

# Eye Thought You Were Sick! Exploring Eye Behaviors for Cybersickness Detection in VR

Phil Lopes

École Polytechnique Fédérale de  
Lausanne  
phil.lopes@epfl.ch

Nana Tian

École Polytechnique Fédérale de  
Lausanne  
nana.tian@epfl.ch

Ronan Boulic

École Polytechnique Fédérale de  
Lausanne  
ronan.boulic@epfl.ch

## ABSTRACT

Cybersickness induced through Virtual Reality (VR) applications is still one of its main barriers as it can induce unwanted side-effects in players, significantly hindering the overall experience. Despite the wealth of research available on this topic, it is still an unsolved problem. Although previous studies have explored methods of cybersickness mitigation in addition to correlating physiological factors, there has been little research on the potential correlation of eye behavior and cybersickness. Thanks to advances in eye-tracking technology within HMDs, detecting eye behavior has become a much easier process. This paper explores the differences of pupil position and blink rate in relation to cybersickness intensity. The latter is measured through the standard and a customized version of the Simulator Sickness Questionnaire (SSQ). A total of 34 participant data was collected from two separate playing sessions of a VR maze game, where each session presented a unique control scheme.

## CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI); Interaction paradigms;**

## KEYWORDS

cybersickness detection, eye tracking, blink rate

### ACM Reference Format:

Phil Lopes, Nana Tian, and Ronan Boulic. 2020. Eye Thought You Were Sick! Exploring Eye Behaviors for Cybersickness Detection in VR. In *Motion, Interaction and Games (MIG '20)*, October 16–18, 2020, Virtual Event, SC, USA. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3424636.3426906>

## 1 INTRODUCTION

Despite many years of development in the field of Virtual Reality (VR), VR-induced sickness remains a serious problem that may hamper individual experience and hinder its widespread adoption[25]. Over the last thirty years, similar terms have been used interchangeably in papers studying VR sickness such as visually induced motion sickness(VIMS), simulator sickness and cybersickness. Cybersickness is mainly accepted as describing discomfort caused by exposure to virtual reality. We use cybersickness in the following sections.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
*MIG '20, October 16–18, 2020, Virtual Event, SC, USA*

© 2020 Association for Computing Machinery.  
ACM ISBN 978-1-4503-8171-0/20/10...\$15.00  
<https://doi.org/10.1145/3424636.3426906>

Through the years, various theories have been developed to explain the cause of cybersickness. Well-received explanations are the visual-vestibular conflict[17] and visually-induced self-motion illusion[1]. People might feel dizziness, nauseous, unsteady and disoriented when they are experiencing moving visual stimulus and interpreting such visual motion as self-motion, while the inner ear vestibular system detecting physical movement, orientation and balance cannot sense actual motion. Thus, such inconsistent signals may be the cause of the unpleasant symptoms. Also, individual factors like age, gender, past VR experience determine the susceptibility and tolerance of induced cybersickness, resulting in a wide variety of individual sensitivity and making it more difficult to find a generic solution to combat cybersickness[24][17].

Regardless of the recent progress in the ability to track VR user movements over larger spaces such as with the HTC-Vive Lighthouse 2.0, it is highly likely that a significant number of applications will continue to suffer from a spatial discrepancy between the physical space and the virtual space. In the absence of an affordable omnidirectional treadmill, such type of applications generally have to put in place navigation metaphors that may trigger the visual-vestibular conflict. The main concern of this paper is the early detection of cybersickness by exclusively studying the pupil position and blinking patterns of the eye. Taking into account that good quality eye tracking devices are now being integrated within VR Head-Mounted Displays (HMDs), such systems may provide a minimal intrusion method of detecting cybersickness (i.e. without physiological devices). Any early signal would be beneficial for either providing an explicit feedback to the user (e.g. having a break) or for dynamically adjusting navigation parameters.

As such this paper intends to isolate and explore the relationship of two specific eye conditions: Blink Rate Frequency and Pupil Position Variance; with cybersickness. Results are observed between two diverging groups, where one group consists solely of sessions where individuals reported to be cybersick while the other consists of sessions where individuals reported no cybersickness influence. This paper is an extension study of previous work by the authors [3].

## 2 RELATED WORK

Cybersickness triggered by atypical visual stimuli refers to visually induced motion sickness specifically in virtual environments, which could lead to physiological symptoms such as nausea, dizziness, sweating, headache, among others [5]. In previous studies, different conflicting theories have been postulated and discussed to explain the cause of cybersickness including sensory conflict theory, postural instability theory, and poison theory [5, 18]. Yet, as of the time of writing there is still no standard explanation [5, 7].

Previous literature indicates that cybersickness differs from similar terms such as motion sickness and simulator sickness in three different aspects [12, 14, 26]: one being the contributing factors and environment, the other being the types of symptoms and finally the last being the severity of such symptoms.

Unlike motion sickness, cybersickness can be elicited without significant physical motion [11]. According to Stanney [12, 26], the general severity of cybersickness is claimed to be three times greater than that of simulator sickness. The profile of cybersickness is fundamentally different from simulator sickness with disorientation as the dominant symptom, followed by nausea and oculomotor as the least. On the other hand, Gavvani et al. [21] concluded that motion sickness and cybersickness are identical problems while using the Motion Sickness Assessment Questionnaire. Even though the previously mentioned studies employed different questionnaires, it is still inconclusive whether these questionnaires designed for motion sickness or simulator sickness are suitable to assess cybersickness. Currently, there is no standardized questionnaire designed for cybersickness in particular. Hyun et al. [14] attempted to restructure the questions from the popular Simulation Sickness Questionnaire (SSQ) [13] for the specific purpose of cybersickness studies. Unfortunately the tasks used to induce cybersickness within the context of their work did not fully explore movement within VR space, where the tasks did not require the participants to even move within the virtual world. Thus, the tasks are not representative of actual inducement of visual motion in VR, which make this paper insufficient for redefining questions and symptoms for VR sickness studies.

## 2.1 Subjective and Objective Measures

Subjective and objective measures both have their advantages and disadvantages. Different subjective measures such as the motion sickness assessment questionnaire, the fast motion sickness questionnaire, the average discomfort score questionnaire, the motion sickness susceptibility, and motion sickness history questionnaire have been heavily used to evaluate the severity of cybersickness [5].

