Depth Perception Cues in VR Under Sleep Deprivation

Ammaar Zaman* Know Center Research GmbH
James Baumeister† University of South Australia
Ernst Kruijff‡ Bonn-Rhein-Sieg University of Applied Sciences
Eduardo Veas§ Graz University of Technology
Aleksandra Krajnc¶ Know Center Research GmbH
Neven ElSayed‖ Know Center Research GmbH

2026 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) | 979-8-3315-5945-8/26/$31.00 ©2026 IEEE | DOI: 10.1109/VR67842.2026.00053

ABSTRACT

As virtual reality (VR) headsets become more comfortable and accessible, their growing use in high-stakes, time-critical settings raises concerns about fatigue. Fatigue impairs perceptual and cognitive functioning. It reduces oculomotor accuracy and visual focus, and can lead to an early decline in depth estimation performance. By augmenting the visual presentation with depth information cues, adaptive designs can reduce fatigue-related depth perception errors, enhancing safety and task effectiveness. Explicit cues present depth information directly through text or color, while subtle cues adjust scene properties, such as depth-dependent blur, to convey depth information implicitly without drawing overt attention. We examined how fatigue interacts with different cues in a 27-hour within-subject protocol. Across six overnight sessions (20:00–07:00), twenty-three participants completed a VR depth perception task at varying fatigue levels and under four cue conditions: no cue (baseline), text, color, and blur. Over the night, vigilance declined, sleepiness and mental effort increased, and simulator sickness rose before stabilizing, independent of cue condition. All augmented cues reduced depth estimation error relative to baseline. Text yielded the largest and most consistent accuracy gains, especially for farther targets and later sessions. At peak fatigue, response probability dipped for text but remained stable for blur and color, indicating an accuracy versus responsiveness trade-off. These results support mixed adaptive designs that default to subtle cues to preserve responsiveness at low alertness and introduce explicit overlays when precise metric information is needed.

Index Terms: Virtual Reality, Depth Perception, Fatigue, Sleep Deprivation, User Study, Visual Cues, Subtle Cues

1 INTRODUCTION

Modern Virtual Reality (VR) headsets are increasingly used in demanding contexts (e.g., remote operation and navigational decision support) where precision, situational awareness, and rapid decision making are essential. In these contexts, users often remain in VR for extended periods, making fatigue difficult to avoid. While hardware advances reduce some physical discomfort, the gradual fatigue related decline in perceptual performance remains largely unaddressed in VR interface design [67, 12, 68].

Fatigue is a common challenge across operational domains. It can impair perception, cognition, and motor performance [47, 65, 68]. It can result from sleep loss, prolonged task engagement, or sustained mental workload, and it typically reduces visual acuity and attentional control [66, 79]. In immersive displays, such effects matter because accurate distance and depth judgments rely on sustained attention and stable oculomotor control, including binocular coordination and gaze stability. When fatigue is driven by sleep loss, binocular coordination and gaze stability can degrade, increasing errors in spatial judgment [71, 69]. It also increases lapses, so fatigue effects may appear as missed responses in time critical interaction [13, 4, 5]. These effects align with prior research on VR induced visual fatigue, which manifests as blurred vision, eye strain, and reduced focus [22]. Hence, fatigue in VR can reduce spatial understanding and depth estimation, increasing the risk of errors in navigation and object interaction. It can also make visual information harder to interpret, increasing the effort required to read overlays, monitor displays, or judge spatial relationships. But, since VR allows control over the visual presentation, such effects can be averted through adaptive designs that respond to the user's state.

In VR, augmentation refers to adding visual information or modifying scene appearance to support perception and interaction. In this work, we focus on augmented depth cues that support distance judgments. Two cue classes are commonly discussed, explicit cues and subtle cues. Explicit cues convey distance through overlays such as text labels or coded color mappings [14, 25]. In practice, these overlays can improve accuracy, but they may also impose additional cognitive demands during extended use [10, 14]. Subtle cues instead modify scene properties such as depth dependent blur to reinforce depth without overtly capturing attention [28, 35, 45]. This approach can be valuable in operational environments where visual clutter must be minimized and task relevant visual details must remain visible. Prior work has shown benefits for these cues in well rested participants in controlled settings [78]. This remains an open question because prior work does not establish which VR depth augmentation strategies remain robust when attentional stability declines, or whether cue benefits shift from accuracy to responsiveness under fatigue.

We evaluated these cues in a dynamic depth judgment task where participants responded when an approaching target reached a designated depth plane [9]. The task emphasizes depth relevant perception and action under fatigue rather than static depth matching. Participants took part in a 27-hour sleep-deprivation protocol, performing repeated trials under four cue conditions, baseline (no cue), explicit cues (text and coded color), and a subtle cue (blur). We recorded response probability and depth error, alongside subjective sleepiness, perceived effort, simulator sickness, and vigilance.

Across the night, vigilance declined, sleepiness and effort increased, and simulator sickness rose before stabilizing. All augmented cues reduced depth error relative to baseline. Text produced the largest, most consistent accuracy gains, especially for farther targets and later sessions.

At peak fatigue, response probability decreased most for text, while blur and color maintained higher responsiveness. Meanwhile, text remained the most accurate cue on responded trials. This pattern indicates that fatigue can change how cues help performance, so a cue that improves accuracy may still increase lapse risk when attention becomes unstable. This motivates cue designs that prioritize response reliability under low alertness.

*e-mail: azaman@know-center.at
†e-mail: james.baumeister@unisa.edu.au
‡e-mail: ernst.kruijff@h-brs.de
§e-mail: eveas@tugraz.at
¶e-mail: akrajnc@know-center.at
‖e-mail: nelsayed@know-center.at

2642-5254/26/$31.00 ©2026 IEEE
DOI 10.1109/VR67842.2026.00053

Authorized licensed use limited to: Technische Universitaet Graz. Downloaded on April 08,2026 at 16:31:11 UTC from IEEE Xplore. Restrictions apply.

2 Background and Related Work

Research on fatigue, depth perception, and visual augmentation in VR spans domains, from cognitive neuroscience to applied human–computer interaction. We review how fatigue affects attention and spatial judgment, how Head Mounted Display (HMD) constraints distort depth perception, and how augmented cues are used in VR. We then identify the gap in cue robustness under sustained fatigue.

2.1 Effects of Fatigue on Human Performance

Fatigue caused by sleep deprivation, prolonged cognitive effort, or extended task engagement undermines sensory, cognitive, and motor performance [38, 47, 48]. Typical consequences include slower reaction time, reduced attentional control, and declines in working memory [23, 65]. These effects have been reported across safety-critical contexts where sustained alertness is required [38]. In short, fatigue degrades the basic resources on which higher order perceptual judgments rely.

Sleep deprivation is closely linked to vigilance decrement and wake state instability [50, 53]. Under this instability, performance failures often appear as intermittent lapses and missed responses, and sustained attention requires greater effort as time awake increases [29]. These effects have clear consequences in applied perception and action tasks. For example, driving studies report slower responses, more lapses, and degraded vehicle control under sleep loss, with safety relevant outcomes [54]. Similar fatigue mechanisms are expected to affect time critical perception and action in immersive systems. In the next section, we focus on how fatigue influences spatial judgment and depth estimation.

2.2 Fatigue, Spatial Judgment and Depth Estimation

Depth perception is a key component of human spatial cognition, enabling effective interaction in both real and virtual settings. Depth judgments rely on combining binocular cues such as disparity, motion cues like motion parallax, and pictorial cues including occlusion, texture gradients, luminance, and relative size [51, 74]. In this context, fatigue impairs spatial judgment, the ability to estimate distances, positions, and motion trajectories in three dimensional environments. This skill is crucial for tasks such as interception, navigation, targeting, and obstacle avoidance. Fatigue can compromise spatial judgment through physiological and neural disruptions such as oculomotor instability, reduced binocular coordination, and slower visual processing speed [62, 69, 71].

Depth estimation, a subset of spatial judgment, is especially vulnerable because accurate egocentric distance perception depends on the integration of binocular disparity, convergence, and motion parallax. Under fatigue, these mechanisms can be disrupted, for example by reduced convergence accuracy or impaired smooth pursuit eye movements [69, 71]. Such disruptions increase the likelihood of timing and spatial errors in operational tasks that demand precise distance judgments [68]. In VR headsets, these same mechanisms are further constrained by display and rendering factors, which can amplify distance errors.

2.3 Depth Perception Challenges in VR

VR introduces unique challenges for depth perception because technological limitations in current head-mounted displays constrain the available cues. Depth perception in VR can be disrupted by vergence–accommodation conflict, display resolution and sampling, and the absence of stable spatial anchors [2, 15, 52]. Commercial HMD comparisons further document performance differences and simulator-sickness issues on depth tasks [34]. Even without fatigue, VR depth perception is limited by hardware and rendering constraints. Key contributors include the vergence–accommodation conflict, narrow field of view, and display parameters that constrain disparity sampling [33, 39, 41]. In action space, typically spanning 2–30 meters, users often underestimate egocentric distances and respond as though objects are closer than they are [15, 39, 59]. Classic cue weighting theory explains this compression, since depth perception normally relies on binocular, motion, and pictorial cues, but in VR these cues are compromised [16]. Mechanical properties of HMDs, including weight, inertia, and field of view restrictions, further amplify distance errors, with mock-HMD experiments confirming that device design itself contributes to perceptual bias [75]. Although some adaptation occurs with repeated exposure or calibration [15, 42], residual biases persist and remain problematic for precision-critical tasks. Even advanced solutions such as focus-tunable optics [61] or accommodation-invariant displays [44] have not fully resolved the compression of egocentric distance judgments in VR [41, 58].

