Contextual Influences on Simulator Sickness: A Comprehensive Analysis of Demographics, Gaming Experience and Simulation Context Complexity

Melina Bergen[1], Martin Fischer[1]

(1) German Aerospace Centre (DLR), Institute of Transportation Systems, Lilienthalplatz 7, 38108 Braunschweig, Germany, E-Mail: {melina.bergen, ma.fischer}@ dlr.de

Abstract - *Simulator sickness poses a persistent challenge in traffic research, affecting participant well-being and data validity. While individual traits such as age, gender, and gaming experience are known to influence susceptibility, contextual factors, such as simulator type and scenario complexity, are often overlooked or treated too simplistically. This study addresses this gap by introducing the Simulator Sickness Inducing Potential (SSIP) score, a metric that quantifies the contextual demands of a simulation. Data from eight studies involving pedestrian, cycling, and driving simulators (N = 268) were analyzed to examine how personal and contextual factors interact in predicting simulator sickness, measured via the Simulator Sickness Questionnaire (SSQ).*

The findings show consistently higher SSQ levels in female participants, regardless of scenario. In contrast, age and gaming experience interacted with simulation complexity: older participants and those with less gaming experience were particularly vulnerable in high-SSIP scenarios. These findings highlight the importance of considering both person-related and scenario-specific factors. The SSIP score proved valuable not only for retrospective analysis but also holds promise for prospective scenario design and risk assessment in simulation studies.

Keywords: *Simulator sickness, human in the loop simulation, scoring system, gaming experience, scenario complexity*

https://doi.org/10.82157/dsa/2025/9

Introduction

Traffic research is essential for understanding the behavior of road users and implementing safety measures to prevent accidents. Simulators provide a cost-efficient and safe testing environment, where variables can be precisely controlled without exposing participants to real-world risks (Brooks, et al., 2010; Kennedy, et al., 1993). This allows studying a wide range of traffic scenarios, from pedestrian behavior to vehicle interactions, under reproducible conditions. A central aspect of these studies is the generation of realistic participant behavior, to ensure the validity of findings.

However, simulator sickness poses a major challenge. Simulator sickness is a side effect, that can occur when being exposed to a Virtual Environment (VE) (Dużmańska, Strojny, and Strojny, 2018). It is characterized by symptoms such as nausea, dizziness and disorientation (Kennedy, et al., 1993). In addition to causing discomfort for the participants, simulator sickness leads to unnatural behavior, impacting data validity or even causing participants to drop out (Stanney, Mourant, and Kennedy, 1998). To address this issue, it is essential to understand the factors that influence simulator sickness, in order to be able to create scenarios and simulator settings that minimize the symptoms.

Previous research has identified several person-related factors influencing the susceptibility to simulator sickness. Demographic variables, such as gender and age have been shown to affect this susceptibility (Brooks, et al., 2010; Pöhlmann, et al., 2023; Stanney, Mourant, and Kennedy, 1998). For example, findings indicate that women are affected more often than men, potentially due to differences in visual-vestibular processing (Munafo, Diedrick, and Stoffregen, 2017). Similarly, age-related sensory declines might increase symptoms in older participants (Brooks, et al., 2010).

Furthermore, experience with video games has been discussed as a potential protective factor, as people who frequently play video games may be better adapted to VEs and virtual motion (Pöhlmann, et al., 2024). However, findings on the effect of gaming on simulator sickness susceptibility are inconsistent, with some studies reporting reduced sickness with increased gaming experience, and others finding no such relationship (Hunt and Potter, 2018; Kourtesis, et al., 2023; Munafo, Diedrick, and Stoffregen, 2017).

In addition to person-related factors, recent research highlights the importance of contextual influences on simulator sickness. Simulator sickness does not occur in a vacuum: The technological setup (e.g., monitor vs. Virtual Reality (VR) - Head-Mounted Display (HMD)), the movement dynamics (e.g., speed changes, turns), and the visual and interactive complexity of the scenario all shape how participants respond (Rebenitsch and Owen, 2016). Therefore, any evaluation of sickness symptoms must account for scenario-specific characteristics, which may modulate or even override the effects of individual differences. Treating these influences as binary distinctions (e.g., monitor vs. VR-HMD) oversimplifies a multidimensional problem (Rebenitsch and Owen, 2016; Stoner, Fisher, and Mollenhauer, 2011).

