, G. J. Augustine, D. Fitzpatrick, L. C. Katz, A.-S. LaMantia, J. O. McNamara, and S. M. Williams, "Types of eye movements and their functions," *Neuroscience*, 2001.

[167] V. S. Ramachandran, "Perception of shape from shading," *Nature*, vol. 331, no. 6152, pp. 163–166, 1988.

[168] S. Rangelova and E. Andre, "A survey on simulation sickness in driving applications with virtual reality head-mounted displays," *Presence: Virtual and Augmented Reality*, vol. 27, no. 1, pp. 15–31, 2019.

[169] S. Rangelova, D. Decker, M. Eckel, and E. Andre, "Simulation sickness evaluation while using a fully autonomous car in a head mounted display virtual environment," in *International Conference on Virtual, Augmented and Mixed Reality*. Springer, 2018, pp. 155–167.

[170] S. Rangelova, K. Rehm, S. Diefenbach, D. Motus, and E. André, "Gender differences in simulation sickness in static vs. moving platform VR automated driving simulation," in *International Conference on Human-Computer Interaction*. Springer, 2020, pp. 146–165.

[171] J. T. Reason, "Motion sickness adaptation: a neural mismatch model," *Journal of the Royal Society of Medicine*, vol. 71, no. 11, pp. 819–829, 1978.

[172] J. T. Reason and J. J. Brand, *Motion sickness*. Academic press, 1975.

[173] L. Rebenitsch and C. Owen, "Individual variation in susceptibility to cybersickness," in *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM, 2014, pp. 309–317.

[174] L. Rebenitsch and C. Owen, "Review on cybersickness in applications and visual displays," *Virtual Reality*, vol. 20, no. 2, pp. 101–125, 2016.

[175] L. R. Rebenitsch, "Cybersickness prioritization and modeling," Ph.D. dissertation, Michigan State University, 2015.

[176] R. J. Reed-Jones, L. A. Vallis, J. G. Reed-Jones, and L. M. Trick, "The relationship between postural stability and virtual environment adaptation," *Neuroscience letters*, vol. 435, no. 3, pp. 204–209, 2008.

[177] C. Regan, "An investigation into nausea and other side-effects of head-coupled immersive virtual reality," *Virtual Reality*, vol. 1, no. 1, pp. 17–31, Jun 1995.

[178] E. C. Regan, "Some evidence of adaptation to immersion in virtual reality," *Displays*, vol. 16, no. 3, pp. 135–139, 1995.

[179] K. Rehm, "Simulation sickness in virtual autonomous driving simulators," Master's thesis, Ludwig-Maximilians-University Munich, Germany, 2019.

[180] D. Reich and R. Stark, "Head-mounted displays for hci validations while driving," *International Journal of Electrical and Information Engineering*, vol. 10, no. 10, pp. 1887–1891, 2016.

[181] R. Riccelli, I. Indovina, J. P. Staab, S. Nigro, A. Augimeri, F. Lacquaniti, and L. Passamonti, "Neuroticism modulates brain visuo-vestibular and anxiety systems during a virtual rollercoaster task," *Human Brain Mapping*, vol. 38, no. 2, pp. 715–726, 2017.

[182] G. E. Riccio and T. A. Stoffregen, "An ecological theory of motion sickness and postural instability," *Ecological psychology*, vol. 3, no. 3, pp. 195–240, 1991.

[183] I. Rish, "An empirical study of the naive Bayes classifier," in *IJCAI 2001 workshop on empirical methods in artificial intelligence*, 2001, pp. 41–46.

[184] A. Rolnick and R. Lubow, "Why is the driver rarely motion sick? The role of controllability in motion sickness," *Ergonomics*, vol. 34, no. 7, pp. 867–879, 1991.

[185] S. Ropelato, F. Zünd, S. Magnenat, M. Menozzi, and R. Sumner, "Adaptive tutoring on a virtual reality driving simulator," *International SERIES on Information Systems and Management in Creative EMedia (CreMedia)*, vol. 2017, no. 2, pp. 12–17, 2018.

[186] A. H. Rupert and O. I. Kolev, "The use of tactile cues to modify the perception of self-motion," Army Aeromedical Research Lab Fort Rucker AL, Technical Report, 2008.

[187] J. A. Russell, "A circumplex model of affect," *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.

[188] SAE, "Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles," *SAE International: Warrendale, PA, USA*, 2018.

[189] C. Sagnier, E. Loup-Escande, and G. Valléry, "Effects of gender and prior experience in immersive user experience with virtual reality," in *International Conference on Applied Human Factors and Ergonomics*. Springer, 2019, pp. 305–314.

