Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Using Heart Rate Variability and Eye-Tracking to Differentiate Simulator Sickness Severity Rated by Visual Analogue Scale

Guilherme H. Elcadi

guilherme.elcadi@foi.se

Swedish Defence Research Agency

Marcus Lindskog

Swedish Defence Research Agency

Robin Dahlkvist

Swedish Defence Research Agency

Christian K.K. Lindberg

Swedish Defence Research Agency

Tina Karlsson

Swedish Defence Research Agency

Jouni Lindqvist

Swedish Defence Research Agency

Research Article

Keywords: Heart Rate Variability, Eye-tracking, Simulator Sickness, Visual Analogue Scale, pupil coherence

Posted Date: June 14th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6810122/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Virtual Reality on May 13th, 2026. See the published version at https://doi.org/10.1007/s10055-026-01390-7.

Abstract

Simulator sickness (SS) is an obstacle to extensive adoption of virtual reality (VR) systems. While the Simulator Sickness Questionnaire (SSQ) is the standard for assessing SS symptoms, it lacks temporal resolution to monitor symptom progression during exposures. In the present study, we examined heart rate variability (HRV) and eye-tracking metrics as physiological markers of SS, using a Visual Analogue Scale (VAS) to gauge symptoms concurrently. Thirty subjects completed up to five virtual roller coaster rides while electrocardiography (ECG) and eye-tracking data were continuously recorded. VAS ratings were collected at baseline and after each ride and showed a strong positive correlation with post-exposure SSQ total scores ($r = 0.78$, $p < .001$). Among our metrics, from HRV, low-frequency/high-frequency (LF/HF) ratio was significantly associated with VAS ratings ($b = 0.26$, $\chi^2 = 4.98$, $p = .03$), indicating a move toward sympathetic dominance as symptoms intensified. From eye-tracking, pupil coherence emerged as a robust predictor of VAS ($b = 4.31$, $\chi^2 = 4.10$, $p = .04$), while other time-domain HRV (e.g., RMSSD, SDNN) and eye-tracking (e.g., blink rate, mean pupil size) metrics were not significantly related to ratings (all p:s > .05). In a combined linear mixed-effects model, both LF/HF ratio ($b = 0.26$, $p = .02$) and pupil coherence ($b = 5.53$, $p = .02$) remained significant predictors of VAS ratings. These findings support the use of VAS and multimodal physiological metrics, specifically LF/HF and pupil coherence as indicators of SS symptom levels, to aid adaptive strategies for mitigating SS in immersive environments.

Introduction

Simulator sickness (SS), a general cause of discomfort often experienced in virtual reality (VR) and in simulators, is one of the main causes preventing the widespread use of immersive technologies (Voinescu et al., 2023). Characterized by symptoms such as dizziness, nausea, and visual discomfort, SS is a form of cyber sickness arising from conflicts between visual, vestibular, and proprioceptive responses (Kennedy et al., 1993; Reason, 1978). These conflicts impair the brain's ability to integrate motion and orientation signals, triggering the hallmark symptoms (Reason & Brand, 1975).

With the growing use of 3D headsets in training, entertainment, and therapy, SS has become an increasingly important area of study (Weech et al., 2019). Although the precise etiology remains unclear, symptoms of SS can be broadly grouped into three categories: disorientation (D), e.g., dizziness and vertigo, oculomotor (O), such as eyestrain and blurred vision, and nausea (N), including stomach awareness and salivation (Kennedy et al., 1993; Rebenitsch & Owen, 2016). Several theories have attempted to explain the mechanisms behind SS. The most influential, sensory conflict theory, proposes that conflicting inputs, especially between vision and the vestibular system, elicit discomfort (Lackner & DiZio, 2006; Reason, 1978). Previous neuroimaging studies have shown that such conflicts activate regions like the insula and limbic system, responsible for autonomic responses (Napadow et al., 2013; Stephan et al., 2005).

In contrast, the eye movement theory attributes SS to oculomotor strain and extraocular proprioception stimulating the vagus nerve, contributing to emetic stimuli (Palmisano et al., 2020). Additionally, the evolutionary theory posits SS as a protective mechanism against neurotoxic ingestion, although it cannot fully explain non-nausea symptoms like eyestrain or disorientation (Treisman, 1977).

While the Simulator Sickness Questionnaire (SSQ) (Kennedy et al., 1993) is the most commonly used subjective assessment tool for SS, its pre-post format may miss detection of rapid symptom fluctuations, particularly during short VR exposures. In contrast, the Visual Analogue Scale (VAS) enables concurrent monitoring of symptom intensity, offering a more temporally sensitive method for capturing the onset and escalation of SS (Gavgani et al., 2017; Nalivaiko et al., 2015). In addition, the VAS provides a simple, sensitive method to track symptom intensity over time where participants rate their discomfort on a continuous scale, capturing the dynamics of SS more precisely (Garrido et al., 2022; Gavgani et al., 2017; Nalivaiko et al., 2015).

Alongside subjective reports, a growing body of work has explored physiological measures such as heart rate variability (HRV), eye-tracking, and electrodermal activity to better understand SS and its autonomic correlates (Baldini et al., 2024; Ishaque et al., 2023; Tian et al., 2022).

HRV is a well-established, non-invasive marker of autonomic nervous system (ANS) balance. Reductions in HRV, particularly in high-frequency (HF) power or root mean square of successive differences between heart beats intervals (RMSSD), suggest increased sympathetic arousal or parasympathetic withdrawal in response to stress or sensory mismatch (Malińska et al., 2015; Stauffert et al., 2020). Notably, nausea has been linked to vagal withdrawal and brainstem activation (Babic & Browning, 2014), disorientation to reduced HRV and impaired sensory integration, and oculomotor strain to parasympathetic withdrawal tied to visual-vestibular load (Yang et al., 2023). These distinct patterns suggest that HRV could serve as an early physiological marker of specific SS symptom clusters.