Rendering level adaptations can also mitigate discomfort. For example, gaze contingent adaptation of stereoscopic rendering parameters such as interaxial separation and depth of field blur has been shown to reduce cybersickness and improve comfort in immersive environments [70]. These display specific constraints mean that fatigue findings from other operational domains cannot be transferred directly to VR. Since depth in HMD VR is already distorted by factors such as the vergence–accommodation conflict [46]. This motivates testing which augmentation strategies preserve reliable depth based performance when alertness drops.

2.4 Interplay of Fatigue and Depth Perception in VR

While fatigue impairs spatial judgment across environments, VR introduces additional constraints that can amplify these effects. In extended VR use cases such as long duration training, overnight monitoring, teleoperation, and simulation based mission rehearsal, fatigue and workload are difficult to avoid [12, 68]. When fatigue is layered on top of HMD limitations, perceptual challenges become more pronounced. In VR, fatigue has been associated with increased stereoscopic visual fatigue, reduced depth sensitivity, and slower responding [27, 77, 79]. Related work also reports degraded binocular accuracy and changes in depth related neural markers such as visual evoked potentials [26, 77]. The combined effect of hardware related distortions and fatigue induced degradation substantially increases the risk of perceptual failure in VR. This combination is especially detrimental in applications where precise egocentric distance estimation determines safety, such as interception tasks, collision avoidance in remote robotics, or VR assisted procedures. Depth estimation is both a critical and fragile perceptual function in VR for operational contexts. Even small decrements from fatigue can have disproportionate operational consequences.

2.5 Augmented Depth Cues

Augmented depth cues supplement or reinforce natural depth information and are often grouped as explicit or subtle cues. In this paper, *explicit* cues convey distance through a symbolic mapping that must be interpreted, such as reading a numeric label or applying a learned color to depth rule. We treat color as explicit because it functions as a coded mapping rather than a depth signal specified by the scene itself [25].

Explicit cues such as numeric labels or color-coded mappings can improve depth judgments by providing overt metric information [3, 25, 49], but overlays in head-mounted displays can impose additional cognitive strain and attention demands [8].

Subtle cues such as blur, luminance gradients, aerial perspective, and gaze contingent stereo rendering can reinforce depth without drawing overt attention [28, 43, 45, 74]. For example, inverse blur has been shown to reduce depth estimation errors by 36–48% [36].

In practice, text needs focal attention and reading before acting, and coded color needs applying a learned mapping. Blur instead changes image clarity with depth in a way that can complement disparity with less symbolic decoding [28]. Since sleep deprivation

Authorized licensed use limited to: Technische Universitaet Graz. Downloaded on April 08,2026 at 16:31:11 UTC from IEEE Xplore. Restrictions apply.

increases vigilance decrement and attentional lapses [50, 23], cues that rely on sustained decoding may be more vulnerable to missed responses than cues that reinforce depth at a perceptual level.

Fatigue has also been shown to degrade depth processing, visuomotor coordination, and stereoscopic sensitivity in VR [26, 27, 76, 77], while augmented cues can improve performance or comfort in well-rested users [35, 45, 78]. However, direct comparisons of explicit and subtle cue classes under sustained fatigue remain limited. In summary, it remains unclear how these cue classes compare under sustained fatigue, especially when distinguishing accuracy on responded trials from responsiveness. This gap also raises a design question about how cueing should be adjusted as alertness changes. In physiological computing, fatigue and workload are treated as input to a closed loop system, where psychophysiological or behavioral markers can drive real time interface adaptation [21, 20].

3 Research Focus and Contributions

Our study aims to evaluate which depth cue strategies remain robust as alertness declines under fatigue. We treat fatigue as a controlled cognitive and physiological stressor to evaluate cue performance as alertness declines. We compare explicit and subtle VR depth augmentation strategies under controlled overnight fatigue and quantify the trade offs that emerge as alertness declines. We examine two classes of augmented cues: **explicit** (text, coded color) and **subtle** (blur). The aim is to understand how these cues interact with fatigue to affect perceptual accuracy, user experience, and cognitive effort. Our hypotheses draw on vigilance decrement accounts of sleep loss, where reduced attentional stability increases lapses and effort, and where performance failures may appear as missed responses rather than uniform slowing [50, 53, 29].

The study addresses the following research questions:

RQ1: How does fatigue interact with cue type to affect performance in a dynamic depth judgment task, measured by response probability and depth error.

We expected blur to be less vulnerable to fatigue related reductions in response probability than text or color, because explicit cues require symbolic decoding that becomes more fragile under reduced attentional stability [50, 53, 29]. **H1**.

RQ2: Do augmented cues retain accuracy benefits over baseline across the night and across distances.

We expected that all augmented cues (text, color, blur) would outperform the baseline (no cue) in accuracy on responded trials as fatigue accumulates. **H2**.

RQ3: Do explicit overlays produce higher subjective burden than blur during prolonged use under fatigue.

We expected higher subjective burden for text or color than for blur, because head-worn overlays can impose additional parsing demands during task execution [8]. **H3**.

We evaluate RQ1 and H1 using mixed effects models for response probability and depth error. We evaluate RQ2 and H2 using cue against baseline contrasts across sessions and distances. We evaluate RQ3 and H3 using mixed effects models for KSS, mental effort, and SSQ, while PVT tracks vigilance decline across sessions.

This work provides empirical guidance for immersive systems that must sustain spatial perception in high-stakes, extended-use scenarios such as remote robotics, emergency response, and prolonged simulation training. The contributions are as follows:

Framing depth support as a fatigue-sensitive design problem. Sustained wakefulness reduces alertness and attentional capacity, and cues that are effective under normal conditions may incur higher processing cost as fatigue accumulates. In particular, explicit overlays require attentional selection and symbolic decoding, whereas subtle blur strengthens depth information with lower decoding demands. This motivates treating cue selection as fatigue-sensitive, rather than assuming one cue is best overall.

A controlled comparative evaluation under sustained sleep loss. We employ a within-subject overnight protocol with counterbalanced cue conditions and distances in a standardized environment to reduce external influences on alertness. We combine behavioral accuracy, responsiveness, and state measures across six sessions, providing direct evidence on how cue effectiveness changes over time as fatigue increases.

Responsiveness treated as a primary outcome. Sleep deprivation increases attentional fluctuations and missed responses, so fatigue is not fully captured by response speed on responded trials alone [5]. Restricting analyses to completed trials can inflate accuracy estimates and obscure a critical operational risk. To address this, we run two linked mixed-effects models with the same predictors, a logistic model for the probability of responding on each trial and a continuous model for error size on the responded trials. This separates two questions that can diverge under fatigue, whether users respond at all and how accurate those responses are. This framing shows that cue classes can create an accuracy versus responsiveness trade-off under fatigue.

Design opportunities for adaptive augmentation. Our findings suggest that subtle depth support can stay usable during extended sessions, while explicit overlays can be applied selectively when precise metric information is required. Overlay placement and density can also be adjusted based on fatigue state and vigilance indicators to limit clutter and maintain usability. These implications translate the cue comparison into implementable interface guidance by identifying which strategies stay usable as alertness declines and what trade-offs they introduce under fatigue in head-mounted VR.

4 Experimental Setup

4.1 VR-Based Asteroid Task

A previously reported interactive VR-based asteroid game [78] was implemented to assess depth estimation performance in dynamic virtual environments under four cue conditions (no cue, text, color, and blur; see Fig. 1) and at three catching frame distances (near, medium, and far). During the game (further referred to as the *task*), approaching asteroids had to be tagged as soon as they reached a visible planar catching frame at a known egocentric distance, simulating depth-based object interaction in VR. The catching frame served as an explicit spatial reference for the judgment.

This is a dynamic depth judgment task as responses must be timed to a fixed depth plane for an approaching target and therefore performance reflects depth relevant perception and action under fatigue rather than isolated depth cue integration measured by classical static matching. The task mechanics, cue design, and visual encoding followed the non-clutter implementation protocol in [78] to simulate depth perception in minimal-clutter virtual scenes and to isolate the effects of different cue conditions under varying fatigue levels.

Each task session contained one block of each cue condition, with 81 trials per cue, resulting in 324 trials per task session (and 1944 trials over six task sessions; see Section 5.2). In each trial, a single target asteroid spawned at 10 m and moved on a straight path toward the observer. Its speed was resampled every frame within 1.5–2.5 m/s (i.e., instantaneous speed varied continuously throughout the approach from 10 m to 0 m), with speed-dependent cueing updates tracking the instantaneous speed. The asteroid's geometry and texture were intentionally irregular to limit familiar-size inferences. Trajectories were drawn from nine predefined paths spanning the field of view, with a random variation of horizontal and vertical approach angles across different trials. The catching frame appeared at the egocentric distance of 3 m (near), 5 m (medium), or 7 m (far). Within each cue condition, frame distance followed a pre-generated balanced schedule so that each distance occurred 27 times (81 trials = 27 $\times$ 3 distances).

Authorized licensed use limited to: Technische Universitaet Graz. Downloaded on April 08,2026 at 16:31:11 UTC from IEEE Xplore. Restrictions apply.

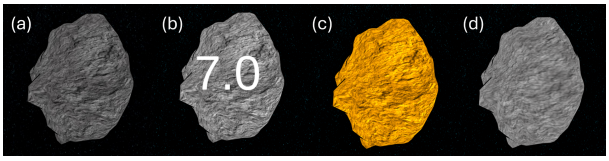


Figure 1: Augmented depth cues. (a) No cue (baseline): asteroid texture unmodified. (b) Text (explicit): egocentric distance (m) overlaid. (c) Color (explicit): fixed hue-to-depth map (yellow=far, red=near). (d) Blur (subtle): texture blur increases with distance. Rendered at 7 m; other scene elements unchanged [78].

As soon as the asteroid entered the catching frame, the observer pressed the index trigger on a handheld controller. We used a single trigger press to keep motor demands constant across conditions and reduce variability from larger movements [55]. No feedback was provided during or after trials. One task session lasted ≈27 minutes (≈6:45 per cue), and including the PVT, pre-task KSS and mental effort ratings, and post-cue SSQs, a full experimental session took ≈37–40 minutes.

