Exploring neural and peripheral physiological correlates of simulator sickness

Jan-Philipp Tauscher¹ | Alexandra Witt² | Sebastian Bosse³ | Fabian W. Schottky¹ | Steve Grogoric¹ | Susana Castillo¹ | Marcus Magnor¹

¹Institut für Computergraphik, Technische Universität Braunschweig, Braunschweig, Germany
²International Max Planck Research School (IMPRS) for Neurosciences, Georg-August-Universität Göttingen, Göttingen, Germany
³Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Berlin, Germany

Correspondence
Jan-Philipp Tauscher, Institut für Computergraphik, Technische Universität Braunschweig, Mühlenpfordtstraße 23, Braunschweig 38106, Germany.
Email: tauscher@cg.cs.tu-bs.de

Funding information
German Science Foundation,
Grant/Award Numbers: INST 188/409-1
FUGG ICG Dome, MA2555/15-1
Immersive Digital Reality

Abstract
This article investigates neural and physiological correlates of simulator sickness (SS) through a controlled experiment conducted within a fully immersive dome projection system. Our goal is to establish a reliable, objective, and in situ measurable predictive indicator of SS. SS is a problem common to all types of visual simulators consisting of motion sickness-like symptoms that may be experienced while and after being exposed to a dynamic, immersive visualization. It leads to ethical concerns and impaired validity of simulator-based research. Due to the popularity of virtual reality devices, the number of people exposed to this problem is increasing and, therefore, it is crucial to find reliable predictors of this condition before any symptoms appear. Despite its relevance and the several theories about its origins, SS cannot yet be quantitatively modeled and predicted. Our results indicate that, while neural correlates did not materialize, physiological measures may be a solid early indicator of oncoming SS.

KEYWORDS
human–computer interaction, EEG, motion sickness, peripheral physiological measures, simulator sickness, virtual reality

1 | INTRODUCTION

Simulations play an increasingly bigger role in our everyday lives. Be it in the form of driving simulators, therapy, or using the ever growing popularity of head-mounted display (HMDs) which allow us to delve into virtual reality (VR), simulations are becoming ubiquitous. However, there exists a problem common to all of these simulations: simulator sickness (SS).

The term SS describes symptoms that are quite similar to those of motion sickness (MS), but are caused by a motion experience solely induced by visual motion cues in an immersive environment and omitting physical cues of movement immersion in a simulation instead of physical movement.¹ Their origin lies on the mismatch between the received information through the vestibular system and our visual cortex. While they also enclose nausea and disorientation, they have a bigger emphasis on oculomotor forms, such as blurry vision, headaches, and dry eyes.² Common terms for this phenomenon also include cybersickness, virtual reality sickness, or the more general term visually induced MS, depending on the exact circumstances of the symptoms. Besides being inconvenient during simulations, in entertainment systems, or...
even more important in fields like telemedicine, SS imposes problems on several levels in research contexts. For one, there are ethical concerns when exposing users to immersive applications knowing they may experience discomfort as a result. In addition, SS limits the validity of results in simulation studies compared with real-world situations, as it is unclear whether differences in user performance are caused by SS rather than the conditions investigated in the study. While the susceptibility to SS varies depending on individual characteristics, it is assumed that any person with an intact vestibular system experiences SS when the conditions in the simulation are sufficiently sickness-inducing. Therefore, SS constitutes a ubiquitous problem.

The goal of this article is twofold. First, we aim to find reliable correlates of SS that are not side effects of the time period spent in a simulation necessary to induce it. Second, we want to assess these correlates in regard to their viability of predicting SS before the symptoms are noticeable to the participants. To this end, electroencephalography (EEG) data and other peripheral physiological data were collected from participants that were immersed in VR. The experimental group experienced a version of the environment that was likely to induce SS, while the control group remained in the normal environment. Once the data was statistically evaluated, we analyzed it using a variety of machine learning techniques to assess the predictive capabilities of the most promising correlates.

2 | RELATED WORK

The theories on the physiological mechanisms of MS can also be applied to SS. The sensory conflict theory is the oldest and describes a congruence conflict between expected signals as the basis of MS. Accordingly, Reason and Brand name the core principle of sensory rearrangement where a distorted perception modality becomes incongruent with the other in the vestibular system. Reason revised this theory and suggested that MS is based on a mismatch between the perceived and expected perceptions rather than between the perceptions themselves. Treisman followed an evolutionary standpoint with his toxin detector theory. MS may not be caused by the conflict itself but motor coordination conflicts that resolve to early warning sign of neurotoxin poisoning that can be removed from the body by vomiting. Riccio and Stoffregen argued that sensory conflicts are common also for nonsickness-inducing situations concluding that sensory conflict cannot be the only reason for MS and suggested that MS is caused by prolonged postural instability instead. In summary, the mechanisms of SS are still unclear although many assume some sort of central sensory conflict.

2.1 | Measuring SS

The simulator sickness questionnaire (SSQ), a widely used version of the MS questionnaire, was adapted for use with SS. It consists of 16 symptoms on the scales oculomotor discomfort, disorientation and nausea rated on a scale from 0 (none) to 3 (severe). While the SSQ is an economic tool to assess SS intensity, problems exist like being highly subjective or interfering with the experimental protocol due to oral feedback during a session.