One explanation for these effects is provided by *Cue Conflict Theory*, which is the most widely accepted explanation of simulator sickness (Reason and Brand, 1975). According to this theory, sickness arises when the brain receives conflicting information from different sensory systems, for example, when visual cues suggest motion but the vestibular system does not register the corresponding physical movement. These mismatches disrupt the body's expectations based on real-world experience and may result in nausea, disorientation, or discomfort. This occurs, for example, in fixed-base driving simulators, which present motion visually while the user remains physically still. In contrast, pedestrian simulators, specifically those using natural locomotion like treadmills or motion capture spaces, minimize such conflict because proprioceptive and vestibular feedback align more closely with visual movement.

Next to the choice of display (VR-HMD, monitor etc.) and hardware characteristics (e.g., Field of View (FOV), image resolution, refresh rate, fixed-based vs. motion-based platforms), factors like scene and scenario design also influence the occurrence of simulator sickness (Stoner, Fisher, and Mollenhauer, 2011). For example, scene design decisions that increase optic flow, such as densely placed roadside objects or highly detailed textures, can enhance the perception of motion. While this adds realism, it also increases the likelihood of cue conflict. Similarly, scenario design elements such as sharp decelerations, tight curves, and especially 90-degree turns have been linked to higher levels of simulator sickness (Chrysler and Williams, 2005; Edwards, et al., 2003; Mourant, et al., 2007).

These insights informed the development of the Simulator Sickness Inducing Potential (SSIP) score, a metric that captures contextual complexity based on measurable scenario characteristics such as movement variability and visual density. This allows simulations to be classified based on their sickness-inducing potential, independent of participant traits.

This work investigates data from eight traffic simulation studies across pedestrian, bicycle, and driving simulators (both VR and monitor setups). The inclusion of multiple studies results in a large overall sample size ($N = 268$). By applying the SSIP score across all datasets, we aim to systematically assess how simulation complexity moderates the relationship between individual characteristics and simulator sickness. In particular, we test whether the effects of age, gender, and gaming experience differ depending on scenario demands.

Taken together, this work contributes to a more nuanced understanding of simulator sickness by emphasizing the interaction between personal and contextual factors. The use of the SSIP framework makes it possible to identify not just who is at risk, but under what circumstances.

Method

The data for this analysis was drawn from eight studies with a total sample size of $N = 268$ participants. The studies were conducted on multiple simulators, namely a driving simulator (with VR-HMD and with a triple-monitor setup), a bicycle simulator and a pedestrian simulator. The primary goal of these studies was not analysis of simulator sickness, but broader traffic-related research questions. Accordingly, the simulation scenarios varied considerably in terms of complexity, simulation duration and simulator types used, which made it difficult to directly compare the simulator sickness values collected.

Simulators

Driving Simulators

Two different versions of the driving simulator were used, which have identical setups, except for the visualization device. The driving simulators use a fixed-base setup, with a force-feedback steering wheel and pedals. The *CarSimDisplay* uses a triple-monitor setup, while the *CarSimVR* uses a VR-HMD. In all studies except two, the *HTC Vive Pro Eye* was used (the others two studies used the *Varjo XR-3* and the *Pimax8K-X)*. The display variant of the driving sim uses three 65” *Samsung* GQ-65Q90R TV's with a resolution of 3840 x 2160 p and a refresh rate of 60 Hz (Temme, et al., 2024).

Bicycle Simulator

The *BikeSim* consists of a modified bicycle mounted on a 2-Degree of Freedom (DOF) motion platform, which can simulate pitch and roll. The *HTC Vive Pro Eye* is used as a visual output device in all studies on the *BikeSim*.

Pedestrian Simulator

The *PedSim* uses an omnidirectional treadmill (Omnideck, (Guilfoyle and Thor, 2023)). The *Omnideck* consists of 16 tapered segments containing motorized spinning tubes, which transport the user back to the static platform at the center of the deck. A safety harness is worn by the participants, which is connected to the ceiling over the center of the deck, to prevent injuries in case of falling. A VR-HMD is used as visualization device (in this case the *HTC Vive Pro Eye*).

Measuring Simulator Sickness

The Simulator Sickness Questionnaire (SSQ, Kennedy, et al., 1993) was used in all studies, to

Figure 1: Simulators of the MoSAIC at the DLR in Brunswick, Germany

measure simulator sickness symptoms. The standardized SSQ is a well established tool to measure symptoms of simulator sickness. In practice, however, there is the shortcoming, that participants who drop out of the study due to simulator sickness, mostly do not fill out the questionnaire anymore. For that reason the average SSQ score is underestimated in studies with high drop-out rates. Since it is important when conducting empirical studies, to have an estimation of the drop-out rates, we introduce a fixed SSQ score for people that dropped out of the study. This approach is based on the assumption that discontinuation due to severe symptoms would typically correspond to a very high SSQ value, even if this could not be explicitly recorded. The resulting average SSQ score including the scores for drop-outs will be referred to as SSQ+ score in the following.

