修士論文

生理学的データを用いた VR 酔いの発生の予測

公立はこだて未来大学大学院　システム情報科学研究科
情報アーキテクチャ領域

曲木 拓朗

指導教員　　ヴァランス マイケル

提出日　　2020 年 3 月 9 日

Master's Thesis

Forecasting the Onset of Cybersickness using Physiological Data

by

Takurou Magaki

Graduate School of Systems Information Science, Future University Hakodate
Media Architecture
Supervisor:　　Dr. Michael Vallance

Submitted on March 9, 2020

Abstract–

Cybersickness is a critical problem when immersed in Virtual Reality (VR) experiences. However, cybersickness has not been resolved because the composition of cybersickness is complicated and its mechanism is not completely understood. In addition, cybersickness has been commonly evaluated using a qualitative, subjective survey called the Simulator Sickness Questionnaire (SSQ) conducted after the VR experience. More recently, studies estimating cybersickness using physiological data as an objective evaluation have been undertaken. However, cybersickness as a dynamically changing condition during the VR experience has not been considered. In this research, the hypothesis is defined as "Cybersickness can be estimated from physiological data captured in real-time during a VR experience." Based on the hypothesis, this research is undertaken as follows. First, a reliable instrument that can capture physiological data was selected; in this research the Empatica E4. Then, validation of physiological data analysis was tested and applied in PC and VR activities called PolyWorld. The metrics of cybersickness were considered from the resulting physiological data. Then an indicator of cybersickness based on SSQ, called Onset of Cybersickness (OCS) determined by the physiological data, was proposed. Moreover, an Android / iOS application named Cybatica, which is a real-time forecasting and visualizing cybersickness application used as a case study of OCS, was developed and tested. In the evaluation, OCS scores were analyzed by correlation with SSQ scores. As a result, the research demonstrated an approach for forecasting cybersickness using physiological data. However, the resulting evaluation was inconclusive and additional research is proposed.

Keywords: application, cybersickness, physiological data, virtual reality

概 要：

バーチャルリアリティ (Virtual Reality, VR) の体験に没入する際に発症する VR 酔い (Cybersickness) は，VR に関するコンテンツにおいて無視できない重大な問題である．しかし，VR 酔いの構成は複雑であり，そのメカニズムは完全には理解されていないため，VR 酔いの問題は未だ完全には解決されていない．VR 酔いは，主に Simulator Sickness Questionnaire (SSQ) と呼ばれる定性的で主観的なアンケートを用いて，VR 体験後に評価される．より最近の研究では，客観的な評価として生理学的データを用いて VR 酔いを推定する研究が行われている．しかし，VR 体験中において状態が動的に変化する VR 酔いを考慮することはできていない．本研究では，仮説として「VR 酔いは VR 体験中にリアルタイムで計測された生理学的データから推定できる」と定義し，仮説に基づいて研究を行った．はじめに，生理学的データを計測できる信頼性の高い機器の考察を行い，Empatica E4 と呼ばれるリストバンド型のデバイスを選択した．次に，生理学的データを用いた VR 酔いの分析，および検証を行い，PolyWolrd と呼ばれる PC および VR アクティビティにおいて実験を行った．そして，結果として得られた生理学的データから，VR 酔いを構成する生理学的指標，および測定基準を考察した．そして，生理学的データによって計算される，Onset of Cybersickness (OCS) と呼ばれる SSQ の各スコアに基づく VR 酔いの指標を提案した．さらに，OCS の活用方法として，OCS を用いた VR 酔いのリアルタイムにおける予測および可視化をする Cybatica と呼ばれる Android および iOS アプリケーションの開発および検証を行った．評価では，SSQ のスコアとの相関によって OCS のスコアを分析した．その結果，本研究では生理学的データを用いた VR 酔いを予測するアプローチを示すことができた．しかし，VR 酔いを予測するには決定的ではなく，追加の研究の必要性を示した．

キーワード： アプリケーション, VR 酔い, 生理学的データ, バーチャルリアリティ

Contents

Chapter 1 Introduction

In this chapter, the background of Virtual Reality technologies and cybersickness will first be introduced. Then, the goal of the research will be highlighted. After that, the organization of the thesis will be summarized.

1.1 Background

Virtual Reality (VR) technologies have been used to solve real-world problems in such domains as military, medical, training, education, commerce, and entertainment. An advantage of VR is its capacity to provide simulated spaces of content and context which would be difficult or impossible in the real world. For example, companies have developed a VR surgeon simulator that can train surgeons to perform surgery[*1] and a VR chemical experiment simulator which can support the learning of chemistry[*2]. These scenarios are difficult to reproduce in the real world as there are risks of failure involved and can be expensive. VR can thus be utilized without the risks. The first VR system called Sensorama, which had a multi-modal interface, was invented by Morton Heilig in the early 1960's. The system is an immersive system that stimulated vision and hearing, and is a pioneer of immersive systems. However, the term 'Virtual Reality' had not been generally considered until used by Jaron Lanier in 1989 when a communication system using a Head Mounted Display (HMD) and Data Glove was introduced. Recently, VR technologies have proliferated as HMD such as Oculus Rift[*3], HTC VIVE[*4], and Windows MR Headsets[*5] have arrived on the market. Also, VR HMD' s have been manufactured with low prices that do not require high-performance computers; examples include the Oculus Go and the Nintendo Labo VR Kit[*6]. Consequently, VR technology that was expensive and only used in limited research has now become a more accessible technology for research, education and consumers. The improvement of peripheral technologies such as graphics and computing has also increased adoption. However, applying VR to a wide target audience such as mainstream education is difficult at present because VR has a health issue that affects each individual differently. The issue is called cybersickness (*aka.* VR sickness). The physiological mechanism of cybersickness is not yet concluded despite numerous projects and research conducted over an extended pe-

[*1] Oxford Medical Simulation (https://oxfordmedicalsimulation.com/)
[*2] Vollege (https://sites.google.com/veyond.tw/vollege)
[*3] Oculus Rift(https://www.oculus.com/)
[*4] HTC VIVE (https://www.vive.com/)
[*5] Windows MR Headset (https://www.microsoft.com/en-us/windows/windows-mixed-reality)
[*6] Nintendo Labo VR kit (https://labo.nintendo.com/kits/vr-kit/)

riod of time because cybersickness is affected by individual differences and thus it is difficult to make a consistently reliable model from quantitative evaluations of captured physiological data.

1.2 Goals and Approaches

To overcome these challenges, in this research a forecastable indicator of cybersickness using physiological data as an Onset of Cybersickness (OCS) is proposed, and a monitoring method of cybersickness in real-time on a native application called Cybatica is developed.

1.3 Organization of the Thesis

The thesis is organized by the following four chapters. First, Chapter 1 introduces the background of Virtual Reality technologies and its problems, the goal of the research, and the organization of the thesis. Then, Chapter 2 will provide a systematic literature review that discusses approaches for cybersickness detection such as mechanism, reducing method, and evaluation method. Next, in Chapter 3, the research methodology will be detailed. The hypothesis, the research method, implementation, and experiments will be explained. Moreover, the data of the experiments will be analyzed and displayed. Then, in Chapter 4, observations and outcomes are discussed from the results and analysis. Finally, Chapter 5 will be concluded with a summary of this research, its limitations, and suggestions for future research.

Chapter 2 Literature Review

In this chapter, the literature of related studies, prerequisite knowledge, and applications in this research field are summarized. The topics covered are cybersickness, reduction method of cybersickness, measuring method of cybersickness, and physiological data.

2.1 Cybersickness

Overcoming cybersickness has been problematic for developers of VR since its inception and remains an obstacle for widespread adoption of VR in such scenarios as mainstream education and industrial training. Symptoms such as discomfort, headache, eye strain, and dizziness during and after VR experiences can be associated with cybersickness. Cybersickness is often compared with motion sickness and simulator sickness. However, all are categorized slightly differently because the above-mentioned symptoms occur in a different manner [1]. Cybersickness is categorized as a subset of motion sickness because cybersickness is considered a form of visually induced motion sickness (VIMS) [2]. Also, simulator sickness is the same form but is almost only applied to vehicle simulators. Moreover, Stanney et al. showed cybersickness is three times more serious compared with simulator sickness [1][3].

2.2 Factors of Cybersickness

One commonly accepted theory is the sensory conflict theory [1][4] where it is proposed that cybersickness occurs due to a mismatching of information between the vision and the vestibular systems. This includes a related phenomenon called vection in which a person feels in motion but is actually stationary. One common example is sitting on a stationary train and then observing a neighboring train moving slowly from the station. There is a temporary sensation of moving, even though still stationary. Vection occurs due to the close association between the processing of visual and vestibular motion. Seno and Suzuki showed that the degree of cybersickness is increased when the intensity of vection increases [5]. There are other theories such as poison theory [6] and postural instability theory [7] when unnatural posture occurs during immersive experiences; however, the occurrence of this condition is limited.

Although there are commonly accepted symptoms and theories of cybersickness, there is still no consensus on how to solve this ongoing problem. Davis et al. state, “the issue is complicated as experiences of cybersickness vary greatly between individuals, the technologies being used, the design of the environment and the tasks being performed” [1].

The individual factors of age, gender, health condition, and posture are also considered. Sensitivity affects to cybersickness is strongly considered as an individual factor. People between the ages of 2 to 12 are said to have high sensitivity, and sensitivity decreases with increasing age. Gender also affects cybersickness in that females are more likely to be to sick than males because female hormones affect sensitivity. Also, sensitivity is increased by taking medicines and alcohol. In addition, posture relates to the postural instability theory. Regarding technology factors, a display's lag, flicker, calibration, and ergonomics are considered. The display's lag relates to the sensory conflict theory because deviation occurs between vision through displays and an actual sense that is based on real life. Flicker causes eye strains. In particular, humans are sensitive to the flicker in a peripheral field of view. As for ergonomics, physical discomfort occurs by the weight of HMD and the calibration of the interpupillary distance. As for environments and tasks, if motion is predictable when operating in the VR activity then there is less possibility of cybersickness. On the other hand, the occurrence of cybersickness is more likely when motion or situations in the VR activity are unpredictable; for example when a viewpoint is forced to change but ignores the user's head movement. The confusion relates to the sensory conflict theory. Another example is a passenger being more likely to be to sick in a car than the driver because the next movements cannot be predicted. This leads to motion sickness. Finally, the degree of cybersickness increases with the immersion time in a VR activity.

2.3 Reduction Method of Cybersickness

There is research that considers the reduction method of cybersickness. For example, Fernandes et al. showed that cybersickness is reduced by modifying the field of view (FoV) range when moving in a VR activity (see Fig. 2.1) [8]. The activity was walking around a VR environment. The Simulator Sickness Questionnaire was used as an evaluation instrument in the experiment.

Also, Whittinghill et al. tested displaying a virtual nose to measure the time that participants could endure in a VR activity (see Fig. 2.2) [9]. The two activities, that are walking around and a roller coaster game, were carried out. As a result, the engaged time in VR activity was longer with a virtual nose than without the virtual nose. The experiment captured and evaluated electrodermal activity (EDA) to measure the physiological effects of the VR activities.

Tambovtsev et al. [10] suggested including a 'frame of reference' such as an horizon or a control panel or dashboard in front of a user' s vision. This method removes the confusion of unpredictable movement due to the placement of an object in a stable

position from the user's vision.

Furthermore, a reduction method using haptic devices is not just a physical expression in the VR activity. There are head-worn haptic devices such as the PhantomLegs, which is proposed by Liu et al. [11]. The device creates a haptic-vestibular illusion by simulating a user's head, which is synchronized to the user's footsteps in VR (see Fig. 2.3). An experiment was carried out to walk (operated moving by a controller) around with a checkpoint in VR by 30 participants. In the experiment, a control, dynamic-FoV applied, and haptic-device-assisted condition was compared and evaluated using the Simulator Sickness Questionnaire. As a result, the discomfort was reduced in the haptic-device assisted condition than in the control condition and the dynamic-FoV condition.

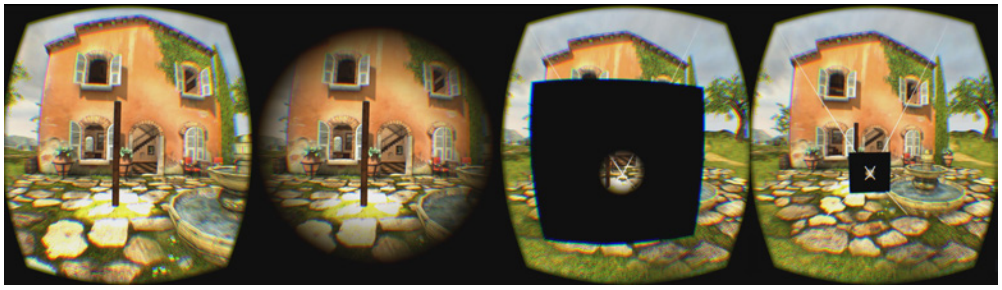


Fig. 2.1: Combating VR sickness through subtle dynamic field-of-view modification [8]

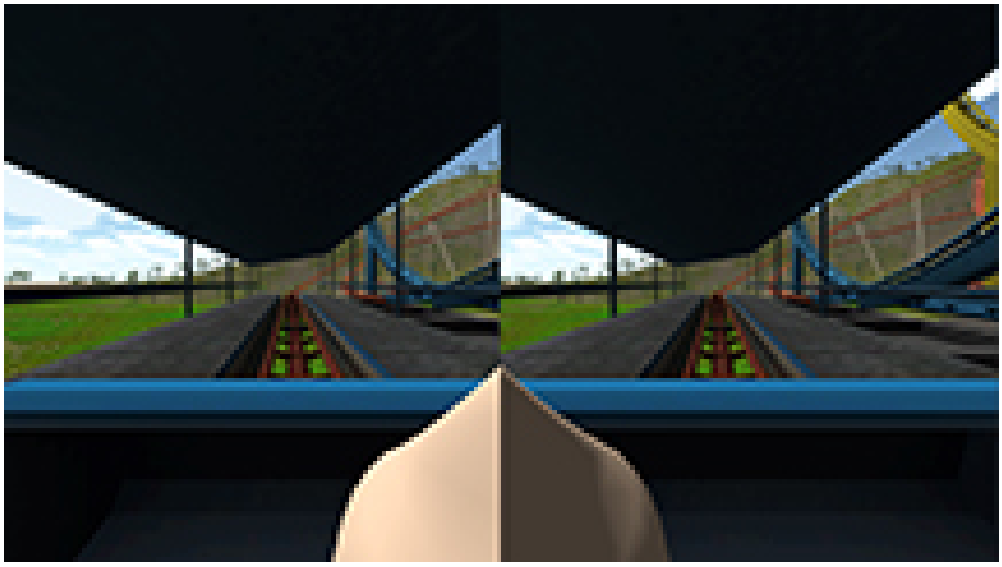


Fig. 2.2: Nasum virtualis: A simple technique for reducing simulator sickness [9]

Fig. 2.3: PhantomLegs: Reducing Virtual Reality Sickness Using Head-Worn Haptic Devices [11]

2.4 Measures of Cybersickness

Methods of measuring cybersickness are similarly inconclusive. Currently, a commonly used method of measuring cybersickness is the Simulator Sickness Questionnaire (SSQ) developed by Kennedy et al. [13]. The Questionnaire calculates the degree of simulator sickness from the sum of 16 symptoms and the symptoms are weighted 0 to 3. In addition, the symptoms are classified from three subclasses, which are nausea, oculomotor, and disorientation. However, even though the SSQ is considered the default instrument for measuring sickness, this method is difficult to use as reliable data because it is self-reported data and captured 'after' the VR experience. SSQ cannot be effectively undertaken during a VR activity as it may affect presence in VR. Table 2.1 shows SSQ symptom items and which subclass they belong.

The Virtual Reality Symptom Questionnaire specializes in Cybersickness [14]. The questionnaire collects data of 13 symptoms that have been filtered from common experiences derived from 47 symptoms and then classified into two subclasses which are Non-ocular and Ocular Symptoms. However, it is not widely used because of insufficient verification.

McHugh et al. proposed a measurement method that uses a physical dial and Fast Motion Sickness (FMS) as a subjective method in real-time [15]. FMS has also been proposed by Keshavarz et al. [16], which scaled motion sickness between 0 to 20, and participants controlled the FMS scale using a physical dial. As a result, the physical dial scale showed a correlation between the SSQ data and was considered easy to capture the

Table 2.1: Computation of Simulator Sickness Questionnaire. x is weight of symptoms (0 to 3)

Parameter	**Nausea**	**Oculomotor**	**Disorientation**
General Discomfort	x	x	
Fatigue		x	
Headache		x	
Eye strain		x	
Difficulty focusing		x	x
Increased salivation	x		
Sweating	x		
Nausea	x		x
Difficulty concentrating	x	x	
Fullness of head			x
Blurred vision		x	x
Dizzy (eyes open)			x
Dizzy (eyes closed)			x
Vertigo			x
Stomach awareness	x		
Burping	x		
Total	a	b	c
$NauseaScore = a \times 9.54$			
$OculomotorScore = b \times 7.58$			
$DisorientationScore = c \times 13.92$			
$TotalScore = (a + b + c) \times 3.74$			

scale of cybersickness. However, this method is subjective and showing the difference between FMS scale and the physical dial scale remains doubtful.

2.5 Physiological Data

More recently, physiological data has been captured to determine cybersickness. Dennison et al. estimated the occurrence of cybersickness from physiological signals such as electrocardiogram (ECG), electrogastrogram, electro-oculogram, photoplethysmogram, breathing rate, and galvanic skin response (GSR) as quantitative evaluations [17]. The physiological data might be able to indicate a dynamically changing cybersickness. However, the degree of cybersickness was not explicitly shown.

Garcia-Agundez et al. tried to detect cybersickness using Heart Rate Variability parameters [18]. The research carried out an experiment that is evaluated by the correlation between ECG parameters and SSQ data through a shooting game as a VR activity. The analysis concluded NN Mean and SDNN are simpler to measure cybersickness. As

a limitation though, there were only 13 experimental participants, and 4 participants interrupted the experiment procedure.

Sun et al. reported the potential of a mental stress classification by gathering ECG, GSR, and accelerometer data from 20 participants across three activities: sitting, standing, and walking [19]. This paper concluded the following. In ECG analysis, mean HR and RR are the most reliable features to recognize mental stress across three physical activities. And, total duration, total magnitude, the total occurrence of the responses, and mean GSR level illustrate an obvious increase from baseline to a stressed segment. However, the standard deviation did not provide a significant change between conditions. As mentioned, the research measured stress and not cybersickness. However, physiological symptoms may be similar.

Care should be taken when analyzing participants with physiological data. Physiological data from participants are irreproducible, even statistically. A living being such as a person can never exactly be set again to its prior state. This change in being and the inevitable change in the physical environment in which data is collected between successive measures leads to limitations pertinent to all experiments. In addition, due to the transmission of measuring signals from source to the device, noise will also inevitably incur. For example, imperfect electric conduction between skin and electrodes makes data acquisition challenging [20]. This limitation was also discovered in a previous study; see below.

For this reason, there is no single physiological measure to capture cybersickness, no established method to determine the severity of cybersickness, and the resultant analysis is often inaccessible to VR users who are not experts in studies in physiological and statistical analysis. It is, therefore, reasoned that due to such perceived complexity, mainstream adoption of VR has not been as ubiquitous as previously predicted.

2.6 Previous Study

In order to alleviate cybersickness in VR and attempt to reduce a user' s anxiety while in VR, specific solutions were identified in the previous graduation study research [12]. The first was the selection and reproduction of the reducing methods of cybersickness. Experiments controlling the FoV [8], and showing the virtual nose [9], were undertaken. One Virtual Environment with two reduction methods was created in order to judge whether the effect of reduction methods differ depending on the information volume within VR. Secondly, the preparation of a biosensor for measuring cybersickness through EDA and

ECG physiological data was undertaken using a custom-made Bitalino device [*1]. The research measured physiological responses of elevation of heart rate and perspiration, which are considered responses caused by confusion, physiological disturbance and being in an anxious state while active in VR. Different conditions of reducing methods were employed. By using the recognized and validated SSQ, undertaken by participants after engaging in the experiment tasks, the data was evaluated to reveal that reducing the FoV had a statistically significant effect in reducing cybersickness. However, it was acknowledged that the study was limited by the number of participants and the unreliable physiological instrument. This prior experiment by the author revealed the complexity of determining cybersickness outside the expertise of medical researchers and was thus a motivation to continue the research.