The most widely accepted and popular one is the Simulator Sickness Questionnaire (SSQ). The SSQ was originally developed by Kennedy et al. [13], and consists of 16 questions on a four-point scale (None-0, Slight-1, Moderate-2, and Severe-3). Each question relates to one of three symptomatic groups being: Nausea, Oculomotor and Disorientation. One of the main downsides of using questionnaires however, is on the problems consisting individual bias and misinterpretation, such as conceptual misunderstandings or the way individuals associate answers with their own physical and psychological bias. Furthermore, simply the act of using a questionnaire or when this questionnaire is presented can also be disputable. For example, Young et al. [29] investigated the impact of using the motion sickness questionnaire differently by either presenting the SSQ only at the end of the experiment (post-SSQ) or presenting the questionnaire both at the beginning and the end of the experiment (pre- and post-SSQ). Results indicate that using exclusively a post-SSQ individuals tend to report less motion sickness. Furthermore, by using a pre-SSQ it is possible to establish an initial baseline value providing a more accurate measure of the overall SSQ score from VR exposure.

Given the recent studies showcasing the superiority of comparative methodologies for questionnaires [20, 28], this paper attempts to apply a similar tactic for the post-SSQ as both a measure of control (i.e. see if there is a consistency between player responses) and its potential viability.

The usage of objective measures is also quite common in the detection and even prediction of cybersickness. Physiological devices such as the Electrogastrogram (EGG) which record gastric activity, Electrocardiogram (ECG), Electrooculogram, Electroencephalogram (EEG), Electrodermal Activity (EDA), Photoplethysmogram (PPG), Eye-Tracking Devices (for the detection of eye blink frequency, fixation, saccades), Respiration and Postural movement have all been applied to estimate levels of cybersickness symptoms [5, 15]. Unlike questionnaires, such measures can capture the dynamic changes of physiology of each participant during VR exposure without interrupting the experience. Studies have also included multi-modal systems, where multiple physiological measures are explored and compared for the task of cybersickness detection [15]. However, such studies can be difficult to process as devices tend to have very different methods of interpreting the shape of the signal, which can provide inconsistent conclusions. Furthermore, the accuracy of physiological measurements lies in the quality of each sensor, its positioning and user movements. Not to mention that such sensors tend to be quite invasive, when compared solely to eye tracking. As such when using physiology, questionnaires tend to also be used in conjunction with such objective measures. The reason being that physiological signals can be ambiguous in the sense that there is little context that can infer that certain bodily reactions are directly correlated to the effect of cybersickness or other potential factors present in the experience. Thus, it is common to use subjective measures in addition to objective ones for the purpose of accuracy and context [6].

## 2.2 Cybersickness and Pupil Position

Human eye movements and their respective changes can reveal how individuals visually perceive the world. However, there are no established connections between pupil movement and cybersickness due to a lack of advanced eye-tracking technology for VR HMDs. Hvass and Berthelsen [9] conducted a preliminary study exploring the correlation of cybersickness and gaze direction for walking tasks, where individuals were directed to walk in a straight line and turn within the VR environment. Three different conditions were experimented per participant: one where participants played seated, another where they stood and the final condition had them physically walking. Even though the overall SSQ scores showed that there was a substantial increase of cybersickness after VR exposure for each of the conditions, no significant difference was found among them. As such the authors were not able to make any noteworthy conclusions between the relation of pupil position or eye saccade periods (i.e. quick simultaneous movements of the eyes), and cybersickness.

## 2.3 Cybersickness and blink rates

Blink rate has often been associated with negative human behaviours such as stress and fatigue [19]. Furthermore, empirical studies have also suggested that changes in blink rate can also be a

reliable measure for eyestrain [4, 15]. Given that eye fatigue is one of the core symptoms of the SSQ [13] it makes sense to explore this factor as a potential variable for cybersickness detection [10].

Previous studies on blink rate and cybersickness did reveal a positive correlation between SSQ Scores and blinking [16]. Kim et al [15] described the significant changes of blink rate over time during a VR navigation task. The authors state that blink rate (blinks per minute) tends to considerably increase over the course of a game, while at the very beginning of the experience the rate is below even the baseline value. Dennison et al. [8] also reported similar findings, where the rate of eye blinks tended to decrease at the initial state of navigation and increase above baseline level around the middle of the task. One conclusion was that the greater changes in blink rate between baseline and the virtual navigation the higher the incidence of cybersickness was reported.

This work intends to further explore the eye blink frequency phenomenon on the effects of cybersickness. More precisely, this work intends to compare the blinking frequencies of sick and non-sick individuals over the course of a VR navigation task. Unlike previous work, which exclusively explored the correlation between SSQ and blink rate, this paper focuses on comparing the blink rate between cybersick and non-cybersick individuals and infer if there is a significant difference between both populations.

### 3 METHODOLOGY

Purposefully inducing cybersickness can be a delicate process as the tolerance can vary substantially between individuals. The principal risk consists of being in either extremes of the “cybersickness spectrum”. More precisely, either not inducing the effect at all (i.e. lack of cybersickness data) or overcompensating and making the experience unplayable for the majority of participants (i.e. forcing the participants to quit midway through an experiment). To balance this risk two alternate versions of the game were developed, where one version of the game was designed to have a higher risk of inducing cybersickness than the other (Section 3.2).

Two types of data are collected from the eye tracking device during play: the *Blink Data Markers* (Section 3.3) and the X and Y *Pupil Position* axes (Section 3.4) of both the left and the right eyes of the participant. The device used to record eye data markers and synchronize with the game is further described in section 3.1.

Finally, the Simulation Sickness Questionnaire (SSQ) was used to assess the intensity of cybersickness felt by each participant. Two types of SSQs were used for the cybersickness assessment. The first consists of the standard non-altered SSQ as stated in [13], where participants are tasked in filling the questionnaire before (Pre-SSQ) and after (Post-SSQ) the game-playing session in order to derive the variation  $\Delta SSQ$  (Section 3.6) over the game session. In the standard formulation the participant indicates an absolute levels for each component of the questionnaire. Although the Post-SSQ items were kept identical to those of the Pre-SSQ, we added one question per item asking for the relative participant state (better, worse, the same). The latter was mostly used as an additional control to analyse the reliability of the answers obtained by comparing with the standard SSQ results.