Simulator Sickness Inducing Potential

Variations in scenario design, exposure duration, and simulator type influence the occurrence of sickness symptoms, limiting direct comparability of SSQ scores across studies. To account for these differences, the SSIP metric was developed in order to provide an estimate of the simulation's inherent likelihood to provoke simulator sickness. The SSIP score incorporates simulator- and scenario-specific factors, such as scenario complexity and simulation duration, to support a more standardized comparison across different experimental contexts.

In this paper, the SSIP score is specifically applied to compare the studies included in our dataset. For this purpose, we focused exclusively on scenario-related aspects, as well as a simulator-specific factor. Certainly, hardware characteristics such as FOV or refresh rate of the HMD affect the occurrence of simulator sickness as well. However, we chose not to include those factors in the SSIP calculation, since nearly all studies in our sample used the same HMD model. Similarly, we excluded performance-related variables (e.g., frame rate stability), since they were consistent across the included studies. However, if the SSIP metric were to be applied in broader contexts with greater technical variability, these factors should be incorporated into the model.

In order to derive the formula for the calculation of the SSIP score a multi-step process was employed. First, a set of scenario characteristics was determined that influence the occurrence of simulator sickness. This was based on expert ratings and literature on the subject. Next, a simulator-specific factor was derived. A subset (75 %) of the available studies was then used to empirically optimize the weights of each feature, aiming to maximize the correlation between the computed SSIP score and the observed SSQ+ values. This optimization was performed using grid search techniques to identify the most predictive weight configuration. The remaining studies, not included in the optimization step, served as a validation set to evaluate the generalizability and robustness of the resulting SSIP model.

In the following, each step of the calculation of the SSIP score is explained in detail:

1. Simulator factor:

To account for differences in simulator setups, a simulator-specific factor was included in the SSIP score. This factor reflects the varying likelihood of simulator sickness across modalities, which is primarily driven by the level of sensory cue conflict inherent in each simulator type (cf. Cue Conflict Theory; Reason and Brand, 1975). Simulators that allow for natural locomotion and provide corresponding force feedback, such as pedestrian simulators, tend to produce less sickness due to better alignment between visual and vestibular cues, while seated simulations with high motion complexity (e.g., bicycle simulators in VR) lack adequate force feedback and therefore tend to induce stronger cue conflicts. The simulator types were evaluated across three key dimensions that are known to be linked to simulator sickness:

- **Visual–vestibular mismatch:** the degree to which visual motion cues conflict with information from the vestibular system
- **Movement expectation conflict:** the extent to which the user actively initiates movement (e.g., via walking, steering, or pedaling) but does not receive consistent proprioceptive or vestibular feedback
- **Balance demands:** the degree of active postural control required to stay upright or stable

Each simulator type was assigned a score from 0–3 in each category. The three dimension scores were summed to obtain a total conflict score (ranging from 0 to 9). The total score was then linearly scaled to a range from 0.0 (lowest sickness potential) to 2 (highest sickness potential), yielding the following simulator factors: *Real* = 0.0, *PedSim* = 0.7, *CarSimDisplay* = 0.8, *CarSimVR* = 1.3, *BikeSim* = 2.0. A value of 0 corresponds to real-world exposure (no risk of sickness), while 2 reflects a maximum risk scenario, such as bicycle simulators with high visual–vestibular mismatch, balance demands, and movement incongruence.

2. Weighted Z-transformation of the scenario characteristics:

Furthermore, scene and scenario design also influence the potential for cue conflict. High visual complexity and sharp directional changes (e.g., 90-degree turns, abrupt decelerations) may increase the mismatch with vestibular cues, thereby enhancing simulator sickness risk (Chrysler and Williams, 2005; Stoner, Fisher, and Mollenhauer, 2011). These effects are especially pronounced in simulator types with higher baseline cue conflict, as outlined in *Cue Conflict Theory* (Reason and Brand, 1975). Therefore, all scenario features were scaled using the simulator factor described above.

Eight objective scenario features were multiplied by the simulator factor, then z-transformed and assigned predetermined weights. The weightings were based on established findings on their relevance to simulator sickness and expert ratings.