Similarly, eye-tracking provides valuable insights into oculomotor strain and related vestibular effects. Measures such as saccade dynamics, fixation stability, and pupil dilation offer objective indicators of visual overload, cognitive load, and sensory conflict (Chen et al., 2017; Dymczyk et al., 2024; Lopes et al., 2020). For example, increased pupil dilation may signal autonomic arousal linked to disorientation or nausea (Balaban, 1999; Kourtesis et al., 2024).

The primary aim of this study was to evaluate HRV and eye-tracking metrics as sensitive physiological markers of SS onset and intensity, using concurrent symptom ratings captured via a VAS. By prioritizing the VAS over traditional measures like the SSQ, we aimed to establish a more temporally precise and responsive framework for assessing SS in immersive environments. Specifically, the study investigated whether changes in autonomic and oculomotor responses align with subjective reports of discomfort collected during short simulator exposures, thus offering a novel way to monitor and interpret SS dynamics as they unfold.

Materials and Methods

Participants

For this study, 30 subjects consisting of 25 males (34.1 ± 10.0 yrs) and 5 females (33.6 ± 12.4 yrs) participated. All subjects were in good health and free of discomfort. Prior to testing, subjects completed a questionnaire concerning their level of general health and none reported any medical condition that required any kind of regular medication. Subjects were also asked about their level of usage of 3D headsets and none reported extensive experience or continuous usage. Subjects with a BMI > 30 were excluded as a proxy for excessive subcutaneous adipose tissue, due to its association with autonomic imbalance, including reduced parasympathetic tone, (Yadav et al., 2017), thus one additional subject was excluded before testing. Subjects signed informed consents prior to the participation in the study and were given verbal and written descriptions of the experimental protocols. Table 1 shows anthropometric as well as other personal data for all subjects. Subjects were instructed not to perform unfamiliar exercise 24 hours before the experiment or drink coffee or highly caffeinated beverages for at least 1 hour prior to experiments. The study was approved by the Swedish Ethical Review Authority (Dnr 2024-04423-01) in accordance to the Declaration of Helsinki on the use of human subjects in research.

Table 1
Anthropometric and personal data of all subjects ± standard deviation (SD)

	Subjects	**SD**
Age (years)	34.1	10.2
Height (cm)	177.1	9.4
Weight (kg)	76.3	12.3
BMI (kg/m^2)	24.3	3.1
SpO2 (%)	98.1	1.1
Rest Pulse (bpm)	71.8	16.0
Blood Pressure Systolic (mmHg)	129.1	10.7
Blood Pressure Diastolic (mmHg)	82.1	6.4

Experimental set-up and Protocol

Experimental set-up and Protocol

Prior to the start of the experimental testing protocol, subjects underwent an array of preparation procedures. That included a short clinical examination to check blood pressure, arterial partial pressure of oxygen with a finger oximeter and to identify the ECG electrodes locations. All subjects had normal

blood pressure levels at the time of the experiment (Table 1). Thereafter, subjects completed the SSQ a first time and ECG electrodes were placed.

Next, the 3D headset (XR HMD Varjo XR-4, Helsinki, FI) was placed on the subjects' head, adjusted to fit each subject individually, and a calibration procedure for the eye-tracker was conducted, including automatic adjustment of the lenses for subjects' interpupillary distance. For the present study, a realistic 3D VR roller coaster simulation, Appalachian Escapade, developed using NoLimits 2 Roller Coaster Simulation (Version 2.6.6.0; Lange O. (2024)) was used to induce SS. It has been shown that roller coasters containing more realistic details have increased chance of inducing SS (Davis et al., 2015). Graphic realism and higher levels of visual flow have been previously underlined as contributing factors to SS (Kennedy, RS et al., 1988). We note however, that the current study was not designed to identify SS problems related to the used headset, specifically, or VR technology in general. We purposely chose a SS-provoking roller coaster simulation in an effort to provide conditions that would invoke SS.

After the 3D VR headset had been placed on the subjects' head, they were seated comfortably and safely in a flight simulator chair and ECG and eye-tracking measurements were started. Both ECG and eye-tracking data were recorded continuously throughout the experiment. The experiment consisted of six segments in VR, a baseline segment and, up to, five rides in the roller coaster simulation. After a baseline segment of 5 minutes, subjects rated their discomfort for the first time with the VAS scale and then the first roller coaster simulation ride began. Each ride consisted of a 2.53 minutes roller coaster simulation that subjects completed a maximum of five times with a short (~ 10 s) pause between. During the pause, subjects submitted VAS ratings with ratings between 0 and 10, where 0 meant "no discomfort at all" and 10 extreme discomfort, to mark their level of discomfort (SS) during the just completed roller coaster simulator ride. Subjects could decide to stop the experiment at any time and without any restrictions. In case of nausea, the experimenter stopped the experiment immediately. At the end of the last roller coaster simulation ride, the 3D headset was removed and subjects rested for approx. 3 minutes before completing the SSQ questionnaire a second time.

Visual Analogue Scale (VAS)

To assess the subjective level of SS with a minimal time lag to the experience, a VAS was employed. The VAS is a widely used psychometric instrument for measuring subjective experiences such as pain, discomfort, or nausea (Gift, 1989; Nalivaiko et al., 2015). In this study, the scale ranged from 0 to 10, where 0 indicated "no discomfort at all" and 10 represented "extreme discomfort." Participants were instructed to mark their perceived level of discomfort after the baseline segment and during short pauses between roller coaster simulation rides. This approach allowed for high temporal resolution tracking of symptom progression. VAS ratings are considered sensitive to small changes in symptom severity and provide a continuous rather than categorical measure, which makes them especially suitable for capturing dynamic experiences such as SS (Gift, 1989; Wewers & Lowe, 1990). The tool has previously been validated for assessing nausea and motion-related discomfort in virtual environments (Gavgani et al., 2017; Nalivaiko et al., 2015). The simplicity of the VAS and its non-intrusive nature also

made it feasible for repeated administration between each simulator exposure, supporting its role as a rapid and intuitive metric in psychophysiological studies of cybersickness.