### 3.1 Apparatus

The head-mounted display used in this study is the *HTC Vive Pro Eye* (HTC 2019). The headset integrated with Tobii Eye tracking includes a dual OLED screen, a resolution of 1440 \* 1600 pixels per eye, a 110-degree field of view and a refresh rate of 90 Hz. Both eyes update data at the sampling rate of 120 Hz, and the estimated accuracy of the eye tracker is 0.5 degrees with a latency of approximately 10 ms.

Given the disparity between the sampling rate of the eye-tracker and Unity’s Frames Per Second (FPS) hard limit cap of 90 FPS, it was not ideal to interface between both systems directly. First it would heavily limit the sampling rate of the eye-tracker as data would only be fetched during each Unity frame, which has the potential of increasing the error rate and lowering the accuracy of blink detection. Thus, to effectively synchronize both data-streams the LabStreaming Layer (LSL)<sup>1</sup> middleware was used syncing both the game with the eye-tracking device.

This synchronization allows to trim the data with substantial accuracy ( $\approx 1\text{-}2$  ms error) based on several markers obtained from the game. For this study, the eye data was delimited exclusively to the 8 min game-play portion of the experiment (i.e. the start of play until the end).

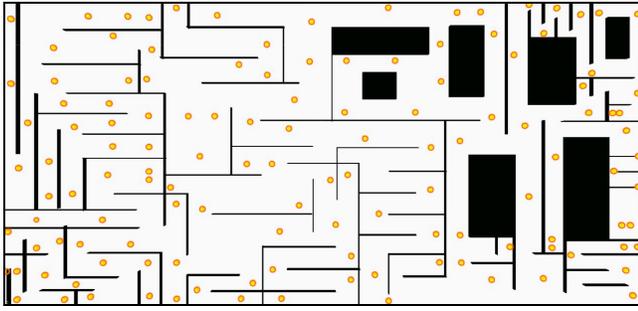
### 3.2 The Game

The game consists of an exploration game where players are tasked in collecting as many coins as possible within a large virtual maze (see Fig. 1) in a specified time limit of 8 minutes, which is always hidden from participants during play. Furthermore, the game was made purposefully simple and straightforward as to mitigate the influence of skill of the different participants as much as possible. It was also advantageous as the game was quite easy to explain and understand by the participants. The game itself is also designed specifically to force players to explore and move within the virtual space, as the main task is to collect as many coins as possible. To avoid any other types of external influences, such as stressing the player, the remaining time left is never shown and there are no competitive elements in the game. Unbeknownst to the players, the maze itself is impossible to fully complete (i.e. collect all coins) within the allocated time. Coins are placed in such a way that it constantly pushes the player to move forward and towards key junctions of the maze. The game itself was fully developed using the *Unity 3D* (Unity Technologies) game engine using the Steam VR plugin (for the VR support) and the provided SRAnipal SDK from HTC for eye tracking. The controllers used for the purpose of this experiment were the *Index Controllers* by Valve (2019), which support motion tracking and provides two joystick controllers. An example of game-play can be seen in the following video<sup>2</sup>.

Movement within VR environments is one of the main contributors of cybersickness [27], as the dissonance between virtual rotation and the lack of physical movement can significantly conflict with the vestibulo-oculomotor reflex [22]. Although many other factors exist that have the potential of increasing the risk of cybersickness, this paper focuses exclusively on manipulating the player

<sup>1</sup><https://github.com/sccn/labstreaminglayer>

<sup>2</sup>Anon

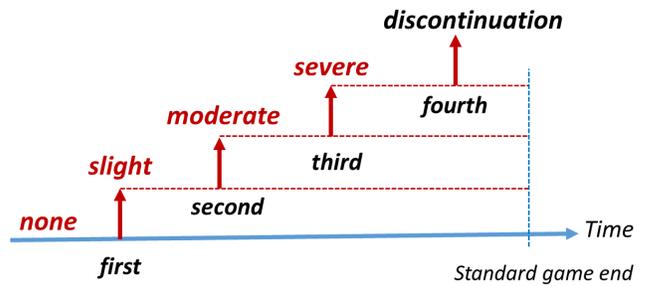


**Figure 1: The maze (or level) played by each participant for both variations of the game. The objective of players is to collect as many coins (yellow circles) as possible in the allotted time.**

displacement metaphor for the sake of simplicity and to avoid additional external variables that can compromise the analysis of the gathered data. Thus, two variants of the game were developed with different types of player displacement mechanics: the Teleportation (TP) and Standard Control (SC) game variants.

**3.2.1 Teleportation Control (TP).** This control scheme relies on the well-known “teleportation” metaphor. More precisely, to move within the virtual world players point to a location within the virtual environment with the aid of a controller capable of tracking player motion, and choose which location they wish to move to. If the location is valid the player is then teleported to that location with a fade-to-black screen in-between the last position and the newest one. Furthermore, to be able to rotate their avatar within the virtual environment players have to physically rotate in place, where the rotation is approximately one-to-one with both the physical and virtual space. Given the latter we expect less conflict with the vestibulo-oculomotor, which helps to reduce the overall risk of cybersickness during play. Furthermore, previous research has suggested that “teleportation-type” control schemes tend to provide the most comfortable experiences and provide less risk of inducing cybersickness [27]. Thus, this control scheme is designed to be the “low-risk” variation for inducing cybersickness, where the expected SSQ result is non-existent to mild cybersickness overall symptoms in the majority of participants.

**3.2.2 Standard Control (SC).** Standard Control relies heavily on the preconceived or “standard” control schemes, which are prevalent in most Third and First Person action games. This control scheme exclusively uses the two joysticks available on the controller allowing the players to effectively move within the environment. The left joystick controls the player displacement allowing them to move forward, backwards and sideways (left or right without rotating the camera). The right joystick allows the players to rotate their virtual avatar left and right; unlike the previous control scheme players do not physically rotate in-place, and as such they are told to use the controllers instead. Thus, the lack of physical rotation has the potential of creating a higher dissonance between the visual and physical body movement, substantially increasing the risk of cybersickness. Unlike teleportation, this control scheme is often avoided due to its high risk of inducing cybersickness during actual



**Figure 2: The different tiers of the cybersickness report button. Once the button is pressed the tier is increased by one, until the end of the game. If the button is pressed 4 times the experiment is aborted.**

play[2]. As such, this control scheme is purposefully designed to be the “higher-risk” variation, where the intensity of the symptoms felt by the players are expected to be more intense than the previous control scheme, while still being tolerable enough for players to fully complete the session.

