Multimedia Systems (2026) 32:645
https://doi.org/10.1007/s00530-026-02733-9

REGULAR PAPER

VR video user experience evaluation under visual and temporal variations and machine learning–based cybersickness prediction

Muhammad Shahid Anwar[1,2] · Amer Mousa Alkharoubi[1] · Mubashir Ahmad[2] · Majed Elwardy[3]

Received: 26 April 2026 / Accepted: 6 October 2026
© The Author(s) 2026

Abstract

Virtual reality (VR) allows users to experience an engaging and interactive virtual environment. Such experience is often limited by cybersickness, a perceptual discomfort that varies among individuals. Understanding and modeling user responses to visual and temporal characteristics of VR content is essential for designing adaptive and user-centered systems. Despite its importance, the influence of variations in content encoding and playback on subjective perception of cybersickness remains underexplored. This study presents a user-centered evaluation of cybersickness in VR videos manipulated across three dimensions: quantization levels, resolutions, and playback delays. The study then conducted a subjective experiment with 34 participants who experienced these conditions via a head-mounted display (HMD). The user's subjective responses were collected along with a simulator sickness questionnaire (SSQ). The subjective results indicate that cybersickness severity increases with higher compression levels, temporal discontinuities, and extreme resolutions, whereas moderate resolution (2.5K) and low compression (QP22) yield the most comfortable viewing experiences. Moreover, the study proposed a Machine Learning-based cybersickness levels prediction model. The subjective data were arranged and used to train multiple machine learning techniques, including ANN, SVM, Decision Tree, CatBoost, and XGBoost. A comparative evaluation among these techniques reveals that XGBoost consistently outperforms the other methods in terms of accuracy, F1-score, and AUROC, demonstrating strong discriminative capability and generalization performance. These findings highlight clear response patterns that can inform user modeling for adaptive VR systems, where content quality and playback conditions are dynamically adjusted to minimize discomfort. This work contributes to the broader goal of personalized, user-adaptive virtual environments that balance immersion and comfort through informed design.

Keywords Virtual reality (VR) videos · Multimedia analysis · User experience · Visual and temporal variations · Immersive media · Cybersickness prediction · Machine learning

Communicated by Bing-kun Bao.

✉ Majed Elwardy
majed.elwardy@bth.se

Muhammad Shahid Anwar
Muhammad.anwar.1@kfupm.edu.sa

Amer Mousa Alkharoubi
kharoubi@kfupm.edu.sa

Mubashir Ahmad
mubashir.ahmad@kfupm.edu.sa

[1] Integrated Design Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

[2] Center for Intelligent Secure Systems (IRC-ISS), King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

[3] Department of Computer Science, Blekinge Institute of Technology, Karlskrona 37179, Sweden

Published online: 16 October 2026

1 Introduction

Virtual reality (VR) provides an immersive medium through which users can experience and interact with computer-generated or captured environments, often from an embodied or viewer-centered perspective [1]. Advances in display technologies, motion tracking, and rendering performance have enabled VR applications across domains such as healthcare [2], entertainment [3], education [4], and cultural heritage [5]. However, cybersickness remains one of the major barriers that prevents users from using VR technology for extended periods, induced by immersive visual effects despite the progress in VR environment [6].

The fundamental concept of cybersickness are often explained through sensory conflict theory, which recommends that user discomfort induces when visual and vestibular systems provide mismatched motion cues. Especially in VR environment, this mismatch can result from several factors, including low frame rates, visual artifacts, poor image quality, or temporal disruptions. When the user's eyes in VR perceive motion that the body does not physically experience, the users' brain confuses and interprets this conflict as an error. This phenomena leads to discomfort or nausea called cybersickness. The degree of cybersickness varies among users [7], partly because individuals differ in their sensitivity to visual motion, discomfort, experience level, and exposure to head movement [8, 9]. This variability motivates user-centered approaches that consider both system-level factors and individual responses to immersive content.

The symptoms of cybersickness, including nausea, disorientation, and oculomotor discomfort, create a negative impact on user experience and decrease system usability [10]. Users experience cybersickness in the VR environment when they become fully immersed and perceive themselves moving through the digital space while their bodies remain still. Previous studies have used different methods to detect cybersickness, but the simulator sickness questionnaire (SSQ) provides the widely used assessment [11, 12]. Evaluating cybersickness in VR video is therefore important for understanding how content quality and playback conditions affect user comfort and for informing the design of immersive systems that minimize discomfort during use. User-adaptive VR systems need developers to understand how cybersickness develops and which factors help reduce its effects to create technologies that will maintain user comfort together with their active engagement. Similarly, prior VR experience also affect the users' subjective responses to immersive environments, influencing perceived usability, control, and overall experience during VR exposure [5]. The level of cybersickness that users experience shows different patterns between individuals, which demonstrates the importance of user-based modeling methods, which need to include both perceptual sensitivity and user interaction behavior.

Although advances in HMD hardware, rendering pipelines, and motion-prediction techniques have contributed to reduced latency and improved display stability [13], user comfort in VR remains affected by content-level factors such as compression, resolution, frame continuity, and playback interruptions. In addition, the visual fidelity of an image, its smooth movement, and the user's perception of motion stability all depend on the specific parameters that need to be addressed. Therefore, there is an opportunity to address how these factors work together to create different levels of cybersickness. The majority of previous research have concentrated on system optimization processes instead of assessing how users experience actual VR content. However, limited attention has been given to the combined effect of compression level, spatial resolution, and playback interruptions on SSQ-based cybersickness responses in VR video viewing. In particular, the possible non-linear relationship between resolution and cybersickness remains insufficiently examined, although it may have important implications for balancing visual quality and user comfort.

This study investigates three main factors through SSQ assessments, which use VR videos with different quantization levels (QP22, QP28, QP34, QP40) to show compression strength, resolution (HD, fHD, 2.5K, and 4K) to demonstrate spatial quality, and temporal delays (initial, middle, and two 5-second interruptions) to show playback interruptions. The thirty-four study participants assessed their discomfort after watching modified video clips through an HMD. The analysis of the data resulted in discovering patterns that showed how content characteristics affected the severity of cybersickness that people experienced. The research results establish user response modeling frameworks that enable adaptive content delivery systems to achieve a real-time balance between VR system immersion and user comfort. This paper makes the following key contributions to the study of cybersickness and user-centered VR systems.

- The study simulates three specific factors, including different resolution effects, loading delay, and quantization effects, to create network-induced impairments and compression artifacts in VR video streams.
- A controlled user-centered subjective evaluation of cybersickness responses to three VR video manipulations, analyzing the effects of video compression, spatial resolution, and playback delays
- An empirical analysis showing that cybersickness increases under strong compression and playback interruptions, while the resolution results indicate a non-linear pattern in which the moderate-resolution condition

produced lower SSQ scores than both lower and higher resolution conditions.
- A comparison of SSQ subcomponents showing that disorientation and oculomotor symptoms are particularly sensitive to visual degradation and temporal discontinuity in the tested VR video conditions.
- A machine-learning analysis evaluating the feasibility of classifying cybersickness severity levels from condition-level VR video parameters, with the limitations of the dataset and validation procedure explicitly discussed.