Simulator Sickness Questionnaire (SSQ)

The SSQ is a standardized self-report tool originally developed by Kennedy et al. (Kennedy et al., 1993) to assess symptoms of SS following exposure to virtual environments. It consists of 16 items covering common symptoms such as dizziness, nausea, eyestrain, and general discomfort, grouped into three subscales: nausea (N), oculomotor (O), and disorientation (D). Each symptom is rated on a four-point scale with accompanying labels (None/Slight/Moderate/Severe), and weighted scores are calculated for each subscale and total severity (Bimberg et al., 2020; Kennedy et al., 1993).

In this study, the SSQ was administered before and after exposure to confirm the presence of SS and to support validation of subjective VAS-ratings. However, our primary interest lies in using the VAS as a more dynamic and responsive tool to capture concurrent symptom progression. While the SSQ offers broad symptom categorization, it may be less suited for detecting rapid onset or subtle changes in discomfort during short simulation exposures.

Heart Rate Variability (HRV)

A bipolar electrocardiogram (ECG) was continuously recorded during the baseline segment and throughout the experimental protocol using the Biopac MP160 system (Biopac Systems Inc., Goleta, CA, USA), sampled at 2000 Hz. Surface electrodes (EL504 Ag/AgCl, Biopac Systems Inc., Goleta, CA, USA) were placed on the sixth left rib and the distal end of the sternum to ensure stable R-wave acquisition. During data acquisition, analog filtering was applied via the ECG100D amplifier to reduce low-frequency drift and high-frequency noise. The ECG data were then exported for offline analysis and processed in MATLAB R2024b (MathWorks, Natick, MA, USA). A second-order Butterworth bandpass filter (0.5–40 Hz) was applied digitally to isolate the QRS complex and suppress residual baseline drift and muscle artifacts.

Empirical Mode Decomposition (EMD) was subsequently performed, and the cleaned signal was reconstructed by summing intrinsic mode functions (IMFs 2–6), preserving QRS morphology while attenuating noise. R-peaks were detected using a hybrid method combining z-score-based peak prominence and amplitude thresholding. Detected peaks were retained only if they were separated by at least 0.3 seconds and passed a width-based filter to exclude broad T-waves. All detected R-peaks and reconstructed signals were visually inspected to confirm detection accuracy and signal quality. RR intervals (inter-beat intervals, IBIs) were computed from the final set of R-peaks. Ectopic or abnormal intervals were identified and corrected using linear interpolation. The resulting cleaned IBI series was then segmented for HRV analysis.

To investigate dynamic changes in autonomic regulation, HRV metrics were computed separately for each simulator ride (maximum five) and for the resting baseline segment. Segment boundaries were identified using event markers, and only R-peaks occurring within each segment were included. For each segment, time-domain metrics were calculated, including the standard deviation of NN intervals (SDNN), the root mean square of successive differences (RMSSD), and the percentage of successive IBIs differing by more than 50 ms (pNN50). Frequency-domain analysis involved interpolating the RR series using cubic spline interpolation at 4 Hz, followed by spectral power estimation using Welch's method. Although absolute LF and HF powers were computed internally, only the LF/HF ratio was retained for analysis as an index of sympatho-vagal balance, using standard frequency bands (LF: 0.04–0.15 Hz; HF: 0.15–0.4 Hz) ("Heart Rate Variability. Standards of Measurement, Physiological Interpretation, and Clinical Use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology," 1996). The LF/HF ratio was only calculated for segments with at least 64 RR intervals and HF power above 0.01 ms^2, to prevent inflated or unstable values resulting from extremely low parasympathetic activity. All HRV metrics were exported for statistical comparison across segments and participants.

Eye-tracking

Eye-tracking data was collected using the built in eye-tracker of the Varjo 3D head-set (XR HMD Varjo XR-4, Helsinki, FI). The eye-tracker has a sampling frequency of 200 Hz, accuracy of ≤ 1° and a precision of ≤ 0.2°. Eye-tracking data were recorded binocularly, but only the right eye was used for analysis, as both eyes showed comparable signal quality across participants (see Nyström et al. (Nyström et al., 2024) for a similar approach). Data were exported for offline analysis in Python 3.12 to calculate metrics related to blinks and pupil size.

To calculate the blink metrics of interest, we used the pupil signal of the right eye (pupil size in mm). In a first step, blinks were detected from the pupil signal using the method proposed by Nyström et al. (2024). From the detected blinks, we calculated the average blink duration and the blink rate (blinks / min) for the six experimental segments (baseline and up to five roller coaster simulation rides), separately.

The pupil metrics of interest were calculated after the pupil signal of the right eye had been submitted to a four step pre-processing pipeline. In the first step, the signal was deblinked by removing samples identified as belonging to a blink, as described above. Next, unreasonably small (< 1mm) and large (> 9mm) values were removed from the signal. In the third step, missing data were linearly interpolated with a maximum gap of 25 samples (125 ms). Finally, in the fourth step a Savitzky-Golay filter with a second order polynomial and 25 sample window length (125 ms) was applied to smooth the signal. From the pre-processed pupil signal, the average pupil size and the standard deviation of the pupil size were calculated for the six experimental segments (baseline and up to five roller coaster simulation rides), separately. Furthermore, using the approach described by Park et al. (Park et al., 2021), a pupil coherence metric was calculated for each segment.