[*1] Bitalino: https://bitalino.com

Chapter 3 Method

The chapter that is informed by the research hypothesis is constructed as follows. Firstly, a valid instrument that can measure physiological data in the research is selected. Secondly, PC and VR applications as experimental material and experimental tasks are developed and considered. Then, experiment 1 is explained and obtained data from experiment 1 is analyzed. Next, a real-time forecastable cybersickness indicator is proposed, and an application which can visualize the proposed indicator and physiological data in real-time is developed. Finally, experiment 2, which is modified by reviewing data of experiment 1, is explained. Experiment 2 is carried out for collecting data to develop and evaluate the indicator. Finally, the physiological data and indicator obtained from experiment 2 is analyzed and evaluated.

3.1 Research Hypothesis

Davis et al. [1] confirm researchers need to "develop a cost-effective physiological measure that quantifies an individual' s susceptibility to cybersickness and also to develop an objective measure of the intensity of the condition." And, some physiological data are changed with the occurrence of cybersickness. This fact was shown from some reviews [17][18]. In contrast, the possibility of the occurrence of cybersickness can be calculated from some physiological data. Therefore, the hypothesis in this research is succinctly defined as: "Cybersickness can be estimated from physiological data captured in real-time during a VR experience." Based on the hypothesis, this research was undertaken as follows:

- Select an instrument that can capture physiological data
- Consider metrics associated with cybersickness
- Develop a forecastable indicator of cybersickness
- Develop a real-time monitoring method of cybersickness
- Evaluate a forecastable indicator of cybersickness with Simulator Sickness Questionnaire Scores

3.2 Instrument

This section tests devices of measuring data to use in the experiments. An instrument to measure physiological data for objective evaluation to estimate cybersickness is needed. The criteria for testing devices are usability, price, sensor to skin contact reliability,

method of data capture and operation, and the type of data metrics available. The considered instruments were Bitalino (see Fig. 3.1), itDEAL bracelet (see Fig. 3.2), Polymate II (see Fig. 3.3), Apple Watch Series 3 (see Fig. 3.4), and Empatica E4 (see Fig. 3.5). Table 3.1 shows the tested instruments' information. First, the custom-made

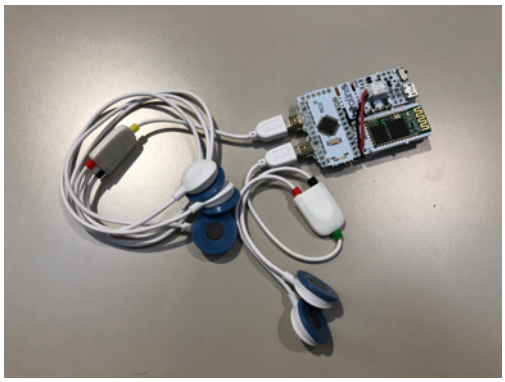


Fig. 3.1: Bitalino

Fig. 3.2: itDEAL baracelet

Fig. 3.3: Polymate II

Fig. 3.4: Apple Watch Series 3

Fig. 3.5: Empatica E4

sensor called Bitalino appeared most scalable and accessible. However, it was unreliable and unstable for capturing experimental data in a previous project [12]. Also, it was

Table 3.1: Tested instruments

Instrument	**Bitalino**	**itDEAL bracelet**	**Polymate II**	**Apple Watch Series 3**	**Empatica E4**
Measure	ECG, EDA, HR	HR	ECG, EDA	HRV	BVP, EDA, IBI, HR, TEMP
JPY (2019)	36,000	6,000	1,400,000	38,000	178,000
USD (2019)	340	56	13,240	360	1,690
EUR (2019)	306	51	11,930	324	1,520

considered difficult to use in the experiment because it was necessary to place a part of the electrode on a person's chest. Next, itDEAL bracelet was tested. It is inexpensive but a time scale of one-hour intervals does not allow for short-term data capture. Then, Polymate II was considered but it is expensive specialized equipment and an external tool was necessary when measuring EDA. It also requires to put an electrode on the skin the same as Bitalino. In addition, Polymate II required software that works on only the Windows 7 PC environment. Next, Apple Watch Series 3 was considered. However, it needs to be set up for each user with an Apple account. And, its time scale does not allow for detailed data observation. Finally, Empatica E4 was selected due to its wristband form being natural in the experiment (other than using electrodes) and the data could be captured using its eco-system.

3.2.1 Empatica E4

Empatica E4 provided by Empatica Inc.[*1] can capture Blood Volume Pulse (BVP) via a photoplethysmograph sensor, Electrodermal Activity (EDA), 3-axis accelerometer, Inter beat interval (IBI), and Heart Rate (HR). Sampling rates are recorded as follows: accelerometer at 32 Hz, BVP at 64 Hz, EDA at 4 Hz, the temperature at 4 Hz, HR in spans of 10 seconds.

There is an eco-system of Empatica E4 which is called E4 manager, E4 connect, and E4 realtime. First, E4 manager is a Windows / Mac OS application (see Fig. 3.6), which can import captured data by Empatica E4 via USB and transfers the data to E4 connect. Next, E4 connect is the web application, which can view and manage captured data on any web browser (see Fig. 3.7 and 3.8). The captured data can be downloaded as a CSV file from this application. Finally, E4 realtime is an Android / iOS application,

[*1] Empatica Inc. (https://www.empatica.com/)

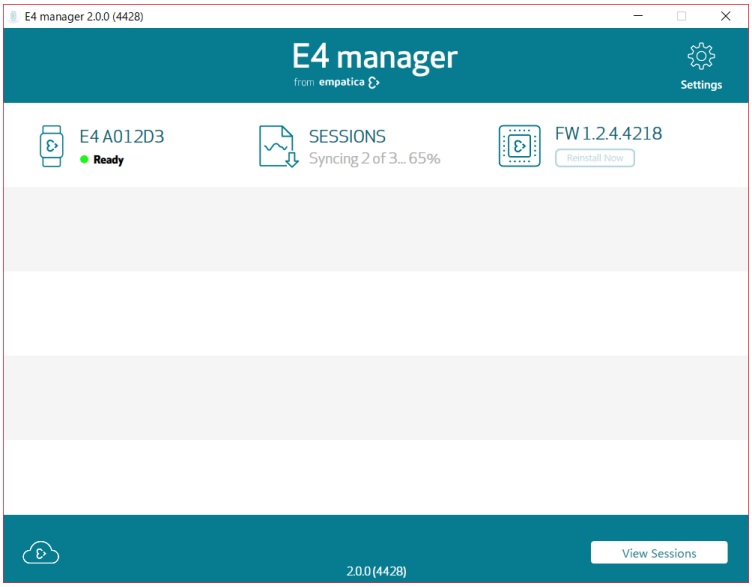


Fig. 3.6: E4 manager

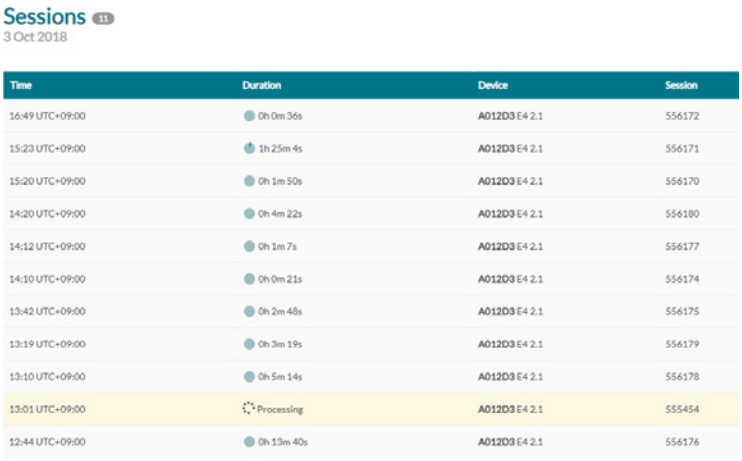


Fig. 3.7: E4 connect data list view

which can manage Empatica E4 via Bluetooth. For example, connecting Empatica E4 and smartphone(see Fig. 3.9), viewing real-time captured data (see Fig. 3.10), and uploading to E4 connect after a session.

Fig. 3.8: E4 connect chart view

Fig. 3.9: E4 realtime

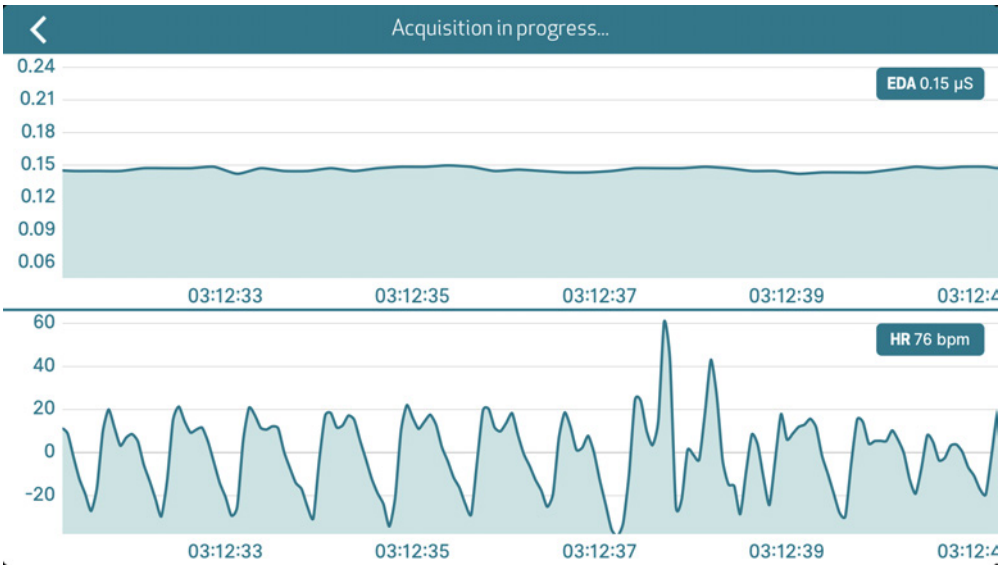


Fig. 3.10: E4 realtime chart view

3.3 PolyWorld and Tasks

A PC / VR application called PolyWorld (see Fig. 3.11) was designed and developed in Unity. PolyWorld was a small island with height differences. A design concept was a simple task to be able to induce weak anxiety in VR.

The goal of a task in PolyWorld was to look for five objects (a bag, lantern, a sword, a tablet, and a bottle) within a 5-minute time period. Fig. 3.12 and 3.13 are viewpoints of a player in PC and VR versions. In PC version, a timer on the top of the screen shows an elapsed time of the task. And, a list of target objects are shown on the bottom left of the screen. Moreover, a red pointer that is used for picking up the target object, is shown. As for the operation of the player, moving is based on a viewpoint that can be pressed by keyboard keys; W key to move forward, S key to move back, A key to move left, and D key to move right. The viewpoint can be changed by moving the PC mouse. The target object can be collected by left-clicking the mouse when the red pointer touches the target object. In VR version, the timer and list target objects are shown above a VR controller; this assumes the Oculus Go and its Hand controller. The red pointer as a laser appears from the controller. As for the operation of the player, to move a direction the controller trackpad is touched. The viewpoint in the HMD is changed by simply moving. The target object can be collected by pulling a controller's trigger when the red pointer touches the target object. To provide a supporting clue, particle effects appear as the player gets closer to an object (see Fig. 3.14).

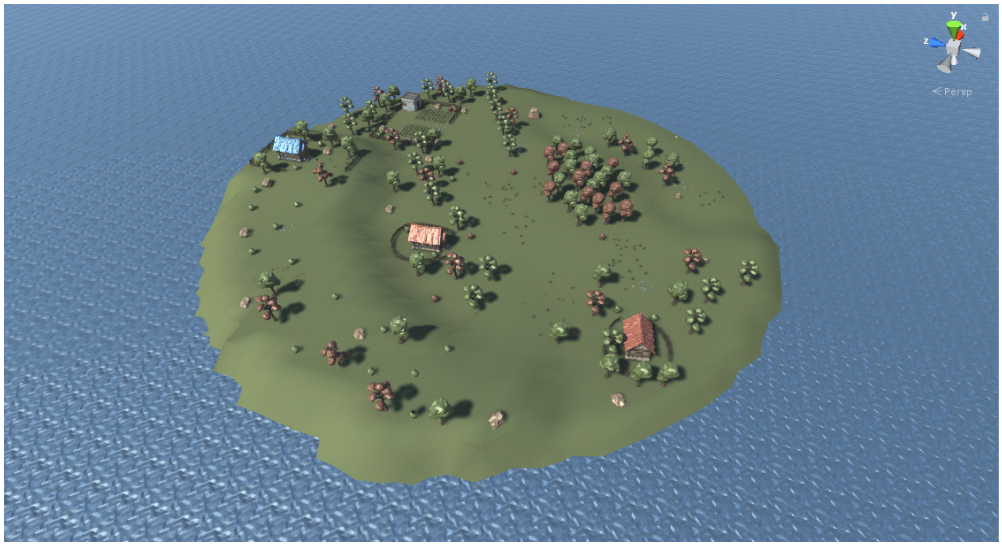


Fig. 3.11: PolyWorld

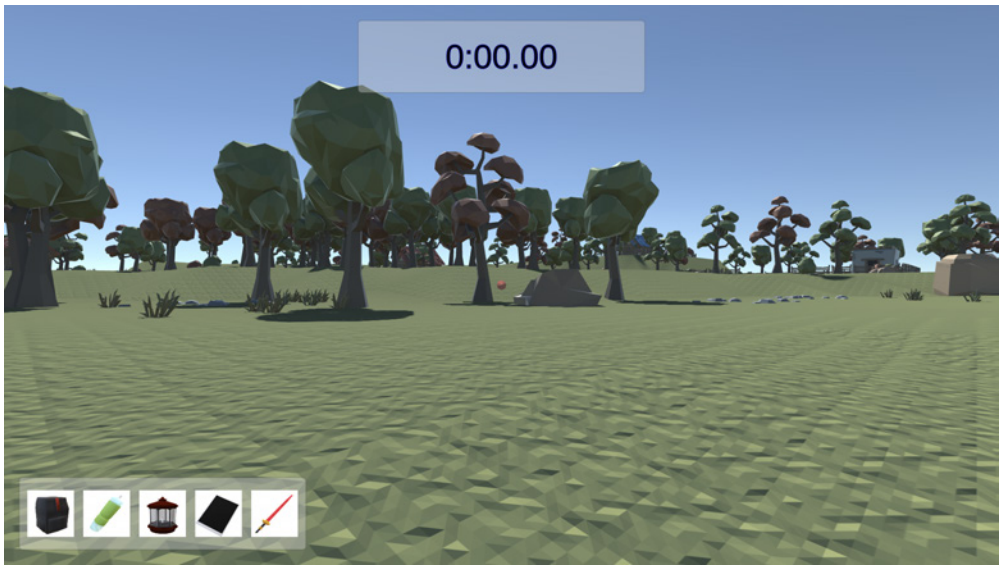


Fig. 3.12: PC PolyWorld

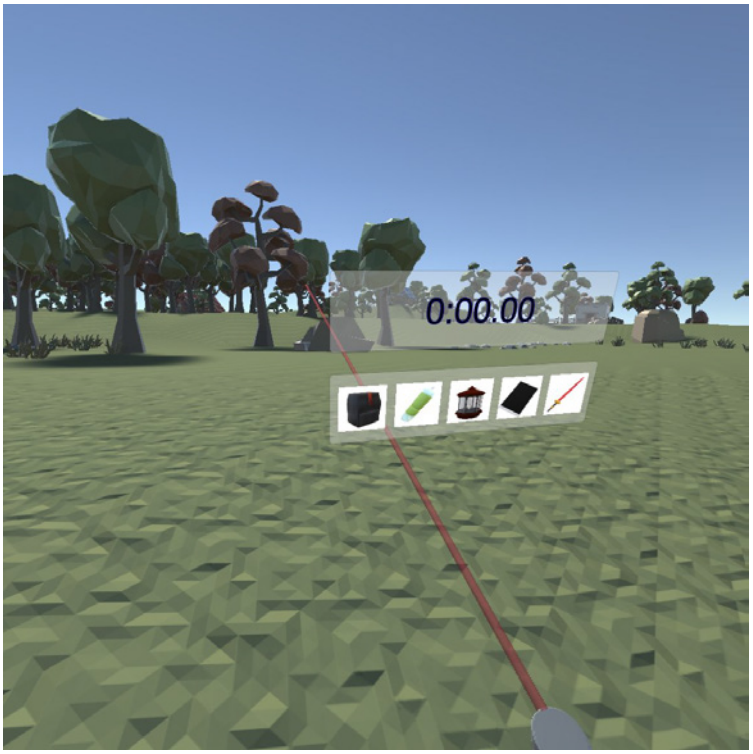


Fig. 3.13: VR PolyWorld

Fig. 3.14: The particle effect appeared when the player approached

3.4 Experimental Materials

As experimental materials, PolyWorld as PC and VR tasks (see Fig. 3.15 and 3.16), Empatica E4 as measuring instrument, the Simulator Sickness Questionnaire (SSQ) [13], a pre-survey, and a post-survey were prepared. PolyWorld task was carried out in the following environments.

- The PC task; A 27inch screen with FHD resolution is connected to PC, which is Windows 10 OS, Intel Core i7 8700k processor, 32GB RAM, and Nvidia GeForce GTX 1080Ti card. And a keyboard and mouse were used to operate in the PC PolyWorld.
- VR task; Oculus Go as a VRHMD. And its hand controller was used to operate the VR PolyWorld.

The pre-survey asked individual information, experiences of VR in the past, and time of playing a digital game in a day. The post-survey asked whether a similar feeling to motion sickness occurred in the experiment.

3.5 Experiment 1

The goal of the experiment was to compare physiological data during Normal (base condition), PC task (non-VR condition), and VR task (VR condition) experiences. The Normal and PC task were undertaken sitting, and the VR task was undertaken standing.

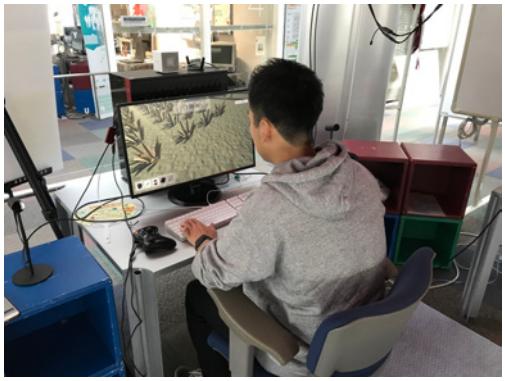


Fig. 3.15: PC task scene

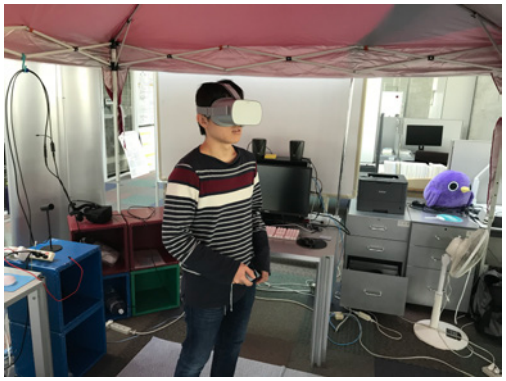


Fig. 3.16: VR task scene

Participants (n = 16; male 15, female 1) undertook the experiment. Informal consent from the participants was obtained for the experiment. And, all communication was undertaken in the participants' first language of Japanese. Physiological measures were recorded using the Empatica E4 device and automatically uploaded the Empatica Sessions to E4 connect via E4 realtime application. The experiment was carried out in the procedure as follows (and Fig. 3.17):

1. Participant completes a pre-survey.
2. Measure physiological data in the base condition (5-minute).
3. Carry out PC or VR task (max 5-minute).
4. Answer SSQ and then rest for 2 minutes.
5. Carry out VR or PC task (max 5-minute).
6. Answer SSQ.
7. Complete a post-survey.

