IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.0429000

Effects of Ambient Illumination and Screen Luminance on Mixed Reality Interaction

RUILIN LIU, TINGHUI LI, AND ZHANNA SARSENBAYEVA
School of Computer Science, The University of Sydney, NSW 2006, Australia
Corresponding author: Ruilin Liu (e-mail: rliu0288@uni.sydney.edu.au).

ABSTRACT As mixed reality (MR) is increasingly adopted in diverse real-world contexts, its usability becomes susceptible to contextual factors known as situational impairments. Among these, ambient illuminance is particularly critical because it directly influences the visibility and legibility of virtual elements superimposed on the real world. To create MR systems that are robust to these conditions, a systematic understanding of the interplay between ambient illuminance and screen luminance is essential. This paper investigated three ambient illuminance levels (dark: 5 lx, medium: 500 lx, and bright: 40,000–72,000 lx) and two screen luminance levels (low, high) on MR interaction across three fundamental tasks (direct target selection, ray-cast target selection, and text entry). We evaluated the effects of ambient light and screen luminance on interaction performance (movement time, pointing offset, error rate, throughput) as well as on visual fatigue (blink rate and subjective questionnaire). Our results indicate that the effects of lighting are task-dependent. Direct selection remained robust across all conditions, whereas ray-casting showed increased movement time and pointing offset under both bright and dark ambience. Text entry was similarly impaired, with throughput dropping substantially in dark and bright conditions. Furthermore, both dark and bright contexts, especially in combination with high screen luminance, significantly increased visual fatigue. Our findings enhance the understanding of how ambient illuminance and screen luminance affect mixed reality interaction and contribute to the growing body of research on situational impairments in MR.

INDEX TERMS Ambient illuminance, Fitts's Law, mixed reality, screen luminance, situational impairments, text entry, visual fatigue.

I. INTRODUCTION

Mixed reality (MR) are rapidly transforming how people interact with digital content by blending virtual elements with the physical world. Unlike virtual reality (VR), which immerses users in entirely synthetic environments, MR systems leverage high-fidelity video pass-through to overlay interactive virtual content onto the user's real-world view [1]. This unique capability enables a wide applications in fields such as design [2], education [3], healthcare [4], and entertainment [3]. As MR experiences increasingly extend into everyday settings, they are becoming more susceptible to real-world contextual challenges. These factors are known as Situationally-Induced Impairments and Disabilities (SIIDs), or situational impairments [5], [6]. SIIDs are temporary, context-dependent challenges that degrade users' abilities to interact with technology [7], [8]. Extensive research has demonstrated that situational impairments, including cold environments [9], [10], ambient noise [11], [12], stress [13], mobility constraints [14], and physical encumbrance [15], [16], significantly impair effective user interaction with mobile devices [17], desktop computers [18], and smartwatches [19].

Although this body of work establishes the importance of SIIDs, their influence within the context of MR remains underexplored. Initial studies have begun to examine mobility-related situational impairments, such as user motion [20], [21] and encumbrance [20], [22]. However, a critical environmental factor, ambient illuminance, remains uninvestigated. Ambient light is known to affect the usability of digital devices, such as smartphones [23], tablets [24], and e-readers [25], by degrading interaction performance [26] and user comfort [25], [27]. In the context of MR, ambient light is particularly important, as for example in outdoor MR applications (e.g., navigation assistance, construction site visualisation), high levels of sunlight can wash out virtual overlays and reduce contrast [28], causing challenges in interaction. Conversely, in dim indoor environments (e.g., healthcare training, emergency response simulations), insufficient light can cause difficulties for users to simultaneously perceive physical surroundings and virtual cues [29]. Although users may attempt to mitigate the effects of ambient light by adjusting display luminance, the effectiveness of such adaptations and their combined impact on both performance and comfort in MR interaction remains unexplored [30].

In this paper, we investigated how the combination of ambient illuminance and screen luminance affects MR interaction performance and visual fatigue. We conducted the experiment under three ambient light conditions: dark (5 lx, dimly lit room), medium (500 lx, typical indoor office lighting), and bright (40,000-72,000 lx, outdoor sunlight), in combination with two screen luminance levels: low (50%) and high (100%). We examined three fundamental MR interactions: direct target selection (touching virtual objects), ray-cast target selection (pointing with a virtual ray), and text entry. Our evaluation combines established performance metrics (movement time, pointing offset, error rate, throughput) with both objective (blink rate) and subjective (visual fatigue questionnaire) measures of visual fatigue.