**3.2.3 Cybersickness Report Button.** To pinpoint the exact moment in which players start to feel the different intensities of cybersickness symptoms, a “report” button (the right control trigger) was provided allowing participants to report the intensity of the symptoms during actual play. This system annotates the data to specific points in time providing the ability to concentrate on the time window and data where the actual phenomena arises for each of the participants during its analysis. To avoid overwhelming players and allow them to concentrate on the game, the report trigger was designed with basic functionalities, while using the same “tier-based” system that participants familiarized with in the Pre-SSQ (see Fig. 2). The first time the participant presses the button the system will annotate from that point forward that the player is feeling “slightly” sick. If the button is pressed again it will increase to the tier above it, annotating from that point forward that the player is feeling “moderately” sick and so on. In total there are three tiers: Slight, Moderate and Severe; on the fourth button press the participant states that they wish to terminate the experiment and the game is immediately aborted. On each button press a small text appears on screen displaying the current tier the player is currently at and subsequently disappears after 10 seconds. Lastly, the report button is enabled for both control scheme variations allowing participants to state the level of motion sickness intensity during play.

### 3.3 Blinking Data Markers

Blinking is a bodily function which consists of the rapid closing and opening of the eyelid, which often happens automatically as a means of spreading tears across the surface of the eye. Given that one of the most prominent symptoms of cybersickness is related to oculomotor functions such as eyestrain, dizziness and blurry vision, such associated phenomena can often be manifested through bodily functions such as blink rate frequency[15]. As such, this work intends to understand if a relation exist between cybersickness and the frequency of blink and/or different blink types.

For the purpose of this work three types of blinks are being tracked as classified according to literature [23]: **Short Blink** (1 to 7 ms), **Medium Blink** (8 to 15 ms) and **Extended Blink** (16 to 25 ms).

### 3.4 Pupil Position

For the purposes of this work pupil position consists of the position of the each pupil (left or right) within the boundaries of the eye tracking screen. Given the hardware used, this means that each eye has its own dedicated tracking screen and is limited by the boundaries of these screens. Pupil position can also relate to gaze direction, which can be estimated with some degree of accuracy the general direction and area that has the player’s attention. For the purpose of this work however gaze direction is not used, as the interest is more on analysing eye behaviors and phenomena that can arise from a state of cybersickness, than attempting to contextualize the data based on what is happening on screen.

Each position value is normalized between  $[-1, 1]$  for each axis, where  $-1$  consists of the leftmost boundary of  $X$  and the lowest boundary of  $Y$ , while  $1$  consists of the rightmost boundary of  $X$  and the upmost boundary of  $Y$ . The average between both the left and right eye pupil positions at each frame is then calculated for both the  $X$  and  $Y$  axes to obtain a general representative value of both eye locations.

### 3.5 Simulation Sickness Questionnaire (SSQ)

Cybersickness elicitation is not guaranteed among participants, even though the different game conditions are intended to increase and decrease this risk factor. As such, to assess the intensity of cybersickness among participants the standard Simulation Sickness Questionnaire (SSQ) was employed [13]. The SSQ can derive a rough numerical estimate of the intensity of cybersickness felt by each individual based on three specific symptomatic groups: Nausea, Oculomotor and Disorientation. Participants are tasked in answering several questions from within each group using a 4-point Likert scale with the following labels: None, Slight, Moderate and Severe.

Calculating the total score of the SSQ depends on each symptomatic group score, which themselves depend on the ratings of each symptom stated by an individual participant (see Table 1). Each of the constants in the following formulae were empirically derived based on [13]. The total SSQ Score ( $SSQ_{Score}$ ) is calculated as such:

$$SSQ_{Score} = \left( \sum N_{Sym} + \sum O_{Sym} + \sum D_{Sym} \right) \times 3.74 \quad (1)$$

where  $N_{Sym}$ ,  $O_{Sym}$  and  $D_{Sym}$  are the ratings of all associated symptoms to the Nausea, Oculomotor and Disorientation symptomatic groups, respectively.

To obtain a more consistent representation of the SSQ score a pre and a post-SSQ is given to the participant, where the only difference is the pre-SSQ is filled before the start of the experiment while the post-SSQ is filled at the end of the experiment. The difference between the post- and pre-SSQ are subsequently calculated ( $\Delta SSQ$ ), which is the score used for each individual session.

Symptoms	Nausea	Oculomotor	Desorient.
General Discomfort	*	*	
Fatigue		*	
Headache		*	
Eye Strain		*	
Focus		*	*
Salivation	*		
Sweating	*		
Nausea	*		*
Concentration	*	*	
Head Fullness			*
Blurry Vision		*	*
Dizzy (Eyes Open)			*
Dizzy (Eyes Closed)			*
Vertigo			*
Stomach	*		
Burp	*		

**Table 1: The Symptomatic Groups of the SSQ: Nausea (N), Oculomotor (O) and Disorientation (D). Asterisks (\*) symbolizes which symptom contributes to the overall Symptomatic Group score [13].**

### 3.6 Applying Comparison Methods to SSQ

Over the past few years there has been a push within the scientific community towards using comparative methods, instead of the commonly used “Likert-style” questionnaires when performing studies that require participant input [20, 28]. The reasoning behind this is how scales are often ambiguous and personal to each individual, which can often lead to inconsistent and noisy answers. For the purposes of this work, an alternative version of the SSQ questionnaire was developed, where all questions were turned into a comparison. Participants are tasked of comparing the current state of each symptom (the current state being after VR stimulus), with that symptoms initial state right before VR exposure, similarly to what is achieved with the  $\Delta SSQ$ .