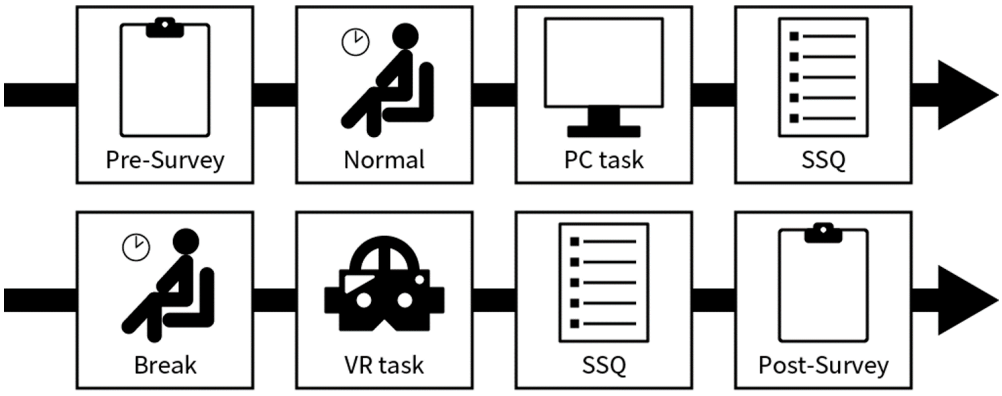


Fig. 3.17: Flowchart of experiment 1

3.6 Analysis 1

The section describes analyzing the results of experiment 1.

First, SSQ scores (Nausea, Oculomotor, Disorientation, and Total score) are analyzed by t-test between PC and VR task conditions. SSQ is a commonly used questionnaire for scaling cybersickness. Participants are asked to rate 16 symptoms on a 4-point scale (0 to 3). These ratings are used to generate scores, which are Nausea, Oculomotor, and Disorientation. Also, the Total score is considered.

Next, the three domains (Time, Frequency, and Non-linear domain) of Heart Rate Variability (HRV) are analyzed by ANOVA and Bonferroni-test between Normal, PC task, and VR task condition. HRV is the variation in the time interval between heartbeats and is measured by the variation in the beat-to-beat, or Inter beat interval (IBI). In general terms, it is an indication of how quickly an autonomic nervous system (ANS) communicates with a heart. Essentially, the ANS helps the brain send signals to the heart, muscles, and glands to help regulate digestion, heart functions and also stress. The ANS is made up of two systems: the sympathetic or 'fight or flight' system and parasympathetic which regulates energy. For example, a high HRV average is an indication that the parasympathetic system is helping the body recover. A low HRV average can be an indication of stress or anxiety. However, unlike heartbeat (beats-per-minute or BPM), HRV is not a single number and its metrics need to be compared to an individual' s baseline. Consequently, there are a number of methods used to analyze HRV: time-domain methods; frequency-domain methods; and non-linear methods.

- Time domain analysis parameters are NN Mean, SDNN, RMSSD. NN Mean is the mean value of the NN (normal beat-to-beat) intervals; SDNN is the standard deviation of NN intervals; RMSSD is the square root of the mean of the squares of successive differences between adjacent NNs.
- Frequency domain analysis parameters are Very Low-frequency (VLF), Low-frequency (LF), High-frequency (HF), and LF/HF.
- Non-linear analysis parameters are SD1, SD2, SD2/SD1, and S and can be viewed as a Poincaré Plot.

Also, BVP data can be used the same as ECG data because Selvaraj et al.'s reported these parameters show an equality [21]. These analyses are calculated by custom-developed programs using 'numpy' and 'astropy' Python libraries.

Finally, EDA data is analyzed by ANOVA and Bonferroni-test between Normal, PC task, and VR task conditions. EDA is the variation in the electrical characteristics of the skin, often referred to as skin conductance, and is an indication of physiological or

psychological arousal. EDA is analyzed using the EDA value, the maximum derivative of SCR (skin conductance response), and amplitude of peak. These parameters are calculated by the online application EDA Explorer*2 [22]. Its filter configurations are: minimum amplitude (threshold) is 0.01 micro-Siemens, offset is 0.5 sec, filter frequency is 1 Hz, filter order is 6, the max rise time is 4 sec and the max decay time is 4 sec.

3.6.1 SSQ

The mean and standard deviation of each SSQ parameter, calculated from the mean of each participant, are shown in Table 3.2. And, a result of SSQ t-test between Normal, PC task and VR task condition in Table 3.3

Table 3.2: SSQ score on each condition ($n = 16$). Values shown are M (SD).

Parameter	**PC task**	**VR task**
Nausea Score	15.50 (38.77)	33.39 (54.75)
Oculomotor Score	15.16 (22.99)	36.95 (44.79)
Disorientation Score	15.66 (48.98)	51.33 (81.28)
Total Score	17.77 (38.57)	45.11 (63.99)

Table 3.3: SSQ paired t-test results between PC task and VR task

Parameter	*df*	*t*	*p*
Nausea Score	15	-2.42	.03 *
Oculomotor Score	15	-3.08	.01 **
Disorientation Score	15	-2.90	.01 **
Total Score	15	-2.98	.01 **

Nausea Score

A paired-samples t-test was conducted to compare Nausea Score in PC task and VR task conditions. There was a significant difference in the scores for PC task ($M = 15.50, SD = 38.77$) and VR task ($M = 33.99, SD = 54.75$) conditions; $t(15) = -2.42, p = .03$.

Oculomotor Score

A paired-samples t-test was conducted to compare Oculomotor Score in PC task and VR task conditions. There was a significant difference in the scores for PC task ($M = 15.16, SD = 22.99$) and VR task ($M = 36.95, SD = 44.79$) conditions; $t(15) = -3.08, p < .01$.

*2EDA Explorer: https://eda-explorer.media.mit.edu/

Disorientation Score

A paired-samples t-test was conducted to compare Disorientation Score in PC task and VR task conditions. There was a significant difference in the scores for PC task ($M = 15.66, SD = 48.98$) and VR task ($M = 51.33, SD = 81.28$) conditions; $t(15) = -2.90, p = .01$.

Total Score

A paired-samples t-test was conducted to compare Total Score in PC task and VR task conditions. There was a significant difference in the scores for PC task ($M = 17.33, SD = 38.57$) and VR task ($M = 45.11, SD = 63.99$) conditions; $t(15) = -2.98, p < .01$.

3.6.2 HRV (Time domain)

The Mean and Standard Deviation of each HRV time domain parameter, calculated from the mean value of each participant, are shown in Table 3.4. And, a result of one-way repeated measures ANOVA and Bonferroni-test as post hoc tests between Normal, PC task and VR task condition in Table 3.5

Table 3.4: Time domain parameter on each condition ($n = 16$). Values shown are M (SD).

Parameter	**Normal**	**PC task**	**VR task**
NN Mean (ms)	860.41 (95.35)	834.15 (94.89)	757.76 (91.64)
SDNN (ms)	77.69 (19.24)	73.64 (19.56)	55.95 (13.35)
RMSSD (ms)	80.82 (25.06)	86.16 (34.80)	57.03 (17.68)

Table 3.5: Time domain ANOVA and paired t-test results between Normal, PC task and VR task

	ANOVA		**N & PC**		**N & VR**		**PC & VR**	
Parameter	F	p	t	p	t	p	t	p
NN Mean (ms)	42.19	.00 **	1.99	.18 ns	8.38	.00 **	8.50	.00 **
SDNN (ms)	19.13	.00 **	.89	1.17 ns	6.95	.00 **	5.20	.00 **
RMSSD (ms)	13.95	.00 **	.98	1.02 ns	5.21	.00 **	4.00	.00 **

NN Mean

A one-way repeated measures ANOVA was conducted to compare the values of NN Mean of Normal, PC task and VR task condition. There was a significant effect of NN Mean, $F(2, 30) = 42.19, p = .00$. Three paired samples t-tests were used to make

post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = 860.41, SD = 95.35$) and PC task ($M = 834.15, SD = 94.89$) conditions; $t(15) = 1.99, p = .18$. A second paired samples t-test indicated that there was a significant difference in the scores for Normal ($M = 860.41, SD = 95.35$) and VR task ($M = 757.76, SD = 91.64$) conditions; $t(15) = 8.38, p = .00$. A third paired samples t-test indicated that there was a significant difference in the scores for PC task ($M = 834.15, SD = 94.89$) and VR task ($M = 757.76, SD = 91.64$) conditions; $t(15) = 8.50, p = .00$.

SDNN

A one-way repeated measures ANOVA was conducted to compare the values of SDNN of Normal, PC task and VR task condition. There was a significant effect of SDNN, $F(2, 30) = 19.13, p = .00$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = 77.69, SD = 19.24$) and PC task ($M = 73.64, SD = 19.56$) conditions; $t(15) = .89, p = .1.17$. A second paired samples t-test indicated that there was a significant difference in the scores for Normal ($M = 77.69, SD = 19.24$) and VR task ($M = 55.95, SD = 13.35$) conditions; $t(15) = 6.95, p = .00$. A third paired samples t-test indicated that there was a significant difference in the scores for PC task ($M = 73.64, SD = 19.56$) and VR task ($M = 55.95, SD = 13.35$) conditions; $t(15) = 5.20, p = .00$.

RMSSD

A one-way repeated measures ANOVA was conducted to compare the values of RMSSD of Normal, PC task and VR task condition. There was a significant effect of RMSSD, $F(2, 30) = 13.95, p = .00$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = 80.82, SD = 25.06$) and PC task ($M = 86.16, SD = 34.8$) conditions; $t(15) = -.98, p = 1.00$. A second paired samples t-test indicated that there was a significant difference in the scores for Normal ($M = 80.82, SD = 25.06$) and VR task ($M = 57.03, SD = 17.68$) conditions; $t(15) = 5.21, p = .00$. A third paired samples t-test indicated that there was a significant difference in the scores for PC task ($M = 86.16, SD = 34.8$) and VR task ($M = 57.03, SD = 17.68$) conditions; $t(15) = 4.00, p = .00$.

3.6.3 HRV (Frequency domain)

The Mean and Standard Deviation of each HRV frequency domain parameter, calculated from the mean value of each participant, are shown in Table 3.6. And, a result of one-way repeated measures ANOVA and Bonferroni-test as post hoc tests between Normal, PC task and VR task condition in Table 3.7.

Table 3.6: Frequency domain parameter on each condition ($n = 16$). Values shown are M (SD).

Parameter	**Normal**	**PC task**	**VR task**
VLF (ms^2/Hz)	900.89 (457.45)	703.37 (612.89)	607.15 (383.69)
LF (ms^2/Hz)	1068.01 (522.85)	919.03 (576.55)	625.26 (374.00)
HF (ms^2/Hz)	1086.39 (638.25)	862.57 (637.75)	465.45 (297.78)
LF/HF	1.09 (0.38)	1.83 (2.37)	1.60 (0.78)

Table 3.7: Frequency domain ANOVA and paired t-test results between Normal, PC task and VR task condition

	ANOVA		**N & PC**		**N & VR**		**PC & VR**	
Parameter	F	p	t	p	t	p	t	p
VLF (ms^2/Hz)	1.94	.16 ns	-	-	-	-	-	-
LF (ms^2/Hz)	6.45	.00 **	1.00	.99 ns	4.42	.00 **	2.40	.09 ns
HF (ms^2/Hz)	8.95	.00 **	1.63	.36 ns	3.95	.00 **	2.64	.06 ns
LF/HF	0.96	.39 ns	-	-	-	-	-	-

VLF

A one-way repeated measures ANOVA was conducted to compare the values of VLF of Normal, PC task and VR task condition. There was no significant effect of VLF, $F(2, 30) = 1.94, p = .16$.

LF

A one-way repeated measures ANOVA was conducted to compare the values of LF of Normal, PC task and VR task condition. There was a significant effect of LF, $F(2, 30) = 6.45, p = .00$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = 1068.01, SD = 522.85$) and PC task ($M = 919.03, SD = 576.55$) conditions; $t(15) = 1.00, p = .99$. A second paired samples t-test indicated that there was a significant difference in the scores for

Normal ($M = 1068.01, SD = 522.85$) and VR task ($M = 625.26, SD = 374.00$) conditions; $t(15) = 4.42, p = .00$. A third paired samples t-test indicated that there was no significant difference in the scores for PC task ($M = 919.03, SD = 576.55$) and VR task ($M = 625.26, SD = 374.00$) conditions; $t(15) = 2.40, p = .09$.

HF

A one-way repeated measures ANOVA was conducted to compare the values of HF of Normal, PC task and VR task condition. There was a significant effect of HF, $F(2, 30) = 8.95, p = .00$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = 1086.39, SD = 638.25$) and PC task ($M = 862.57, SD = 637.75$) conditions; $t(15) = 1.63, p = .36$. A second paired samples t-test indicated that there was a significant difference in the scores for Normal ($M = 1086.39, SD = 638.25$) and VR task ($M = 465.45, SD = 297.78$) conditions; $t(15) = 3.95, p = .00$. A third paired samples t-test indicated that there was no significant difference in the scores for PC task ($M = 862.57, SD = 637.75$) and VR task ($M = 465.45, SD = 297.78$) conditions; $t(15) = 2.64, p = .06$.

LF/HF

A one-way repeated measures ANOVA was conducted to compare the values of LF/HF of Normal, PC task and VR task condition. There was no significant effect of LF/HF, $F(2, 30) = 0.96, p = .39$.

3.6.4 HRV (Non-linear domain)

A Poincaré Plot as a non-linear domain analysis was analyzed. The Poincaré Plot is a method of HRV analysis [23][24]. It is a graphical representation of temporal correlations within the RR intervals derived from ECG. Each data point represents a pair of successive beats, the x-axis is the current NN interval, and the y-axis is the previous NN interval. In other words, the nth RR interval is plotted together with the $n + 1$th RR interval at $(x, y) = ((RR_n, RR_{n+1})$. And, SD1, SD2, SD2/SD1, and S are calculated from plotted data. SD1 is the standard deviation measuring the dispersion of points in the plot across the identity line. On the other hand, SD2 is the standard deviation measuring the dispersion of points along the identity line. And, SD2/SD1 is the ratio of SD2 to SD1 which shows analogy to LF/HF from HRV frequency domain analysis. Moreover, S is corresponding to the area of an imaginary ellipse (S = SD1 $\times$ SD2) with the axes of lengths. In this analysis, RR interval values are altered by IBI which is the same

meaning. An example of the Poincaré Plot of Normal condition, PC Task condition, VR Task condition, and combined data is shown in Fig. 3.18. The plot represents a changing time series. Table 3.8 shows the mean and standard deviation of each Poincaré Plot parameter, as calculated from the mean of each participant. And, a result of one-way repeated measures ANOVA and Bonferroni-test as post hoc tests between Normal, PC task and VR task condition in Table 3.9.

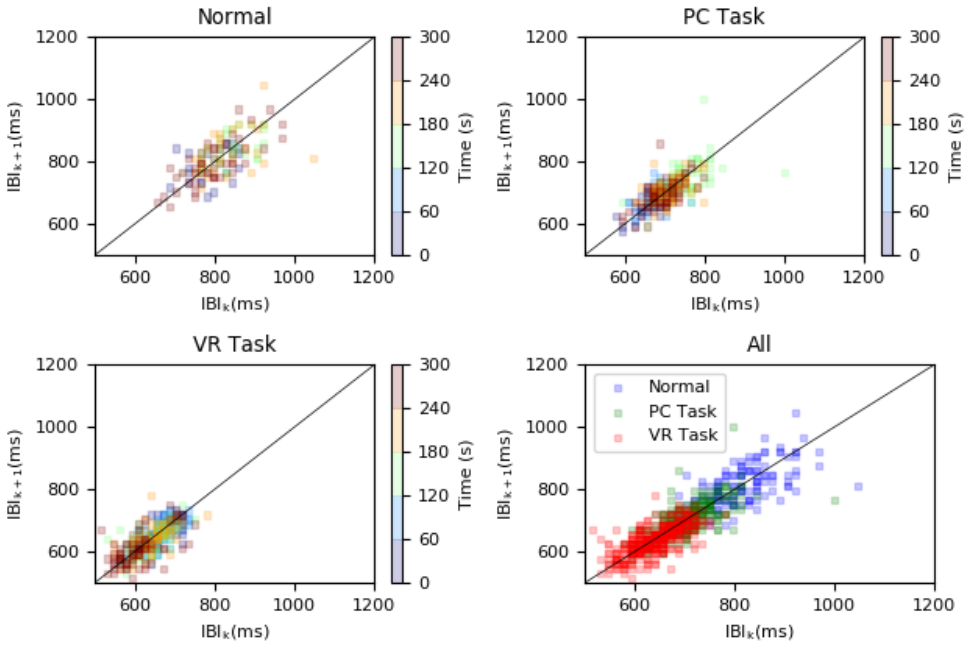


Fig. 3.18: Poincaré Plot

Table 3.8: Poincaré Plot parameter on each condition ($n = 16$). Values shown are M (SD).

Parameter	**Normal**	**PC task**	**VR task**
SD1 (ms)	57.42 (17.87)	62.38 (27.07)	40.45 (12.59)
SD2 (ms)	93.16 (23.34)	82.40 (22.01)	67.39 (16.75)
SD2/SD1	1.68 (0.35)	1.46 (0.48)	1.76 (0.45)
S (ms^2)	17676.44 (8974.11)	16666.49 (8532.65)	8920.94 (4169.64)

SD1

A one-way repeated measures ANOVA was conducted to compare the values of SD1 of Normal, PC task and VR task condition. There was a significant effect of SD1, $F(2, 30) = 12.67, p = .00$. Three paired samples t-tests were used to make post hoc test comparisons

Table 3.9: Poincaré Plot ANOVA and paired t-test results between Normal, PC task and VR task condition

	ANOVA		**N & PC**		**N & VR**		**PC & VR**	
Parameter	F	p	t	p	t	p	t	p
SD1 (ms)	12.67	.00 **	-1.15	.81 ns	5.21	.00 **	3.80	.00 **
SD2 (ms)	14.76	.00 **	1.81	.27 ns	6.53	.00 **	3.61	.00 **
SD2/SD1	4.26	.02 *	2.39	.09 ns	-.85	1.23 ns	-2.27	.12 ns
S (ms^2)	15.18	.00 **	.56	1.77 ns	5.04	.00 **	4.65	.00 **

using Bonferroni-test between conditions. A first paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = 57.42, SD = 17.87$) and PC task ($M = 62.38, SD = 27.07$) conditions; $t(15) = -1.15, p = .81$. A second paired samples t-test indicated that there was a significant difference in the scores for Normal ($M = 57.42, SD = 17.87$) and VR task ($M = 40.45, SD = 12.59$) conditions; $t(15) = 5.21, p = .00$. A third paired samples t-test indicated that there was a significant difference in the scores for PC task ($M = 62.38, SD = 27.07$) and VR task ($M = 40.45, SD = 12.59$) conditions; $t(15) = 3.80, p = .00$.

SD2

A one-way repeated measures ANOVA was conducted to compare the values of SD2 of Normal, PC task and VR task condition. There was a significant effect of SD2, $F(2, 30) = 14.76, p = .00$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = 93.16, SD = 23.34$) and PC task ($M = 82.40, SD = 22.01$) conditions; $t(15) = 1.81, p = .27$. A second paired samples t-test indicated that there was a significant difference in the scores for Normal ($M = 93.16, SD = 23.34$) and VR task ($M = 67.39, SD = 16.75$) conditions; $t(15) = 6.53, p = .00$. A third paired samples t-test indicated that there was a significant difference in the scores for PC task ($M = 82.40, SD = 22.01$) and VR task ($M = 67.39, SD = 16.75$) conditions; $t(15) = 3.61, p = .00$.

SD2/SD1

A one-way repeated measures ANOVA was conducted to compare the values of SD2/SD1 of Normal, PC task and VR task condition. There was a significant effect of SD2/SD1, $F(2, 30) = 4.26, p = .02$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test indicated that there was no significant difference in the scores for Normal ($M =$

$1.68, SD = .35$) and PC task ($M = 1.46, SD = .48$) conditions; $t(15) = 2.39, p = .09$. A second paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = 1.68, SD = .35$) and VR task ($M = 1.76, SD = .45$) conditions; $t(15) = -.85, p = 1.00$. A third paired samples t-test indicated that there was no significant difference in the scores for PC task ($M = 1.46, SD = .48$) and VR task ($M = 1.76, SD = .45$) conditions; $t(15) = -2.27, p = .12$.