The weighted characteristics were:

- Total Turns (1.0–2.0, dynamically weighted)
- Avg. Speed Variability (1.0–2.0, dynamically weighted)
- Avg. Number Non-Player Character (NPC) Interactions per Scenario (1.0–2.0, dynamically weighted)
- Total Different Tasks / Interactive Targets (1.0–2.0, dynamically weighted)
- Adjusted Path Length (1.6)
- Avg. Number of Intersections per Scenario (2.0)
- Number of Scenario Executions (1.0)
- Number of Different Scenarios (1.5)

Whether a feature was assigned a fixed or dynamic weight depended on the assumption that its contribution to simulator sickness may vary depending on the feature's own intensity or the simulator context. Features that were expected to interact with simulator configuration (such as turns, speed variability, or frequent task changes) were dynamically weighted. In contrast, features with a more linear or consistent effect, such as the number of different scenarios or the adjusted path length, were assigned fixed weights. This distinction helps reflect the context-sensitivity of some scenario elements.

The dynamic weights were assigned using a percentile-based approach within the dataset, ensuring that relatively extreme values contributed more strongly to the final score. Features that were expected to reflect higher visual-vestibular conflict (e.g., directional turns, speed fluctuations), increased task complexity (e.g., number of tasks or NPC interactions), or cognitive load (e.g., decision points at intersections) received higher weights if their values fell into the upper quartile of the observed distribution. Each of these features could receive a dynamic weight ranging from 1.0 to 2.0, where 1 represents a moderate (baseline) influence of a given feature, and 2 represents a strong influence. This range was chosen to ensure that all features have a minimum influence, avoiding "zeroing out" effects, while still allowing meaningful differentiation between features.

Fixed weights were assigned to scenario features with more stable and context-independent effects. These weights were derived based on theoretical considerations and expert judgment regarding the features' contributions to simulator sickness. Specifically, each feature was evaluated across six dimensions known to influence simulator sickness risk: social interactivity, motor demands, attentional load, cognitive complexity, vestibular/proprioceptive conflict, and visual load. Features with high relevance across multiple dimensions received higher weights, while those with limited or indirect influence were weighted lower.

3. Linear scaling of the SSIP to SSQ+ level:

The sum of the weighted and z-transformed scenario characteristics, adjusted by the dynamic simulator factor, resulted in the scaled SSIP score. The final SSIP value for a scenario was calculated as follows (see Eq. 1):

$$\mathit{SSIP}_i = \alpha \cdot \sum_k w_k \cdot \frac{(F_k \cdot \mathit{SimFac}_i - \mu_{F_k})}{\sigma_{F_k}} \tag{1}$$

where F_k denotes the k-th scenario-specific value of a given feature F, while μ_{F_k} and σ_{F_k} represent the mean and standard deviation of that feature across all studies i, w_k represents the weighting of a given feature, and SimFac_i the dynamic simulator factor for study i. The scalar α linearly maps the computed values onto the empirical SSQ+ (including dropout adjustment) scale for comparability.

To validate the SSIP approach, average SSIP values per study were correlated with mean SSQ+ values. The result was a high *Spearman's correlation* ($\rho = 0.90$, $p < .001$), confirming that the SSIP effectively reflects scenario-induced sickness potential at the study level.

Individual Factors

To explore the interplay between individual and contextual influences on simulator sickness, we analyzed whether scenario complexity (as captured by the SSIP score) moderates the effects of age, gender and gaming experience on sickness symptoms. Gaming experience was assessed using a 6-point *Likert scale*, ranging from "never" to "very often." Age was treated as a continuous variable.

Results

Individual-Level Correlation between SSIP on SSQ and Subscale Analysis

After validating the SSIP score on the study level (see Method), we next examined how well the SSIP predicts actual symptom severity on the individual level. A regression analysis across all participants confirmed a strong relationship between SSIP and SSQ scores ($\beta = 0.3365, p < .001$). The SSQ score increased with SSIP, but the relationship was non-linear: Scores remained stable at low to moderate SSIP levels, then rose disproportionately once SSIP exceeded approximately 15, as shown in Figure 2.

Further differentiation by SSQ subscales revealed that:

- **Disorientation** was the most sensitive to SSIP ($R^2 = 0.0788, \beta = 0.3973, p < 0.001$), highlighting its responsiveness to immersive or spatially incoherent environments.

- **Nausea** showed the highest explained variance ($R^2 = 0.1307$, $\beta = 0.3427, p < 0.001$), suggesting it reflects broader aspects of discomfort in complex scenarios.
- **Oculomotor** symptoms exhibited the weakest association with SSIP ($R^2 = 0.0804, \beta = 0.2037, p < 0.001$), indicating lower contextual sensitivity.

