

## **Analysis of physiological responses induced by motion sickness and its detection based on ocular parameters**

**Auteur :** Pyre, Floriane

**Promoteur(s) :** Drion, Guillaume

**Faculté :** Faculté des Sciences appliquées

**Diplôme :** Master en ingénieur civil biomédical, à finalité spécialisée

**Année académique :** 2020-2021

**URI/URL :** <http://hdl.handle.net/2268.2/11510>

---

### *Avertissement à l'attention des usagers :*

*Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.*

*Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.*

---



University of Liège  
Faculty of Applied Sciences  
Academic year 2020-2021

---

# Analysis of physiological responses induced by motion sickness and its detection based on ocular parameters

---

MASTER THESIS CONDUCTED BY

**FLORIANE PYRE**

WITH THE AIM OF OBTAINING THE DEGREE OF MASTER IN  
BIOMEDICAL ENGINEERING

Conducted in collaboration with



Under the supervision of

Clémentine François, PhD  
Prof. Guillaume Drion



# Abstract

One of the promises of autonomous cars is that these will allow drivers to become passengers, and therefore to be engaged in many tasks other than driving such as working, reading or relaxing. However, there exists an increased risk of motion sickness incidence in self-driving cars, thus preventing people suffering from this state from devoting themselves to these tasks. As a consequence, the user acceptance and uptake of autonomous cars could be negatively affected, limiting the benefits this emerging technology may provide. To avoid the negative impact motion sickness could have on autonomous car adoption, this problem has to be investigated and appropriate countermeasures have to be developed.

A first step in the development of a solution consists in detecting early signs of motion sickness, allowing so to initiate the triggering of various processes intended to alleviate the symptoms associated with motion sickness. The aim of this thesis is to identify physiological parameters that are indicative of motion sickness, and to determine the relevance of ocular parameters for predicting this state. Indeed, ocular data could be easily recorded in autonomous cars through integrated high precision cameras.

A protocol to acquire data in context is first designed. This protocol aims at inducing motion sickness in 2 different ways. The first one consists in driving in a fixed-base driving simulator. The second one consists in performing some tasks on paper while being a passenger in a moving car. Twenty subjects took part to the protocol. Severe motion sickness was reported by 3 and 9 participants, during the session in simulator and the session in car respectively. The analysis of the collected data shows that heart rate, electrodermal and gastric activities increase with motion sickness. Machine learning models are then trained with ocular data as inputs, and a 3-level score, reflecting the severity of motion sickness, as ground truth. The results suggest that ocular data alone cannot predict motion sickness, but that it may be appropriate to combine it with other physiological data in order to predict motion sickness.



# Acknowledgments

This project represents my first long term work and the completion of it would probably not have been possible without the support of many people.

First of all, I would like to express my sincere gratitude to my industrial supervisor Clémentine François, PhD, for giving me the opportunity to work on such an interesting and innovative subject. Besides her dynamism and motivation that have deeply inspired me, she has always provided invaluable guidance throughout this work.

Furthermore, I would like to thank my academic promotor Prof. Guillaume Drion for his support regarding my project and his availability for answering any of my concerns.

Moreover, I would like to express my thanks to the people working at Phasya and with who I had the opportunity to work during my internship. From the very beginning, they included me in their team and guided me in my work. Despite the Covid-19 pandemic and the social distancing, they were always available to answer my questions and to give me advice for my work as if it was their own.

Finally, I am extremely thankful to my family and friends for their continuous encouragement during this project and more generally during my studies. Particularly, I would like to thank my parents and my sister for their endless support, always pushing me to give my best. A special thanks to my boyfriend, Arthur, for his unconditional love and continuous support throughout all the steps of my studies.

# Contents

<b>Introduction</b>	<b>1</b>
<b>1 Background</b>	<b>4</b>
1.1 Definition and characterization of motion sickness . . . . .	4
1.1.1 What does motion sickness mean ? . . . . .	4
1.1.2 Why does motion sickness occur ? . . . . .	4
1.1.3 What stimuli cause motion sickness ? . . . . .	6
1.2 Methods for detecting motion sickness . . . . .	8
1.2.1 Objective and physiological methods . . . . .	8
1.2.2 Subjective methods . . . . .	14
1.3 State of the art in the field . . . . .	16
1.3.1 Exclusion criteria of participants . . . . .	16
1.3.2 What kind of equipment is used ? . . . . .	17
1.3.3 How to induce motion sickness ? . . . . .	17
1.3.4 What is recorded ? . . . . .	18
<b>2 Motion sickness protocol</b>	<b>20</b>
2.1 Recruitment of participants . . . . .	20
2.2 Session progress and performed tasks . . . . .	21
2.2.1 Driving simulator session . . . . .	21
2.2.2 Car session . . . . .	24
2.3 Collected data and equipment . . . . .	25
2.3.1 Physiological data . . . . .	25
2.3.2 Subjective data . . . . .	32
<b>3 Data analysis</b>	<b>34</b>
3.1 Synchronization and processing . . . . .	34
3.1.1 Relevant physiological data . . . . .	34
3.2 Results and discussion . . . . .	41
3.2.1 Cardiac activity . . . . .	42

3.2.2	Electrodermal activity . . . . .	59
3.2.3	Gastric activity . . . . .	64
3.3	Ground truth definition . . . . .	70
3.3.1	Cardiac activity . . . . .	72
3.3.2	Electrodermal activity . . . . .	75
3.3.3	Gastric activity . . . . .	76
3.4	Conclusions and limitations . . . . .	78
<b>4</b>	<b>Detection of motion sickness based on ocular parameters</b>	<b>80</b>
4.1	Motivation . . . . .	80
4.2	Method . . . . .	81
4.2.1	Dataset . . . . .	81
4.2.2	Method description . . . . .	82
4.3	Results . . . . .	86
4.3.1	First consideration . . . . .	86
4.3.2	Second consideration . . . . .	91
4.3.3	Third consideration . . . . .	93
4.4	Conclusions and limitations . . . . .	97
	<b>Conclusion</b>	<b>99</b>
	<b>Appendices</b>	<b>101</b>
<b>A</b>	<b>Protocol details</b>	<b>102</b>
A.1	Route of the car session . . . . .	102
A.2	Tasks proposed to the participants during the motion sickness section of the car session . . . . .	103
A.3	List of common physical responses during motion sickness . . . . .	114
<b>B</b>	<b>Data analysis</b>	<b>116</b>
B.1	Cardiac activity . . . . .	116
B.2	Electrodermal activity . . . . .	118
<b>C</b>	<b>Pseudo-code of the considered method for training machine learning models</b>	<b>119</b>
	<b>Bibliography</b>	<b>120</b>

# List of Figures

1.1	Illustration of the interaction between the sensory systems providing the perception of self-motion. The vestibular system, represented by the semi-circular canals and otolith organs, distinguishes neck rotational and translational motion. The eyes are stimulated by the movement of a large visual field. Non-vestibular proprioceptive inputs, besides motor control commands contribute to spatial orientation [10]. . . . .	5
1.2	Typical appearance of an ECG [20]. . . . .	9
1.3	Two consecutive typical BVP pulse waveforms [21]. . . . .	10
1.4	Brain lobes [36]. . . . .	13
2.1	Fixed-base driving simulator at the University of Liège [44]. . . . .	22
2.2	Illustration of the 6 sections composing the driving simulator scenario. . . .	23
2.3	MP160 data acquisition unit and transducer module (AMI100D) [47]. . . .	27
2.4	Wireless BioNomadix module [48]. . . . .	27
2.5	Lead set [49]. . . . .	27
2.6	Disposable electrodes [50]. . . . .	27
2.7	Illustration of the 4 parts composing the solution provided by <i>BIOPAC</i> [48].	28
2.8	B-Alert X10 wireless EEG system [51]. . . . .	28
2.9	Sample channels at 9 brain sites [51]. . . . .	29
2.10	Mastoids illustration, fixing the gain of the channels of the 9 brain EEG sites. . . . .	29
2.11	EGG100D Smart Amplifier [53]. . . . .	30
2.12	<i>E4</i> wristband [54]. . . . .	31
2.13	Phasya Glasses and laptop on which the <i>Drowsilogic</i> software is installed [44]. . . . .	31
2.14	Use and outputs of the Drowsimeter R100 [44]. . . . .	32
3.1	Examples of downsampled ECG (500 Hz), band-pass filtered ECG (0.5-35 Hz) and template signal obtained using <i>AcqKnowledge</i> . . . . .	36

3.2	Example of a R-R interval tachogram, illustrating the evolution of R-R interval with time. The tachogram is displayed with the beat number along the horizontal axis and R-R time interval on the vertical axis. . . . .	36
3.3	Example of raw and phasic EDA data. . . . .	38
3.4	Peak detection on raw data without (top) and with (bottom) artifact detection. . . . .	39
3.5	Part of a stepped EGG signal (left) and its corresponding smoothed signal (right). . . . .	40
3.6	Example of local maximum detection on a smoothed EGG signal part. . .	41
3.7	Means of normalized HR values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 7 (left) and 16 (right), who suffered from motion sickness during simulator environment exposure. . . . .	43
3.8	Means of normalized HR values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 13 (left) and 17 (right), who did not suffer from motion sickness during simulator environment exposure. . . . .	44
3.9	Means of normalized HR values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants. . . . .	45
3.10	Means of normalized pNN20, pNN50, RMSSD and SDRR values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 7 (left) and 16 (right), who suffered from motion sickness during simulator environment exposure. . . . .	46
3.11	Means of normalized pNN20, pNN50, RMSSD and SDRR values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 13 (left) and 17 (right), who did not suffer from motion sickness during simulator environment exposure. . . . .	48
3.12	Means of normalized pNN20 values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants. . . . .	49
3.13	Means of normalized pNN50 values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants. . . . .	49
3.14	Means of normalized RMSSD values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants. . . . .	50

3.15	Means of normalized SDRR values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants. . . . .	50
3.16	Means of normalized HR values over fixed time periods and collected MS scores for participant 3 (left) and 8 (right), who suffered from motion sickness during the car session. . . . .	51
3.17	Means of normalized HR values over fixed time periods and collected MS scores for participant 5 (left) and 18 (right), who did not suffer from motion sickness during the car session. . . . .	52
3.18	Means of normalized HR values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants. . . . .	53
3.19	Means of normalized pNN20, pNN50, RMSSD and SDRR values over fixed time periods and collected MS scores for participant 3 (left) and 8 (right), who suffered from motion sickness during the car session. . . . .	54
3.20	Means of normalized pNN20, pNN50, RMSSD and SDRR values over fixed time periods and collected MS scores for participant 5 (left) and 18 (right), who did not suffer from motion sickness during the car session. . . . .	55
3.21	Means of normalized pNN20 values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants. . . . .	56
3.22	Means of normalized pNN50 values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants. . . . .	57
3.23	Means of normalized RMSSD values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants. . . . .	57
3.24	Means of normalized SDRR values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants. . . . .	58
3.25	Means of numbers of peaks per minute over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 7 (left) and 16 (right), who suffered from motion sickness during simulator environment exposure. . . . .	60

3.26	Means of numbers of peaks per minute over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 13 (left) and 17 (right), who did not suffer from motion sickness during simulator environment exposure. . . . .	60
3.27	Means of numbers of peaks per minute over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all EDA respondent participants with good quality signals. . . . .	61
3.28	Means of numbers of peaks per minute over the different sections composing the car session and collected MS scores for participant 3 (left) and 8 (right), who suffered from motion sickness during the car session. . . . .	62
3.29	Means of numbers of peaks per minute over the different sections composing the car session and collected MS scores for participant 5 (left) and 18 (right), who did not suffer from motion sickness during the car session. . .	63
3.30	Means of numbers of peaks per minute over the different sections composing the car session in function of the session period (left) and the MS score (right), for all EDA respondent participants with good quality signals. . . .	64
3.31	Means of CPM values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 7 (left) and 16 (right), who suffered from motion sickness during simulator environment exposure. . . . .	66
3.32	Means of CPM values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 13 (left) and 17 (right), who did not suffer from motion sickness during simulator environment exposure. . . . .	66
3.33	Means of CPM values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants. . . . .	67
3.34	Means of CPM values over fixed time periods and collected MS scores for participant 3 (left) and 8 (right), who suffered from motion sickness during the car session. . . . .	68
3.35	Means of CPM values over fixed time periods and collected MS scores for participant 5 (left) and 18 (right), who did not suffer from motion sickness during the car session. . . . .	68
3.36	Means of CPM values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants. . . . .	70

3.37	Pie chart illustrating the percentage of 3-level MS scores for the baseline (left), the motion sickness (center) and the rest (right) sections, for the simulator sessions. . . . .	71
3.38	Pie chart illustrating the percentage of 3-level MS scores for the baseline (left), the motion sickness (center) and the rest (right) sections, for the car sessions. . . . .	72
3.39	Means of normalized HR values over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels, for all participants . . . . .	73
3.40	Means of normalized HRV parameter values over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels, for all participants. . . . .	74
3.41	Means of normalized HR values over the different sections composing the car session in function of the MS score brought down to 3 levels, for all participants. . . . .	74
3.42	Means of normalized HRV parameter values over the different sections composing the car session in function of the MS score brought down to 3 levels, for all participants. . . . .	75
3.43	Means of numbers of peaks per minute over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels, for all EDA respondent participants with good quality signals. . . . .	76
3.44	Means of numbers of peaks per minute over the different sections composing the car session in function of the MS score brought down to 3 levels, for all EDA respondent participants with good quality signals. . . . .	77
3.45	Means of CPM values over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels, for all participants. . . . .	77
3.46	Means of CPM values over the different sections composing the car session in function of the MS score brought down to 3 levels, for all participants. . .	78
4.1	Evolution of the percentage of time where the eye is opened with the MS score (left) and with the period (right) of the car session, for all participants.	85
A.1	Route of the car session. . . . .	102
B.1	Raw and processed IBI and deduced HR for participant 7 (left) and 16 (right) during simulator sessions. . . . .	116



B.2	Raw and processed IBI and deduced HR for participant 3 (left) and 8 (right) during car sessions. . . . .	117
B.3	EDA signals of one non-responder, for both simulator (left) and car (right) sessions. . . . .	118

# List of Tables

1.1	Misery scale [42]. . . . .	15
1.2	Borg rating scale [43]. . . . .	16
3.1	Ranges of average HR values and considered HRV parameter values over a time period of all the 20 participants, during both simulator and car sessions.	37
3.2	Means of CPM values derived for the 4 different experimental conditions. .	65
4.1	Mean over the 4 subsets of the outer cross-validation of the means of the AUPRC over the 5 subsets of the inner cross-validation, associated weighted accuracy, means of the AUPRC and of the weighted accuracy computed on the 4 test sets, for all the learning algorithms trained on the entire car sessions dataset. . . . .	88
4.2	Optimal values of tuned hyperparameters, for all the learning algorithms trained on the entire car sessions dataset. . . . .	89
4.3	Contingency tables for the first (a), second (b), third (c) and fourth (d) test set, resulting from the bagging technique with the optimal hyperparameter value, for the first consideration. . . . .	90
4.4	Mean over the 4 subsets of the outer cross-validation of the means of the weighted accuracy over the 5 subsets of the inner cross-validation, means of the weighted accuracy and of the AUPRC computed on the 4 test sets, for all the learning algorithms trained on the reduced car sessions dataset, composed of observations belonging to the motion sickness section. . . . .	92
4.5	Optimal values of tuned hyperparameters, for all the learning algorithms trained on the reduced car sessions dataset, composed of observations belonging to the motion sickness section. . . . .	93
4.6	Mean over the 4 subsets of the outer cross-validation of the means of the weighted accuracy over the 5 subsets of the inner cross-validation, means of the weighted accuracy and of the AUPRC computed on the 4 test sets, for all the learning algorithms trained on the simulator sessions dataset. . .	94
4.7	Optimal values of tuned hyperparameters, for all the learning algorithms trained on the simulator sessions dataset. . . . .	95

4.8	Contingency tables for the first (a), second (b), third (c) and fourth (d) test set, resulting from the SVM technique with the optimal hyperparameters values, for the third consideration. . . . .	96
-----	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----

# List of Acronyms

**ANS** Autonomic Nervous System

**AUPRC** Area Under the Precision-Recall Curve

**BPM** Beats Per Minute

**BVP** Blood Volume Pulse

**CPM** Cycles Per Minute

**ECG** ElectroCardioGram

**EDA** ElectroDermal Activity

**EEG** ElectroEncephaloGram

**EGG** ElectroGastroGram

**EMG** ElectroMyoGram

**EOG** ElectroOculoGram

**GSR** Galvanic Skin Response

**HR** Heart Rate

**HRV** Heart Rate Variability

**IBI** InterBeat Interval

**MISC** Misery SScale

**MRMR** Minimum Redundancy Maximum Relevance

**MSQ** Motion Sickness Questionnaire

**P-P** Peak-to-Peak

**pNN20** percentage of successive R-R intervals that differ by more than 20 ms

**pNN50** percentage of successive R-R intervals that differ by more than 50 ms

**PNS** Parasympathetic Nervous System

**PPG** PhotoPletysmoGram

**PR** Precision-Recall

**RMS** Root Mean Square

**RMSSD** Root Mean Square of Successive Differences

**ROC** Receiver Operating Characteristic

**RSA** Respiratory Sinus Arrhythmia

**SCL** Skin Conductance Level

**SCR** Skin Conductance Response

**SDL** Scenario Definition Language

**SDRR** Standard Deviation of R-R intervals

**SNS** Sympathetic Nervous System

**SVM** Support Vector Machine

# Introduction

The progressive arrival of autonomous cars on the market will revolutionize the daily life of millions of people who spend hours behind the steering wheel. Drivers will become passengers, and will therefore have the opportunity to perform non-driving tasks. In addition, it appears to be an excellent mean of considerably reducing road accidents, mainly caused by human errors. Indeed, self-driving cars will be equipped with the latest technologies for obstacle detection and will so make the roads safer.

However, self-driving cars face multiple challenges that must be overcome in order to ensure the acceptance and uptake of this technology on our roads. Among these challenges, a comfortable experience for passengers performing non-driving tasks has to be ensured. While drivers do generally not develop symptoms such as nausea, headaches, sweating or disorientation, they might be susceptible to motion sickness when riding as passengers in autonomous cars. It is therefore necessary to detect motion sickness and to develop measures aiming at reducing the intensity of motion sickness symptoms. Motion sickness in autonomous cars can be detected via various physiological data sensors. More and more sensors are indeed integrated in vehicles with the aim of capturing information on the state of the users, making it possible to improve safety, experience, well-being and comfort of the users.

According to the National Institutes of Health, about one in three people are highly susceptible to motion sickness [1]. A lot of studies have shown that low frequency acceleration is associated with motion sickness response [2, 3, 4, 5]. However, these have been conducted in laboratory motion platforms, therefore failing to reproduce the vehicle movements in real conditions that is related to the development of motion sickness in road vehicles. Indeed, while motion profiles generated in motion platforms are mainly oscillatory, typical vehicle dynamics include stochastic combinations of lateral and longitudinal accelerations, pitch rotations and vertical accelerations, resulting from curved roadways, braking and accelerating events as well as road surface geometry.

The realization of this project was mainly motivated by the need for a protocol in which the movements of the vehicle would be closer to the movements in real conditions, thus making it possible to reliably identify the physiological parameters associated with autonomous car sickness. The goal of this work is to develop such a protocol, then identify physiological indicators of motion sickness, and use these to finally develop a model aiming to predict whether or not a person suffers from motion sickness. In particular, this model will be based exclusively on ocular parameters at the request of Phasya. Indeed, if ocular parameters are shown to be indicators of motion sickness, constituting also an objective of this thesis, detection of motion sickness early signs could be easily performed through various high precision cameras that would be integrated in self-driving cars. Counter-measures mitigating motion sickness symptoms could then be triggered, thus ensuring well-being within autonomous cars.

This thesis is divided into 4 main chapters. Chapter 1 introduces the background on motion sickness and reviews the methods for detecting and causing this state. Chapter 2 presents the data acquisition protocol, as well as the equipment used for collecting the data. Chapter 3 is dedicated to the processing and the analysis of a part of the acquired data. Chapter 4 describes the development and presents the results of machine learning models that aim at identifying potential motion sickness indicators, among ocular parameters. Finally, this thesis concludes on the perspectives on the subject.

The present work includes the following main personal contributions:

- Literature review and identification of:
  - parameters that are indicators of motion sickness among the physiological data;
  - key elements for the definition of a protocol, induction of motion sickness and data collection;
- Definition of a protocol for data acquisition;
- Implementation of a driving scenario in a fixed-base driving simulator;
- Participation in the data acquisitions;
- Processing of cardiac and gastric signals;
- Analysis of the evolution of cardiac, electrodermal and gastric activities with motion sickness;

- Development of training and assessing procedure of machine learning models that aim at predicting motion sickness via ocular parameters;
- Analysis of the results of machine learning models.



# Chapter 1

## Background

The goal of this first chapter is to provide a general context for the project. First of all, a definition of motion sickness is provided, the reasons of its occurrence as well as the stimuli causing this state are explained. Then, the methods aiming to detect motion sickness, including objective and subjective methods, are described. Afterwards, some protocols of previous studies are presented.

### 1.1 Definition and characterization of motion sickness

#### 1.1.1 What does motion sickness mean ?

Motion sickness is a state of sickness resulting from unfamiliar or specific motion stimuli, as well as from the absence of expected motion. This state is characterized by many symptoms, which may appear separately or in combination and may vary considerably between people regarding their occurrence and degree of intensity. Some of the most common symptoms are headaches, nausea, excessive fatigue, sweating, excessive saliva production, disorientation, postural instability, and dizziness [6, 7]. The ultimate manifestation of motion sickness is vomiting and is typically preceded by some of the symptoms listed above.

Although the most common environment in which motion sickness occurs is on boats, this natural automatic response also manifests in ground transportation such as cars, buses, trains, as well as in air. The increasing use of simulators and virtual environments has also demonstrated occurrence of motion sickness.

#### 1.1.2 Why does motion sickness occur ?

There are various theories trying to explain why motion sickness occurs. Among these, the most accepted theory is the one presented by Reason and Brand in 1975 [8], further

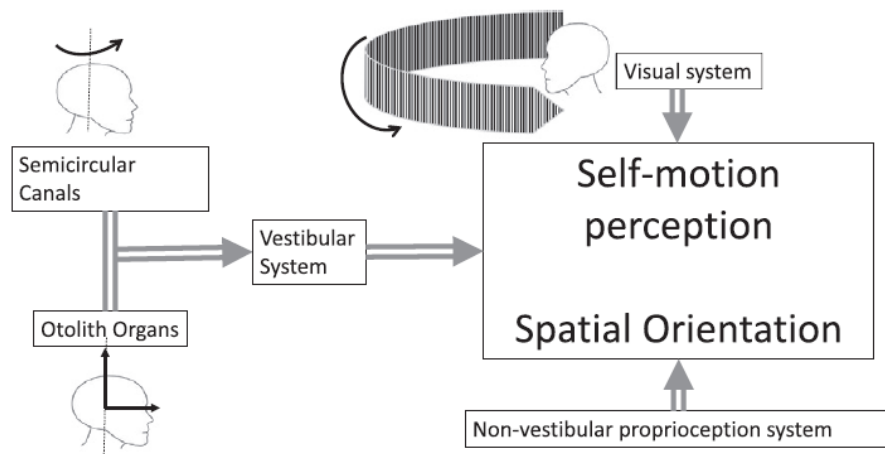


Figure 1.1: Illustration of the interaction between the sensory systems providing the perception of self-motion. The vestibular system, represented by the semicircular canals and otolith organs, distinguishes neck rotational and translational motion. The eyes are stimulated by the movement of a large visual field. Non-vestibular proprioceptive inputs, besides motor control commands contribute to spatial orientation [10].

developed by Benson [9]: the sensory conflict theory, also known as the neural mismatch hypothesis. The general assumption of this theory is that motion sickness occurs due to a conflict in perceptual information coming from visual, vestibular and proprioceptive receptors, which is not in accordance with our expectations. In particular, our nervous system inputs that concern motion originate from 3 different sensory systems, each of them being specific to particular aspects of the motion stimulus. All sensory signals are combined in the brain in order to estimate our motion. However, it sometimes appears that the combination of sensory signals coming from the vestibular, visual and somatosensory systems is judged impossible by our brain, because it deviates from the stored patterns from recent transactions, referred as the sensory conflict. An illustration of the interaction between the sensory systems providing the perception of self-motion is presented in Figure 1.1.

Treisman developed the so-called evolutionary theory, in which he suggests that people experience motion sickness due to the continuous misalignment between the signals from the spatial framework defined by the visual, vestibular and proprioceptive inputs [11]. Treisman theory differs from the sensory conflict one by 2 elements. Firstly, it answers the question of why this misalignment causes symptoms of motion sickness, nausea and vomiting especially. Secondly, it does not consider past experience or expectations, but states that all the 3 sensory systems have to work continuously in parallel, considering only the present sensory inputs. He explains the continuous need for neural inputs from the sensors of the 3 sensory systems, and how a disruption of this activity can result in nausea and vomiting. Actually, unexpected movement or illusion of movement could

be the cause of what would be interpreted as physiological disturbance produced by absorbed toxins. As a consequence, the human body would react with nausea and vomiting for preventing poison from getting into the stomach.

Another theory, known as the postural instability theory, attributes motion sickness to behavior rather than sensory stimulation. Contrary to many theories of motion sickness, this one, developed by Riccio and Stoffregen, does not originate from the sensory conflict theory. The 2 researchers have criticized this last theory, claiming that no sensory conflict exists. Instead, they have proposed that the motion sickness symptoms may be the result of a lost postural and stability control following prolonged instability [12]. Motion sickness would therefore not be due to the movement itself or to the perception of movement by our senses, but related to behavioral issues.

A theory describing visually induced motion sickness has been proposed by Bos, Bles and Groen [13]. It assumes that the onset of associated symptoms is the consequence of a discrepancy existing between the sensed and the expected vertical regarding the head position.

In the following, the sensory conflict theory will be considered to be the one explaining the onset of motion sickness symptoms.

### 1.1.3 What stimuli cause motion sickness ?

As mentioned previously, motion sickness can occur in various modes of transportation as well as in simulators or other virtual environments. While the onset of motion sickness symptoms on boats, planes, cars or trains originates from motion perceived by sensory receptors of the vestibular system and not by the ones of the visual system, it can be the opposite situation in a simulator which is not equipped with a moving platform. In this last case, visual system is signaling a self-motion while vestibular and somatosensory systems are indicating that the person is sitting still.

A useful set of rules that should be observed to avoid motion sickness has been proposed by Stott [14]. These rules aim to avoid sensory conflict and are the following:

- Rule 1: Motion of the head in one direction must result in motion of the external visual scene in the opposite direction. This first rule concerns the accordance between visual and vestibular inputs.

- Rule 2: Rotation of the head, other than in the horizontal plane, must be accompanied by appropriate angular change in the direction of the gravity vector. This second rule aims to avoid the mismatch between canal and otolith inputs.
- Rule 3: Any sustained linear acceleration is due to gravity, has an intensity of 1  $g$  and defines downwards. This last rule deals with the accordance of the inputs of the 2 otolith organs in the vertebrate inner ear, the utricle and saccule.

Although these 3 rules seem simple, it can be difficult to explain the onset of motion sickness in all kinds of environments using them, as the number of stimuli and conflicts that exist can be large. Despite this complexity, it has been shown that low frequency translational motion is a major source of motion sickness in land vehicles, aircraft and boats. In particular, motion frequencies lower than 1 Hz have been proven to be more nauseogenic than higher frequencies, with a peak around 0.2 Hz [15].

Whereas motion sickness susceptibility is related to the motion type and provocative property of stimulus, it is also determined by individual differences. Besides the initial sensitivity to motion sickness, rate of natural adaptation and physiology of sensory systems which are specific to each individual, age and sex are proven to be determinant factors of motion sickness susceptibility. Many surveys indeed indicate that women are more susceptible to motion sickness than men, while infants and very young children do not suffer from motion sickness [14]. The increased motion sickness susceptibility in women is supposed to come in part from the female hormonal cycle. Similarly, there are only hypotheses explaining why young children are immune to motion sickness. Among those, the fact that the vestibular organ is not fully formed could explain the absence of sensory mismatch detection.

Some experimental means of provoking motion sickness exist and are based on relevant disorienting environments. While some methods tend to mimic situations encountered in real conditions [4, 16, 17], others use specific processes in order to induce the onset of motion sickness symptoms. Among this last category, an optokinetic drum can be used. It is a rotating instrument in which individuals are seated facing the wall of the drum. The interior surface of the drum is composed of stripes, usually black and white, and long enough to cover the entire visual field of the person in the center of the drum. Thus, the eyes of this person are subject to a moving visual field while he/she is not moving. This situation often induces motion sickness and its intensity can be increased by varying the speed of the drum and the duration of the test [18].

In autonomous cars, passengers can be engaged in various non-driving activities. These occupations have many consequences and can result in an increase of the occurrence rate of motion sickness. Indeed, engagement in other tasks than driving can lead to an increased number of head movements in different directions of the ones of the vehicle. Moreover, the view of the outside world is precluded, leading to the loss of vehicle controllability and the inability to anticipate the future motion trajectory. The intensity of symptoms could be further increased if the horizontal acceleration variations and the duration exposure are important.

## 1.2 Methods for detecting motion sickness

### 1.2.1 Objective and physiological methods

The main symptom associated with motion sickness is vomiting. The directly preceding signs and symptoms are pallor, sweating, stomach awareness and nausea. These can either be observed with an external eye or be felt by the person suffering from motion sickness. On the contrary, initial events leading finally to vomiting are generally not detectable. This initial disturbance in the vestibular, visual and proprioceptive systems actually triggers a set of autonomic responses that lead to a disorder between 2 subdivisions of the autonomic nervous system (ANS): the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS).

The primary process of the SNS is to stimulate the body fight or flight response [19], it prepares the organism to stress. It is antagonistic to the PNS which leads the body to a relaxation state. Each of these 2 divisions of the ANS initiates a series of physiological processes leading to the desired state. These 2 systems modulate many bodily processes that occur automatically, such as blood circulation, respiration, digestion, temperature maintenance, etc. Together, they control the activity of several organs and functions through their opposing actions. When sensory conflict occurs, the balance between the 2 subdivisions of the ANS is disrupted, resulting in a set of specific autonomic responses.

In the following, some of the most common physiological signals and the corresponding autonomic responses associated with motion sickness are presented.

#### **Cardiac activity**

The activity of the heart can be measured through electrocardiography. This non-invasive technique measures the electrical activity of the heart using electrodes placed at the torso

on the skin, and produces a tracing called electrocardiogram (ECG). The typical appearance of an ECG is presented in Figure 1.2. It typically contains 5 waves. The first wave, the P wave, is due to the depolarization of the atria triggered by the sinus node. The QRS complex formed by Q, R and S waves follows the P wave. It is due to ventricular depolarization and precedes contraction of the ventricles. The T wave is the last wave and it corresponds to ventricular repolarization.

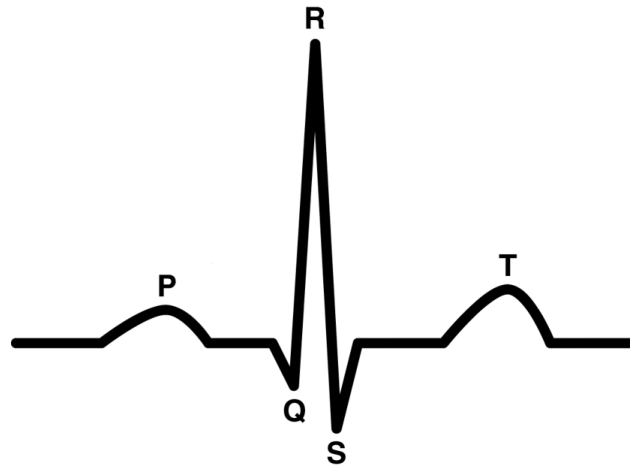


Figure 1.2: Typical appearance of an ECG [20].

Heart rate (HR) is a measurement of cardiac activity defined as the speed of the heart-beat, and it is derived from the ECG. It is calculated by counting the average number of contractions (beats) within a time frame, the most current measure is beats per minute (BPM). The average beat is derived from the R waves of the ECG, representing a single heart beat.

Another famous method for obtaining the heart rate is by using blood volume pulse (BVP) measurement, obtained by the use of photoplethysmogram (PPG) sensor placed at pads of fingers or feet. This method is for instance preferred rather than ECG in biofeedback training situations, because the PPG sensor is easier to apply than ECG. It is, however, less precise than an ECG and subject to measurement errors. The BVP measures heart rate based on the volume of blood passing through the tissues in a localized area with each heart beat. Two consecutive typical BVP waveforms are presented in Figure 1.3. The first peak of a waveform is the systolic peak and results from the ejection of blood into the aorta and pulmonary trunk. The second peak of a waveform is the diastolic peak and results from the filling of the heart chambers with blood. Heart rate is derived from the BVP signal by measuring the interval between the systolic peaks [21].

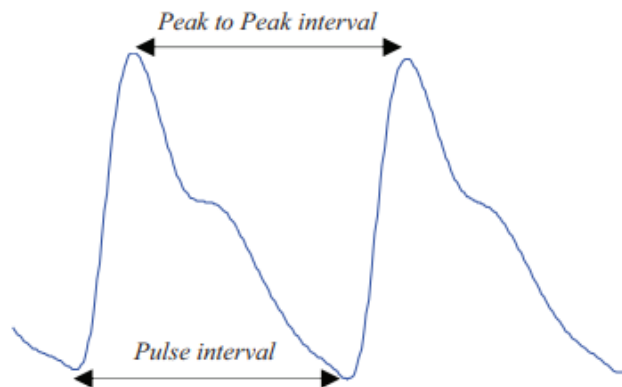


Figure 1.3: Two consecutive typical BVP pulse waveforms [21].

The heart activity can also be studied according to its dominant frequency range through an ECG power spectrum analysis. Low and medium frequencies ( $[0.04-0.15 \text{ Hz}]$ ) are mainly associated to sympathetic activity, while high frequencies ( $[0.15-0.40 \text{ Hz}]$ ) are associated to parasympathetic activity [22].

The R-R variability constitutes another measure of heart rate variability (HRV) and reflects the variability in distance between the R waves. Therefore, a slow heart rate is characterized by a longer distance between the R waves, and vice versa.

During motion sickness, heart rate has been observed to increase as a response of the SNS, and its magnitude depends on susceptibility and duration of the stimulus [23]. On the opposite, a decrease in HRV can be expected since an increase in HR is generally combined with regular heart beats [24].

### Skin conductance

Electrodermal activity (EDA) defines autonomic changes in the electrical properties of the skin, whose typical unit is the microsiemens ( $\mu\text{S}$ ). One of the common studied properties is the skin conductance, a measure of the electrical conductivity of the skin. The EDA complex is composed of 2 main components. The first one is the general tonic-level EDA and refers to the background and slower characteristics of the signal, such as the overall level, slow climbing and declination over time. This component is measured by the skin conductance level (SCL). The second component is the phasic one and refers to faster changing elements of the signal resulting from sympathetic activity. The skin conductance response (SCR) measures this second component. These 2 components vary with the activity of eccrine sweat glands, distributed all over the body, due to stress, arousal

or emotional excitement.