A *response* was the first trigger press at any time from asteroid spawn (10 m) until the asteroid reached 0 m (trial end); a *lapse* was no trigger press during this window. The *response probability* was $\hat{p} = \frac{n_{\text{response}}}{n_{\text{response}}+n_{\text{lapse}}}$ (i.e., 1− lapse rate). Accuracy for response trials used the signed depth error $\Delta = z_{\text{asteroid}}(t_{\text{response}}) - z_{\text{frame}}$ with $z_{\text{frame}} \in \{3,5,7\}$ m, where $\Delta < 0$ indicates an early response and $\Delta > 0$ indicates a late response relative to the catching frame. We analyze Δ^2 as the primary accuracy outcome to model error magnitude while avoiding cancellation of early and late responses when averaging. Squaring also penalizes larger misses more disproportionately, which matches the operational motivation of the task. Separating response probability and error magnitude supports interpreting cue effects as an accuracy versus responsiveness trade-off under fatigue.

5 Experimental Design and Procedure

5.1 Sample Selection Criteria and Procedure

Only healthy adult participants were eligible to participate in the study. Inclusion criteria were age 18–40 years, no chronic health conditions, no regular medication use, non-smoking, and normal or corrected-to-normal vision (excluding glasses). Participants were also required to have an established and consistent sleep routine to reduce variability in baseline *homeostatic sleep pressure*, so that they began the protocol from comparable alertness levels. These criteria were chosen to minimize between-subject variability in susceptibility to sleep loss and in the study's measured outcomes, including PVT vigilance and depth-judgment performance.

Screening captured major moderators of sleep-loss vulnerability, including shift-work history, recent time-zone travel, chronotype and sleep-risk indicators, and habitual caffeine and alcohol use. Baseline sleep habits were verified using a sleep and wake diary during the week preceding the laboratory night. In the screening questionnaire, participants reported a habitual sleep duration of $M = 7.87$ h ($SD = 0.88$; $N = 23$). The target sample size was set by feasibility constraints inherent to a 27-hour, continuously monitored sleep-deprivation protocol with repeated overnight testing, and is comparable to prior in-laboratory total sleep-deprivation designs with repeated vigilance assessment (e.g., [32, 13]).

The study was approved by UniSA Human Research Ethics Committee (HREC), and all participants received monetary compensation. Prior immersive-technology experience was assessed at screening. 7/23 (30.4%) reported no prior VR/MR use, and 15/23 (65.2%) reported limited prior use, while one participant did not report VR/MR experience. To account for individual differences in vulnerability to sleep loss, all mixed-effects analyses included participant random effects.

5.2 Repeated-Measures Design

We examined cue performance across six overnight sessions during a 27-hour sleep-deprivation protocol at 20:00, 22:00, 00:00, 03:00, 05:00, and 07:00 (see Section 5.4). Each session included a PVT run (Section 6.1), pre-task ratings (KSS and mental effort; Section 6.2), one VR asteroid task run (Section 4.1) with four cue blocks, and post-cue ratings (KSS, mental effort, and SSQ) after each cue block.

To limit learning and carry-over effects across sessions, participants completed a supervised familiarization block before Session 1 (structurally identical to a task block) and familiarization trials were excluded from analysis. No performance feedback was provided during or after trials to reduce explicit reinforcement and strategy shaping. Within each VR task session, cue order (no cue, text, color, blur) was counterbalanced using a 4×4 Latin-square schedule so that each cue appeared equally often in each ordinal position across participants.

5.3 Implementation and Equipment

The task was implemented in Unity3D (v2021.3.15f1) and ran on a Meta Quest Pro headset via Oculus Link. Rendering used 4128×2096 at 72 Hz over a cabled connection to maintain stable visual fidelity. Audio and haptic feedback were disabled. Interpupillary distance was adjusted using the Quest fit procedure prior to each session. The scene used a dark, space-themed background, and directional lights, shadows, and texture compression were disabled to suppress incidental spatial anchors outside the experimental manipulation context.

The task was performed in a seated-upright position using an armless chair. Participants were instructed to keep the head and torso still while looking straight ahead and minimizing head movements during task execution. Responses were made with the right-hand controller using the index trigger as the only input. A single trigger press was used to keep motor demands and movement amplitude constant across conditions, helping isolate cue-related effects under fatigue [55]. Touch-based responses would require arm transport and introduce additional motor variability and a different latency budget, which can affect both performance and perception in time-critical tasks [55, 30, 31].

5.4 Sleep-Deprivation Protocol

To ensure a consistent starting point for fatigue onset, participants maintained a regular 8-hour sleep schedule for seven days before the laboratory night. Compliance was verified using sleep diaries, time-stamped text messages before sleep and after waking, and in-lab questioning. This reduced the risk that participants entered the protocol with different baseline sleep pressure that could confound cue-dependent performance trends.

On the night before the laboratory session, participants restricted sleep to 5 hours from 01:00 to 06:00. This ensured a common wake time of 06:00 and aligned circadian phase at study onset. Starting from this wake time, the protocol spanned 27-hours, so the final session occurred after more than 24 hours awake. Participants arrived at the laboratory at 16:00 and were monitored continuously in a controlled environment until release the following morning.

5.5 Experimental Protocol

5.5.1 Onboarding Session

Upon their arrival to the lab at 16:00, participants received briefing on the study schedule, sleep restriction requirements, and behavioral guidelines. All participants had to provide a written informed consent to continue with the experimental procedure.

Authorized licensed use limited to: Technische Universitaet Graz. Downloaded on April 08,2026 at 16:31:11 UTC from IEEE Xplore. Restrictions apply.

5.5.2 Training Session

Following onboarding, participants received instructions for the PVT (see Section 6.1) and the subjective state assessments (see Section 6.2). The researcher then introduced the VR-based asteroid task (see Section4.1), oriented participants to the four cue conditions, and demonstrated how cue appearance changes with frame distance.

Training consisted of a supervised familiarization block immediately before Session 1 that was structurally identical to a task block. Participants practiced using the catching frame as a reference and aligning the cue information on the asteroid with the frame to respond at the correct depth. This familiarization ensured comprehension and reduced early learning effects.

5.5.3 Experimental Sessions

Six experimental sessions were administered at 20:00, 22:00, 00:00, 03:00, 05:00, and 07:00, with ∼50 min breaks in between sessions and a longer meal break following the 00:00 session. Break timing and session procedures were standardized to reduce uncontrolled recovery or activity differences across participants. Within each experimental session, the protocol execution order was: (1) PVT; (2) pre-task assessment (i.e., KSS and SSQ); (3) VR-based asteroid task; and (4) post-task assessment after delivery of each asteroid task condition (i.e., KSS, Mental Effort Rating Scale, and SSQ).

Environmental conditions were held constant across participants and sessions. Room lighting was fixed to reduce circadian drift, ambient noise was kept stable, and activity levels were standardized. Caffeine and vigorous activity were prohibited, and meal timing was controlled to avoid energy dips unrelated to sleep pressure. Participants were monitored continuously by staff, with two researchers per shift enforcing alertness rules and documenting deviations. Monitoring included verifying that participants remained awake, upright, and engaged during required tasks.

Access to personal electronic devices was restricted throughout the study to reduce external stimulation and preserve the internal validity of attentional and perceptual measures. During designated breaks, participants were permitted quiet leisure activities such as reading or puzzles. Interpersonal interaction was allowed but monitored to prevent discussion of experimental protocol content, ensuring the integrity of individual performance. Non-caffeinated beverages, snacks, and heating packs were offered upon request to support comfort without influencing alertness.

Meal times were fixed at 18:45, 02:00, and 08:30 to avoid energy dips unrelated to sleep deprivation. Pre-selection of meals ensured consistency and minimized metabolic confounds. Experimental sessions were spaced at approximately 90–100 minute intervals, with a longer 00:00–03:00 interval to accommodate the longer meal break. This spacing was designed to balance two goals: (1) capturing performance fluctuations across fatigue progression, and (2) avoiding task saturation or excessive recovery that could mask fatigue effects. Six experimental sessions at the planned intervals provided sufficient granularity to observe time-of-night effects and fatigue-associated cue interactions.

The final experimental session was completed around 08:00, after more than 24 hours of continuous sleep deprivation. Participants were debriefed and escorted home by third-party transport, as driving in this state posed ethical and safety concerns. All participants who entered the laboratory completed all six experimental sessions.

6 Measures

6.1 Psychomotor Vigilance Test

The Psychomotor Vigilance Test (PVT) was used to assess overall sustained attention (vigilance) to visual stimulation [18]. A simple, three-minute computer-based reaction test measures response speed at a millisecond level by means of button presses to visual stimuli presented at random time intervals. The psychomotor vigilance metric is defined in terms of the mean reciprocal reaction time (RRT), computed as $\mathrm{RRT} = \mathrm{mean}(1/\mathrm{RT})$ in s^{-1} over valid responses. The so-called false starts ($\mathrm{RT} < 100$ ms) and lapses ($\mathrm{RT} \geq 500$ ms) are treated as invalid and are not included in the RRT metric computation.

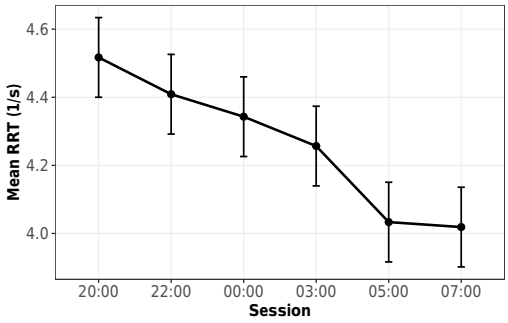


Figure 2: Estimated marginal means (EMMs) of PVT reciprocal reaction time (RRT; 1/s; higher values indicate faster responding) across sessions (20:00–07:00). Error bars show ±1 SE

6.2 Subjective State Assessment

Karolinska Sleepiness Scale (KSS): Perceived level of drowsiness was measured with the single-item 9-point Karolinska Sleepiness Scale (KSS; 1 = "extremely alert", 9 = "very sleepy, fighting sleep") [1, 37].