Our results indicate that ambient light significantly impacts MR interaction performance, but its effects vary substantially across different interaction tasks. We found that direct selection remained robust across all light conditions. In contrast, ray-cast selection was highly vulnerable to extreme lighting, showing longer movement times and larger pointing offsets in both bright sunlight and dark ambience. Text entry was also sensitive to lighting variations, with throughput dropping substantially in dark and bright conditions. Screen luminance had minimal impact on target selection but improved text entry throughput at higher levels, though it also increased the uncorrected error rates. Most critically, our visual fatigue analysis revealed that both extreme ambient light conditions and high screen brightness substantially increased eye strain, with the combination of dark ambient lighting and maximum screen brightness producing the most severe fatigue.

In summary, our study makes the following contributions:

- We quantify the effects of ambient illumination and screen luminance on MR interaction performance;
- We evaluate the effects of ambient illumination and screen luminance on visual fatigue;
- We expand our research community's understanding of situational impairments in MR and propose potential strategies to mitigate their effect in real-world use.

II. RELATED WORK

A. SITUATIONAL IMPAIRMENTS IN MR

Although the studies of SIIDs are well-documented for mobile and desktop computing, research into their impact on MR interaction is still in its early stages. For instance, Li et al. [20] investigated the impact of encumbrance and walking on performance in canonical MR interaction tasks. Their study showed that carrying a 1.0 kg weight increased selection movement time by 28% and decreased text entry throughput by 17%. Walking imposed even greater challenges, leading to a 63% increase in ray-cast movement time and a 51% reduction in text entry throughput. Complementing this, a study by Li et al. [21] examined how varying motion intensity affects target selection. By comparing user performance while standing, walking, running, and jumping, the authors found that as the intensity of movement increased, both selection time and precision decreased substantially.

B. AMBIENT LIGHT AS A SITUATIONAL IMPAIRMENT

Despite aforementioned contributions exploring mobility-related SIIDs in MR, a critical environmental factor, ambient light, remains underexplored. Ambient lighting conditions are known to play a critical role in visual perception, user comfort, and interaction quality. Research on smartphones, tablets, and e-readers has consistently shown that suboptimal lighting, both too dim and excessively bright, degrades performance by increasing task completion times and error rates [23], [26]. Beyond objective performance, lighting conditions profoundly impact user experience, with factors such as poor screen contrast and glare contributing to significant visual fatigue and discomfort [24], [25], [27].

This challenge is amplified in the context of head-mounted displays (HMDs), where the additive light model makes virtual content highly susceptible to being "washed out" under high ambient illuminance. Early work by Gabbard et al. [28] demonstrated that brighter lighting significantly degrades text legibility in augmented reality (AR) by reducing the contrast between virtual content and real-world backgrounds. More recent studies have confirmed that the visibility and contrast of virtual elements in AR decline sharply as ambient light increases, becoming almost unusable in bright conditions (e.g., 20,900 lx) [31]. Choi et al. [32] further reported that bright outdoor environments induce substantial losses in presence, usability, and immersion for AR displays. Although studies have investigated how ambient light affects content visibility and user experience with AR HMDs, the distinct challenges posed by MR have been overlooked. Unlike AR, which typically overlays information onto the users' view, MR enables users to directly and actively manipulate virtual objects within the real world [1]. This complex spatial interactivity introduces unique perceptual and usability challenges that have yet to be investigated.

A primary method for users to counteract the negative effects of ambient illuminance is by adjusting screen luminance (brightness). Most modern displays, including MR headsets, support this adjustment to accommodate varying environmental conditions [33]. However, the interaction between ambient light and display brightness is complex, and inappropriate settings may even degrade visual fatigue and performance rather than mitigate it [34], [35]. Research in AR shows that for typical indoor lighting, both excessively low and high screen brightness settings increased visual fatigue and degraded task performance [36]. Conversely, in dark environments, default HMD brightness is often too high, and lowering it can maintain visual search task performance while reducing eye strain [35]. Therefore, we speculate that different screen luminance levels may exacerbate or alleviate MR interaction across different ranges of ambient illuminance.