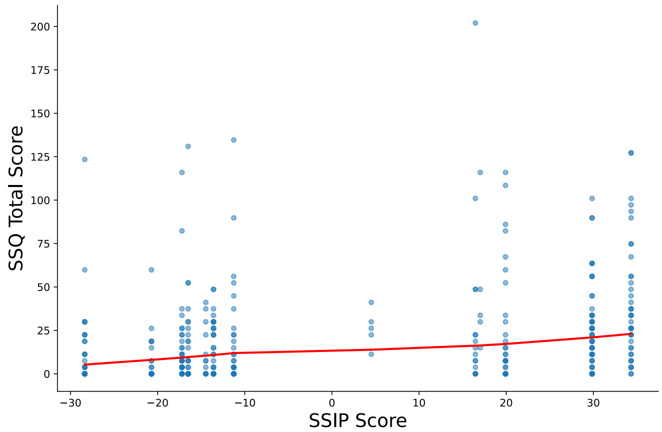

Figure 2: Relationship between SSIP Score and SSQ score. The red line represents a locally weighted regression (LOWESS)

Gender Differences in Simulator Sickness

The analysis of gender differences across simulators revealed significantly higher average SSQ scores for female than for male participants. The median SSQ score for women was 18.70, while for men it was 7.48. A *Mann-Whitney U-test* confirmed that this difference was statistically significant ($U = 18626.5, p < 0.001$). These findings are in line with previous literature suggesting that there may be a gender-specific vulnerability to simulator sickness, caused by physiological and vestibular sensitivity differences (Pöhlmann, et al., 2023). Notably, this gender effect remained stable across simulation contexts and SSIP levels.

Age and Simulation Complexity

The relationship between age, SSIP and SSQ scores was analyzed using a linear regression model with an interaction term between age and SSIP. While neither age ($\beta = 0.138, p = 0.293$) nor SSIP alone ($\beta = -0.024, p = 0.905$) significantly predicted the SSQ scores, the interaction term ($\beta = 0.0129, p = 0.037$) was significant, which means that SSQ scores increase more steeply with age under high-complexity conditions. This interaction is shown in Figure 3.

Simulator-specific regression plots (see Figure 4) further contextualize this finding. When looking at high-SSIP scenarios, for the *BikeSim*, *CarSimDisplay* and *CarSimVR* non-significant positive relationships were found between age and SSQ scores. However, for the *PedSim* a significant slightly negative correlation was found (*Spearman's* $\rho = -0.190, p = 0.035$). In contrast, in the low-SSIP scenarios, the relationship was

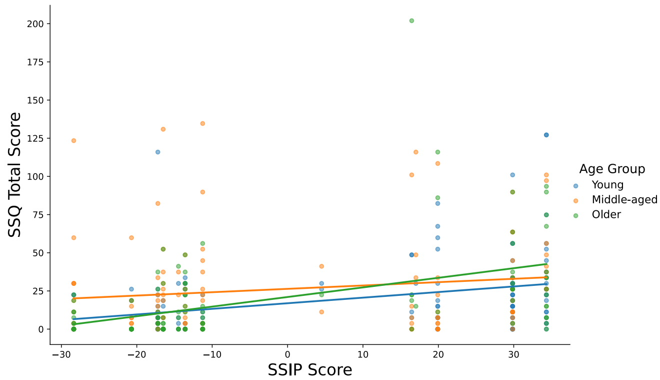

Figure 3: SSQ Scores by SSIP Score and Age Group

non-significant and close to zero or slightly negative. These non-parametric results, however, do not capture potential interaction effects. Additional *Spearman correlations* conducted across the full dataset revealed a significant negative correlation between age and SSQ scores in high-SSIP conditions ($\rho = -0.146,\ p = 0.017$), which appears contradictory to the regression results.

This discrepancy is probably due to methodological differences. *Spearman's* ρ looks at whether age and SSQ generally correlate within each group (e.g., high or low SSIP), but it does not take possible interactions with SSIP into account. The regression model, on the other hand, takes the interaction effect into account and shows that age becomes more relevant when SSIP is high. So even if the simple correlations don't show a clear pattern, the regression suggests that older participants are more affected by simulator sickness, especially in complex scenarios. This supports the idea that age-related sensitivity isn't consistent across all conditions but depends on the context, specifically, on how demanding the simulation is.

Additionally, a regression analysis of the SSQ subscales was conducted. A significant interaction between age and SSIP was found for the *Nausea* subscale ($\beta = 0.017, p = 0.003$). This indicates that older individuals are particularly susceptible to nausea in high-complexity simulations. No significant interactions were found for the *Disorientation* or *Oculomotor* subscales.