As sweating generally increases during motion sickness, the skin conductance, estimated by measuring skin potential between 2 electrodes, increases as well [25]. It is important to note that EDA is used to register relative changes with respect to a baseline measurement, as the degree of arousal the measurement induces is variable between subjects.

A particularly interesting EDA parameter showing an increase in sympathetic activity is the number of peaks per minute, on the EDA signal [26, 27]. However, extraction of this measure requires signal post-processing, and it is therefore impossible to evaluate it in real-time.

### **Skin temperature**

The skin temperature measurement is, as its name indicates, performed on the skin surface. Normal skin temperature on the body trunk varies between 33.5 and 36.9°C [28]. This measurement is often preferred to the core temperature one because of its simplicity, although this last is not a clear indicator of body internal temperature. The skin temperature is evaluated as change over time, and is varying depending on the considered body part and ambient temperature.

A decrease of the skin temperature is usually observed during motion sickness [29, 30], since sweating leads to an increase of heat loss.

### **Respiration**

The term "respiration" encompasses several concepts: cellular respiration and external respiration. In this context, we are interested in external respiration, which refers to the exchange of gases between the environment and the cells of the body. It can be subdivided into 4 processes [31]:

- Ventilation. The exchange of air between the atmosphere and the lungs is necessary to constantly renew the air in the alveoli. This process is ventilation where inspiration corresponds to the entry of air into the lungs and expiration to its exit from the lungs.
- Lung-blood gas exchange. Oxygen diffuses from the alveoli to the blood, while carbon dioxide ( $CO_2$ ) follows the opposite path.
- The transport of oxygen and  $CO_2$  by the blood.



- Gas exchange between blood and cells. Oxygen diffuses from the blood to the tissues and  $CO_2$  from the tissues to the blood.

When dealing with physiological research, we are interested in ventilation, where the chest expansion and respiration rate are usually measured using a strain gage. Under normal conditions, respiration rate of an adult person at rest ranges from 12 to 15 breaths per minute, while it can reach 50 breaths per minute when physical activity is performed. As a consequence, when measuring respiration responses to a stimulus, one has to ensure that the participant is not under influence of physical activity. In addition, breathing rate can be negatively affected by the instructions that are given to the participant, often leading to non-regular breathing patterns. It therefore complicates the detection of ANS responses, but it still often used for measuring autonomic activity in response to motion stimulus [32, 33, 34].

During motion sickness, the respiratory frequency has been observed to initially increase, then to decrease between the elevated initial response and the end criteria, and can again increase during vomiting [32].

Respiration influences heart rate through the respiratory sinus arrhythmia (RSA). It is a fluctuation of heart rate in synchrony with respiration, by which the R-R interval on an ECG is shortened during inspiration and prolonged during expiration.

### **Eye movements**

As other organs, the human eye responds to changes in autonomic activity. The major contribution to that response is brought by the PNS through the third cranial nerve, inducing a constriction of the pupil. Dilation of the pupil occurs when the SNS dominates.

Vision constituting a major component in the onset of motion sickness, possible ways of reducing the symptoms of motion sickness through eye movements have been identified [35]. However, very few studies deal with the possible relationship between eyelid movements and motion sickness. This point will be addressed in chapter 4.

### **Brain activity**

The brain activity is measured through electroencephalography. It is a non-invasive functional brain imaging technique that measures the electrical activity generated by nerve cells through electrodes placed on the scalp. These electrodes cover the frontal, parietal, occipital and temporal lobes (Figure 1.4) with a variable space between them depending on the number of electrodes used. The electroencephalogram (EEG) is the transcription

in the form of a tracing of the variations in time of the brain electrical activity.

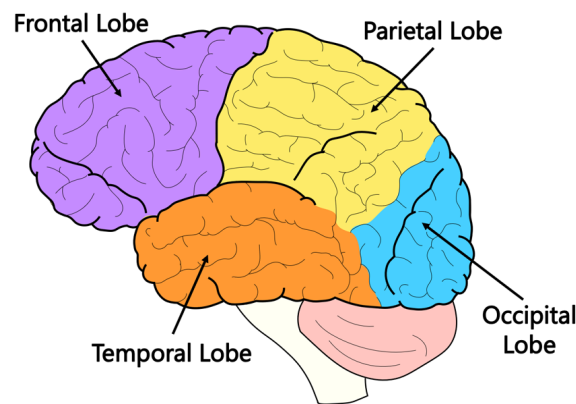


Figure 1.4: Brain lobes [36].

On an EEG tracing, one can observe the cerebral rhythms, which constitute spontaneous signals, i.e., not induced by an external stimulation. Among the cerebral rhythms, different frequency profiles can be identified:

- Alpha rhythms:  $\in [8, 13]$  Hz. These waves are present during a state of alertness or light meditation. They are associated with the coordination of mental activity and learning.
- Beta rhythms:  $\in [13, 35]$  Hz. These waves are present during a waking state when our attention is engaged by cognitive tasks (e.g. making a decision, solving a problem) or by the outside world.
- Delta rhythms:  $< 4$  Hz. These are slow waves. They are present during a state of deep meditation or a state of dreamless sleep.
- Theta rhythms:  $\in [4, 8]$  Hz. These are waves present during deep sleep. These waves play a role in the learning and consolidation of memories.
- Gamma rhythms:  $> 35$  Hz. These are the fastest waves. They are associated with the processing of information by different regions of the brain (i.e., the synchronization of several regions of the brain). They are also present during states requiring a high level of attention or concentration.

EEG power responses have been observed to be highly correlated with subjective sickness levels in different brain regions. Firstly, in the left and right motor brain areas, located behind the frontal lobe, before the central groove separating the frontal lobe from the parietal one. Secondly, in the parietal brain area, and finally in the lateral occipital and occipital midline brain areas [37].

## Gastric activity

Gastric activity is a measurement of stomach movement performed through an electrogastrograph. More precisely, the electrogastrograph records the electrical signals traveling through the stomach muscles and controlling the muscle contractions, and produces an electrogastrogram (EGG) [38]. The activity of gastric smooth muscle is controlled by the ANS and its frequency is around 0.05 Hz (3 cycles per minute) under normal conditions [39].

The EGG appears to be very sensitive to motion and physical activity. Indeed, when gastric activity is studied, electrodes are placed on the stomach surface, and abdominal muscles are therefore separating the electrodes and the stomach. As a consequence, any movement of these muscles results in an artifact on the recorded signal, thus complicating the analysis. This is why this signal is always used together with other physiological measurements as well as with self-reported measures. Relations between an increased gastric activity frequency and subjective scales of motion sickness have been reported in previous research [40].

### 1.2.2 Subjective methods

The need to collect subjective measures of motion sickness in addition to physiological data is evident in order to clearly identify the development of motion sickness in a person subjected to motion sickness stimuli. Indeed, the subject may show autonomic activity related to the one motion sickness can give rise to while reporting no motion sickness. In this case, it cannot be excluded that the observed autonomic activity is the result of something else.

There are various subjective methods for assessing motion sickness severity. While some of them are used to assess the susceptibility to motion sickness based on previous experiences and are generally performed before motion sickness stimuli exposure, others are used to collect dynamic changes in intensity of symptoms during motion exposure. Obviously, a trade-off between the questions specificity and the obtaining of a valid motion sickness-related measurement has to be made, explaining the existence of different scales. All these methods have been recently adapted to simulator sickness.

## Questionnaires

The most common motion sickness questionnaire (MSQ) used is the Pensacola questionnaire. As others, this questionnaire records individual exposure to motion in diverse forms

of transports, including cars, buses, coaches, trains, small boats, ships and airplanes, as well as the occurrence of illness and vomiting in these forms of transports. Depending on the questionnaire that is used, the responses must relate to childhood, the previous 10 years, or only the year preceding the completion of the questionnaire. The most commonly developed symptoms (e.g. headaches, sweating, drowsiness, dizziness, nausea, vomiting), the types of transport avoided as well as a self-assessment of susceptibility to motion sickness are usually collected in these questionnaires. The questionnaire responses can be used for calculating several measures related to motion exposure and motion sickness susceptibility [41].

### Self-evaluation scales

Illness rating scales are mainly used during motion sickness exposure to evaluate the severity variability of motion sickness symptoms. Generally, these scales are presented to subjects who respond verbally. Depending on the scale, the participants determine the severity of symptoms either through numbers only, either through words associated with a rating.

A commonly used rating scale is the misery scale (MISC). It is a 0-10 scale that has been adapted several times, the 2011 version is shown in Table 1.1.

Symptom		Score
No problems		0
Uneasiness (no typical symptoms)		1
Dizziness, warmth, headache, stomach awareness, sweating	Vague	2
	Slight	3
	Fairly	4
	Severe	5
Nausea	Slight	6
	Fairly	7
	Severe	8
	Retching	9
Vomiting		10

Table 1.1: Misery scale [42].

The Borg scale constitutes another famous rating scale and is presented in Table 1.2.

•	Absolute maximum
10	Extremely strong
9	
8	
7	Very strong
6	
5	Strong
4	
3	Moderate
2	Weak
1	Very weak
0.5	Extremely weak
0	Nothing at all

Table 1.2: Borg rating scale [43].

## 1.3 State of the art in the field

A lot of studies have been conducted to better understand the physiological phenomena caused by motion sickness. The increasingly frequent use of driving simulators has also led to the study of simulator sickness. In the following, various protocols that have caused motion sickness are presented. Recruitment criteria for participants are first discussed. The tests are then presented, including the equipment used, the physiological signals recorded as well as the subjective scales considered.

### 1.3.1 Exclusion criteria of participants

In order to ensure the smooth running of the experiment and the representativeness of the population, some criteria must be met by the people wishing to participate in the protocol. In particular, in a protocol inducing motion sickness, participants should have no history of gastrointestinal, cardiovascular and vestibular disorders, no history of drug or alcohol abuse, and should have normal or corrected-to-normal vision [17, 37].

Before the test, potential participants are often asked to complete various questionnaires dealing with their susceptibility to motion sickness, their past experiences as well as their gender and age [7, 16], in order to ensure that the sample of participants has a non-zero susceptibility to motion sickness, and is composed of a balanced number of men and women, as well as of people of very different ages. Once the participants have been chosen, they are required to complete and sign documents related to the progress of the protocol and the processing of their data, as well as related to the guaranty of anonymity.

### 1.3.2 What kind of equipment is used ?

Studies conducted to investigate motion sickness in autonomous cars mainly use 2 types of equipment, either one or both in the same study. Indeed, the experience can take place in a car, where the participant is usually a passenger, but also in a driving simulator, where the participant acts as either the driver either the passenger.

The driving simulator presents several advantages compared to a real car [29]. First of all, the study can be conducted in a highly controlled environment, while in a car, the experimenter has no full control of the dynamic events that happen during the scenario. Then, repeatable scenarios eventually tailored to a particular research question can be created in a driving simulator. As a result, a scenario either with a lot of strong lateral and longitudinal accelerations or with only few accelerations can be carried out. On the contrary, external factors including traffic or weather conditions affect the scenario in a real car. Indeed, these factors cannot be kept constant and therefore induce a scenario variability affecting the onset of motion sickness symptoms. Another advantage of the driving simulator is the possibility to create any driving style, including some that would not be possible in real conditions. This makes it possible to test innovative scenarios while ensuring the safety of all. Also, some data such as velocity, acceleration, surrounding traffic and road geometry can be monitored and recorded in a driving simulator and not in a real car.

The major disadvantage of the driving simulator is that generated motion sickness can be confounded with simulator sickness. This syndrome is similar to motion sickness in many ways, but occurs in simulated environments and can be induced without motion. Actually, the sensation of self-motion in absence of physical motion, calledvection, is usually induced in a driving simulator with no motion platform because of the large visual field observed by an immobile observer, and can result in motion sickness.

### 1.3.3 How to induce motion sickness ?

Various protocols exist to induce motion sickness. Independently of the equipment used, all protocols generally include at least one part which aims to record the physiological data of the participants under normal conditions, i.e., without the induction of motion sickness. When the participant is a passenger in a car, he/she usually has no task to perform during this section, allowing him/her to look at the road, which has usually a straight trajectory. If the participant is the driver of a driving simulator, the driving is generally done on a straight line, and without events causing the participant to brake. Depending on the protocol, this session can be part of the motion sickness scenario [37]

or be conducted at a separate time [16]. In the first case, the protocol usually begins and ends with a session of this type, allowing on one hand the participant to acclimatize to the environment, and, on the other hand, to recover from his/her sickness.

Concerning the part of the scenario aiming to induce the onset of motion sickness symptoms, it is very different depending on the equipment used. In a real car, the participant has to perform some tasks and has no longer any vision of the external world. The car trajectory includes a lot of braking events and turns, further inducing conflict between all the sensory signals [16, 17, 29]. The tasks can either be performed on a paper or on a handheld device and can include reading comprehension, visual search, text entry, pattern recognition, video watching, etc. In order to ensure that the scenario experienced by one participant is similar to the one experienced by another subject, the driver of the car can use, if available, advanced driver assistance systems allowing to automate certain parts of the driving experience. Moreover, if several drivers are designated, they are trained beforehand so that their driving behavior is similar.

When the scenario takes place in a driving simulator, the session aiming to induce motion sickness can be composed of long winding road, with short turns and braking events. A distinction has been made between the type of simulators. Indeed, some simulators are equipped with a motion platform that has generally 6 degrees of freedom, allowing therefore to mimic the motion of a real car. In that case, the participant can be, as in the real car, a passenger and has some tasks to perform [29]. When the simulator is not equipped of a motion platform, there is no sense to perform tasks for distracting the participant from the road, the situation would then revert to performing tasks seated, without motion stimuli. As a consequence, the participant has to be the driver in that case, andvection is induced since he/she is exposed to a large visual field without physical motion.

### 1.3.4 What is recorded ?

As mentioned previously, both physiological (therefore objective) data and subjective rating are generally collected during motion sickness studies. The physiological data collected varies from one study to another, as well as the way of collecting this data and the moment of subjective ratings assessment.

Physiological data are usually collected through wearable sensors, whose electrodes are placed in a non-invasive way, at the body surface. Among others, heart rate, skin

impedance and temperature, brain and gastric activities are commonly recorded physiological signals. Moreover, facial features, head movements and participant posture can be recorded thanks to 3D depth cameras [16].

Concerning the collection of subjective data, the completion of motion sickness questionnaires is performed before the scenario in some studies [16], while some others recommend that this document be completed after the scenario [29]. The subjective rating collection is also subject to variation depending on the considered study. First of all, participants have to give a rating of their symptoms at regular intervals in some studies, while some other studies ask the participants to give a rating only when there is an intensity variation of the symptoms of motion sickness. Then, the way of collecting this data is also different from one study to another. Sometimes, a joystick is provided to the participant, and they have to encode their level of motion sickness using this joystick. More often, participants have to give a motion sickness score orally.

In general, it naturally appears that the motion sickness score is higher when the participant is performing tasks [16, 17]. It is also higher in people under the age of 60 [16]. In addition, this score is often correlated with the level of activity of the sympathetic nervous system, resulting in variations in the signals discussed in section 1.2.1.



# Chapter 2

## Motion sickness protocol

This second chapter concerns the protocol, including equipment and methods, used for motion sickness induction and data acquisition. The recruitment of participants is first discussed. After that, the course of acquisition sessions is covered, followed by a description of the collected data and related equipment.

### 2.1 Recruitment of participants

The experimental protocol was divided into 2 sessions. The first one was conducted in the fixed-base driving simulator at the University of Liège, where the participant was the driver. The second session took place in a real car, where the participant was the passenger. A total of 12 women and 8 men participated in each of these 2 sessions.

The inclusion and exclusion criteria for participants in this acquisition were as follows:

- Be between 20 and 70 years old;
- Have a driving license for at least 2 years and drive more than 5,000 km per year;
- Have a good state of health (not be epileptic, not have heart, vestibular or gastrointestinal problems, not be suffering from a chronic disease, etc.);
- Do not wear prescription glasses;
- Do not have skin allergies relating to cosmetics or lotions.

The 2 last criteria are especially related to the acquisition of physiological data. Indeed, ocular parameters have been recorded through glasses, presented in section 2.3.1, and several physiological data have been collected using electrodes with a sticky substance on the skin.

In addition, it has been asked to the participants to not have eye makeup and to not eat during the 3 hours preceding the 2 acquisition sessions, this to ensure optimal physiological data collection.

## 2.2 Session progress and performed tasks

As mentioned previously, the experimental protocol consisted in 2 distinct sessions: a simulator session and a car session. Actually, the way of inducing motion sickness is very different from one session to the other. While the car session aimed to induce motion sickness in the same way that it would be induced in self-driving cars, and so to induce autonomous car sickness in a sense, the simulator session aimed to observe the similarity or not of the symptoms and physiological responses of the participants who demonstrated a state of sickness resulting from the simulator driving. As a consequence, for the following, we will keep in mind that the car session is more related to autonomous car sickness than the simulator session.

Both sessions was divided into 3 sections. The aim of the first and the last sections, called the baseline section and the rest section respectively, was to record the physiological data of the participants under normal conditions, i.e., without induction of motion sickness. These 2 sections lasted approximately 15 minutes. On the contrary, the purpose of the second section, called the motion sickness section, was to induce the onset of motion sickness, and therefore physiological responses that are specific to this state. The motion sickness section lasted 30 minutes. The whole session therefore lasted about 1 hour.

### 2.2.1 Driving simulator session

#### Simulation apparatus

The driving simulation experiment was performed using the fixed-base driving simulator at the University of Liège, presented in Figure 2.1. The driving simulator cabin consists in a car seat, a manual gearbox, a steering wheel, and foot pedals. Three front-view large screens provide 135 degrees driver field of view. These components are interfaced with STISIM Drive®<sup>®</sup>, a specialized simulation environment developed over many years.

In addition to the driving simulator, there is also a user station allowing the user to configure the graphics system, sounds, driving controls and scenarios via the STISIM Drive software. Moreover, it allows the user to retrieve the wide range of data that can be obtained from the simulator such as speed, lateral and longitudinal positions, acceleration, crashes, red-light violations, speed violations, inputs of pedals, etc.



Figure 2.1: Fixed-base driving simulator at the University of Liège [44].

### Scenario design

A STISIM Drive simulation run is composed of 2 primary elements:

- The configuration file, setting all the global parameters for the simulator;
- The scenario file, describing what is happening in the roadway scene.

The configuration file allows to specify more than a hundred parameters. Among those, vehicle dynamics, driver controls, simulation sounds, crash options and roadway settings can be controlled and modified through a configuration dialog window. As a consequence, it is for instance possible to use an autopilot mode, to adjust the way the steering wheel will work during a simulation drive, to define the position of the vehicle cab or to use different transmission gears.

The most important component of the simulator is the roadway scenario that is used. Indeed, it defines the simulation experience through dictating what the driver sees during the simulation run. It also specifies the measures that will be collected during the run.

STISIM Drive® Scenario Definition Language (SDL) is a software allowing to design custom roadway environments and situations. The definition of a scenario is basically composed of 6 steps. The first one consists in defining the data to collect. It actually determines the critical events and sequences of the scenario. For instance, it could be necessary to collect the driver reaction time, an event such as a pedestrian who suddenly crosses the road could therefore be defined. The second step is about the design of the critical events, so that these can lead to the type of data the user wants to collect. After

that, the layout of the roadway has to be determined. This third step consists in defining the locations and characteristics of intersections, curves and hills. The next step is the addition of critical events to the roadway, with some iterations so that roadway and events are designed properly. Events include for instance the sudden stop of the car in front of the driver, pedestrians crossing the road, red lights, etc. A lot of parameters can be adjusted, among those the speed of cars and pedestrians, the type and the color of cars or the distance from the user at which these appear. When these critical events are defined, the infrastructure has to be added to the scenario. It consists in various static objects such as buildings, trees, road signs, etc., and allows the scenario to be realistic. Again, a lot of parameters can be adjusted for adding these static objects, such as the distance from the road, the type of buildings and trees, their density. Finally, the last step consists in adding some benign events, i.e., that must not cause a driver reaction, in order to make the scenario interesting.

The scenario that has been proposed to the participants was composed of 6 distinct sections (see Figure 2.2), and lasted approximately 1 hour, depending on the speed at which the participants were driving. The simulation context was replicating the European road layout and surrounding environment, i.e., road markings, roadside signs, etc., such that the participant was immersed in an as real as possible environment. The whole scenario was designed with daylight and favorable weather conditions enabling a distant view.

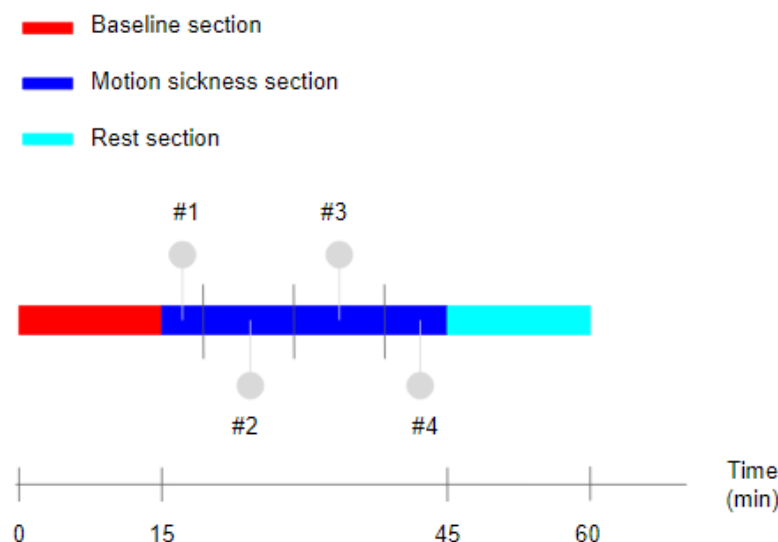


Figure 2.2: Illustration of the 6 sections composing the driving simulator scenario.

The first and the last sections consisted in the baseline and the rest sections, respectively. The first section was composed of 20 km on a 3-lane roadway, with a speed limit fixed at 120 km/h, followed by 7.5 km on a 2-lane roadway, where the speed limit dropped from 120 to 90 km/h. The last section was similarly composed but in the reverse order, i.e., the 90 km/h section was before the 120 km/h one. In the 2 cases, the roadway was mainly straight, with very few wide turns.

The 4 sections between the baseline section and the rest one were composing the motion sickness section. Therefore, the roadway was composed of sharp turns and braking, also induced by pedestrians crossing the road or red lights. The speed of the first and third parts of this motion sickness section was limited to 70 km/h, while the one of 2 others was reduced to 50 km/h. The first motion sickness section was 4 km long, and many country houses were by the side of the road. The second section was 8 km long and took place in a city with many big buildings. The scenery of the next section, which was 10 km long, reproduced that of mountains, with hills and steep slopes on the sides of the road. Finally, the last part of the motion sickness section included a passage through a construction zone, with a lot of construction materials and workers, and was 7 km long.

### 2.2.2 Car session

During the car session, the subject was installed on the front passenger seat. As mentioned previously, the car session was also divided into 3 parts, the car trip was therefore determined based on these parts. While the road was mainly straight and flat during the baseline and the rest sections, it was composed of a lot of sharp turns and hills during the motion sickness part. The route is presented in Appendix A.1.

During the baseline and the rest sections, the participant had no task to perform, allowing him/her to look wherever he/she wants. On the contrary, he/she had to perform some tasks on paper during the motion sickness section, preventing him/her from looking at the road. The proposed tasks were relatively simple, because their only purpose was to hold the attention of the participants. These are presented in Appendix A.2. They had to complete the tasks in the following order, and were not under pressure to complete all the tasks during the motion sickness section: reading comprehension, hidden words, text entry, message decoding, pattern recognition and rebus.

## 2.3 Collected data and equipment

Various data have been collected during the acquisition sessions in order to identify motion sickness in participants. In particular, physiological and subjective data have been recorded through different means, these are presented in the following. While the collection of subjective data can be done very simply, the collection of physiological data requires specialized equipment. Those used at Phasya are also presented in this section.

### 2.3.1 Physiological data

Several physiological parameters appear to be strongly affected during motion sickness. As mentioned previously, heart rate, skin impedance and temperature, breathing rate, and brain and gastric activities constitute a set of physiological signals that are affected when a sensory conflict occurs. As a consequence, all these data were recorded during the acquisitions, except breathing rate because of the superimposition of the breathing rate sensor and one of the electrodes used for gastric activity recording. This would indeed lead to bad quality gastric activity recordings. The choice to record the gastric activity rather than the breathing rate comes from my personal experience and that of the Phasya team, we have indeed felt more disorder in the stomach activity than in breathing rate when we experienced motion sickness. In addition, eye features have been collected since a main objective of this work is to determine whether or not ocular parameters are indicators of motion sickness.

Equipment for collecting physiological data that Phasya owns has demonstrated reliable and good quality data recordings during previous acquisitions. This equipment was therefore used for motion sickness acquisitions.

#### *BIOPAC*

Most of the equipment used at Phasya comes from *BIOPAC Systems, Inc.* company. This biotechnology company has been founded in 1985, and is recognized around the world as a premier choice for life science hardware and software [45]. *BIOPAC* provides, among others, research system solutions for acquiring and analyzing biometric data for a full range of applications and signals [46].

The solution provided by *BIOPAC* to record and analyze physiological data usually consists of 4 parts:

- *AcqKnowledge*, software running on a stationary PC or a laptop. It is used for real-time monitoring and physiological data analysis. This software allows to easily

edit data, cut and paste sections of data, perform mathematical and statistical transformations, export data to other applications, etc.

- Research System, such as the MP160 system or B-Alert. It performs acquisition and analysis of life science data by using *AcqKnowledge* software. The MP160 research system includes the MP160 data acquisition unit and the AMI100D transducer module, and is compliant with any Ethernet ready 64-bit computer running Windows or Mac. Both B-Alert and MP160 are fully integrated within *AcqKnowledge*, and one click of button starts both devices.
- Module optimized for acquisitions of one or two types of physiological data, either composed of a matched transmitter and receiver if the data transmission is wireless (BioNomadix wireless transmitter and receiver, with Bluetooth transmission mode), or of a smart amplifier. The receiver is directly connected to the MP160 unit, while the smart amplifier also but via the AMI100D transducer module.
- Disposable electrodes that have to be placed at different locations of the body according to the physiological signal that has to be recorded, and corresponding lead set, directly connected either to the transmitter or to the smart amplifier.

Except the software, all the different parts composing a solution provided by *BIOPAC* are presented in Figures 2.3, 2.4, 2.5, 2.6, while a sample system overview of the MP160 with BioNomadix is illustrated in Figure 2.7.

## EEG

For EEG recordings, B-Alert X10 wireless EEG system is used at Phasya. On the contrary to all the following physiological signals, the record of EEG does not require the MP160 acquisition unit, the AMI100D transducer and the acquisition module. Indeed, B-Alert research system includes all these parts in a single device. The B-Alert wearable system uses Bluetooth as transmission mode. An illustration of the solution for EEG recordings is presented in Figure 2.8.

The system includes a sensor strip that applies all sensors simultaneously with coverage of the brain executive functions (prefrontal cortex), motor control (motor-strip), and visual system activity (parietal region) [51]. The B-Alert X10 sensor headset is composed of 10 sample channels at 256 Hz each: 9 for brain sites (see Figure 2.9) and one optional channel for ECG, electrooculogram (EOG) or electromyogram (EMG). The gain of the channels of the 9 brain EEG sites is fixed and referenced to the linked mastoids (see Figure 2.10), while the differential channel has a programmable gain.



Figure 2.3: MP160 data acquisition unit and Figure 2.4: Wireless BioNomadix module transducer module (AMI100D) [47]. [48].

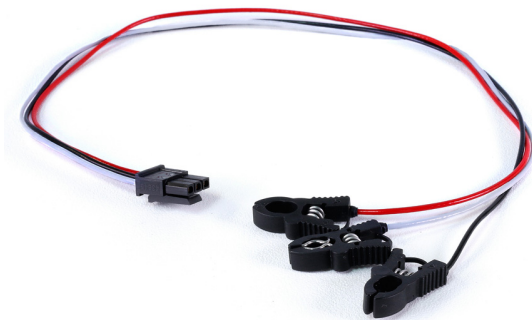


Figure 2.5: Lead set [49].



Figure 2.6: Disposable electrodes [50].

## ECG

The dual wireless Respiration and ECG BioNomadix module pair is used for recording the electrical activity generated by the heart (and for respiration records, when it is needed). It consists of a matched transmitter and receiver pair optimized for respiration and ECG data. Physiological signal data is transmitted at a rate of 2,000 Hz, providing high resolution waveforms at the receiver module output. Moreover, raw ECG data is band-limited from 0.05 to 150 Hz in order to provide a high quality recording. The largest ECG signal than can be measured is 10 mV peak to peak (P-P), while the smallest is less than 1  $\mu$ V root mean square (RMS). The module pair has a fixed gain of 2,000 [48]. These units interface with the MP160 data acquisition and analysis platform and *AcqKnowledge* software.



### Sample System Overview MP160 with BioNomadix

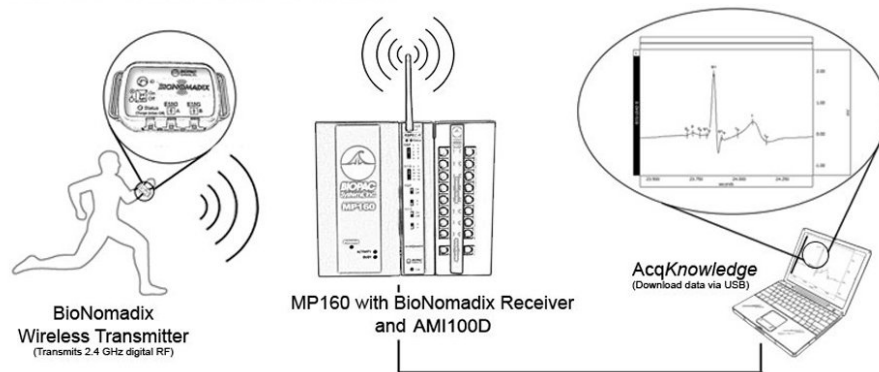


Figure 2.7: Illustration of the 4 parts composing the solution provided by *BIOPAC* [48].

### Sample System Overview B-Alert X10

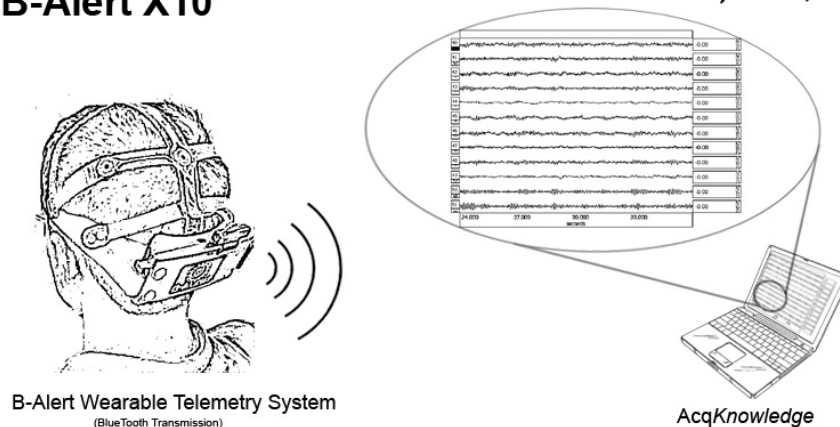


Figure 2.8: B-Alert X10 wireless EEG system [51].

The electrical activity of the heart is recorded through a 3-lead set, and disposable electrodes. The lead set is directly connected to the wireless BioNomadix transmitter, attached to the subject through a Velcro strap.

### Electrodermal activity

The dual wireless PhotoPlethysmogram and Electrodermal Activity BioNomadix module pair is used for recording skin sweating activity (and blood volume pulse when it is needed) and consists in a matched transmitter and receiver. As the ones concerning the ECG, these units interface with the MP160 platform and the *AcqKnowledge* software. The data is also transmitted at a rate of 2,000 Hz, and raw data is band-limited from 0 to 10 Hz, in order to provide rapid response measurement [52]. A 2-lead set and disposable electrodes

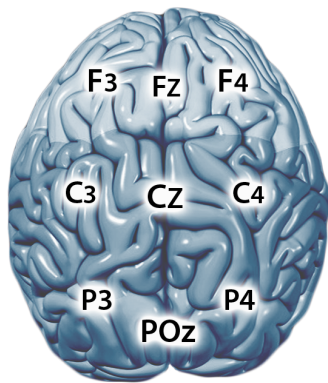


Figure 2.9: Sample channels at 9 brain sites [51].



Figure 2.10: Mastoids illustration, fixing the gain of the channels of the 9 brain EEG sites.

are used for recording the EDA and are generally placed on the fingertips of the left or right hand, and the transmitter on the subject wrist. The module pair can also be used to measure EDA data on the toes or other body locations. At Phasya, the electrodes are generally placed on the right scapula, and the transmitter on the upper torso, in order to avoid as much as possible the movement artifacts generated by the movement of the arms.

### Gastric activity

Electrogastrogram signals are recorded through the EGG100D Smart Amplifier, illustrated in Figure 2.11. It amplifies the electrical signal resulting from stomach and intestinal smooth muscle activity and monitors the potential on the skin surrounding, or surface of, the intestine and stomach, which is indicative of the degree of slow wave contraction [53]. Raw EGG data is band-limited from 0.005 to 1 Hz. The smart amplifier has a fixed gain of 2,000. It features a 3-meter cable that connects to the AMI100D Amplifier Input Module, itself directly connected to the MP160 platform. As for ECG recording, a 3-lead set and disposable electrodes are used for EGG recording.

As mentioned earlier, it has been asked to the participants to not eat during the 3 hours preceding an acquisition session. Indeed, the purpose of this study was to identify physiological responses resulting from motion sickness. Gastric activity being modified during digestion process, this could have been considered as a response to motion sickness when it was not.

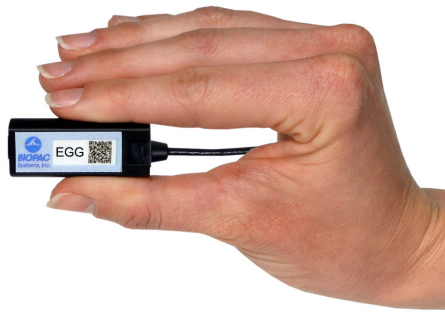


Figure 2.11: EGG100D Smart Amplifier [53].

### *E4*

In addition to the solutions provided by *BIOPAC*, Phasya also owns the *E4* wristband, presented in Figure 2.12. It is medical-grade wearable device offering real-time physiological data acquisition [54]. The *E4* is equipped with various sensors:

- A PPG sensor measuring blood volume pulse and from which heart rate variability can be derived;
- An EDA sensor measuring the constantly fluctuating changes in certain electrical properties of the skin;
- A 3-axis accelerometer, capturing motion-based activity;
- An infrared thermopile, reading peripheral skin temperature.

Moreover, it has an event mark button that allows to tag events and link them to physiological signals, as well as an internal real-time clock.