Because no scale exists within this version of the questionnaire, the scoring of each question is slightly different than the standard version of the SSQ. For this version scoring consists of either a 1 for “positive” answers or a 0 for “negative” answers. Positive answers refers specifically if the participant strongly felt that particular symptom during or after the play session, while the negative answer states that the participant did not feel any difference or anything related to that symptom. To calculate the total score and each symptomatic group score the same methodology is used as the one previously described, which is the same from the standard SSQ method[13]. The main difference being scores will have considerably smaller values due to being bounded to 0 or 1.

### 3.7 Experimental Protocol

The experiment takes place over the course of two different sessions of around 40 minutes each (see Fig. 3). Each session is conducted on two different days (with a maximum interval of  $\approx 1$  to 3 days). For each session a participant plays one variation of the game (SC or TP). To ensure an equal balance between both variations and

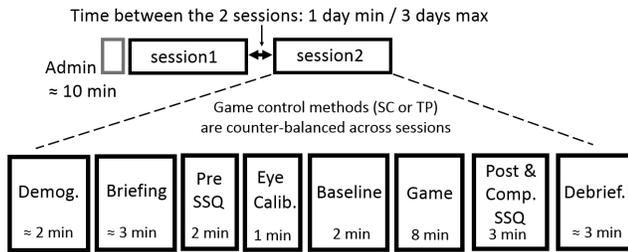


Figure 3: The Experimental Protocol.

mitigate an “order”-bias, the first version played by each participant alternates from one participant to the next. It is important to keep in mind that all participants play both variations of the game.

Each session employs the same sequence (see Fig. 3). As the name suggests each participant is tasked in filling a demographics survey and the Pre-SSQ before the experiment begins. Once completed, participants are given a brief explanation of the control mechanisms for the session and the cybersickness report button. Once the headset is placed the eye calibration software is run to ensure the eye tracking device is functioning according to the participant’s individual specification. Once calibration is completed, the baseline method is run consisting of a black screen with text stating the game commences in  $X$  seconds. At the end of the timer the game starts and participants play for a total duration of 8 minutes (unless the game is aborted by the participant). Once finished the Post and Comparative SSQs are filled and if necessary an experimental debriefing is held allowing the experimenter to ask questions to the participants.

## 4 RESULTS

Experimentation was fully approved by the University ethics committee where the only restriction was on the types of participants allowed to be recruited, which consisted of fully healthy individuals between the ages of 18 to 40. In total 36 participants were recruited for the experiment. Out of the 36 participants 2 participants had to be fully discarded, because one participant’s data was incomplete due to a recording error while the other was not able to finish the SC session and aborted midway through the level. As a result a total of 34 participants presented valid sessions for both SC and TP making it a total of 68 sessions, where 34 are SC and 34 are TP. Out of the 34 participants 12 are females and 22 are males aged between 18 to 36. Among them, 17 participants rarely play digital games; 12 participants stated that they play games a few hours per month; 2 participants spend a few hours per week (1 to 4h) playing games; 2 stated they play several hours per week (5 to 10 h), and 1 participant was a regular player (10h+). In relation to previous VR experiences the majority stated that they had a few VR experiences beforehand but no regular use (25 participants in total), one participant stated that they played VR games often (a few hours per month), while the remaining 8 participants stated that they had never used VR before.

Variation	Sick	Non-Sick
SC	21	13
TP	11	23
<b>Total</b>	<b>32</b>	<b>36</b>

Table 2: Distribution of Sick ( $\Delta SSQ_{Score} \geq 15$ ) and Non-Sick ( $\Delta SSQ_{Score} < 15$ ) between the TP and SC game variations.

### 4.1 Standard and Comparative SSQ Correlation

To ensure consistency during the SSQ phase two different variations of the questionnaire are used: The Standard and the Comparative versions. For the purposes of this work the comparative version serves mostly as a control to ensure that there is a consistency among the answers given by each participant. Thus, to observe this consistency a Pearson Correlation is calculated on the SSQ Standard and Comparative scores. The observed coefficient is  $\approx 0.912$ , which confirms that the answers given by participants are at the very least consistent among the two questionnaires being used.

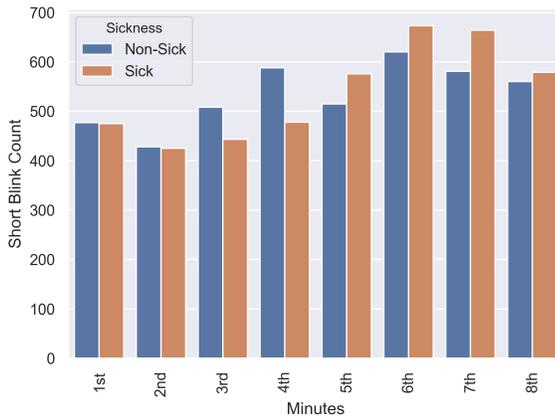
### 4.2 SSQ Sickness Distribution

Inducing cybersickness is not a straightforward task and may vary extensively between participants. The purpose of the different game variations presented in this work is to manipulate the risk factor by using methods which are known to either lower (TP) or increase (SC) the risk of cybersickness symptoms. Table 2 presents the distribution of sessions where participants reported they were sick ( $\Delta SSQ \geq 15$ ) or non-sick ( $\Delta SSQ < 15$ ). This value was chosen based on previous literature [12], which states that scores above this value are “sufficiently” high enough to feel the effects of motion sickness.

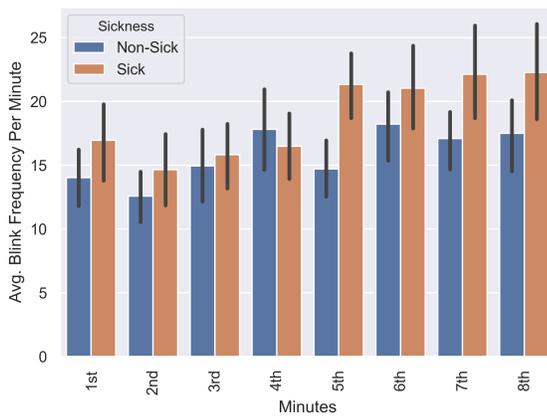
The distribution suggests that the risk factors predicted from both variations are validated, where the risk of cybersickness is higher for the SC variation ( $\approx 62\%$ ), while for the TP variation there is a higher risk for not getting sick ( $\approx 68\%$ ). Surprisingly when looking at the sickness factor for both session types there does seem to be a balance between non-sick and sick sessions, where a minor skew exists towards non-sick.