To find an objective measurement for SS, various physiological measurements have been explored. Several authors investigated the correlation of cardiac measurements and MS or SS. Doweck et al. induced MS by a rotatory chair and simultaneous head movements and suggested the parasympathetic nervous system may be involved in MS. Gianaros et al. also investigated the involvement of the parasympathetic nervous system in MS and used a rotating optokinetic drum to induce MS and found a correlation between respiratory sinus arrhythmia. Martin et al. let participants play VR games using an HMD to induce SS with predictors being heart rate (HR) and electrodermal activity (EDA). Dennison et al. investigated the connection between SS and EDA, breathing, HR, blinking, skin conductance, stomach activity, and time between heart beats. They found HMD-viewed virtual environments to cause significantly higher SS scores than desktop-viewed.

Another proposed physiological measurement is EEG. In an early study by Wood et al., MS was induced by rotary chair while performing head movements. There was an increase in theta wave activity over the premotor cortex compared with baseline data gathered before the experiment. Numerous authors investigated MS in a driving simulator: Lin et al. found an increase in alpha and beta power over the parietal lobe and, for some participants, motor cortex. Chen et al. found a correlation of alpha power in the occipital, parietal and motor areas and MS. Lin et al.’s top three classifiers were beta and delta waves over the left motor area and gamma-waves over the occipital midline. Lin et al.’s best classifier was a broadband power increase at the occipital midline, especially alpha and gamma power. The following studies investigated SS in contexts that are most similar to the one used in this article. Kim et al. investigated SS while navigating a virtual environment using a steering wheel, an accelerator, and brakes while seated. The baseline was collected while participants
looked at a freeze-frame of the environment. They found an increase in delta power over F3 and T3, while there was a decrease in slow beta waves (12–20 Hz) at F3 and in all beta waves at T3 compared with the baseline. Naqvi et al.\textsuperscript{21} showed their participants a movie in two-dimensional (2D) and three-dimensional (3D). There was a decrease of theta power over the frontal and beta power over the temporal lobe during the 3D-movie, which elicited higher simulator-sickness scores than the 2D-movie. Pane et al.\textsuperscript{22} let participants play the video game Mirror’s edge. They only gathered occipital EEG data and found a decrease in beta power and unspecified change in theta waves.

3 | EXPERIMENTAL DESIGN

We conducted a controlled study on SS in VR collecting EEG data and other peripheral physiological measurements. Our setup allows us to compare changes caused by SS (experimental group, $G_{\text{exp}}$) to those caused by time spent in a nearly identical environment (control group, $G_{\text{ctrl}}$).

The experiment followed a single-factorial design. The independent variable was the virtual environment varying between “normal” for the ($G_{\text{ctrl}}$) and “sickness-inducing” for the ($G_{\text{exp}}$). The concurrent EEG-measurements constituted the dependent variable. A between-subjects-design was used as the experiment had to be terminated after the onset of severe SS for both ethical and practical reasons, which prohibited balancing the order of conditions. This also prevented influences stemming from demand characteristics. Data on SS susceptibility was collected to control for potential side effects.

3.1 | Stimuli

The virtual environment had two modes: normal, and highly likely to induce SS, enabling a properly controlled design. In the normal condition as well as in the baseline for the experimental condition, the camera slowly (0.2 m/s) moved through the environment along the street of a virtual city in the desert (see Figure 1b). In the sickness-inducing mode used for the ($G_{\text{exp}}$), the camera sped up significantly (7 m/s) for random amounts of time between 4 and 7.5 s at random intervals between 4 and 6 s. In addition, the floor of the virtual environment swayed up and down slightly to induce MS according to previous research.\textsuperscript{23} This sickness induction would be increased by the camera speed-up according to sensory conflict theory (conflict caused by the visual perception of movement that is not perceived by the vestibular system, enhanced by speed in- and decreases and ground sway), neural mismatch theory (after the baseline period, the slow camera speed is expected and at odds with random peaks and ground sway), as well as postural instability theory (postural stability is compromised as one tries to compensate for the seeming movement of the environment, especially with the ground sway).

Several test runs confirmed the efficacy of this induction method. The randomized aspects of the environment were identical for all participants. While the camera movement in the control condition alone could be suspected to cause SS, a stationary control condition would cause SS to be confounded with the moving environment that may trigger, for example, motor-cortex responses or effects caused by the perception of moving objects in visual processing. Prior research shows that SS symptoms can occur in regular virtual environments after 8–10 min.\textsuperscript{1,14} To avoid sickness in the ($G_{\text{ctrl}}$), and as participants in the ($G_{\text{exp}}$) were likely resistant to SS if they did not experience symptoms after 5 min in an environment designed to cause it, this was chosen as the maximum duration of the experiment.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{(a) Dome projection system, (b) snapshot of a participant during the simulation, (c) illustration of our placement of the electrodes}
\end{figure}
3.2 | Apparatus

EEG data was measured using a 16-electrode g.tec Nautilus system with a sampling rate of 250 Hz with an earlobe electrode as reference (Figure 1c). Peripheral physiological data was collected using the Nexus-10 MK II system by Mind Media. We used a finger-mounted blood volume pulse (BVP) sensor and an abdominal belt to collect cardiac and respiration data, respectively. To measure EDA, two electrodes attached to the participant’s fingers.