The *E4* wristband can either be used in recording mode, either in streaming mode. In the first case, the wristband is worn and simply records data during at most 60 hours with 5 seconds synchronization resolution. Recorded data can then be imported via USB and transferred to the cloud platform, where the user can view and manage the collected data. It is also possible to download raw data in CSV format for processing and analysis. For streaming mode, the *E4* wristband connects to a smartphone or a tablet via Bluetooth. It is therefore possible to view collected data in real time. Data is automatically uploaded on the cloud platform after a session ends.

At Phasya, the *E4* wristband is generally used in the same time than the *BIOPAC* solutions. Since these 2 systems collect similar data, this ensures in particular that the data will still be collected if *BIOPAC* should encounter an operating problem. It is also



Figure 2.12: *E4* wristband [54].

sometimes interesting to compare the recording of the same physiological data by the 2 solutions. Indeed, the data is not collected at the same body location, and this can lead to signal differences.

### Drowsimeter

For recording the eye features, Phasya has its own solution, called the Drowsimeter R100 [44]. Indeed, due to its great interest in detecting drowsiness, the company has developed a complete solution for measuring drowsiness and eye metrics in an objective, automatic and accurate way. The Drowsimeter R100 is composed of Phasya Glasses, a laptop and the *Drowsilogic* software. The hardware is illustrated in Figure 2.13.



Figure 2.13: Phasya Glasses and laptop on which the *Drowsilogic* software is installed [44].

The Phasya Glasses are equipped with a high-speed camera, allowing the acquisition of eye images at 120 Hz. Positioned so as not to obstruct the user view, this camera ensures accurate and real time measurements of various eye metrics, such as eyelid movements

and eyeball movements. The data outputted by the Drowsimeter R100 includes a lot of metrics such as the mean blink duration, the blink frequency, the eyelid gap, the pupil position and the pupil diameter, and also eye images. An illustration of the use and outputs of the Drowsimeter R100 is presented in Figure 2.14.

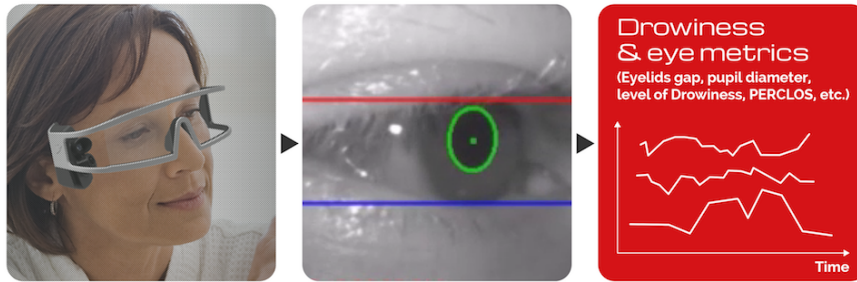


Figure 2.14: Use and outputs of the Drowsimeter R100 [44].

### 2.3.2 Subjective data

#### Questionnaire

When recruiting participants, they were asked to complete a questionnaire about their susceptibility to motion sickness. Most of the questions included in this questionnaire were related to the year 2019, for which participants were asked to fill in the number of journeys as a passenger in different types of transport, the number of times they suffered from motion sickness, and the symptoms commonly developed. In addition, they had to self-assess their susceptibility to motion sickness compared to the general population. The main purpose of this questionnaire was to identify 20 potential participants with different sensitivities to motion sickness, in order to ensure the representativeness of the population within the selected sample.

#### Self-evaluation scale

As mentioned previously, there is a need to collect subjective measures of motion sickness during motion sickness exposure in addition to the physiological data, in order to verify that the autonomic responses observed are resulting from motion sickness. Therefore, for each participant, the self-assessment of motion sickness level, on a scale of 0 (nothing) to 10 (absolute maximum, need to stop the session) has been collected every 5 minutes, starting at the beginning of the motion sickness section and ending at the end of the rest section. This self-assessment measure will be referred as MS score in the following. In addition, they were asked in the same time to describe their feelings and sensations through

a list of common physical responses during motion sickness that was presented to them. This list is included in Appendix A.3. This collection of self-evaluation measures made it possible to directly identify the participants developing motion sickness symptoms and to measure their evolution along the acquisition session.

From this subjective measure, it appeared that the car session has caused much more onsets of motion sickness than the simulator session. Indeed, only 3 participants reported motion sickness scores equal to 8 or more in the simulator, while this number increases to 9 for car sessions. Considering motion sickness scores from 4 to 7, 5 participants reported such a score during simulator sessions, while 8 participants are concerned for the car sessions.

# Chapter 3

## Data analysis

This chapter deals with the analysis of acquired data, including physiological and subjective data. First of all, the processing that has been performed on physiological signals is discussed. After that, analysis of resulting signals is performed, by considering by-participant results as well as global ones, for both simulator and car sessions. Finally, a ground truth for machine learning models that will be considered in chapter 4 is defined, based on the preceding analysis.

### 3.1 Synchronization and processing

Raw data have been collected from all the equipment used for the acquisition sessions. The clocks of the 2 laptops running the *AcqKnowledge* and *Drowsilogic* software respectively have been synchronized at the beginning of each acquisition day.

The processing of the data recorded with *BIOPAC* has been done either directly on *Matlab*, or first with *AcqKnowledge* and then on *Matlab*. The processing of the data recorded with *E4* has been done directly on *Matlab*. Only the most relevant physiological data have been processed, it will be detailed in the following. For confidentiality reasons, *Matlab* codes are not included in this report.

#### 3.1.1 Relevant physiological data

Among all the physiological data collected, only the most commonly correlated with motion sickness were processed and analyzed. Therefore, cardiac, electrodermal and gastric activities have been considered. Although skin temperature is recorded and used as an indicator of motion sickness in many studies [29, 30], it was not analyzed in view of the non-repeatable conditions between participants. Indeed, the temperature of the room in which the simulator is located was not controlled, nor was the temperature inside the

car, and the car windows were sometimes opened, when participants suffered from severe motion sickness symptoms.

### Cardiac activity

ECG signals recorded with *BIOPAC* have been first processed with *AcqKnowledge*. In particular, the goal of this processing was to extract the time interval between successive beats of the heart, called interbeat interval (IBI). This physiological parameter is usually expressed in milliseconds (ms), and its value varies from beat to beat under normal conditions. This variation is referred as heart rate variability. However, cardiac conditions and illnesses can result in a nearly null HRV because of nearly constant IBI values.

To extract IBI from the ECG signal, the data has been first downsampled from 2,000 Hz to 500 Hz, allowing to reduce the time needed for the next processing operations while keeping all the relevant information of the signal. Several steps have then been performed, based on the *Application Note 233 – Heart Rate Variability – Preparing Data for Analysis Using Acqknowledge* [55] manual. A band-pass filter has been firstly applied, with low and high cutoff frequencies respectively equal to 0.5 and 35 Hz, with a Blackman window. Secondly, the data has been transformed by using the template correlation function, resulting in a new signal where QRS complexes can clearly be identified. The R-R interval tachogram, a graph illustrating the evolution of R-R intervals with time, is thirdly created by performing a HRV spectral analysis on the new signal, from which corresponding IBIs and related parameters can be obtained either directly with *AcqKnowledge*, or on *Matlab*. This second method has been preferred since this is the one used at Phasya. An illustration of the signals obtained after the first and second steps is presented in Figure 3.1, while the signal obtained at the third step is illustrated in Figure 3.2.

R-R interval tachograms are then processed on *Matlab*, from which HR and HRV parameters are extracted. From Figure 3.2, it appears that there are some artifacts in the generated tachogram, illustrated by spikes almost reaching 2 seconds of R-R intervals. This would lead to a rapid drop of subject heart rate from approximately 60 BPM to 30 BPM before settling back to 60 BPM, which is not physiologically possible. These artifacts can actually be explained by the fact that a beat was not detected. The time between 2 beats has therefore been doubled. In order to get rid of these artifacts, IBI values that are more than 4 times the standard deviation away from the median and from the mean have not been considered for the extraction of HR and HRV parameters respectively.





Figure 3.1: Examples of downsampled ECG (500 Hz), band-pass filtered ECG (0.5-35 Hz) and template signal obtained using *AcqKnowledge*.

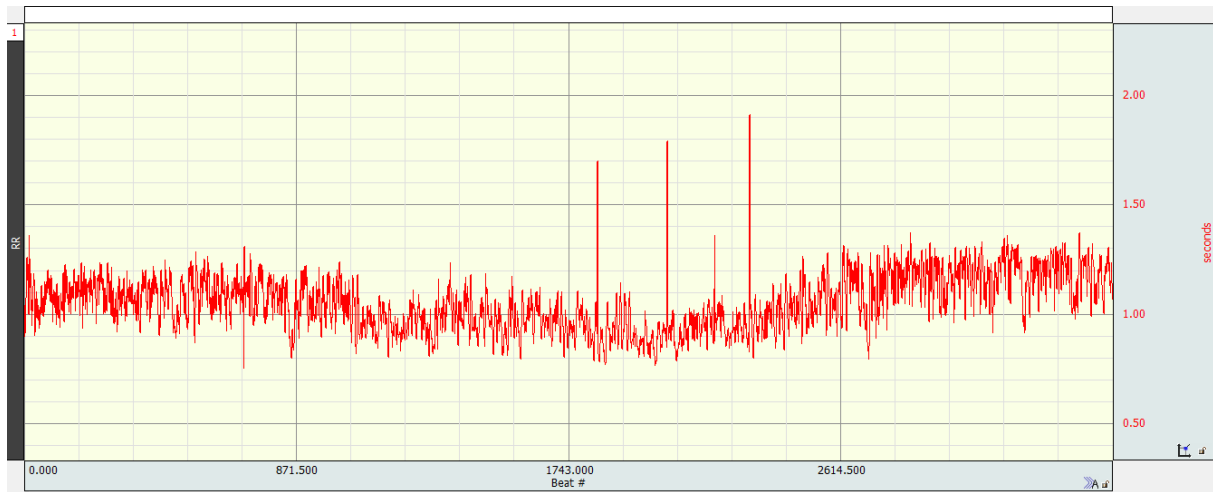


Figure 3.2: Example of a R-R interval tachogram, illustrating the evolution of R-R interval with time. The tachogram is displayed with the beat number along the horizontal axis and R-R time interval on the vertical axis.

Heart rate can easily be deduced from IBI for any pair of beats. Indeed, one has simply to divide the number of milliseconds in 1 minute (60,000) by the IBI (expressed in ms) for obtaining the number of beats per minute, referred as the heart rate. Illustrations of raw and processed IBI and deduced HR are presented for several participants in Appendix B.1, Figures B.1 and B.2 for simulator and car sessions, respectively.

Several time-domain HRV parameters have been extracted from the tachogram [56], these quantify the amount of HRV observed during recording periods, and have been computed every minute, over a 3-minute window:

- pNN20: percentage of successive R-R intervals that differ by more than 20 ms;

- pNN50: percentage of successive R-R intervals that differ by more than 50 ms;
- RMSSD (root mean square of successive differences): square root of the mean square of the sum of all differences between successive R-R intervals.
- SDRR: standard deviation of R-R intervals.

Higher values of all these 4 parameters indicate increased parasympathetic activity. HRV indeed increases during relaxing and recovering activities, and decreases during stress, e.g. during motion sickness. There are many more HRV parameters, including time and frequency domain measurements, but many of them are highly correlated and redundant.

It is important to keep in mind that there is a variability between subjects in the values of the HR and HRV parameters. To give an idea about these values, the ranges of average HR values and the 4 HRV parameter values over a time period (later detailed) that have been observed in this study are presented in Table 3.1. As a consequence, relative differences existing between cardiac parameters of a subject according to a particular section of the protocol will be considered in the following.

Cardiac measure	Range
HR (BPM)	51 - 102
pNN20 (%)	0.37 - 89.9
pNN50 (%)	0 - 70.4
RMSSD (ms)	6.6 - 120
SDRR (ms)	12.5 - 143

Table 3.1: Ranges of average HR values and considered HRV parameter values over a time period of all the 20 participants, during both simulator and car sessions.

### Electrodermal activity

Both EDA signals coming from *BIOPAC* and *E4* have been processed and then analyzed. Indeed, this made it possible to compare the 2 signals, and to consider the best of the 2 for the global analysis. All the processing steps have been performed on *Matlab* and are based on the *GSR pocket guide* from *iMotions* [57].

EDA signals coming from *BIOPAC* have been downsampled from 2,000 Hz to 10 Hz, while the sample rate of EDA signals captured by *E4* is 4 Hz and did not require a downsampling. The phasic component of each EDA signal has then been extracted by applying a basic median filter. In particular, for each sample, the median galvanic skin response (GSR) score of the surrounding samples has been computed based on a  $\pm 4$

second time interval centered on the current sample. The GSR score average has then been subtracted from the current sample, resulting in the phasic data of the signal. This subtraction procedure acts actually as a form of normalization of the EDA data. Raw and phasic data of a subject are illustrated in Figure 3.3. From this Figure 3.3, negative values of the phasic component can be observed, meaning actually that the tonic skin conductance is decreasing [58].

After that, onset and peak detection has been performed. To do this, peak onsets ( $> 0.01 \mu\text{S}$ ) and offsets ( $< 0 \mu\text{S}$ ) have been identified in the phasic data, and the maximum GSR value within each pair of onsets and offsets has been found in the unfiltered data. This processing step is illustrated in the top graph of Figure 3.4. The GSR peak amplitude is finally obtained by subtracting the amplitude at onset to the amplitude at the peak. All the peaks with a GSR peak amplitude smaller or equal to  $0.05 \mu\text{S}$  have been considered as artifacts [59]. One can observe the result of artifact detection in Figure 3.4.

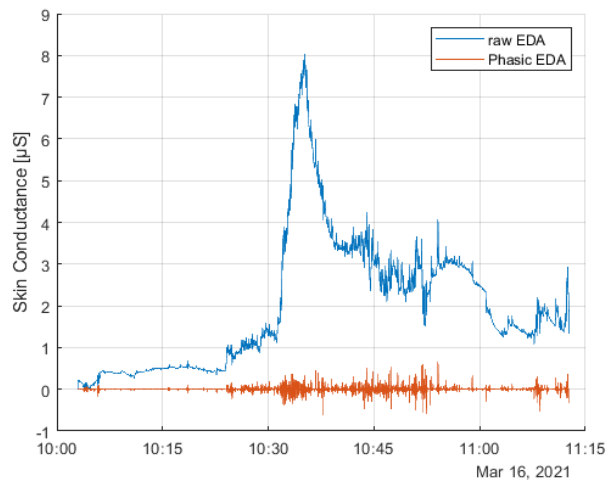


Figure 3.3: Example of raw and phasic EDA data.

### Gastric activity

As for ECG signals, EGG signals coming from *BIOPAC* have been first processed with *AcqKnowledge*. These have been downsampled from 2,000 Hz to 1.953 Hz in order to allow the correct use of automatic gastric wave analysis tool implemented in *AcqKnowledge*. This tool actually uses a time-frequency analysis to determine the percentage of gastric waves that fall within the frequency bands corresponding to normal, bradygastric, and tachygastric waves, as well as the percentage of waves falling outside of these boundaries and are arrhythmias. However, this analysis is designed to run on 1 Hz sampled data. 1.953 Hz being the closest (and the minimum) sample rate proposed by *AcqKnowledge*

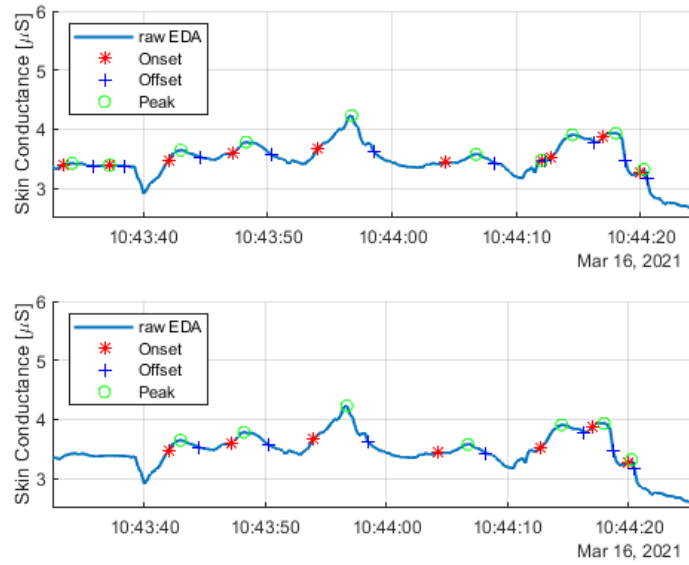


Figure 3.4: Peak detection on raw data without (top) and with (bottom) artifact detection.

when starting from a sample rate of 2,000 Hz, it has been chosen to be the sample rate of the downsampled signal. Actually, too large sample rates with respect to 1 Hz lead to arrhythmia classification of all the gastric waves composing the EGG. The downsampled signal has then been band-pass filtered, with low and high cutoff frequencies respectively equal to 0.00833 and 0.15 Hz. These values have been determined according to the gastric wave analysis tool. Indeed, for humans, the default frequency bands of gastric waves appear to be from 0.5 cycle per minute (CPM) to 9 CPM. More precisely, gastric waves ranging from 0.5 to 2 CPM are bradygastric waves, those ranging from 2 to 4 CPM are normal waves, and those ranging from 4 to 9 CPM are tachygastric waves. As a consequence, the resulting signal bandwidth was limited to the minimum necessary, increasing therefore the proportion of gastric waves being correctly classified by the gastric wave analysis tool. In order to easily compute the subject gastric activity in a particular time range (baseline, motion sickness or rest section), the gastric wave analysis tool has not been used, and the processing has been continued on *Matlab*.

The final goal being to extract the CPM mean value for different sections of the protocol, the downsampled and filtered EGG signal has been smoothed in order to ensure a proper detection of local maximums. By looking at the signal closely, it indeed appeared that the signal was stepped, thus leading to a distorted and excessive maximum detection. This can be explained as follows: when the file containing the signal of interest is exported from *AcqKnowledge* to *Matlab*, the highest sampling frequency of all the signals composing the file is considered. Since the signal of interest has the smallest sampling fre-

quency due to the downsampling, the same points are repeated until reaching the highest sampling frequency. Each smoothed value has been calculated based on data points over 4 seconds, and with the *moving* method consisting in applying a low-pass filter on the signal. An illustration of the stepped EGG signal and its corresponding smoothed EGG signal is presented in Figure 3.5.

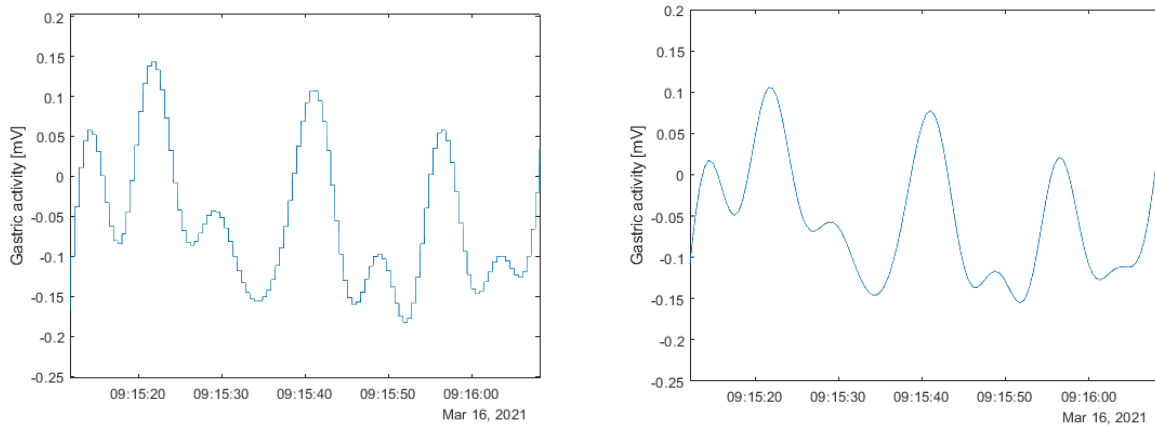


Figure 3.5: Part of a stepped EGG signal (left) and its corresponding smoothed signal (right).

All the relevant local maximums of the signal have finally been extracted by imposing a condition on the minimum separation between 2 maximums, as well as on their minimum prominence. In particular, a minimum of 5 seconds had to separate 2 detected maximums, and the minimum prominence has been fixed to 0.01 mV. An example of maximum detection on part of an EGG signal is illustrated in Figure 3.6.

An important note concerns the detection of artifacts. Under normal conditions, the EGG signal amplitude can go up to 400  $\mu\text{V}$  [60, 61]. Therefore, an artifact detection had been implemented in order to subtract all the maximums larger than 500  $\mu\text{V}$  from the total number of detected local maximums. However, it has been remarked that the number of detected artifacts in the motion sickness section was much larger than the one in the baseline and the rest sections, especially for car sessions. Actually, the majority of detected artifacts was due to the course trajectory of the car during the motion sickness section. Indeed, as mentioned previously, it contained a lot of sharp turns and braking, causing significant movements of the participant. As a consequence, real gastric activity of the participant was sometimes detected as artifacts, leading to a very small number of CPM. As a result, the detection of artifacts has not been considered for EGG signals, in order to detect gastric activity that is as close as possible to the real one. One knows

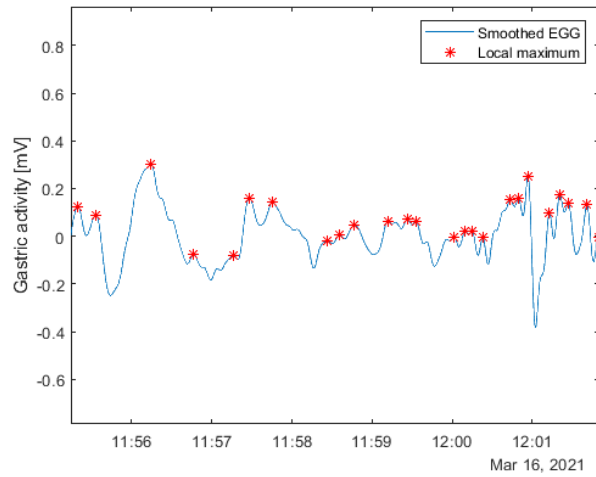


Figure 3.6: Example of local maximum detection on a smoothed EGG signal part.

that the EGG signal is very sensitive to motion and physical activity, it will therefore be necessary to be very careful in the conclusions drawn from the analysis of this signal.

## 3.2 Results and discussion

In the following, results of 2 subjects who suffered from motion sickness and 2 subjects who did not will be considered and analyzed, for both the simulator and the car sessions. In addition, global results considering all the subjects will also be presented and will allow conclusions to be drawn for the rest of this work.

For the simulator sessions, participants 7, 13, 16 and 17 will be considered. Participants 7 and 16 suffered from motion sickness during the simulator environment exposure. In particular, participant 7 assessed a maximum score of 9 and suffered from severe nausea and headaches; and participant 16 reported a maximum score of 10 and therefore stopped the experiment before the end of the session. He had severe nausea and hot flashes. As a consequence, no data was collected for the fourth part of the motion sickness section and the rest section for that participant, for the simulator session. The last MS score for that participant has been collected few minutes after the beginning of the fourth part of the motion sickness section and has so been attributed to the thirtieth minute of the motion sickness part. Participants 13 and 17 did not suffer from motion sickness during the simulator environment exposure. Indeed, the maximum MS score was equal to 0 and 1 for participant 13 and 17 respectively.

For the car sessions, participants 3, 5, 8 and 18 will be considered. Participants 3 and 8 suffered from motion sickness during the car session. In particular, participant 3 assessed a maximum MS score of 9 while it is equal to 8 for participant 8. They actually stopped executing the tasks and looked at the road when reporting this score, from minutes 28 and 25 of motion sickness section for participants 3 and 8, respectively. Indeed, they suffered from severe motion sickness symptoms and therefore wanted to stop the task execution before the normally scheduled end so as not to vomit. Participants 5 and 18 did not suffer from motion sickness during the car session, they reported very low MS scores during all the car session.

Some additional remarks about the car sessions need to be specified. First of all, no data was collected for the last part of the rest section for participant 18 because a battery problem causing the MP160 data acquisition unit to shut down was encountered. A similar problem occurred for 2 other participants. Moreover, it appeared several times that some of the participants looked briefly at the road during the session in order to reduce their symptoms of motion sickness. These one-off events have not been noted, but we will keep them in mind for the rest of this report.

### 3.2.1 Cardiac activity

Cardiac variables assumed to indicate motion sickness are the mean values of HR, as well as the mean values of pNN20, pNN50, RMSSD and SDRR during a precise amount of time. When simulator sessions are considered, these variables have been computed over each of the 6 different sections composing the scenario. This method has been preferred rather than considering fixed amount of time because all the 20 participants did not drive at the same speed in the simulator, therefore causing a shift in the observed scenario at a fixed time. On the contrary, for car sessions, cardiac variables have been computed over 5-minute periods except for the baseline section, for which only one value has been computed. This constitutes an appropriate method since all the participants were subjected to the same course at the same time of their session, all the drivers have indeed been trained to drive in a similar way. This method has been chosen in order to observe in a more accurate way the time evolution of these variables and the relationship existing between these variables and the MS score that has been collected every 5 minutes from the beginning of the motion sickness section. Other variables, later presented in this report, have been computed in the same way.

In view of the great variability of cardiac parameters existing between the different subjects, they were normalized as follows before computing the mean over a time period:

$$HR_{norm}(t) = \frac{HR(t) - p_{10}(HR)}{p_{10}(HR)} \cdot 100$$

$$HRV_{norm}(t) = \frac{HRV(t) - p_{90}(HRV)}{p_{90}(HRV)} \cdot 100$$

with  $p_{10}(HR)$  and  $p_{90}(HRV)$  respectively equal to the percentile 10 and percentile 90 of all HR values and considered HRV parameter values collected during a session. Indeed, HR tends to increase during motion sickness while HRV parameters tend to decrease during this state.

### Simulator sessions

Means of normalized HR values over the 6 different sections composing the driving simulator scenario as well as collected MS scores are presented in Figure 3.7 and 3.8 for 4 participants. Figure 3.7 concerns participants 7 and 16, who suffered from motion sickness during the simulator environment exposure, while Figure 3.8 concerns participants 13 and 17, who did not suffer from motion sickness during the simulator environment exposure.

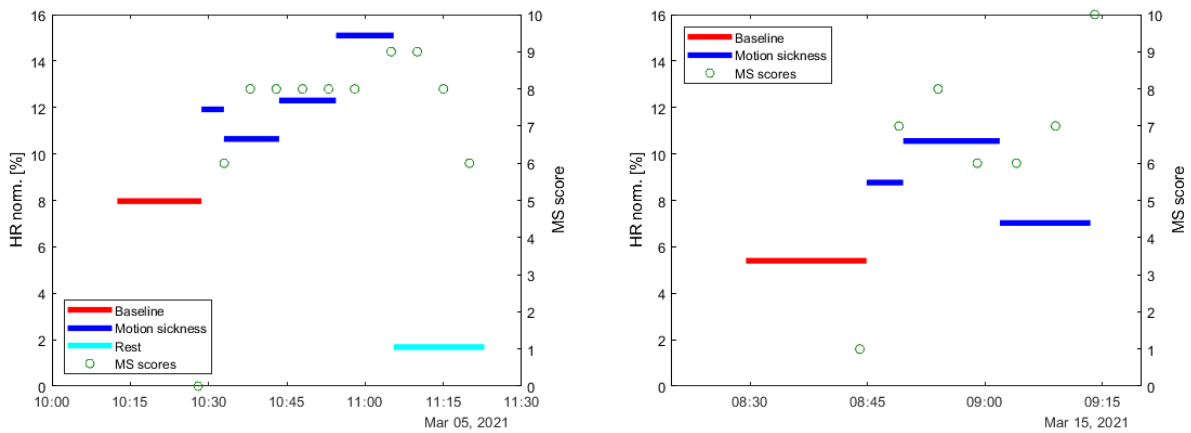


Figure 3.7: Means of normalized HR values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 7 (left) and 16 (right), who suffered from motion sickness during simulator environment exposure.

One can first observe that HR increases during motion sickness section compared to the baseline and the rest sections for all the 4 participants. However, the augmentation of HR seems more pronounced when participants suffered from motion sickness, since the percentage difference between the baseline or the rest HR value and the motion sickness one is larger for participant 7 than for participants 13 and 17. Concerning the participant



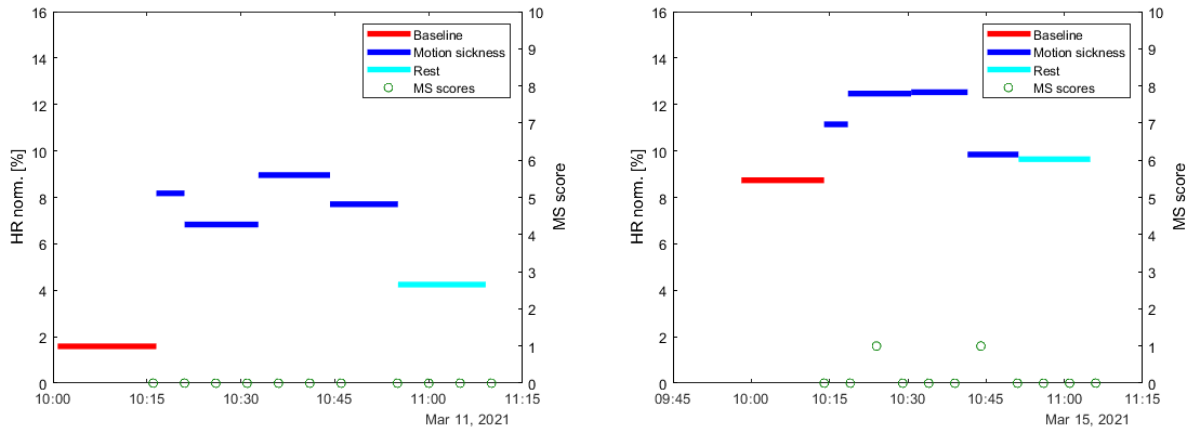


Figure 3.8: Means of normalized HR values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 13 (left) and 17 (right), who did not suffer from motion sickness during simulator environment exposure.

16, who, as a reminder, had to stop precociously the experiment because of severe symptoms, one can suppose that the average of normalized HR values during the rest section would be similar or even smaller than the one of the baseline section.

HR value under normal conditions (i.e., during the baseline and the rest sections) can be considered as the minimum of baseline and rest HR values. Indeed, depending on each participant level of stress, HR value is minimum either during the baseline section, either during the rest section. For most participants, the minimum HR value is the one of the baseline section since the one of the rest section is influenced by the motion sickness section they just went through, the normal cardiac rhythm taking some time to be recovered. On the contrary, the minimum HR value of several participants is the one of the rest section, probably because, during the baseline section, they were stressed by the acquisition session, since they did not know what to expect.

Means of normalized HR values in function of the session period and the MS score are presented in Figure 3.9, for all participants. MS scores collected after 5 minutes from the beginning of the motion sickness section were attributed to belong in the first part composing the motion sickness scenario section (MS Part1), MS scores collected after 10 and 15 minutes from the beginning of the motion sickness section were attributed to belong to the second part (MS Part2), the ones collected after 20 and 25 minutes were attributed to belong to the third part (MS Part3), and the ones collected after 30 minutes and at the end of the motion sickness section were attributed to belong to the last part composing the motion sickness scenario section (MS Part4). Indeed, on one hand, the experienced scenario at a fixed time (e.g. when asking a participant to evaluate his/her

motion sickness level) was not the same for all the participants given the different speed at which participants were driving, leading to the necessity of verifying that MS scores collected at a given time interval from the beginning of motion sickness scenario section belonged to the same motion sickness section of the scenario. On the other hand, most of the participants were driving at a lower speed than the allowed one, thus increasing the duration of the motion sickness section of the scenario, and therefore the total time of the simulator session.

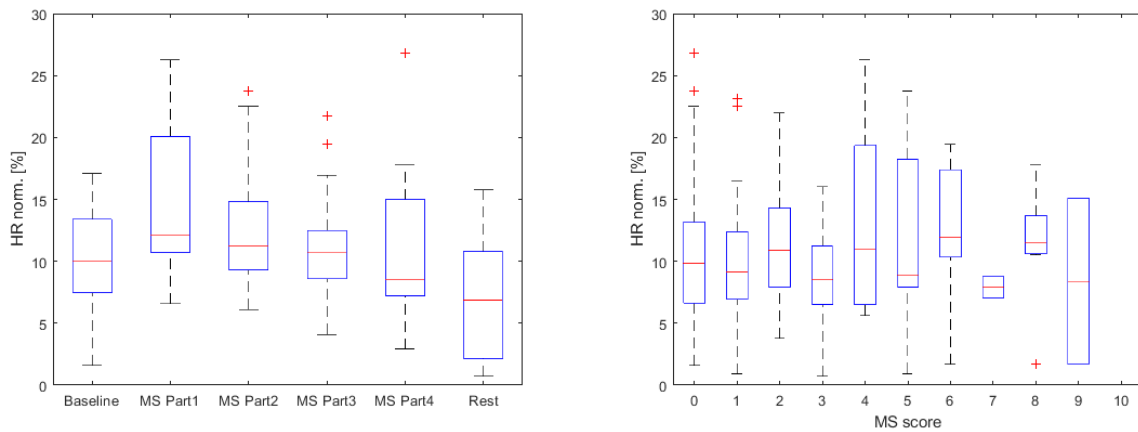


Figure 3.9: Means of normalized HR values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants.

One can observe from the left part of Figure 3.9 that means of normalized HR values are larger in the motion sickness section of the scenario than those of the baseline and the rest sections. It is particularly true for the 3 first motion sickness parts composing the motion sickness section. However, one cannot deduce an increase in HR due to motion sickness, because most of the participants did not suffer from motion sickness during the simulator sessions, but presented increased HR values during motion sickness section. This is confirmed by the right part of that figure, from which an increase in HR with the MS score cannot be concluded.

Means of normalized HRV parameter values over the 6 different sections of the driving simulator scenario and the collected MS scores are presented in Figure 3.10 for participants 7 and 16, and in Figure 3.11 for participants 13 and 17. In the same way as for HR, one can assume that each HRV parameter value under normal conditions is the maximum of baseline and rest HRV parameter values, since high HRV reflects increased parasympathetic activity, i.e., leading the body to a calm and composed state.