Statistical analyses

Statistical analyses were done in R 4.4.1 (R Foundation for Statistical Computing, Vienna, Austria). For correlational analyzes we used the psych package (Revelle, 2020). Linear-mixed models were ran using the *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) packages. Statistical plots (Figs. 1 and 2) were produced using the *ggstatsplot* package (Patil, 2021).

Results

To verify that the roller coaster simulator rides induced SS in subjects, we compared pre-exposure to post-exposure scores on the SSQ. A paired t-tests indicated a significant pre-post difference ($t(29) = 6.9$, $p < .001$) with higher total SSQ scores post-exposure ($M = 65.2$, $SD = 42.53$) than pre-exposure ($M = 8.98$, $SD = 10.83$). The same statistically significant pattern was found for all three subscales of the SSQ (all p:s < .001). Next, to validate that the VAS-scale captures SS in a similar manner as the SSQ, we correlated participants VAS-ratings on the last coaster simulator ride before termination with their post-exposure SSQ-scores. As illustrated in Fig. 1A, there was a positive, significant correlation between VAS-ratings and total SSQ scores post-exposure ($r(27) = .79$, $p < .001$). Furthermore, the same pattern of a positive, statistically significant correlation was found between VAS and the three subscales of the SSQ (Figs. 1B-D, r:s > .67, all p:s < .001). These results indicate that the VAS-ratings do indeed capture SS in a similar manner as the SSQ.

We next evaluated the effect of the roller coaster simulator rides on VAS-ratings. Since not all subjects completed all five roller coaster simulator rides, we submitted our data to a between-subjects ANOVA with VAS-ratings as dependent variable and segment (Baseline/Ride 1–5) as independent variable. This analysis, illustrated in Fig. 2, found a main effect of segment ($F_{Welch}(5, 49.33) = 29.98$, $p < .001$). Follow-up post hoc analysis showed that all differences between each ride segment and the baseline segment was statistically significant (all p:s < .05), while none of the differences between the ride segments were. Accordingly, in subsequent analyses we collapsed the data from the five roller coaster simulator ride segments into a single, ride category.

To evaluate if the eye-tracking or HRV metrics could predict VAS ratings we ran separate linear mixed-effects models for each metric. Each model had VAS ratings as outcome variable, segment category (Baseline / Ride) as a fixed-effect and random intercepts for subjects, and finally, the metric of interest (e.g., mean heart rate) as a fixed-effect variable. In a final model, metrics from single metric models that contributed significantly to VAS ratings were combined. We used this approach in an effort not to saturate a complex initial model with metrics that were not related to VAS, and to avoid the risk of entering highly collinear metrics into the same model at an early stage. Note that this approach is similar to only entering predictors with a significant zero-order correlation into a linear regression model, but with the added benefit of taking the repeated measures approach of multiple segments into account. The results from these analyses are summarized in Table 2. As is evident from the table, only two metrics were significantly related, to VAS-ratings in the single metric models. Out of the HRV metrics, the

LF/HF ratio was a significant predictor ($\chi^2(1) = 4.98$, $p = .026$, $b = 0.26$) while pupil coherence was the only significant eye-tracking metric ($\chi^2(1) = 4.10$, $p = .043$, $b = 4.31$). Combining these two metrics in a final model revealed that both parameters remained significant.

We note that an LF/HF ratio could not be calculated for all segments for all participants. To verify that there is no systematic difference, with respect to VAS ratings, between segments where a LF/HF ratio could be computed and those where the ratio could not be computed, we ran a linear-mixed model with VAS ratings as outcome, segment category (Baseline / Ride) and missing LF/HF ratio (Yes / No) as a fixed-effects and random intercepts for subjects. This model revealed no effect of missing LF/HF ratio ($\chi^2(1) = 0.32$, $p = .57$).

Table 2

Single metric and combined linear mixed-effects models of HRV and Eye-tracking metrics predicting VAS. Statistical results are from likelihood ratio tests comparing each model with and without the metric of interest. The table reports unstandardized coefficients (*b*), chi-square values (χ^2), degrees of freedom (*df*), and p-values. Metrics that were significant in single-metric models ($p < .05$) were included in a final combined model, in which both remained significant predictors of VAS.

	Metric	***b***	χ^2	***df***	***p***
Single metric models					
HRV metrics					
	Mean HR	-0.0001	0.0005	1	.99
	Mean RR	0.00024	0.032	1	.86
	SDNN	-0.00005	0.0001	1	.99
	RMSSD	0.0014	0.13	1	.72
	pNN50	-0.003	0.027	1	.88
	LF/HF	0.26	4.98	1	.026*
Eye-tracking metrics					
	Blinks duration	0.86	0.25	1	.62
	Blinks rate	0.018	0.394	1	.53
	Mean pupil size	0.38	0.19	1	.66
	Std. pupil size	-6.49	2.43	1	.12
	Pupil coherence	4.31	4.10	1	.043*
Combined model					
	LF/HF	0.26	5.69^{+}	1	.017*
	Pupil coherence	5.53	5.58^{+}	1	.018*
* $p < .05$					

$^{+}$ Compared to a reduced model with the other metric included.

Discussion

The present study aimed to evaluate HRV and eye-tracking metrics as physiological markers of SS, using a VAS to concurrently assess symptom intensity. Our results demonstrated that the VAS is a valid and sensitive tool for monitoring SS onset and progression, with VAS ratings correlating strongly with the

post-exposure Questionnaire (SSQ) scores, including its subscales for nausea, disorientation, and oculomotor discomfort. These findings are in line with previous work suggesting that VAS ratings provide a more temporally precise, intuitive alternative to the SSQ for capturing dynamic changes in SS symptoms (Garrido et al., 2022; Gavgani et al., 2017; Nalivaiko et al., 2015).