Gaming Experience and Simulation Complexity

Generally, gaming experience showed no significant main effect on SSQ scores across all studies. Again, a linear regression analysis was conducted. The findings showed a non-significant negative relationship between gaming experience and SSQ score ($\beta = -1.89,\ p = 0.031$), indicating that participants with more gaming experience showed less simulator sickness. Also, the interaction between gaming experience and SSIP was not significant ($\beta = -0.01, p = 0.867$), indicating that the protective effect does not systematically differ between low and high complexity contexts.

Spearman's correlations provided partial support for this pattern. In high-SSIP simulations, a significant

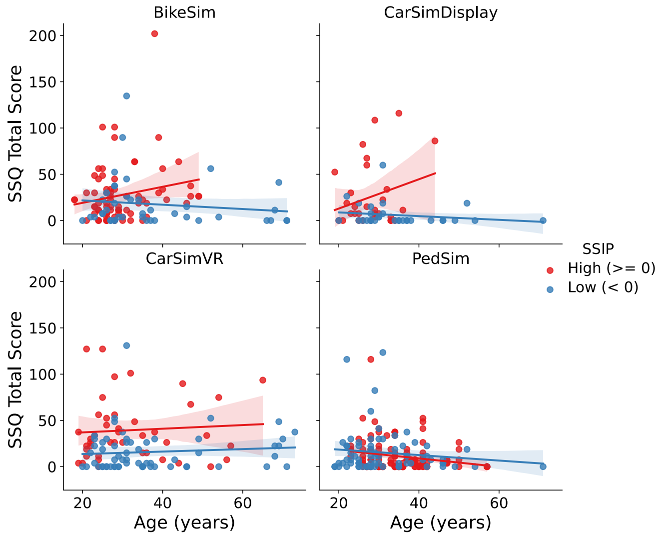

Figure 4: Age and SSQ Scores by Simulator Type and SSIP Scores

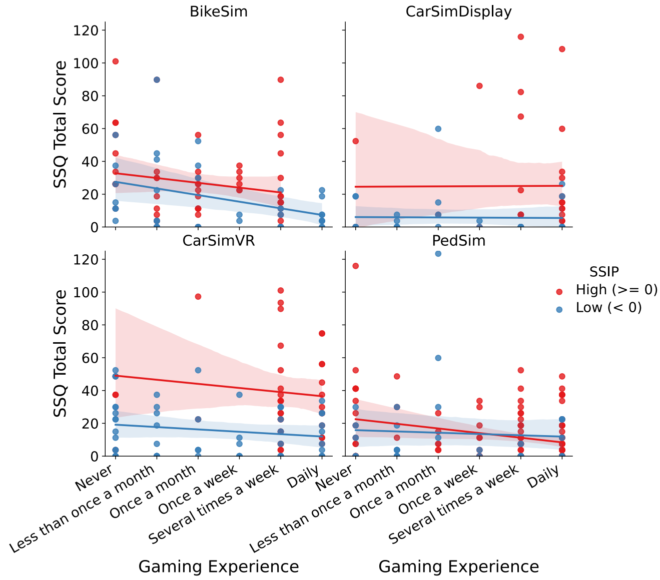

Figure 5: Gaming Experience and SSQ Scores by Simulator Type and SSIP Scores

negative correlation was found between gaming experience and SSQ scores ($\rho = -0.197$, $p = 0.001$), while the trend in low-SSIP scenarios was weaker and not significant ($\rho = -0.130$, $p = 0.092$). When examining individual simulators (see Figure 5), the protective effect of gaming was most evident in *BikeSim* under low-SSIP conditions ($\rho = -0.284$, $p = 0.069$), though this did not reach statistical significance. Other simulators showed weaker or inconsistent associations. In *CarSimDisplay*, no meaningful relationship was observed between gaming experience and SSQ scores under either SSIP condition.

These findings suggest that a high amount of gaming experience may help reduce simulator sickness symptoms, but the effect is weak and context-dependent and differs across simulators. The analysis of SSQ subscales did not reveal any significant associations with gaming experience. Moreover, no significant interaction effects with age were found.

Discussion

This research investigated how individual characteristics such as gender, age, and gaming experience influence simulator sickness, and how these effects vary depending on simulation complexity. In order to systematically rate contextual differences the SSIP score was developed and applied across all studies.

Simulator Sickness as a Function of Complexity

A key finding is the non-linear relationship between SSIP and total SSQ scores. While symptoms remained relatively stable at lower levels of complexity, SSQ scores rose disproportionately once SSIP values surpassed a threshold around 15. This pattern is consistent with theories of sensory conflict and cognitive overload, where additional load or immersion may exceed individual tolerance limits. Subscale analyses further revealed that *Disorientation* and *Nausea* are the most responsive to increases in SSIP. *Disorientation* showed the steepest increase, while *Nausea* accounted for the highest explained variance.