S

A one-way repeated measures ANOVA was conducted to compare the values of S of Normal, PC task and VR task condition. There was a significant effect of S, $F(2, 30) = 15.18, p = .00$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = 17676.44, SD = 8974.11$) and PC task ($M = 16666.49, SD = 8632.65$) conditions; $t(15) = .56, p = 1.00$. A second paired samples t-test indicated that there was a significant difference in the scores for Normal ($M = 17676.44, SD = 8974.11$) and VR task ($M = 8920.94, SD = 4169.64$) conditions; $t(15) = 5.04, p = .00$. A third paired samples t-test indicated that there was a significant difference in the scores for PC task ($M = 16666.49, SD = 8632.65$) and VR task ($M = 8920.94, SD = 4169.64$) conditions; $t(15) = 4.65, p = .00$.

3.6.5 EDA

The mean and standard deviation of each EDA parameter, calculated from the mean of each participant, are shown in Table 3.10. And, a result of one-way repeated measures ANOVA and Bonferroni-test as post hoc tests between Normal, PC task and VR task condition in Table 3.11.

Table 3.10: EDA parameter on each condition ($n = 16$). Values shown are M (SD).

Parameter	Normal	PC task	VR task
Peak EDA (μS)	1.58 (2.81)	2.12 (4.06)	3.48 (4.78)
Max deriv. ($\mu S/s$)	0.09 (0.06)	0.07 (0.05)	0.15 (0.14)
Amplitude (μS)	0.06 (0.06)	0.04 (0.05)	0.13 (0.13)
SCR width (s)	1.14 (0.32)	1.01 (0.35)	1.43 (0.49)

Peak EDA

A one-way repeated measures ANOVA was conducted to compare the values of peak EDA of Normal, PC task and VR task condition. There was a significant effect of peak

Table 3.11: EDA ANOVA and paired t-test results between Normal, PC task and VR task condition

	ANOVA		**N & PC**		**N & VR**		**PC & VR**	
Parameter	F	p	t	p	t	p	t	p
Peak EDA (μS)	7.61	.00 **	-1.43	.51 ns	-3.24	.03 *	-2.62	.06 ns
Max deriv. ($\mu S/s$)	5.02	.01 *	1.62	.39 ns	-1.78	.27 ns	-2.81	.03 *
Amplitude (μS)	5.94	.01 *	1.04	.93 ns	-2.23	.12 ns	-2.89	.03 *
SCR width (s)	8.85	.00 **	1.51	.45 ns	-2.82	.03 *	-3.65	.00 **

EDA, $F(2, 30) = 7.61, p = .00$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = 1.58, SD = 2.81$) and PC task ($M = 2.12, SD = 4.06$) conditions; $t(15) = -1.43, p = .51$. A second paired samples t-test indicated that there was a significant difference in the scores for Normal ($M = 1.58, SD = 2.81$) and VR task ($M = 3.48, SD = 4.78$) conditions; $t(15) = -3.24, p = .03$. A third paired samples t-test indicated that there was no significant difference in the scores for PC task ($M = 2.12, SD = 4.06$) and VR task ($M = 3.48, SD = 4.78$) conditions; $t(15) = -2.62, p = .06$.

Maximum derivation of SCR

A one-way repeated measures ANOVA was conducted to compare the values of max derivative of SCR of Normal, PC task and VR task condition. There was a significant effect of max derivative of SCR, $F(2, 30) = 5.02, p = .01$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = .09, SD = .06$) and PC task ($M = .07, SD = .05$) conditions; $t(15) = 1.62, p = .39$. A second paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = .09, SD = .06$) and VR task ($M = .15, SD = .14$) conditions; $t(15) = -1.78, p = .27$. A third paired samples t-test indicated that there was a significant difference in the scores for PC task ($M = .07, SD = .05$) and VR task ($M = .15, SD = .14$) conditions; $t(15) = -2.81, p = .03$.

Amplitude

A one-way repeated measures ANOVA was conducted to compare the values of amplitude of Normal, PC task and VR task condition. There was a significant effect of amplitude, $F(2, 30) = 5.94, p = .01$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired sam-

ples t-test indicated that there was no significant difference in the scores for Normal ($M = .06, SD = .06$) and PC task ($M = .04, SD = .05$) conditions; $t(15) = 1.04, p = .93$. A second paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = .06, SD = .06$) and VR task ($M = .13, SD = .13$) conditions; $t(15) = -2.23, p = .12$. A third paired samples t-test indicated that there was a significant difference in the scores for PC task ($M = .04, SD = .05$) and VR task ($M = .13, SD = .13$) conditions; $t(15) = -2.89, p = .03$.

SCR width

A one-way repeated measures ANOVA was conducted to compare the values of SCR width of Normal, PC task and VR task condition. There was a significant effect of SCR width, $F(2, 30) = 8.85, p < .001$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test indicated that there was a significant difference in the scores for Normal ($M = 1.14, SD = .32$) and PC task ($M = 1.01, SD = .35$) conditions; $t(15) = 1.51, p = .45$. A second paired samples t-test indicated that there was no significant difference in the scores for Normal ($M = 1.14, SD = .32$) and VR task ($M = 1.43, SD = .49$) conditions; $t(15) = -2.82, p = .03$. A third paired samples t-test indicated that there was a significant difference in the scores for PC task ($M = 1.01, SD = .35$) and VR task ($M = 1.43, SD = .49$) conditions; $t(15) = -3.65, p = .00$.

3.7 Reflection of Experiment 1

Firstly, all SSQ scores (Nausea Score, Oculomotor Score, Disorientation Score) showed a significant difference between VR task and PC task conditions. It is conceivable that the condition of VR task was more likely to generate anxiety of a higher intensity than PC task condition. And, PolyWorld can be considered as a valid activity to trigger cybersickness in VR task compared with PC task.

Next, in the result of HRV analysis, NN Mean, SDNN, RMSSD of the Time domain, and SD1, SD2, and S of the Non-linear analysis Poincaré Plot were shown to be significant in the VR task condition compared with the PC task condition. In the parameters of the Frequency domain, there was no significant difference between the respective conditions in the VLF and the LF/HF, and the results showed that there was a significant difference only between the Normal and the VR task conditions for LF and HF. In addition, although SD2/SD1 of Poincaré Plot showed a significant difference in the analysis of variance, no significant difference was found between the respective conditions in a post hoc test under each condition. For this reason, it is proposed that NN Mean, SDNN and

RMSSD of the Time domain parameters of HRV, and SD1, SD2, S of the Poincaré Plot parameters can be an index of cybersickness. However, the intensity of cybersickness was uncertain. Also, since the value of S is obtained from SD1 and SD2 data, the index can be supplemented with SD1 and SD2 values. LF and HF in the Frequency domain showed a significant difference only in Normal and VR conditions. This can be an effective index for evaluating the use of a reduction method to lower cybersickness when compared with the Normal state [17]; although it may not be effective for comparing different reduction methods.

From the results of EDA, the SCR and SCL parameters tended to rise when comparing the Normal condition with the VR task condition. Statistical analysis showed that SCR width in Normal condition and PC task conditions were significantly different. On the other hand, EDA peak Amplitude in Normal condition and PC task condition showed no significant difference. In addition, Max derivate and Amplitude of SCR in VR task condition showed a significant difference when compared to Max derivate and Amplitude of SCR in Normal task condition. However, there was no significant difference when compared with PC task condition. Consequently, Max derivate and Amplitude of SCR may not be suitable for an index as there were no significant differences between the Normal condition and the PC task. However, although the data is inconclusive there is a possibility that SCR width could be an index of cybersickness.

There are obviously limitations in experiment 1. Firstly, it is acknowledged that the experiment was limited by the number of participants. Also, since the standard deviation for each data is large, it is not possible to generalize individual differences. And the EDA analysis showed a significant difference in the SCR width, even though the parameters were obtained from the EDA Explorer in this analysis. However, some parameters analyzed from EDA Explorer were later found to be difficult to embed in the Cybatica application to work in real-time. Therefore, the last one minute of Peak EDA and Mean EDA will be used. Moreover, there is a problem with the effect of the order in the analysis, which is not considered in the procedure of experiment 1. This problem should be improved in the next experiment.

3.8 Onset of Cybersickness

An indicator of cybersickness called Onset of Cybersickness (OCS) is proposed. OCS is based on SSQ, with the indicator replacing SSQ items by physiological parameters. OCS Total Score is the same as SSQ Total score. Also, OCS Nausea Score, Oculomotor Score, and Disorientation Score are the same SSQ Nausea Score, Oculomotor Score, and Disorientation Score as a subset of OCS. Therefore, OCS scores uses the SSQ score equations

(see Equation 3.8.1, 3.8.2, 3.8.3, and 3.8.4). However, a value of a' is physiological data that replaces SSQ symptoms scores (see Equation 3.8.5). b' and c' also are calculated to replace Oculomotor and Disorientation of SSQ (see Equation 3.8.6 and 3.8.7).

$$OCSNauseaScore = a' \times 9.54 \tag{3.8.1}$$

$$OCSOculomotorScore = b' \times 7.58 \tag{3.8.2}$$

$$OCSDisorientationScore = c' \times 13.92 \tag{3.8.3}$$

$$OCSTotalScore = (a' + b' + c') \times 3.74 \tag{3.8.4}$$

$$a' = \sum_{i=1}^{n} Proxy(NauseaParameter_i) \tag{3.8.5}$$

$$b' = \sum_{i=1}^{n} Proxy(OculomotorParameter_i) \tag{3.8.6}$$

$$c' = \sum_{i=1}^{n} Proxy(DisorientationParameter_i) \tag{3.8.7}$$

The use of indicators replacing SSQ items by physiological parameters is justified because SSQ is an established formula generated over many years and experiments. And, OCS scores can be compared with SSQ score, if OCS scores are made the same scale as SSQ scores. The use of SSQ for OCS will be detailed by the result of the next experiment in section 3.12 and 3.13 in which the validity of OCS scores is considered by analyzing correlations between OCS scores and SSQ scores.

3.9 Cybatica

In order to display OCS and associated physiological data, an Android / iOS application of a real-time cybersickness monitoring system called Cybatica was developed by using Xamarin.Forms and C#. Cybatica can record BVP and EDA when connected to the Empatica E4 and can calculate analysis data such as OCS, SDNN, NN Mean, Mean EDA, and Peak EDA from obtained data. And, Cybatica can visualize a real-time chart of physiological data and analysis data. Moreover, the captured data can be stored in any storage such as Google Drive when stopped measuring. A library of controlling Empatica E4, Empalink for Android and E4link for iOS, which were provided by the official developer portal, were used.

Table 3.12: Computation of weight of symptoms score (0 to 3)

Parameter	**Nausea**	**Oculomotor**	**Disorientation**
General discomfort	x	x	
Fatigue		x	
Headache		x	
Eye strain		x	
Difficulty focusing		x	x
Increased salivation	x		
Sweating	x		
Nausea	x		x
Difficulty concentrating	x	x	
Fullness of head			x
Blurred vision		x	x
Dizzy (eyes open)			x
Dizzy (eyes closed)			x
Vertigo			x
Stomach awareness	x		
Burping	x		
Total	a'	b'	c'
$OCSNauseaScore = a' \times 9.54$			
$OCSOculomotorScore = b' \times 7.58$			
$OCSDisorientationScore = c' \times 13.92$			
$OCS = (a' + b' + c') \times 3.74$			

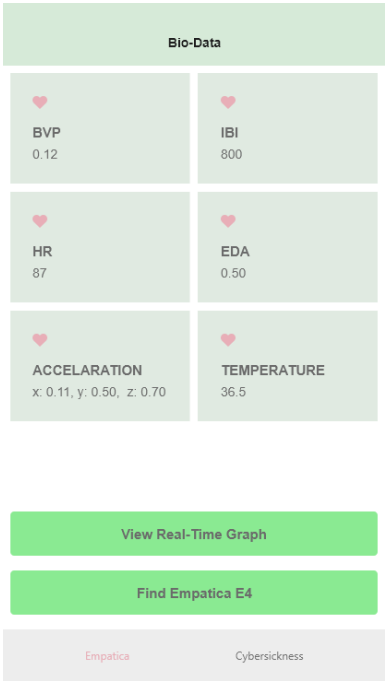
3.9.1 Design

First, a prototype screen design is considered before developing. Fig. 3.19 is designed using Adobe XD. The application requirement is defined as follows.

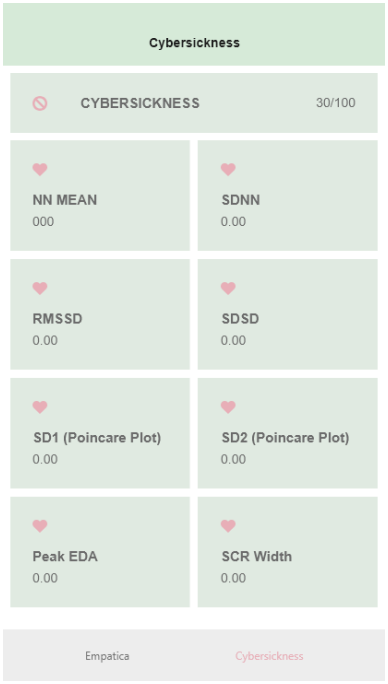
- The connection between Empatica E4 can be managed
- The physiological data can be shown
- The analysis data can be shown
- The realtime updatable chart can be shown

Next, the first prototype of Cybatica is developed by C# and Xamarin.Forms (see Fig. 3.20).

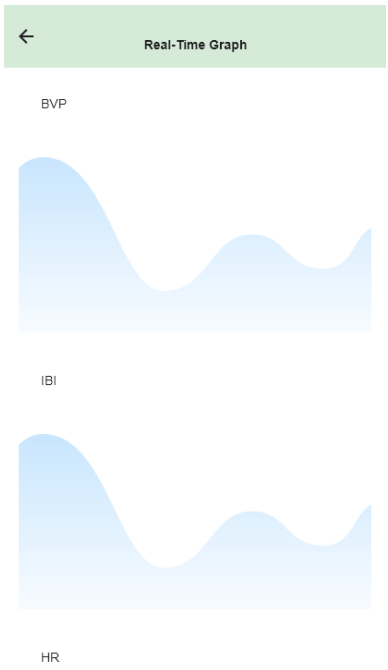
Finally, the base session and data session are separated because to implement OCS the ratio from the mean value in the base session (Normal condition in Experiment 1) is used. Fig. 3.21 is the final design of Cybatica. See also Appendix A.



(a) Bio Data view

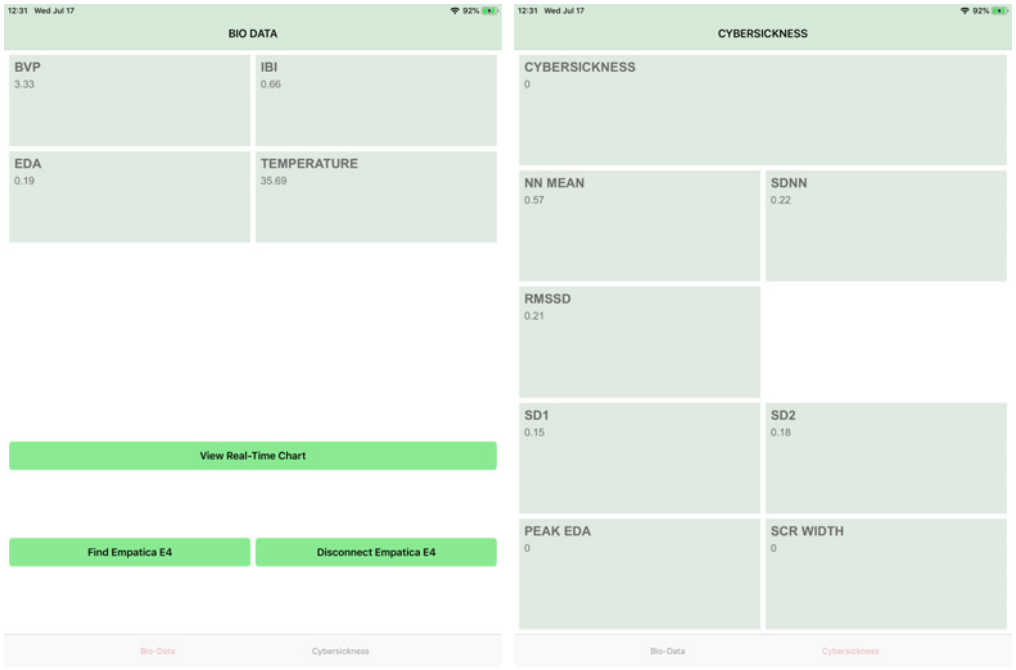


(b) Analysis view

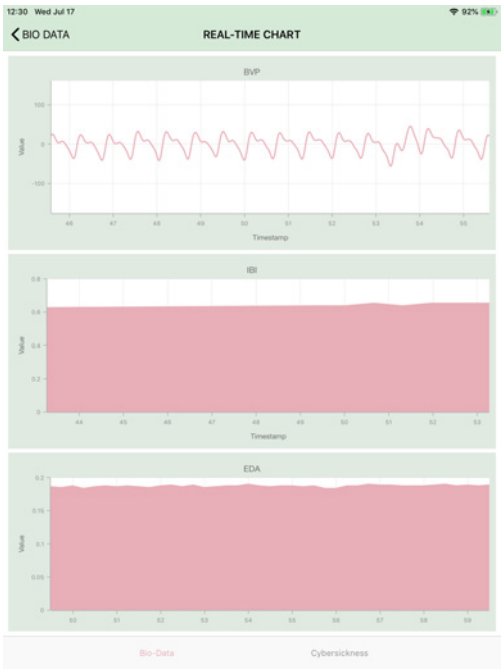


(c) Chart view

Fig. 3.19: The screen design of Cybatica

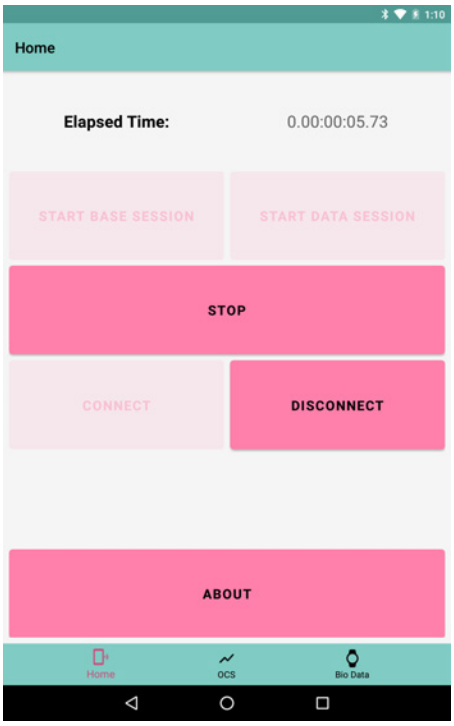


(a) Home and Bio Data view (b) Analysis view

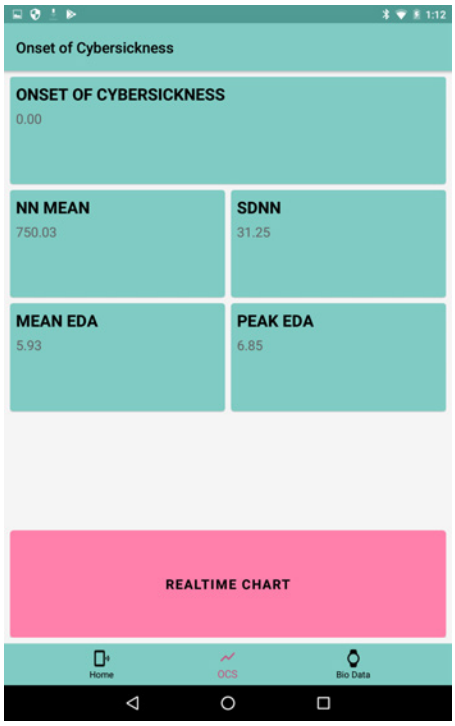


(c) Chart view

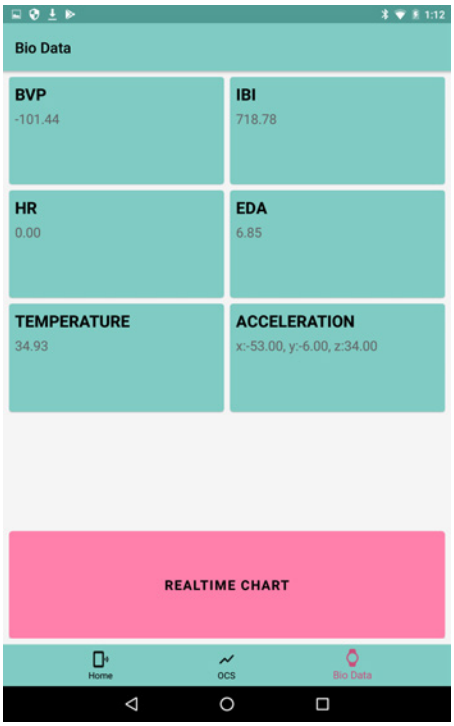
Fig. 3.20: Cybatica prototype



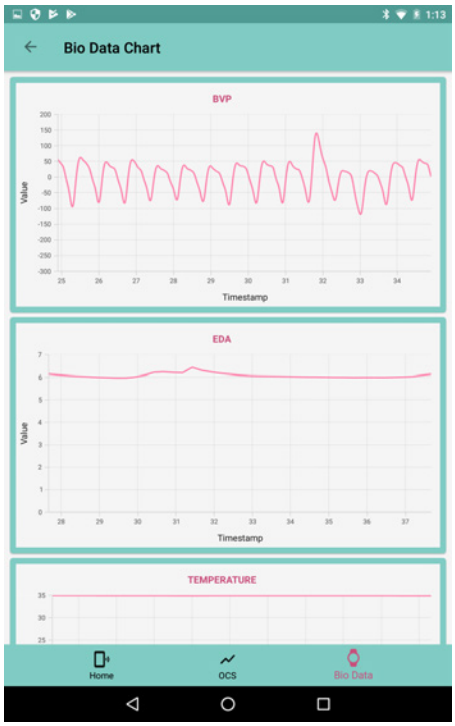
(a) Home view



(b) Analysis view



(c) Bio Data view



(d) Bio Data Chart view

Fig. 3.21: Cybatica

3.10 Experiment 2

The goal of the experiment was to collect data to develop OCS and evaluate OCS. The experiment was the same as experiment 1, but the procedure was modified after the reflection of experiment 1. Changed points were: 1. The Normal condition was