The experiment was conducted in a dome projection system (dome) for an immersive experience (Figure 1a). The dome is a tilted full-dome real-time video projection system with a diameter of 5 m and six projectors with 2,560 × 1,600 px at up to 120 fps.24 This setup presents several advantages against using a HMD. First, the dome’s natural full field of view is not achievable with any available HMD. Second, the straps on the HMD would interfere with the EEG-measurements, lowering the signal to noise ratio.25 Finally, HMDs are known to be likely to cause SS,2 which while being an advantage in the (Gexp), would be detrimental to the (G ctrl). While the seated position in our dome setup allows for the cleanest measurements of physiological indicators, it is also the hardest subset of HMD experiences to induce sickness in. Therefore, the insights from our setup will also apply to most other computer graphics-relevant VR-contexts.

An X-Box controller was used by participants to provide feedback when they start to experience symptoms of SS and stopping the simulation in case of severe discomfort. Repeated feedback about symptom changes could be given by pushing the analog stick up or downwards. We also used these signals as markers for the EEG data. In order to assess SS, we used the SSQ9,26 prior to and postimmersion.

3.3 | Psychophysical methodology

3.3.1 | Participants

The sample consisted of 37 participants (19 females) with an average age of 22 (range: 19–34, SD = 2.91). Participants were compensated with 20€ and randomly assigned to the conditions. Fifteen of the participants suffered from SS throughout the experiment. Twelve of the sick participants were members of the (Gexp). Five participants chose to terminate the experiment early and all of them were members of the (Gexp). In consonance with prior results, the proportion of females (47.36%) who reported SS was slightly higher than that of males (33.33%).

3.3.2 | Procedure

The session started with an informed consent form, a demographic questionnaire including factors influencing susceptibility to SS, and a German translation of the SSQ26 (all participants were native German speakers). They were told to be testing the influence of a virtual environment run on a new graphics engine on EEG data. They were allowed to look around in the scene but at the same time instructed to minimize movement while sitting on a chair as to not compromise the EEG data, and not to speak during the experiment. They were asked to move the joystick on the controller they were holding if they noticed the onset of any symptom listed on the SSQ or press the emergency stop button to terminate the simulation in case of severe sickness. Then, both groups spent a baseline period of 110 s viewing the initial slow-ride virtual environment to let participants familiarize themselves with environment to diminish arousal- and novelty-reactions. After 110 s, the sickness-inducing mode set in for the (Gexp). Both groups then spent up to 5 min after onset of the experimental condition viewing their respective virtual environment if they did not choose to terminate the experiment at an earlier point in time. After the experiment, participants filled in the SSQ again and received information on the actual goal of the study and on how to ease SS.

4 | RESULTS AND ANALYSIS

4.1 | Self-assessment SSQ

We analyzed the SSQ-scores gathered after the experimental run through a mixed-design ANOVA. There was a significant interaction between time (before and after the experiment) and sickness occurrence regarding \(F(1, 35) = 12.596, p = .001\),
that is, there was a significant increase in symptoms in the sick group, while there was no such difference in the nonsick group, with both groups starting at a comparable sickness level. This validates the use of the nonsick group as control, as well as our sickness induction stimuli. There were no significant differences between the groups regarding factors influencing SS susceptibility.

The criteria for inclusion in the sick group were self-reported sickness during the experiment and an above average (average SSQ change was 0.237) SSQ-score increase. The inclusion criteria for the nonsick group were no self-reported sickness during the experiment and a below average SSQ-score increase. Twenty-eight participants fulfilled these requirements. Six participants reported sickness during the experiment but had below average SSQ-increases and four participants did not report sickness during the experiment but had above average SSQ-increases. Thus, all 10 participants were excluded from the analysis.

The resulting sample consisted of 27 participants (51.85% females) with an average age of 22 (range: 19–34, SD = 3.07). A third of participants reported sickness during the experiment, and 18.52% chose to terminate the experiment early. In addition, SS susceptibility was measured before the experiment for every participant. There were no significant differences between the sick and nonsick group in terms of SS susceptibility. A mixed-design ANOVA shows a significant interaction between time and group ($F(1,25) = 31.991, p < .001$) validating the use of the nonsick group as ($G_{\text{ctrl}}$) again.

### 4.2 Electroencephalography

We performed a time-frequency analysis using a multitaper method to extract the power spectral densities (PSDs) of alpha (8–13 Hz), beta (13–30 Hz), gamma (30–45 Hz), delta (1–4 Hz), and theta (4-8 Hz) waves from the raw data. All PSDs were decibel-normalized in comparison to the data from 15 to 5 s before the end of the baseline period to prevent any spillover-effect from changes resulting from the condition change. As not all members of the ($G_{\text{exp}}$) experienced SS, while some members of the ($G_{\text{ctrl}}$) did, the data were split by whether a participant reported SS during the experiment rather than condition. PSDs from the sick and nonsick participants were compared with determine changes in EEG data unique to SS. To determine significance, we used the pixel-based multiple comparisons correction\textsuperscript{27} for explorative EEG-analysis. This method is more suitable for explorative EEG-analysis than Bonferroni correction as the latter assumes the data to be independent which is not the case for EEG and increases false negatives. For the first analysis, the PSDs extracted from 5 to 15 s after self-reported sickness onset were used as the data for the sick group to prevents any motor spillover-effect from the button press itself. The control data was extracted from 5 to 15 s after the average sickness onset (190.66 s after the start of the experiment, resulting in the samples starting at 196 and ending at 206 s after simulation start). This period was chosen to parallel any time effects in the sickness data as good as possible. When analyzing the mean differences of the PSDs of the sick and nonsick groups, we observed that the alpha waves were less prevalent in the sick group over all electrodes. Beta waves were slightly more prevalent over the temporal lobe and slightly less prevalent everywhere else. Gamma waves were more prevalent over all electrodes. Delta waves were less prevalent over all electrodes in sick individuals. Theta waves were slightly less prevalent over FP2, F4, Cz, P4, and PO8, while they were slightly more prevalent over FP1, T7, P3, and PO7. The pixel-based multiple comparisons correction with $10^5$ permutations per iteration resulted in cut-off values of $-2.86$ and $2.74$ dB for significant de- and increases, respectively, at a 5% level. None of the group differences reached these thresholds.