C. MR INTERACTION ASSESSMENT

To evaluate the effects of ambient light and screen brightness on MR interaction, we focused on three fundamental interaction techniques that represent the essential components of MR experiences: direct target selection, ray-cast target selection, and text entry [20], [37]. These tasks have been shown to demonstrate distinct performance characteristics and are influenced differently by varying ambient lighting conditions [26], walking, and encumbrance [20].

Target selection is a foundational interaction in graphical user interfaces [21], and its performance is commonly modelled by Fitts' Law [38]. Fitts' Law describes the trade-off between speed and accuracy in aimed movements [39] and has been widely used to evaluate pointing devices and interaction techniques [13], [20], [21]. In MR, there are two main techniques used to select a target, namely direct selection and ray-casting [40]. **Direct selection**, which simulates reaching out to physically touch virtual objects, is the primary method for near-field interaction within arm's reach [41]. The success of direct selection relies heavily on accurate depth perception and precise hand-eye coordination [42]. In contrast, **ray-casting** is the standard technique for far-field interaction, allowing users to select distant targets with less physical movement [43]. This is typically achieved by emitting a ray from the user's hand and performing a pinch gesture to confirm selection. Ray-casting offers advantages in selecting targets beyond the user's reach and is widely adopted for remote interaction. It demands sustained visual attention and fine motor control to maintain a steady hand and accurately align the ray [40]. Performance in target selection tasks is commonly measured through movement time (the time taken to select a target) [15], [16], pointing offset (offset size from target centre) [11], [26], and error rate (frequency of incorrect selections) [44].

Beyond target selection, **text entry** is another common and critical task on any computing platform, essential for productivity, communication, and search [45]. In our study, we employ a standard MR text entry method where users type on a virtual QWERTY keyboard via air-tapping [20]. It requires continuous visual attention, switching between individual keys and text output areas. Text entry performance is typically evaluated using uncorrected error rate (errors remaining in final text) [46], [47], and corrected error rate (errors identified and fixed during typing) [46], [47]. To account for the inherent speed-accuracy trade-off in typing, researchers often use a unified throughput metric that combines speed and error rates based on information theory, providing a more holistic measure of text entry efficiency [47].

In addition to objective performance measures, visual fatigue represents a critical subjective outcome that can impact user experience and the long-term adoption of MR systems. Visual fatigue, often referred to as Digital Eye Strain (DES), encompasses a range of symptoms including eye strain, dry eyes, blurred vision, and headaches that arise from extended use of digital displays [48]. Research indicates that DES can be more severe when using HMDs compared to conventional displays such as smartphones [49], due to the visual demands of stereoscopic imagery and the complex interplay between virtual content and real-world lighting conditions [50].

Visual fatigue can be assessed through both physiological signals and subjective scales. Eye blink is a well-known indicator for DES in general [51], [52]. A large body of studies suggests that blink patterns are highly sensitive to lighting conditions and screen brightness [53], [54]. For instance, in low ambient illumination, increasing screen luminance often leads to a decrease in blink frequency [25]. This reduction in blinking can impair the quality of the eye's tear film and stress the cornea, causing dry eye symptoms, which is a primary contributor to visual fatigue [50].

As for subjective measurements, one of the most commonly used questionnaires for assessing DES and general discomfort in HMDs is the Simulator Sickness Questionnaire (SSQ) [21], [55], [56]. Although the SSQ includes four oculomotor symptoms (headache, eye strain, difficulty focusing, and blurred vision), it was originally designed to assess simulator sickness related to motion-induced visuo-vestibular conflicts. Recent work by Hirzle et al. [56] highlights its limitations in capturing the broader spectrum of discomfort prevalent in modern HMDs, particularly symptoms of DES and ergonomic fatigue. Furthermore, their findings show that such symptoms are more common and severe than motion-related sickness in contemporary HMD use. Moreover, the tasks in our study do not involve user motion, head-turning, or virtual object movement. Thus, we adopted the Visual Fatigue Questionnaire (VFQ) developed by Hayes et al. [57]. This questionnaire has been widely used in the subjective assessment of visual fatigue [48], [58]. The VFQ consists of ten items covering multiple dimensions of DES, including accommodation difficulties (e.g., refocusing slowness, distance blur), ocular surface symptoms (e.g., dryness, irritation), and general visual discomfort (e.g., eye strain, photosensitivity).