Fig. 3.22: PC task scene

Fig. 3.23: VR task scene

undertaken sitting, and PC task and the VR task were undertaken standing (see Fig. 3.22 and 3.23); 2. The rest time between tasks was increased to 5-minute; 3. The data obtained method of physiological data was changed to Cybatica from Empatica E4 ecosystems; 4. The order of the procedure of PC and VR tasks were inversed. The modified experiment procedure was as follows (and Fig. 3.24)

1. Participant completes a pre-survey.
2. Measure physiological data in the base condition (5-minute).
3. Carry out PC or VR task (5-minute).
4. Answer SSQ and rest for 5-minute.
5. Carry out VR or PC task (5-minute).
6. Answer SSQ.
7. Complete a post-survey.

Participants (n = 21; male 18, female 3) undertook the experiment. Before the experiment, written consent from (see Appendix B) the participants was obtained for the experiment. And, all communication was undertaken in the participants' first language of Japanese. Physiological measures were recorded using the Empatica E4 device and uploaded the sessions to Google Drive via Cybatica and its function.

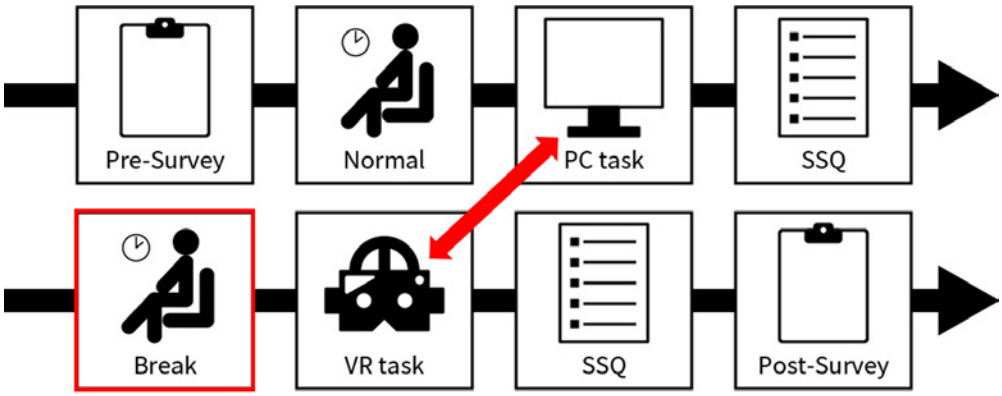


Fig. 3.24: Flowchart of experiment 2

3.11 Analysis 2

In this section, a result of the obtained data in experiment 2 will be shown. First, SSQ scores will be analyzed to confirm the validation of activity by t-test similar to experiment 1. Second, differences of physiological data such as NN Mean, SDNN, Mean EDA, and Peak EDA will be analyzed by ANOVA between three conditions (Normal, PC task, and VR task conditions). Finally, a specific composition of OCS will be calculated by multiple linear regressions using physiological data.

3.11.1 SSQ

The mean and standard deviation of each SSQ parameter, calculated from the mean of each participant, are shown in Table 3.13. And, a result of SSQ t-test between Normal, PC task and VR task condition are shown in Table 3.14

Table 3.13: SSQ score on each condition ($n = 21$). Values shown are M (SD).

Parameter	PC task	VR task
Nausea Score	7.27 (9.00)	27.71 (26.11)
Oculomotor Score	9.02 (10.34)	24.18 (22.44)
Disorientation Score	9.94 (15.96)	41.10 (44.99)
Total Score	9.97 (11.61)	33.84 (30.16)

Nausea Score

A paired-samples t-test was conducted to compare Nausea Score in PC task and VR task conditions. There was a significant difference in the scores for PC task ($M = 7.27, SD = 9.00$) and VR task ($M = 27.71, SD = 26.11$) conditions; $t(20) = -3.76, p = .00$.

Table 3.14: SSQ paired t-test results between PC task and VR task

Parameter	*df*	*t*	*p*
Nausea Score	20	-3.76	.00 **
Oculomotor Score	20	-2.84	.01 *
Disorientation Score	20	-2.89	.01 *
Total Score	20	-3.37	.00 **

Oculomotor Score

A paired-samples t-test was conducted to compare Oculomotor Score in PC task and VR task conditions. There was a significant difference in the scores for PC task ($M = 9.02, SD = 10.34$) and VR task ($M = 24.18, SD = 22.44$) conditions; $t(20) = -2.84, p = .01$.

Disorientation Score

A paired-samples t-test was conducted to compare Disorientation Score in PC task and VR task conditions. There was a significant difference in the scores for PC task ($M = 9.94, SD = 15.96$) and VR task ($M = 41.10, SD = 44.99$) conditions; $t(20) = -2.89, p = .01$.

Total Score

A paired-samples t-test was conducted to compare Total Score in PC task and VR task conditions. There was a significant difference in the scores for PC task ($M = 9.97, SD = 11.61$) and VR task ($M = 33.84, SD = 30.16$) conditions; $t(20) = -3.37, p = .00$.

3.11.2 HRV

The Mean and Standard Deviation of each HRV time domain parameter, calculated from the mean value of 5-minute in the experiment 2 of each participant, are shown in Table 3.15. And, a result of one-way repeated measures ANOVA and Bonferroni-test as post hoc tests between Normal, PC task and VR task condition are shown in Table 3.16

Table 3.15: HRV parameter on each condition ($n = 21$). Values shown are M (SD).

Parameter	**Normal**	**PC task**	**VR task**
NN Mean (ms)	712.85 (98.16)	694.09 (86.92)	678.46 (86.92)
SDNN (ms)	55.79 (23.05)	61.80 (16.83)	56.01 (16.83)

Table 3.16: HRV ANOVA and paired t-test results between Normal, PC task and VR task

	ANOVA		**N & PC**		**N & VR**		**PC & VR**	
Parameter	F	p	t	p	t	p	t	p
NN Mean (ms)	6.27	0.00 **	1.98	0.18 ns	3.76	0.00 **	1.49	0.46 ns
SDNN (ms)	0.87	0.43 ns	-	-	-	-	-	-

NN Mean

A one-way repeated measures ANOVA was conducted to compare the values of NN Mean of Normal, PC, and VR condition. There was a significant effect of NN Mean, $F(2, 40) = 6.27, p = .00$. Three paired samples t-tests were used to make post hoc test comparisons using Bonferroni-test between conditions. A first paired samples t-test was conducted to compare NN Mean in Normal and PC conditions. There was no significant difference in the scores for Normal ($M = 712.85, SD = 98.16$) and PC ($M = 694.09, SD = 117.47$) conditions; $t(20) = 1.98, p = .18$. A second paired samples t-test was conducted to compare NN Mean in Normal and VR conditions. There was a significant difference in the scores for Normal ($M = 712.85, SD = 98.16$) and VR ($M = 678.46, SD = 86.92$) conditions; $t(20) = 3.76, p = .00$. A third paired samples t-test was conducted to compare NN Mean in PC and VR conditions. There was no significant difference in the scores for PC ($M = 694.09, SD = 117.47$) and VR ($M = 678.46, SD = 86.92$) conditions; $t(20) = 1.49, p = .46$.

SDNN

A one-way repeated measures ANOVA was conducted to compare the values of SDNN of Normal, PC, and VR condition. There was no significant effect of SDNN, $F(2, 40) = .87, p = .43$.

3.11.3 EDA

The mean and standard deviation of each EDA parameter, calculated from the mean value of 5-minute in the experiment 2 of each participant, are shown in Table 3.17. And, a result of one-way repeated measures ANOVA and Bonferroni-test as post hoc tests between Normal, PC task and VR task condition are shown in Table 3.18.

Mean EDA

A one-way repeated measures ANOVA was conducted to compare the values of Mean EDA of Normal, PC, and VR condition. There was no significant effect of Mean EDA,

Table 3.17: EDA parameter on each condition ($n = 21$). Values shown are M (SD).

Parameter	Normal	PC task	VR task
Mean EDA ($\boxtimes$ S)	3.35 (5.50)	4.06 (4.07)	4.07 (4.07)
Peak EDA ($\boxtimes$ S)	0.78 (1.78)	0.61 (1.08)	0.84 (1.08)

Table 3.18: EDA ANOVA and paired t-test results between Normal, PC task and VR task

	ANOVA		N & PC		N & VR		PC & VR	
Parameter	F	p	t	p	t	p	t	p
Mean EDA ($\boxtimes$ S)	0.42	0.66 ns	-	-	-	-	-	-
Peak EDA ($\boxtimes$ S)	0.35	0.71 ns	-	-	-	-	-	-

$F(2, 40) = .42, p = .66$.

Peak EDA

A one-way repeated measures ANOVA was conducted to compare the values of Peak EDA of Normal, PC, and VR condition. There was no significant effect of Peak EDA, $F(2, 40) = .35, p = .71$.

3.11.4 Construction of SSQ Symptoms by Physiological Data

OCS formulas will be composed of physiological data corresponding to SSQ scores. Therefore, SSQ scores are shown by coefficient values of NN Mean, SDNN, Mean EDA, and Peak EDA as predictors that will be calculated by multiple linear regression analysis. In this time, these predictors will be applied as a mean of the ratio between VR condition and Normal condition that was obtained in experiment 2. And, intercepts will be constant to zero. Table 3.19 shows a summary of the multiple linear regression to predict each SSQ symptom on NN Mean, SDNN, Mean EDA and Peak EDA values.

General discomfort

A multiple linear regression was calculated to predict the general discomfort based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was found ($F(4, 17) = 5.64, p = .00$), with an R^2 of .57. Physiological data that predicted general discomfort is equal to .52 (NN Mean) + .16 (SDNN) + .09 (Mean EDA) + -.00 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition. Consequently, NN Mean ($p = .34$), SDNN ($p = .65$), Mean EDA ($p = .48$), and Peak EDA ($p = .74$) were not

Table 3.19: The result of multiple linear regression by experiment 2 data

Parameter	F	p	R^2
General discomfort	5.64	0.00 **	0.57
Fatigue	2.88	0.05 ns	0.40
Headache	2.88	0.05 ns	0.40
Eye strain	3.60	0.03 *	0.46
Difficulty focusing	3.68	0.02 *	0.46
Increased salivation	3.46	0.03 *	0.45
Sweating	1.99	0.14 ns	0.32
Nausea	2.07	0.13 ns	0.33
Difficulty concentrating	3.34	0.03 *	0.44
Fullness of head	3.15	0.04 *	0.43
Blurred vision	5.53	0.00 **	0.57
Dizzy (eyes open)	1.98	0.14 ns	0.32
Dizzy (eyes closed)	1.88	0.16 ns	0.31
Vertigo	4.15	0.02 *	0.49
Stomach awareness	1.08	0.40 ns	0.20
Burping	-	-	-

significant predictors of the general discomfort.

Fatigue

A multiple linear regression was calculated to predict the fatigue based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was not found ($F(4, 17) = 2.88, p = .05$), with an R^2 of .40. Physiological data that predicted fatigue is equal to .97 (NN Mean) + -.22 (SDNN) + -.03 (Mean EDA) + .01 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition.

Headache

A multiple linear regression was calculated to predict the headache based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was not found ($F(4, 17) = 2.88, p = .05$), with an R^2 of .40. Physiological data that predicted headache is equal to -.17 (NN Mean) + .29 (SDNN) + -.04 (Mean EDA) + .00 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition.

Eye strain

A multiple linear regression was calculated to predict the eye strain based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was found ($F(4, 17) = 3.60, p = .03$), with an R^2 of .46. Physiological data that predicted eye strain is equal to .78 (NN Mean) + -.15 (SDNN) + -.11 (Mean EDA) + .01 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition. Consequently, NN Mean ($p = .04$) was significant predictor of the eye strain. On the other hand, SDNN ($p = .53$), Mean EDA ($p = .24$), and Peak EDA ($p = .09$) were not significant predictors of the eye strain.

Difficulty focusing

A multiple linear regression was calculated to predict the difficulty focusing based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was found ($F(4, 17) = 3.68, p = .02$), with an R^2 of .46. Physiological data that predicted difficulty focusing is equal to .24 (NN Mean) + .34 (SDNN) + -.07 (Mean EDA) + .01 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition. Consequently, NN Mean ($p = .57$), SDNN ($p = .22$), Mean EDA ($p = .51$), and Peak EDA ($p = .32$) were not significant predictors of the difficulty focusing.

Increased salivation

A multiple linear regression was calculated to predict the increased salivation based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was found ($F(4, 17) = 3.46, p = .03$), with an R^2 of .45. Physiological data that predicted increased salivation is equal to .88 (NN Mean) + -.11 (SDNN) + -.06 (Mean EDA) + .00 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition. Consequently, NN Mean ($p = .06$), SDNN ($p = .71$), Mean EDA ($p = .60$), and Peak EDA ($p = .91$) were not significant predictors of the increased salivation.

Sweating

A multiple linear regression was calculated to predict the sweating based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was not found ($F(4, 17) = 1.99, p = .14$), with an R^2 of .32. Physiological data that predicted sweating is equal to .34 (NN Mean) + .06 (SDNN) + -.05 (Mean EDA) + .01 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition.

Nausea

A multiple linear regression was calculated to predict the nausea based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was not found ($F(4, 17) = 2.07, p = .13$), with an R^2 of .33. Physiological data that predicted nausea is equal to .60 (NN Mean) + .09 (SDNN) + -.00 (Mean EDA) + -.00 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition.

Difficulty concentrating

A multiple linear regression was calculated to predict the difficulty concentrating based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was found ($F(4, 17) = 3.34, p = .03$), with an R^2 of .44. Physiological data that predicted difficulty concentrating is equal to .29 (NN Mean) + -.10 (SDNN) + -.02 (Mean EDA) + .01 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition. Consequently, NN Mean ($p = .19$), SDNN ($p = .49$), Mean EDA ($p = .65$), and Peak EDA ($p = .08$) were not significant predictors of the difficulty concentrating.

Fullness of head

A multiple linear regression was calculated to predict the fullness of head based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was found ($F(4, 17) = 3.15, p = .04$), with an R^2 of .43. Physiological data that predicted fullness of head is equal to .15 (NN Mean) + .10 (SDNN) + .00 (Mean EDA) + .01 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition. Consequently, NN Mean ($p = .62$), SDNN ($p = .61$), Mean EDA ($p = .97$), and Peak EDA ($p = .33$) were not significant predictors of the fullness of head.

Blurred vision

A multiple linear regression was calculated to predict the blurred vision based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was found ($F(4, 17) = 5.53, p = .00$), with an R^2 of .57. Physiological data that predicted blurred vision is equal to .33 (NN Mean) + .18 (SDNN) + -.14 (Mean EDA) + .02 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition. Consequently, Peak EDA ($p = .01$) was significant predictor of the blurred vision. On the other hand, NN Mean ($p = .33$), SDNN ($p = .42$), and Mean EDA ($p = .11$) were not significant predictors of the blurred

vision.

Dizzy (eyes open)

A multiple linear regression was calculated to predict the dizzy (eyes open) based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was not found ($F(4, 17) = 1.98, p = .14$), with an R^2 of .32. Physiological data that predicted dizzy (eyes open) is equal to .56 (NN Mean) + -.02 (SDNN) + -.06 (Mean EDA) + .00 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition.

Dizzy (eyes closed)

A multiple linear regression was calculated to predict the dizzy (eyes closed) based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was not found ($F(4, 17) = 1.88, p = .16$), with an R^2 of .31. Physiological data that predicted dizzy (eyes closed) is equal to .42 (NN Mean) + -.10 (SDNN) + -.06 (Mean EDA) + .01 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition.

Vertigo

A multiple linear regression was calculated to predict the vertigo based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was found ($F(4, 17) = 4.15, p = .02$), with an R^2 of .49. Physiological data that predicted vertigo is equal to .60 (NN Mean) + .04 (SDNN) + -.15 (Mean EDA) + .02 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition. Consequently, Peak EDA ($p = .02$) was significant predictor of the vertigo. On the other hand, NN Mean ($p = .19$), SDNN ($p = .88$), and Mean EDA ($p = .20$) were not significant predictors of the vertigo.

Stomach awareness

A multiple linear regression was calculated to predict the stomach awareness based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was not found ($F(4, 17) = 1.08, p = .40$), with an R^2 of .20. Physiological data that predicted stomach awareness is equal to .44 (NN Mean) + -.22 (SDNN) + .04 (Mean EDA) + .00 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition.

Burping

A multiple linear regression was calculated to predict the burping based on NN Mean, SDNN, Mean EDA, and Peak EDA values. A significant regression equation was not found ($F(4, 17) = nan, p = nan$), with an R^2 of nan. Physiological data that predicted burping is equal to .00 (NN Mean) + .00 (SDNN) + .00 (Mean EDA) + .00 (Peak EDA), where NN Mean, SDNN, Mean EDA, and Peak EDA values are calculated by the ratio from the mean value in the Normal condition.

3.12 Specific Composition of OCS

From the result of the analysis, the OCS equations are constructed by the coefficient value of physiological parameters. Equations 3.8.5, 3.8.6, and 3.8.7 are applied by symptom scores that are calculated by the coefficient values of each physiological data. And, each symptom score is set between 0 to 3. The coefficient values of each physiological data by each symptom, are summarized in Table 3.20.