For the second analysis, the time frame of interest were the last 15–5 s before termination of the experiment, as the postimmersion SSQ-scores used for the inclusion criteria were obtained immediately after the end of the experiment, making the last seconds of the experiment most representative of these values. This data was compared with the last 15–5 s before the average time of termination (i.e., starting 382.15 s after the start of the experiment) from the ($G_{\text{ctrl}}$). Looking at the mean differences of the PSDs of the sick and nonsick group we can observe that Alpha waves were more prevalent over the all electrodes except Cz and P4, where they hardly differed between sick and nonsick participants. Beta waves were notably more prevalent over the temporal lobe, FP1, F3, and Oz and slightly more prevalent over Fz, F4, the motor cortex, P3, Pz, and PO7. There was barely any difference concerning beta waves between the groups over the other electrodes. Gamma waves were more prevalent over all electrodes except FP2, over which they were less prevalent. Delta and theta waves were both more prevalent in the sick group over all electrodes. The pixel-based multiple comparisons correction with $10^5$ permutations/iteration resulted in cut-off values of $-2.31$ and $2.96$ dB for significant de- and increases, respectively, at a 5% level. The differences failed to reach significance again.

Third, we also hypothesized sickness to be related to other physiological markers and found skin conductance to be a relevant physiological correlate. Here, we analyzed the EEG frequency data for a duration of 15 s after the highest peak
in tonic level after the sickness mode start. This allowed us to set the analysis region to relevant positions even when a participant did not actively report any sickness. Figure 2a shows an example for a single participant of the sickness condition analysis region shifted from the medium user reported sickness onset (196 s after experiment start) towards the individual tonic peak (both shown in red).

### 4.2.1 Time-dependent effects

EEG is a highly nonstationary process and its effects may evolve and change over time. The individually defined analysis areas above did not show significance, therefore the subsequent analysis focuses on the entire time-course of the experimental sessions instead of just the single windows as previously described. Thus, we used the pixel-based multiple comparisons correction method over the full spectrum with 1,000 permutations per iteration resulting in cut-off values of −1.68 and 1.80 dB for significant de- and increases, respectively, at a 5% level. No clusters were found for any electrode as no regions were statistically significant.

Common spatial pattern (CSP) is a supervised data-driven method that decomposes the signal and projects it into a subspace in which the variance of one class is maximized while the other is minimized. This increases the discriminability between conditions and bypasses the channel selection problem. A thorough overview of CSP is given in Reference 28. The CSP were computed for the individual EEG wave bands and used for classification using linear regression. The gamma waveband exhibits highest classification accuracy during the sickness condition (see Figure 2b) compared conventional channel-based statistics indicating a possible presence of a neural marker.

### 4.3 Peripheral physiological signals

#### 4.3.1 Electrodermal activity

It is also known as skin conductance that varies with the activation of sweat glands. EDA can be considered as one of the most common observation channels of sympathetic nervous system activity, and manifests itself as a change in electrical properties of the skin, such as skin conductance.

The model $cvxEDA$ we employ describes skin conductance as the sum of three terms: the phasic component (skin conductance response, SCR), the tonic component (skin conductance level, SCL), and an additive white Gaussian noise term incorporating model prediction errors as well as measurement errors and artifacts. This model is physiologically inspired and fully explains EDA through a methodology based on Bayesian statistics, convex optimization, and sparsity. The modeled EDA generation process includes the following assumptions that make it suitable for the analysis conducted as described further below. Bursts from the sudomotor nerves (sudomotor nerve activity, SMNA) that control the sweat
glands precede the SCRs. These bursts are temporally discrete episodes,\(^\text{30}\) that is, SCRs are generated by a sparse, non-negative neural signal due to the nature of a nerve activity. There is a linear relationship between the number of active sweat glands and the amplitude of a firing burst\(^\text{30}\) and a single a single neural burst that evokes a SCR is not influenced by previous ones.\(^\text{31}\) This phasic activity is superimposed to a slowly varying tonic activity with spectrum below 0.05 Hz,\(^\text{32}\) so its information content can be represented by samples spaced every 10 s.\(^\text{33}\)

These assumptions allow us to analyse the SCR in two ways: the discrete integration (Simpson’s rule) of the SMNA peaks and the count of the peaks. For comparable time frames, we used the entire baseline duration of 110 s and a corresponding time window of 110 s length after the sickness onset for both \((G_{\text{exp}})\) and \((G_{\text{ctrl}})\). Then, the ratios of condition/base were compared in subsequent analyses. We found both integral \((p = .006, t = −2.95)\) and count \((p = .049, t = −2.06)\) to be significantly different between \((G_{\text{exp}})\) and \((G_{\text{ctrl}})\) (ind. \(t\)-test).