III. METHOD

Our study followed a 3 × 2 within-subjects design (three ambient illuminance levels × two screen brightness settings). This design controls for individual differences and systematic biases [59], as each participant experienced all six experimental conditions. To minimise the impact of confounding factors such as fatigue and learning effects, the order of conditions was counterbalanced using a Balanced Latin Square [60], which effectively reduces sequential effects that could influence performance measurements [26]. The experiment was approved by the University of Sydney's Human Research Ethics Committee (Application identifier: 2025/HE000045).

We conducted our experiment using a Meta Quest Pro headset (1800 × 1920 px), running a custom-built application developed in Unity (2022.3.12f1). The Meta Quest Pro was chosen for two primary reasons: its high-resolution colour pass-through, which is essential for creating a convincing mixed reality experience, and its integrated eye-tracking capabilities [61]. The Quest Pro's eye-tracker has demonstrated accuracy and precision comparable to leading devices such

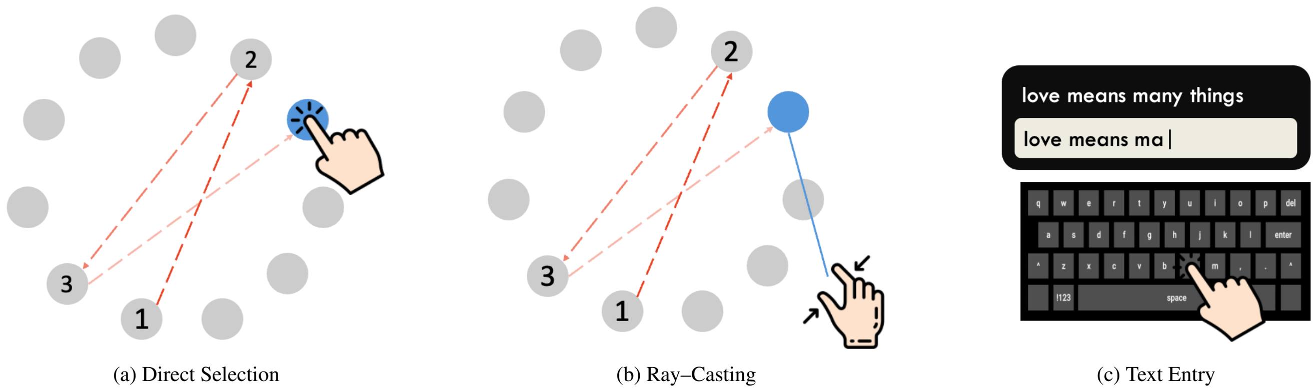


FIGURE 1. **The visual representation of the tasks used in the experiment. In the target selection tasks (a) and (b), participants followed the red dashed arrows to sequentially select the targets along the indicated path.**

as HTC Vive Pro Eye and HoloLens 2 [62], making it well-suited for precise blink rate measurement [63], [64].

A. TASKS

We investigated the effect of ambient illuminance and screen brightness on performance through 3 common MR interactions: (1) direct target selection, (2) ray-cast target selection, and (3) text entry. These tasks were adapted from the prior study by Li et al. [20], which investigated the effects of walking and encumbrance on MR interaction. We extended them by integrating eye-blink data collection using the Meta Movement SDK [65]. This integration ran in the backend and did not alter the user-facing interaction. All three tasks were integrated into a single MR application and presented in a counterbalanced order to mitigate sequence effects [11].