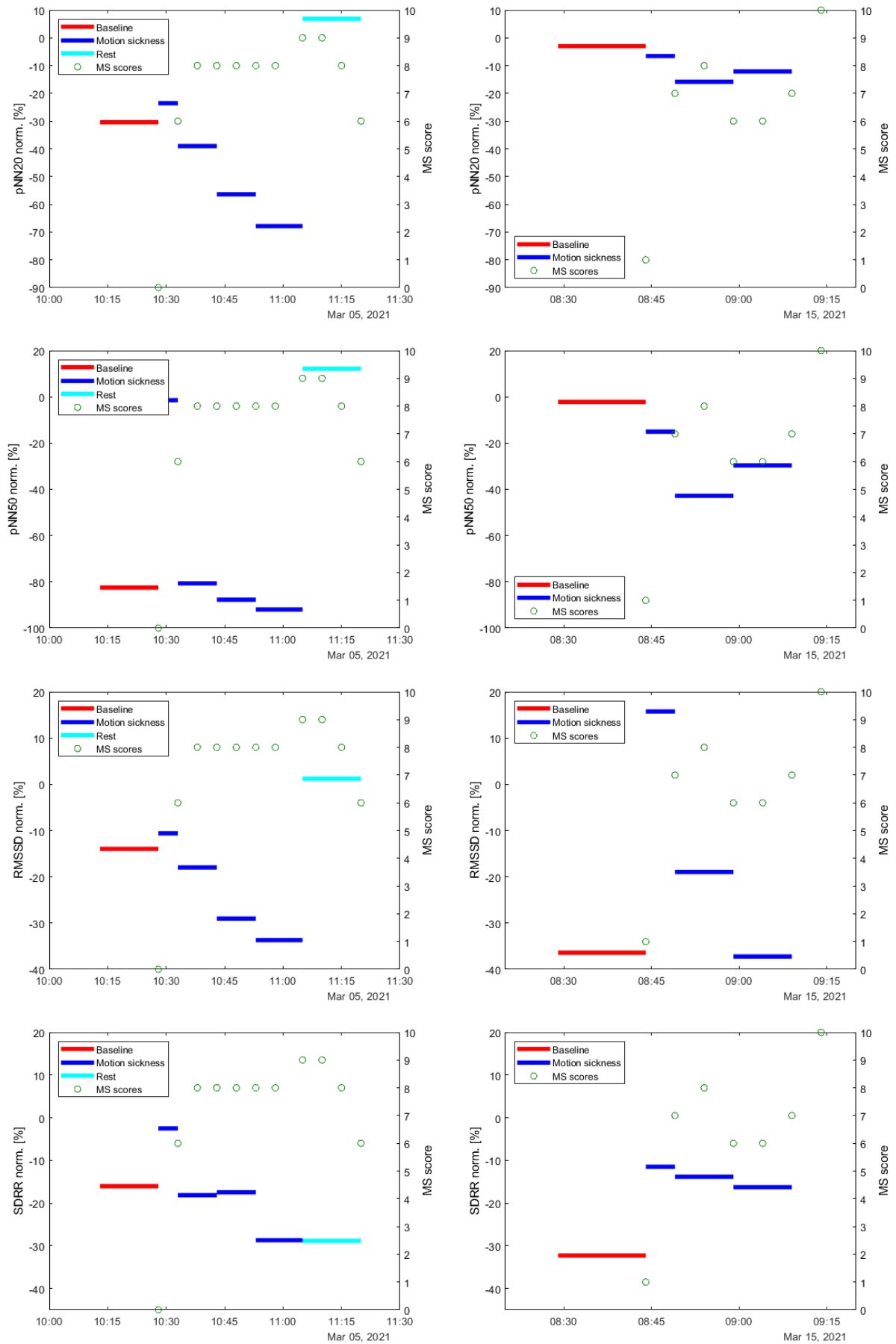


Figure 3.10: Means of normalized pNN20, pNN50, RMSSD and SDRR values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 7 (left) and 16 (right), who suffered from motion sickness during simulator environment exposure.

From Figure 3.10, it can be firstly observed that means of normalized pNN20 and pNN50 values of the motion sickness section are smaller than those of either the baseline section or the rest one. One usually sees that these values are smaller as the MS score is higher. Concerning means of normalized RMSSD, a similar conclusion can be drawn. Indeed, although 2 out of the 3 mean values of the motion sickness section are larger than the one of the baseline section for the participant 16, one can remark that mean value is decreasing as the intensity of motion sickness symptoms increased, leading finally to the need to stop the session. This suggests that the rest section mean value would have been larger than the ones of the motion sickness section. Finally, one also sees that the mean of normalized SDRR values is decreasing as the MS score increases, reflecting an increased sympathetic activity surely due to motion sickness.

The results presented in Figure 3.11, that concern 2 participants who did not suffer from motion sickness during the simulator session, seem to be similar to the ones of Figure 3.10, that concern 2 participants who suffered from motion sickness during the same session. Indeed, means of normalized HRV parameter values during motion sickness section are most of the time smaller than the ones either of the baseline section or of the rest section. However, it must be observed that these parameter values are either increasing or decreasing (except for pNN50 which has null values for participant 13, meaning that no successive R-R intervals differ by more than 50 ms) depending on the participant considered, while this was not the case when considering participants who suffered from motion sickness. As a consequence, one cannot infer a decrease in HRV parameter values due to the section of the scenario only from the results of these 2 participants, suggesting that the latter might decrease when the intensity of motion sickness symptoms increases.

Means of normalized pNN20, pNN50, RMSSD and SDRR parameter values in function of the session period and the MS score are presented in Figure 3.12, 3.13, 3.14 and 3.15, for all participants. By looking at the mean values of normalized HRV parameters in function of the period of the scenario, one can remark that these are usually smaller in the motion sickness section than in the baseline and the rest sections, except for the RMSSD, for which the baseline value is the smallest one. However, the decrease in HRV parameter values as the MS score increases is not so easily observed, meaning that an intensity increase of motion sickness symptoms is not highly correlated with a decrease in HRV parameter values.

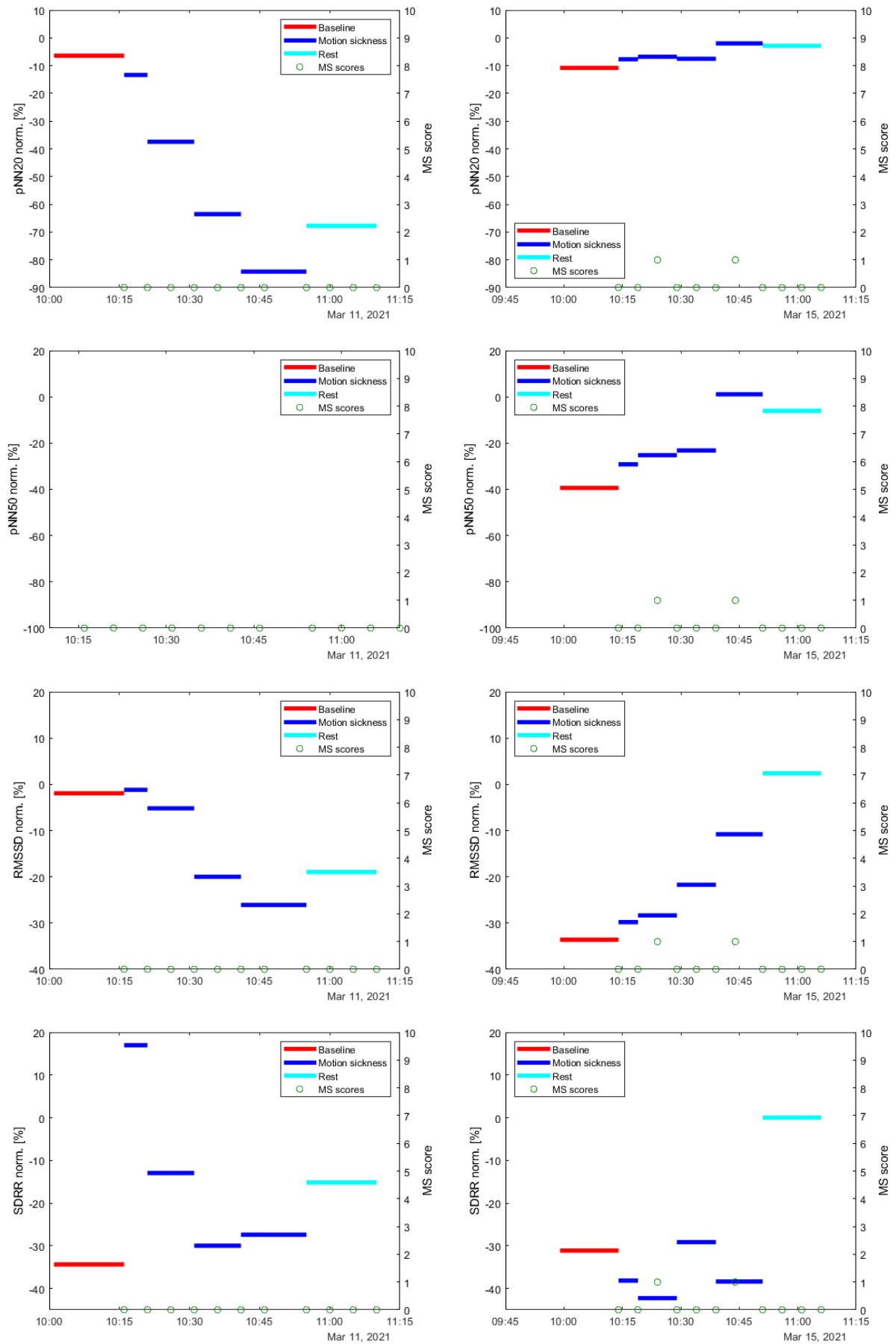


Figure 3.11: Means of normalized pNN20, pNN50, RMSSD and SDRR values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 13 (left) and 17 (right), who did not suffer from motion sickness during simulator environment exposure.

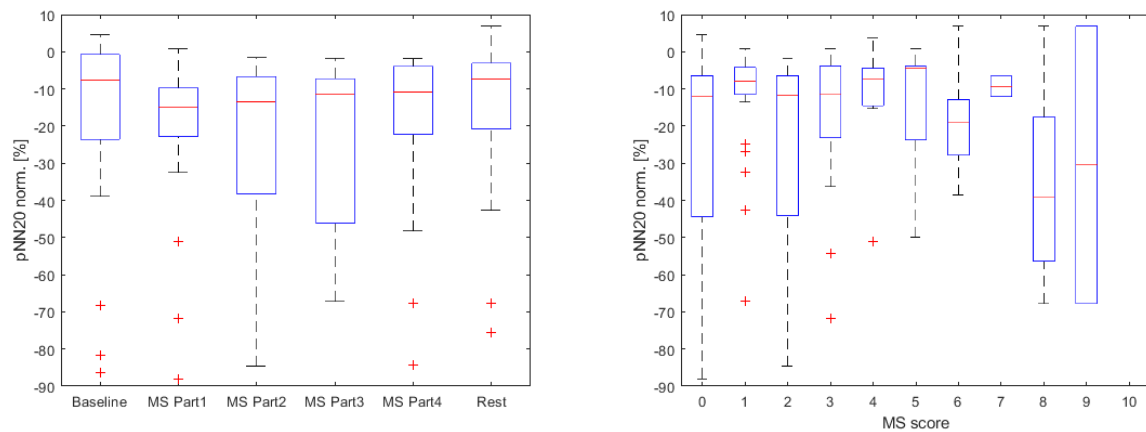


Figure 3.12: Means of normalized pNN20 values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants.

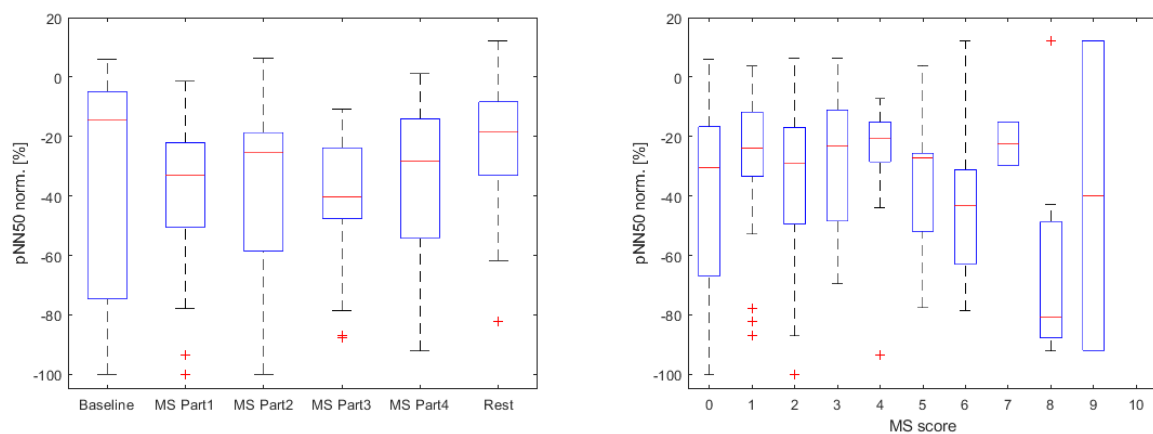


Figure 3.13: Means of normalized pNN50 values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants.

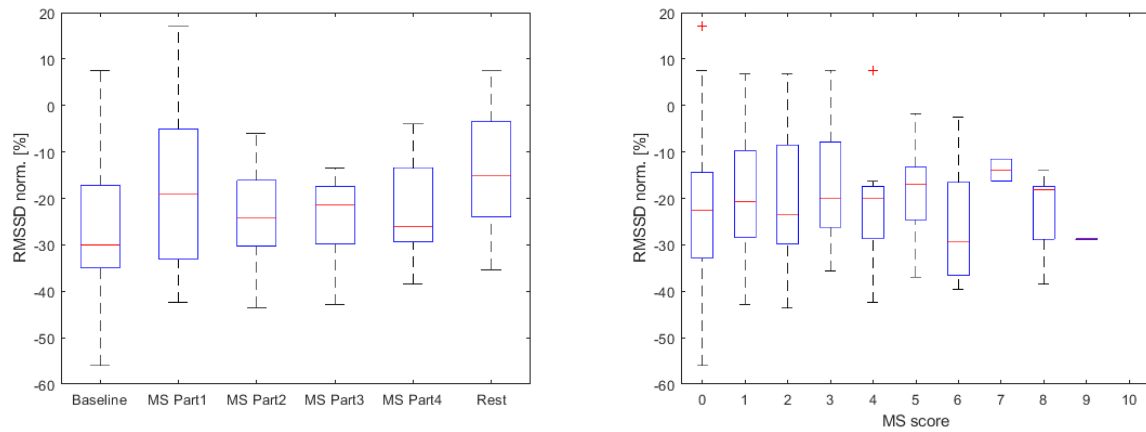


Figure 3.14: Means of normalized RMSSD values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants.

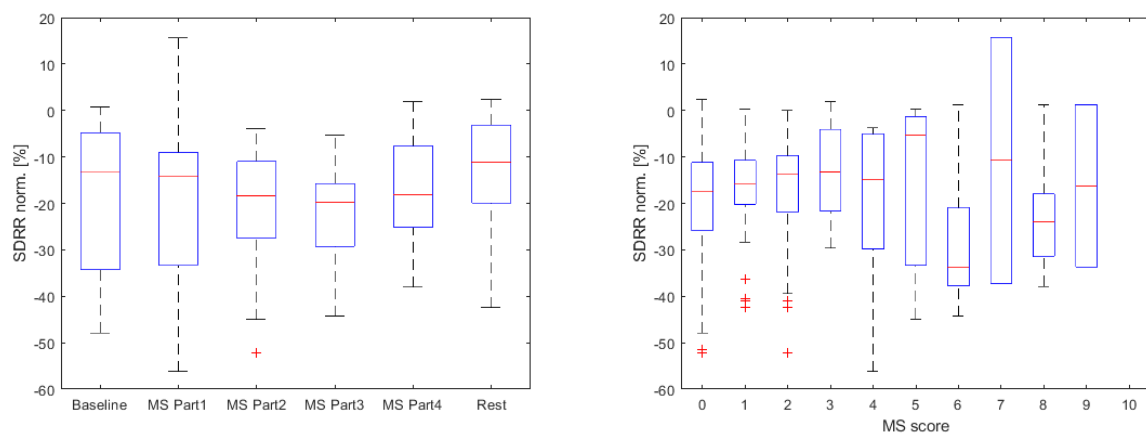


Figure 3.15: Means of normalized SDRR values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants.

It is important to mention that the number of values included in a boxplot for a particular MS score can be very different from one MS score to another. Indeed, as mentioned previously, simulator sessions have only caused a few onsets of motion sickness compared to the car sessions. As a consequence, the number of physiological measures associated with a low MS score is much larger than the one associated with a high MS score. Therefore, data values of one specific participant can have an important influence on the location of the boxplot for high MS scores, while this influence is much smaller when considering low MS scores.

### Car sessions

Means of normalized HR values over, either 5-minute periods or over the 15 minutes composing the baseline section, as well as collected MS scores, are presented in Figure 3.16 and 3.17. Figure 3.16 concerns participants 3 and 8, who suffered from motion sickness during the car session, while Figure 3.17 concerns participants 5 and 18, who did not suffer from motion sickness during the car session.

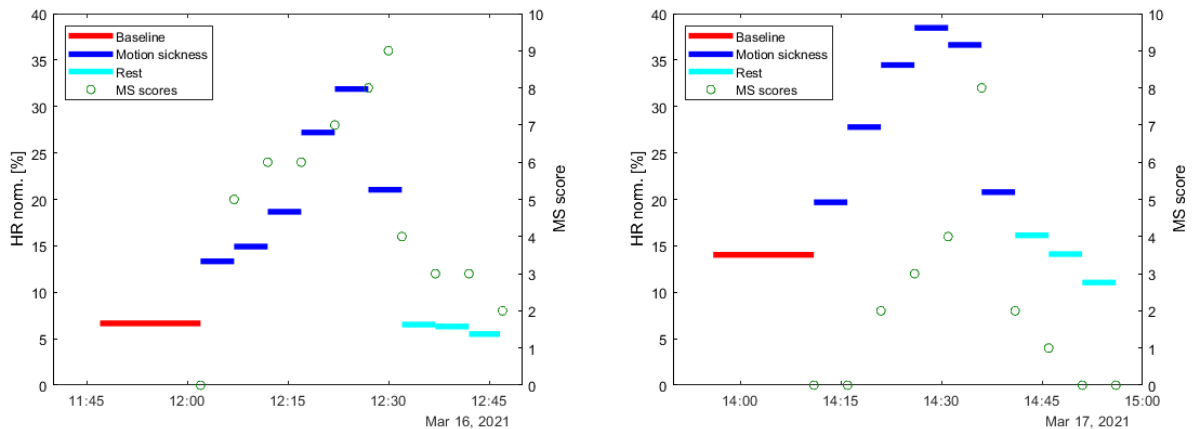


Figure 3.16: Means of normalized HR values over fixed time periods and collected MS scores for participant 3 (left) and 8 (right), who suffered from motion sickness during the car session.

As for the simulator sessions, one can observe an increase in means of normalized HR values during the motion sickness section for all the 4 participants. However, the increase is much more pronounced for participants who suffered from motion sickness than for participants who did not. Also, while mean of normalized HR values during the motion sickness section is randomly varying for participants 5 and 18, a continuous increase in HR value is observed for participants 3 and 8, as the MS score increases, except during the last part of this motion sickness section. The sudden decrease can be explained by



the stopping of the task execution and the related reduction in the severity of motion sickness symptoms. Then, by comparing the augmentation of HR values for sick participants during simulator and car sessions, one observe that it is much more significant for car sessions than for simulator ones.

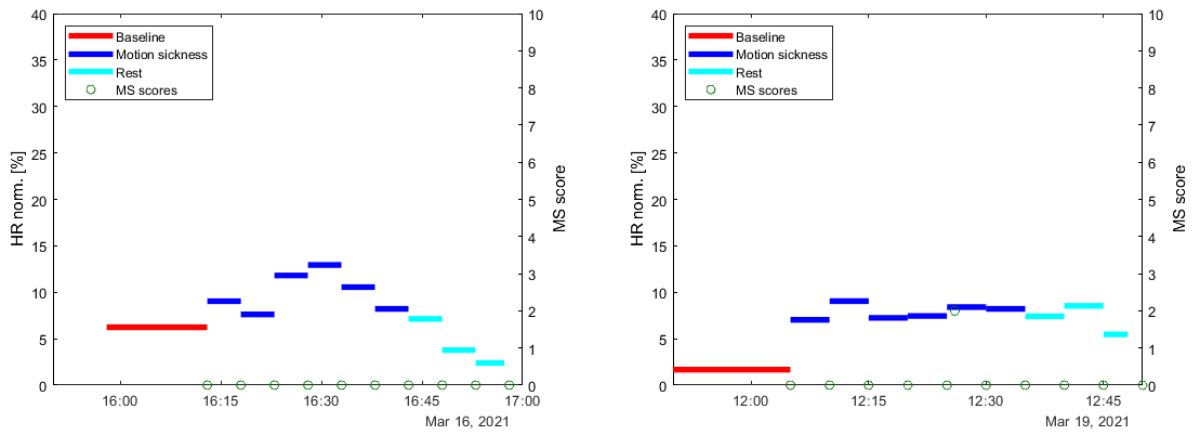


Figure 3.17: Means of normalized HR values over fixed time periods and collected MS scores for participant 5 (left) and 18 (right), who did not suffer from motion sickness during the car session.

Means of normalized HR values in function of the session period and the MS score are presented in Figure 3.18, for all participants. MSx and Restx refer to motion sickness and rest part respectively spanning from minute 5x-5 to minute 5x of the motion sickness and the rest sections. For instance, MS2 refers to the motion sickness part spanning the fifth to the tenth minute of the motion sickness section, while Rest3 refers to the rest part spanning the tenth to the fifteenth minute of the rest section.

From the left part of Figure 3.18, it can be observed that means of normalized HR values are larger in the motion sickness section than in the baseline and the rest sections. Moreover, this trend is much more pronounced for car sessions than for simulator ones. As previously, an increase in HR due to the motion sickness cannot be deduced only from that graph, because not all the participants suffered from motion sickness during the motion sickness section. However, the number of participants that suffered from that state is much larger for car sessions than for simulator ones. As a consequence, it is more likely that HR increases as severity motion sickness symptoms increases. This correlation is actually well observable in the right part of Figure 3.18, since means of normalized HR values are higher as the MS score is.

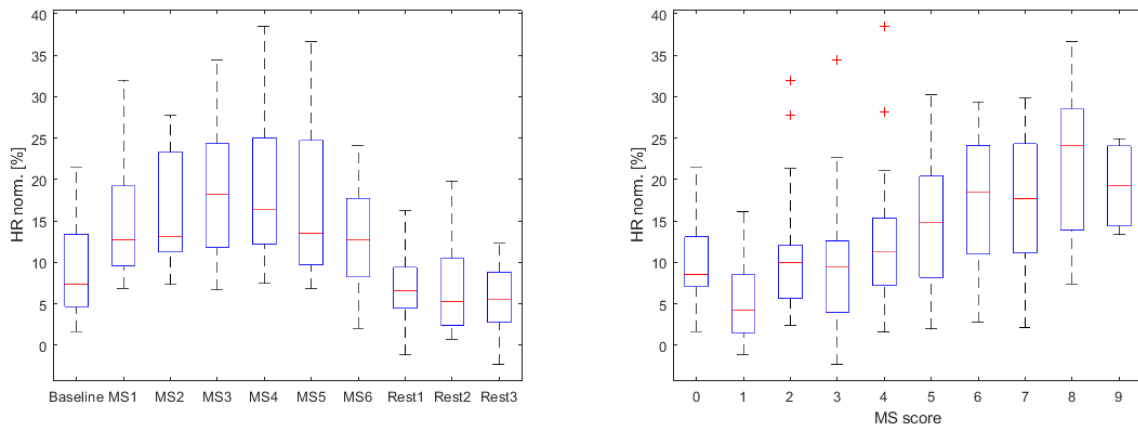


Figure 3.18: Means of normalized HR values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants.

One can also remark that means of normalized HR values are usually higher at a MS score of 0 than at MS scores of 1 or 2. As previously, this can be explained by the presence of stress at the beginning of the session, the MS scores of 0 having been generally collected during the baseline section.

Means of normalized HRV parameter values over either 5-minute periods or over the 15 minutes composing the baseline section, as well as the collected MS scores, are presented in Figure 3.19 for participants 3 and 8, and in Figure 3.20 for participants 5 and 18.

One can observe in Figure 3.19 that means of normalized pNN20 values generally decrease as the MS score increases, and then increase when the MS score decreases. Although a trend is present, the same observation cannot be made when considering pNN50, RMSDD and SDRR parameters, since mean of normalized values sometimes increases when the MS score increases. Actually, this curious increase could be explained by the fact that participants took a look at the road during the session in order to decrease the severity of their symptoms, although their HR has not been observed to decrease at that time.

Again, the sudden increase in means of normalized HRV parameter values for the last part of the motion sickness section can be explained by a diminished intensity of motion sickness symptoms due to the stopping of the task execution and attention to the road. Finally, means of normalized HRV parameter values usually increase as the rest section progresses, showing a progressive return towards the baseline value.

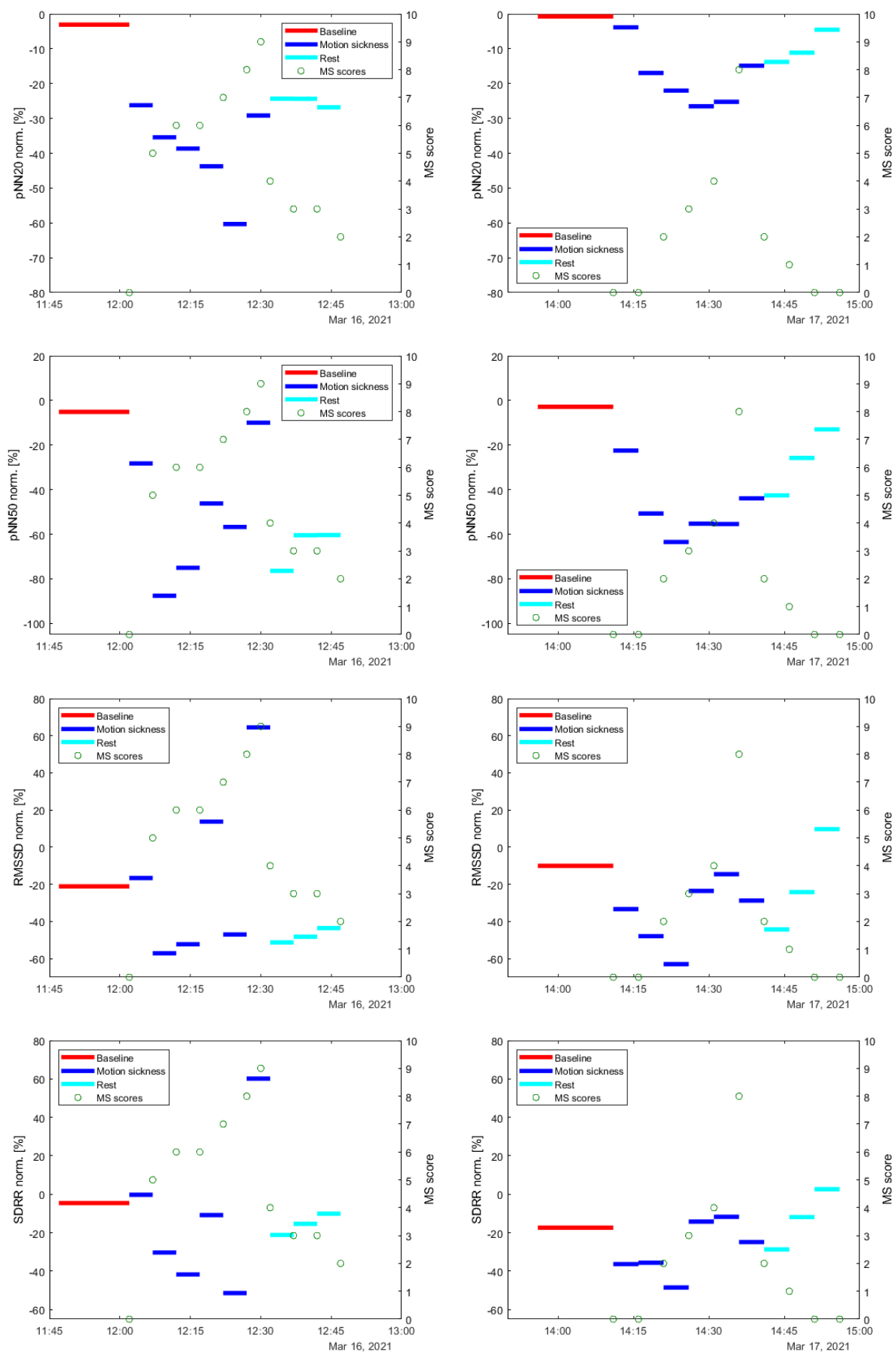


Figure 3.19: Means of normalized pNN20, pNN50, RMSSD and SDRR values over fixed time periods and collected MS scores for participant 3 (left) and 8 (right), who suffered from motion sickness during the car session.

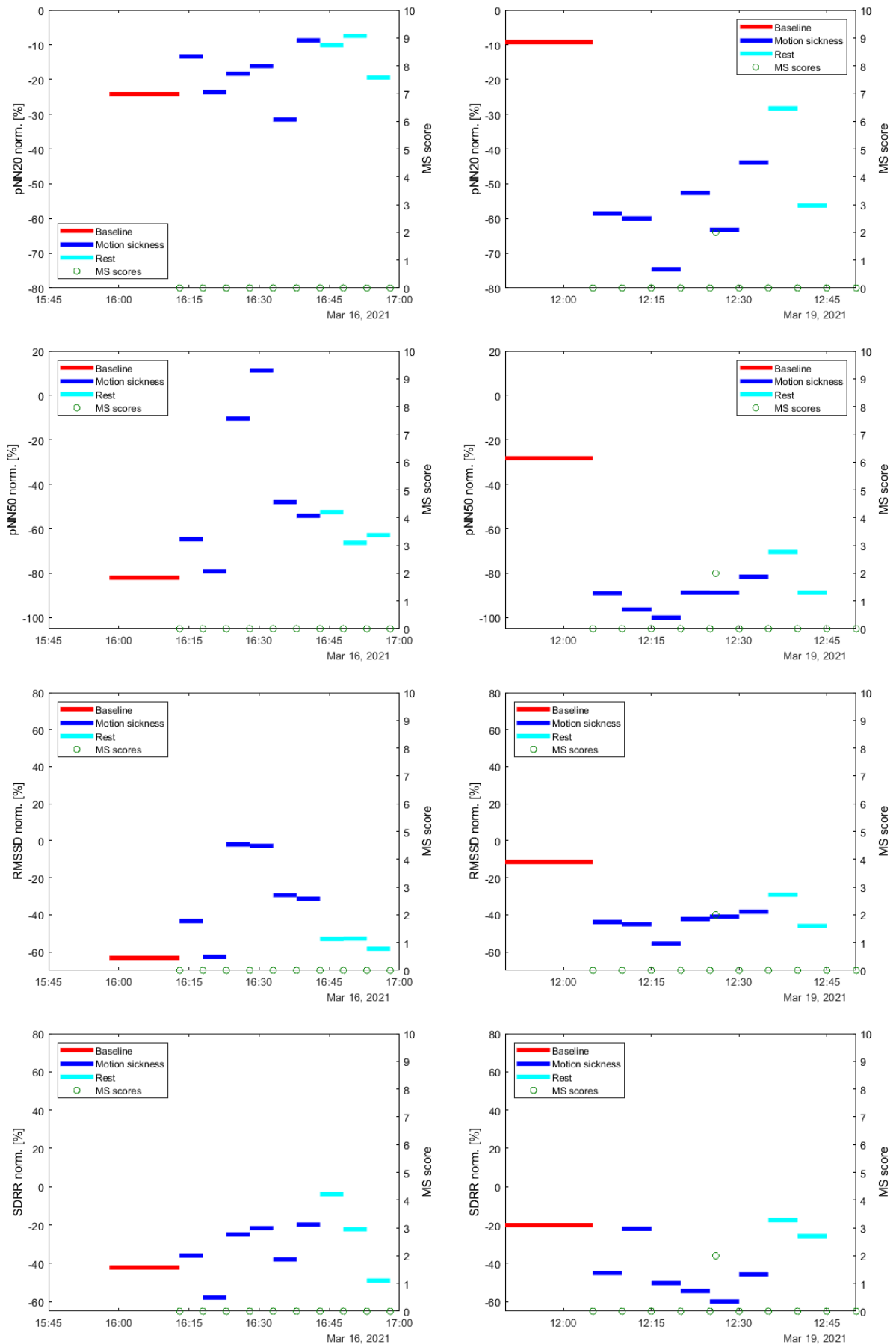


Figure 3.20: Means of normalized pNN20, pNN50, RMSSD and SDRR values over fixed time periods and collected MS scores for participant 5 (left) and 18 (right), who did not suffer from motion sickness during the car session.

The results presented in Figure 3.20 might suggest again that HRV parameter values decrease as a result of motion sickness. Indeed, one can remark in this Figure 3.20 that the means of normalized HRV parameter values evolve in different directions depending on the participant that is considered. For instance, considering the motion sickness section, the means of normalized pNN50 values first increase for participant 5, and decrease in a second time, while this is the opposite phenomenon that is observed for the participant 18.

Means of normalized pNN20, pNN50, RMSSD and SDRR parameter values in function of the session period and the MS score are presented in Figure 3.21, 3.22, 3.23 and 3.24, for all participants. On all of these figures, one first sees the presence of very large positive outliers, reducing the space allocated to the boxplots. These outliers are due to the way the normalization of HRV parameters is performed. Indeed, by considering the percentile 90 instead of the maximum value, normalized HRV values of values larger than the percentile 90 can reach very large value.

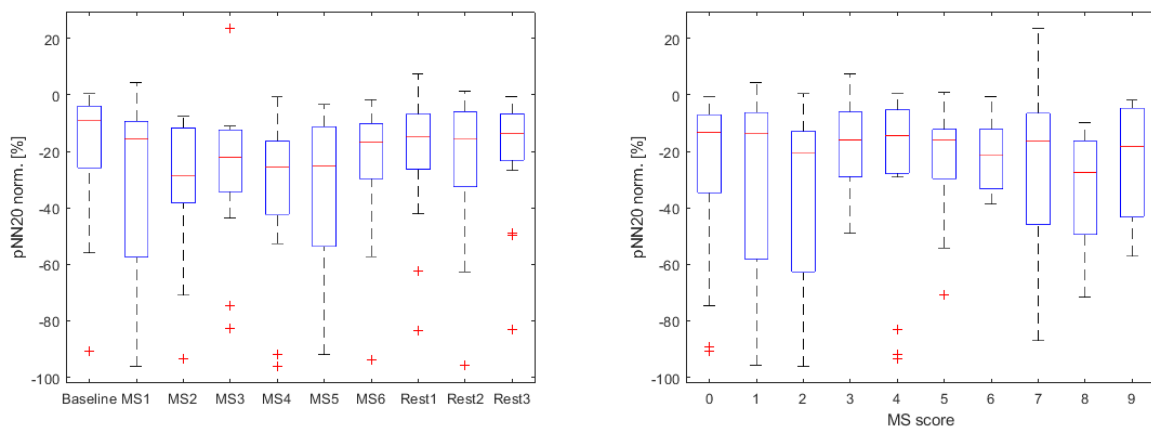


Figure 3.21: Means of normalized pNN20 values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants.

Considering now only the boxplots, one remarks that means of normalized pNN20 parameter values are usually smaller during the motion sickness section than during the 2 others. However, although a trend is observable, one cannot conclude that means of normalized pNN20 parameter values decrease as the MS score increases, since there is sometimes an increase of the percentile 50 of these values between 2 adjacent MS scores.