Now, we found a change in overall skin conductance between the two groups, so the next step is to test it for signs of SS before the user actually reports discomfort. For this, we analyse the participants’ data from the \((G_{\text{exp}})\), that actually reported sickness during the experiment. We paired the duration of this nonsick period during the experimental condition with a time frame before the condition start (sickness start) of the same length. Again, ratios of condition/base were compared. For the \((G_{\text{ctrl}})\) the same duration (110 s) as before has been used because \((G_{\text{ctrl}})\) and \((G_{\text{exp}})\) are not directly paired. These numbers also show a significant difference \((p = .008, t = 2.96)\).

The slower tonic SCL response might also be an indicator of SS. Again both groups were baseline corrected (condition/base). We found the tonic level to be significantly higher during the sickness mode as compared with \((G_{\text{ctrl}})\) \((p = .002, t = 3.33)\).

### 4.3.2 Heart data

Various measures derived from the BVP were analyzed. The BVP amplitude is computed from the raw BVP signal and is an indicator of relative blood flow. We found the BVP to significantly differ between \((G_{\text{exp}})\) and \((G_{\text{ctrl}})\) (ind. \(t\)-test \(p = .01, t = −2.90)\).

The HR is derived from the raw BVP signal. In our experiment, we could not find a correlation between HR and SS (ind. \(t\)-test \(p = .09, t = −1.71)\).

HRV is variations between subsequent heartbeats modulated by sympathetic and parasympathetic nerves of the autonomic nervous system. Our results suggest a tendency but fail to reach significance for HRV amplitude (ind. \(t\)-test \(p = .06, t = −1.88)\). In addition, a combination of the low-to-high frequency ratio did not show significant differences.

### 4.3.3 Respiration data

It was analyzed using time domain derived features (respiration rate and amplitude), phase space diagrams, Fourier spectra, and Lyapunov exponents (LLEs). Again, we compare the baseline corrected mean values of \((G_{\text{exp}})\) and \((G_{\text{ctrl}})\), with both time windows being 110 s. A normal respiration rate is shown in our results with no significant difference between \((G_{\text{exp}})\) and \((G_{\text{ctrl}})\) \((p = .36, t = −0.92)\). This seems to be consistent with previous research, which has indicated that respiration quality and rhythm relates to susceptibility to MS but not to occurrence or severity.\(^\text{34}\) Second, respiration stability was assessed by analyzing the power spectrum using Welch’s method. We found no regularity changes across the \((G_{\text{exp}})\) and \((G_{\text{ctrl}})\). Respiration amplitude was visually analyzed using phase space diagrams that show velocity as a function of sensor displacement and is useful for visualizing oscillatory processes like respiration.\(^\text{35}\) Visually, as well as quantitatively, respiration amplitude is not affected by the experimental manipulation \((p = .69, t = 0.41)\). Finally, using Rosenstein’s LLE method, we analyse respiration wave forms by considering the respiratory system as a dynamical and chaotic system.\(^\text{35}\) A positive LLE is a sign of chaos and shows instability in a particular direction whereas a negative LLE represents the tendency to converge towards a stable state. Our results show LLE values close to zero (0.005) with a positive tendency. We could not find a significant difference between the LLEs for the \((G_{\text{exp}})\) and \((G_{\text{ctrl}})\) \((p = .70, t = 0.39)\).

### 5 Predictive capabilities

After finding several peripheral physiological correlates to be statistically significant, we are interested in their potential as predictors for the onset of SS before the participants are aware of any symptoms. Here, we train several classifiers on
### TABLE 1
Classification results sorted by accuracy for: EDA, BVP amplitude, and EDA and BVP

| #  | EDA       | Heart                        | EDA+ Heart                   |
|----|-----------|------------------------------|------------------------------|
|    | Phasic    | Tonic                        | BVP amplitude                 |                             |
|    | Classifier | Accuracy                     | Classifier                    | Accuracy                     |
| 1  | NB-PCA    | 0.933                        | LR + PT-YJ                    | 0.9                         | NB + Normalizer              | 0.85                        | CART-PCA + RS                | 0.967                        |
| 2  | MLP-PCA   | 0.917                        | MLP                           | 0.833                       | LDA-PCA + QT-N               | 0.75                        | RF-PCA + QT-N                | 0.967                        |
| 3  | RF-PCA    | 0.917                        | MLP-PCA + RS                  | 0.783                       | LR + Normalizer              | 0.75                        | CART                          | 0.933                        |

Abbreviations: BVP, blood volume pulse; EDA, electrodermal activity; LDA, linear discriminant analysis; LR, logistic regression; MLP, multilayer perceptron; PCA, principal component analysis; QT-N, quantile-transformer (normal); RF, random forest.

the data to verify its predictive abilities. Specifically, for training, we used the most commonly used, well-known learning techniques: logistic regression (LR), linear discriminant analysis, k-nearest neighbor, decision tree (CART), Gaussian Naive Bayes (NB), support vector machine, random forest, and MLP. As scaling may influence data processing, especially with machine learning methods, we also employed a number of different scalers: standard scaler, min–max scaler, max–abs scaler, robust scaler (RS), quantile-transformer (normal), quantile-transformer (uniform), power transformer (Yeo Johnson), and the normalizer. In addition, principal component analysis (PCA) is known to be influenced by different scaling methods, so we added PCA to the pipeline. For training, the data was split into train (80%) and test sets (20%) and results are reported as accuracy scores from a 10-fold random cross-validation splits.