Consistent with our hypotheses, our physiological data for both HRV and eye-tracking were associated with self-reported SS levels. Specifically, for HRV metrics, the LF/HF ratio, reflecting sympatho-vagal balance, was positively correlated with VAS ratings during the roller coaster simulator rides. This aligns with previous findings that associated increased LF/HF ratios with elevated sympathetic arousal or parasympathetic withdrawal in response to SS-inducing stimuli (Malińska et al., 2015; Stauffert et al., 2020). While other time-domain HRV measures such as RMSSD and SDNN were not significantly related to VAS ratings, the selective sensitivity of LF/HF may suggest that acute changes in autonomic modulation, rather than general variability, are more relevant to identify concurrent SS occurrence.

As for eye-tracking metrics, pupil coherence emerged as the strongest predictor of VAS ratings. While still a relatively new metric in SS research, this novel finding suggests that coherence in pupil dynamics, possibly reflecting autonomic regulation or cognitive-emotional integration (Kourtesis et al., 2024; Park et al., 2021), may provide a sensitive marker of simulator-induced discomfort. Interestingly, traditional metrics such as blink rate or average pupil diameter did not show significant relationships with SS, though consistent with some prior work indicating that pupil variability may better reflect complex interactions between sensory conflict and autonomic output (Al-Ashwal et al., 2024; Lopes et al., 2020). However, albeit interesting this finding should be interpreted with caution, as pupil coherence is a derived signal-level metric and may be influenced by general pupil size dynamics or measurement noise. Furthermore, similar to Park et al. (Park et al., 2021), we found a reduction in pupil coherence after exposure to a SS inducing VR experience. However, they also documented a negative correlation between SSQ-scores and pupil coherence, in contrast to the positive relation we found in this study. Notably, Park et al. (Park et al., 2021) exposed participants to a combination of 2D and a 3D settings. In our study, we used a continuous, fully immersive roller coaster designed to provoke SS, which may explain the discrepancy between results. Thus, our findings may not necessarily contradict theirs but rather reflect task-specific differences in how pupil coherence relates to perceived discomfort. Taken these findings together the role of pupil coherence as a predictor of self-reported sickness-levels deserves further investigation.

Together, our results support a multi-system response to SS involving autonomic imbalance and oculomotor strain, in line with both the sensory conflict and oculomotor theories of SS (Palmisano et al., 2020; Reason & Brand, 1975). In the combined model, including both LF/HF ratio and pupil coherence, our results showed that both metrics remained as significant predictors of VAS ratings, even when accounting for the other metric, reinforcing the value of integrating complementary physiological signals to capture different aspects of cybersickness.

This study also demonstrated the practical advantage of using VAS over SSQ in high-temporal-resolution SS research. The VAS was sensitive enough to detect discomfort changes ride-by-ride, enabling precise alignment with physiological fluctuations. In contrast, the SSQ, while validated and useful for pre/post comparisons, it may miss transient symptom variations or onset dynamics in short-duration exposures.

A few limitations should be acknowledged. First, not all subjects completed all five roller coaster simulation rides, and in some segments the LF/HF ratio could not be calculated due to extremely low HF power. However, statistical analysis confirmed no systematic differences in VAS ratings between included and excluded segments. Second, while eye-tracking and HRV were effective predictors, additional physiological measures such as electrodermal activity or EEG may further improve classification accuracy and provide deeper insight into SS mechanisms (Baldini et al., 2024; Yang et al., 2023).

Conclusion

This study provides compelling evidence that both HRV (specifically the LF/HF ratio) and eye-tracking (pupil coherence) can serve as physiological markers of SS, when aligned with concurrent VAS symptom ratings. These findings support the use of the VAS as a dynamic and sensitive alternative to traditional questionnaires like the SSQ for capturing SS onset and severity. By integrating autonomic and oculomotor data with subjective reports, this approach enables a richer understanding of the physiological underpinnings of SS and may inform the design of adaptive systems and interventions to reduce SS in VR and immersive environments.

Declarations

Funding

This work was supported by grants from the Swedish Defence Research Agency, FOI-2025-1008

Conflicts of interest/Competing interests

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Author Contribution

G.E. conceived of the study. G.E. and M.L. designed and performed the experiments assisted by R.D., C.K.K.L and T.K. R.D., C.K.K.L., T.K, and J.L. worked out the Virtual Reality technical implementation. G.E. and M.L. performed data analyses, R.D., and C.K.K.L. aided in subjective scales analysis. G.E. and M.L. wrote the manuscript. All authors commented and contributed to the final manuscript.