Mental-Effort Rating Scale: Perceived mental effort was measured with the single-item 9-point Mental-Effort Rating Scale by Paas (1 = "very, very low mental effort", 9 = "very, very high mental effort") [60].

Simulator Sickness Questionnaire (SSQ): The perceived severity of cybersickness was assessed with the Simulator Sickness Questionnaire (SSQ) that measures the severity of nausea, oculomotor, and disorientation symptoms on 16 4-point scales (0 = "none", 3 = "severe") [40]. For the analysis, only the Total SSQ Score was used.

7 Results

A total of 23 participants took part in the study (13 women, 9 men, and one genderqueer participant), aged 18–38 years ($M = 24.4$, $SD = 6.6$). Statistical analyses were conducted in R [63] using mixed-effects modeling. Linear mixed-effects models (LMMs) were performed using LME4 package [6], and generalized linear mixed-effect models (GLMMs) using glmmTMB package [11]. Where outcome distributions were approximately linear and symmetric, we used LMMs. Where outcomes were non-Gaussian, we used GLMMs with an appropriate family and link function.

Fixed effects were tested using Wald χ^2 tests (Type II/III), with estimated marginal means (EMMs) for post-hoc tests using Sidak or Holm adjustment for multiple comparisons. Where relevant, effects are reported on the natural scale, i.e., odds ratios (OR) for logistic models, ratios of estimated means for Gamma models, and mean differences for linear mixed models. All values were back-transformed from the model link function. Two-tailed tests used $\alpha = 0.05$ with 95% confidence intervals (CIs). All (G)LMM analyses included a random effect of *participant* to account for individual differences. Relevant distributions and link functions for GLMMs, fixed effects, and reference levels selected for individual analyses are reported at the beginning of each subsection. Fatigue was operationalized as *session* (time-of-night), reflecting the protocol progression across distinct testing times. KSS was collected once per session and is reported as a subjective state measure alongside performance outcomes.

Authorized licensed use limited to: Technische Universitaet Graz. Downloaded on April 08,2026 at 16:31:11 UTC from IEEE Xplore. Restrictions apply.

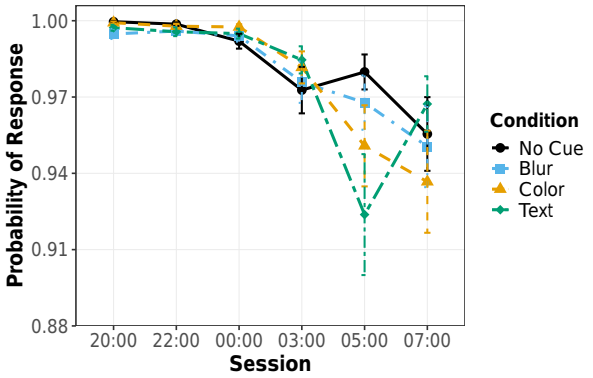


Figure 3: Asteroid task—response probability by session and cue. Lines and points show estimated marginal means (EMMs) of response probability from a binomial GLMM (logit link) across sessions (20:00–07:00) for each cue (None, Blur, Color, Text). Error bars show ±1 SE.

7.1 Psychomotor Vigilance Test Performance

Mean reciprocal reaction time (RRT; in 1/s) was analyzed using LMMs. Fixed effect was *session* (20:00, 22:00, 00:00, 03:00, 05:00, 07:00). Reference level was set to 20:00.

7.1.1 Reciprocal Reaction Time (RRT)

A Type II Wald χ^2 test confirmed a significant effect of *session*, $\chi^2(5) = 60.18, p < .001$. The model explained a small proportion of variance through fixed effects, large proportion when including both fixed and random effects ($R^2_{\text{marginal}} = .10$, $R^2_{\text{conditional}} = .78$).

Relative to the 20:00 baseline (EMM ≈ 4.52, SE ≈ 0.12), RRT was *lower at 03:00* ($p < .05$), and *lower at 05:00 and 07:00* (both $p < .001$), while 22:00 and 00:00 did not differ from baseline (ns.). Among non-baseline sessions, early-evening levels (22:00, 00:00) were *higher than late-night levels*: 22:00 > 05:00 ($p < .001$) and 07:00 ($p < .001$); 00:00 > 05:00 ($p < .01$) and 07:00 ($p < .01$). Adjacent late-night sessions (03:00 vs. 05:00; 05:00 vs. 07:00) did not differ (ns.); all other pairwise differences were not significant (ns.). Performance declined over the night (Fig. 2), consistent with accumulating sleep pressure and the early-morning circadian trough [50].

7.2 VR-Based Asteroid Task Performance

Both performance measures were analyzed using GLMM. Response probability was fitted with a binomial distribution and logit link. Depth-estimation accuracy (squared depth error, Δ^2) was fitted with a Gamma distribution and log link. The Gamma distribution was selected due to the positive, right-skewed nature of the data. We use Δ^2 to model error magnitude, and frame distance was included as a fixed effect with cue and session interactions to estimate cue effects conditional on distance. Fixed effects for both models were *cue* (no cue, text, color, blur), *session* (20:00, 22:00, 00:00, 03:00, 05:00, 07:00), and *frame distance* (3, 5, 7 m), including their two- and three-way interactions. Reference levels were *no cue* and 20:00.

7.2.1 Response Probability

A Type III Wald χ^2 test indicated that the three-way interaction of *cue* × *session* × *distance* was not significant, $\chi^2(30) = 25.88$, $p = .68$. Accordingly, we examined the two-way interactions. A Type III Wald χ^2 test showed a *significant cue* × *session interaction*, $\chi^2(15) = 225.78$, $p < .001$, whereas *cue* × *distance* and *session* × *distance* were not significant ($\chi^2(6) = 4.96$, ns.; $\chi^2(10) = 9.19$, ns.).

The model explained a moderate proportion of variance ($R^2_{\text{marginal}} = .27$, $R^2_{\text{conditional}} = .58$).

The post-hoc results show that under lower-fatigue sessions (20:00–00:00), response probability was near ceiling across all four cue conditions (*EMMs* > .99). A clear divergence emerged at *05:00: the text condition yielded the lowest response probability* (EMM = .950) relative to no cue (.987), blur (.980), and color (.971), see Fig. 3. Response probability for the text condition, whilst comparatively accurate in the depth-estimation analysis, was lower than for no cue, color, and blur. Response probabilities somewhat recovered by 07:00, consistent with circadian influences on alertness [50]. These results suggest that while textual depth cues may enhance depth-estimation accuracy, they may also increase lapses late in the night.

To further investigate differences at the 05:00 session, we fit a focused binomial GLMM (logit link) including *cue condition* and *frame distance* as fixed effects, with a random intercept for participant. A Type II Wald χ^2 test showed a significant main effect of *condition*, $\chi^2(3) = 125.12$, $p < .001$, and a main effect of *frame distance*, $\chi^2(2) = 26.67$, $p < .001$, while the *condition* × *distance* interaction was not significant, $\chi^2(6) = 4.76$, $p = .575$. Participants were more likely to respond in the no cue condition compared to color *no cue* > *color* (OR = 2.36, $p < .001$), blur *no cue* > *blur* (OR = 1.61, $p < .01$), and text *no cue* > *text* (OR = 4.05, $p < .001$). Among augmented cues, responses were more likely with blur than color *blur* > *color* (OR ≈ 1.46, $p < .05$), and more likely with both color *color* > *text* (OR = 1.72, $p < .001$) and blur *blur* > *text* (OR = 2.51, $p < .001$) than with text. (Here, OR>1 favors the first-listed condition; e.g., OR= 4.05 for none/text means the odds of responding were about four times higher with *none* than with *text*.) In addition, post hoc tests on *frame distance* indicated that response probability was lower at 3 m than at both 5 m *3 m* < *5 m* (OR = 0.70, $p = .005$) and 7 m *3 m* < *7 m* (OR = 0.57, $p < .001$), while the difference between 5 m and 7 m was not significant (OR = 0.805, $p = 0.197$). Taken together, response probabilities were uniformly high early in the night and then diverged at 05:00, with *text* showing the greatest vulnerability to lapses at 05:00, even though it improved depth-estimation accuracy.

7.2.2 Depth-Estimation Accuracy (Squared Δ)

A Type III Wald χ^2 test showed a significant *cue* × *session* × *distance* interaction, $\chi^2(30) = 188.70$, $p < .001$.

The model explained a moderate proportion of variance ($R^2_{\text{marginal}} = .27$, $R^2_{\text{conditional}} = .44$).

The post-hoc results show that depth estimation error (i.e., squared Δ) increased with fatigue (especially from 05:00 onward) and with greater distance (see Fig. 4), and that *cue effectiveness depended on both factors*. Relative to no cue, depth estimation error was *consistently lower for text across sessions and distances* (e.g., *text* < *no cue* at 07:00 and 7 m (ratio = 0.337, $p < .001$), with *color* and *blur* also conferring benefits under high fatigue and farther distance (e.g., *color* < *no cue* at 07:00 and 7 m (ratio = 0.653, $p < .001$), *blur* < *no cue* at 07:00, 7,m (ratio = 0.770, $p < .05$). At the 05:00 session, *text* remained significantly better than no cue (e.g., *text* < *no cue* at 05:00, 3,m (0.631, $p < .001$); *text* < *no cue* at 05:00, 5,m (0.431, $p < .001$), whereas *blur* and *color* often did not differ from no cue (ns.). These patterns highlight that depth-cue benefits are *largest under more demanding conditions (i.e., late night, far targets)*. Unless otherwise stated, results refer to post-hoc contrasts versus 'none'; contrasts not mentioned were ns.