Table 3.20: Computation of symptoms score with the coefficient value of physiological parameter

Parameter	**NN Mean**	**SDNN**	**Mean EDA**	**Peak EDA**
General discomfort	0.52	0.16	0.09	-0.00
Fatigue	0.97	-0.22	-0.03	0.01
Headache	-0.17	0.29	-0.04	0.00
Eye strain	0.78	-0.15	-0.11	0.01
Difficulty focusing	0.24	0.34	-0.07	0.01
Increased salivation	0.88	-0.11	-0.06	0.00
Sweating	0.34	0.06	-0.05	0.01
Nausea	0.60	0.09	-0.00	-0.00
Difficulty concentrating	0.29	-0.10	-0.02	0.01
Fullness of head	0.15	0.10	0.00	0.01
Blurred vision	0.33	0.18	-0.14	0.02
Dizzy (eyes open)	0.56	-0.02	-0.06	0.00
Dizzy (eyes closed)	0.42	-0.10	-0.06	0.01
Vertigo	0.60	0.04	-0.15	0.02
Stomach awareness	0.44	-0.22	0.04	0.00
Burping	0.00	0.00	0.00	0.00

3.13 Correlation between OCS Scores and SSQ Scores

The validity of OCS scores is considered by analyzing correlations between OCS scores and SSQ scores. The correlation between OCS Total Scores and SSQ Total Scores is calculated. OCS Total Score is calculated by obtained data in experiment 2 for each participant. In this time, obtained data in experiment 1 and experiment 2 are reused because further experiments could not be carried out and additional data could not be collected.

3.13.1 OCS scores and SSQ scores by Experiment 2 Data

Experiment 2 physiological data are applied to OCS formula to compare with SSQ scores. Table 3.22 shows correlation analysis parameters, and Fig. 3.25 shows data points of OCS scores and SSQ scores and their regression line.

Table 3.21: Correlation of OCS scores and SSQ scores by Experiment 2 Data

Parameter	*df*	*r*	*p*
Nausea Score	19	-0.09	0.71 ns
Oculomotor Score	19	0.13	0.58 ns
Disorientation Score	19	0.01	0.97 ns
Total Score	19	0.06	0.80 ns

Nausea Score

OCS Nausea Score is calculated by Equation 3.8.1 and physiological data obtained in experiment 2. There is no statistically significant relationship between OCS Nausea Score and SSQ Nausea Scores, $r(19) = -.09, p = .71$.

Oculomotor Score

OCS Oculomotor Score is calculated by Equation 3.8.2 and physiological data obtained in experiment 2. There is no statistically significant relationship between OCS Oculomotor Score and SSQ Oculomotor Scores, $r(19) = .13, p = .58$.

Disorientation Score

OCS Disorientation Score is calculated by Equation 3.8.3 and physiological data obtained in experiment 2. There is no statistically significant relationship between OCS Disorientation Score and SSQ Disorientation Scores, $r(19) = .01, p = .97$.

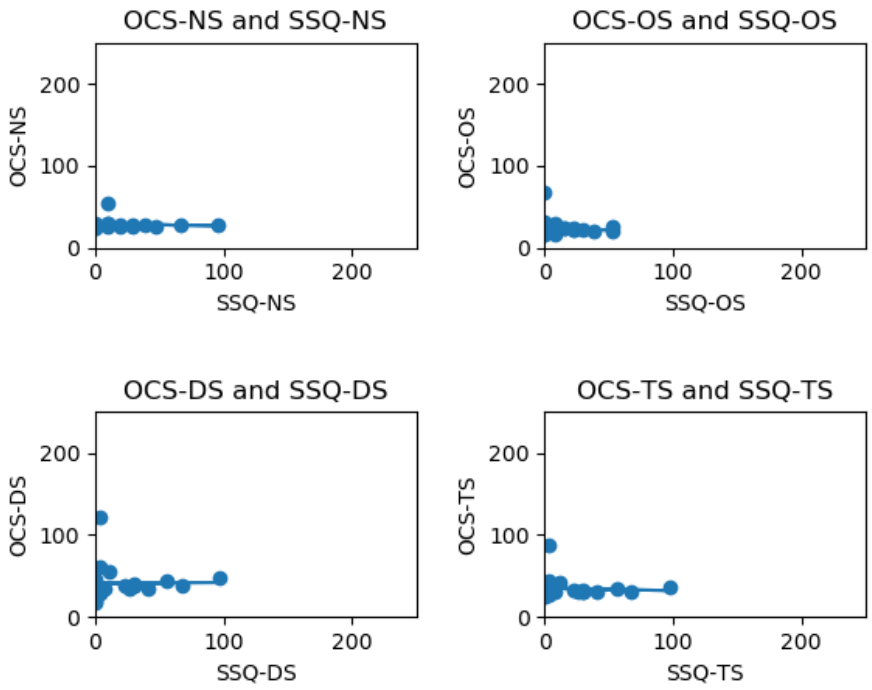


Fig. 3.25: Correlation between OCS scores and SSQ Total Scores by experiment 2

Total Score

OCS Total Score is calculated by Equation 3.8.4 and physiological data obtained in experiment 2. There is no statistically significant relationship between OCS Total Score and SSQ Total Scores, $r(19) = .06, p = .80$.

3.13.2 OCS scores and SSQ scores by Experiment 1 Data

Experiment 1 physiological data are applied to OCS formula to compare with SSQ scores. The analysis is tested to confirm that the OCS formula can be adapted under different conditions. Table 3.22 shows correlation analysis parameters, and Fig. 3.26 shows data points of OCS scores and SSQ scores and their regression line.

Table 3.22: Correlation of OCS scores and SSQ scores by Experiment 1 Data

Parameter	*df*	*r*	*p*
Nausea Score	14	-0.15	0.57 ns
Oculomotor Score	14	0.43	0.10 ns
Disorientation Score	14	0.27	0.31 ns
Total Score	14	0.41	0.11 ns

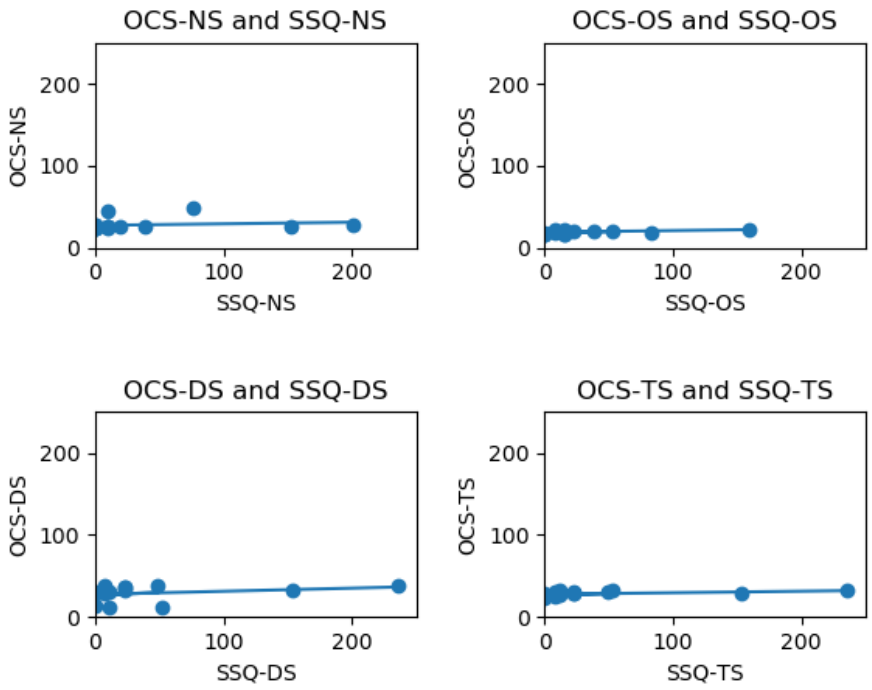


Fig. 3.26: Correlation between OCS scores and SSQ Total Scores by experiment 1

Nausea Score

OCS Nausea Score is calculated by Equation 3.8.1 and physiological data obtained in experiment 1. There is no statistically significant relationship between OCS Nausea Score and SSQ Nausea Scores, $r(14) = -.15, p = .57$.

Oculomotor Score

OCS Oculomotor Score is calculated by Equation 3.8.2 and physiological data obtained in experiment 1. There is no statistically significant relationship between OCS Oculomotor Score and SSQ Oculomotor Scores, $r(14) = .43, p = .10$.

Disorientation Score

OCS Disorientation Score is calculated by Equation 3.8.3 and physiological data obtained in experiment 1. There is no statistically significant relationship between OCS Disorientation Score and SSQ Disorientation Scores, $r(14) = .27, p = .31$.

Total Score

OCS Total Score is calculated by Equation 3.8.4 and physiological data obtained in experiment 1. There is no statistically significant relationship between OCS Total Score

and SSQ Total Scores, $r(14) = .41, p = .11$.

3.14 Reflection of Experiment 2

Experiment 2 was aimed to confirm the working of Cybatica, reexamining experiment 1, and collecting data to construct OCS equations.

First, Cybatica could be used to measure physiological data with Empatica E4 because physiological data and analysis data were shown as charts and numerical data in real-time. Cybatica could record multiple sessions in one connection. For example, firstly Cybatica captures base-session (Normal condition), and, then captures data-session (PC task, and VR task conditions) seamlessly. Empatica E4 real-time App, on the other hand, has to re-connect for each session. However, the connection to Empatica E4 was sometimes hard because Android Cybatica could not detect the Empatica E4 over a long time. On the other hand, iOS Cybatica could smoothly connect to the Empatica E4 but E4link SDK could not receive physiological data from the Empatica E4 after the version of iOS 13.

Secondly, all SSQ scores (Nausea Score, Oculomotor Score, Disorientation Score, and Total Score) showed a significant difference between VR task and PC task conditions similar to experiment 1. Therefore, PolyWorld can claim to be a valid experimental activity.

Next, in the result of physiological data, NN Mean and SDNN were analyzed. NN Mean showed a significant difference between Normal and VR task conditions; which is the same as the result of experiment 1. However, SDNN showed no significant difference in ANOVA analysis; although SDNN showed a significant difference in ANOVA analysis in experiment 1. From this result, SDNN may not be considered a predictable metric. In the analysis of EDA, Mean EDA and Peak EDA were used instead of the EDA Explorer's parameters to embed the analysis of EDA to the Cybatica in the experiment. Mean EDA and Peak EDA did not show a significant difference between all conditions though. Peak EDA showed a significant difference in experiment 1 but did not show a significant difference in experiment 2. In either value, noises that are filtered by the EDA Explorer in experiment 1 should be considered. Therefore, OCS parameter in EDA should be changed.

Finally, OCS scores were analyzed. The proposed equations could not show a significant relationship between all OCS scores and SSQ scores. The result means the proposed OCS equations are difficult to apply to the forecasting of the onset of cybersickness using physiological data.

3.15 OCS Composited by Experiment 1 Data

From the analysis of experiment 2 data, it was found that the OCS metric was not a valid indicator of cybersickness. This finding is stated because the data of experiment 2 showed no significant difference when tested by ANOVA between each of the experiment 2 conditions. And, the result could not be considered valid when there is a difference between conditions. Therefore, OCS was next analyzed with experiment 1 data and the correlations tested.

3.15.1 Multiple Linear Regression

OCS equations are next calculated by multiple linear regression analyses that are the same as when determined by experiment 2 data. Table 3.23 shows analysis parameters by regression, and Table 3.24 shows coefficient values of parameters.

Table 3.23: The result of multiple linear regression by experiment 1 data

Parameter	F	p	R^2
General discomfort	10.55	0.00 **	0.78
Fatigue	13.87	0.00 **	0.82
Headache	21.24	0.00 **	0.88
Eye strain	12.44	0.00 **	0.81
Difficulty focusing	14.63	0.00 **	0.83
Increased salivation	8.99	0.00 **	0.75
Sweating	5.05	0.01 *	0.63
Nausea	13.01	0.00 **	0.81
Difficulty concentrating	10.77	0.00 **	0.78
Fullness of head	9.09	0.00 **	0.75
Blurred vision	18.53	0.00 **	0.86
Dizzy (eyes open)	8.86	0.00 **	0.75
Dizzy (eyes closed)	34.12	0.00 **	0.92
Vertigo	34.64	0.00 **	0.92
Stomach awareness	210.18	0.00 **	0.99
Burping	8.96	0.00 **	0.75

Table 3.24: Computation of symptoms score with the coefficient value of physiological parameter

Parameter	**NN Mean**	**SDNN**	**Mean EDA**	**Peak EDA**
General discomfort	-0.45	1.33	0.08	-0.06
Fatigue	1.50	-0.71	0.09	-0.02
Headache	-1.32	1.73	0.09	-0.04
Eye strain	0.82	0.07	0.10	-0.07
Difficulty focusing	-0.56	0.89	0.09	-0.05
Increased salivation	-1.76	2.43	0.06	-0.04
Sweating	-0.28	0.74	0.08	-0.01
Nausea	-0.06	0.22	0.11	-0.05
Difficulty concentrating	-1.68	2.20	0.08	-0.04
Fullness of head	-0.95	2.04	0.06	-0.01
Blurred vision	-0.89	1.21	0.09	-0.04
Dizzy (eyes open)	-0.21	0.46	0.09	-0.05
Dizzy (eyes closed)	-1.21	1.47	0.09	-0.04
Vertigo	-1.24	1.52	0.09	-0.04
Stomach awareness	-1.47	1.69	0.09	-0.04
Burping	-3.20	4.07	0.03	-0.03

3.15.2 Correlation

First, experiment 1 physiological data are applied to new OCS formula to compare with SSQ scores. The analysis is tested to confirm that the new OCS formula can be adapted under different conditions. Table 3.25 shows correlation analysis parameters, and Fig. 3.27 shows data points of OCS scores and SSQ scores and their regression line.

Table 3.25: Correlation of new OCS scores and SSQ scores by Experiment 1 Data

Parameter	*df*	*r*	*p*
Nausea Score	14	-0.01	0.97 ns
Oculomotor Score	14	-0.18	0.49 ns
Disorientation Score	14	-0.08	0.78 ns
Total Score	14	-0.08	0.77 ns

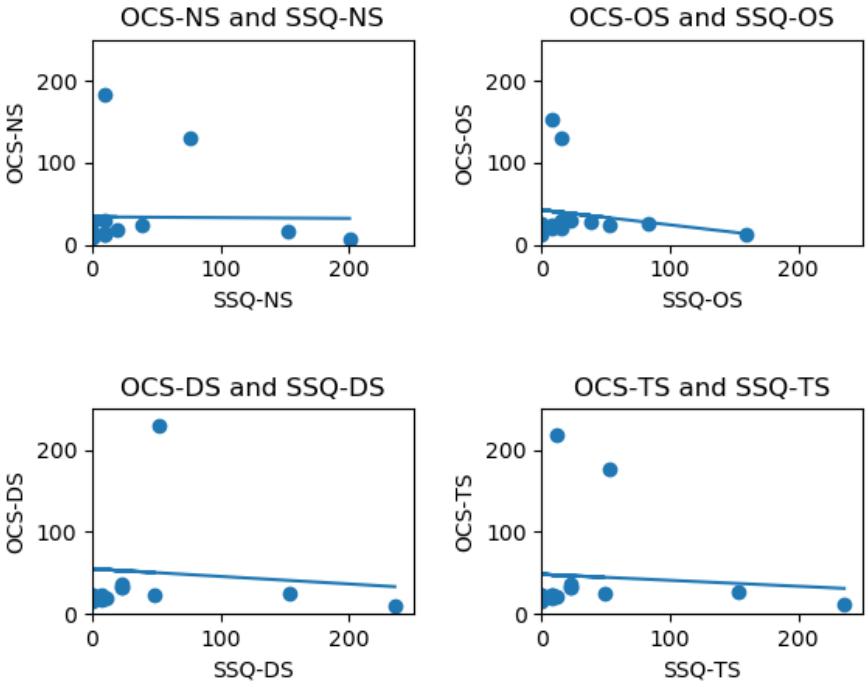


Fig. 3.27: Correlation between new OCS scores and SSQ Total Scores by experiment 1

Next, experiment 2 physiological data are applied to the new OCS formula to compare with SSQ scores. The analysis is tested to confirm that the new OCS formula can be adapted under different conditions. Table 3.26 shows correlation analysis parameters, and Fig. 3.28 shows data points of OCS scores and SSQ scores and their regression line.

Table 3.26: Correlation of new OCS scores and SSQ scores by Experiment 2 Data

Parameter	*df*	*r*	*p*
Nausea Score	19	-0.02	0.93 ns
Oculomotor Score	19	-0.05	0.84 ns
Disorientation Score	19	-0.08	0.72 ns
Total Score	19	-0.05	0.84 ns

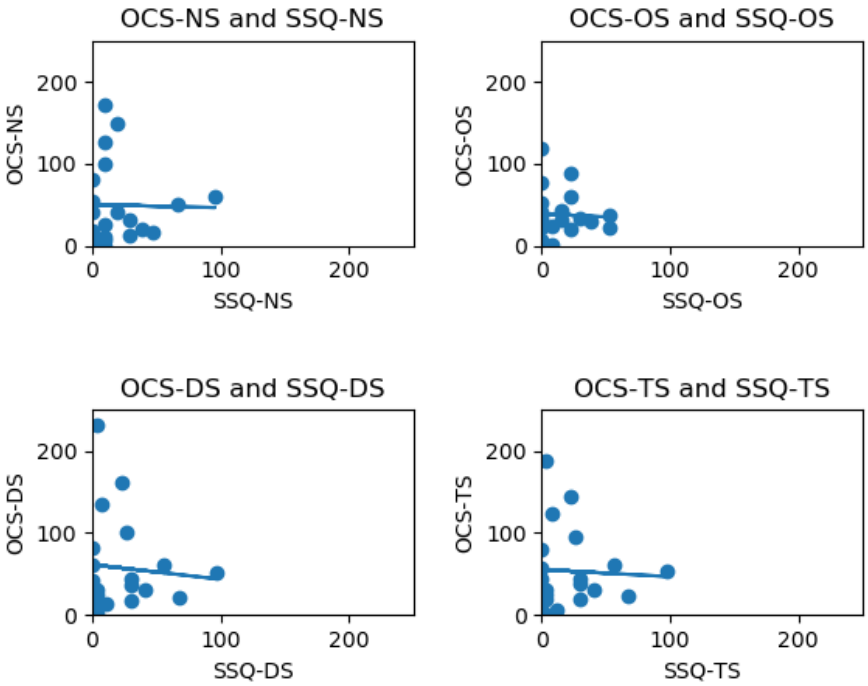


Fig. 3.28: Correlation between new OCS scores and SSQ Total Scores by experiment 2

3.15.3 Reflection

OCS equations were re-calculated by experiment 1 physiological data and it was analyzed by the correlation between OCS scores and SSQ scores of each experiment. As a result, a significant regression equation was found in all symptoms in the multiple linear regression analyses. However, actual OCS scores did not show a significant correlation when compared with SSQ scores in the correlation analysis.

Chapter 4 Discussion

4.1 General Observation

In the development of experiment 1, it was first needed to ascertain an affordable, portable and valid instrument for recording physiological data. Consumer devices such as the Apple Watch (measuring SDNN as a proxy for HRV) currently do not record a wide enough range of data. On the other hand, expensive medical equipment such as Polymate II requires specialized operation. This study determined that the Empatica E4 physiological sensor, along with its smartphone App and Empatica' s online data acquisition portal, provided an eco-system that could be utilized with minimal specialized training.

The next challenge was to consider the metrics associated with cybersickness. It was necessary to determine how to interpret the metrics' associated data in order to ascertain if participants moved from their so-called normal state to a state which may potentially lead to cybersickness. In addition, data representation and analysis had to be understood by non-specialists; ideally, educators and trainers considering adopting VR.

A custom-designed activity called PolyWorld was then developed in Unity and ported as a PC activity and as a VR activity for collecting experimental data. The tasks in both the PC condition and the VR condition were identical; i.e. maneuver around PolyWorld to locate and collect five objects within a five-minute time frame. In experiment 1 sixteen participants (15 male, 1 female) volunteered and data was collected in three states (Resting, PC condition, VR condition) using the Empatica E4 worn by each participant. HRV and EDA data were collected in the Normal (i.e. Resting) condition, the PC task condition, and the VR task condition. Time domain analysis, Frequency domain analysis and Non-linear analysis of HRV, the parameters of SCR and SCL of EDA, and each SSQ score were then statistically analyzed. Subsequently, experiment 1 only succeeded in determining a valid instrument (i.e. the Empatica E4) and specific metrics for analysis.