References

1. Al-Ashwal, W., Asadi, H., Alsanwy, S., Chalak Qazani, M., Mohamed, S., & Nahavandi, S. (2024). Objective Assessment of Cybersickness in a Stationary Helicopter Simulator Using Pupil Diameter as a Quantitative Indicator. *2024 IEEE International Systems Conference (SysCon)*, 1–7. https://doi.org/10.1109/SysCon61195.2024.10553536
2. Babic, T., & Browning, K. N. (2014). The role of vagal neurocircuits in the regulation of nausea and vomiting. *European Journal of Pharmacology*, *722*, 38–47. https://doi.org/10.1016/j.ejphar.2013.08.047
3. Balaban, C. D. (1999). Vestibular autonomic regulation (including motion sickness and the mechanism of vomiting): *Current Opinion in Neurology*, *12*(1), 29–33. https://doi.org/10.1097/00019052-199902000-00005
4. Baldini, A., Patron, E., Gentili, C., Scilingo, E. P., & Greco, A. (2024). Novel VR-Based Biofeedback Systems: A Comparison Between Heart Rate Variability- and Electrodermal Activity-Driven Approaches. *IEEE Transactions on Affective Computing*, *15*(3), 1580–1589. IEEE Transactions on Affective Computing. https://doi.org/10.1109/TAFFC.2024.3352424
5. Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01
6. Bimberg, P., Weissker, T., & Kulik, A. (2020). On the Usage of the Simulator Sickness Questionnaire for Virtual Reality Research. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 464–467. https://doi.org/10.1109/VRW50115.2020.00098
7. Chen, H., Dey, A., Billinghurst, M., & Lindeman, R. W. (2017). *Exploring pupil dilation in emotional virtual reality environments*. https://ir.canterbury.ac.nz/bitstream/10092/15404/2/icat-egve2017-HaoChen.pdf
8. Davis, S., Nesbitt, K., & Nalivaiko, E. (2015). *Comparing the onset of cybersickness using the Oculus Rift and two virtual roller coasters*. *167*.
9. Dymczyk, M., Przekoracka-Krawczyk, A., Kapturek, Z., & Pyżalska, P. (2024). Effect of a vergence-accommodation conflict induced during a 30-minute Virtual Reality game on vergence-accommodation parameters and related symptoms. *Journal of Optometry*, *17*(4), 100524. https://doi.org/10.1016/j.optom.2024.100524
10. Garrido, L. E., Frías-Hiciano, M., Moreno-Jiménez, M., Cruz, G. N., García-Batista, Z. E., Guerra-Peña, K., & Medrano, L. A. (2022). Focusing on cybersickness: Pervasiveness, latent trajectories, susceptibility, and effects on the virtual reality experience. *Virtual Reality*, *26*(4), 1347–1371. https://doi.org/10.1007/s10055-022-00636-4
11. Gavgani, A. M., Nesbitt, K. V., Blackmore, K. L., & Nalivaiko, E. (2017). Profiling subjective symptoms and autonomic changes associated with cybersickness. *Autonomic Neuroscience*, *203*, 41–50. https://doi.org/10.1016/j.autneu.2016.12.004

12. Gift, A. G. (1989). Visual Analogue Scales: Measurement of Subjective Phenomena. *Nursing Research*, *38*(5), 286.
13. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). *European Heart Journal*, *17*(3), 354–381.
14. Ishaque, S., Khan, N., & Krishnan, S. (2023). Physiological Signal Analysis and Stress Classification from VR Simulations Using Decision Tree Methods. *Bioengineering*, *10*(7), Article 7. https://doi.org/10.3390/bioengineering10070766
15. Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, *3*(3), 203–220. https://doi.org/10.1207/s15327108ijap0303_3
16. Kennedy, RS, Hettinger, LJ, & Lilienthal, M. G. (1988). Simulator sickness. *Motion and Space Sickness*, 317–341.
17. Kourtesis, P., Papadopoulou, A., & Roussos, P. (2024). Cybersickness in Virtual Reality: The Role of Individual Differences, Its Effects on Cognitive Functions and Motor Skills, and Intensity Differences during and after Immersion. *Virtual Worlds*, *3*(1), Article 1. https://doi.org/10.3390/virtualworlds3010004
18. Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13
19. Lackner, J. R., & DiZio, P. (2006). Space motion sickness. *Experimental Brain Research*, *175*(3), 377–399. https://doi.org/10.1007/s00221-006-0697-y
20. Lange, O. (2024). *NoLimits 2—Roller Coaster Simulation—Home*. https://nolimitscoaster.com/
21. Lopes, P., Tian, N., & Boulic, R. (2020). Eye Thought You Were Sick! Exploring Eye Behaviors for Cybersickness Detection in VR. *Proceedings of the 13th ACM SIGGRAPH Conference on Motion, Interaction and Games*, 1–10. https://doi.org/10.1145/3424636.3426906
22. Malińska, M., Zużewicz, K., Bugajska, J., & Grabowski, A. (2015). Heart rate variability (HRV) during virtual reality immersion. *International Journal of Occupational Safety and Ergonomics*, *21*(1), 47–54. https://doi.org/10.1080/10803548.2015.1017964
23. Nalivaiko, E., Davis, S. L., Blackmore, K. L., Vakulin, A., & Nesbitt, K. V. (2015). Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time. *Physiology & Behavior*, *151*, 583–590. https://doi.org/10.1016/j.physbeh.2015.08.043
24. Napadow, V., Sheehan, J. D., Kim, J., LaCount, L. T., Park, K., Kaptchuk, T. J., Rosen, B. R., & Kuo, B. (2013). The Brain Circuitry Underlying the Temporal Evolution of Nausea in Humans. *Cerebral Cortex*, *23*(4), 806–813. https://doi.org/10.1093/cercor/bhs073
25. Nyström, M., Andersson, R., Niehorster, D. C., Hessels, R. S., & Hooge, I. T. C. (2024). What is a blink? Classifying and characterizing blinks in eye openness signals. *Behavior Research Methods*, *56*(4), 3280–3299. https://doi.org/10.3758/s13428-023-02333-9