Skin conductance data exhibits the best results for classification when using Gaussian NB with PCA for phasic response with an accuracy of 93.3% (Table 1). It is closely followed by its tonic level counterpart with an accuracy of 90% when using LR with the power transformer (Yeo-Johnson). The significant measure is BVP amplitude that leads to a classification accuracy of 85% with the NB approach using normalization (Table 1). The most informative correlates found so far are phasic response and tonic level from EDA data and BVP amplitude derived from cardiac data. While BVP amplitude and both EDA measures provide good classification results on their own, the question remains whether less significant data may improve classification results. For this purpose, features were constructed from the data above by computing the means over conditions and their respective ratio of sickness condition to baseline for the \((G_{exp})\) and \((G_{ctrl})\). Fusion of the data slightly improved classification accuracy as compared with the strongest classifier learnt from BVP amplitude. Therefore, joining both data modalities, a decision tree with PCA and the RS yield an accuracy of 96.7% (Table 1).

### 6 DISCUSSION AND CONCLUSION

The aim of this study was to find neural and peripheral physiological correlates of SS and to investigate the potential of these correlates as predictors of SS. The current study indicates the feasibility of certain peripheral physiological correlates to act as a predictor of SS, but was unable to find any correlates of SS in EEG data.

SS is a complex, multisymptomatic phenomenon that may not be distinguishable using frequency-band-activity on its own. Nevertheless, other analysis methods that take fluctuations over time into account rather than averaging neural activity over a set period of time also did not provide further insights. Even assuming the existence of a set number of neural correlates of SS that are unique to it, EEG may not be suited to measuring them. Considering the potential involvement of the sympathetic nervous system, the neural activity related to SS would likely stem from the hypothalamus, which deep location within the brain makes its activity hardly measurable on the surface of the skull. This could result either on EEG being unable to find any correlates, or to be clouded by randomly activation patterns at the surface of the brain rendering false positives on neural correlates. Thus, it could be advisable to look into SS using other imaging techniques with a focus on spatial rather than temporal precision, such as fMRI or fNIRS. Nevertheless and despite the absence of indicators present in the EEG data, we found correlates within the peripheral physiological data. Since it cannot be told from the EEG data alone whether SS was present, the other biological signals provide better insight. As previous studies suggest, skin conductance and heart data might not properly distinguish between just an aroused state or sickness itself. Assuming that the experimental manipulation actually did induce SS as represented within the biofeedback, the hypothesis holds that EEG is not the right tool. Otherwise, if we actually just measured arousal stemming from, for example, novelty reaction to scene change or just the fast simulator ride instead of the sickness state, this
would not help quantifying SS under real-world condition, for example, while playing an immersive action shooter or horror game.

Our results also indicate higher skin conductance present in the signal between the time of the onset of the sickness condition as compared with the time after participants reported sickness by pressing a button. If it is not the sickness state present in the data after the condition start but solely arousal due to the sudden speed onset, then this change to even higher skin conductance values indicate the presence of sickness after the participant reported discomfort. In addition, classification rates at an acceptable level were achieved already for EDA data alone. Therefore, the data strongly suggests the presence of SS.

6.1 | CONCLUSION

On the one hand, our experimental results suggest that both EDA and HR data (BVP Amplitude) have great potential as SS predictors.

On the other hand, the current findings indicate a dire need of more controlled studies in this field of research to exclude time effects as the actual cause for the neural correlates found so far in EEG. To have a better insight into neural components involved in SS, using other analysis paradigms with EEG data or considering fMRI or fNIRS studies could help identifying the neural foundation of SS.

ACKNOWLEDGEMENT

This work is funded by the German Science Foundation (DFG MA2555/15-1 Immersive Digital Reality and DFG INST 188/409-1 FUGG IC Dome). Open access funding enabled and organized by Projekt DEAL.

ORCID

Jan-Philipp Tauscher https://orcid.org/0000-0003-2407-391X
Susana Castillo https://orcid.org/0000-0003-1245-4758