By looking at the other HRV parameters, one cannot definitely conclude in a decrease in HRV with a increased severity of motion sickness symptoms. Indeed, considering

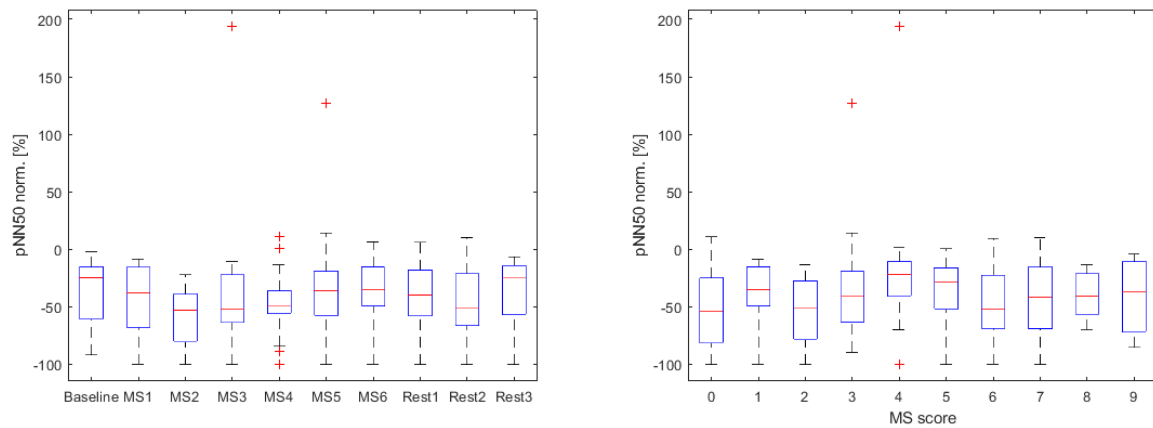


Figure 3.22: Means of normalized pNN50 values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants.

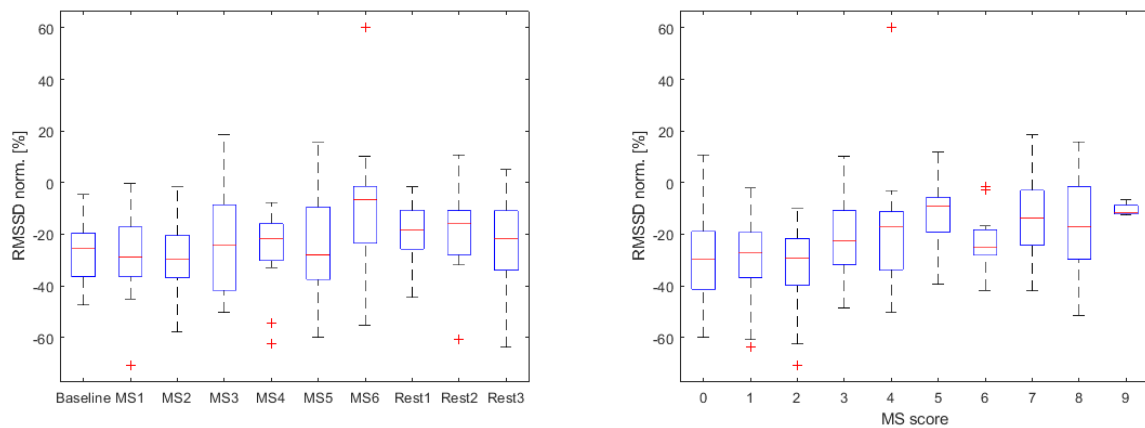


Figure 3.23: Means of normalized RMSSD values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants.

firstly the pNN50 parameter, one can see that the normalized pNN50 parameter values of high MS scores are similar to the ones of low MS scores. It is also the case for the SDRR parameter, while the normalized RMSSD values seem to increase as the MS score increases.

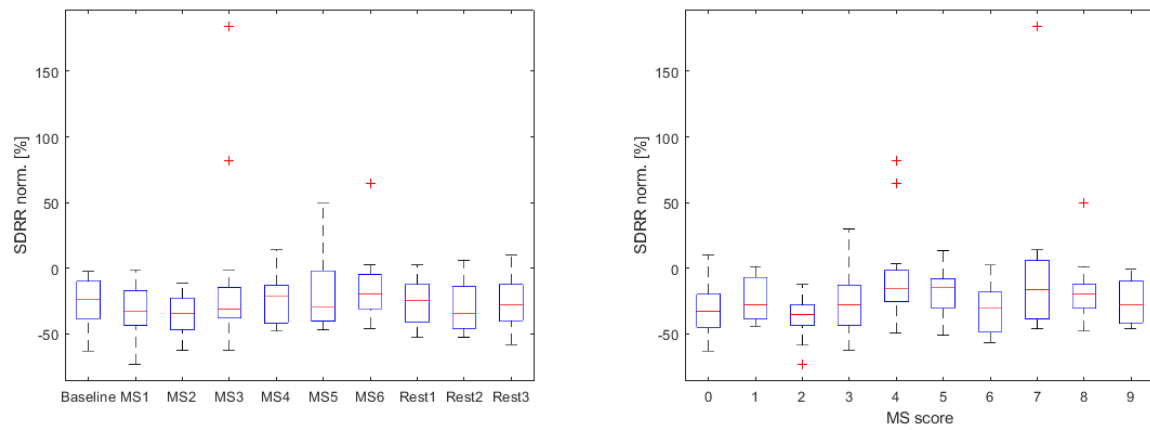


Figure 3.24: Means of normalized SDRR values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants.

### 3.2.2 Electrodermal activity

In this report, the mean value of numbers of EDA peaks per minute over the same periods as the ones considered for cardiac activity is considered as the variable indicating motion sickness. Again, the same 8 participants will be considered for showing detailed results, and global results will also be presented.

Approximately 10 % of people are estimated to be non-responders (hypo-responsive) in terms of their EDA [62]. In this experiment, 3 out 20 participants have been identified to be non-responders. They indeed presented very weak EDA signals during the 2 sessions of data acquisition, with both *BIOPAC* and *E4* recording units. EDA signals of one non-responder are presented in Appendix B.2, Figure B.3, for both simulator and car sessions<sup>1</sup>. Moreover, several participants, who were not identified to be non-responders, have presented an EDA signal of bad quality either during the simulator session or the car session, surely due to a poor contact between the skin and the electrodes. As a consequence, these signals have not been included in global results, these last therefore consider signals of EDA responders that are of good quality. Signals from respectively 12 and 10 participants are so considered, for the simulator sessions and the car ones.

As both EDA signals coming from *BIOPAC* and *E4* have been considered, the way the results are presented is a little different from the one for cardiac activity. Indeed, for each of the 8 participants, means of numbers of peaks derived from both *BIOPAC* and *E4* signals will be presented in a same graph. Also, when considering simulator sessions, the median of MS scores collected over the considered period of time has been computed and is displayed such as the mean number of peaks can be related to only one MS score<sup>2</sup>. Obviously, this operation is not needed for car sessions since there are as many periods as MS scores collected. However, for car sessions, only the MS score at regular intervals is displayed for sake of simplicity. It indeed happened that the severity of the symptoms increased or decreased suddenly during an interval of 5 minutes, the participant then reported it and the score at that time was considered for cardiac activity and will also be for gastric activity, but not here. It actually does not modify the conclusions that will be drawn. Finally, the scales of the y-axis were not constrained to be identical for the different participants, in order to observe the evolution of the electrodermal activity with the MS score in a more obvious way.

<sup>1</sup>Note that the EDA signal recorded with *BIOPAC* during the car session is completely noisy (due to poor contact between the skin and an electrode), it does not allow to conclude that the participant is a non-responder.

<sup>2</sup>The median of a vector containing only 2 values is equivalent to the mean of these 2 values, therefore leading to non-integer numbers.



### Simulator sessions

Means of numbers of peaks per minute over the 6 different sections of the simulator session and collected MS scores are presented in Figures 3.25 and 3.26, for participants 7 and 16, and 13 and 17 respectively.

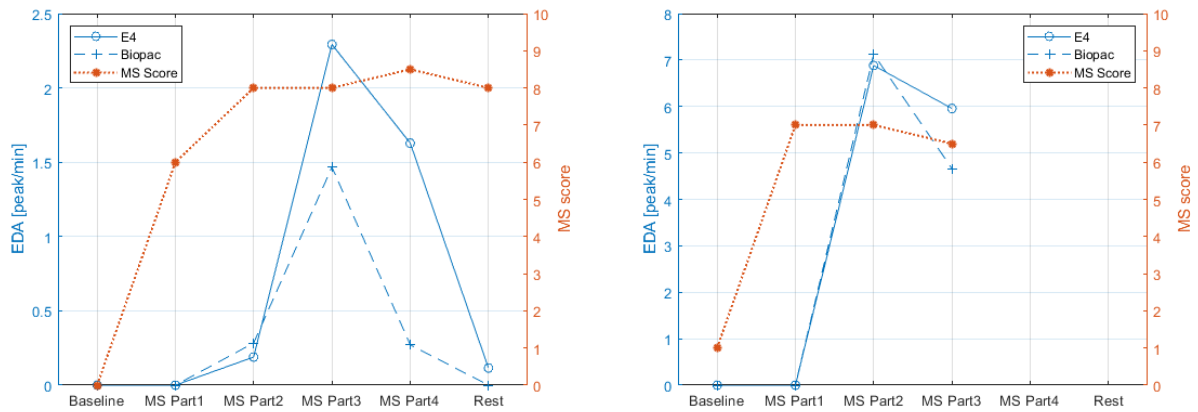


Figure 3.25: Means of numbers of peaks per minute over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 7 (left) and 16 (right), who suffered from motion sickness during simulator environment exposure.

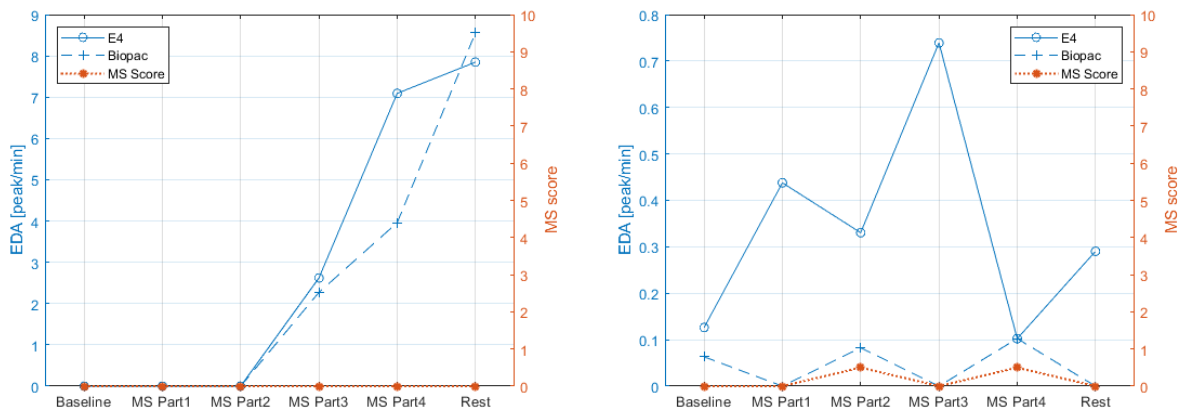


Figure 3.26: Means of numbers of peaks per minute over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 13 (left) and 17 (right), who did not suffer from motion sickness during simulator environment exposure.

From Figure 3.25, one can first remark that the number of peaks evolves in the same direction for both *BIOPAC* and *E4* signals. The same observation can be made in Figure 3.26 for the participant 13. On the contrary, for the participant 17, the number of peaks evolves in the opposite direction when considering the 2 recording units. By looking at

the peak value, one sees that these are very small, one cannot therefore conclude that the 2 recording units do not agree.

Means of numbers of peaks per minute for the 2 considered participants who suffered from motion sickness seem to be positively correlated with the MS score. One remarks however that means of the numbers of peaks per minute are smaller for the participant 7 than for the participant 16. Concerning participants who did not suffer from motion sickness, one sees that the number of peaks per minute considerably increases for the participant 13, while this number is nearly constant when the participant 17 is considered.

The global results are presented in Figure 3.27. These considered *E4* EDA signals since these last were usually of better quality than the ones coming from *BIOPAC*.

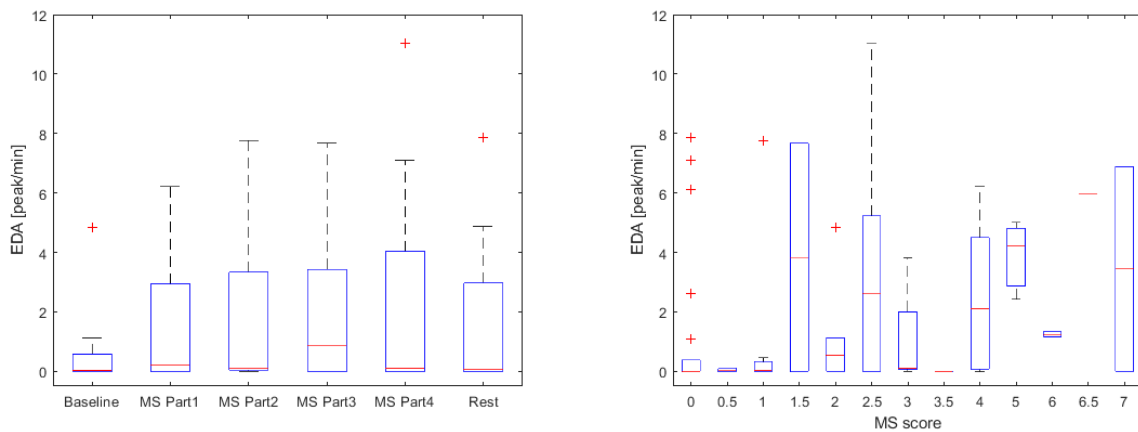


Figure 3.27: Means of numbers of peaks per minute over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all EDA respondent participants with good quality signals.

By looking at the left plot of that Figure 3.27, one sees that the means of numbers of peaks per minute are generally higher during the parts composing the motion sickness section of the scenario, although the percentile 50 values are very close from one section to the other. Again, no conclusion can be made by looking only at this graph since most of the participants did not suffer from motion sickness during the simulator sessions. One has to look at the right graph.

As a reminder, some MS scores are non-integer numbers since the median of collected MS scores over a period has been computed. Also, there is no MS scores higher than 7 because the results presented in Figure 3.27 are considering EDA respondent participants

with good quality signals, and either participants who reported a high MS score were not EDA respondent, or their EDA signals were of bad quality.

From the right part of this Figure 3.27, conclusion about the evolution of the number of peaks per minute with the intensity of motion sickness symptoms cannot be drawn. Although, for instance, the percentile 50 increases between the MS scores 4 and 5, and 6 and 6.5, it decreases between MS scores 5 and 6, and 6.5 and 7. The high values of means of the numbers of peaks per minute for low MS scores can be explained by the fact that participants who reported a low MS score may be affected by another condition, such as stress or high cognitive load. A less likely explanation is that some participants suffered from motion sickness without reporting high MS scores.

### Car sessions

Means of numbers of peaks per minute over the different sections of the car session and collected MS scores are presented in Figures 3.28 and 3.29, for participants 3 and 8, and 5 and 18 respectively.

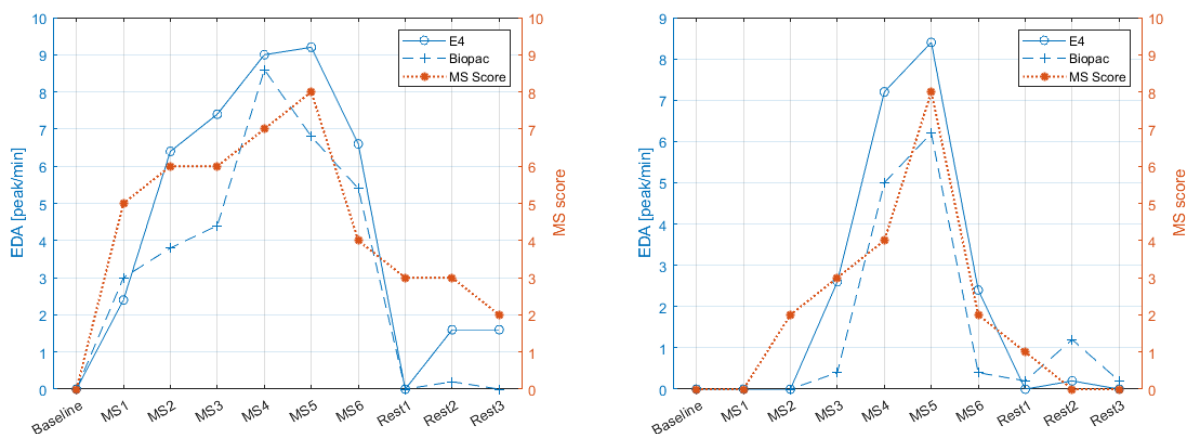


Figure 3.28: Means of numbers of peaks per minute over the different sections composing the car session and collected MS scores for participant 3 (left) and 8 (right), who suffered from motion sickness during the car session.

As for the simulator sessions, one first observes that the number of peaks evolves in the same direction for both *BIOPAC* and *E4* signals, except for participant 18 who is actually a non-responder. Then, one clearly remarks in Figure 3.28 an increase in the number of peaks as the MS score increases, thus for the 2 participants who suffered from motion sickness during the car session. In particular, the maximum number of peaks is reached when the MS score is the highest. The observed decrease in the mean of numbers of peaks in the last part of the motion sickness section can again be explained by the fact that

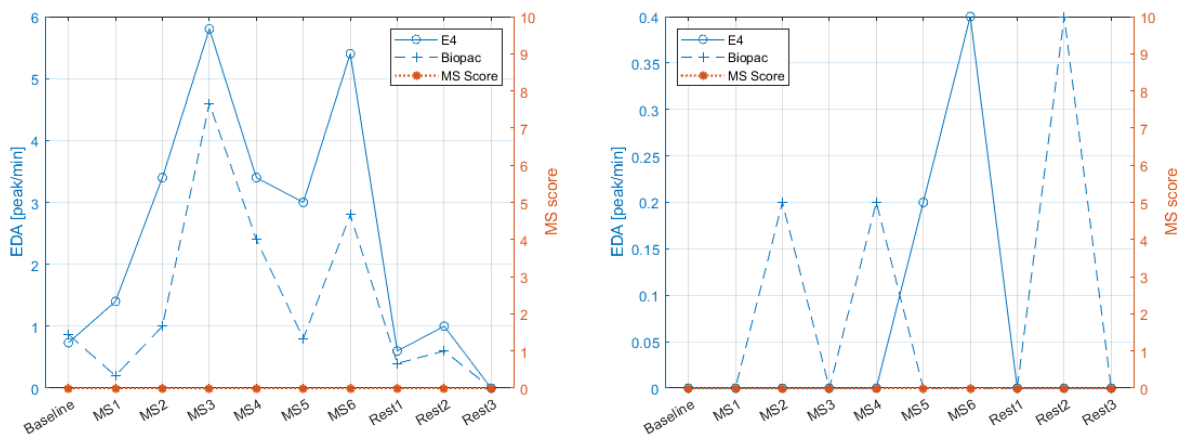


Figure 3.29: Means of numbers of peaks per minute over the different sections composing the car session and collected MS scores for participant 5 (left) and 18 (right), who did not suffer from motion sickness during the car session.

these 2 participants stopped the task execution due to severe symptoms of motion sickness.

Concerning participant 5, whose results are presented in Figure 3.29, one can observe that the number of peaks is generally higher during the motion sickness section than during the baseline and the rest sections, that could suggest that an elevation of the number of peaks per minute is due to the section rather than to the onset of motion sickness. However, global results have to be analyzed before drawing conclusion. These are presented in Figure 3.30. As for simulator sessions, these global results are derived from EDA signals recorded with *E4*. Indeed, these were usually of better quality than the one coming from *BIOPAC*, and recordings were never interrupted since these were not dependent of the MP160 unit.

From this Figure 3.30, one observes globally that the mean of numbers of peaks per minute increases during the motion sickness section, and as the MS score increases. The increase with the MS score seems to be done by plateau, and is particularly pronounced between the MS scores 5 and 6. This increase of the mean of numbers of peaks per minute with the MS score is true until considering the MS score equals to 9. At first glance, it may seem strange, but it can actually be easily explained. From all participants who reported a MS score of 9, only participants 3 and 10 presented a good quality EDA signal. Therefore, only these 2 are included in the boxplot of MS score 9. In particular, both participants reported a MS score of 9 only once, and it was at the end of the motion sickness section. At that time, they stopped the execution of the tasks because of too severe motion sickness symptoms, which potentially had the impact of drastically reducing the

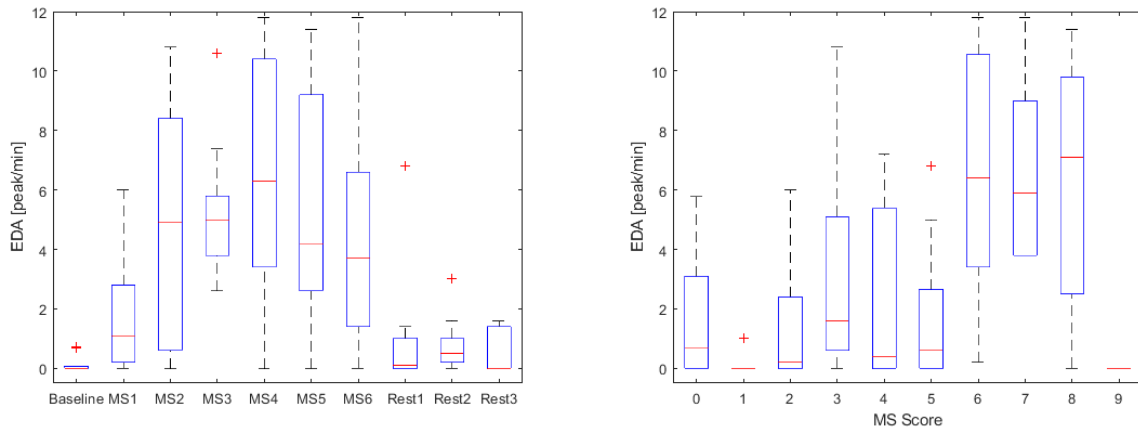


Figure 3.30: Means of numbers of peaks per minute over the different sections composing the car session in function of the session period (left) and the MS score (right), for all EDA respondent participants with good quality signals.

number of peaks in the last period of the motion sickness section.

From all the figures presented in this section 3.2.2, it appeared that means of numbers of peaks per minute derived from both *BIOPAC* and *E4* signals were evolving in the same way. However, their values were often different. As mentioned in chapter 2, section 2.3.1, the data is not collected at the same body location and that can lead to signal differences. In that case, *E4* acquisition unit provided EDA signals of better quality than those collected with *BIOPAC* surely because there are more sweat glands at the location the wristband was worn than at the scapula, where the surface electrodes were placed for recording EDA through *BIOPAC* system. Note that signals recorded with *BIOPAC* would be preferably considered in the case where the developed protocol asked participants to move their forearms for instance, in which case *E4* signals would suffer from a lot of movement artifacts.

### 3.2.3 Gastric activity

Gastric variable that has been considered for motion sickness detection is the mean value of CPM over the same periods as the ones considered for cardiac and electrodermal activities.

To record an EGG in the best possible way, it is strongly recommended by *BIOPAC* that the subject get in supine position and be completely still during the recording. However, the developed protocol did not allow to meet these recommendations. Indeed, during both simulator and car sessions, participants were in a sitting position, either they

were driving in the simulator or they were performing some tasks in the car, and they regularly spoke to assess their level of motion sickness. So, wanting to make sure that the obtained results were not affected by these conditions, I carried out EGG on me in 4 different conditions, without having eaten during the 3 hours preceding the acquisitions, as asked to the participants:

1. In supine position and completely still;
2. In supine position and speaking;
3. In sitting position and completely still;
4. In sitting position and speaking.

Each EGG has been performed during 12 minutes, and the resulting means of CPM are presented in Table 3.2, for each of the 4 experimental conditions.

Conditions	CPM
1	4.25
2	4.833
3	4.25
4	4.75

Table 3.2: Means of CPM values derived for the 4 different experimental conditions.

From this Table 3.2, it appears that the number of CPM is not affected by the position. However, this number is increased when I was speaking compared to when I was completely still. It appears obvious that general conclusions cannot be drawn from these results since these experiments were only carried out on one person given the sanitary conditions and newly imposed rules. However, one can keep in mind for the following that the derived means of CPM might be a little higher than those that would be obtained if the participants were in supine position and completely still during the recording. As a consequence, one can suppose that the frequency bands characterizing bradygastric (0.5-2 CPM), normal (2-4 CPM) and tachygastric (4-9 CPM) waves could be shifted up from a fraction of one CPM for this protocol.

As previously, results from participants 7, 13, 16, 17 of simulator sessions and 3, 5, 8, 18 of car sessions will be presented, as well as global results that consider all the participants.

### Simulator sessions

Means of CPM values over the 6 different sections composing the driving simulator scenario as well as collected MS scores are presented in Figures 3.31 and 3.32 for participants 7 and 16, and 13 and 17 respectively.

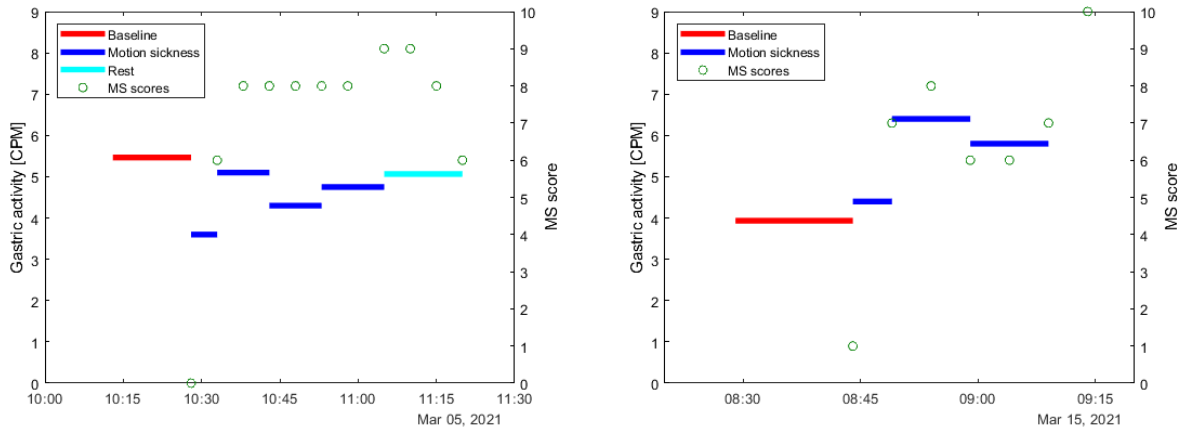


Figure 3.31: Means of CPM values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 7 (left) and 16 (right), who suffered from motion sickness during simulator environment exposure.

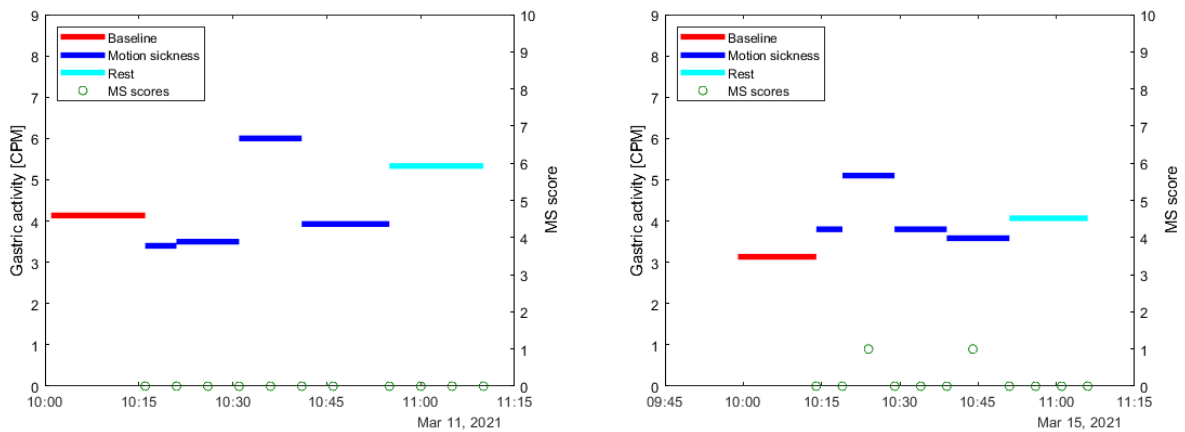


Figure 3.32: Means of CPM values over the 6 different sections composing the driving simulator scenario and collected MS scores for participant 13 (left) and 17 (right), who did not suffer from motion sickness during simulator environment exposure.

From Figure 3.31, one observes that the evolution of mean of CPM values within the motion sickness section is not similar for the 2 participants. Indeed, for participant 7, one sees that the gastric activity first increases between the first and second parts of the motion sickness section, as the MS score, and then decreases between the second and

third part of that section, while keeping a constant MS score. For participant 16, it seems to exist a positive correlation between the number of CPM and the MS score, since mean of CPM number fluctuates as the MS score does.

Considering participants who did not suffer from motion sickness during the simulator session, whose results are presented in Figure 3.32, one observes that gastric activity varies during the session while the MS score does not change. However, it does not vary in the same way for the 2 participants. For instance, between the second and third parts composing the motion sickness section, it increases for participant 13 and it decreases for participant 17. This variation in opposite ways is also observed when considering the baseline section and the first part of the motion sickness section.

One actually has to observe the results that consider all the 20 participants to conclude or not of a positive correlation between the number of CPM and the MS score. These are presented in Figure 3.33. First, one sees that means of CPM values are generally higher during the parts 2 and 3 of the motion sickness section than during the baseline and the rest sections. However, this is not true for the 2 other parts composing the motion sickness section. More importantly, one globally observes an increase in the gastric activity as the severity of motion sickness symptoms increases, although it sometimes decreases from one MS score to its directly superior one. Actually, this general increase would be better observable by grouping adjacent MS scores. Moreover, one can remark that means of CPM values for low MS scores are generally either in the upper part of the normal frequency band, or in the tachygastric frequency band. As mentioned previously, this could be due to the non-recommended conditions in which the EGG signals have been recorded.

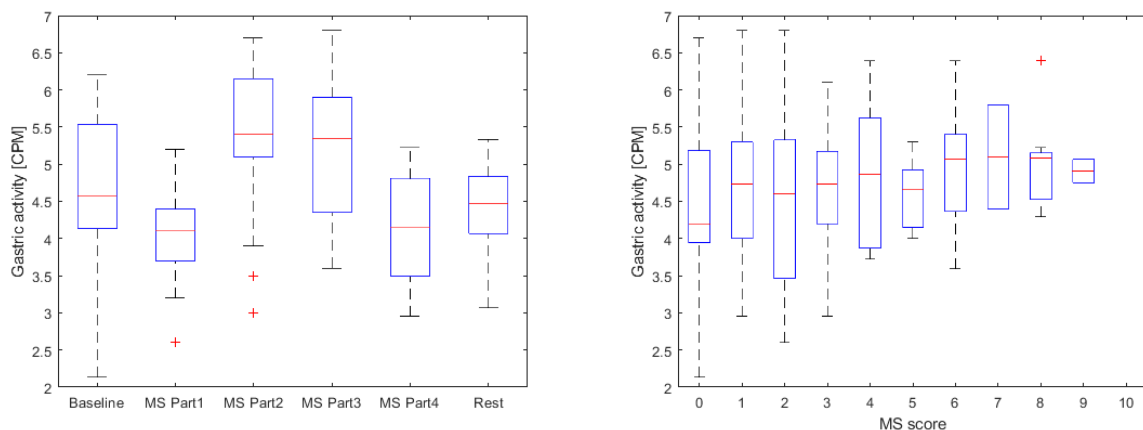


Figure 3.33: Means of CPM values over the 6 different sections composing the driving simulator scenario in function of the session period (left) and the MS score (right), for all participants.



### Car sessions

Means of CPM values over, either 5-minute periods or over the 15 minutes composing the baseline section, as well as the collected MS scores, are presented in Figures 3.34 and 3.35, for participants 3 and 8, and 5 and 18 respectively.

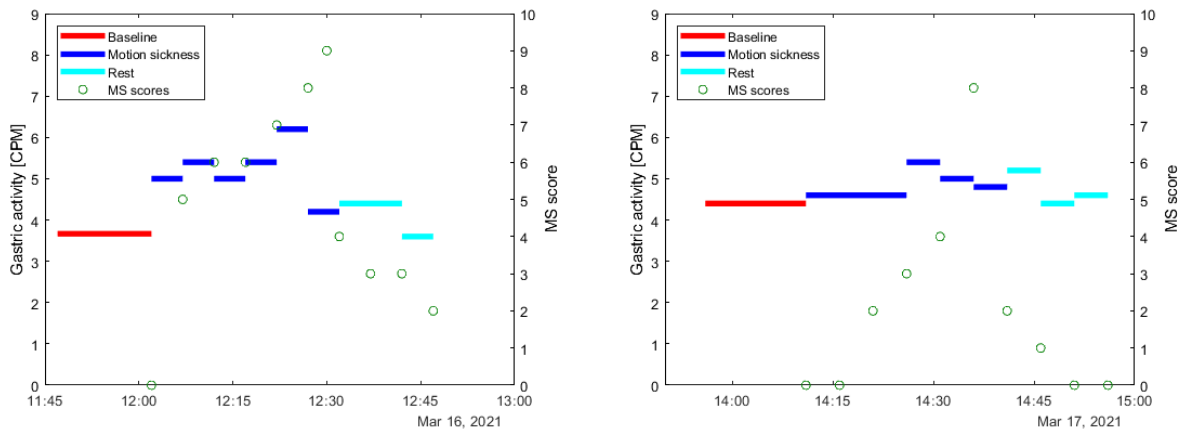


Figure 3.34: Means of CPM values over fixed time periods and collected MS scores for participant 3 (left) and 8 (right), who suffered from motion sickness during the car session.

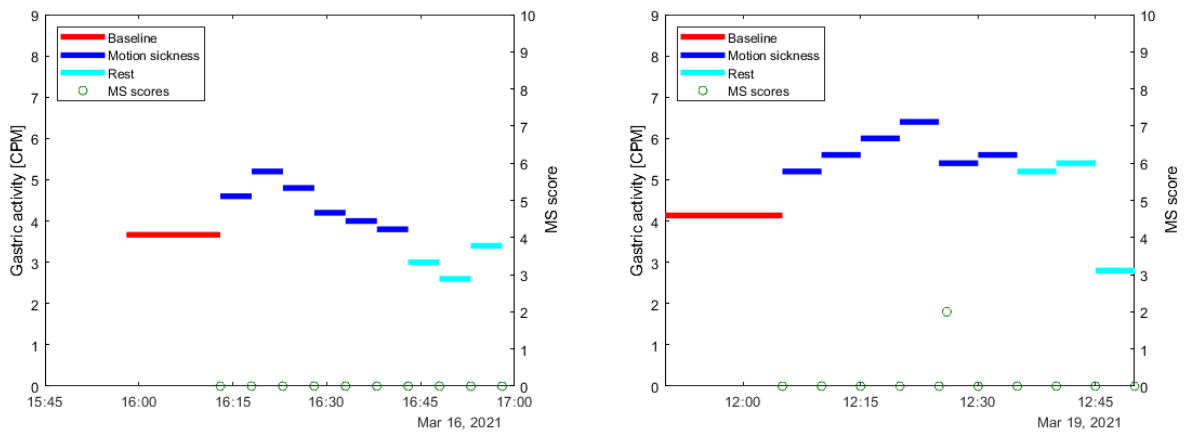


Figure 3.35: Means of CPM values over fixed time periods and collected MS scores for participant 5 (left) and 18 (right), who did not suffer from motion sickness during the car session.