26. Palmisano, S., Allison, R. S., & Kim, J. (2020). Cybersickness in Head-Mounted Displays Is Caused by Differences in the User's Virtual and Physical Head Pose. *Frontiers in Virtual Reality*, *1*. https://www.frontiersin.org/articles/10.3389/frvir.2020.587698
27. Park, S., Mun, S., Ha, J., & Kim, L. (2021). Non-Contact Measurement of Motion Sickness Using Pupillary Rhythms from an Infrared Camera. *Sensors*, *21*(14), Article 14. https://doi.org/10.3390/s21144642
28. Patil, I. (2021). Visualizations with statistical details: The "ggstatsplot" approach. *Journal of Open Source Software*, *6*(61), 3167. https://doi.org/10.21105/joss.03167
29. Reason, J. T. (1978). Motion Sickness Adaptation: A Neural Mismatch Model. *Journal of the Royal Society of Medicine*, *71*(11), 819–829. https://doi.org/10.1177/014107687807101109
30. Reason, J. T., & Brand, J. J. (1975). *Motion sickness* (pp. vii, 310). Academic Press.
31. Rebenitsch, L., & Owen, C. (2016). Review on cybersickness in applications and visual displays. *Virtual Reality*, *20*(2), 101–125. https://doi.org/10.1007/s10055-016-0285-9
32. Revelle, W. (2020). psych: Procedures for psychological, psychometric, and personality research. *R Package Version*, *2*(5).
33. Stauffert, J.-P., Niebling, F., & Latoschik, M. E. (2020). Latency and Cybersickness: Impact, Causes, and Measures. A Review. *Frontiers in Virtual Reality*, *1*. https://doi.org/10.3389/frvir.2020.582204
34. Stephan, T., Deutschländer, A., Nolte, A., Schneider, E., Wiesmann, M., Brandt, T., & Dieterich, M. (2005). Functional MRI of galvanic vestibular stimulation with alternating currents at different frequencies. *NeuroImage*, *26*(3), 721–732. https://doi.org/10.1016/j.neuroimage.2005.02.049
35. Tian, N., Lopes, P., & Boulic, R. (2022). A review of cybersickness in head-mounted displays: Raising attention to individual susceptibility. *Virtual Reality*, *26*(4), 1409–1441. https://doi.org/10.1007/s10055-022-00638-2
36. Treisman, M. (1977). Motion Sickness: An Evolutionary Hypothesis. *Science*, *197*(4302), 493–495. https://doi.org/10.1126/science.301659
37. Voinescu, A., Petrini, K., & Stanton Fraser, D. (2023). Presence and simulator sickness predict the usability of a virtual reality attention task. *Virtual Reality*, *27*(3), 1967–1983. https://doi.org/10.1007/s10055-023-00782-3
38. Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and Cybersickness in Virtual Reality Are Negatively Related: A Review. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00158
39. Wewers, M. E., & Lowe, N. K. (1990). A critical review of visual analogue scales in the measurement of clinical phenomena. *Research in Nursing & Health*, *13*(4), 227–236. https://doi.org/10.1002/nur.4770130405
40. Yadav, R. L., Yadav, P. K., Yadav, L. K., Agrawal, K., Sah, S. K., & Islam, M. N. (2017). Association between obesity and heart rate variability indices: An intuition toward cardiac autonomic alteration – a risk of CVD. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy*, *10*, 57–64. https://doi.org/10.2147/DMSO.S123935

41. Yang, A. H. X., Kasabov, N. K., & Cakmak, Y. O. (2023). Prediction and detection of virtual reality induced cybersickness: A spiking neural network approach using spatiotemporal EEG brain data and heart rate variability. *Brain Informatics*, *10*(1), 15. https://doi.org/10.1186/s40708-023-00192-w