Gender, Age, and Experience

Gender effects were robust across simulations: Female participants consistently reported higher SSQ scores than males. These findings replicate prior research and may be related to sex-based differences in vestibular sensitivity or motion processing. Importantly, the effect remained stable across SSIP levels, indicating a general vulnerability rather than a context-dependent one.

The relationship between age and simulator sickness depended on the complexity of the simulation. Although age alone was not a significant predictor, the interaction between age and SSIP score was significant. This indicates that in more demanding simulations, older individuals were more susceptible to simulator sickness. This effect was particularly visible in the *Nausea* subscale, which might be explained by physiological changes with age (e.g., sensory integration, motion sensitivity).

Gaming experience showed a weak negative relationship with SSQ scores, particularly in high-SSIP conditions. While the interaction with SSIP was not statistically significant, *Spearman's correlations* in high-complexity scenarios suggest a mild protective effect. This aligns with the idea that regular exposure to interactive VEs may enhance motion tolerance and reduce discomfort, although the effect is small. No significant interaction with age was found.

Value of the SSIP Framework

The SSIP score proved to be a useful framework for disentangling otherwise weak or inconsistent effects. For instance, the age effect would have remained undetected without considering the interaction with SSIP. Similarly, the effect of gaming experience only became visible when separating by scenario complexity. This suggests that the SSIP can serve as a meaningful contextual variable in simulator research. The SSIP offers a more nuanced and precise way to describe the complexity of a simulation, rather than just relying on binary distinctions (e.g. monitor vs. VR - HMD). Beyond that, the SSIP also has the potential to be used as a predictive tool. In the future, it could be used to estimate expected simulator sickness levels prior to the execution of a study, based solely on scenario and simulator parameters. This would allow researchers to classify and adapt simulation setups in advance — for instance, by adjusting high-SSIP scenarios or by recruiting only certain participant groups. However, it is important to note that the current version of the SSIP is based on relative scaling within the context of this dataset. Accurate prediction for individual studies would require either a standardized reference distribution or a more generalized model. With further refinement and broader validation, the SSIP could be developed into a more universally applicable tool to support planning decisions and scenario design across diverse simulation contexts. In this way, the SSIP would not only be able to help explain differences in symptom severity, but could also be used for planning decisions and scenario design.

Limitations and Outlook

Several limitations must be considered. First, while the SSIP score is based on objectively measurable scenario parameters (e.g., number of turns, speed variability), however certain steps in its construction involved subjective decisions, for example, how to define a 'turn' based on angular thresholds, or how to deal with scenarios where no predefined path existed. Additionally, the selection of scenario features, although grounded in prior literature and expert input, inevitably involve a degree of interpretation. Therefore, while the SSIP is internally consistent and systematically constructed, it would still benefit from further validation. Second, the study combined data from different simulators and studies, which increases ecological validity but limits causal inference. Third, some subgroups (e.g., older adults with low gaming experience) may have been underrepresented, limiting the generalizability of interaction effects.

Future work should focus on validating and refining the SSIP score. While the current version already integrates objective scenario metrics grounded in expert judgment and empirical correlations, the process of determining feature weights could benefit from more systematic approaches. For instance, future studies might explore alternative weighting strategies, data-driven optimization techniques to improve generalizability. Although initial machine learning-based approaches did not yield meaningful improvements in this project, further work is needed to evaluate whether other methods (e.g., *Bayesian model* selection or larger training datasets) could enhance predictive performance.

Conclusion

This study examined how individual characteristics such as gender, age, and gaming experience interact with simulation complexity to influence simulator sickness. By introducing the SSIP, we provide a framework to quantify scenario complexity and better account for contextual influences. The findings reinforce the idea that simulator sickness results from the interaction between individual characteristics and the demands of a simulation.

While gender remained a stable risk factor across all simulations, the influence of age and gaming experience became more pronounced under higher-complexity conditions. These insights underscore the importance of considering simulation context when evaluating susceptibility to simulator sickness.

Although the current implementation is tailored to the datasets included in this study, it lays the foundation for more generalizable predictive models. With further refinement, the SSIP could support researchers in both the design and interpretation of simulation studies, helping to identify high-risk scenarios and inform participant selection or scenario adjustments in advance.

Taken together, this work contributes to a more nuanced understanding of simulator sickness by bridging person- and context-related factors, and offers a promising direction for future simulation design, evaluation, and planning.