### 4.3 Report Button Effectiveness

Over the course of several experiments a particular pattern was observed from the participants interaction with the cybersickness report button. Participants who claimed to be sick by the end of the experiment would scarcely report it using the interface. Once inquired about this fact several participants suggested that it was not necessary to press it as they could simply “power-through” the experience and complete the task without the need of reporting it. This may have been a miscommunication on the part of the experimenters, however the psychological factor of pride or “not appear weak” during the game should also not be ignored. In fact, when comparing between the SSQ Scores of individuals who are within the sickness category only 46.9% of individuals utilized the button prompt. Thus, less than 50% of individuals who were in-fact sick did not use the available reporting mechanism at their disposal. Due to the unreliability of the reporting mechanism, it was not used for the subsequent analyses.



**Figure 4: The Total Number of Short Blinks Per Game Minute. As the game progresses the amount of short blinks tends to be slightly higher in the “Sick” group than that of the “Non-Sick” group.**



**Figure 5: The Average and Standard Error of the Blink Frequency per Minute across all participants for both Sick and Non-Sick Groups.**

#### 4.4 Blink Pattern Analysis

In the context of this work blink data consists specifically of markers which are triggered during play and denote the exact time within the game a type of blink occurred. As such blink data has no “numerical” value aside from its quantitative measure (i.e. number of blinks occurred at a time interval). Furthermore, given the large skew that exists within each game variation between sick and non-sick, a quantitative evaluation in such conditions would heavily bias the data towards the skewed group, which was confirmed during preliminary analyses. Thus, it was decided to present and analyse the data from a global perspective combining both game variations yielding a more balanced and fair comparison.

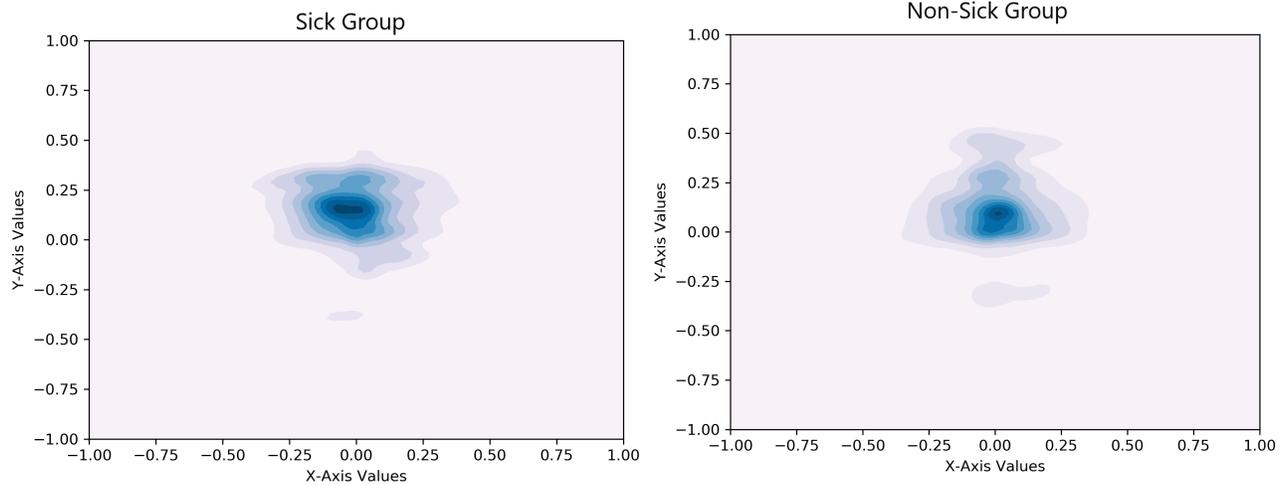
Figure 4 presents the quantity of “short blinks” per in-game minute over the course of the game. Although, in the latter stages of the first half of the game it can be seen that non-sick tend to have a higher blink-rate this pattern is suddenly broken on the second half of the game. This pattern may arise due to a slight habituation period where the participants eyes are still adjusting to the on-screen action, which can explain why in the latter stages of the game the blink-rate seems to stabilize.

Interestingly for the Sick group a more gradual progression of blinks is observed over the course of the game. During the first half of the game the number of blinks stays relatively stable constantly in-between the 400 to 500 interval. This initial phenomena may arise due to a natural reaction to cybersickness, where the first hints of disorientation forces a participant to concentrate on key elements in the level. Such focus can lead to a stabilization or even reduction of eye-blink frequency. In the second half of the game there is a sudden increase of blinks, even surpassing the Non-Sick group. This observation may be due to participants starting to feel a higher intensity of the eye-strain symptom, which can lead to a higher frequency of blinks. The Kolmogorov-Smirnov test (KS-Test) was conducted on both distributions to verify if there was a statistical difference between both, however no such difference was observed as the  $pvalue$  was considerably above the typical 0.05 threshold.

To further investigate the previous phenomenon the average blink rate per minute was plotted across all participants for both sick and non-sick conditions (see Fig. 5). A KS-Test was also conducted for this distribution to verify if there was a significant difference between both conditions, although it did yield a lower  $pvalue$  than the previous version ( $\approx 0.28$ ) it was still not significantly different. This being said an interesting pattern does emerge from the blink frequency where sick individuals tend to present a higher blinking rate than those who stated they were not sick. This does go inline with previous literature [15], where a correlation was observed between cybersickness and blink rate. Similarly it also does explain the previous observation, given that the blink frequency per minute tends to extensively increase in the latter stages of the game. To conclude, it is also important to note that both conditions are not entirely balanced, even though the skew is minor towards non-sick sessions, it may be enough to contribute to the overall inconclusiveness obtained within these results. It also showcases that even with this skew, participants with the sick condition have a higher blink rate and count which can be even more prominent if the data was more equally balanced among the two groups.