2 Related work

Cybersickness has long been a major barrier to the widespread adoption of immersive VR systems [7, 10]. It has been recognized and investigated for its theoretical foundations, sensory integration, conflicts, and the contributing factors that trigger it. Nevertheless, recently the spotlight has been on discovering how media-specific factors like visual quality, compression, and temporal stability affect people's tolerance, especially in the case of VR video-based experiences [14, 15]. This section summarizes previous studies on cybersickness theory, visual and temporal influencing factors, and user-adaptive VR systems, thereby placing the current study within the more extensive context of user-oriented VR quality evaluation.

2.1 Cybersickness and sensory conflict

The study of cybersickness is based on the theory of sensory conflict [16], which states that motion sickness arises from the inconsistency of sensory input between the visual and vestibular systems. In VR, the user sees the simulated motion of self, but he or she is physically stationary, leading to discrepancies in perception that can cause nausea, disorientation, and discomfort in the eye. This theory has been generally accepted for explaining the sickness caused by simulators in flight simulators, driving simulators, and HMD-based VR systems.

Later theories like the postural instability hypothesis [17] have put more emphasis on the factor of prolonged difficulty in the control of posture rather than sensory mismatch alone. In this way, cybersickness is thought to occur through the user's inability to maintain the preferred body posture over time, which is a result of the unpredictable or unnatural motion cues. Research has shown that increased head movement, prolonged exposure, and unfamiliar locomotion patterns worsen instability and discomfort [18].

Recent models take into account not only sensory conflict but also individual susceptibility factors such as previous VR experience, gender, age, and motion sensitivity [19, 20]. These factors are relevant because they may influence how different users respond to the same VR video condition. However, the present study focuses primarily on content-level and playback-level parameters rather than on individual-difference modeling; therefore, prior VR experience, gaming experience, and motion-sickness susceptibility were not independently measured. Theoretical models help explain why cybersickness may differ across individuals, but they do not fully explain how specific media parameters, such as encoding quality, spatial resolution, or playback continuity, affect subjective discomfort in modern VR video applications.

2.2 Visual factors influencing cybersickness

Poor visual fidelity and low image quality are additional reasons for inducing cybersickness in VR video. According to earlier research, conditions such as low resolution, reduced frame rate, and compression artifacts impair spatial perception and increase visual discomfort [21, 22]. In VR videos, the users actively explore the scene through head movements and hence the major impact of these effects because of the constant change of visual input during head movements.

High compression ratios lead to the occurrence of visual impairments such as blocking, blurring, flickering, and inconsistencies over time, which are very much disruptive to the optic flow. Also, the continuous pattern of motion is very important for perceiving the motion and deciding about the space around [23]. In such a situation, the visual system gives unstable motion cues, which leads to an increased strain on the eye movement control system as well as a feeling of being lost. Many research works have concluded that the presence of visual artifacts in the immersive video will result in higher SSQ scores when compared to the situation where content is visually pristine [24, 25].

The resolution is another factor that performs dual functions. Higher resolutions are advantageous because they improve visual clarity, depth perception, and presence. They may cause even more motion in virtual environments and consequently increase nausea in sensitive users [26]. While very low resolutions have the opposite effect, they lack detail and therefore, the user has to rely more on eye accommodation and head movement, which may in turn raise discomfort. The study in [15] indicated that moderate resolutions in VR videos might provide a sweet spot between immersion and comfort, thus indicating a thorough investigation across different resolution levels along with compression parameters.

2.3 Temporal disruptions and playback delays

Temporal stability is considered an essential factor in assessing cybersickness in immersive environments. Frame drops, freezes, latency, and playback interruptions all disrupt the continuity of motion, which is a prerequisite for continuous sensorimotor integration [27]. In VR, even a temporary disruption can cause a sudden shift in the user's perception of motion, resulting in increased perception of nausea and disorientation.

The research on VR latency has found that the higher the motion-to-photon, the more the sickness symptoms, particularly in the case of head rotations [13, 28]. Another issue besides latency is the buffering of events and the stalling [29] of playback in streaming-based VR systems. Studies say that interruptions during the experience are more troubling than loading delays at the beginning because the users have already gotten used to the motion dynamics of the scene [30, 31].

Irregular or repeated playback interruptions increase the discomfort level even more, as they do not allow the users to set up stable motion expectations again [29]. In [32, 33] found that in the case of VR videos, the disorientation effect from temporal artifacts according to the SSQ score was higher than that of spatial artifacts only. Thus, it can be inferred that in the case of VR, temporal consistency could be a more important factor in mitigating cybersickness than video quality in interactive and head-tracked situations.

2.4 User modeling and adaptive VR systems

Recent studies in user modeling and adaptable VR systems have focused on personalizing immersive experiences to minimize discomfort and keep the user immersed. Among the adaptive strategies used, resolution, bitrate, frame rate, and field of view are dynamically adjusted based on the user's behavior, system limitations, or anticipated sickness levels [34, 35]. These approaches align with Quality of Experience (QoE) frameworks, which emphasize subjective user perception rather than purely technical performance.

Physiological and behavioral signals, such as heart rate variability, galvanic skin response, and head movement patterns, have also been investigated for real-time cybersickness prediction in VR [36, 37]. These approaches are valuable for continuous monitoring, but they often require additional sensors, calibration, and signal-processing pipelines, which may limit their practicality in controlled user studies or large-scale evaluations. Therefore, the present study focuses on SSQ-based subjective assessment and condition-level VR video parameters, providing a reproducible basis for examining how compression, resolution, and playback interruptions affect perceived cybersickness.

By connecting the SSQ outcomes to variable video parameters, the data-driven user modeling can be done without the need for intrusive Sensors. Recent studies have addressed that AI models trained with user ratings have the ability to predict the user discomfort and therefore influence streaming adjustments in VR [38, 39]. Nonetheless, current research tends to emphasize one parameter like bitrate or frame rate instead of thoroughly considering the interrelations between compression, resolution, and temporal stability. This limitation hinders the application of adaptive techniques in actual VR video scenarios.

The existing research has established the theoretical framework for cybersickness while identifying personal visual and temporal factors that require further investigation through a user-centered study that assesses VR video using different compression levels and resolutions and playback continuity. Very few studies explicitly associate these parameters with SSQ-based subjective outcomes and intend to provide the adaptive content strategies accordingly. The current research fills this void by conducting a controlled assessment of visual and temporal changes and revealing how the interaction of these changes can guide the design of user-centered VR systems that offer both immersion and comfort.

3 Subjective user study

Figure 1 shows the complete workflow of the proposed subjective user study, which begins with preparing VR videos featuring controlled visual and temporal changes and ends with collecting user data and conducting cybersickness assessments. Participants viewed the rendered VR content on an HTC Vive headset and provided subjective ratings using the standardized SSQ questionnaire.