Detailed Sidak-adjusted post-hoc contrasts (cue vs. *no cue*) across all session-by-distance cells are summarized in Table 1. Briefly, *text* outperformed *no cue* in all 18 cells ($p < .001$); *blur* improved 10/18 (mainly 3–5 m); *color* improved 8/18 (concentrated at 7 m), with one worse-than-none case at 00:00 and 3 m ($p < .05$).

Authorized licensed use limited to: Technische Universitaet Graz. Downloaded on April 08,2026 at 16:31:11 UTC from IEEE Xplore. Restrictions apply.

Table 1: Depth-estimation accuracy (squared Δ): ratio of estimated marginal means (cue / no cue) from the Gamma GLMM (log link), by session and frame distance. Ratios < 1 indicate lower squared error (better). Significance from Sidak-adjusted post-hoc contrasts vs. *no cue*: *** $p < .001$, ** $p < .01$, * $p < .05$, **ns** = nonsignificant. † = significantly worse than *no cue* (ratio > 1, $p < .05$).

Frame distance	Session	Text	Color	Blur
3 m	20:00	0.242***	1.220 **(ns)**	0.586***
	22:00	0.122***	0.857 **(ns)**	0.642***
	00:00	0.302***	1.309*†	0.469***
	03:00	0.229***	0.859 **(ns)**	0.736**
	05:00	0.631***	0.868 **(ns)**	1.009 **(ns)**
	07:00	0.341***	1.221 **(ns)**	0.606***
5 m	20:00	0.194***	0.804 **(ns)**	0.640***
	22:00	0.130***	0.867 **(ns)**	0.517***
	00:00	0.121***	0.959 **(ns)**	0.567***
	03:00	0.182***	0.683***	0.786*
	05:00	0.431***	0.886 **(ns)**	0.875 **(ns)**
	07:00	0.268***	0.761*	0.798 **(ns)**
7 m	20:00	0.141***	0.558***	0.856 **(ns)**
	22:00	0.192***	0.724**	0.858 **(ns)**
	00:00	0.136***	0.663***	0.911 **(ns)**
	03:00	0.194***	0.639***	0.921 **(ns)**
	05:00	0.319***	0.543***	0.852 **(ns)**
	07:00	0.337***	0.653***	0.770*

7.3 Subjective State Assessment

The Karolinska Sleepiness Scale (KSS), the perceived mental effort scale, and the Simulator Sickness Questionnaire (SSQ) were analyzed using LMMs. Fixed effects for all three LMMs were *[cue] condition* ([pre-task], no cue, text, color, blur), *session* (20:00, 22:00, 00:00, 03:00, 05:00, 07:00), and their two-way interaction (*[cue] condition* × *session*). Note that only KSS and SSQ include the *pre-task* condition, which was collected before the VR-based asteroid task execution and serves as a session-level state baseline. Accordingly, we use the term *condition* (rather than *cue condition*) for KSS/SSQ. Reference levels were set to pre-task/no cue and 20:00 for *condition*/*cue condition* and *session*, respectively.

7.3.1 Karolinska Sleepiness Scale (KSS)

A Type II Wald χ^2 test showed significant effects of *condition*, $\chi^2(4) = 28.89$, $p < .001$, and *session*, $\chi^2(5) = 880.61$, $p < .001$; the *condition* × *session* interaction was ns., $\chi^2(20) = 11.99$.

The model explained a substantial proportion of variance ($R^2_{\text{marginal}} = .42$, $R^2_{\text{conditional}} = .69$).

The post-hoc results show that *KSS ratings were lower during the pre-task condition (EMM = 5.45, 95% CI [4.82, 6.08]) than during all four task conditions (No Cue = 6.11, Text = 6.13, Color = 6.27, Blur = 6.14; all $p < .001$)*, with no differences among the task conditions (all ns., see Fig. 5a). These results suggest that perceived sleepiness increased significantly after engaging with the asteroid task, but was not affected by the type of augmented cue presented or whether a cue was presented at all.

Furthermore, KSS ratings also show an increase over sessions (see Fig. 5b). EMMs were 3.85 (20:00), 4.31 (22:00), 5.21 (00:00), 7.50 (03:00), 7.88 (05:00), and 7.37 (07:00). Pairwise tests showed: within the early evening, *20:00 < 22:00* ($p < .05$) and *22:00 < 00:00* ($p < .001$) (hence *20:00 < 00:00*, $p < .001$). Among late sessions, 03:00 vs. 05:00 and 03:00 vs. 07:00 were ns., while *05:00 > 07:00* showed a small decline by the final session ($p < .05$). *All early (20:00–00:00) vs. late (03:00–07:00) comparisons were significant* ($p < .001$). Overall, sleepiness rose from evening through night, peaked around 05:00, then slightly declined by 07:00.

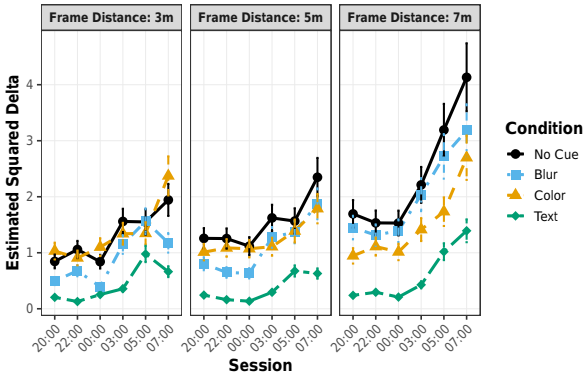


Figure 4: Asteroid task-depth estimation accuracy. Lines and points show EMMs of squared depth error (Δ^2, in m^2) across sessions (20:00–07:00), faceted by frame distance (3, 5, 7 m) and colored by cue, from a Gamma GLMM with a log link; error bars are ±1 SE of the EMMs. Lower values indicate more accurate depth estimation.

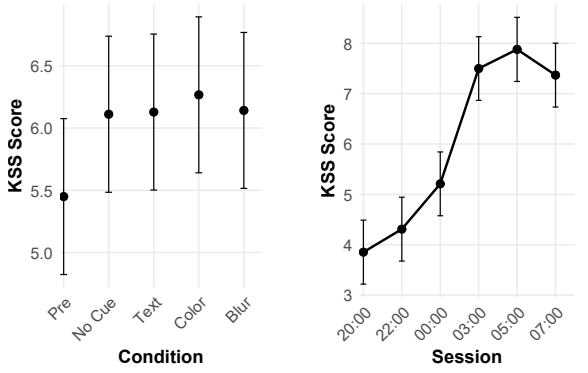


Figure 5: Estimated marginal means (EMMs) of Karolinska Sleepiness Scale (KSS) from LMMs. Left: by condition (collapsed across sessions). Right: by session (collapsed across conditions). Error bars show ±1 SE.

7.3.2 Mental-Effort Rating Scale

A Type II Wald χ^2 test revealed a significant main effect of *session*, $\chi^2(5) = 218.58$, $p < .001$. Neither *condition* nor the *condition* × *session* interaction was significant ($\chi^2(3) = 2.42$, ns.; $\chi^2(15) = 15.11$, ns.).

The model explained a modest proportion of variance ($R^2_{\text{marginal}} = .17$, $R^2_{\text{conditional}} = .61$).

Perceived mental effort increased across sessions (Fig. 6a). Pairwise comparisons showed: 20:00 ≈ 22:00 (ns.); *both < 00:00 ($p < .05$); 00:00 < 03:00, 05:00, 07:00 (all $p < .001$)*; and late-session comparisons (03:00, 05:00, 07:00) were ns. Cue pairwise contrasts were ns. Table 2a summarizes Post-hoc EMMs with 95% CIs. Perceived mental effort *rose across the night and plateaued toward morning*, with *no evidence of cue-related differences*.

7.3.3 Simulator Sickness Questionnaire (SSQ)

A Type II Wald χ^2 test revealed significant effects of *condition*, $\chi^2(4) = 33.30$, $p < .001$, and *session*, $\chi^2(5) = 549.35$, $p < .001$; the *condition* × *session* $\chi^2(20) = 13.75$, interaction was ns.

The model explained a substantial proportion of variance ($R^2_{\text{marginal}} = .25$, $R^2_{\text{conditional}} = .72$).

Post-hoc contrasts showed *pre-task* < each task condition (all $p < .001$) with no differences among task conditions (ns.) (Fig. 6b).

Authorized licensed use limited to: Technische Universitaet Graz. Downloaded on April 08,2026 at 16:31:11 UTC from IEEE Xplore. Restrictions apply.

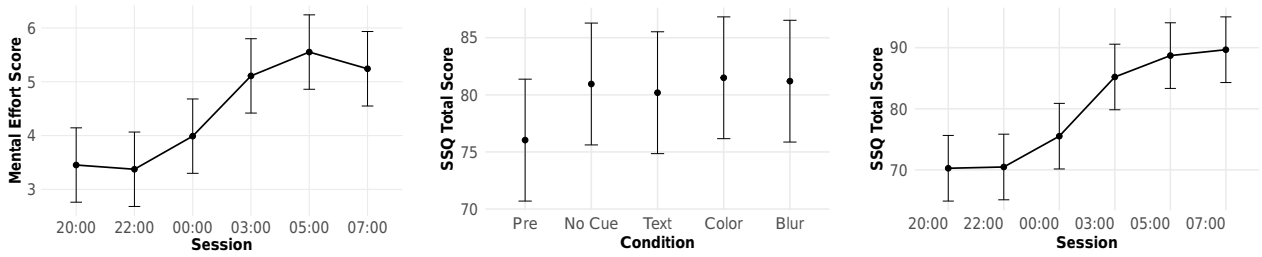


Figure 6: Estimated marginal means (EMMs; 95% CIs) from LMMs for subjective state measures. Left: mental effort (Paas, 1–9) by session (collapsed across conditions). Middle: SSQ Total by condition (collapsed across sessions). Right: SSQ Total by session (collapsed across conditions).