#### 4.5 Pupil Position – Data Spread and Variability

Over the course of experimentation a large amount of pupil data was collected where each session yielded approximately 57600 data points. For the purposes of this paper the core interest is to observe the spread and variation of this data. The hypothesis being that during a cybersickness event an individual tend to center their gaze considerably more as a mechanism to overcome this phenomenon, while individuals who are not cybersick will allow themselves to observe their surroundings creating more variability in the data. As such the following section analyzes the spread of the total pupil position data between both Sick and Non-Sick groups.



**Figure 6: Pupil Position Density Plot of Sick and Non-Sick Groups. The distribution of all gaze data points over the course of an 8 minute session of each individual whose  $\Delta SSQ_{Score} < 15$ . Darker colors represent a higher concentration of data points within that locations.**

**4.5.1 Analyzing Spread.** All collected data is split into two diverging groups similarly to previous analyses based on the  $\Delta SSQ_{Score}$ . Figure 6 shows the density plots of all the collected gaze data obtained over the course of 8 minute sessions for both the “Sick” and “Non-Sick” groups, respectively.

Overall the sick group does present a consistent area of concentration in the middle of the data cluster when compared to the non-sick group. It is evident that some degree of stability exists for individuals of this group, while for non-sick this stability and concentration is less present due to a smaller concentration area. The latter does suggest that due to this the data is more spread and less concentrated at specific areas. The spread within each axes also slightly differ between both groups. For the sick group the data points seem to be more spread along the X-Axis while more compact along the Y-Axis, meaning that individuals tended to look more to the sides than the upward periphery. On the other hand for the Non-Sick group the data points were heavily more spread-out along the Y-Axis, although on the X-Axis it was quite similar to the Sick group comparatively. Finally, a Kolmogorov-Smirnov test (KS-Test) was also conducted on both distributions for both the X- and Y-Axes to verify their statistical significance. For both a  $p_{value} < 0.01$  was observed.

Although the spread of the distributions does suggest that there is some slight variability between both conditions and a statistical difference was observed, it is important to note that the data itself is not devoid of variability and noise. For example the inherent nature of the task used (i.e. a game) provides a substantial amount of freedom to participants allowing them to take actions and be at different locations at different points in time, which can impact what they are looking at or what they are concentrating on. Although such tasks are not fully controllable, they are the easiest way of evoking cybersickness given that the participant has an “active” role in the task at hand. Furthermore it is also important to mention the

granularity of the pupil position data can be especially “unrefined”, meaning that within the data itself a lot of minor variations can also be present due to the inherent nature of eye-tracking itself.

## 5 DISCUSSION

Early results based on the experiment conducted within this paper suggests that some aspects of eye-tracking data can be difficult to associate with cybersickness. Particularly eye pupil data proved to be the more complex type of eye data used within this work to process. Overall it provided inconclusive evidence that any link exists between pupil position and the effects of cybersickness. Furthermore, if there is a link between this type of data and cybersickness, one of the core challenges is to separate the noise obtained from regular play and the actual data capable of detecting cybersickness with some degree of accuracy. Although an attempt was made to constrain the data by allowing participants to report the specific point in time any cybersickness symptom was felt, the method was unsuccessful due to the lack of use by the participants themselves. Constraining the data through annotation markers can be a potential solution of refining the search instead of using large chunks of data whereas the majority of that data is noise.

Observing the plot density of both the X- and Y-Axes of the pupil position data did provide some interesting results, however it is particularly hard to discern an objective observation from such an amount of data that occurs during a task that allows so much participant freedom. Even so, the data point density observed does suggest that there may be some truth in the hypothesis proposed as for the sick group tends to present a higher concentration area, while the non-sick group presents a smaller high density area suggesting that for non-sick individuals this data is more spread out than that of the latter group.

On the other hand, the blink data also yielded interesting results when analysing the number of “short blinks” along the game

session. Although the difference between both distributions are not statistically significant the blink count progression observed from both sick and non-sick groups are not entirely similar and some diverging patterns do emerge. In fact this pattern tends to be substantially more prominent once we observe the blink frequency per minute, where in the second half of the game individuals with cybersickness tended to blink considerably more than those without. Based on this, it would be highly valuable to continue exploring the eye blink phenomenon.

Finally as a follow-up to this research, the authors intend to explore the concept of eye-tracking with the added contribution of human physiological signals (i.e. Electrocardiogram, Electrodermal Activity and Electrogastrogram). The hypothesis being that a multi-modal system can potentially improve the accuracy of cybersickness detection due to the added value of bodily signals, and also aid in restricting the eye-tracking data to the specific time-windows where bodily functions have a higher “activation-rate”. Given the unsuccessful attempt in using the report button functionality previously presented in this paper, the use of physiology (being an objective measure) can potentially provide a way of constraining the eye-tracking data in some capacity.

## 6 CONCLUSIONS

This paper presented a study of potential features obtained from eye-tracking data: pupil position and blink type behavior; for the detection of cybersickness. The eye tracking data of a total of 34 participants was collected over the course of two different sessions. Each session presented a variation of a game-playing task where one presented a higher risk of inducing cybersickness, while the other presented a lower risk of inducing it.

Results show some hints towards the hypotheses put forward within this paper but overall the results obtained were inconclusive. Given the large amounts of data processed the risk of noise is extensively higher due to the nature of the data itself, which is a common problem when analyzing human derived signals. Even so, the hints found within the data could potentially be refined with more substantial data constraining methods which would reduce the total search space allowing for a more focused analysis. Although this paper did not find conclusive evidence that points towards the usage of such eye-tracking features as good indicators of cybersickness, it does not necessarily dismiss them as some hints are present within the data that can signal to their potential use with a more refined analysis.

## ACKNOWLEDGMENTS

This research is supported by the SNF grant CRSII5\_180319 / 1.