3.1 Experimental design

Before conducting the subjective user study, a controlled within-subject experimental procedure was designed to evaluate how VR video manipulations affect perceived cybersickness. Initially, 37 participants took part in the experiment, including 21 males and 16 females, aged between 25 and 40 years. Before the experiment, all participants were briefed about the study objectives, experimental procedure, viewing tasks, and their right to discontinue participation at any time. Visual acuity and color vision were screened using Snellen 20/20 and Ishihara charts, respectively, and all participants reported normal visual acuity and normal color vision. Prior VR experience, gaming experience, and motion-sickness susceptibility were not independently measured; therefore, their possible influence on

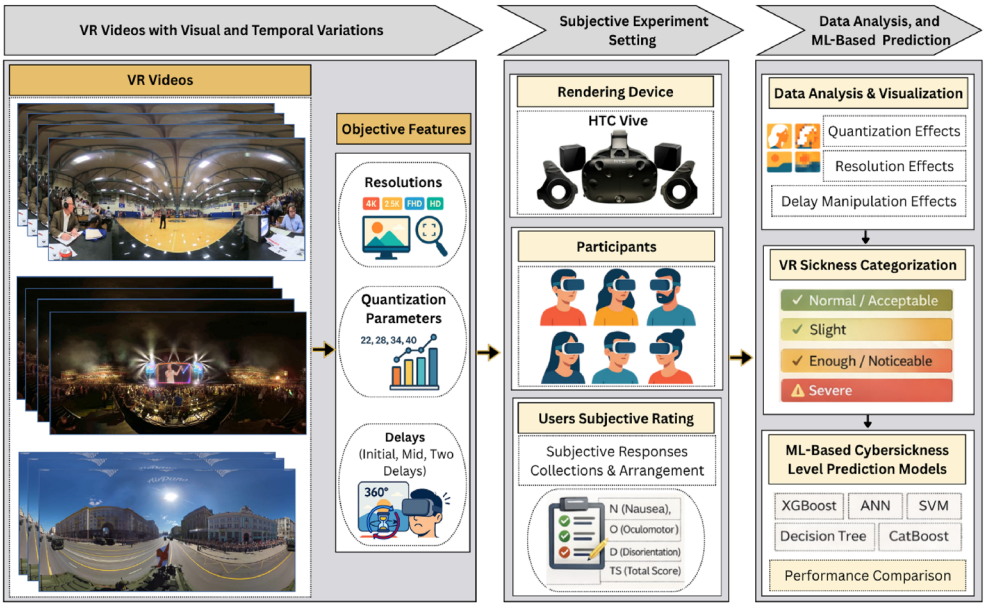


Fig. 1 Overview of the proposed framework, from VR video manipulation and a subjective user study to data analysis and ML-based prediction of cybersickness levels

cybersickness responses is acknowledged as a limitation of the study.

Before the main experiment, participants completed an approximately 30 min familiarization session with the HTC Vive HMD. During this session, the instructor explained how to wear, adjust, and use the headset, including proper fitting and visual adjustment. Participants were also given time to become familiar with the VR viewing environment before data collection began. Any issues related to headset fitting, visual comfort, or handling were resolved before the experimental trials started.

After the familiarization session, each participant viewed a series of VR video stimuli with controlled system-level manipulations. A total of eleven VR videos were shown to each participant in a fixed sequential order. The video material consisted of three 360° VR scenes: a basketball match, a music-hall concert, and a Russian Victory Day military parade. The experimental conditions varied in terms of quantization parameter (QP), spatial resolution, and playback delay. The videos were encoded using H.264 at 30 fps and presented through the HTC Vive HMD with a 110° field of view and a 90 Hz refresh rate.

The playback-delay conditions were implemented to simulate network-buffering events during VR video streaming. For the middle-delay and two-delay conditions, the video was temporarily paused for 5 s while a rotating buffering/loading spinner was displayed at the center of the participant's view. After the 5-second interval, normal video playback resumed automatically from the paused frame. The same procedure was used for the initial-delay condition, except that the buffering spinner was displayed for 5 s before the video started.

To reduce fatigue and mitigate the cumulative effects of cybersickness, a minimum 5-minute rest period was scheduled after each VR video. Participants were allowed to take longer breaks whenever needed based on their comfort level. They were also permitted to discontinue the experiment at any point without penalty if they experienced excessive discomfort or chose to withdraw from the study.

3.2 User data collection

After experiencing each VR video, the participants indicated their perceived cybersickness symptoms by using a standardized SSQ. The answers were taken under the direction of an instructor to make sure that they were accurate and complete. The analysis of data after the experiment revealed that three participants' responses were considered as statistical outliers and thus excluded from further analysis. Hence, the results of this study are based on the subjective scores obtained from 34 participants. The adoption of this structured experimental process was aimed at receiving dependable subjective feedback while concurrently safeguarding and comforting the participants throughout the study.

During the user study, the level of cybersickness was calculated using a standardized 16-item SSQ questionnaire [40] to measure the severity of cybersickness. The symptoms of cybersickness are divided into three categories: Nausea (N), Oculomotor (O), and Disorientation (D), with each category being assigned a score using a 4-point scale

(0; no sickness, 1; mild sickness, 2; considerable sickness, 3; severe sickness).

4 Subjective results and discussion

The subjective results based on 34 participants, illustrate the influence of three primary manipulations, such as quantization, resolution, and delay, on cybersickness severity when viewing VR videos through HDMs. The TS ranges were categorized as follows: 0–10 indicating normal or acceptable sickness levels [7, 41], 11–31 indicating slight sickness [7], 32–40 indicating enough or noticeable sickness [42, 43], and values above 40 representing severe sickness [7]. The analysis in this section is primarily descriptive. Mean SSQ scores and 95% confidence intervals are reported to summarize the observed cybersickness response patterns across the tested conditions. Therefore, the results are interpreted as descriptive trends within the present dataset rather than as confirmatory statistical evidence of condition effects. The following subsections discuss the observed results in terms of each manipulation type.

4.1 Quantization effects

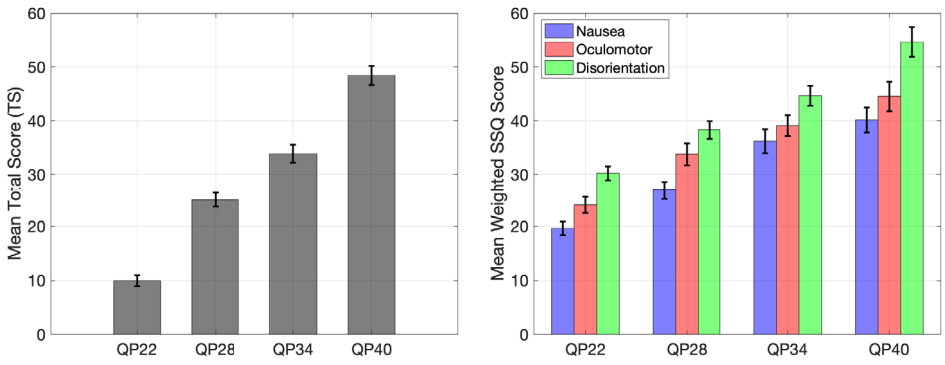
The quantization manipulation demonstrated a clear increase in cybersickness severity with higher compression levels shown in Fig. 2a. At QP22, the video quality was visually stable and smooth, resulting in a low mean TS of 9.99, categorized as normal or acceptable. The quantization parameter increase to QP28 and QP34 resulted in mean TS values of 24.96 and 33.89, which showed slight to noticeable sickness. The highest compression level at QP40 produced a mean TS value of 48.43, which indicates severe cybersickness. This progressive increase can be attributed to the loss of spatial and temporal fidelity caused by higher quantization. As the video becomes more compressed, artifacts such as blockiness, flicker, and smearing distort the optic flow, which the visual system relies on for motion perception. These visual distortions increase oculomotor strain and disrupt the synchronization between visual and vestibular cues, ultimately leading to sensory conflict. The presence of sensory conflicts in highly immersive virtual reality environments results in nausea, disorientation, and fatigue, which explains why higher sickness scores occur with increased QP values with 95% confidence intervals. The Fig. 2b analysis shows that all three SSQ subscales N O and D displayed increasing results with higher QP values. The study found that disorientation displayed the strongest reaction because it served as the primary driver of cybersickness development during extreme compression tests.