SSQ increased from *20:00 / 22:00 to 00:00* ($p < .001$) and from *00:00 to 03:00 / 05:00 / 07:00* (all $p < .001$), with 05:00 $\approx$ 07:00 (ns.) (Fig. 6c) Full EMMs and 95% CIs are in Table 2b. Overall, cybersickness increased with time-of-night; the *pre-task baseline was lower than all task conditions* (None, Text, Color, Blur), with no differences among the task conditions.

Table 2: Simulator Sickness Questionnaire (SSQ) Total Score and mental effort (Paas, 1–9; higher scores = greater effort): estimated marginal means (EMMs) with 95% CIs from LMMs. Top panel: EMMs by session (collapsed across conditions). Bottom panel: EMMs by condition (collapsed across sessions; SSQ includes a pre-task baseline).

(a) By session

Session	SSQ (EMM [95% CI])	Mental effort (EMM [95% CI])
20:00	70.3 [64.9, 75.6]	3.45 [2.76, 4.14]
22:00	70.5 [65.1, 75.8]	3.37 [2.68, 4.07]
00:00	75.5 [70.1, 80.9]	3.99 [3.30, 4.68]
03:00	85.2 [79.8, 90.6]	5.11 [4.42, 5.80]
05:00	88.7 [83.3, 94.1]	5.55 [4.86, 6.24]
07:00	89.7 [84.3, 95.0]	5.24 [4.55, 5.93]

(b) By condition

Condition	SSQ (EMM [95% CI])	Mental effort (EMM [95% CI])
Pre	76.0 [70.7, 81.4]	—
None	80.9 [75.6, 86.3]	4.32 [3.65, 4.99]
Text	80.2 [74.9, 85.5]	4.56 [3.88, 5.23]
Color	81.5 [76.2, 86.8]	4.41 [3.74, 5.09]
Blur	81.2 [75.9, 86.5]	4.52 [3.85, 5.19]

8 Discussion

8.1 Fatigue and circadian timing drive depth loss

Fatigue from extended wakefulness interacts with circadian phase and can degrade visual perception and visuomotor control; performance is poorest when high homeostatic sleep pressure coincides with the biological night [17, 24, 56]. This pattern is consistent with the two-process framework and wake-state instability accounts, which predict slower responses and intermittent lapses as time awake increases and the circadian trough around 02:00–06:00 is reached [19]. Sleep deprivation can also produce vigilance decrement and attentional state instability, with transient lapses and higher response variability rather than only uniform slowing [13, 4, 19, 5]. Our overnight pattern, PVT slowing after midnight with a circadian trough near 05:00, rising KSS, and greater response variability tracks these predictions. This pattern is evident in the MRRT decline and KSS increase across sessions, with the circadian trough near 05:00 (Fig. 5b).

Even a single night without sleep disrupts binocular coordination and smooth pursuit [69, 71], degrading the fidelity of disparity/convergence and motion signals needed for precise egocentric distance judgments. In HMD-VR, depth is already constrained by disparity sampling, field of view (FOV), and the vergence–accommodation conflict; in the 2–30 m "action space" users commonly show distance compression and elevated variance [15, 39, 52, 57, 64]. It follows that fatigue should most harm far-distance judgments late at night—precisely where we observed the largest errors and where augmentation helped most.

8.2 Augmented cues improve accuracy, while subtle cueing better preserves responsiveness

All augmented cues reduced depth error relative to no cue, with the largest gains at 7 m and in late sessions. Text produced the strongest and most consistent accuracy benefits across distances and times, but around 05:00 it showed a dip in response probability, whereas blur and color maintained higher response probability at 05:00 (blur $>$ color $>$ text; Fig. 3). One explanation is vigilance decrement under sleep loss, where performance failures often appear as intermittent lapses and attentional instability rather than only slower responses[13, 4, 23, 5]. In this task, responding at the correct depth plane still requires sustained attention and temporal prediction because the target keeps moving. When attentional stability drops near the circadian trough, cues that require active decoding can become less reliable for responding even if they remain accurate on trials where a response is made. This can explain why explicit cues show strong accuracy on responded trials while response probability drops under high fatigue. Text requires reading and verification, coded color requires applying a learned mapping, and blur changes image structure in a depth-contingent way that can be used with less symbolic decoding [28, 8].

This trade-off aligns with evidence that head-mounted displays can elevate cognitive load and that reading dynamic numerals recruits controlled attention, costly when sleepy [8, 10]. Blur retains responsiveness while improving accuracy under fatigue, consistent with its low parsing cost and prior reports of comfort benefits via foveated/depth-of-field blur [28, 35]. We interpret these cue effects as differences in behavioral performance under fatigue rather than evidence about cue integration weights. The takeaway is that augmentation cues robustness varies under fatigue, leading to a trade off between responsiveness and accuracy on responded trials.

8.3 Fatigue amplifies underestimation at far distances and increases variability

Egocentric distance in HMD-VR is typically underestimated in the 2–30 m action space [16], often 70–80% of physical distance [39, 57, 64], due to limited and conflicting depth cues and optical constraints [15, 52]. Since, the primary analyses used Δ^2, we interpret results as error magnitude rather than directional bias.

Authorized licensed use limited to: Technische Universitaet Graz. Downloaded on April 08,2026 at 16:31:11 UTC from IEEE Xplore. Restrictions apply.

Given the literature and our finding that fatigue disproportionately harms far-distance estimation, a plausible interpretation is that sleep loss amplified underestimation at 7 m and increased variance late at night [80]. Future analyses should report signed Δ alongside Δ^2 to test whether cues primarily reduce compression (bias) or variance. The disproportionate degradation at 7 m late in the night is consistent with this account (Fig. 4).

8.4 Cybersickness did not drive performance changes

SSQ increased across the night and then stabilized, with no differences between cue conditions and no condition $\times$ session interaction. No reliable association between SSQ and depth performance (error or lapses) was observed, suggesting cue effects and overnight degradation primarily reflect fatigue rather than cybersickness [40]. This interpretation is reinforced by the absence of cue-specific SSQ differences despite distinct performance profiles. The SSQ plateau by 07:00 contrasts with the response probability dip around 05:00 (Figs. 3, 6c).

8.5 Adaptive cueing should default to subtle support and escalate explicit overlays on demand

Cues improved how effectively participants performed at a given fatigue level, but did not alter the level of fatigue. KSS rose and PVT slowed similarly across cue conditions, consistent with evidence that neuro-behavioral impairment accumulates with additional wakefulness and is not mitigated by interface changes [72]. Augmentation therefore shapes performance expression at a given fatigue level rather than altering fatigue itself.

A pragmatic policy defaults to subtle support, to preserve responsiveness and comfort at low alertness, transiently switching explicit overlays when precise metric information is needed (decision points, on demand, or under gaze dwell). This mixed strategy mirrors findings from fatigue-prone AR tasks, where scheduling predictive overlays around critical events stabilized performance near the circadian trough without chronically increasing load [7, 73]. If available, gaze-contingent stereo/focus tuning and display-level mitigation of the vergence–accommodation conflict (varifocal/light-field) can further reduce oculomotor strain and improve depth fidelity [33, 44, 45, 61, 70].

This policy reflects the responsiveness-accuracy trade-off observed on responded trials (Fig. 3, Fig. 4). It also aligns with physiological computing perspectives where interface parameters adapt based on operator state [21, 20]. In deployed systems, lapses or sustained drops in responsiveness could trigger reduced use of decoding-heavy overlays and selective escalation of explicit metric support at decision-critical moments. Future work can test whether adding state proxies such as KSS or online lapse markers improves prediction of within-block instability beyond time-of-night alone.

8.6 Reflection of research questions

RQ1 (with H1). Fatigue reduced both responsiveness (response probability) and accuracy (squared depth error), with a reduction dependent on cue type. Consistent with this, the condition $\times$ session interaction showed that fatigue reduced responsiveness and increased squared error overall, with blur exhibiting less degradation in responsiveness than text at peak fatigue (Figs. 3 and 4). Subjective state followed the expected overnight trajectory: KSS rose and then plateaued, effort increased after 03:00, and SSQ totals rose before stabilizing (Figs. 2, 5b, 6a and 6c). Thus, *H1 is partially supported*: blur better preserves responsiveness at peak fatigue, whereas text retains the strongest accuracy on responded trials.

RQ2 (with H2). All augmented cues retained accuracy benefits over baseline across the night and across distances. As predicted in *H2*, text, color, and blur outperformed no cue on accuracy across sessions (Fig. 4). Relative advantages shifted with task difficulty and time of night: text produced the largest accuracy gains, especially for far targets late in the night, while blur and color maintained moderate gains and higher responsiveness during the circadian trough (Figs. 3 and 4). *H2 is supported*.

RQ3 (with H3). Explicit overlays did not produce higher subjective burden than blur during prolonged use under fatigue. Self-reports (KSS, Paas effort, SSQ) did not differ by cue (Figs. 5a, 6a, 6b). Thus, *H3 is not supported*. Cue-by-fatigue differences mainly appeared behaviorally (accuracy and responsiveness trade-offs) rather than in self-report measures.

9 Limitations

Our findings are best interpreted within the scope of the task and measurement choices in this overnight protocol. The asteroid task is a dynamic depth judgment task, so response error reflects a combination of depth relevant perception, temporal prediction, and motor execution. For this reason, we frame the results as depth relevant perception and action performance under fatigue rather than as isolated cue integration.