[190] C. Sas and G. O'Hare, "Impact of cognitive style upon sense of presence," in *Proceedings of HCI International*, 2003, pp. 696–700.

[191] T. Schubert, F. Friedmann, and H. Regenbrecht, "The experience of presence: Factor analytic insights," *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 3, pp. 266–281, 2001.

[192] M. J. Schuemie, P. van der Straaten, M. Krijn, and C. A. van der Mast, "Research on presence in virtual reality: A survey," *CyberPsychology & Behavior*, vol. 4, no. 2, pp. 183–201, 2001, pMID: 11710246.

[193] A. F. Seay, D. M. Krum, L. Hodges, and W. Ribarsky, "Simulator sickness and presence in a high field-of-view virtual environment," in *CHI'02 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2002, pp. 784–785.

[194] N. Setiowati, T. Wijayanto, and F. Trapsilawati, "Identifying cybersickness when wearing a head-mounted display through heart rate variability data," in *IOP Conference Series: Materials Science and Engineering*. IOP Publishing, 2020, p. 012069.

[195] K. Sharma, "Prevalence and correlates of susceptibility to motion sickness," *Acta geneticae medicae et gemellologiae: twin research*, vol. 46, no. 2, pp. 105–121, 1997.

[196] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson, "Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems," *Displays*, vol. 29, no. 2, pp. 58–69, 2008.

[197] L. Shu, J. Xie, M. Yang, Z. Li, Z. Li, D. Liao, X. Xu, and X. Yang, "A review of emotion recognition using physiological signals," *Sensors*, vol. 18, no. 7, p. 2074, 2018.

[198] M. Sieber, G. Berg, I. Karl, K. Siedersberger, A. Siegel, and B. Färber, "Validation of driving behavior in the vehicle in the loop: Steering responses in critical situations," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 2013, pp. 1101–1106.

[199] L. J. Smart Jr, T. A. Stoffregen, and B. G. Bardy, "Visually induced motion sickness predicted by postural instability," *Human factors*, vol. 44, no. 3, pp. 451–465, 2002.

[200] R. H. So, "The search for a cybersickness dose value," in *HCI (1)*, 1999, pp. 152–156.

[201] R. H. So and S. Yuen, "Comparing symptoms of visually induced motion sickness among viewers of four similar virtual environments with different color," in *International Conference on Virtual Reality*. Springer, 2007, pp. 386–391.

[202] R. H. So, A. Ho, and W. Lo, "A metric to quantify virtual scene movement for the study of cybersickness: Definition, implementation, and verification," *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 2, pp. 193–215, 2001.

[203] R. H. So, W. Lo, and A. T. Ho, "Effects of navigation speed on motion sickness caused by an immersive virtual environment," *Human factors*, vol. 43, no. 3, pp. 452–461, 2001.

[204] A. Spinks and J. Wasiak, "Scopolamine (hyoscine) for preventing and treating motion sickness," *Cochrane database of systematic reviews*, no. 6, 2011.

[205] K. M. Stanney and P. Hash, "Locus of user-initiated control in virtual environments: Influences on cybersickness," *Presence*, vol. 7, no. 5, pp. 447–459, 1998.

[206] K. M. Stanney and R. S. Kennedy, "The psychometrics of cybersickness," *Communications of the ACM*, vol. 40, no. 8, pp. 66–68, 1997.

[207] K. M. Stanney and R. S. Kennedy, "Aftereffects from virtual environment exposure: How long do they last?" in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, 1998, pp. 1476–1480.

[208] K. M. Stanney, R. S. Kennedy, and J. M. Drexler, "Cybersickness is not simulator sickness," in *Proceedings of the Human Factors and Ergonomics Society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA, 1997, pp. 1138–1142.

[209] K. M. Stanney, R. S. Kennedy, J. M. Drexler, and D. L. Harm, "Motion sickness and proprioceptive aftereffects following virtual environment exposure," *Applied ergonomics*, vol. 30, no. 1, pp. 27–38, 1999.

[210] K. M. Stanney, D. S. Lanham, R. S. Kennedy, and R. Breaux, "Virtual environment exposure drop-out thresholds," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, 1999, pp. 1223–1227.

[211] K. M. Stanney, K. S. Kingdon, D. Graeber, and R. S. Kennedy, "Human performance in immersive virtual environments: Effects of exposure duration, user control, and scene complexity," *Human performance*, vol. 15, no. 4, pp. 339–366, 2002.

[212] K. M. Stanney, K. S. Hale, I. Nahmens, and R. S. Kennedy, "What to expect from immersive virtual environment exposure: Influences of gender, body mass index, and past experience," *Human factors*, vol. 45, no. 3, pp. 504–520, 2003.