4.2 Resolution effects

The study found that different display resolutions affected the level of cybersickness presented in Fig. 3a. The 2.5K resolution display showed the least sickness symptoms with a mean TS value of 15.86, which falls within the normal range. The FHD and 4K display resolutions showed a small increase in TS values, which reached 20.96 and 19.85, respectively, that demonstrated UX mild to moderate sickness symptoms. The HD resolution produced a mean TS result of 24.87 which showed a tendency toward reaching the “enough” category.

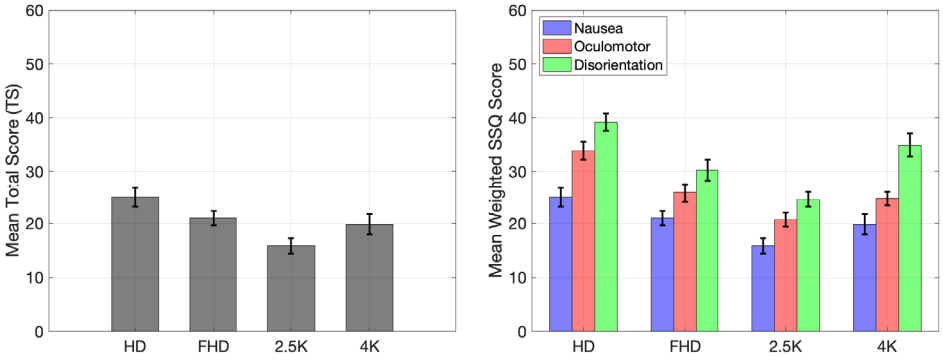
On the other hand, lower resolutions degrade the picture quality and introduce pixelation,forcing the eyes to overaccommodate and refocus, leading to increased oculomotor stress. The 4K resolution provides visual detail and motion perception cues, which create a stronger sensation of self-movement. The visual enhancement improves realism, yet it also strengthens the visual-vestibular mismatch that produces dizziness and nausea. The mid-level resolution (2.5K) provided the best balance between clarity and comfort, producing minimal cybersickness symptoms. The most enjoyable VR experience results from maintaining enough visual detail while controlling visual motion elements.

The Fig. 3b shows average weighted SSQ subscale scores for N O D which demonstrate that all three subscale scores



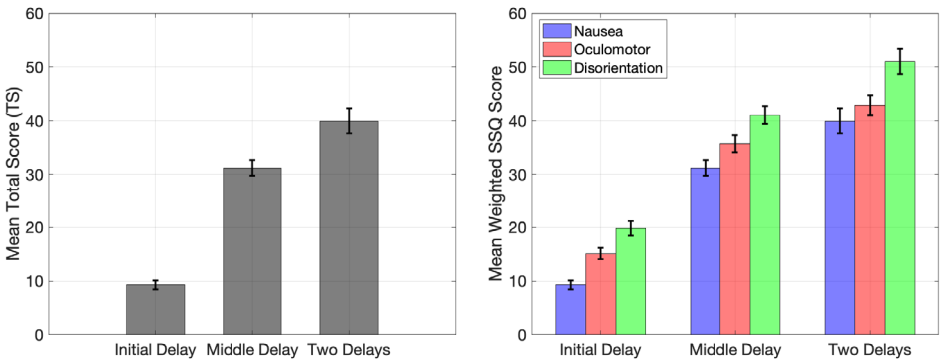
(a) Total scores (b) Mean weighted SSQ scores

Fig. 2 Effect of quantization parameters on cybersickness: **a** mean TS and **b** mean weighted SSQ subscale scores for nausea (N), oculomotor strain (O), and disorientation (D). Error bars indicate 95% confidence intervals



(a) Total scores (b) Mean weighted SSQ scores

Fig. 3 Effect of VR video resolution on cybersickness: **a** mean TS and **b** mean weighted SSQ subscale scores for nausea (N), oculomotor strain (O), and disorientation (D). Error bars indicate 95% confidence intervals



(a) Total scores (b) Mean weighted SSQ scores

Fig. 4 Effect of different types of delay in VR video on cybersickness: **a** mean TS and **b** mean weighted SSQ subscale scores for nausea (N), oculomotor strain (O), and disorientation (D). Error bars indicate 95% confidence intervals

decline from HD to FHD and the scores reach their minimum point at 2.5K resolution before showing a significant increase at 4K resolution. The U-shaped pattern shows that users experience less visual discomfort at moderate spatial resolution, while both low (HD) and high (4K) resolution cause them noticeable cybersickness.

The most profound response to resolution changes occurs with disorientation because its peak performance occurs at 4K resolution which produces sensory conflicts from enhanced visual detail and motion intensity during immersive experience. The results demonstrate that high resolutions create additional demands on both perceptual abilities and oculomotor functions, which emphasizes the need to choose the best resolution for maintaining visual quality and user comfort in VR environments.

4.3 Effect of delays in VR videos on users

The most significant changes in cybersickness symptoms occurred due to different types of delays in the VR videos while experiencing by users. Figure 4 presents the effect of different types of delay in VR video on users' cybersickness, where Fig. 4a shows the Mean TS for initial, middle and two 5-seconds delays and Fig. 4b presents the mean weighted SSQ sub-scale scores for N, O, D under different delay conditions. From Fig. 4a, the initial 5-second delay, occurring before motion, showed negligible impact (mean TS = 9.28), corresponding to normal levels of sickness. However, a 5-second delay occurring during video play significantly increased discomfort (mean TS = 31.21), placing it in the "enough" category. The two 5-second delays at different intervals in teh VR video resulted in the highest mean TS of 39.97, categorized as severe sickness. These results align with the sensory conflict theory, which explains that cybersickness arises when the brain receives mismatched motion cues from the visual and vestibular systems. Interruptions or freezes during movement break the temporal continuity of optic flow, causing users to experience abrupt mismatches between perceived and actual motion. Repeated interruptions amplify this mismatch, making it more difficult for users to reorient after each delay. Consequently, disorientation and nausea increase sharply, producing the highest sickness ratings under the two-delay condition.