Eye movements were not recorded, so physiological links to binocular coordination remain inferential. The scene was also intentionally minimal and trajectories were linear to isolate cue effects, whereas operational environments include clutter and competing signals that can change cue use and cognitive load.

Repeated exposure across six sessions may have introduced practice effects or strategy drift. Since fatigue progresses with time of night in a sleep-deprivation design, fatigue and session order cannot be fully separated. We therefore interpret cue differences primarily through cue by session patterns and recommend future multi-day counterbalancing or additional baseline sessions.

Depth perception work often reports signed error and variability as complementary outcomes. Given that the primary analysis uses Δ^2, we interpret results as error magnitude, and plan to report signed and proportional metrics in follow-up work. Responses were discrete trigger presses, so depth at response reflects decision time, motor actuation, and system input latency. Although this motor component is constant across cue conditions, future studies should further separate perceptual and motor contributions using eye or motion tracking or alternative response modalities.

10 Conclusion

This study examined how sleep deprivation shapes the usefulness of explicit and subtle depth cues in VR across a 27-h overnight protocol. Across sessions and distances, all augmented cues improved depth performance relative to the no-cue baseline. *Text* provided the strongest accuracy gains on trials with a response, with the largest benefits for far targets late in the night. However, response probability decreased around the circadian trough, indicating that the most accurate cue was not always the most reliable for sustaining responses under fatigue. *Blur* and *color* maintained higher responsiveness at the same time while still improving accuracy.

These results show that fatigue can change how cue benefits appear in practice, and that cue evaluation should consider both accuracy on responded trials and the likelihood of responding at all. From a design perspective, subtle support such as depth dependent blur can serve as a robust default during extended use, while explicit overlays can be introduced selectively when precise metric information is needed. Future work can evaluate whether these patterns generalize to more complex scenes and operational tasks, and can use this evidence to guide state-aware cue adaptation.

11 Acknowledgements

The Know Center Research GmbH is funded within the Austrian COMET Program, supported by the BMVIT, BMWFJ, and the State of Styria, and managed by the FFG.

Authorized licensed use limited to: Technische Universitaet Graz. Downloaded on April 08,2026 at 16:31:11 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] T. Åkerstedt and M. Gillberg. Subjective and objective sleepiness in the active individual. *International journal of neuroscience*, 52(1-2):29–37, 1990. 5

[2] C. Armbrüster, M. Wolter, T. Kuhlen, W. Spijkers, and B. Fimm. Depth perception in virtual reality: distance estimations in peri-and extrapersonal space. *Cyberpsychology & Behavior*, 11(1):9–15, 2008. 2

[3] O. Ashtiani, H.-J. Guo, and B. Prabhakaran. Impact of motion cues, color, and luminance on depth perception in optical see-through ar displays. *Frontiers in Virtual Reality*, 4:1243956, 2023. 2

[4] S. Banks and D. F. Dinges. Behavioral and physiological consequences of sleep restriction. *Journal of clinical sleep medicine*, 3(5):519–528, 2007. 1, 8

[5] M. Basner, T. M. Moore, J. Nasrini, R. C. Gur, and D. F. Dinges. Response speed measurements on the psychomotor vigilance test: how precise is precise enough? *Sleep*, 44(1):zsaa121, 2021. 1, 3, 8

[6] D. Bates, M. Mächler, B. Bolker, and S. Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015. doi: 10.18637/jss.v067.i01 5

[7] J. Baumeister, J. Dorrian, S. Banks, A. Chatburn, R. T. Smith, M. A. Carskadon, K. Lushington, and B. H. Thomas. Augmented reality as a countermeasure for sleep deprivation. *IEEE transactions on visualization and computer graphics*, 22(4):1396–1405, 2016. 9

[8] J. Baumeister, S. Y. Ssin, N. A. ElSayed, J. Dorrian, D. P. Webb, J. A. Walsh, T. M. Simon, A. Irlitti, R. T. Smith, M. Kohler, et al. Cognitive cost of using augmented reality displays. *IEEE transactions on visualization and computer graphics*, 23(11):2378–2388, 2017. 2, 3, 8

[9] R. J. Bootsma and C. M. Craig. Information used in detecting upcoming collision. *Perception*, 32(5):525–544, 2003. 1

[10] P. Breves and J.-P. Stein. Cognitive load in immersive media settings: the role of spatial presence and cybersickness. *Virtual Reality*, 27(2):1077–1089, 2023. 1, 8

[11] M. E. Brooks, K. Kristensen, K. J. Van Benthem, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Mächler, and B. M. Bolker. glmmtmb balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. 2017. 5

[12] Y. Chen and Z. Wu. A review on ergonomics evaluations of virtual reality. *Work*, 74(3):831–841, 2023. 1, 2

[13] E. C.-P. Chua, S.-C. Yeo, I. T.-G. Lee, L.-C. Tan, P. Lau, S. Cai, X. Zhang, K. Puvanendran, and J. J. Gooley. Sustained attention performance during sleep deprivation associates with instability in behavior and physiologic measures at baseline. *Sleep*, 37(1):27–39, 2014. 1, 4, 8

[14] F. A. Conners. Attentional control and the simple view of reading. *Reading and Writing*, 22:591–613, 2009. 1

[15] S. H. Creem-Regehr, J. K. Stefanucci, and B. Bodenheimer. Perceiving distance in virtual reality: theoretical insights from contemporary technologies. *Philosophical Transactions of the Royal Society B*, 378(1869):20210456, 2023. 2, 8

[16] J. E. Cutting and P. M. Vishton. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *Perception of space and motion*, pp. 69–117. Elsevier, 1995. 2, 8

[17] D.-J. Dijk and S. N. Archer. Circadian and homeostatic regulation of human sleep and cognitive performance and its modulation by period3. *Sleep medicine clinics*, 4(2):111–125, 2009. 8

[18] D. F. Dinges and J. W. Powell. Microcomputer analyses of performance on a portable, simple visual rt task during sustained operations. *Behavior research methods, instruments, & computers*, 17(6):652–655, 1985. 5

[19] S. M. Doran, H. Van Dongen, and D. F. Dinges. Sustained attention performance during sleep deprivation: evidence of state instability. *Archives italiennes de biologie*, 139(3):253–267, 2001. 8

[20] S. Fairclough. A closed-loop perspective on symbiotic human-computer interaction. In *International Workshop on Symbiotic Interaction*, pp. 57–67. Springer, 2015. 3, 9

[21] S. H. Fairclough. Fundamentals of physiological computing. *Interacting with Computers*, 21(1-2):133–145, 2009. 3, 9

[22] L. Fan, J. Wang, Q. Li, Z. Song, J. Dong, F. Bao, and X. Wang. Eye movement characteristics and visual fatigue assessment of virtual reality games with different interaction modes. *Frontiers in neuroscience*, 17:1173127, 2023. 1

[23] A. García, J. Del Angel, J. Borrani, C. Ramirez, and P. Valdez. Sleep deprivation effects on basic cognitive processes: which components of attention, working memory, and executive functions are more susceptible to the lack of sleep? *Sleep Science*, 14(2):107, 2021. 2, 3, 8

[24] N. Goel, M. Basner, H. Rao, and D. F. Dinges. Circadian rhythms, sleep deprivation, and human performance. *Progress in molecular biology and translational science*, 119:155–190, 2013. 8

[25] C. R. Guibal and B. Dresp. Interaction of color and geometric cues in depth perception: When does "red" mean "near"? *Psychological Research*, 69(1):30–40, 2004. 1, 2

[26] M. Guo, H. Gao, S. Yang, K. Yue, Y. Liu, and Y. Wang. Evaluation of stereoscopic visual fatigue in virtual reality with exploration of brain dynamics. *Displays*, 87:102898, 2025. 2, 3

[27] M. Guo, K. Yue, H. Hu, K. Lu, Y. Han, S. Chen, and Y. Liu. Neural research on depth perception and stereoscopic visual fatigue in virtual reality. *Brain Sciences*, 12(9):1231, 2022. 2, 3

[28] R. T. Held, E. A. Cooper, and M. S. Banks. Blur and disparity are complementary cues to depth. *Current biology*, 22(5):426–431, 2012. 1, 2, 8

[29] W. S. Helton and P. N. Russell. Rest is best: The role of rest and task interruptions on vigilance. *Cognition*, 134:165–173, 2015. 2, 3

[30] N. Henze, M. Funk, and A. S. Shirazi. Software-reduced touchscreen latency. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 434–441, 2016. 4

[31] N. Henze, S. Mayer, H. V. Le, and V. Schwind. Improving software-reduced touchscreen latency. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 1–8, 2017. 4

[32] K. Hoedlmoser, H. Griessenberger, R. Fellinger, R. Freunberger, W. Klimesch, W. Gruber, and M. Schabus. Event-related activity and phase locking during a psychomotor vigilance task over the course of sleep deprivation. *Journal of sleep research*, 20(3):377–385, 2011. 4

[33] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, 8(3):33–33, 2008. 2, 9

[34] J. Hurter, C. Maraj, and S. Murphy. Commercial virtual reality displays: Issues of performance and simulator sickness from exocentric depth-perception tasks. *Displays*, 70:102095, 2021. 2

[35] R. Hussain, M. Chessa, and F. Solari. Mitigating cybersickness in virtual reality systems through foveated depth-of-field blur. *Sensors*, 21(12):4006, 2021. doi: 10.3390/s21124006 1, 3, 8

[36] R. Hussain, M. Chessa, and F. Solari. Improving depth perception in immersive media devices by addressing vergence-accommodation conflict. *IEEE Transactions on Visualization and Computer Graphics*, 2023. 2

[37] K. Kaida, M. Takahashi, T. Åkerstedt, A. Nakata, Y. Otsuka, T. Haratani, and K. Fukasawa. Validation of the karolinska sleepiness scale against performance and eeg variables. *Clinical neurophysiology*, 117(7):1574–1581, 2006. 5

[38] K. C. Kayser, V. A. Puig, and J. R. Estepp. Predicting and mitigating fatigue effects due to sleep deprivation: A review. *Frontiers in Neuroscience*, 16:930280, 2022. 2

[39] J. W. Kelly. Distance perception in virtual reality: A meta-analysis of the effect of head-mounted display characteristics. *IEEE transactions on visualization and computer graphics*, 2022. 2, 8

[40] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993. 5, 9

[41] J. M. Knapp and J. M. Loomis. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators & Virtual Environments*, 13(5):572–577, 2004. 2

Authorized licensed use limited to: Technische Universitaet Graz. Downloaded on April 08,2026 at 16:31:11 UTC from IEEE Xplore. Restrictions apply.