Then, the indicator of cybersickness named the Onset of Cybersickness (OCS) equations (OCS Nausea Score, Oculomotor Score, Disorientation Score, and Total Score) were proposed. OCS scores were based on Simulator Sickness Questionnaire [13] scores to be indicated by physiological data. The indicators aimed to scale the same as SSQ scores (an established formula over many studies) so that OCS could be compared with SSQ.

Next, a Cybatica App that connects with Empatica E4, and visualizes physiological

data and analysis data, was designed, programmed, and developed for Android / iOS as a real-time cybersickness monitoring method. Firstly, the application was designed with the aim of embedding OCS as a method of forecasting and visualizing cybersickness. Consequently, physiological data, analysis data, and associated line charts were shown and updated in the Cybatica application's screens in real-time.

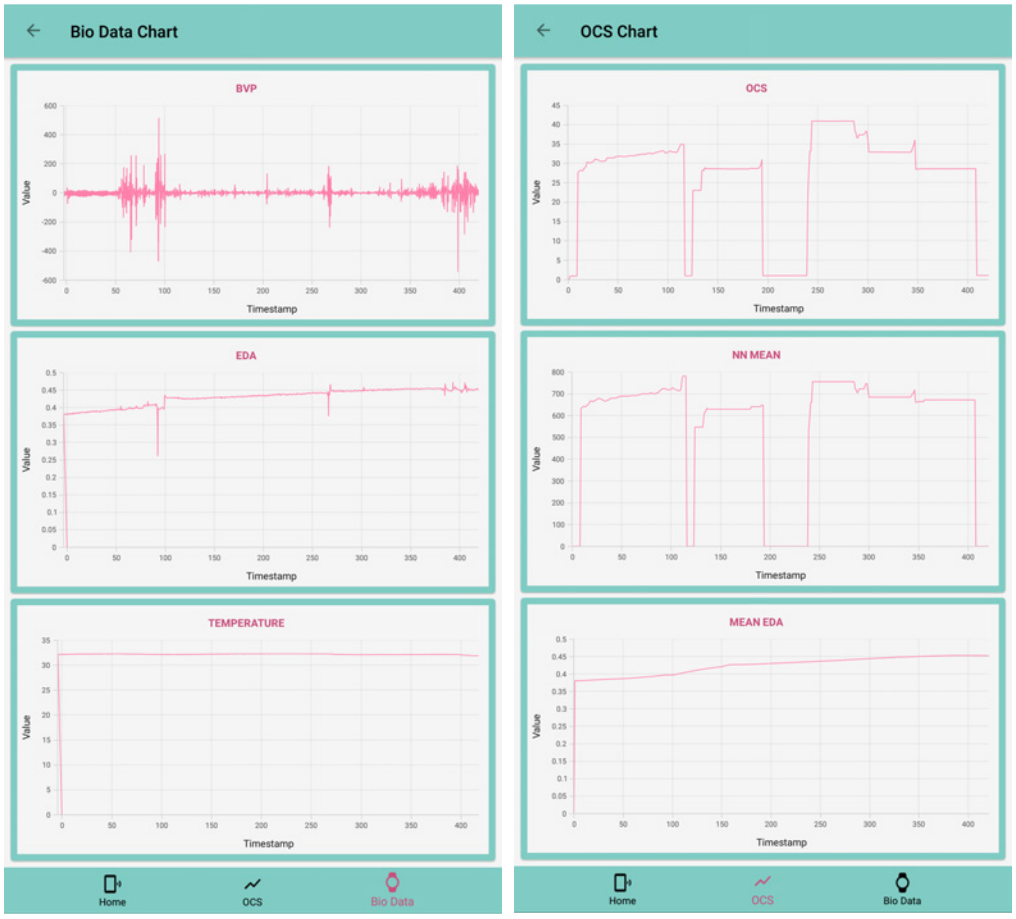
To collect data to further develop OCS, experiment 2 was carried out with twenty-one participants (18 male, 3 female). The experiment was carried out in Polyworld, the same as experiment 1, but the experimental procedure was modified: 1. The Normal condition was undertaken sitting, and PC task and the VR task were undertaken standing; 2. The rest time between tasks was increased to 5 minutes; 3. The method to obtain physiological data was changed to Cybatica from Empatica E4 eco-systems; 4. The order of the procedure of PC and VR tasks were inversed. In addition, the measuring method was changed to the Cybatica App from E4 real-time to confirm the work of Cybatica. Physiological measures, BVP, IBI, EDA, and Temperature were captured. And, NN Mean, SDNN, Mean EDA, and Peak EDA in one minute's window were captured as analysis measures at the same time.

In the analysis of experiment 2, firstly, SSQ scores were analyzed to confirm the validation of activity by *t*-test; similar to experiment 1. The result meant PolyWorld was a valid experimental activity, due to a significant difference between PC task and VR task in all scores which were the same as the result of experiment 1. Secondly, differences in physiological data such as NN Mean, SDNN, Mean EDA, and Peak EDA were analyzed by ANOVA between three conditions (Normal, PC task, and VR task conditions). The result showed a significant difference in only NN Mean between Normal condition and VR task condition. The difference of experiment 2 EDA with experiment 1 EDA was that noise had been filtered in experiment 1 but not in experiment 2. This difference should be considered for embedding OCS to Cybatica but it was not possible to apply EDA Explorer's function of filtering noise; this is a limitation of the study. In addition, differences in participants' past experiences of VR application might affect the data. For instance, the result of pre-surveys (see Appendix C and E) showed experiment 2 participants had more VR experiences than experiment 1 participants. Finally, a specific composition of OCS was calculated by multiple linear regressions using physiological data. And, the correlation between OCS scores and SSQ scores was analyzed as an evaluation of the validity of OCS using composited OCS formula. As a result, the proposed equations could not show a significant relationship between all OCS scores and SSQ scores.

In addition, a new OCS was composited by experiment 1 data that showed a significant

difference, and then tested for correlation with SSQ scores. As a result, a significant regression equation was found in all symptoms in the multiple linear regression analyses. However, actual OCS scores did not show a significant correlation when compared with SSQ scores in the correlation analysis.

Contrary to the hypothesis "Cybersickness can be estimated from physiological data captured in realtime during a VR experience" a proposed Onset of Cybersickness (OCS) formula was difficult to apply to the forecasting of the onset of cybersickness using physiological data when users are immersed in Virtual Reality. Although OCS scores did not show a significant relationship between SSQ scores in the statistical analysis, how the OCS formula behaves when applied to Cybatica was tested as a case study of the application of OCS. For instance, real time Bio Data Chart and OCS Chart Cybatica screen captures, which are actually working to apply the OCS formula (using Section 3.12's formula), are shown in Fig. 4.1. Real time BVP, EDA, and Temperature data are shown in Bio Data Chart. And, real time OCS (Total Score), NN Mean, and Mean EDA are shown in OCS Chart. The sudden drops within the chart shows lost physiological data; most likely due to a poor connection. The sequence within the chart shows resting until timestamp 60, carrying out VR PolyWorld task during timestamp 60 to timestamp 360, and resting from timestamp 360 to the end of the data capture. Referring to Fig. 4.2 BVP, EDA, NN Mean, and Mean EDA data overlaid with OCS data, OCS shows a chart similar to NN Mean but OCS looks like it is adjusted with increasing Mean EDA. OCS shows dynamic changes with fluctuations of physiological data over time in the Figures. This detailed fluctuation is difficult to be captured by only a subjective evaluation. Therefore, OCS has the potential of a valuable indication of real-time forecastable cybersickness if OCS can be modified by reliable data and valid construction.



(a) Bio Data Chart (b) OCS Chart

Fig. 4.1: Bio Data Chart and OCS Chart Cybatica screen captures

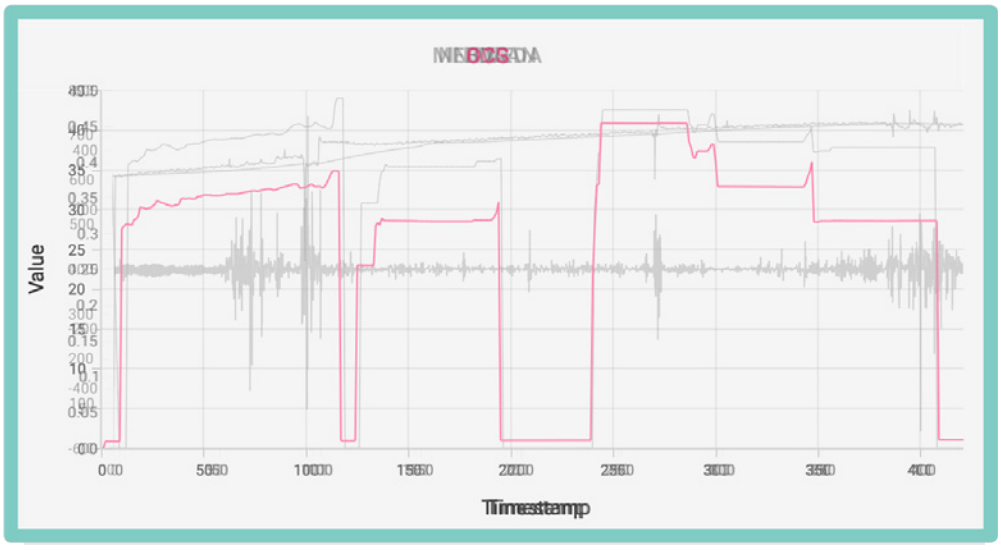


Fig. 4.2: NN Mean, Mean EDA, BVP, and EDA data overlaid with OCS data

4.2 Outcomes

The outcomes from the research, the measuring method of cybersickness, the way of analysis of cybersickness, the monitoring method of cybersickness and physiological data in VR activities, and forecasting the onset of cybersickness using physiological data, can be shown. OCS as a forecastable indicator of cybersickness is a useful guide because OCS observes a user' s health during VR experiences, and it is able to warn about cybersickness before it becomes serious. OCS may apply for not only forecasting cybersickness but also for forecasting motion sickness, and observing the physiological impact of VR activities. However, after detailed analysis and as all OCS scores were analyzed, the proposed equations could not show a significant relationship between all OCS scores and SSQ scores in this research.

4.3 Limitation

As a limitation, less data is definitely a problem with few participants. In this research, there were only 16 and 21 participants in experiment 1 and experiment 2, respectively. These are statistically low. Also, gender, age, and past VR experience were not evenly obtained. Therefore, more data should be collected over a longer study.

Next, the construction of OCS was applied to multiple linear regression analyses in this research. However, the result could not be shown effectively. Therefore, the combination of descriptive variables by physiological parameters should be modified in multiple linear regression. And, the reproduction of the result is difficult because an insufficient number of data and physiological data are used. Therefore, this result might change depending on the data.

In addition, measuring cybersickness from the composed OCS formulas in the research proved to be difficult because the OCS scores of the high SSQ score participants were the same as the SSQ scores of the low SSQ score participants in experiment 1. Subsequently, the possibility of a gap between subjective reports by the participant and objective results by physiological data should be considered. Therefore, a further determination of objective criteria for cybersickness will be required in order to evaluate cybersickness via objective methods that are associated with subjective methods.

In addition, E4link (SDK for iOS) had a technical problem that could not receive physiological data from the Empatica E4 on iOS 13; which was an update for the iPad during the latter stages of the research period. This was not a problem when used on Android or iOS prior to iOS 13. The problem should be fixed by the provider or develop a new SDK for iOS 13.

Chapter 5 Conclusion

5.1 Summary

The research showed the progress of developing a forecastable cybersickness indicator called the Onset of Cybersickness (OCS). First, the experiment was carried out to analyze physiological data using Empatica E4 and its eco-system in PC and VR activities. Next, OCS was proposed and an Android / iOS application monitoring real-time cybersickness called Cybatica was developed and OCS was embedded. Then, Cybatica was used in the second experiment, and specific OCS formulas were developed and evaluated. As a result, the research illustrated one approach of forecasting cybersickness using physiological data, but the data analysis revealed it was not a reliable method to forecast cybersickness. Therefore, the resulting evaluation was inconclusive and additional research is proposed.

5.2 Future work

As for future work, the concept of the Onset of Cybersickness will be analyzed by trying other statistical methods using physiological data because the OCS equations calculated by multiple linear regression did not show significant relationships with SSQ scores in this research. And, the analysis should be carried out with more data. As for the experimental material, the combination of Cybatica and Empatica E4 is an effective method that could measure physiological data in the experiment. It should be updated with valid physiological parameters to measure physiological data.

Acknowledgement

Lots of appreciation to Prof. Vallance, who is the supervisor of this research, and Prof. Fujino, Prof. Mikami, Prof. Sato, and Prof. Tominaga who provided valuable opinions in the presentations.

Publications

Peer Reviewed

[1] T. Magaki and M. Vallance, "Measuring Reduction Methods for VR Sickness in Virtual Environments," *The International Journal of Virtual and Personal Learning Environments (IJVPLE)*, Vol. 7, No. 2, 2017. doi:10.4018/IJVPLE.2017070103.

[2] T. Magaki and M. Vallance, "Developing an Accessible Evaluation Method of VR Cybersickness," in *Proceedings of 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Osaka, Japan, 2019, pp. 1072-1073. doi:10.1109/VR.2019.8797748.

[3] T. Magaki and M. Vallance, "Real time Monitoring Method for Cybersickness using Physiological Signals," in *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology (VRST)*, Parramatta, NSW, Australia, 2019. doi:10.1145/3359996.3364700.

Non-Peer Reviewed

[1] T. Magaki and M. Vallance, "Real-time Forecasting of Cybersickness," in *the 24th Annual Conference of Virtual Reality Society of Japan*, Osaka, Japan, 2019.

Bibliography

[1] S. Davis, K. Nesbitt and E. Nalivaiko, "A Systematic Review of Cybersickness," in *Proceeding of the Conference on Interactive Entertainments (IE)*, pp. 1-9, 2014. doi:10.1145/2677758.2677780

[2] S. Weech, S. Kenny, M. Barnett-Cowan, "Presence and Cybersickness in Virtual Reality Are Negatively Related: A Review," *Frontiers in Psychology*, vol. 10, pp. 158, 2019. doi:10.3389/fpsyg.2019.00158.

[3] K. M. Stanney, R. S. Kennedy, and J. M. Drexler, "Cybersickness is not simulator sickness," in *Proceedings of the Human Factors and Ergonomics Society annual meeting*, vol. 41, pp. 1138 - 1142, Los Angeles, CA: Sage Publications, Oct. 1997, doi:10.1177/107118139704100292.

[4] S. Tachi, M. Sato and M. Hirose, *Study of Virtual Reality*, Tokyo, Japan: Corona-Sha, 2011.

[5] T. Seno and H. Suzuki, *What is Vection !?*, Tokyo, Japan: Kyoritsu Shuppan, 2017.

[6] S. Bouchard, G. Robillard, R. Renaud, and F. Bernier, "Exploring new dimensions in the assessment of virtual reality induced side effects," *Journal of Computer and Information Technology*, vol. 1 pp. 20-32, 2011.

[7] G. Riccio, and T. Stoffregen, "An Ecological Theory of Motion Sickness and Postural Instability," *Ecological Psychology (ECOL PSYCHOL)*, vol. 3. pp. 195-240, 1991. doi:10.1207/s15326969eco0303_2.

[8] A. S. Fernandes and S. K. Feiner, "Combating VR sickness through subtle dynamic field-of-view modification," in *Proceeding of the IEEE International Symposium on 3D User Interfaces (3DUI)*, Greenville, South Carolina, USA 2016, pp. 201-210, 2016.

[9] W. D. Whittinghill, B. Ziegler, J. Moore and T. Case "Nasum virtualis: A simple technique for reducing simulator sickness," 2015 [Online]. Available: http://www.purdue.edu/newsroom/releases/2015/Q1/virtual-nose-may-reduce-simulator-sickness-in-video-games.html. [Accessed: 12-November-2018].

[10] D. Tambovtsev, N. Floksy and O. Peshé, "How to Avoid the Effect of Motion Sickness in VR," VRScout, Vol. 1, No. 7, 2016. [Online]. Available: https://vrscout.com/news/avoid-motion-sickness-developing-for-vr/. [Accessed: May 12, 2017].

[11] S. Liu, S. Liu, N. Yu, L. Chan, Y. Peng, W. Sun, and M. Y. Chen, "PhantomLegs: Reducing Virtual Reality Sickness Using Head-Worn Haptic Devices," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Osaka, Japan, 2019,

pp. 817-826. doi:10.1109/VR.2019.8798158.

[12] T. Magaki and M. Vallance, "Measuring Reduction Methods for VR Sickness in Virtual Environments," *The International Journal of Virtual and Personal Learning Environments (IJVRPLE)*, vol. 7, no. 2, pp. 27-42, 2017. doi:10.4018/IJVPLE.2017070103.

[13] R. S. Kennedy, N. E. Lane, K. S. Berbaum and M. G. Lilienthal, "Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness," *Tthe International Journal of The International Journal of Aviation Psychology*, vol. 3, no. 3, pp. 203-220, 1993.

[14] S. L. Ames, J. S. Wolffsohn, and N. A. McBrien, "The development of a symptom questionnaire for assessing virtual reality viewing using a head-mounted display," *Optometry and Vision Science*, vol. 82, no. 3, pp. 168-176, 2005.

[15] N. McHugh, S. Jung, S. Hoermann, and R. W. Lindeman. "Investigating a Physical Dial as a Measurement Tool for Cybersickness in Virtual Reality," in *25th ACM Symposium on Virtual Reality Software and Technology (VRST ' 19)*, Parramatta, NSW, Australia, 2019. doi:10.1145/3359996.3364259.

[16] B. Keshavarz and H. Hecht, "Validating an Efficient Method to Quantify Motion Sickness," *Human Factors* vol. 53, no. 4, pp. 415 - 426, 2011. doi: 10.1177/0018720811403736.

[17] M. S. Dennison, A. Z. Wisti, M. D'Zumura, "Use of physiological signals to predict cybersickness," *Display*, vol. 44, pp. 42-52, 2016. doi:10.1016/j.displa.2016.07.002.

[18] A. Gurcia-Agundez, C. Reuter, P. Caserman, R. Konrad and S. Göbel. "Identifying Cybersickness through Heart Rate Variability alterations," *The International Journal of Virtual Reality (IJVR)*, vol. 19, no. 1 pp. 1-10, 2019.

[19] F. T. Sun, C. Kuo, H. T. Cheng, S. Buthpitiya, P. Collins and M. Griss, "Activity-aware Mental Stress Detection Using Physiological Sensors," in *Mobile Computing, Applications, and Services*, M. Gris, G. Yang, Eds. Berlin: Springer, pp. 211-230. doi:10.1007/978-3-642-29336-8_12.

[20] H. Kantz, J. Kurths, and G. Mayer-Kress, Eds., *Nonlinear Analysis of Physiological Data*, 1st ed, Berlin: Springer, 1998.

[21] N. Selvaraj, A. K. Jaryal, J. Santhosh, K. K. Deepak and S. Anand, "Assessment of heart rate variability derived from finger-tip photoplethysmography as compared to electrocardiography," *The Journal of Medical Engineering and Technology*, vol. 33, no. 6, pp. 479-484, 2008. doi:10.1080/03091900701781317.

[22] S. A. Taylor, N. M. Jaques, W. Chen, S. Fedor, A. Sano and R. W. Picard, "Automatic Identification of Artifacts in Electrodermal Activity Data",

in *Proceeding of the International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, August 2015. pp. 1934-1937. DOI:10.1109/EMBC.2015.7318762.

[23] P. Guzik, J. Piskorski, T. Krauze, R. Schneider, K. H. Wesseling, A. Wykretowicz, and H. Wysocki, "Correlations between the Poincaré Plot and Conventional Heart Rate Variability Parameters Assessed during Paced Breathing," *The Journal of Physiological Sciences* vol. 57, no. 1, pp. 63-71, doi:10.2170/physiolsci.RP005506.

[24] M. B. Tayel, and A. I. Eslam, "Poincaré Plot for Heart Rate Variability," *The International Journal of Biomedical and Biological Engineering*, vol 9, no. 9, 2015.