In general, the findings from above three Figs. 2, 3, and 4, indicate that the severity of cybersickness is primarily governed by the stability and consistency of visual motion cues in VR. Conditions that maintain smooth, high-fidelity, and temporally continuous motion such as QP22, 2.5K resolution, and an initial delay produce minimal sickness. In contrast, conditions that introduce visual artifacts, low clarity,

or temporal disruptions such as QP40 and multiple delays exacerbate sensory conflict and increase discomfort. Ensuring smooth frame rates, moderate resolution, and minimal compression is therefore critical to maintaining comfort and reducing cybersickness in immersive VR video experiences.

5 ML-based cybersickness levels prediction

5.1 Data arrangement and prediction model

The cybersickness data collected after the subjective user study were organized for machine-learning-based classification. To perform supervised learning, the continuous Total SSQ (TS) scores were categorized into four ordinal cybersickness severity levels, which served as the target variable during model training. Specifically, TS values ranging from 0 to 10 were labeled Acceptable Sickness (Class 0) [7, 41], values from 11 to 31 were labeled Light Sickness (Class 1) [7], values from 32 to 40 were categorized as Noticeable Sickness (Class 2) [42, 43], and scores greater than 40 were classified as Severe Sickness (Class 3) [7]. These thresholds are based on the work of Stanney et al. [44]. This categorization transforms the continuous TS scores into discrete cybersickness severity categories suitable for multi-class classification. Each machine-learning sample represented one participant under one experimental condition. The input features consisted of the experimental condition representing the applied VR video manipulation together with the corresponding Total SSQ score, whereas the output variable was the four-class cybersickness severity level. The participant identifier was used only for data organization and was excluded from model training. Model performance was evaluated using stratified 10-fold cross-validation with multiple complementary evaluation metrics.

Each machine-learning sample represented one participant under one experimental condition. The input features consisted of the experimental condition representing the applied VR video manipulation together with the corresponding Total SSQ score, whereas the output variable was the four-class cybersickness severity level derived from the predefined TS thresholds. The participant identifier was used only for data organization and was excluded from model training. The purpose of the machine-learning analysis was to evaluate the capability of different machine-learning algorithms to classify predefined cybersickness severity categories under the tested VR conditions. To ensure robustness and generalization, model performance was evaluated using 10-fold cross-validation, and multiple complementary evaluation metrics were employed. These sickness levels are considered as the target output to train the XGBoost classifier. In order to guarantee the robustness and generalization of the model, its performance was assessed through 10-fold cross-validation, and furthermore, several different evaluation metrics were used.

5.2 Performance evaluation and result analysis

Figure 5a presents the average Precision–Recall (PR) curve, illustrates the trade-off between precision and recall across different decision thresholds. Figure 5b presents the average ROC curve obtained through 10-fold cross-validation, where the true positive rate (TPR) is plotted against the false positive rate (FPR). The average PR curve displayed in Fig. 5a demonstrates the trade-off between precision and recall for different decision thresholds. The model manages to have a high precision throughout a wide recall range, which means that the XGBoost classifier is effective in detecting the different levels of cybersickness while minimizing false positives. This trait is crucial in cybersickness prediction as misclassification especially under the case of

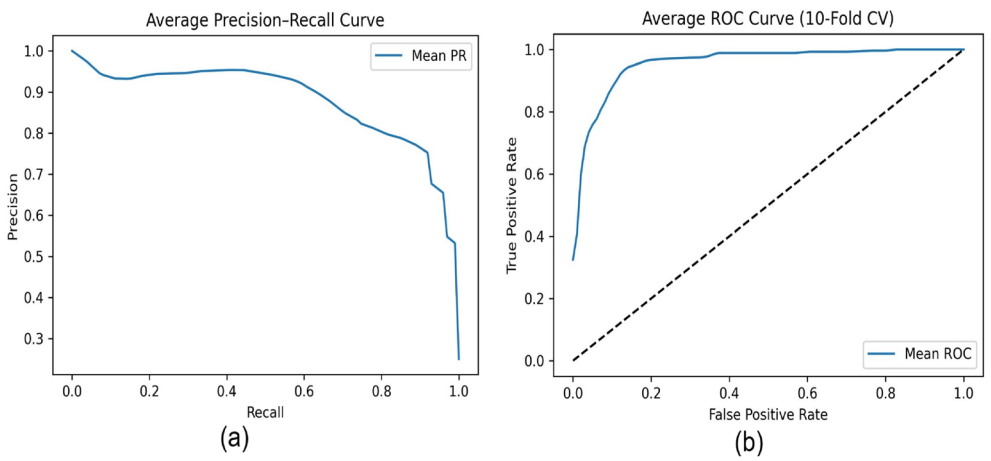


Fig. 5 Performance evaluation of the XGBoost-based cybersickness prediction model: **a** average Precision–Recall (PR) curve and **b** average Receiver Operating Characteristic (ROC) curve obtained using 10-fold cross-validation

severe sickness being wrongly identified as mild can lead to safety hazards for the user. The gradual decline in precision at very high recall points is an indication of the difficulty involved in separating adjacent sickness levels when all the samples are being classified, which is a common issue in multi-class, ordinal prediction tasks.

From Fig. 5b, the ROC curve mostly keeps its position near the upper-left corner, which signifies a great discriminative power of the XGBoost model throughout all classes of cybersickness. The substantial distance from the diagonal reference line indicating random classification adds to the certainty with which the model can discriminate between the different levels of sickness. Furthermore, the near-unity area under the ROC curve further confirms the robustness and reliability of the proposed prediction framework.

Figure 6a presents the average training and validation loss curves while Fig. 6b depicts the t-SNE visualization of the predicted class probability space obtained from the XGBoost model across 10-fold cross-validation. The average training along with validation losses shown in Fig. 6a, reveal the XGBoost model's learning behavior and convergence characteristics. Both curves show a quick fall in the loss during the first few training iterations and then a stable convergence as the training progresses. The training and validation losses maintained a small and consistent gap which indicates that the model performed well in terms of generalization and it is not likely that the model suffers from overfitting. This interaction of the learning phases (balanced) proves that the hyperparameters choice and cross-validation strategy are right for the dataset used.

In Fig. 6b, every single point stands for a data sample that has been projected into a 2D space with the colors indicating its predicted cybersickness class. The visual shows clusters of different sickness levels that are especially well-defined for the Acceptable and Severe Sickness classes. There is some overlapping of classes Light and Noticeable Sickness that are adjacent, but this is anticipated to some extent because of their ordinal nature. The overall separation qualitatively suggests that the learned feature representations preserve meaningful relationships among the cybersickness severity categories. However, the t-SNE visualization is intended only as an illustrative representation of the learned feature space and should not be interpreted as quantitative evidence of model generalization or predictive performance. Rather, it complements the quantitative evaluation by providing an intuitive visualization of the learned representations. The conclusions regarding classifier performance and robustness are based on the quantitative evaluation metrics obtained through the stratified 10-fold cross-validation experiments.