[42] K. Kohm, S. V. Babu, C. Pagano, and A. Robb. Objects may be farther than they appear: depth compression diminishes over time with repeated calibration in virtual reality. *IEEE transactions on visualization and computer graphics*, 28(11):3907–3916, 2022. 2

[43] R. Konrad, A. Angelopoulos, and G. Wetzstein. Gaze-contingent ocular parallax rendering for virtual reality. *ACM Transactions on Graphics (TOG)*, 39(2):1–12, 2020. 2

[44] R. Konrad, N. Padmanaban, K. Molner, E. A. Cooper, and G. Wetzstein. Accommodation-invariant computational near-eye displays. *ACM Transactions on Graphics (TOG)*, 36(4):1–12, 2017. 2, 9

[45] B. Krajancich, P. Kellnhofer, and G. Wetzstein. Optimizing depth perception in virtual and augmented reality through gaze-contingent stereo rendering. *ACM transactions on graphics (TOG)*, 39(6):1–10, 2020. 1, 2, 3, 9

[46] G. Kramida. Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE transactions on visualization and computer graphics*, 22(7):1912–1931, 2015. 2

[47] A. J. Krause, E. B. Simon, B. A. Mander, S. M. Greer, J. M. Saletin, A. N. Goldstein-Piekarski, and M. P. Walker. The sleep-deprived human brain. *Nature Reviews Neuroscience*, 18(7):404–418, 2017. 1, 2

[48] K. Kunasegaran, A. M. H. Ismail, S. Ramasamy, J. V. Gnanou, B. A. Caszo, and P. L. Chen. Understanding mental fatigue and its detection: a comparative analysis of assessments and tools. *PeerJ*, 11:e15744, 2023. 2

[49] Z. Li, Y. Cui, T. Zhou, Y. Jiang, Y. Wang, Y. Yan, M. Nebeling, and Y. Shi. Color-to-depth mappings as depth cues in virtual reality. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, pp. 1–14, 2022. 2

[50] J. Lim and D. F. Dinges. Sleep deprivation and vigilant attention. *Annals of the new York Academy of Sciences*, 1129(1):305–322, 2008. 2, 3, 6

[51] M. A. Livingston, A. Dey, C. Sandor, and B. H. Thomas. *Pursuit of "X-Ray Vision" for Augmented Reality*, pp. 67–107. Springer New York, New York, NY, 2013. doi: 10.1007/978-1-4614-4205-9_4 2

[52] J. M. Loomis, J. M. Knapp, et al. Visual perception of egocentric distance in real and virtual environments. *Virtual and adaptive environments*, 11:21–46, 2003. 2, 8

[53] C. J. Lowe, A. Safati, and P. A. Hall. The neurocognitive consequences of sleep restriction: a meta-analytic review. *Neuroscience & Biobehavioral Reviews*, 80:586–604, 2017. 2, 3

[54] J. Lowrie and H. Brownlow. The impact of sleep deprivation and alcohol on driving: a comparative study. *BMC public health*, 20(1):980, 2020. 2

[55] I. S. MacKenzie. Motor behaviour models for human-computer interaction. *HCI models, theories, and frameworks: Toward a multidisciplinary science*, pp. 27–54, 2003. 4

[56] M. Maire, C. F. Reichert, V. Gabel, A. U. Viola, C. Phillips, C. Berthomier, S. Borgwardt, C. Cajochen, and C. Schmidt. Human brain patterns underlying vigilant attention: impact of sleep debt, circadian phase and attentional engagement. *Scientific reports*, 8(1):970, 2018. 8

[57] R. Messing and F. H. Durgin. Distance perception and the visual horizon in head-mounted displays. *ACM Transactions on Applied Perception (TAP)*, 2(3):234–250, 2005. 8

[58] A. Murgia and P. M. Sharkey. Estimation of distances in virtual environments using size constancy. *International Journal of Virtual Reality*, 8(1):67–74, 2009. 2

[59] P. E. Napieralski, B. M. Altenhoff, J. W. Bertrand, L. O. Long, S. V. Babu, C. C. Pagano, J. Kern, and T. A. Davis. Near-field distance perception in real and virtual environments using both verbal and action responses. *ACM Transactions on Applied Perception (TAP)*, 8(3):1–19, 2011. 2

[60] F. G. Paas. Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach. *Journal of educational psychology*, 84(4):429, 1992. 5

[61] N. Padmanaban, R. Konrad, T. Stramer, E. A. Cooper, and G. Wetzstein. Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays. *Proceedings of the National Academy of Sciences*, 114(9):2183–2188, 2017. 2, 9

[62] Z. Peng, C. Dai, X. Cai, L. Zeng, J. Li, S. Xie, H. Wang, T. Yang, Y. Shao, and Y. Wang. Total sleep deprivation impairs lateralization of spatial working memory in young men. *Frontiers in neuroscience*, 14:562035, 2020. 2

[63] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2025. 5

[64] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert. The perception of egocentric distances in virtual environments-a review. *ACM Computing Surveys (CSUR)*, 46(2):1–40, 2013. 8

[65] J. Roca, L. J. Fuentes, A. Marotta, M.-F. López-Ramón, C. Castro, J. Lupiáñez, and D. Martella. The effects of sleep deprivation on the attentional functions and vigilance. *Acta psychologica*, 140(2):164–176, 2012. 1, 2

[66] M. Sabat, B. Haładus, M. Klincewicz, and G. J. Nalepa. Cognitive load, fatigue and aversive simulator symptoms but not manipulated zeitgebers affect duration perception in virtual reality. *Scientific reports*, 12(1):15689, 2022. 1

[67] A. D. Souchet, D. Lourdeaux, J.-M. Burkhardt, and P. A. Hancock. Design guidelines for limiting and eliminating virtual reality-induced symptoms and effects at work: a comprehensive, factor-oriented review. *Frontiers in psychology*, 14:1161932, 2023. 1

[68] A. D. Souchet, D. Lourdeaux, A. Pagani, and L. Rebenitsch. A narrative review of immersive virtual reality's ergonomics and risks at the workplace: cybersickness, visual fatigue, muscular fatigue, acute stress, and mental overload. *Virtual Reality*, 27(1):19–50, 2023. 1, 2

[69] L. S. Stone, T. L. Tyson, P. F. Cravalho, N. H. Feick, and E. E. Flynn-Evans. Distinct pattern of oculomotor impairment associated with acute sleep loss and circadian misalignment. *The Journal of physiology*, 597(17):4643–4660, 2019. 1, 2, 8

[70] B. Terzioğlu, U. Celikcan, and T. K. Capin. Gaze-contingent adaptation of vr stereo parameters for cybersickness prevention. *The Visual Computer*, 40(7):5017–5028, 2024. 2, 9

[71] J. Tong, J. Maruta, K. J. Heaton, A. L. Maule, U. Rajashekar, L. A. Spielman, and J. Ghajar. Degradation of binocular coordination during sleep deprivation. *Frontiers in neurology*, 7:90, 2016. 1, 2, 8

[72] H. P. Van Dongen, G. Maislin, J. M. Mullington, and D. F. Dinges. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep*, 26(2):117–126, 2003. 9

[73] B. Volmer, J. Baumeister, R. Matthews, L. Grosser, S. Von Itzstein, S. Banks, and B. H. Thomas. A comparison of spatial augmented reality predictive cues and their effects on sleep deprived users. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 589–598. IEEE, 2022. 9

[74] L. R. Wanger, J. A. Ferwerda, D. P. Greenberg, et al. Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics and Applications*, 12(3):44–58, 1992. 2

[75] P. Willemsen, A. A. Gooch, W. B. Thompson, and S. H. Creem-Regehr. Effects of stereo viewing conditions on distance perception in virtual environments. *Presence: Teleoperators and Virtual Environments*, 17(1):91–101, 2008. 2

[76] C. J. Wilson and A. Soranzo. The use of virtual reality in psychology: A case study in visual perception. *Computational and mathematical methods in medicine*, 2015(1):151702, 2015. 3

[77] K. Yue, M. Guo, Y. Liu, H. Hu, K. Lu, S. Chen, and D. Wang. Investigate the neuro mechanisms of stereoscopic visual fatigue. *IEEE Journal of Biomedical and Health Informatics*, 26(7):2963–2973, 2022. 2, 3

[78] A. Zaman, E. Kruijff, E. Veas, A. Krajnc, and N. ElSayed. Subtle cueing for improving depth perception in virtual reality. In *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 730–739. IEEE, 2024. 1, 3, 4

[79] L. H. Zeuwts, E. Iliano, M. Smith, F. Deconinck, and M. Lenoir. Mental fatigue delays visual search behaviour in young cyclists when negotiating complex traffic situations: A study in virtual reality. *Accident Analysis & Prevention*, 161:106387, 2021. 1, 2

[80] J. Zhang, X. Yang, Z. Jin, and L. Li. Distance estimation in virtual reality is affected by both the virtual and the real-world environments. *i-Perception*, 12(3):20416695211023956, 2021. 9

Authorized licensed use limited to: Technische Universitaet Graz. Downloaded on April 08,2026 at 16:31:11 UTC from IEEE Xplore. Restrictions apply.