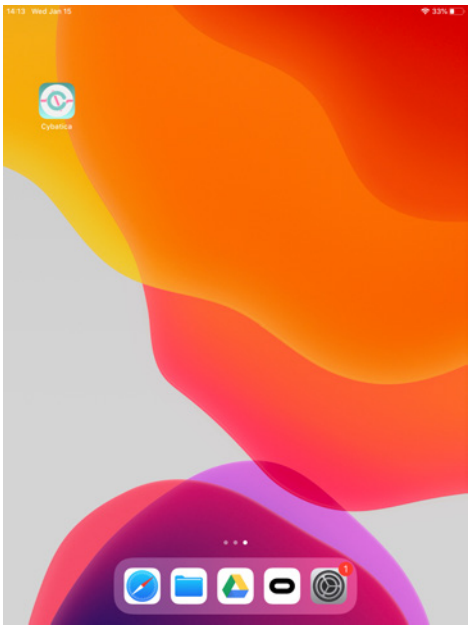
Appendix A Cybatica Application

(a) Android icon

(b) iOS icon

Fig. A.1: Cybatica icons

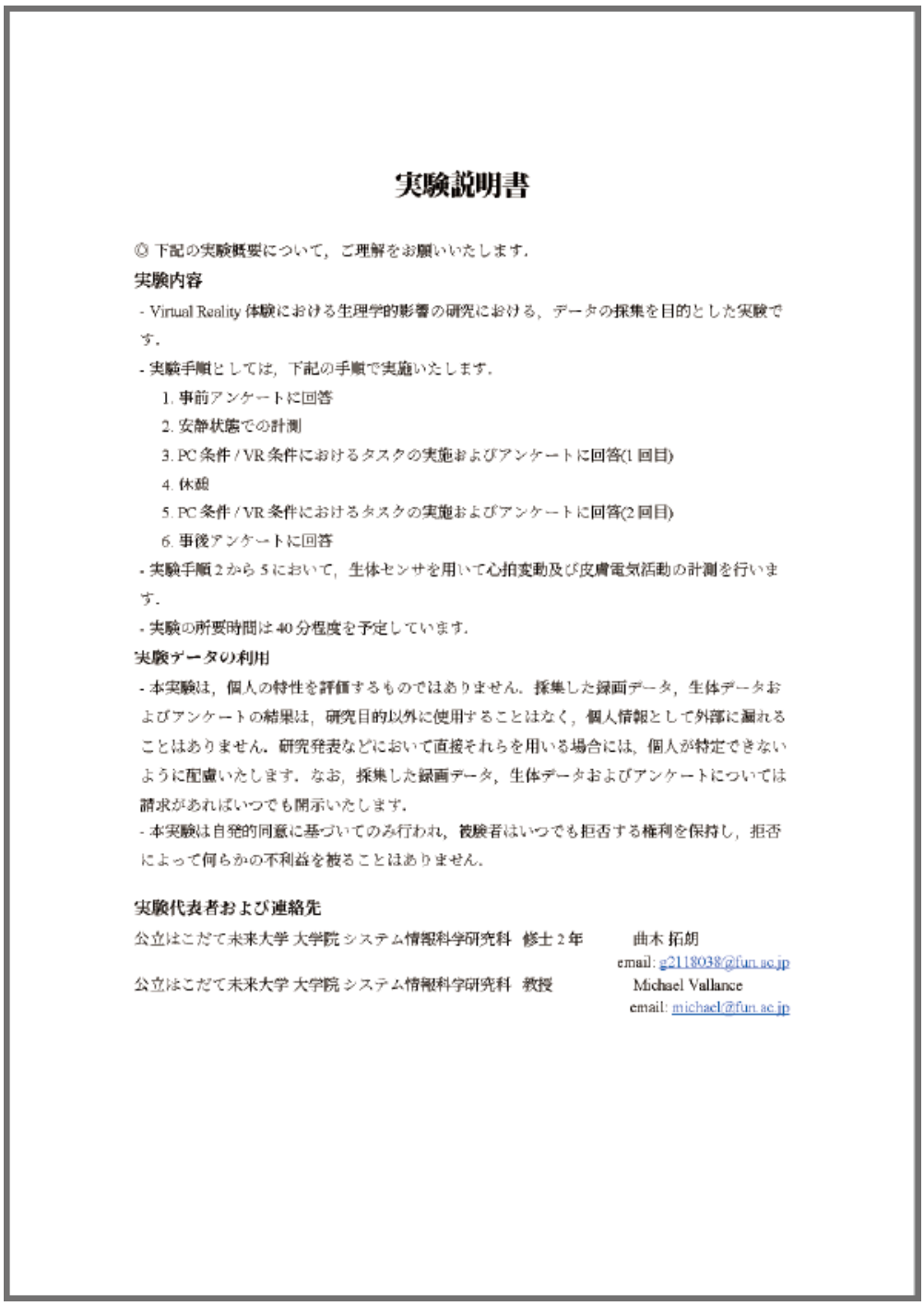
(a) Android



(b) iOS

Fig. A.2: Application Home screenshots

Appendix B Consent Form



実験説明書

◎ 下記の実験概要について，ご理解をお願いいたします．

実験内容

- Virtual Reality 体験における生理学的影響の研究における，データの採集を目的とした実験です．
- 実験手順としては，下記の手順で実施いたします．
  1. 事前アンケートに回答
  2. 安静状態での計測
  3. PC 条件 / VR 条件におけるタスクの実施およびアンケートに回答(1 回目)
  4. 休憩
  5. PC 条件 / VR 条件におけるタスクの実施およびアンケートに回答(2 回目)
  6. 事後アンケートに回答
- 実験手順 2 から 5 において，生体センサを用いて心拍変動及び皮膚電気活動の計測を行います．
- 実験の所要時間は 40 分程度を予定しています．

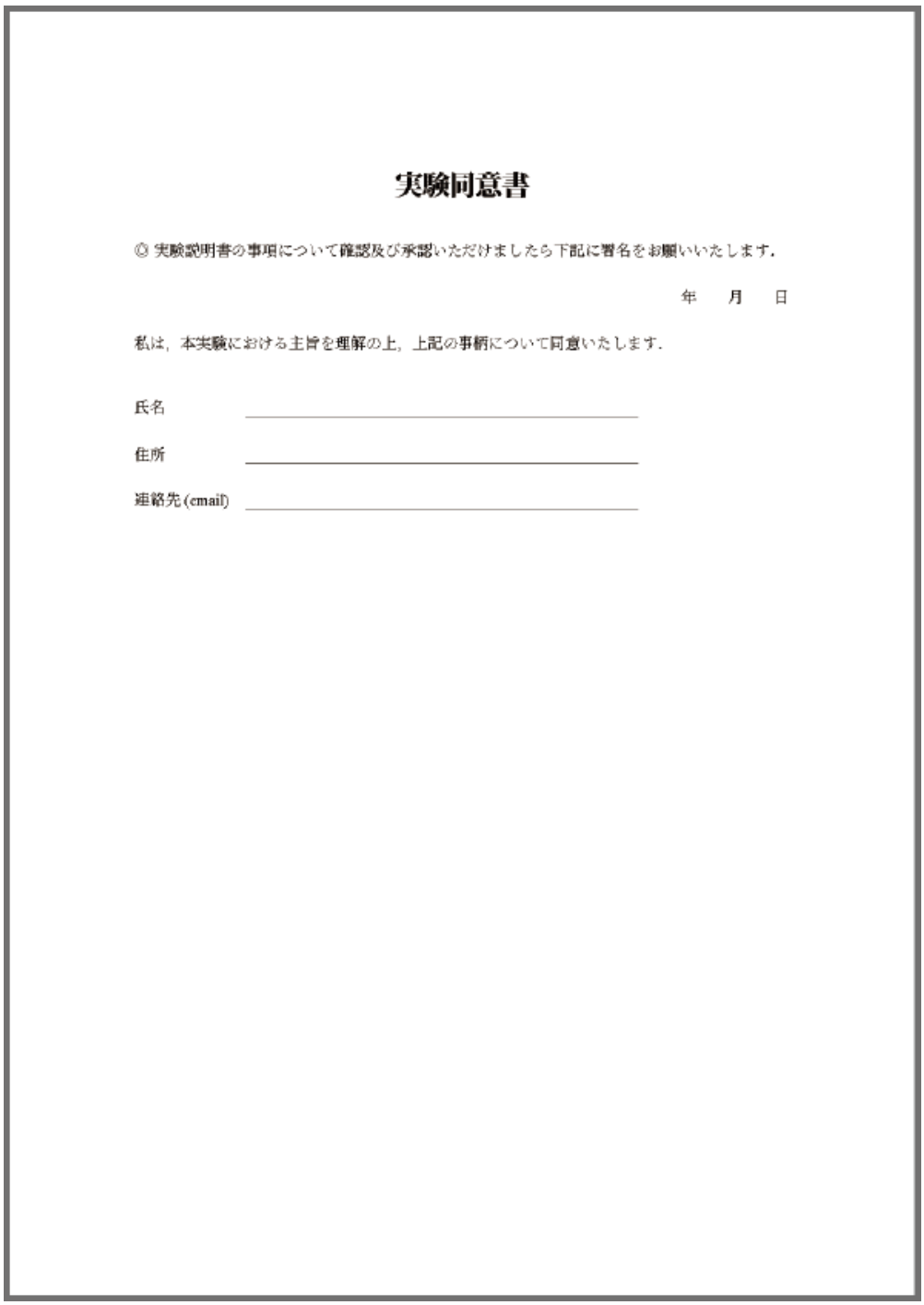
実験データの利用

- 本実験は，個人の特性を評価するものではありません．採集した録画データ，生体データおよびアンケートの結果は，研究目的以外に使用することはなく，個人情報として外部に漏れることはありません．研究発表などにおいて直接それらを用いる場合には，個人が特定できないように配慮いたします．なお，採集した録画データ，生体データおよびアンケートについては請求があればいつでも開示いたします．
- 本実験は自発的同意に基づいてのみ行われ，被験者はいつでも拒否する権利を保持し，拒否によって何らかの不利益を被ることはありません．

実験代表者および連絡先

公立はこだて未来大学 大学院 システム情報科学研究科 修士 2 年　曲木 拓朗
email: g2118038@fun.ac.jp
公立はこだて未来大学 大学院 システム情報科学研究科 教授　Michael Vallance
email: michael@fun.ac.jp

Fig. B.1: Consent form page 1



実験同意書

◎ 実験説明書の事項について確認及び承認いただけましたら下記に署名をお願いいたします．

年　月　日

私は，本実験における主旨を理解の上，上記の事柄について同意いたします．

氏名 _____

住所 _____

連絡先(email) _____

Fig. B.2: Consent form page 2

Appendix C Pre-Survey 1

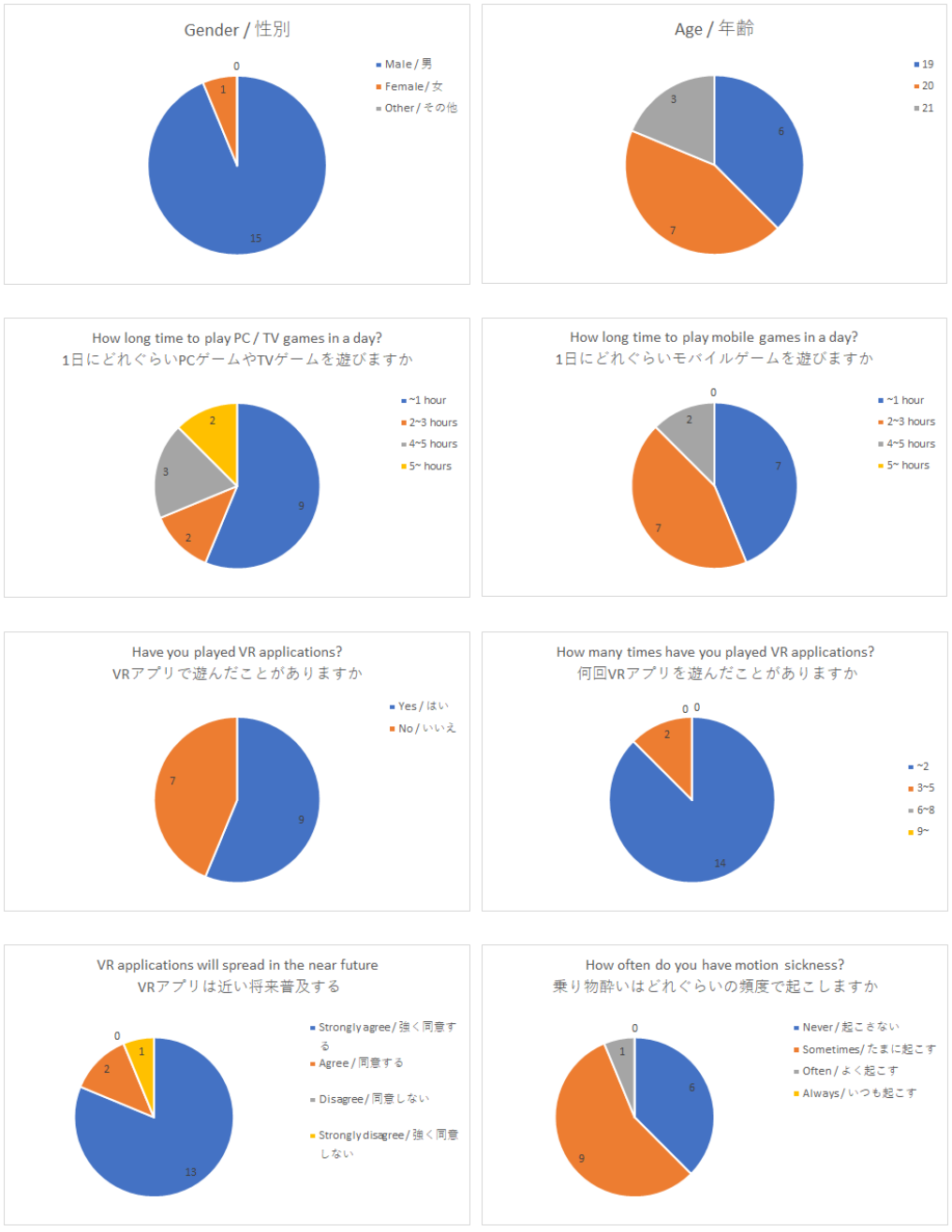


Fig. C.1: Result of Pre-Survey in Experiment 1

Appendix D Post-Survey 1

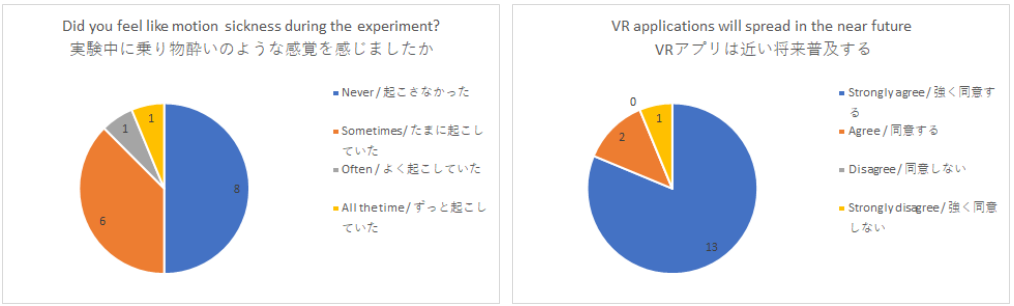


Fig. D.1: Result of Post-Survey in Experiment 1

Appendix E Pre-Survey 2

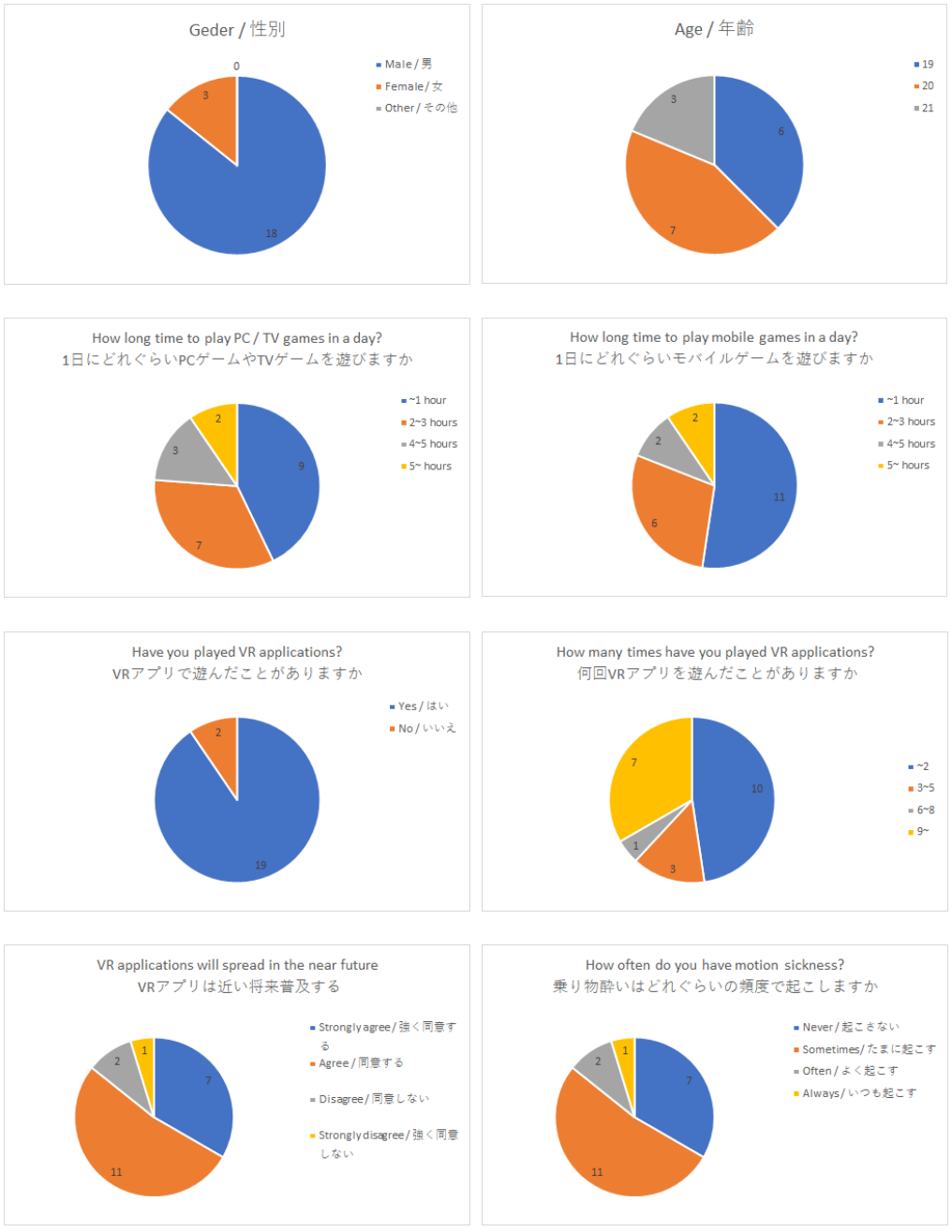


Fig. E.1: Result of Pre-Survey in Experiment 2

Appendix F Post-Survey 2

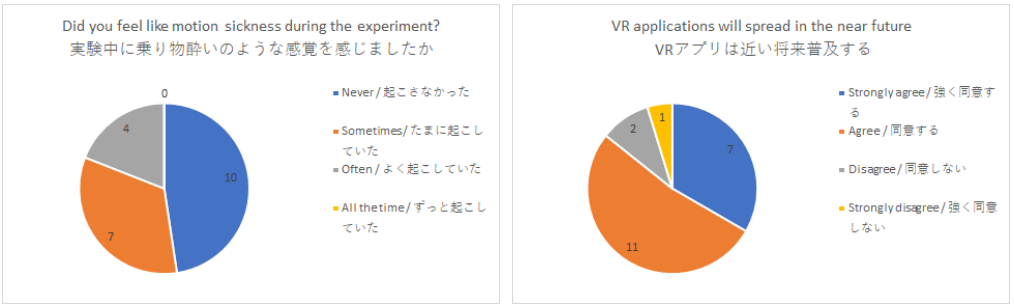


Fig. F.1: Result of Post-Survey in Experiment 2

List of Figures

List of Tables