The t-SNE projections in Fig. 7 provide qualitative visualizations of the predicted class probability representations generated by the different machine-learning models under the stratified 10-fold cross-validation framework. Visually, the XGBoost and CatBoost models exhibit relatively more compact class groupings than the remaining classifiers, particularly for the Acceptable and Severe classes. Nevertheless, these visualizations are intended only to provide an intuitive illustration of the learned representation space and should not be interpreted as quantitative evidence of classifier superiority or model generalization. The comparative performance of the evaluated models is determined from the quantitative evaluation metrics reported in Table 1, including Accuracy, Precision, Sensitivity, F1-score, Specificity, and AUROC.

To further demonstrate the performance and robustness of the XGBoost-based cybersickness prediction model, we conducted a performance assessment that compared its results against commonly used machine learning frameworks such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees and CatBoost. Table 1 present the summary of the classification

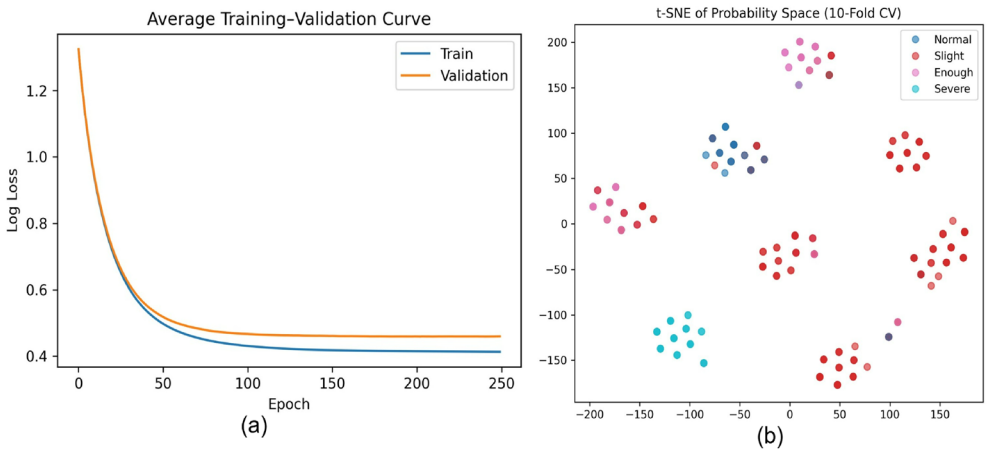


Fig. 6 Model learning behavior and feature space visualization: **a** average training and validation loss curves and **b** t-SNE visualization of the predicted class probability space across 10-fold cross-validation

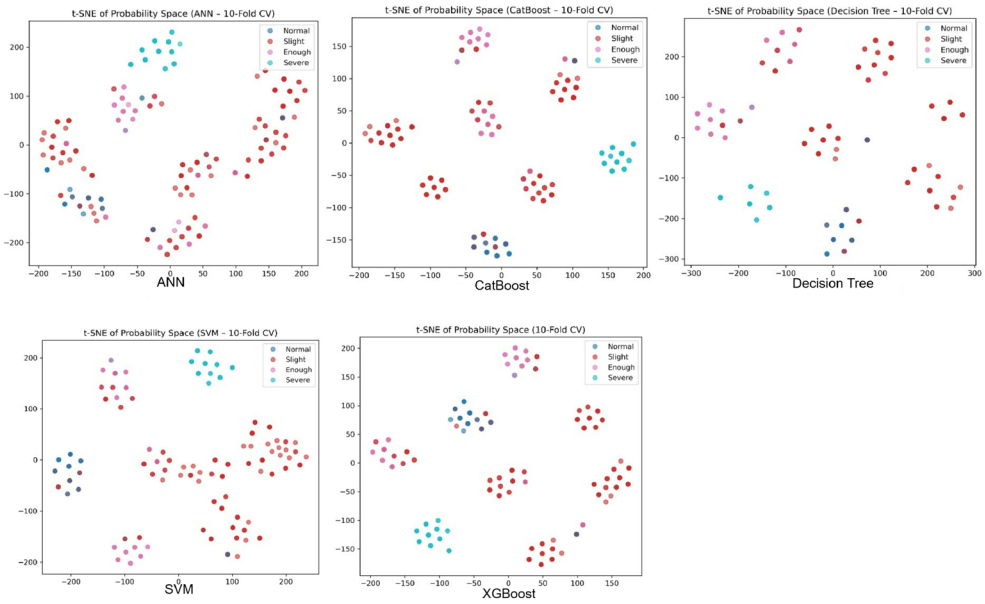


Fig. 7 t-SNE visualization of the predicted class probability space for ANN, CatBoost, Decision Tree, SVM, and XGBoost under 10-fold cross-validation, showing cluster separability across cybersickness levels

Table 1 Performance comparison of machine learning-based cybersickness level prediction using accuracy, precision, sensitivity, F1-score, specificity, and AUROC

Method	Accuracy	Precision	Sensitivity	F1	Specificity	AUROC
ANN	0.749868	0.581696	0.749868	0.649418	0.833049	0.830951
CatBoost	0.797487	0.858471	0.797487	0.809132	0.930204	0.915824
Decision Tree	0.797487	0.858471	0.797487	0.809132	0.930204	0.919392
SVM	0.797487	0.858471	0.797487	0.809132	0.930204	0.879501
XGBoost	0.812037	0.826448	0.812037	0.809363	0.913411	0.920514

performance across multiple evaluation metrics. The proposed XGBoost model achieves the highest overall accuracy (0.812) and AUROC (0.921), showing a superior discriminative capability across different levels of cybersickness. The model demonstrates a reasonable trade-off between sensitivity and F1-score, which is significant for avoiding underestimation of higher sickness levels. The consistent high AUROC results shows that the proposed model can reliably differentiate between different levels of cybersickness under various VR video quality and playback delay conditions. These results indicate that the proposed XGBoost-based cybersickness prediction model effectively captures complex, non-linear relationships between video parameters and user-perceived cybersickness. This make the proposed model a suitable choice for cybersickness-aware prediction in adaptive VR systems.

6 Conclusion

This work presented a comprehensive user-centered cybersickness evaluation in immersive VR video experiences while addressing the effects of video compression, spatial resolution, and playback delays. The study conducted a controlled laboratory subjective experiment to examine how different VR video compression methods, spatial resolution settings, and playback delays affected user-perceived cybersickness according to the standard SSQ Questionnaire. The results from the subjective assessment suggested that increased quantization levels together with temporal discontinuities and extreme resolution induce higher cybersickness symptoms. The research found that moderate resolution (2.5K) and low compression (QP22) delivered users an optimal experience which combined both visual quality and comfort. The analysis of SSQ subcomponents demonstrated that disorientation symptoms and oculomotor strain underwent higher sensitivity to quality degradation and temporal inconsistency which resulted in multiple symptoms of cybersickness.

The study findings also discussed and highlight the critical role of temporal continuity in VR experiences. The outcomes also suggested that active motion playback delays resulted in higher cybersickness symptoms than initial playback delays. This suggested that continuous visual display maintains stability of perception and user comfort in VR environments. Therefore, these findings suggest practical insights for VR content delivery and system design, where both visual quality adaptation and delay-aware playback strategies must be carefully considered to minimize user discomfort.