It can be observed in Figure 3.34 that the gastric activity usually increases and decreases as the MS score does, although this activity sometimes decreases during the motion sickness section while the MS score increases. The smaller means of CPM values at the end of the motion sickness section are again explained by the fact that the participants

had to stop the task execution because of severe symptoms.

From Figure 3.35, one can first remark that means of CPM values of the motion sickness section vary from one part to another, and often in a different direction for the 2 participants. Moreover, one can remark that, for both participants, means of CPM values of motion sickness section are larger than those of the baseline and the rest sections, as for the 2 participants that suffered from motion sickness during the car session. Since the 2 participants considered on that Figure 3.35 did not suffer from motion sickness during the car session, the results may suggest that the motion sickness period of the section is responsible for this increased gastric activity. However, this increase can actually be explained by the curved route trajectory and braking events during the motion sickness section, having probably induced several artifacts that have been confounded with real gastric activity cycles during that section. As a consequence, this explanation suggests that the motion sickness section on itself is not responsible for an increased gastric activity, but onset of motion sickness symptoms could be related to this increase.

The global results presented in Figure 3.36 seem to go in that direction. Indeed, one sees a general increase of mean of CPM values as the MS score increase, although a decrease in gastric activity sometimes appears from one MS score to its superior adjacent. Again, this correlation would be better observed if some groups of MS scores were defined. One can also remark that means of CPM values for MS score from 7 to 9 are lower than those for a MS score of 6. This could actually be explained by the sudden stopping of the task execution when symptoms of motion sickness were too severe. Indeed, when this situation occurs, participants gave a very high MS score and it was suggested to them soon after to stop the task execution and to look at the road.

Again, the evolution of means of CPM values with the session period cannot be used on its own to conclude or not from a correlation between the evolution of the gastric activity and the severity of motion sickness symptoms, since a non negligible proportion of the participants did not suffer from motion sickness during the car sessions.

Despite the observation of a correlation between gastric activity and the intensity of motion sickness symptoms, it cannot be concluded that this physiological data is really correlated with motion sickness, since, as a reminder, the EGG signal is strongly affected by movement and physical activity.

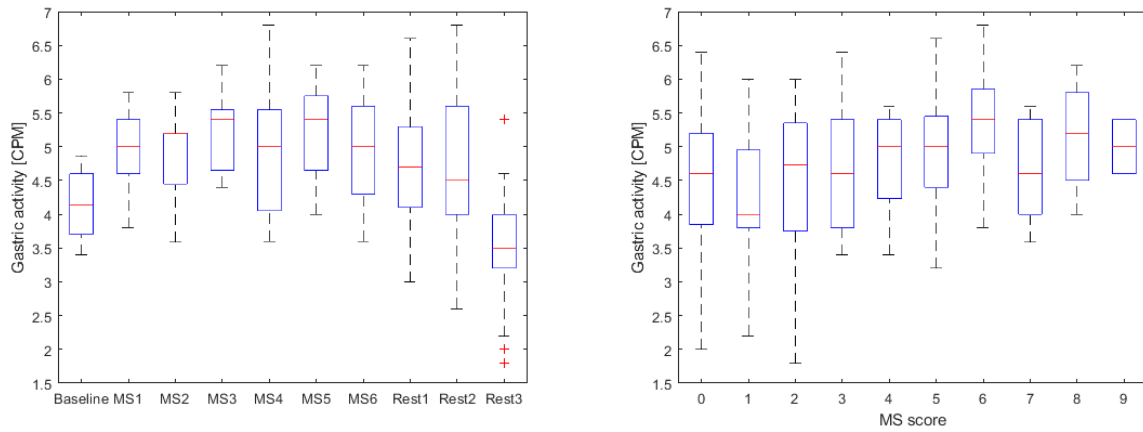


Figure 3.36: Means of CPM values over the different sections composing the car session in function of the session period (left) and the MS score (right), for all participants.

### 3.3 Ground truth definition

In the previous sections, the relationship between certain physiological parameters and the severity of motion sickness symptoms has been discussed. Ideally, these signals should be used in order to redefine a score that would be objective, and no longer subjective, and which could then be used as ground truth in machine learning models. However, this data labeling often requires multiple iterations and is therefore very long in time.

Following the analysis carried out in the previous sections, correlations between certain physiological signals and the MS score have been suggested. Thus, in order to remain pragmatic, especially within the framework of my master thesis, this subjective measure might be used as ground truth. However, in practice, it is very unlikely to see an 11-level ground truth. Moreover, it has sometimes been proposed that correlations could have been observed in a more evident way if some of the MS scores were grouped together. As a consequence, the subjective scores collected on 11 levels were grouped together on 3 levels. In particular, MS scores ranging from 0 to 3 have been associated with a score of 0, those ranging from 4 to 7 have been associated with a score of 1, and those ranging from 8 to 10 have been associated with a score of 2. This classification was made by observing carefully all the data and in such a way that the physiological data corresponding to different levels are distinguished as much as possible. These MS scores on 3 levels will constitute the ground truth.

In the following, the physiological parameters previously considered will be presented, for all the participants, no longer grouped according to the MS score varying from 0 to 10, but according to this same score reduced to 3 levels. This analysis will thus make

it possible to observe the consistency as well as the limitations of this method of data labeling.

Before considering physiological parameters, it can be useful to observe the percentage of MS scores collected during the baseline, the motion sickness and the rest sections, brought down to 3 levels. These are presented in Figure 3.37 and 3.38 for the simulator and the car sessions, respectively. The proportion of MS scores equal to 0 is much smaller for the motion sickness sections of the sessions than for the baseline and the rest sections. In particular, this proportion is even smaller for car sessions, confirming what was previously said: car sessions have caused much more onsets of motion sickness symptoms than simulator sessions.

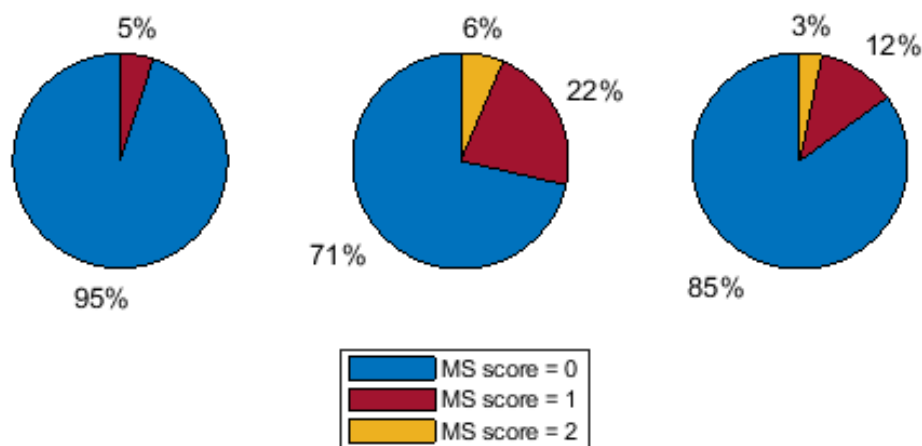


Figure 3.37: Pie chart illustrating the percentage of 3-level MS scores for the baseline (left), the motion sickness (center) and the rest (right) sections, for the simulator sessions.

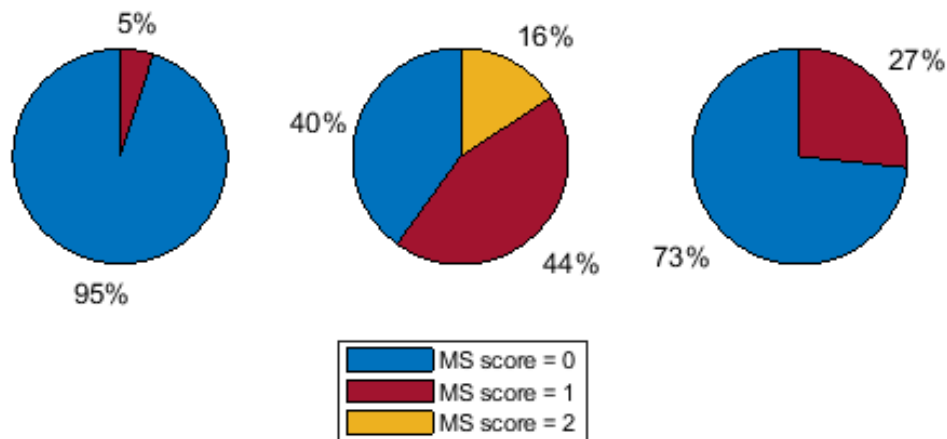


Figure 3.38: Pie chart illustrating the percentage of 3-level MS scores for the baseline (left), the motion sickness (center) and the rest (right) sections, for the car sessions.

### 3.3.1 Cardiac activity

#### Simulator sessions

Means of normalized HR values over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels are presented in Figure 3.39, for all participants. From that Figure 3.39, one observes that means of normalized HR values increase as the MS score increases. However, this increase is not very pronounced, since percentile 50 of means of normalized HR values increases by about 3% from MS score of 0 to MS score of 2.

Means of normalized HRV parameter values over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels are presented in Figure 3.40, for all participants. By looking at the general aspect that Figure 3.40, one observes that means of normalized HRV parameter values usually decrease as the MS score increases, although the percentiles 50 of means of normalized pNN20, pNN50

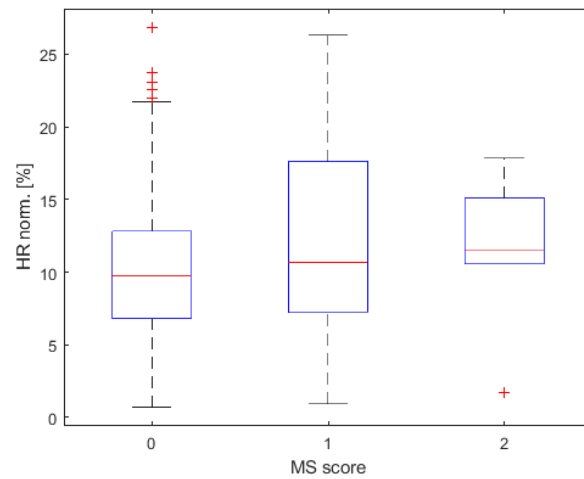


Figure 3.39: Means of normalized HR values over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels, for all participants

and RMSSD for MS score of 1 are a little above the ones for MS score of 0.

### Car sessions

Means of normalized HR values over the different sections composing the car session in function of the MS score brought down to 3 levels are presented in Figure 3.41, for all participants. From that Figure 3.41, it clearly appears that means of normalized HR values increase with the MS score. In particular, this increase is much larger for car sessions than for simulator ones, since percentile 50 of means of normalized HR values increases by about 13% from MS score of 0 to MS score of 2 in this case.

Means of normalized HRV parameter values over the different sections composing the car session in function of the MS score brought down to 3 levels are presented in Figure 3.42, for all participants.

While means of normalized HRV parameter values usually decrease as the MS score increases for simulator sessions, it is not the same for car sessions. Indeed, one observes from Figure 3.42 that all the means of normalized HRV parameter values tend to increase as the MS score increases, except for the pNN20 parameter whose values are decreasing when considering increasing MS scores. The observed behavior is not that expected, since the heart rate variability normally decreases as the heart rate increases, i.e., under conditions causing physiological stress such as motion sickness.

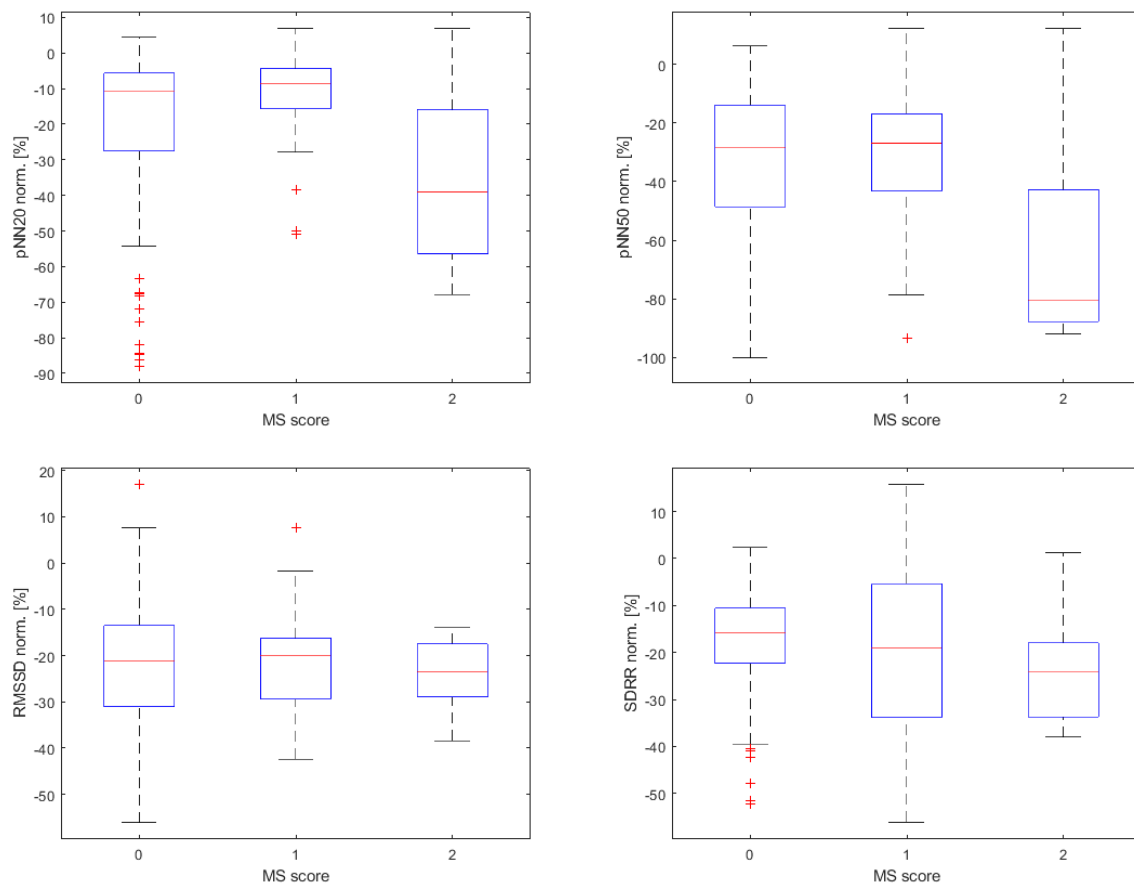


Figure 3.40: Means of normalized HRV parameter values over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels, for all participants.

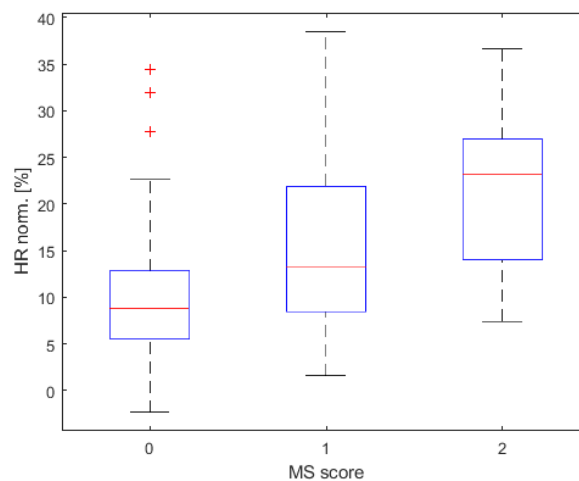


Figure 3.41: Means of normalized HR values over the different sections composing the car session in function of the MS score brought down to 3 levels, for all participants.

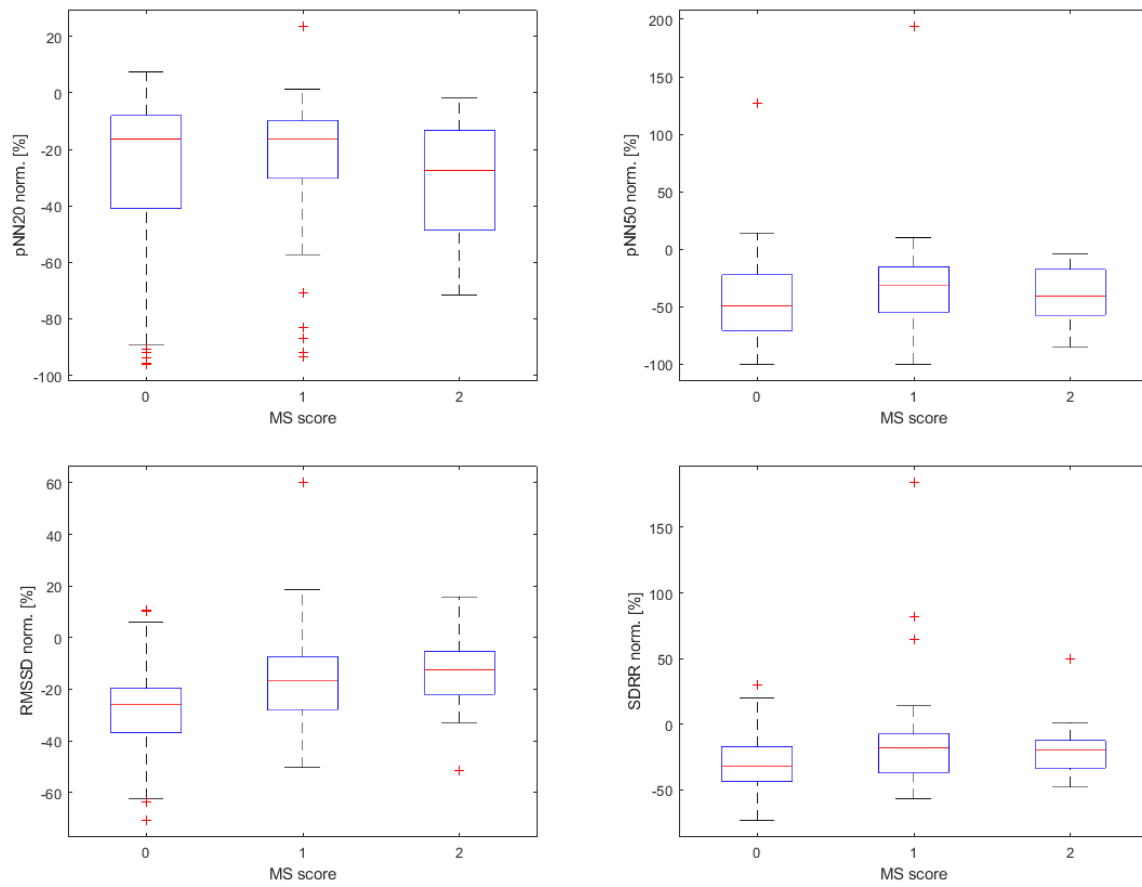


Figure 3.42: Means of normalized HRV parameter values over the different sections composing the car session in function of the MS score brought down to 3 levels, for all participants.

### 3.3.2 Electrodermal activity

#### Simulator sessions

Means of numbers of peaks per minute over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels are presented in Figure 3.43, for all EDA respondent participants with good quality signals. From that Figure 3.43, one first remarks that there is no data for MS score of 2, explained again by the fact that only EDA respondent participants with good quality signals were considered. Then, one sees that means of numbers of peaks per minute are larger when MS scores of 1 are considered. It is also interesting to remark that the number of peaks per minute does usually not exceed 8.



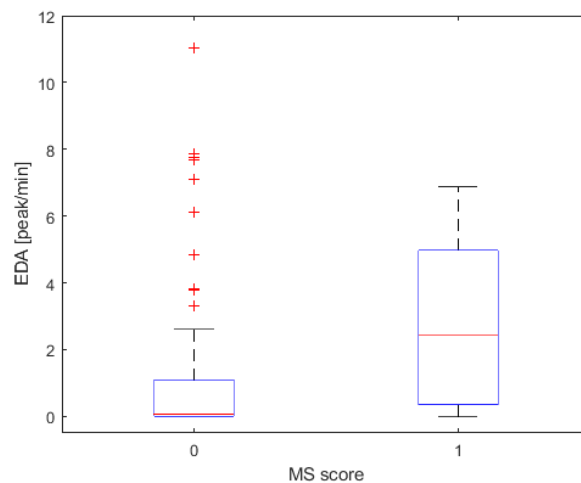


Figure 3.43: Means of numbers of peaks per minute over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels, for all EDA respondent participants with good quality signals.

### Car sessions

Means of numbers of peaks per minute over the different sections composing the car session in function of the MS score brought down to 3 levels are presented in Figure 3.44, for all EDA respondent participants with good quality signals. It clearly appears from that Figure 3.44 that means of numbers of peaks per minute increase as the severity of motion sickness symptoms increases. One also observes that the maximum number of peaks per minute corresponding to MS score of 1 is much larger for the car sessions than for simulator ones: it is almost double. This suggests that the motion sickness felt by the participants during simulator and car sessions are of different natures, the one induced during car sessions inducing a more important EDA autonomic response.

### 3.3.3 Gastric activity

#### Simulator sessions

Means of CPM values over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels are presented in Figure 3.45, for all participants. By looking at the percentile 50, a small increase of means of CPM values is observed between values associated with an increasing MS score, although this tendency is not evident when looking at the general appearance of that graph.

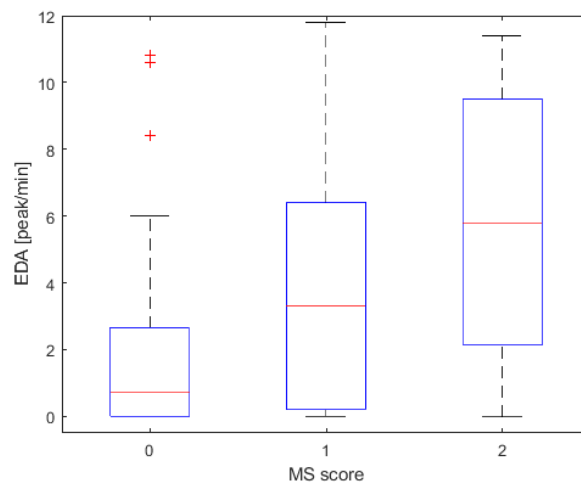


Figure 3.44: Means of numbers of peaks per minute over the different sections composing the car session in function of the MS score brought down to 3 levels, for all EDA respondent participants with good quality signals.

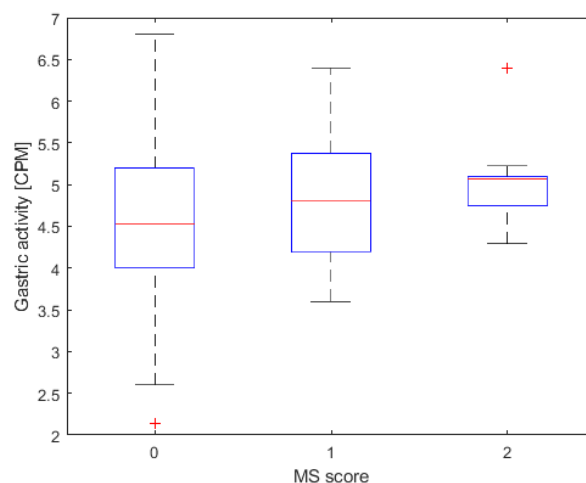


Figure 3.45: Means of CPM values over the 6 different sections composing the driving simulator scenario in function of the MS score brought down to 3 levels, for all participants.

### Car sessions

Means of CPM values over the different sections composing the car session in function of the MS score brought down to 3 levels are presented in Figure 3.46, for all participants. From that Figure 3.46, one remarks that the means of CPM values increase as the MS score increases. This evolution is more easily observable for car sessions than for simulator ones since the size of the boxplot is nearly constant in this case, and all the percentiles 25, 50 and 75 are increasing with the MS score. However, it is important to remember that the EGG signal is strongly affected by the activity of abdominal muscles, since these last separate the stomach from the recording electrodes. As these muscles are used a lot in

the usual movements, one must keep in mind that the trends observed may not exactly reflect the evolution of gastric activity.

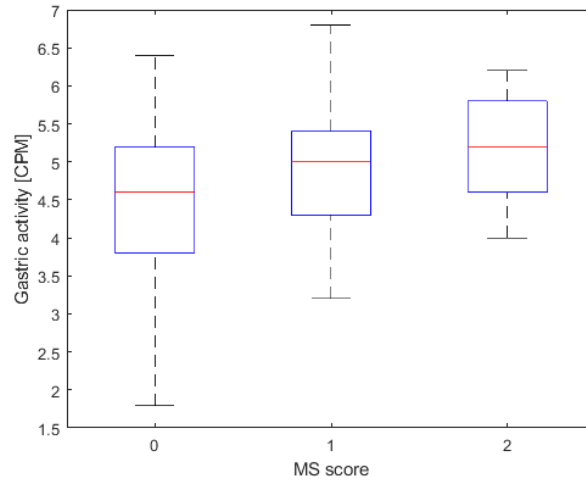


Figure 3.46: Means of CPM values over the different sections composing the car session in function of the MS score brought down to 3 levels, for all participants.

### 3.4 Conclusions and limitations

In this chapter, the evolution of physiological parameters as a function of the subjective MS score, varying from 0 to 10, has been approached. Although not always obvious, some relationships between these parameters and the level of motion sickness have been highlighted. However, one knows that this score is very subjective, consequently inducing that for the same intensity of motion sickness symptoms, 2 participants reported different MS scores, as well as different perceptions of symptom severity. Of course, the difference between these scores is usually not huge, but is well present. As a consequence, it was necessary to implement an objective MS score.

Despite the lack of physiological data labeling, the MS score on 3 levels is fairly objective. Indeed, by grouping the collected MS scores into 3 classes, an important part of the subjectivity has been removed, since the relationships existing between the physiological parameters and the severity of the motion sickness symptoms appear in a much more obvious way.

In particular, the evolution of physiological parameters with the severity of the motion sickness symptoms is, for most of the parameters, that expected. Indeed, heart rate, number of EDA peaks per minute and CPM values have all been observed to increase

with the MS score. This correlation has been observed to be much more pronounced for car sessions than for simulator ones, suggesting that the motion sickness occurring in simulator environments may be different from the one occurring in cars. It is, of course, certain that the little data collected does not allow to draw global conclusions about this difference between car sickness and simulator sickness, in particular in view of the small number of participants who suffered from motion sickness during the simulator sessions, but this thesis provides additional information in this domain that could be used to draw conclusions if combined with other studies.

As motion sickness occurring in autonomous cars is the main interest of this thesis, machine learning models aiming to predict whether or not a person is suffering from motion sickness will be at a first time based on ocular data that was collected during the car sessions. Indeed, the protocol that was developed for car sessions was intended to mimic what may happen in autonomous cars, where the driver became passenger could be engaged in many non-driving tasks.

## Chapter 4

# Detection of motion sickness based on ocular parameters

The purpose of this chapter is to identify potential motion sickness indicators through machine learning models. In particular, this research is focused on eye features, for reasons that are described in the first section. The method that has been developed is then described, including a description of the considered algorithms and of the way the relevant features have been selected as predictors. The results are finally presented and discussed.

### 4.1 Motivation

The development of autonomous cars was originally motivated by the increase of the driver comfort and productivity during transportation. The increased risk of motion sickness incidence has a significant negative impact on the original purpose of autonomous cars. As a consequence, countermeasures for decreasing the occurrence of motion sickness have to be proposed, and triggered as soon as a passenger begins to suffer from motion sickness. For instance, the posture of the passenger could be modified through a postural control device installed on the seat [63].

It is necessary to identify signs of motion sickness for countermeasures to be triggered. Thus, means of identifying such signs must be installed within the autonomous cars. In order to ensure user comfort, these means must be non-intrusive. It is therefore impossible to imagine that passengers must equip themselves with various electrodes so that certain physiological parameters can be recorded and monitored. It is precisely for this reason that this work aims to see if certain ocular parameters are indicative of motion sickness. Indeed, if this is the case, motion sickness in self-driving cars could be detected via high-precision cameras that would capture some eye parameters and that could be integrated

in the car without disturbing the passengers.

## 4.2 Method

### 4.2.1 Dataset

From the ocular data that has been recorded via the Drowsimeter R100, 150 eye features have been extracted every second. Each extracted value is computed on a 1-minute window, leading therefore to very similar data from one second to the next. The extracted features are all based on some very common eye parameters, such as the eyelid gap, the blink frequency, the blink duration, the closing duration, etc. For instance, the mean, median, mode and standard deviation of the eyelid gap are computed, as well as the percentages of time during which the eye is opened, closed, or blinking.

As values of ocular parameters are computed every second, one has to have a ground truth every second. As a reminder, MS scores were collected at the end of the baseline section, and then every 5 minutes. Since the MS score collected at a fixed time is related to the symptoms of motion sickness of the directly preceding 5 minutes (or 15 minutes for the baseline section), the 3-level MS scores have been replicated for all the seconds composing the preceding 5 (or 15) minutes, allowing therefore to have a ground truth for each second.

#### **Car sessions dataset**

As mentioned previously, ocular data that has been collected during car sessions will be first considered, since it is closely related to data that would be recorded in autonomous cars. Among all the collected car sessions data, data from 53,632 seconds (more than 14 hours) are valid and can therefore be used for training machine learning models.

This dataset is actually very unbalanced, since 34,430 observations are related to a MS score of 0, 15,320 observations are related to a MS score of 1, and only 3,882 observations are related to a MS score of 2.

#### **Simulator sessions dataset**

Ocular data collected during the simulator sessions will be considered in a second time. Data from 68,707 seconds (more than 19 hours) are valid and can be used for training machine learning models.

## 4.2.2 Method description

Different machine learning classification algorithms have been considered, including decision tree, support vector machine (SVM) and bagging and boosting ensemble methods. All these algorithms have been developed on *Matlab*.

The same procedure has been followed for all of these. Given the size of the dataset and the similarity between the observations, a nested cross-validation, i.e., an inner loop cross-validation nested in an outer loop cross-validation, has been performed. The outer cross-validation partitions the dataset into 4 subsets, each subset being composed of data from 5 participants and constituting a test set. The tuning of model hyperparameters is performed on the inner cross-validation. This last partitions the learning dataset into 5 subsets, each subset being therefore composed of data from 3 participants. For each of these 5 subsets, the model is learned on the objects that are not in the subset, and is evaluated on the samples in the subset. The evaluation metric will be discussed later in this report. The mean of the evaluation metric is then computed, and the hyperparameters maximizing this mean are used for training the model on the training and the validation sets. The performances of the model are estimated on the test set. The mean of the evaluation metric over the 4 subsets is finally computed and constitutes the expected test performance. Among all the different models, the one maximizing the mean over the 4 subsets of the outer cross-validation of the means of the evaluation metric over the 5 subsets of the inner cross-validation is considered as the best model. A pseudo-code of the procedure is presented in Appendix C.

## Algorithms description

### Decision tree

The decision tree is a tree where each internal node tests an attribute, each branch corresponds to an attribute value, and each leaf node is labeled with a class [64]. For growing the tree, the learning sample is used for choosing the attribute maximizing the class separation, the learning sample is split accordingly, and the process is repeated for the created sub-trees.

In order to get the best model of this learning technique, the split criterion (*SplitCriterion*) and the maximum number of internal nodes (*MaxNumSplits*) have been tuned. In particular, all the 3 possible split criteria have been considered: the Gini diversity index (*gdi*), the twoing rule (*twoing*) and the cross entropy (*deviance*). The maximum number of internal nodes was varied from 5 to 100, each multiple of 5.

## SVM

The SVM algorithm computes the optimal margin separating the dataset projected into a particular kernel. The optimal margin is defined as the hyperplane that is the furthest away from the closest data points of the learning dataset, if this dataset is separable (hard margin). If non-linear kernels are considered, the dataset is first projected in some new feature space thanks to a mapping. If the learning dataset is not separable, discrepancies with respect to the margin are allowed (soft margin) [64].

The kernel type (*KernelFunction*) has been tuned in order to get the best SVM algorithm as possible. In particular, polynomial and gaussian kernels have been considered. The degree of the polynomial (*PolynomialOrder*) as well as the kernel size (*KernelScale*) have been tuned, for polynomial and gaussian kernels, respectively. In particular, polynomial degrees 2 and 3 have been considered, and kernel size of 3, 6, 12, 24 and 48 have been tested.

## Bagging

The bagging ensemble method of *Matlab* uses random forest classifier by default [65]. This algorithm uses trees for predicting the output [64]. For that, it first selects a bootstrap sample (i.e., a random sampling with replacement) with which a tree is created as follows: at each node, the data is split according to the predictor maximizing the expected reduction of impurity, among a subset of all variable predictors, this last being randomly selected among all predictors at each node. A new bootstrap sample is then created and a new tree is computed according to the same process, until the forest is composed of a certain number of trees (*NumLearningCycles*). Once the forest is created, the bagging technique is used to make prediction. This technique consists in taking the data and running it down in all the trees of the forest. The output of the forest is the majority vote among all outputs from the trees. Again, the probability of belonging to the majority vote class is computed by the number of outputs of trees belonging to this class over *NumLearningCycles*. Note that there is no overfitting in this estimator due to the random subset of features which is selected at each node, the tuning of the hyperparameters is useful only to search the best evaluation metric value.

The maximum number of internal nodes (*MaxNumSplits*) has been tuned as for decision trees. The number of trees composing the forest, *NumLearningCycles*, has been fixed to 100. Indeed, in general, the more trees are grown, the more reliable estimate. However, the improvement naturally decreases as the number of trees increases. Thus, for computational time reasons, this number was limited to 100.



## Boosting

The boosting ensemble method of *Matlab* that has been considered is the *AdaBoostM2*, an adaptive boosting supporting multiclass classification problems [65]. As the random forest method, it uses trees for predicting the output. It differentiates from the bagging technique by growing *NumLearningCycles* models sequentially. In particular, it creates a strong learner by iteratively adding weak learners. At each step of the training, a new weak learner is added to the ensemble and the weights of cases being misclassified by the last model are increased, such that the algorithm focuses on the difficult cases of the learning sample [64].

The maximum number of internal nodes (*MaxNumSplits*) has been tuned, and was varied from 2 to 20, each multiple of 2. The number of models that have been grown sequentially, *NumLearningCycles*, was fixed to 100, for the same reasons as for the random forest algorithm.

## Features selection

As there are some 150 eye features, it was relevant to select a part of them for training machine learning models. Indeed, using all the eye features as predictor variables is not appropriate. Firstly, it could lead to overfitting. Secondly, there are relationships between some of these features, leading so to redundancy. Thirdly, it is likely that not all eye features have an impact on the 3-level MS score. Finally, even if all the eye features were relevant predictor variables, the time needed for training the models increases as the number of predictor variables increases, constituting also an argument for not considering all the eye features as predictor variables.