1. Used Acronyms

DOF	Degree of Freedom
FOV	Field of View
HMD	Head-Mounted Display
NPC	Non-Player Character
SSQ	Simulator Sickness Questionnaire
SSIP	Simulator Sickness Inducing Potential
VE	Virtual Environment
VR	Virtual Reality

2. Author contributions

Contributions as suggested by the CRediT system (Allen, O'Connell, and Kiermer, 2019):

Melina Bergen: Conceptualization, Project administration, Investigation, Methodology, Software, Visualization, Writing - Original Draft.
Martin Fischer: Supervision, Conceptualization, Resources, Funding acquisition.

References

Allen, L., O'Connell, A., and Kiermer, V., 2019. How can we ensure visibility and diversity in research contributions? How the Contributor Role Taxonomy (CRediT) is helping the shift from authorship to contributorship. *Learned Publishing*, 32(1), pp. 71–74. https://doi.org/https://doi.org/10.1002/leap.1210.

Brooks, J. O., Goodenough, R. R., Crisler, M. C., Klein, N. D., Alley, R. L., Koon, B. L., and Wills, R. F., 2010. Simulator sickness during driving simulation studies. *Accident Analysis & Prevention*, 42(3), pp. 788–796.

Chrysler, S. T. and Williams, A. A., 2005. Driving performance in a simulator as a function of pavement and shoulder width, edge line presence, and oncoming traffic. In: *Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.* Rockport, ME, pp. 370–375.

Dużmańska, N., Strojny, P., and Strojny, A., 2018. Can simulator sickness be avoided? A review on temporal aspects of simulator sickness. *Frontiers in Psychology*, 9, p. 2132. https://doi.org/10.3389/fpsyg.2018.02132.

Edwards, C. J., Creaser, J. I., Caird, J. K., Lamsdale, A. M., and Chisholm, S. L., 2003. Older and younger driver performance at complex intersections: Implications for using perception-response time and driving simulation. In: *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*. Iowa City: University of Iowa, pp. 33–38.

Guilfoyle, R. and Thor, P., 2023. *Omnideck – Product Sheet*. Available at: ¡http://bit.ly/2VpakmL¿.

Hunt, X. and Potter, L. E., 2018. High computer gaming experience may cause higher virtual reality sickness. In: *Proceedings of the 30th australian conference on computer-human interaction*, pp. 598–601.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G., 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3), pp. 203–220.

Kourtesis, P., Amir, R., Linnell, J., Argelaguet, F., and MacPherson, S. E., 2023. Cybersickness, cognition, & motor skills: The effects of music, gender, and gaming experience. *IEEE Transactions on Visualization and Computer Graphics*, 29(5), pp. 2326–2336.

Mourant, R., Prasanna, R., Cox, D., Lin, Y., and Jaeger, B. K., 2007. The Effect of Driving Environments on Simulator Sickness. In: *Proceedings of the 51st Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors and Ergonomics Society.

Munafo, J., Diedrick, M., and Stoffregen, T. A., 2017. The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. *Experimental Brain Research*, 235, pp. 889–901.

Pöhlmann, K. M., Li, G., McGill, M., Pollick, F., and Brewster, S., Mar. 2023. Can Gender and Motion Sickness Susceptibility Predict Cybersickness in VR? In: *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, pp. 277–282.

Pöhlmann, K. M., Li, G., Wilson, G., McGill, M., Pollick, F., and Brewster, S., 2024. Is Video Gaming a Cure for Cybersickness? Gamers Experience Less Cybersickness Than Non-Gamers in a VR Self-Motion Task. *IEEE Transactions on Visualization and Computer Graphics*. Early Access.

Reason, J. T. and Brand, J. J., 1975. Motion sickness. In: *Human Factors*. London: Academic Press, pp. 135–171.

Rebenitsch, L. and Owen, C., 2016. Review on cybersickness in applications and visual displays. *Virtual Reality*, 20(2), pp. 101–125. https://doi.org/10.1007/s10055-016-0285-9.

Stanney, K. M., Mourant, R. R., and Kennedy, R. S., 1998. Human factors issues in virtual environments: A review of the literature. *Presence*, 7(4), pp. 327–351.

Stoner, H. A., Fisher, D. L., and Mollenhauer, M., 2011. Simulator and Scenario Factors Influencing Simulator Sickness. In: *Handbook of Driving Simulation for Engineering, Medicine, and Psychology*. CRC Press. Chap. 14.

Temme, G., Bergen, M., Nicolay, E., Rehm, M. J., and Fischer, M., 2024. Validation of Hardware for looking around in a VR based Driving Simulation (VALHaLa).