REFERENCES

1. Dużmańska N, Strojny P, Strojny A. Can simulator sickness be avoided? a review on temporal aspects of simulator sickness. Front Psychol. 2018;9:2132.
2. Kennedy R, Drexler J, Kennedy R. Research in visually induced motion sickness. Appl Ergonomics. 2010;41(4):494–503. Special Section - The First International Symposium on Visually Induced Motion Sickness, Fatigue, and Photosensitive Epileptic Seizures (VIMSS2007).
3. Stoner H, Fisher D, Mollenhauer M. Simulator and scenario factors influencing simulator sickness. D Fisher, M Rizzo, J Caird, J Lee, (eds.) Handbook of driving simulation for engineering, medicine, and psychology. Boca Raton, Florida: CRC Press; 2011:14.11-14.24.
4. Golding J. Motion sickness susceptibility. Autonom Neurosci. 2006;129(1):67–76. Nausea and Vomiting: An Interdisciplinary Approach.
5. Reason J, Brand J. Motion sickness. London, UK: Academic Press, 1975.
6. Reason J. Motion sickness adaptation: A neural mismatch model. J Royal Soc Med. 1978;71(11):819–829.
7. Treisman M. Motion sickness: An evolutionary hypothesis. Science. 1977;197(4302):493–495.
8. Riccio G, Stoffregen T. An ecological theory of motion sickness and postural instability. Ecol Psychol. 1991;3(3):195–240.
9. Kennedy R, Lane N, Berbaum K, Lilienthal M. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. Int J Aviat Psychol. 1993;3(3):203–220.
10. Keshavarz B, Hecht H. Validating an efficient method to quantify motion sickness. Hum Factors. 2011;53(4):415–426.
11. Doweck I, Gordon C, Shlittner A, et al. Alterations in R–R variability associated with experimental motion sickness. J Autonom Nervous Syst. 1997;67(1-2):31–37.
12. Gianaras P, Quigley K, Muth E, Levine M, Vasko R Jr, Stern R. Relationship between temporal changes in cardiac parasympathetic activity and motion sickness severity. Psychophysiology. 2003;40(1):39–44.
13. Martin N, Mathieu N, Pallamin N, Ragot M, Diverrez JM. Automatic recognition of virtual reality sickness based on physiological signals. IBC 2018. Amsterdam, Netherlands, 2018.
14. Dennison M, Wisti A, D’Zmura M. Use of physiological signals to predict cybersickness. Displays. 2016;44:42–52. Contains Special Issue Articles – Proceedings of the 4th Symposium on Liquid Crystal Photonics (SLCP 2015).
15. Wood C, Stewart J, Wood M, et al. Habituation and motion sickness. J Clin Pharmacol. 1994;34(6):628–634.
16. Lin C, Chuang S, Chen Y, Ko L, Liang SF, Jung T. EEG effects of motion sickness induced in a dynamic virtual reality environment. Poster session presented at: Proceedings of the 29th Annual International Conference of IEEE-EMBS, Engineering in Medicine and Biology Society, EMBC’07. Annual International Conference of the IEEE Engineering in Medicine and Biology, Lyon, France; 2007. p. 3872–3875. https://doi.org/10.1109/EMBS.2007.4353178
17. Chen YC, Duann JR, Chuang SW, et al. Spatial and temporal EEG dynamics of motion sickness. NeuroImage. 2010;49(3):2862–2870.
18. Lin C, Tsai S, Lee H, Huang H, Ho S, Ko L. Motion sickness estimation system. Proceedings of the 2012 International Joint Conference on Neural Networks (IJCNN), Brisbane, Australia; 2012. p. 1–6. https://doi.org/10.1109/IJCNN.2012.6252580.

19. Lin C, Tsai S, Ko L. EEG-based learning system for online motion sickness level estimation in a dynamic vehicle environment. IEEE Trans Neural Netw Learn Syst. 2013;24(10):1689–1700.

20. Kim Y, Kim H, Kim E, Ko H, Kim H. Characteristic changes in the physiological components of cybersickness. Psychophysiology. 2005;42(5):616–625.

21. Naqvi S, Badruddin N, Jatoi M, Malik A, Hazabbah W, Abdullah B. EEG based time and frequency dynamics analysis of visually induced motion sickness (VIMS). Austr Phys Eng Sci Med. 2015;38(4):721–729.

22. Pane E, Khoirunnisa A, Wibawa A, Purnomo M. Identifying severity level of cybersickness from EEG signals using CN2 rule induction algorithm. Proceedings of the 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), vol 3. Bangkok, Thailand; 2018. October 3, p. 170–176. https://doi.org/10.1109/ICIIBMS.2018.8549968.

23. Golding J, Finch M, Stott J. Frequency effect of 0.35-1.0 Hz horizontal translational oscillation on motion sickness and the somatogravic illusion. Aviat Space Environ Med. 1997;68(5):396–402.

24. Grogoric S, Ueberheide M, Tauscher J, Bittner P, Magnor M. Gaze and motion-aware real-time dome projection system. Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan; 2019. p. 1780–1783. https://doi.org/10.1109/VR.2019.8797902.

25. Tauscher J, Schottky F, Grogoric S, Bittner P, Mustafa M, Magnor M. Immersive EEG: Evaluating electroencephalography in virtual reality. Proceedings of the IEEE Virtual Reality (VR) Workshop. PerGraVAR, Osaka, Japan; 2019. p. 1794–1800. https://doi.org/10.1109/VR.2019.8797858.

26. Neukum A, Grattenthaler H. Kinetose in der Fahrsimulation (Projekt: Simulation von Einsatzfahrten im Auftrag des Präsidiums der Bayerischen Bereitschaftspolizei, Abschlussbericht - Teil II). Würzburg, Germany: Interdisziplinäres Zentrum für Verkehrswissenschaften an der Universität Würzburg, 2006.