A major challenge associated to the car sessions dataset is to differentiate changes in ocular parameters due to motion sickness from changes related to the task performed, and so to determine whether or not a particular eye feature is relevant for predicting the MS score. Indeed, during the motion sickness section of the car sessions, the participants had to perform some tasks on paper, which caused them to look down. As a consequence, most of the ocular parameters evolved differently in the motion sickness section compared to those of the baseline and the rest sections. Since motion sickness was mainly induced during the motion sickness section, ocular parameters that could possibly be indicative of motion sickness may have been impacted during the same period than the one during which participants looked down.

One way to determine the relevance of an eye feature to predict the level of motion sickness is to look at its evolution with the MS score and with the period of the car session. If the eye feature value increases (resp. decreases) with the MS score, one might think that this feature is correlated with motion sickness and thus can be considered as a predictor variable. However, if this feature value is also larger (resp. smaller) during the motion sickness section than during the baseline and the rest sections, one might think that this feature is correlated to the task performed during the motion sickness section. One has therefore to compare the eye feature evolution with the MS score and with the period. If that evolution is more pronounced as the MS score increases than as the period varies, one could conclude that the considered eye feature is correlated to the MS score. On the contrary, if the increase (or decrease) of eye feature value is larger as the period varies than as the MS score increases, one might suppose that this eye parameter varies with the task performed and not with the level of motion sickness.

As an example, the evolution of the percentage of time during which the eye is open with the MS score and with the period of the car session is presented in Figure 4.1, considering all the participants. From this Figure 4.1, one can observe that the percentage of time during which the eye is open increases with the MS score, but also during the motion sickness period. Because of the relative increase of this parameter in the 2 graphs, one cannot conclude that this eye feature is relevant for predicting the level of motion sickness. It is actually very difficult to reliably identify the ocular parameters which depend on the severity of the motion sickness compared to those depending on the task performed.

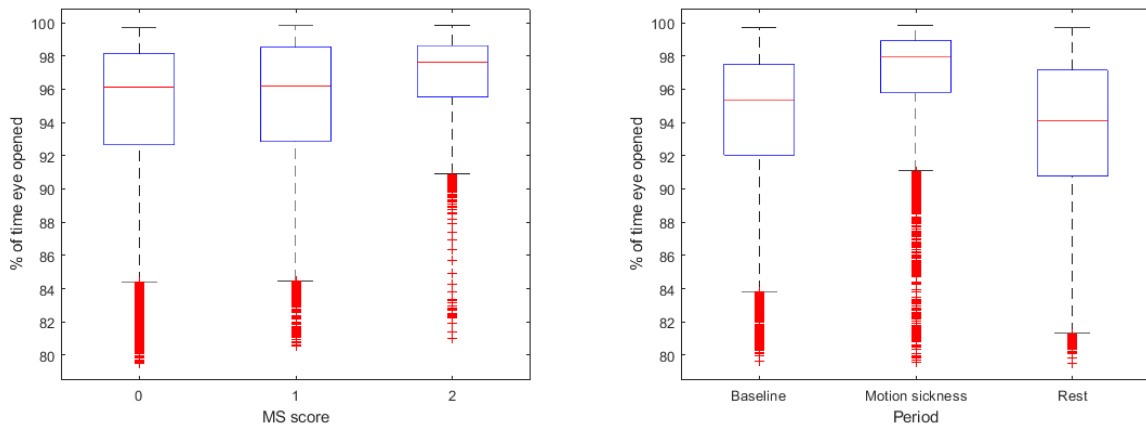


Figure 4.1: Evolution of the percentage of time where the eye is opened with the MS score (left) and with the period (right) of the car session, for all participants.

As a consequence, the relevance between features and labels (3-level MS scores) has been analyzed through the *fscmr* selection function of *Matlab*, for each training dataset.

This function ranks features using minimum redundancy maximum relevance (MRMR) algorithm. This last finds an optimal set of features that is mutually and maximally dissimilar and can represent the response variable effectively [66]. All the features with a score superior to 0.05 have been selected as predictor variables. Since this set can be different according to the considered training dataset, the set of features used for training the models on the set composed of the training and validation sets is the union of all the 5 sets.

## 4.3 Results

The results are presented in this section. In particular, 3 different considerations are analyzed in a sequential way, each as a result of the previous one.

### 4.3.1 First consideration

All the dataset of the car sessions is first used for training machine learning models. As mentioned previously, the 3 classes are not represented equally in this dataset. In order to compensate this problem, the matrix of misclassification costs,  $M$ , has been modified according to the distribution of the different classes in the dataset. The rows of this matrix are the true classes, while its columns are concerned with the predicted classes. The elements below the main diagonal have been modified as follows:

$$M = \begin{bmatrix} 0 & 1 & 1 \\ x & 0 & 1 \\ y & x & 0 \end{bmatrix}$$

where  $x$  and  $y$  are respectively equal to the ratio of the number of observations corresponding to a MS score of 1 divided by the number of observations corresponding to a MS score of 0, and to the ratio of the number of observations corresponding to a MS score of 2 divided by the number of observations corresponding to a MS score of 0. The matrix has been modified so that the penalty of misclassifying an observation corresponding to a MS score of 1 or 2 is greater than that of misclassifying an observation corresponding to a MS score of 0. This choice was motivated by the ultimate goal of triggering countermeasures when a passenger suffers from motion sickness in an autonomous car. Indeed, these measures have to be triggered as soon as a passenger begins to feel motion sickness symptoms so that they do not intensify, corresponding to a MS score of 1. The countermeasures have obviously to be triggered if a passenger suffers from severe motion sickness symptoms, corresponding to a MS score of 2, while they have to be not triggered if the passenger feels perfectly well, i.e., with a MS score of 0, but do not lead to negative

consequences if triggered.

For decision trees, these misclassification costs influence the way the data is split. Indeed, the split maximizing the class separation is chosen at each step, i.e., making successors as pure as possible. The sum of the number of examples in each class within each node has to be computed for calculating the purity measure. This sum can actually be replaced by a weighted sum, in which higher weights are assigned to instances coming from a class with a higher value of misclassification cost [67]. In SVM models, different weights are assigned to different data points such that the training algorithm learns the decision surface according to the relative importance of data points in the training dataset [68]. In particular, the soft margin parameter,  $C$ , can be computed as a weighting of the global  $C$ -value for each example of the training dataset. The margins are therefore harder for the minority classes.

The area under the Precision-Recall (PR) curve (AUPRC) has been chosen to be the evaluation metric. The PR curve is a plot of the recall (x-axis) versus the precision (y-axis) for a number of different candidate threshold values between 0 and 1. The use of this metric actually requires to deal with a binary classification problem, since it considers a positive class and a negative one. Indeed, precision is the ratio of the number of true positives divided by the sum of the true positives and false positives [64]. It describes how good a model is at predicting the positive class. Recall, also called sensitivity, is the ratio of the number of true positives divided by the sum of the true positives and false negatives [64]. Thus, calculations of precision and recall are only concerned with the correct prediction of the positive class.

This evaluation metric has been preferred rather than the area under the receiver operating characteristic (ROC) curve since this last is not suited for unbalanced datasets, on the contrary to the AUPRC. Indeed, a model that cannot discriminate between the classes and would predict a random class is characterized by an area under the ROC curve value of 0.5, whatever the distribution of the dataset. On the contrary, the AUPRC value of a no-skill classifier depends on the distribution of the dataset, and is equal to the proportion of positive cases in the dataset [69].

Since the considered classification task is a 3-class classification problem, 2 of them have to be grouped for calculating the evaluation metric value. In particular, the calculation of the metric value is performed by considering the confidence probabilities associated with the positive class and the negative one. It was preferred to group the probabilities

of the classes MS score = 1 and MS score = 2 rather than those of the classes MS score = 0 and MS score = 1, thinking again to the ultimate goal of triggering countermeasures when a passenger suffers from motion sickness in an autonomous car. The positive class was naturally defined as the one corresponding to MS scores equaling 1 and 2, while the negative class was the one that corresponds to MS scores of 0.

The results are presented in Table 4.1, for each machine learning algorithm. In particular, the mean over the 4 subsets of the outer cross-validation of the means of the AUPRC over the 5 subsets of the inner cross-validation (AUPRC val.) is presented, as well as the associated weighted accuracy (accuracy val.). Moreover, the means of the AUPRC and of the weighted accuracy computed on the 4 test sets (AUPRC test and accuracy test) are presented. The weighted accuracy is computed as follows:

$$accuracy = \frac{1}{N} \sum_{i=1}^N \frac{\text{predictedclass} = i | \text{trueclass} = i}{\text{trueclass} = i} \quad (4.1)$$

where N is the number of classes in the considered dataset.

The table 4.2 contains the optimal values of hyperparameters. More specifically, as the optimal hyperparameters can be different from one subset of the outer cross-validation to the other, either the mean value or the mode (depending on the nature of the considered hyperparameter) of the 4 hyperparameters values (Opt. hyperparam.) is presented in Table 4.2 for the sake of clarity.

Algorithm	AUPRC val.	accuracy val.	AUPRC test	accuracy test
Decision Tree	0.4021	0.3614	0.3470	0.3666
SVM	0.4132	0.3601	0.3468	0.3847
Bagged Tree	0.4313	0.3309	0.3684	0.3509
Boosted Tree	0.4273	0.3383	0.3579	0.3551

Table 4.1: Mean over the 4 subsets of the outer cross-validation of the means of the AUPRC over the 5 subsets of the inner cross-validation, associated weighted accuracy, means of the AUPRC and of the weighted accuracy computed on the 4 test sets, for all the learning algorithms trained on the entire car sessions dataset.

From Table 4.1, one can first observe that the mean over the 4 subsets of the outer cross-validation of the means of the AUPRC over the 5 subsets of the inner cross-validation (AUPRC val.) is smaller than 0.5, for all the models. As mentioned previously, one cannot conclude that these models cannot discriminate between the classes, since such a model is characterized by an AUPRC value equals to the proportion of positive cases in

Algorithm	Opt. hyperparam.
Decision Tree	<ul style="list-style-type: none"> <li>• <i>SplitCriterion</i> = <i>gdi</i></li> <li>• <i>MaxNumSplits</i> = 30</li> </ul>
SVM	<ul style="list-style-type: none"> <li>• <i>KernelFuntion</i> = <i>gaussian</i></li> <li>• <i>KernelScale</i> = 27</li> </ul>
Bagged Tree	<ul style="list-style-type: none"> <li>• <i>MaxNumSplits</i> = 5</li> </ul>
Boosted Tree	<ul style="list-style-type: none"> <li>• <i>MaxNumSplits</i> = 10</li> </ul>

Table 4.2: Optimal values of tuned hyperparameters, for all the learning algorithms trained on the entire car sessions dataset.

the dataset. As there is an AUPRC value for all the validation sets, one would have to compute this proportion for all of these, and to report the corresponding AUPRC. For the sake of simplicity, one can assume that the proportion of positive cases in the entire dataset is an appropriate approximation, it is equal to 0.358. One therefore sees that all the models AUPRC val. values are a little larger than this value. One might therefore conclude that the models are hardly better than no-skill classifiers.

By looking at the corresponding weighted validation accuracy, one actually understands that all the models perform very poorly. The accuracy val. values are indeed very close to one third, meaning that the predictions are random.

Based on the evaluation criteria, the bagging algorithm is the best one since it maximizes the AUPRC val. However, by looking at the value of *MaxNumSplits*, one sees that it is actually the lower bound of the interval considered for tuning, meaning that this complex model does not gain skill when grown trees are deeper. For knowing the number of internal nodes maximizing the AUPRC val. value, the lower limit of the considered interval of *MaxNumSplits* should have been reduced. However, it is actually very surprising that this number was the lowest considered, since bagging ensemble methods usually grow deep trees, these last being grown independently. The number of internal nodes has been thus fixed to 1,000 in order to understand what was happening during the tuning. The AUPRC val. value obtained with these huge number of *MaxNumSplits* was equal to 0.3921, and the corresponding validation accuracy was equal to 0.3464. Given the fact that these values are very close to the ones with a *MaxNumSplits* value of 5, one deduces that the model is not able to differentiate between classes.

The expected performances on the test sets also seem very poor in view of the value of the AUPRC test and that of the accuracy test. In particular, the performances are even worse on the test sets than on the validation ones. This observation was obviously

expected, since the optimization process is performed on the validation sets. Note that the value of the AUPRC test corresponding to a no-skill classifier would be exactly equal to the average over the test sets of the proportion of positive cases in each test set. This value is actually very close to the proportion of positive cases in the entire dataset since the test sets contain nearly the same number of observations. One can look at the contingency tables presented in Table 4.3, for each of the 4 test sets, resulting from the bagging technique. The class 0 refers to a MS score of 0, the class 1 refers to a MS score of 1 and the class 2 refers to a MS score of 2.

(a)				(b)			
True class	Predicted class			True class	Predicted class		
	0	1	2		0	1	2
0	6,127	866	357	0	6,310	54	664
1	2,278	169	149	1	5,028	407	311
2	704	36	24	2	1,407	18	272

(c)				(d)			
True class	Predicted class			True class	Predicted class		
	0	1	2		0	1	2
0	4,359	4,029	253	0	3,118	2,807	5,486
1	1,635	1,569	157	1	1,605	1,022	990
2	272	486	79	2	109	165	310

Table 4.3: Contingency tables for the first (a), second (b), third (c) and fourth (d) test set, resulting from the bagging technique with the optimal hyperparameter value, for the first consideration.

From all of these 4 contingency tables, one generally remarks that the predictions are random. Indeed, the weighted accuracies are respectively equal to 0.310, 0.376, 0.356 and 0.362 for the first, second, third and fourth test set. While the class 0 seems to be well predicted for the first and second test set, it is clearly not the same for the 2 others. Observations corresponding to MS score 1 are generally very badly classified, since it is never the most predicted class, whatever the test set. A similar observation can be made for observations related to a MS score of 2, except for the fourth test set for which more than a half of these observations are well classified.

Thinking about the triggering of the countermeasures for reducing the intensity of motion sickness symptoms, several important observations can be made from the Table 4.3. Ideally, countermeasures should be triggered in a gradual fashion, i.e., specific measures would be triggered for MS scores of 1 and 2, respectively. Thus, if it turns out that the intensity of the symptoms increases despite the initiation of the first countermeasures, other countermeasures must be initiated. Also, if a passenger suddenly develops severe symptoms, all countermeasures should be initiated. By looking at the Table 4.3, one deduces that countermeasures for severe motion sickness symptoms would almost never be triggered when the person suffers from severe symptoms of motion sickness. Assuming now that the same countermeasures were triggered for moderate and severe motion sickness symptoms, i.e., MS scores of 1 and 2, these would only be triggered less than once in 2 (about 43%) when the passenger was suffering from severe symptoms, on average. Under the same assumption, the countermeasures would be triggered less than once in 3 (33%) when the passenger was suffering from moderate symptoms. Finally, countermeasures would be triggered about 37% of time when it is not needed.

From all these observations, it can be concluded that machine learning models cannot classify correctly the observations of the considered dataset, MS scores are predicted in a random way.

### 4.3.2 Second consideration

Since the previous obtained results were not concluding, it was thought to only consider the observations of the same dataset that belong to the motion sickness section. Indeed, the participants had to look down during the motion sickness section for performing the tasks on paper, while it was not the case during the baseline and the rest sections. As a consequence, ocular parameters have undergone changes due to the performed tasks in addition to those possibly induced by motion sickness. It therefore makes sense to only consider the values of the eye parameters during the motion sickness section, so that the machine learning algorithms deal with very similar values of ocular parameters.

The considered dataset is composed of 21,627 observations. More precisely, 7,922 observations are related to a MS score of 0, 9,823 observations are related to a MS score of 1, and 3,882 observations are related to a MS score of 2. Although the 3 classes are still not equally represented, the proportion of observations that correspond to a MS score either equal to 1 or equal to 2 is much larger than in the entire dataset. One even observes that the number of observations related to a MS score of 1 is larger than that related to a MS



score of 0.

The matrix of misclassification costs has been updated according to these proportions. In particular, only the first element of the last line has been modified, corresponding to an observation related to the MS score 2 that would be predicted as a MS score of 0.

The same method as the one for the entire dataset was used, except that the evaluation metric had to be modified. Indeed, as most of the observations that correspond to a MS score of 0 have been removed by considering only the observations of the motion sickness section, some validation sets no longer contained observations corresponding to a MS score of 0. No observation therefore belonged to the negative class, and the AUPRC could not be computed. As a consequence, the weighted accuracy defined in Equation 4.1 has been used as the evaluation metric.

The results are presented in Table 4.4 for each machine learning algorithm. The mean over the 4 subsets of the outer cross-validation of the means of the weighted accuracy over the 5 subsets of the inner cross-validation (accuracy val.) is presented. The means of the accuracy and of the AUPRC computed on the 4 test sets (accuracy test and AUPRC test) are also presented. The optimal values of hyperparameters are presented in Table 4.5, in the same way as in the Table 4.2.

Algorithm	accuracy val.	accuracy test	AUPRC test
Decision Tree	0.3393	0.2958	0.6370
SVM	0.3177	0.3434	0.5709
Bagged Tree	0.3164	0.3244	0.5738
Boosted Tree	0.3151	0.2837	0.5426

Table 4.4: Mean over the 4 subsets of the outer cross-validation of the means of the weighted accuracy over the 5 subsets of the inner cross-validation, means of the weighted accuracy and of the AUPRC computed on the 4 test sets, for all the learning algorithms trained on the reduced car sessions dataset, composed of observations belonging to the motion sickness section.

From Table 4.4, it actually appears that the results are not better than the ones obtained when all the dataset from car sessions was considered. Indeed, one remarks at first that the accuracy val. values are again close to one third, reflecting random predictions. It is even worse than previously, since some of the considered validation sets have observations corresponding to only 2 classes. Their weighted accuracy has so to be close to 0.5 if the predictions were random. As a consequence, the mean over the 4 subsets of the outer cross-validation of the means of the weighted accuracy over the 5 subsets

Algorithm	Opt. hyperparam.
Decision Tree	<ul style="list-style-type: none"> <li>• <i>SplitCriterion</i> = <i>gdi</i></li> <li>• <i>MaxNumSplits</i> = 44</li> </ul>
SVM	<ul style="list-style-type: none"> <li>• <i>KernelFuntion</i> = <i>gaussian</i></li> <li>• <i>KernelScale</i> = 39</li> </ul>
Bagged Tree	<ul style="list-style-type: none"> <li>• <i>MaxNumSplits</i> = 30</li> </ul>
Boosted Tree	<ul style="list-style-type: none"> <li>• <i>MaxNumSplits</i> = 7</li> </ul>

Table 4.5: Optimal values of tuned hyperparameters, for all the learning algorithms trained on the reduced car sessions dataset, composed of observations belonging to the motion sickness section.

of the inner cross-validation could be larger than one third for random predictions. In particular, 3 out of 20 validation sets contain only 2 classes. Then, the accuracy val. value for a no-skill classifier would be equal to 0.358.

Moreover, the decision tree algorithm is the best one since it maximizes the accuracy val. value. This observation again confirms that the classifiers are not able to discriminate between classes by considering the ocular parameters, since very simple models perform better than the complex ones.

Note that the larger AUPRC test values than those considering all the dataset can be explained by the larger proportion of positive cases in this reduced dataset: 0.634. Thus, the expected test performance is very bad in view of the values of the AUPRC test.

### 4.3.3 Third consideration

In view of the previous results, it appears that the models cannot differentiate the 3 classes on the basis of the ocular parameters considered. This could in part be explained by the specific task performed by participants during the motion sickness section, causing different eye behavior. For instance, this could have led to wrong extractions of the ocular parameters, such that the association of some downward gazes with an eye closure. Although the passengers in self-driving cars will be engaged in various tasks which will also cause them to look down, one can try to see if the eye features extracted from the data recorded by the Drowsimeter R100 are indicative of motion sickness, by considering the data collected during the simulator sessions. As a reminder, the induction of motion sickness in the driving simulator was done via a driving scenario, so the participants did not have to perform tasks other than driving during the whole session.

As mentioned previously, this dataset is composed of 68,707 observations, in which respectively 56,304, 9,107 and 3,296 are related to MS scores 0, 1 and 2. This unbalance of the dataset was expected since much more participants suffered from motion sickness during the car sessions than during the simulator ones. Note that this dataset contains observations of only 19 participants, since the ocular data of one participant has not been recorded.

The matrix of misclassification costs has been updated according to this distribution of the simulator sessions dataset, and the same method as the one used for the second consideration was used. Indeed, the weighted accuracy was again used as the evaluation metric because of the non-possible computation of the AUPRC. However, while it was because no observation belonged to the negative class (MS score = 0) for several validation sets, it is here due to the absence of observations belonging to the positive class (MS scores equaling 1 and 2).

The results are presented in Table 4.6 for each learning technique. As previously, the mean over the 4 subsets of the outer cross-validation of the means of the weighted accuracy over the 5 subsets of the inner cross-validation (accuracy val.) is presented. The means of the accuracy and of the AUPRC computed on the 4 test sets (accuracy test and AUPRC test) are also presented. The optimal values of hyperparameters are presented in Table 4.7.

Algorithm	accuracy val.	accuracy test	AUPRC test
Decision Tree	0.5442	0.5063	0.2453
SVM	0.5520	0.4369	0.3570
Bagged Tree	0.5483	0.5139	0.4103
Boosted Tree	0.5314	0.4864	0.3591

Table 4.6: Mean over the 4 subsets of the outer cross-validation of the means of the weighted accuracy over the 5 subsets of the inner cross-validation, means of the weighted accuracy and of the AUPRC computed on the 4 test sets, for all the learning algorithms trained on the simulator sessions dataset.

In view of the values of the accuracy val. of the Table 4.6, one might think that the models manage to classify the observations in a satisfying way. However, it must be kept in mind that a no-skill model would have an accuracy val. value greater than one third, since all the validation sets do not contain observations that belong to the 3 classes. In particular, a no-skill classifier would be characterized by an accuracy val. value equal to 0.5. Therefore, one observes that all of the considered classifiers perform slightly better than a no-skill one, but this does not allow to suggest that the provided ocular parameters

Algorithm	Opt. hyperparam.
Decision Tree	<ul style="list-style-type: none"> <li>• <i>SplitCriterion</i> = <i>gdi</i></li> <li>• <i>MaxNumSplits</i> = 25</li> </ul>
SVM	<ul style="list-style-type: none"> <li>• <i>KernelFuntion</i> = <i>gaussian</i></li> <li>• <i>KernelScale</i> = 32</li> </ul>
Bagged Tree	<ul style="list-style-type: none"> <li>• <i>MaxNumSplits</i> = 31</li> </ul>
Boosted Tree	<ul style="list-style-type: none"> <li>• <i>MaxNumSplits</i> = 5</li> </ul>

Table 4.7: Optimal values of tuned hyperparameters, for all the learning algorithms trained on the simulator sessions dataset.

could be indicators of motion sickness.

As for the accuracy val. value, the accuracy test value of a no-skill classifier would not be equal to one third, since some of the test sets contain only 2 classes. Actually, it would be equal to about 0.416. As a consequence, one sees that the algorithms perform again slightly better than a no-skill classifier, on the test sets. This can be confirmed by looking at the AUPRC test values. These are all larger than 0.195, which is the AUPRC value corresponding to a no-skill classifier, by considering the average of the proportions of positive cases in each test set.

Based on the accuracy val. values, the SVM algorithm would be considered to be the best one, since its accuracy val. value is the largest. One can look at the contingency tables presented in Table 4.8 , for each of the 4 test sets, resulting from the SVM technique. As previously, the class 0 refers to a MS score of 0, the class 1 refers to a MS score of 1 and the class 2 refers to a MS score of 2.

As for the first consideration, one usually sees that the class 0 is well predicted, except for the first test set. Observations belonging to classes 1 and 2 are unfortunately usually very badly classified. Assuming again that same countermeasures were triggered for moderate and severe motion sickness symptoms, these would only be triggered about once in 5 when the passenger was suffering from severe symptoms, on average. Under the same assumption, the countermeasures would be triggered less than once in 3 when the passenger was suffering from moderate symptoms. Finally, countermeasures would be triggered about 19% of time when it is not needed.

(a)					(b)				
True class	Predicted class				True class	Predicted class			
	0	1	2			0	1	2	
0	11,732	6,330	57		0	11,344	3,000	49	
1	815	0	0		1	1,289	1,681	95	
2	0	0	0		2	2,175	351	118	

(c)					(d)				
True class	Predicted class				True class	Predicted class			
	0	1	2			0	1	2	
0	13,615	1,809	964		0	7,252	150	2	
1	820	171	190		1	2,317	1,711	18	
2	0	0	0		2	525	111	16	

Table 4.8: Contingency tables for the first (a), second (b), third (c) and fourth (d) test set, resulting from the SVM technique with the optimal hyperparameters values, for the third consideration.

## 4.4 Conclusions and limitations

From the 3 considerations discussed, it appears that the ocular parameters extracted from the data collected by the Drowsimeter R100 do not make it possible to predict the level of motion sickness. Indeed, the 2 considerations that concern the car sessions show that the algorithms predict the level of motion sickness in a totally random way, and the consideration concerning the simulator sessions does not prove that the predictions are not random, even if slightly better.

After assuming that data collected under different conditions could be responsible for the inability of the models to classify the data, the second consideration showed that performing the tasks, causing the participant to look down, also prevented classifiers from classifying correctly the data. The third consideration was carried out to see if eye parameters could be indicative of motion sickness. Indeed, the considered data does not contain (or very few) glances downwards, since the participants looked at the screen in front of them. Thus, the results of this last consideration were not affected by varying values of ocular parameters and wrong extraction of ocular data. Although slightly better than the previous ones, the results obtained do not allow to confirm that the ocular parameters extracted from the data collected by the Drowsimeter R100 are indicators of motion sickness. Given the very low proportion of observations showing motion sickness in this dataset, one cannot, of course, conclude with certainty that there is no correlation between eye parameters and motion sickness.

The extraction of the eye features could explain in part why the models are not able to classify the observations into 3 classes, corresponding to the severity of motion sickness symptoms. Indeed, the extraction of ocular parameters values requires an important processing of the ocular data collected by the Drowsimeter R100. In particular, assumptions and conditions are introduced during the processing in order to extract coherent values. However, all these conditions were defined before this work, and are thus not based on the collected data. This could have led to non coherent extracted values. It might therefore be appropriate to adapt the conditions to the present dataset.

The ocular parameters used to train the models may also be responsible for the inability to correctly classify the data. They were indeed selected by an algorithm, and not by visual inspection. The selection of ocular parameters and their correlation with motion sickness could therefore be further investigated.

Finally, it is obvious that the data that has been collected for this work does not allow to draw global conclusions about the relationship between eye parameters and motion sickness, and their ability to predict this state. It would be wise to collect more data and investigate different conditions that can lead to the onset of motion sickness in autonomous cars.

# Conclusion

The aim of this work was to identify physiological indicators of motion sickness, and more particularly to consider the relationship between certain ocular parameters and the severity of the symptoms of motion sickness. This identification is the first step in the solution aiming to develop countermeasures against motion sickness in autonomous cars, in order to ensure their acceptance by as many people as possible.

Data from 20 participants were collected for this study. Certain physiological parameters, such as heart rate, number of peaks per minute of electrodermal activity, and number of cycles per minute of gastric activity, have been shown to be correlated with the severity of motion sickness. This observation made it possible to define a ground truth for machine learning models. The purpose of these last was to see if the prediction of a level of motion sickness was possible on the basis of ocular parameters. The results showed that the eye features alone cannot predict motion sickness.

Further investigations are required regarding the prediction of the severity of motion sickness symptoms with physiological parameters. Concerning the eye data, the evolution of pupil parameters with motion sickness could be considered, the pupillary diameter being affected by the activity of the sympathetic and parasympathetic nervous systems. However, the recording of these parameters in real conditions seems compromised, since a high measurement precision would be necessary. It also appears that the diameter of the pupil is subject to variations depending on the degree of light, making it difficult to identify motion sickness via this parameter.

It may be appropriate to consider ocular parameters with other physiological data in order to predict motion sickness. The most promising investigation seems to be the prediction of the level of motion sickness using the heart rate, for several reasons. First of all, an important correlation between changes in heart rate and the severity of motion sickness has been shown. Then, the recording of the heart rate of the passengers in the autonomous car could be done easily, in particular thanks to smartwatches, high-resolution cameras detecting the beats by the movements of the skin, or even thanks to sensors



placed in the seats. Finally, the heart rate can be extracted in real-time from the ECG or PPG, and therefore be used for the detection of motion sickness.

# Appendices

# Appendix A

## Protocol details

### A.1 Route of the car session

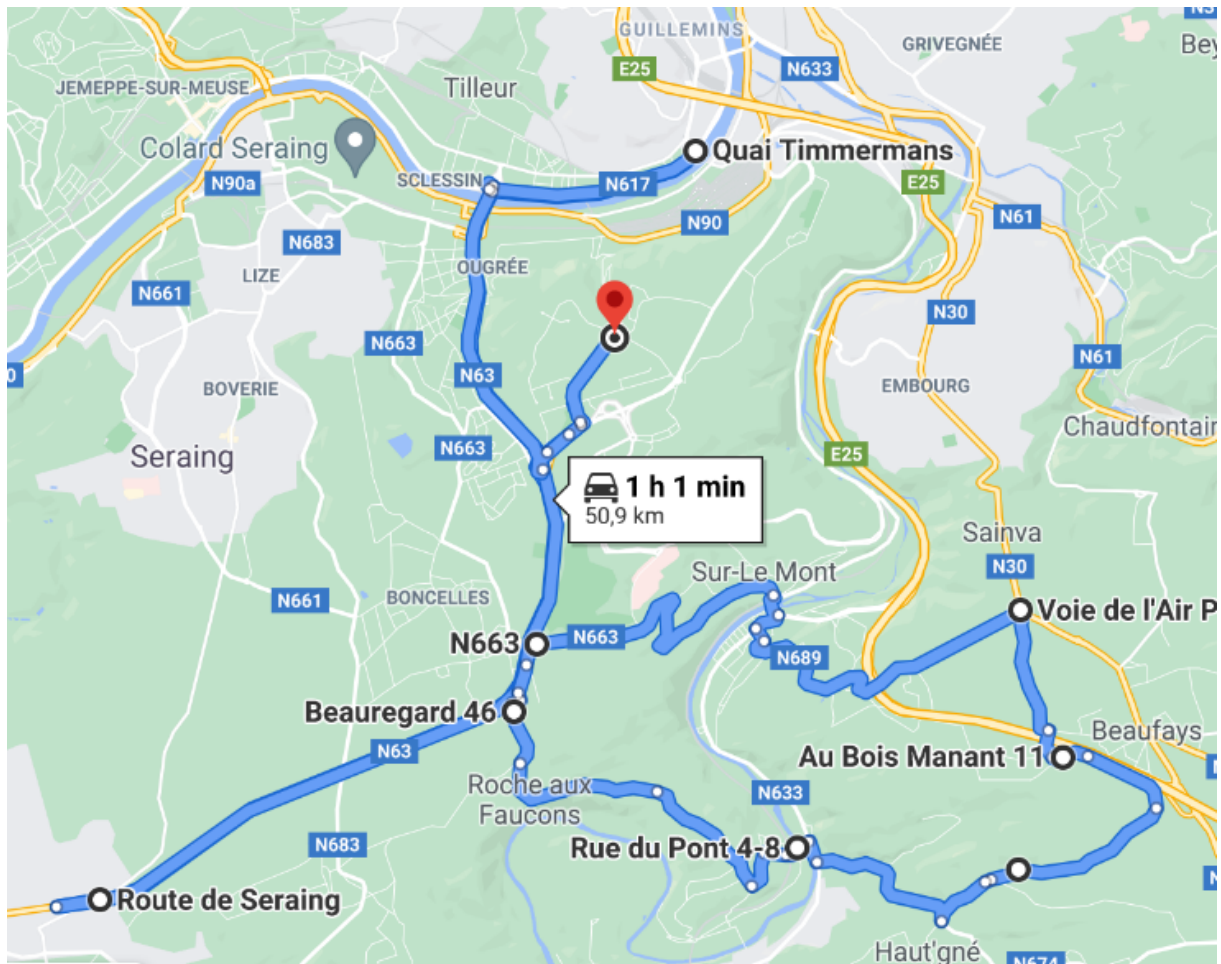


Figure A.1: Route of the car session.

## A.2 Tasks proposed to the participants during the motion sickness section of the car session

SUJET:

DATE:

REPONDEZ AUX QUESTIONS CI-DESSOUS APRES AVOIR LU L'HISTOIRE  
ATTENTIVEMENT. COCHEZ LA PROPOSITION CORRECTE.

### Le bonheur de la vie de famille

Chaque matin, mon fils Jacques se réveille en premier et prend son petit-déjeuner à sept heures. Pour l'instant, la maison est encore calme. Puis, mon mari, Edgar, se lève. Dès qu'il entre dans la cuisine, il allume la radio parce qu'il aime écouter les actualités. Mais Jacques n'est pas d'accord. Il préfère boire son café et lire le journal dans le calme et la tranquillité. Alors, la dispute quotidienne entre le père et le fils commence.

Au premier étage, on peut entendre ma belle-fille -la femme de Jacques- crier après ses deux enfants, Félix et Jeannette, parce qu'ils vont être en retard s'ils ne se lèvent pas tout de suite.

Ensuite mon second fils Robert, qui est au chômage, sort de sa chambre très en colère car il y a trop de bruit et il ne peut pas faire la grasse matinée. Donc, il prend ses cigarettes et quitte la maison avec notre chien Snoby.

Puis, comme d'habitude, mon petit-fils Félix s'énerve car il ne peut pas se laver les dents : sa sœur a la mauvaise habitude de passer beaucoup de temps dans la salle de bains. Il se demande si elle va à l'école pour draguer les garçons ou pour obtenir son diplôme.

Eh bien moi, pendant ce temps, je prends ma tasse de thé et je m'installe dans le salon pour regarder la télé ; mais, parfois, il y a tellement de bruit dans cette maison que je suis obligée de monter le volume au maximum. Vous savez, à mon âge, on n'entend plus très bien !