[213] K. M. Stanney, C. Fidopiastis, and L. Foster, "Virtual reality is sexist: But it does not have to be," *Frontiers in Robotics and AI*, vol. 7, p. 4, 2020.

[214] T. A. Stoffregen and L. J. Smart Jr, "Postural instability precedes motion sickness," *Brain research bulletin*, vol. 47, no. 5, pp. 437–448, 1998.

[215] H. A. Stoner, D. L. Fisher, and M. Mollenhauer, "Simulator and scenario factors influencing simulator sickness," *Handbook of driving simulation for engineering, medicine, and psychology*, pp. 14–11, 2011.

[216] D. C. D. Suwarno, T. Wijayanto, and F. Trapsilawati, "Cybersickness Evaluation While using Driving Simulator in a Head-Mounted Display Environment," in *2019 International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE)*. IEEE, 2019, pp. 103–106.

[217] O. Theobald, *Machine learning for absolute beginners: a plain English introduction*. Scatterplot press, 2017.

[218] K. Uliano, E. Lambert, R. Kennedy, and D. Sheppard, "The effects of asynchronous visual delays on simulator flight performance and the development of simulator sickness symptomatology," Essex Corp Orlando FL, Technical Report, 1986.

[219] R. Venkatakrishnan, A. Bhargava, R. Venkatakrishnan, K. M. Lucaites, M. Volonte, H. Solini, A. C. Robb, C. Pagano, and S. V. Babu, "Towards an immersive driving simulator to study factors related to cybersickness," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019, pp. 1201–1202.

[220] I. Viaud-Delmon, Y. P. Ivanenko, A. Berthoz, and R. Jouvent, "Sex, lies and virtual reality," *Nature Neuroscience*, vol. 1, no. 1, p. 15, 1998.

[221] S. J. Villard, M. B. Flanagan, G. M. Albanese, and T. A. Stoffregen, "Postural instability and motion sickness in a virtual moving room," *Human factors*, vol. 50, no. 2, pp. 332–345, 2008.

[222] S. Von Mammen, A. Knote, and S. Edenhofer, "Cyber sick but still having fun," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*. ACM, 2016, pp. 325–326.

[223] T. Wada, "Motion sickness in automated vehicles," in *In Advanced Vehicle Control: Proceedings of the 13th International Symposium on Advanced Vehicle Control (AVEC'16)*, 2016, pp. 169–174.

[224] J. Wagner, F. Lingenfelser, T. Baur, I. Damian, F. Kistler, and E. André, "The social signal interpretation (SSI) framework: multimodal signal processing and recognition in real-time," in *Proceedings of the 21st ACM international conference on Multimedia*, New York, NY, USA, 2013, pp. 831–834.

[225] M. Walch, J. Frommel, K. Rogers, F. Schüssel, P. Hock, D. Dobbelstein, and M. Weber, "Evaluating VR driving simulation from a player experience perspective," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, New York, NY, USA, 2017, pp. 2982–2989.

[226] H. Wan, S. Hu, and J. Wang, "Correlation of phasic and tonic skin-conductance responses with severity of motion sickness induced by viewing an optokinetic rotating drum," *Perceptual and motor skills*, vol. 97, no. 3_suppl, pp. 1051–1057, 2003.

[227] L. Warwick-Evans, N. Symons, T. Fitch, and L. Burrows, "Evaluating sensory conflict and postural instability. theories of motion sickness," *Brain research bulletin*, vol. 47, no. 5, pp. 465–469, 1998.

[228] S. Weech, J. Moon, and N. F. Troje, "Influence of bone-conducted vibration on simulator sickness in virtual reality," *PloS one*, vol. 13, no. 3, 2018.

[229] S. Weech, S. Kenny, and M. Barnett-Cowan, "Presence and cybersickness in virtual reality are negatively related: a review," *Frontiers in psychology*, vol. 10, p. 158, 2019.

[230] F. Weidner, A. Hoesch, S. Poeschl, and W. Broll, "Comparing VR and non-VR driving simulations: An experimental user study," in *2017 IEEE Virtual Reality (VR)*, 2017, pp. 281–282.

[231] D. M. Whittinghill, B. Ziegler, T. Case, and B. Moore, "Nasum virtualis: A simple technique for reducing simulator sickness," in *Games Developers Conference (GDC)*, 2015.

[232] J. Wilding and R. Meddis, "A note on personality correlates of motion sickness," *British Journal of Psychology*, vol. 63, no. 4, p. 619, 1972.

[233] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, no. 3, pp. 225–240, June 1998.

[234] R. Wright, "Helicopter simulator sickness: A state-of-the-art review of its incidence, causes, and treatment," Technical Report, 1995.