Unlike the quantization and delay results, which follow the expected pattern of increased cybersickness under stronger compression and more severe temporal disruption, the resolution results indicate a more nuanced non-linear pattern. The 2.5K condition produced the lowest mean Total SSQ score among the tested resolutions, whereas both HD and 4K were associated with higher discomfort. This suggests that, under the tested conditions, moderate resolution may provide a balance between sufficient visual clarity and reduced motion-related perceptual intensity. However, this result should be interpreted cautiously because it may depend on the specific video content, HMD configuration, encoding parameters, and participant sample.

Furthermore, the study utilizes XGBoost method for predicting user's different sickness levels achieved from subjective analysis. The proposed XGBoost model achieved better performance over other ML technique such as ANN, SVM, Decision Tree, and CatBoost through its evaluation of multiple classifiers which used SSQ scores to define sickness levels. The proposed models demonstrated their performance ability to predict user-perceived cybersickness by establishing clear class boundaries. This allows the model complex non-linear relationships between VR system parameters and user-perceived cybersickness. Therefore, this integration of subjective evaluation and predictive modeling shows the relevance and feasibility of cybersickness-aware and user-adaptive VR systems. The proposed study framework recommends informed, data-driven adaptation of VR content quality and playback conditions to balance user immersion and comfort while experiencing the VR content.

6.1 Limitations and future work

This study has several limitations that should be considered when interpreting the findings. First, the final analysis was based on 34 participants, which provides useful initial evidence but limits the generalizability of the results to broader VR user populations. Although participants were screened for visual acuity and color vision, individual factors such as prior VR experience, gaming experience, and motion-sickness susceptibility were not independently measured. These factors may influence how users respond to VR video quality degradation and playback interruptions.

Second, a baseline pre-exposure SSQ was not administered. Therefore, the analysis focuses on post-exposure cybersickness responses under different experimental conditions rather than on changes relative to each participant's baseline state. In addition, the VR videos were presented in a fixed sequential order. Although rest breaks were provided after each video, fixed presentation order may introduce order effects, fatigue effects, or carryover effects that cannot be fully excluded.

Third, the study used a limited set of 360° VR videos and a single HMD configuration. The results should therefore be interpreted within the tested content, hardware, encoding, and playback conditions. In particular, the observed non-linear resolution-related pattern may depend on the specific video scenes, motion characteristics, HMD display properties, and encoding settings used in this experiment. Further validation is needed before generalizing this pattern to other VR content types, interactive VR environments, or different headset configurations.

Fourth, the study relied primarily on SSQ-based subjective assessment. Although SSQ is widely used for cybersickness evaluation, it captures self-reported symptoms after exposure and does not provide continuous physiological or behavioral monitoring during VR use. Finally, the machine-learning analysis should be interpreted as an initial classification investigation based on the available experimental dataset. The relatively small dataset, class imbalance, and lack of external validation limit the extent to which the model can be generalized to unseen users, different VR content, or real-time adaptive VR systems.

There are some future work opportunities to extend this approach by addressing other evaluation methodologies such as physiological measurements, increasing number of users, more diverse user populations, and real-time adaptation mechanisms, enabling scientists and researchers to develop personalized and safe VR experiences.

Future studies could also collect pre-exposure individual-difference measures, such as the Motion Sickness Susceptibility Questionnaire (MSSQ), prior VR experience, and gaming experience, to examine how user susceptibility interacts with compression, resolution, and playback interruptions.

Acknowledgements This work was partly supported by the Knowledge Foundation, Sweden, through the Human-Centered Intelligent Realities (HINTS) Profile Project (contract 20220068). The authors would also like to thank IRC-FDE KFUPM for supporting this research.

Author contributions MSA: investigation, writing—original draft, software, visualization, formal analysis, data curation, conceptualiza-

tion, project administration, methodology. AMA: supervision, investigation, writing—review and editing, project administration, validation. MA: data analysis, investigation, validation, writing—review and editing. ME: data analysis, writing—review and editing, funding acquisition, resources, validation, conceptualization.