1. Qui raconte l'histoire ?

- ☐ la tante
- ☐ la grand-mère
- ☐ la voisine
- ☐ la cousine

2. Chaque matin, Jacques se dispute avec son père car:

- ☐ Jacques n'aime pas les actualités
- ☐ Edgar veut lire son journal
- ☐ Jacques aime prendre son petit-déjeuner dans le calme
- ☐ Edgar aime chanter à la radio

3. La mère crie après ses enfants parce que:

- ☐ Félix et Jeannette se lèvent trop tôt
- ☐ Jeannette se dispute avec son frère
- ☐ ils vont être en retard
- ☐ le petit-déjeuner est prêt

4. Robert est:

- ☐ le cousin de Félix et Jeannette
- ☐ leur père
- ☐ leur oncle
- ☐ le beau-frère d'Edgar

SUJET:

DATE:

5. Robert quitte la maison parce que:

- ☐ il veut acheter le journal
- ☐ il ne peut pas dormir à cause du bruit
- ☐ il doit promener le chien
- ☐ il veut fumer une cigarette

6. A votre avis, quel âge a Jeannette ?

- ☐ 8-10 ans
- ☐ 12-14 ans
- ☐ 17-18 ans
- ☐ 25-30 ans

7. Félix n'est pas content car:

- ☐ il ne sait pas se laver les dents
- ☐ il ne peut pas utiliser la salle de bains
- ☐ sa soeur aime draguer les garçons
- ☐ il va être en retard à l'école

8. La grand-mère s'installe dans le salon pour:

- ☐ boire son café et regarder la télévision
- ☐ boire son thé et jouer à un jeu vidéo
- ☐ boire son thé et regarder la télévision
- ☐ se reposer

9. La grand-mère augmente le volume du son de la télévision car:

- ☐ elle est folle et ne voit plus très bien
- ☐ elle est sourde et il y a trop de bruit
- ☐ elle n'arrive pas à dormir
- ☐ c'est son programme de télévision préféré

10. A votre avis, quels adjectifs correspondent à cette drôle de famille ?

- ☐ nombreuse et calme
- ☐ nombreuse et bruyante
- ☐ joyeuse et organisée
- ☐ triste et sérieuse

SUJET:

DATE:

MOTS CACHÉS: Retrouvez les mots présentés ci-dessous dans la grille et trouvez le mot formé par les lettres restantes de la grille.

AORTE	MENTON	PYLORE
BRONCHES	MOELLE	REIN
CERVELET	NERF	SEIN
CLAVICULE	NEZ	SQUELETTE
DOS	OESOPHAGE	TESTICULE
ESTOMAC	OMOPATE	TETE
FOIE	OREILLE	THALAMUS
INTESTIN	ORTEIL	TIBIA
JOUE	PANCREAS	VERTEBRE
LIGAMENT	POUMON	
MACHOIRE	PROSTATE	

— — — — —

E	G	B	R	O	N	C	H	E	S	E	N	N
T	U	E	E	S	T	O	M	A	C	E	I	I
A	E	O	G	N	A	O	M	L	Z	M	T	E
L	T	T	J	A	E	E	A	U	A	O	S	R
P	E	O	T	L	H	V	R	C	O	U	E	B
O	L	R	L	E	I	P	H	C	M	P	T	E
M	E	E	O	C	L	O	O	A	N	A	N	T
O	V	I	U	L	I	E	L	S	I	A	I	R
R	R	L	O	R	Y	A	U	B	E	U	P	E
T	E	L	E	F	H	P	I	Q	S	O	D	V
E	C	E	E	T	E	T	A	T	S	O	R	P
I	A	O	R	T	E	S	T	I	C	U	L	E
L	I	G	A	M	E	N	T	O	N	E	R	F

SUJET:

DATE:

COMPLETEZ LE TEXTE AVEC LES MOTS CI-DESSOUS.

raconter - table - attraper - cadeaux - chambre - merveilleux - lourd - jeter - souhaiter - paquet - échelle - sauter - bonshommes - triste - papiers - attendre - indice - flocons - train - cuisine

Au soir venu, on s'amusait à \_\_\_\_\_ le plus haut possible pour \_\_\_\_\_ des étoiles. Avec mon frère, sur le mur au bout du jardin, on avait juré un jour d'attraper la Lune, mais chaque année c'était en vain. On se contentait de nos étoiles. Quand le panier était trop \_\_\_\_\_, on le portait à deux jusqu'à la maison et on décorait le sapin avec notre butin. Il arrivait même qu'on en ait trop, alors on ouvrait la fenêtre et on les relâchait dans l'air, elles filaient dans le ciel et ça nous rendait fous, comme les feux d'artifice. Le mieux c'était quand il neigeait, car les \_\_\_\_\_ de neige et les flocons d'étoile se mélangeaient et c'était drôlement beau. Surtout, à la fenêtre, on croquait nos \_\_\_\_\_ de pain d'épice avec délice. Maman nous avait donné l'idée de mettre nos étoiles dans des boules en verre coloré. L'année suivante nous avons donc confectionné des photophores et mis les étoiles dedans : des verts, des bleus, des rouges, partout pendus aux branches du sapin ; c'était autre chose, simplement \_\_\_\_\_.

Cette année, Maman préparait ses biscuits à la cannelle et à la pomme dans la \_\_\_\_\_, Papa finissait d'emballer les cadeaux dans sa \_\_\_\_\_. Mon frère et moi étions au pied du sapin, à jouer avec notre \_\_\_\_\_ en bois. Il y avait notre chien, Pavot, qui dormait sur le tapis en mordillant sa balle de tennis. Je crois que c'est Papa qui avait eu l'idée de faire un trou dans la balle de Pavot pour y mettre une étoile. Le chien devait aimer, parce que quand sa balle redevenait normale à la fin des fêtes, il devenait tout \_\_\_\_\_. Mais on lui expliquait, mon frère et moi, qu'il fallait relâcher les étoiles sinon il n'y en aurait plus pour l'année prochaine. Je crois que Pavot comprenait tout cela, qu'il fallait \_\_\_\_\_ mais que ça valait le coup, vraiment. Cette année, on avait eu des bonbons à l'école le dernier jour avant les vacances, et Maman nous avait permis de les manger n'importe quand, à condition de \_\_\_\_\_ le papier à la poubelle directement pour éviter le retour des rats. L'an dernier, des rats avaient ruiné notre réveillon. C'étaient eux qui avaient d'ailleurs convaincu Maman d'avoir un chien, car les rats avaient peur des aboiements.

Quand on était passé à \_\_\_\_\_, Papa avait versé un petit peu de vin dans nos verres à pied de Grands. C'était la première fois que ça nous arrivait, on avait fait les fiers, mon frère et moi. Et puis on avait eu le même repas que les parents, alors qu'avant Maman nous faisait quelque chose de plus simple et de plus enfantin. C'était un autre \_\_\_\_\_ que les parents trouvaient qu'on était assez grands ! On avait porté des chemises, avec des bretelles et des chaussures comme Papa. Ah, on avait rigolé à table, surtout quand l'Oncle Gérard avait appelé pour nous \_\_\_\_\_ bonne fête. C'était notre oncle préféré, il avait toujours des histoires à nous \_\_\_\_\_, le frère de Papa, et souvent on apprenait les bêtises de notre père par sa bouche. Cette année, nous n'eûmes droit à rien, mais il ne nous déçut pas car il raconta une histoire très drôle de cueillette d'étoiles avec son fils, quand ils avaient planté une très grande \_\_\_\_\_ dans le jardin pour décrocher la Lune. Dans les yeux de mon frère et dans les miens, des paillettes avaient surgi à ce moment. On n'y avait jamais pensé, à l'échelle ! On s'était mis à rire.

Quand le moment des cadeaux était venu, on s'était dépêché d'aller au salon. Il y avait quatre tas, un pour chacun, à côté de nos pantoufles. Cette année, entre le tas de mon frère et le mien, un gros \_\_\_\_\_ était posé. Papa et Maman nous avaient dit que c'était pour nous deux, mais que c'était aussi pour eux ; un cadeau pour toute la famille en quelque sorte. Je débballai mes \_\_\_\_\_, arrachant les \_\_\_\_\_ avec impatience, et l'excitation s'amplifia au fur et à mesure des jouets en bois et des bandes dessinées que je découvrais ; l'excitation de mon frère s'accéléra quand il eut son bateau pirate en bois, celui de ses rêves. On était fous de joie. Quand il ne resta plus que le dernier cadeau, on se regarda tous pour savoir qui devait l'ouvrir. Maman et



SUJET:

DATE:

Papa nous désignèrent en se prenant la main. Chacun d'un côté, nous déballions le papier soigneusement, en prenant soin de ne pas l'abîmer. On devait avoir l'impression que notre premier cadeau commun était précieux, plus gros, plus fragile aussi. On se regardait de temps en temps, en souriant. Et quand mon frère dénoua le lacet qui retenait le tout, le colis enfin s'ouvrit. Dans les yeux de mon frère, à l'excitation avait suivi la stupéfaction. Dans les miens aussi d'ailleurs. Et puis ce fut un feu d'artifice de bonheur : nous avions une petite sœur !

SUJET:

DATE:

CODES A GRILLES: SAURIEZ-VOUS RETROUVER LE MESSAGE ?

### 1. Code des templiers

A	B	C	J	K	L
D	E	F	M	N	O
G	H	I	P	Q	R
S			W		
V		T	Z		X
U			Y		

Pour décoder, on utilise la forme de la case (et éventuellement le point) correspondant à une lettre.

Exemple:



correspond à

«Bonjour»

A VOUS !

- a) = \_\_\_\_\_
- b) = \_\_\_\_\_
- c) = \_\_\_\_\_
- d) = \_\_\_\_\_

SUJET:

DATE:

## 2. Le TicTacToc

• • • A B C	• • • D E F	• • • G H I
• • • J K L	• • • M N O	• • • P Q R
• • • S T U	• • • V W X	• • Y Z

Pour décoder, on utilise la forme de la case et un point. Le point se place à la position de la lettre correspondante dans la case de la bonne forme.

Example:



correspond à

«Bonjour»

A VOUS !

a)  $\cdot \neg \cup \cap \subseteq \subset \setminus \times$  = \_\_\_\_\_

b)  $\begin{array}{|c|c|c|c|c|c|c|c|c|c|} \hline \cdot & \cdot & & \cdot & \cdot & & \cdot & & \cdot & \cdot \\ \hline \end{array} = \underline{\hspace{2cm}}$

[illegible]

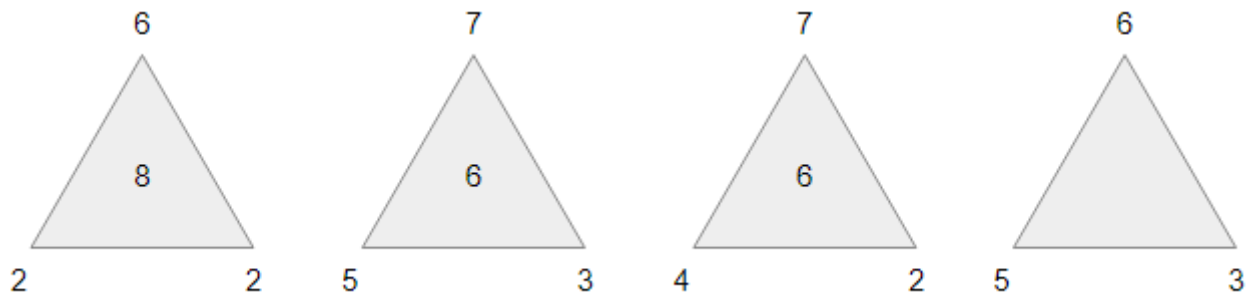
d)  $\square \cdot \square \cdot \neg \cdot \neg \cdot \neg \cdot \square \cdot \neg \cdot \square \cdot \neg$  = \_\_\_\_\_

SUJET:

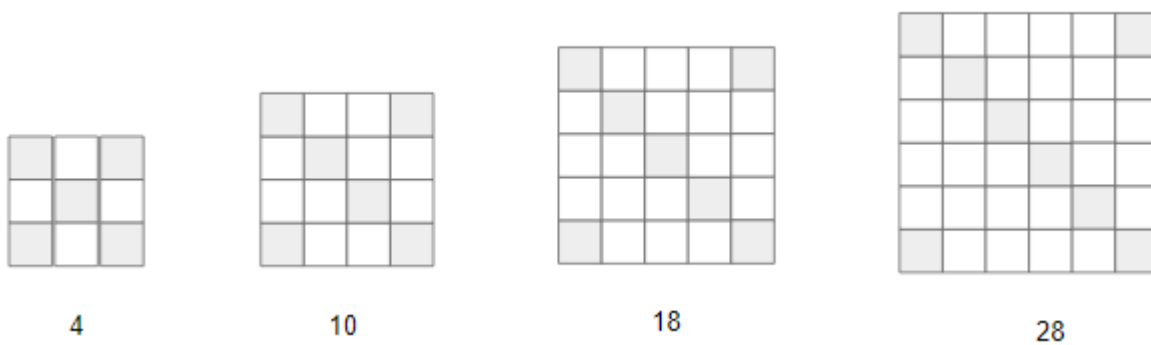
DATE:

COMPLETEZ LES LISTES SUIVANTES.

1. Quel nombre doit être placé dans le triangle vide ? \_\_\_\_



2. Quel nombre vient après ? \_\_\_\_



3. Quel est le nombre manquant ? Entourez la lettre correspondante.

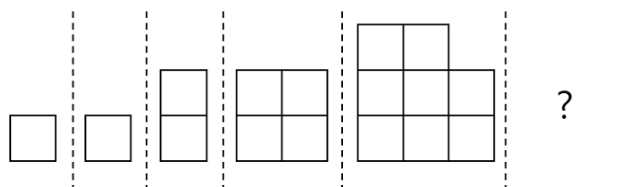
11, 12, 14, \_\_, 21, 26

A. 17      B. 18      C. 19      D. 20

4. Quels sont les nombres manquants ? Complétez les cases.

4	8	3	6	7	14	1	2	5
32	40	18	24	98	112	2	4	

5. Combien de carrés contient le prochain dessin ? Entourez la lettre correspondante.







A. 9      B. 12      C. 16      D. 21

SUJET:





DATE:

6. Quel code correspond à la figure 4 ? Entourez la lettre correspondante.

 9 7	 9 6	 8 7	 ?
1	2	3	4





9	8	8	7
5	5	6	5
A	B	C	D

7. Quel code correspond à la figure 4 ? Entourez la lettre correspondante.

 T U	 T W	 X U	 ?
1	2	3	4

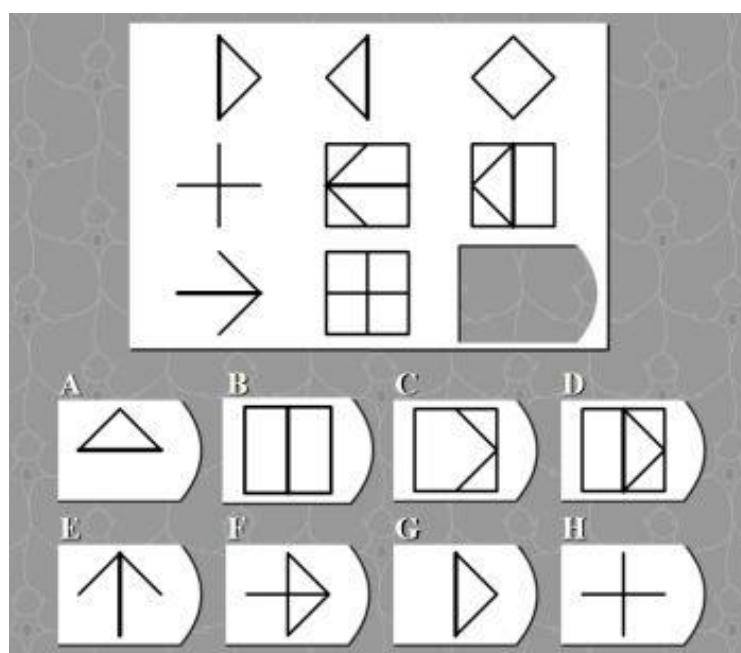
X	X	T	U
T	W	X	W
A	B	C	D

8. Quel code correspond à la figure 4 ? Entourez la lettre correspondante.

 F G	 H J	 H G	 ?
1	2	3	4

J	J	F	F
H	G	J	H
A	B	C	D

9. Quelle figure complète la série ? Entourez la lettre correspondante.



SUJET:

DATE:

SAURIEZ-VOUS RETROUVER LE MESSAGE ?

nous,  
???,  
ils

k'

il,  
elle,  
??



a

M

il,  
elle,  
??

t'



v'



2



2



2



t r'



il,  
elle,  
??

### A.3 List of common physical responses during motion sickness

## **DESCRIPTION DU RESSENTI**

MAL DE TETE

CHANGEMENT DE LA TEMPERATURE DU CORPS

- FRISSONS
- BOUFFEES DE CHALEUR
- ...

SOMNOLENCE

ETOURDISSEMENTS

SENSATIONS AU NIVEAU DE LA BOUCHE

- SECHERESSE
- SALIVATION
- ...

NAUSEES

DIFFICULTE DE CONCENTRATION

IRRITABILITE

FATIGUE DES YEUX



# Appendix B

## Data analysis

### B.1 Cardiac activity

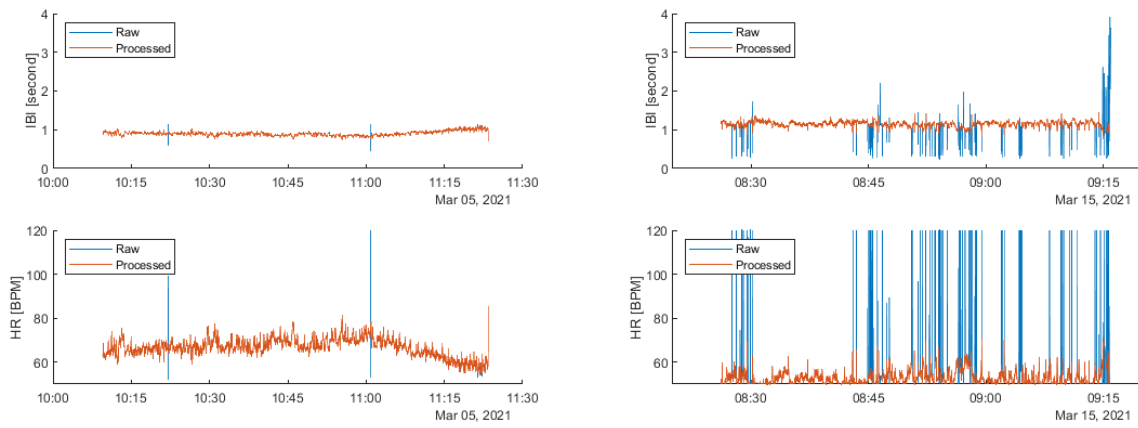


Figure B.1: Raw and processed IBI and deduced HR for participant 7 (left) and 16 (right) during simulator sessions.

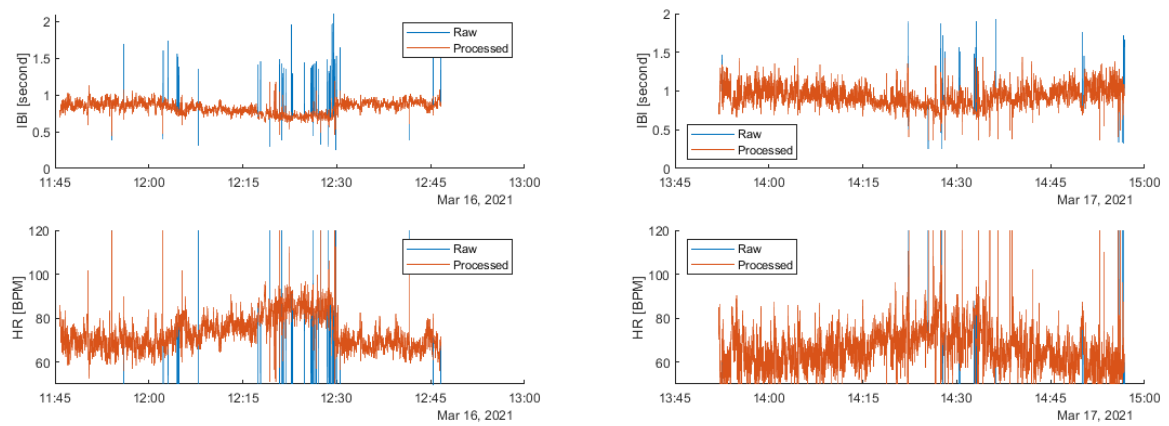


Figure B.2: Raw and processed IBI and deduced HR for participant 3 (left) and 8 (right) during car sessions.

## B.2 Electrodermal activity

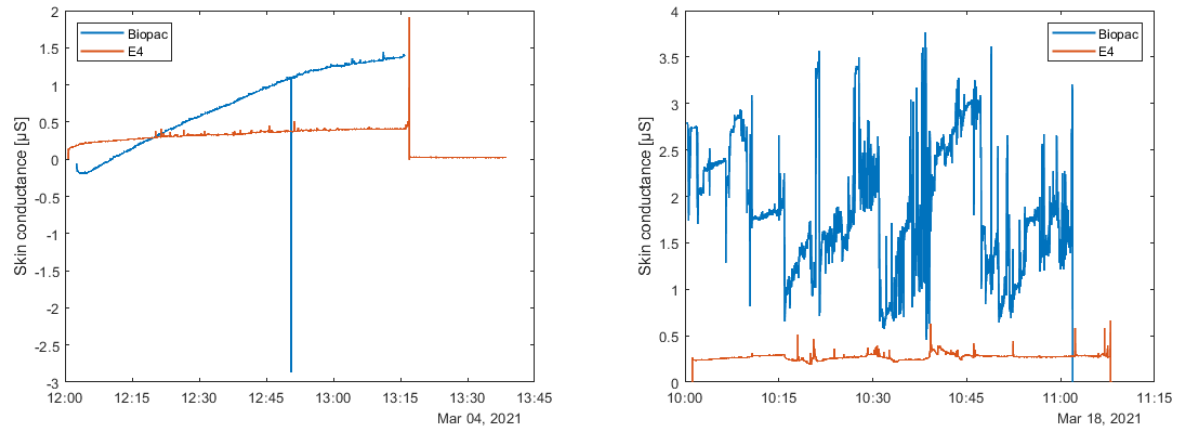


Figure B.3: EDA signals of one non-responder, for both simulator (left) and car (right) sessions.

## Appendix C

### Pseudo-code of the considered method for training machine learning models

```
1 % for all test datasets
2
3 % participants_test definition (5)
4 % participants_training and validation (15)
5
6 % for hyperparameters to tune in the considered ML model
7
8     % for all training and validation datasets
9
10         % participants_val definition (3)
11         % participants_train definition (12)
12
13         % Features selection on the training dataset
14
15         % Training classifier on the training dataset
16
17         % Evaluation on the validation dataset
18         % Calculation of the evaluation metric
19
20         % Calculation of the mean of the evaluation metric
21         % if mean of the evaluation metric > max of the evaluation metric
22             % max of the evaluation metric = mean of the evaluation metric
23             % associated values of hyperparameters
24
25     % Features for training on the training and validation sets = union of
    the 5 sets of selected features
26
27     % Training of the model on training and validation datasets, with the
    best hyperparameters
28
29     % Evaluation on the test dataset
30     % Calculation of the evaluation metric on the test dataset
31
32
33 % Calculation of the average of the maximum of the evaluation metric on the
    validation datasets
34
35
36 % Calculation of the average of the evaluation metric on the test datasets
```

# Bibliography

- [1] National Library of Medicine. *Motion sickness*. URL: <https://medlineplus.gov/genetics/condition/motion-sickness/#frequency>. Last accessed on May 24, 2021.
- [2] B.E. Donohew; M.J Griffin. “Motion Sickness: Effect of the Frequency of Lateral Oscillation.” In: *Aviation, Space, and Environmental Medicine* 75(8) (2004), pp. 649–656.
- [3] C.A. Butler; M.J Griffin. “Motion Sickness during Fore-and-Aft Oscillation: Effect of the Visual Scene.” In: *Aviation, Space, and Environmental Medicine* 77(12) (2006), pp. 1236–1243.
- [4] J.A. Joseph; M.J Griffin. “Motion Sickness: Effect of the Magnitude of Roll and Pitch Oscillation.” In: *Aviation, Space, and Environmental Medicine* 79(4) (2008), pp. 390–396.
- [5] G.F. Beard; M.J Griffin. “Discomfort Caused by Low-Frequency Lateral Oscillation, Roll Oscillation and Roll-Compensated Lateral Oscillation.” In: *Ergonomics* 56(1) (2013), pp. 103–114.
- [6] G. Bertolini; D. Straumann. “Moving in a Moving World: A Review on Vestibular Motion Sickness”. In: *Front. Neurol* (2016).
- [7] Eike. A. Schmidt et al. “An international survey on the incidence and modulating factors of carsickness.” In: *Transportation Research Part F: Traffic Psychology and Behaviour* 71 (2016), pp. 76–87.
- [8] J.T. Reason and J.J. Brand. *Motion Sickness*. New York: Academic press, 1975.
- [9] A.J. Benson. *Motion Sickness*. in *Aviation medicine*, J. Ernsting and P. King, Editors. Butterworths: London, 1988.
- [10] J. Iskander et al. “From car sickness to autonomous car sickness: A review”. In: *Transportation Research Part F: Traffic Psychology and Behaviour* 62 (2019), pp. 716–726.
- [11] M. Treisman. “Motion sickness: An evolutionary hypothesis”. In: *Science* 197(4302) (1977), pp. 493–495.
- [12] G.E. Riccio; T.A. Stoffregen. “An ecological theory of motion sickness and postural instability”. In: *Ecological psychology* 3(3) (1991), pp. 195–240.
- [13] J.E. Bos; W. Bles; E. Groen. “A theory on visually induced motion sickness”. In: *Displays* 29 (2008), pp. 47–57.

- 
- [14] John F. Golding. “Motion sickness susceptibility”. In: *Displays* 129 (2006), pp. 67–76.
  - [15] John F. Golding; A.G. Mueller; M.A. Gresty. “A motion sickness maximum around the 0.2 Hz frequency range of horizontal translational oscillation”. In: *Aviation Space and Environmental Medicine* 72(3) (2001), pp. 188–192.
  - [16] Monica L.H. Jones. “Queasy passengers: A testbed for motion sickness in driverless vehicles.” In: *Mcity - University of Michigan* (2019).
  - [17] N. Isu et al. “Quantitative analysis of time-course development of motion sickness caused by in-vehicle video watching”. In: *Displays* 35 (2014), pp. 90–97.
  - [18] *Optokinetic drum*. URL: [https://en.wikipedia.org/wiki/Optokinetic\\_drum](https://en.wikipedia.org/wiki/Optokinetic_drum). Last accessed on April 10, 2021.
  - [19] *Sympathetic nervous system*. URL: [https://en.wikipedia.org/wiki/Sympathetic\\_nervous\\_system](https://en.wikipedia.org/wiki/Sympathetic_nervous_system). Last accessed on April 10, 2021.
  - [20] *ECG typical aspect*. URL: <https://thephysiologist.org/study-materials/the-normal-ecg/>. Last accessed on April 10, 2021.
  - [21] M. Elgendi. “On the Analysis of Fingertip Photoplethysmogram Signals”. In: *Current Cardiology Reviews* 8(1) (2012), pp. 14–25.
  - [22] K. Li; H. Rüdiger; T. Ziemssen. “Spectral Analysis of Heart Rate Variability: Time Window Matters”. In: *Front. Neurol.* (2019).
  - [23] P.S. Cowings et al. “General autonomic components of motion sickness”. In: *Psychophysiology* 23(5) (1986), pp. 542–551.
  - [24] *Interbeat interval*. URL: [https://en.wikipedia.org/wiki/Interbeat\\_interval](https://en.wikipedia.org/wiki/Interbeat_interval). Last accessed on May 3, 2021.
  - [25] *Electrodermal activity*. URL: [https://en.wikipedia.org/wiki/Electrodermal\\_activity](https://en.wikipedia.org/wiki/Electrodermal_activity). Last accessed on May 3, 2021.
  - [26] J. Bakker; M. Pechenizkiy; N. Sidorova. “What’s your current stress level? Detection of stress patterns from GSR sensor data.” In: *IEEE International Conference on Data Mining* (2011), pp. 573–580.
  - [27] M. Bradley; P. Lang. “Measuring emotion: Behavior, feeling and physiology”. In: *In Cognitive Neuroscience of Emotion, R. Lane and L. Nadel Eds. New York, USA: Oxford University Press* (2000), pp. 242–276.
  - [28] *Skin Temperature*. URL: [https://en.wikipedia.org/wiki/Skin\\_temperature](https://en.wikipedia.org/wiki/Skin_temperature). Last accessed on April 10, 2021.
  - [29] D. Mühlbacher et al. “Methodological considerations concerning motion sickness investigations during automated driving”. In: *Information* 11 (2020), p. 265.
  - [30] G. Nobel et al. “Motion sickness increases the risk of accidental hypothermia”. In: *European Journal of Applied Physiology* 98 (2006), pp. 48–55.
  - [31] P. Kolh. *GBIO0026-1 Physiologie des systèmes*. 2017-2018.
  - [32] J.C. Miller et al. “Autonomic Physiological Data Associated with Simulator Discomfort”. In: *Aviation, Space and Environmental Medicine* 64(9) (1993), pp. 813–819.

- 
- [33] P.S. Cowings et al. “General autonomic components of motion sickness”. In: *Psychophysiology* 23(5) (1986), pp. 542–551.
  - [34] C.S. Stout; W.B. Toscano; P.S. Cowings. “Reliability of psychophysiological responses across multiple motion sickness stimulation tests”. In: *Journal of Vestibular Research* 5(1) (1995), pp. 25–33.
  - [35] M.J. Griffin; M.M. Newman. “Visual field effects on motion sickness in cars”. In: *Aviation, Space and Environmental Medicine* 75(9) (2004), pp. 739–747.
  - [36] K. Guenther. *Brain Lobes*. URL: <https://brainframe-kids.com/brain/facts-lobes.htm>. Last accessed on April 12, 2021.
  - [37] C. Lin; S. Tsai; L. Ko. “EEG-Based Learning System for Online Motion Sickness Level Estimation in a Dynamic Vehicle Environment”. In: *IEEE* 24 (2013), pp. 1689–1700.
  - [38] MedicineNet. *Electrogastrogram (EGG)*. URL: [https://www.medicinenet.com/electrogastrogram/article.htm#what\\_is\\_an\\_electrogastrogram](https://www.medicinenet.com/electrogastrogram/article.htm#what_is_an_electrogastrogram). Last accessed on May 31, 2021.
  - [39] J. Yin; J.D. Chen. “Electrogastrography: Methodology, Validation and Applications”. In: *J Neurogastroenterol Motil* 19(1) (2013), pp. 5–17.
  - [40] R.M. Stern et al. “Spectral analysis of tachygastria during motion sickness”. In: *Gastroenterology* 92 (1987), pp. 92–97.
  - [41] M.J. Griffin; H.V.C. Howarth. “Motion Sickness History Questionnaire”. In: *ISVR Technical Report* 283 (2000).
  - [42] M.L. Van Emmerik et al. “Internal and external fields of view affect cybersickness”. In: *Displays* 32(4) (2011), p. 169.
  - [43] G. Borg; E. Borg. “A New Generation of Scaling Methods: Level-anchored Ratio Scaling”. In: *Psychologica* 28 (2001), pp. 15–45.
  - [44] *Phasya website*. URL: <https://www.phasya.com/en/>. Last accessed on April 20, 2021.
  - [45] Inc. BIOPAC Systems. *About BIOPAC*. URL: <https://www.biopac.com/corporate/about-biopac/>. Last accessed on April 20, 2021.
  - [46] insidescientific. *BIOPAC Systems, Inc.* URL: <https://insidescientific.com/suppliers/biopac-systems-inc/>. Last accessed on April 20, 2021.
  - [47] Inc. BIOPAC Systems. *MP160 data acquisition analysis system*. URL: <https://www.biopac.com/product/mp150-data-acquisition-systems/>. Last accessed on April 20, 2021.
  - [48] Inc. BIOPAC Systems. *BioNomadix Wireless RSP with ECG Amplifier*. URL: <https://www.biopac.com/product/bionomadix-rsp-with-ecg-amplifier/>. Last accessed on April 20, 2021.
  - [49] Inc. BIOPAC Systems. *Electrode Leads – BioNomadix Wireless*. URL: <https://www.biopac.com/product/electrode-leads-bionomadix/>. Last accessed on April 20, 2021.

- 
- [50] Inc. BIOPAC Systems. *Cloth Base Electrodes*. URL: <https://www.biopac.com/product/cloth-base-electrodes/>. Last accessed on April 20, 2021.
  - [51] Inc. BIOPAC Systems. *B-ALERT WIRELESS EEG SYSTEM*. URL: <https://www.biopac.com/product/b-alert-wireless-eeeg-system/>. Last accessed on April 20, 2021.
  - [52] Inc. BIOPAC Systems. *BioNomadix Wireless PPG and EDA Amplifier*. URL: <https://www.biopac.com/product/bionomadix-ppg-and-eda-amplifier/>. Last accessed on April 20, 2021.
  - [53] Inc. BIOPAC Systems. *EGG Smart Amplifier*. URL: <https://www.biopac.com/product/egg-smart-amplifier/>. Last accessed on April 20, 2021.
  - [54] Empatica. *E4 wristband*. URL: <https://www.empatica.com/en-eu/research/e4/>. Last accessed on April 20, 2021.
  - [55] Inc. BIOPAC Systems. *233 – Heart Rate Variability – Preparing Data for Analysis*. URL: <https://www.biopac.com/application-note/heart-rate-variability-preparing-data-for-analysis/>. Last accessed on April 26, 2021.
  - [56] F. Shaffer; J.P. Ginsberg. “An Overview of Heart Rate Variability Metrics and Norms”. In: *Front Public Health* 5:258 (2017).
  - [57] iMotions. *Galvanic Skin Response - The Complete Pocket Guide*. URL: <https://imotions.com/guides/eda-gsr/>. Last accessed on April 27, 2021.
  - [58] Inc. BIOPAC Systems. *Negative EDA (GSR)*. URL: <https://www.biopac.com/knowledge-base/negative-eda-gsr/>. Last accessed on April 28, 2021.
  - [59] M. Benedek; C. Kaernbach. “A continuous measure of phasic electrodermal activity”. In: *Journal of Neuroscience Methods* 190 (2010), pp. 80–91.
  - [60] J. Chen; R.W. McCallum. “Electrogastrography: measurement, analysis and prospective applications.” In: *Med. Biol. Eng. Comput.* 29(4) (1991), pp. 339–350.
  - [61] K.L. Koch. *Handbook of electrogastronomy*. Oxford University Press, 2004.
  - [62] Dr Jason J. Braithwaite et al. *A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments*. URL: <https://www.birmingham.ac.uk/Documents/college-les/psych/saal/guide-electrodermal-activity.pdf>. Last accessed on April 10, 2021.
  - [63] T. Wada. *Motion sickness in automated vehicles*. 2016.
  - [64] L. Wehenkel; P. Geurts. *ELEN062-1 Introduction to Machine Learning*. 2020-2021.
  - [65] MathWorks. *fitcensemble*. URL: <https://nl.mathworks.com/help/stats/fitcensemble.html>. Last accessed on May 18, 2021.
  - [66] MathWorks. *fscmrnr*. URL: <https://nl.mathworks.com/help/stats/fscmrnr.html>. Last accessed on May 14, 2021.
  - [67] A. Fernandez et al. *Learning from Imbalanced Data Sets*. Springer, 2018.
  - [68] X. Yang et al. “A weighted support vector machine for data classification”. In: *International Journal of Pattern Recognition and Artificial Intelligence* 21.05 (2007), pp. 961–976.



- [69] T. Saito; M. Rehmsmeier. “The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets”. In: *PLoS ONE* 10(3) (2015).