## REFERENCES

- [1] Majed Al Zayer, Isayas B. Adhanom, Paul MacNeilage, and Eelke Folmer. 2019. The Effect of Field-of-View Restriction on Sex Bias in VR Sickness and Spatial Navigation Performance. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI '19). ACM, New York, NY, USA, Article 354, 12 pages. <https://doi.org/10.1145/3290605.3300584>
- [2] Majed Al Zayer, Paul MacNeilage, and Eelke Folmer. 2018. Virtual Locomotion: a Survey. *IEEE Transactions on Visualization and Computer Graphics* PP (12 2018), 1–1. <https://doi.org/10.1109/TVCG.2018.2887379>
- [3] Anon Authors. [n.d.]. Anonymized Title.
- [4] Berk Cebeci, Ufuk Celikcan, and Tolga K Capin. 2019. A comprehensive study of the affective and physiological responses induced by dynamic virtual reality environments. *Computer Animation and Virtual Worlds* (2019).
- [5] Simon Davis, Keith Nesbitt, and Eugene Nalivaiko. 2014. A systematic review of cybersickness. In *Proc. of the 2014 Conf. on Interactive Entertainment*. ACM, 1–9.
- [6] Simon Davis, Keith Nesbitt, and Eugene Nalivaiko. 2015. Comparing the onset of cybersickness using the Oculus Rift and two virtual roller coasters. In *Proc. of the 11th Australasian Conf. on Interactive Entertainment (IE 2015)*, Vol. 27.
- [7] Mark Stephen Dennison and Michael D’Zmura. 2017. Cybersickness without the wobble: Experimental results speak against postural instability theory. *Applied ergonomics* 58 (2017), 215–223.
- [8] Mark S Dennison, A Zachary Wisti, and Michael D’Zmura. 2016. Use of physiological signals to predict cybersickness. *Displays* 44 (2016), 42–52.
- [9] Jonatan S Hvass and Theis Berthelsen. 2018. *A preliminary exploration in the correlation of cybersickness and gaze direction in VR*. Master’s thesis.
- [10] Julie Iskander, Mohammed Hossny, and Saied Nahavandi. 2018. A review on ocular biomechanic models for assessing visual fatigue in virtual reality. *IEEE Access* 6 (2018), 19345–19361.
- [11] Robert S Kennedy, Julie Drexler, and Robert C Kennedy. 2010. Research in visually induced motion sickness. *Applied ergonomics* 41, 4 (2010), 494–503.
- [12] Robert S Kennedy, Julie M Drexler, Daniel E Compton, Kay M Stanney, D Susan Lanham, and Deborah L Harm. 2003. Configural Scoring of Simulator Sickness, Cybersickness and Space Adaptation Syndrome: Similarities and Differences. *Virtual and adaptive environments: Applications, implications, and human performance issues* (2003), 247.
- [13] Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lillenthal. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology* 3, 3 (1993), 203–220.
- [14] Hyun K Kim, Jaehyun Park, Yeongcheol Choi, and Mungyeong Choe. 2018. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied ergonomics* 69 (2018), 66–73.
- [15] Young Youn Kim, Hyun Ju Kim, Eun Nam Kim, Hee Dong Ko, and Hyun Taek Kim. 2005. Characteristic changes in the physiological components of cybersickness. *Psychophysiology* 42, 5 (2005), 616–625.
- [16] Afsaneh Koohestani, Darius Nahavandi, Houshyar Asadi, Parham M Kebria, Abbas Khosravi, Roollah Alizadehsani, and Saied Nahavandi. 2019. A Knowledge Discovery in Motion Sickness: A Comprehensive Literature Review. *IEEE access* 7 (2019), 85755–85770.
- [17] Joseph J. LaViola, Jr. 2000. A Discussion of Cybersickness in Virtual Environments. *SIGCHI Bull.* 32, 1 (Jan. 2000), 47–56. <https://doi.org/10.1145/333329.333344>
- [18] Joseph J LaViola Jr. 2000. A discussion of cybersickness in virtual environments. *ACM Sigchi Bulletin* 32, 1 (2000), 47–56.
- [19] Eui Chul Lee, Hwan Heo, and Kang Ryoung Park. 2010. The comparative measurements of eyestrain caused by 2D and 3D displays. *IEEE Transactions on Consumer Electronics* 56, 3 (2010), 1677–1683.
- [20] Hector P Martinez, Georgios N Yannakakis, and John Hallam. 2014. Don’t classify ratings of affect; rank them! *IEEE transactions on affective computing* 5, 3 (2014), 314–326.
- [21] Alireza Mazloumi Gavgani, Frederick R Walker, Deborah M Hodgson, and Eugene Nalivaiko. 2018. A comparative study of cybersickness during exposure to virtual reality and “classic” motion sickness: are they different? *Journal of Applied Physiology* 125, 6 (2018), 1670–1680.
- [22] Michael E McCauley and Thomas J Sharkey. 1992. Cybersickness: Perception of self-motion in virtual environments. *Presence: Teleoperators & Virtual Environments* 1, 3 (1992), 311–318.
- [23] George W Ousler 3rd, Mark B Abelson, Patrick R Johnston, John Rodriguez, Keith Lane, and Lisa M Smith. 2014. Blink patterns and lid-contact times in dry-eye and normal subjects. *Clinical ophthalmology (Auckland, NZ)* 8 (2014), 869.
- [24] Lisa Rebenitsch and Charles Owen. 2014. Individual Variation in Susceptibility to Cybersickness. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) (UIST '14). ACM, New York, NY, USA, 309–317. <https://doi.org/10.1145/2642918.2647394>
- [25] Lisa Rebenitsch and Charles Owen. 2016. Review on cybersickness in applications and visual displays. *Virtual Reality* 20 (04 2016). <https://doi.org/10.1007/s10055-016-0285-9>
- [26] Kay M Stanney, Robert S Kennedy, and Julie M Drexler. 1997. Cybersickness is not simulator sickness. In *Proceedings of the Human Factors and Ergonomics Society annual meeting*, Vol. 41. SAGE Publications Sage CA: Los Angeles, CA, 1138–1142.
- [27] Tim Weißker, André Kunert, Bernd Fröhlich, and Alexander Kulik. 2018. Spatial updating and simulator sickness during steering and jumping in immersive virtual environments. In *2018 IEEE Conference on Virtual Reality and 3d User Interfaces (VR)*. IEEE, 97–104.
- [28] Georgios N Yannakakis and Héctor P Martínez. 2015. Ratings are overrated! *Frontiers in ICT* 2 (2015), 13.
- [29] Sean D Young, Bernard D Adelstein, and Stephen R Ellis. 2007. Demand characteristics in assessing motion sickness in a virtual environment: Or does taking a motion sickness questionnaire make you sick? *IEEE transactions on visualization and computer graphics* 13, 3 (2007), 422–428.