27. Cohen M. Analyzing neural time series data: Theory and practice. Cambridge, MA: MIT Press, 2014.

28. Bosse S, Acqualagna L, Samek W, et al. Assessing perceived image quality using steady-state visual evoked potentials and spatio-spectral decomposition. IEEE Trans Cirt Syst Video Technol. 2017;28(8):1694–1706.

29. Greco A, Valenza G, Lanata A, Scilingo EP, Citi L. cvxEDA: A convex optimization approach to electro-dermal activity processing. IEEE Trans Biomed Eng. 2016;63(4):797–804.

30. Nishiyama T, Sugenoya J, Matsumoto T, Iwase S, Mano T. Irregular activation of individual sweat glands in human sole observed by a videomicroscopy. Autonom Neurosci. 2001;88(1-2):117–126.

31. Bach D, Flandin G, Friston K, Dolan R. Modelling event-related skin conductance responses. Int J Psychophysiol. 2010;75(3):349–356.

32. Bach D, Friston K, Dolan R. An improved algorithm for model-based analysis of evoked skin conductance responses. Biol Psychol. 2013;94(3):490–497.

33. Boucsein W. Electrodermal activity. New York, NY: Springer Science+ Business Media, 2012.

34. Dennison M Jr, D’Zmura M, Harrison A, Lee M, Raglin A. Improving motion sickness severity classification through multi-modal data fusion. Artificial intelligence and machine learning for multi-domain operations applications. Volume 11006 Bellingham, Washington: International Society for Optics and Photonics, 2019; p. 110060T.

35. Son H, Jeong Y, Yoon H, Park JH, Kang DY. Visual and quantitative analysis methods of respiratory patterns for respiratory gated PET/CT. BioMed Res Int. 2016;2016:1-11. https://doi.org/10.1155/2016/7862539.

AUTHOR BIOGRAPHIES

Jan-Philipp Tauscher received his M.Sc. degree in computer science with a special emphasis on visual computing and musical neurophysiology from the Technische Universität Braunschweig (TU Braunschweig), Germany and Queen Mary University of London, UK in 2015. He is currently working towards his PhD degree at the Computer Graphics Lab, TU Braunschweig. His research interests include a combination of computer graphics and neuroscience for understanding human behavior and computer interaction.

Alexandra Witt received her B.Sc. degree in psychology with special emphasis on electrophysiology from Technische Universität Braunschweig (TU Braunschweig), Germany in 2019. She pursues her M.Sc. at the International Max Planck Research School (IMPRS) for Neurosciences at the Georg-August-Universität Göttingen, where she is currently investigating neural correlates of free choice.
Sebastian Bosse is a senior researcher with the Machine Learning group and the Image Video Coding group in the Video Coding & Analytics Department of Fraunhofer Institute for Telecommunications—Heinrich Hertz Institute, Berlin, Germany. He studied Electrical Engineering and Information Technology at RWTH Aachen University, Aachen, Germany and Polytechnic University of Catalonia, Barcelona, Spain. He received the Dipl.-Ing. degree in from RWTH Aachen University and the Dr.-Ing. degree in Electrical Engineering and Computer Science from Berlin University of Technology, Germany. As a student, Sebastian was a visiting researcher at Siemens Corporate Research, Princeton, USA. In 2014, he was a guest researcher at Stanford University, USA. His major research interests are image and video compression, human visual perception for image communications, and neural correlates of perceived image quality and quality of experience, as well as signal processing.

Fabian W. Schottky received his B.Sc. degree in psychology with special emphasis on cognition and behavior from Technische Universität Braunschweig (TU Braunschweig), Germany in 2019. His research interests include human social interaction and personality.

Steve Grogorick received his M.Sc. degree in computer science with a special emphasis on graphical systems and psychophysics from the Technische Universität Braunschweig (TU Braunschweig), Germany, in 2015. He is currently working toward his PhD degree at the Computer Graphics Lab, TU Braunschweig. His current research interests include visual perception and gaze-contingent rendering, with an emphasis on virtual environments.

Susana Castillo Alejandre is working as a post-doc for the Computer Graphics Lab of the Computer Science Department at TU Braunschweig since 2019. She got her computer engineering degree from the University of Zaragoza (Spain) where she worked as a Junior Researcher with the Graphics and Imaging Lab. She received her PhD in computer science while working with the Department for Graphical Systems at BTU Cottbus-Senftenberg (Germany). Her research work to date has focused on working toward a tighter integration of psychology into computer science. Her fields of interest span facial animation, multimodal communication, perception, image and video manipulation, and virtual reality.

Marcus Magnor is full professor of computer science and director of the Computer Graphics Lab at Technische Universität Braunschweig (TU Braunschweig), Germany. He also holds an appointment as adjunct professor in the Department of Physics and Astronomy at the University of New Mexico, USA. He is alumnus of the Alexander von Humboldt Foundation, a former Fulbright scholar, elected member of the Braunschweigische Wissenschaftliche Gesellschaft, and laureate of the Wissenschaftspreis Niedersachsen. His research areas include, but are not limited to, computer graphics, vision, visual perception, image processing, computational photography, astrophysics, imaging, optics, visual analytics, and visualization.

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Tauscher J-P, Witt A, Bosse S, et al. Exploring neural and peripheral physiological correlates of simulator sickness. Comput Anim Virtual Worlds. 2020;31:e1953. https://doi.org/10.1002/cav.1953