Data availability The data will be available on request.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Anwar, M.S., Yang, J., Frnda, J., Choi, A., Baghaei, N., Ali, M.: Metaverse and XR for cultural heritage education: applications, standards, architecture, and technological insights for enhanced immersive experience. Virtual Reality **29**(2), 51 (2025)
2. Laghari, A.A., Estrela, V.V., Li, H., Shoulin, Y., Khan, A.A., Anwar, M.S., Wahab, A., Bouraqia, K.: Quality of experience assessment in virtual/augmented reality serious games for healthcare: a systematic literature review. Technol. Disabil. **36**(1–2), 17–28 (2024)
3. Anwar, M.S., Choi, A., Ahmad, S., Aurangzeb, K., Laghari, A.A., Gadekallu, T.R., Hines, A.: A moving metaverse: Qoe challenges and standards requirements for immersive media consumption in autonomous vehicles. Appl. Soft Comput. **159**, 111577 (2024)
4. Anwar, M.S., Ali, M., Elwardy, M., Ahmad, S., Ahmad, P.N., Ahmad, M.: Gamified immersive learning and user experience evaluation of metaverse-based cultural heritage education. In: Proceedings of the International Workshop on Intelligent Immersification in the Metaverse: AI-Driven Immersive Multimedia, pp. 44–52 (2025)
5. Elwardy, M., Anwar, M.S., Ali, M., Hu, Y., Ahmad, P.N.: Evaluating the impact of participants' prior VR experience in a multi-user metaverse-based cultural heritage game. In: Proceedings of the 2025 31st ACM Symposium on Virtual Reality Software and Technology, pp. 1–2. (2025)
6. Davis, S., Nesbitt, K., Nalivaiko, E.: A systematic review of cybersickness. In: Proceedings of the 2014 Conference on Interactive Entertainment, pp. 1–9. (2014)
7. Shahid Anwar, M., Wang, J., Ahmad, S., Ullah, A., Khan, W., Fei, Z.: Evaluating the factors affecting qoe of 360-degree videos and cybersickness levels predictions in virtual reality. Electronics **9**(9), 1530 (2020)
8. Elwardy, M., Zepernick, H.-J., Hu, Y., Chu, T.M.C., Sundstedt, V.: Evaluation of simulator sickness for 360 videos on an hmd subject to participants' experience with virtual reality. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), pp. 477–484 (2020). IEEE
9. Elwardy, M., Zepernick H.-J., Hu, Y.: On head movements in repeated 360° video quality assessment for standing and seated viewing on head mounted displays. In: Proc. IEEE Conf. on Virtual Reality and 3D User Interfaces Abstracts and Workshops, Lisbon, Portugal, pp. 71–74 (2021)
10. Hasan, S., Wang, J., Anwar, M.S., Zhang, H., Liu, Y., Yang, L.: Investigating the potential of VR in language education: a study of cybersickness and presence metrics. In: 2024 13th International Conference on Educational and Information Technology (ICEIT), pp. 189–196. IEEE (2024)
11. Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G.: Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int. J. Aviat. Psychol. **3**(3), 203–220 (1993)
12. Gianaros, P.J., Muth, E.R., Mordkoff, J.T., Levine, M.E., Stern, R.M.: A questionnaire for the assessment of the multiple dimensions of motion sickness. Aviat. Space Environ. Med. **72**(2), 115–119 (2001)
13. Ryu, Y., Ryu, E.-S.: Overview of motion-to-photon latency reduction for mitigating VR sickness. KSII Trans. Internet Inf. Syst. (TIIS) **15**(7), 2531–2546 (2021)
14. Sameri, J., Damme, S., Schwarzmann, S., Wei, Q., Trivisonno, R., Turck, F., Vega, M.T.: Physiology-driven user perception prediction for networked collaborative virtual reality. IEEE Access **13**, 201183–201196 (2025)
15. Anwar, M.S., Wang, J., Khan, W., Ullah, A., Ahmad, S., Fei, Z.: Subjective qoe of 360-degree virtual reality videos and machine learning predictions. IEEE Access **8**, 148084–148099 (2020)
16. Reason, J.T., Brand, J.J.: Motion sickness, Academic press (1975)
17. Riccio, G.E., Stoffregen, T.A.: An ecological theory of motion sickness and postural instability. Ecol. Psychol. **3**(3), 195–240 (1991)
18. Munafo, J., Diedrick, M., Stoffregen, T.A.: The virtual reality head-mounted display oculus rift induces motion sickness and is sexist in its effects. Exp. Brain Res. **235**(3), 889–901 (2017)
19. LaViola, J.J., Jr.: A discussion of cybersickness in virtual environments. ACM Sigchi Bull. **32**(1), 47–56 (2000)
20. Rebenitsch, L., Owen, C.: Review on cybersickness in applications and visual displays. Virtual Reality **20**(2), 101–125 (2016)
21. Jerald, J.: The VR book: human-centered design for virtual reality. Morgan & Claypool (2015)
22. Chang, E., Kim, H.T., Yoo, B.: Virtual reality sickness: a review of causes and measurements. Int. J. Human-Comput. Interact. **36**(17), 1658–1682 (2020)
23. Koulieris, G.A., Drettakis, G., Cunningham, D., Mania, K.: Gaze prediction using machine learning for dynamic stereo manipulation in games. In: 2016 IEEE Virtual Reality (VR), pp. 113–120 (2016). IEEE
24. Kim, H.K., Park, J., Choi, Y., Choe, M.: Virtual reality sickness questionnaire (vrsq): motion sickness measurement index in a virtual reality environment. Appl. Ergon. **69**, 66–73 (2018)
25. Singla, A., Göring, S., Keller, D., Rao, R.R.R., Fremerey, S., Raake, A.: Assessment of the simulator sickness questionnaire for omnidirectional videos. In: 2021 IEEE Virtual Reality and 3D User Interfaces (VR), pp. 198–206. IEEE (2021)
26. Fernandes, A.S., Feiner, S.K.: Combating VR sickness through subtle dynamic field-of-view modification. In: 2016 IEEE Symposium on 3D User Interfaces (3DUI), pp. 201–210. IEEE (2016)
27. Dennison, M.S., Wisti, A.Z., D'Zmura, M.: Use of physiological signals to predict cybersickness. Displays **44**, 42–52 (2016)
28. Choi, S.-W., Lee, S., Seo, M.-W., Kang, S.-J.: Time sequential motion-to-photon latency measurement system for virtual reality head-mounted displays. Electronics **7**(9), 171 (2018)
29. Anwar, M.S., Wang, J., Ullah, A., Khan, W., Ahmad, S., Fei, Z.: Measuring quality of experience for 360-degree videos in virtual reality. SCIENCE CHINA Inf. Sci. **63**(10), 202301 (2020)
30. Won, J., Kim, Y.S., et al.: A new approach for reducing virtual reality sickness in real time: design and validation study. JMIR Serious Games **10**(3), 36397 (2022)

31. Singla, A.: Assessment of visual quality and simulator sickness for omnidirectional videos. Springer (2024)
32. Guna, J., Geršak, G., Humar, I., Krebl, M., Orel, M., Lu, H., Pogačnik, M.: Virtual reality sickness and challenges behind different technology and content settings. Mobile Netw. Appl. **25**(4), 1436–1445 (2020)
33. Geršak, G., Lu, H., Guna, J.: Effect of VR technology matureness on VR sickness. Multimed. Tools Appl. **79**(21), 14491–14507 (2020)
34. Keshavarz, B., Ramkhalawansingh, R., Haycock, B., Shahab, S., Campos, J.: Comparing simulator sickness in younger and older adults during simulated driving under different multisensory conditions. Transport. Res. F: Traffic Psychol. Behav. **54**, 47–62 (2018)
35. Kuo, P.-C., Chuang, L.-C., Lin, D.-Y., Lee, M.-S.: Vr sickness assessment with perception prior and hybrid temporal features. In: 2020 25th International Conference on Pattern Recognition (ICPR), pp. 5558–5564. IEEE (2021)
36. Padmanaban, N., Ruban, T., Sitzmann, V., Norcia, A.M., Wetzstein, G.: Towards a machine-learning approach for sickness prediction in 360 stereoscopic videos. IEEE Trans. Visual Comput. Graphics **24**(4), 1594–1603 (2018)
37. Islam, M.S., Lim, S.: Sex and age differences in virtual reality (vr) sickness susceptibility in forklift driving simulation. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. **68**, pp. 1206–1208. SAGE Publications Sage CA, Los Angeles, CA (2024)
38. Sameri, J., Coenegracht, H., Van Damme, S., De Turck, F., Torres Vega, M.: Physiology-driven cybersickness detection in virtual reality: a machine learning and explainable ai approach. Virtual Reality **28**(4), 174 (2024)
39. Zaidi, S.F.M., Shafiabady, N., Beilby, J.: Identifying presence of cybersickness symptoms using ai-based predictive learning algorithms. Virtual Reality **27**(4), 3613–3620 (2023)
40. Walter, H., Li, R., Munafo, J., Curry, C., Peterson, N., Stoffregen, T.: A brief explanation of the simulator sickness questionnaire (ssq). APAL Coupling Study, 9–10 (2019)
41. Rebenitsch, L.R.: Cybersickness prioritization and modeling. Michigan State University (2015)
42. Kim, Y.Y., Kim, H.J., Kim, E.N., Ko, H.D., Kim, H.T.: Characteristic changes in the physiological components of cybersickness. Psychophysiology **42**(5), 616–625 (2005)
43. Bruck, S., Watters, P.A.: Estimating cybersickness of simulated motion using the simulator sickness questionnaire (ssq): a controlled study. In: Sixth International Conference on Computer Graphics, Imaging and Visualization, pp. 486–488. IEEE (2009)
44. Stanney, K.M., Kennedy, R.S., Drexler, J.M.: Cybersickness is not simulator sickness. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. **41**, pp. 1138–1142. SAGE Publications Sage CA, Los Angeles, CA (1997)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.