Figures

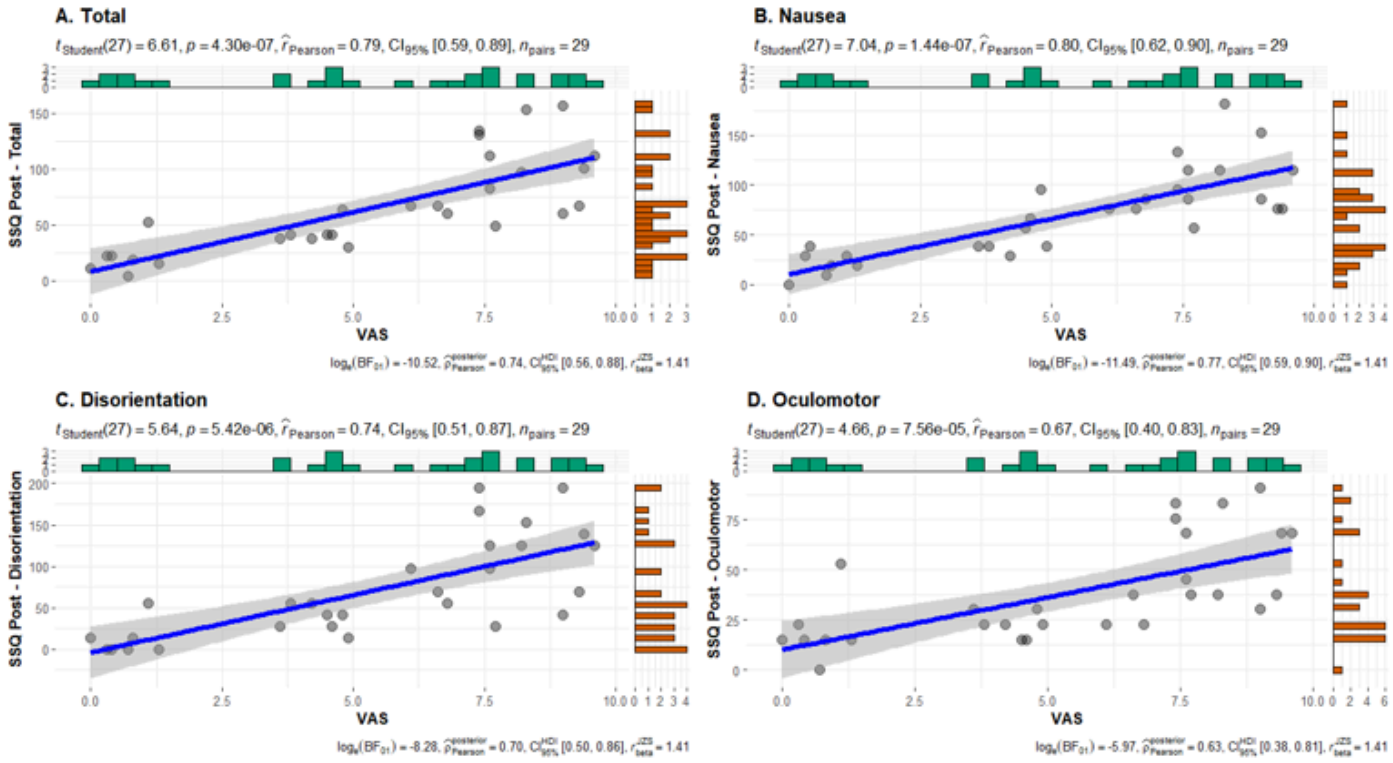


Figure 1

Scatterplots with linear regression lines (±95% CI) illustrating the association between subjects' VAS ratings from their final VR ride and post-exposure Simulator Sickness Questionnaire (SSQ) scores. Panels show correlations for:(A) Total SSQ score, (B) Nausea subscale, (C) Disorientation subscale, and (D) Oculomotor subscale. Each plot includes Pearson's r, confidence intervals, p-values, and log-transformed Bayes factors ($log_{10}[BF_{01}]$), indicating significant positive correlations across all SSQ components, demonstrating strong agreement between VAS ratings and established measures of simulator sickness. Marginal histograms display the distributions of VAS and SSQ scores.

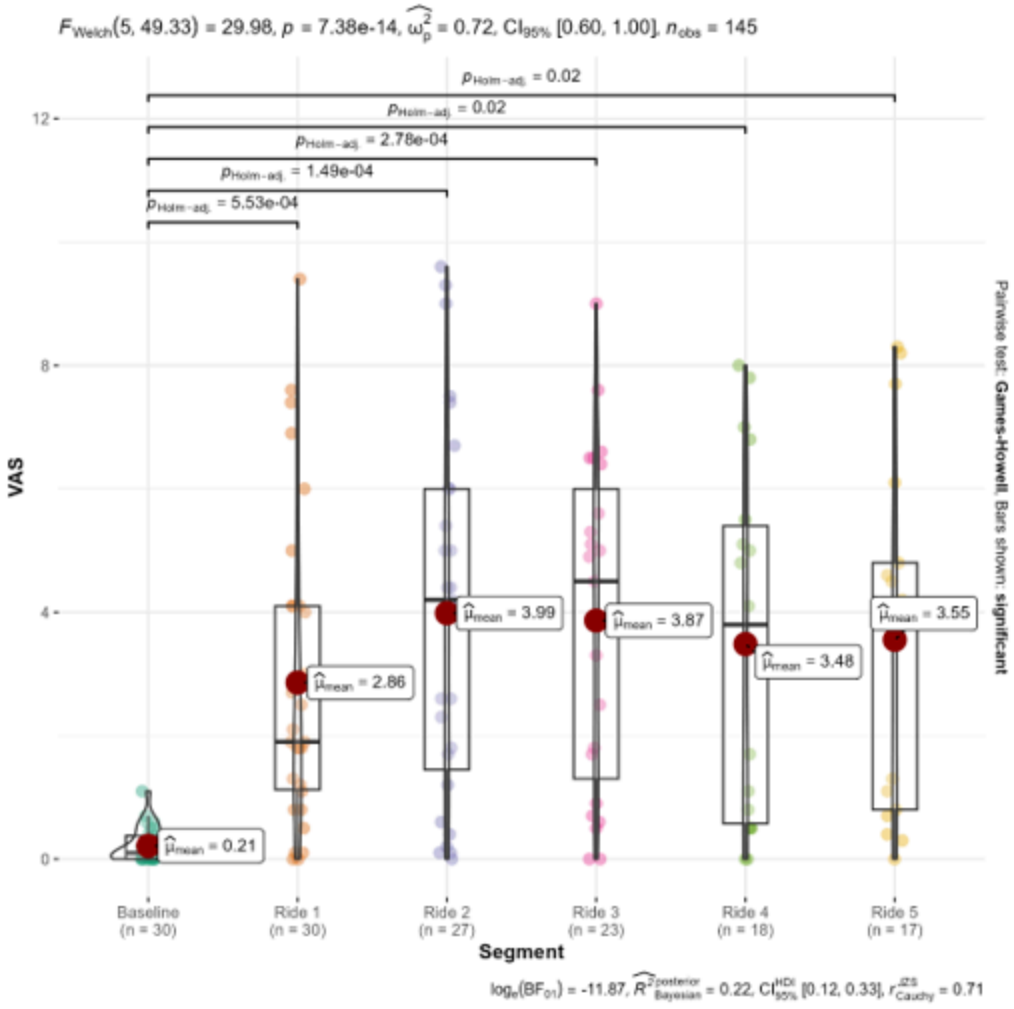


Figure 2

Subjects' VAS ratings as a function of segments. Boxplots show individual data points of participants' Visual Analogue Scale (VAS) ratings (0–10) across segments: baseline and five successive roller coaster rides. Individual observations are positionally adjusted along the x-axis to prevent overplotting. Boxes represent the interquartile range (IQR); horizontal lines within each box indicate the median; whiskers extend to 1.5 × IQR. Red points indicate group means. A Welch's ANOVA revealed a significant main effect of segment, FWelch(5, 49.33) = 29.98, p = 7.38e−14, ω² = 0.72, 95% CI [0.60, 1.00], indicating a large effect size. Games–Howell post hoc comparisons (Holm-adjusted) showed significantly higher VAS ratings in all ride segments compared to baseline (p < .05), while no significant differences were observed between the ride segments.