1) Baseline Visual Fatigue Measurement

At the beginning of each experimental condition, participants completed a 3-minute baseline measurement to establish pre-task visual fatigue under the given ambient light and screen luminance. During this phase, participants wore the headset, which displayed only the real-world pass-through (no virtual content), and were asked to relax and observe a distance of more than 6 meters [66]. The first minute served as an adaptation period to adjust to the new lighting and brightness levels, with blink data collected only from the final two minutes [67]. To prevent participants from consciously altering their blink patterns, they were given a spurious reason (e.g., measuring facial micro-movements) and were unaware that their blink rate was being recorded until the post-study debriefing [68]. Immediately after the baseline, participants removed the headset and completed the Visual Fatigue Questionnaire (VFQ). This baseline measurement was essential for establishing individualised reference points for each participant's visual fatigue levels. It enabled more accurate comparisons of pre- and post-task fatigue under different experimental conditions by controlling for the baseline effects of simply wearing the headset without any interactive engagement.

2) Task 1: Direct Target Selection

For the first task, participants completed a selection by touching the target with the index finger of their dominant hand. We adopted the widely-used Fitts' Law [20], [37], presenting 11 circular targets arranged in a Fitts' ring layout, with a central start target. The ring was positioned 45 cm from the headset's central lens, following Meta developer guidelines [69]. Each trial began when the participant tapped the central start target. Subsequently, one of the peripheral targets was highlighted in blue to indicate the next selection (Figure 1a). This sequence continued until all 11 peripheral targets were selected.

Following standard practice [39], the initial movement from the centre to the first target was omitted from our analysis to ensure all recorded movements shared consistent amplitude and attack velocity. We manipulated the task's Index of Difficulty (ID) by varying target width (W) and the distance between targets (D) [21]. Specifically, we used two target diameters (4 cm and 6 cm) and two distances (3 cm and 5 cm), resulting in four combinations. Thus, participants had to complete all four combinations, each with 11 targets.

3) Task 2: Ray-Cast Target Selection

In the second task, participants interacted with targets using a ray-casting technique. Participants aimed a virtual ray, which originated from their dominant hand, and performed a pinch gesture to select targets (Figure 1b). The task structure was similar to the direct selection task, but the circle was positioned 200 cm away, followed by the Meta developer guidelines. Participants began by selecting a central target and then proceeded to select the peripheral targets as they were sequentially highlighted in blue. To accommodate the increased distance and maintain consistent visual sizes, we doubled the target widths (8 and 12 cm) and distances (6 and 10 cm), creating four IDs [21]. Similar to the first task, participants completed a full sequence of 11 target selections for each ID condition.

4) Task 3: Text Entry

In text entry, participants transcribed three different phrases for each experimental condition [20], [70]. The phrases were

(a) Dark (5 lx) (b) Medium (500 lx) (c) Bright (40,000 - 72,000 lx)

FIGURE 2. The ambient illuminance experimental conditions.

randomly selected from a standardised phrase set developed by MacKenzie and Soukoreff [71], ensuring consistent difficulty across trials. In each trial, a phrase was displayed, and participants typed it verbatim into a virtual text box (Figure 1c). Participants were permitted to submit phrases with errors [20], enabling us to measure both uncorrected and corrected error rates for throughput assessment [72].

B. INDEPENDENT VARIABLES

The two independent variables of this study were ambient illuminance and screen luminance of the headset. These variables were selected to examine their combined effects on MR interaction and visual fatigue, reflecting usage scenarios in real-world settings.

1) Ambient Illuminance

We used three levels of ambient illuminance: dark (5 lx), medium (500 lx), and bright (40,000 - 72,000 lx). The illuminance levels were verified using an LM-3000 handheld digital light meter, positioned flat on the task surface adjacent to the participant, following IESNA recommendations [73]. The dark and medium conditions were conducted in a controlled laboratory environment with shaded windows and consistent artificial overhead fluorescent lighting. The 5 lx level simulated a dimly lit room (Figure 2a), and the 500 lx level represented a typical indoor office environment (Figure 2b) [74]. For the bright condition, most prior research has relied on simulated lighting conditions, typically capping at a maximum illuminance of approximately 32,500 lx [23], [75]. However, direct outdoor sunlight commonly ranges between 50,000 and 100,000 lx, as noted by Halsted [76]. Therefore, we decided to capture the effects of real outdoor sunlight. This condition was conducted in a secluded courtyard adjacent to the laboratory building (Figure 2c), specifically chosen to be away from high-traffic areas, thereby avoiding potential distractions or interruptions, such as vehicles, pedestrians, or curious seekers [28]. To maintain experimental validity, we only performed the experiments on clear or partly cloudy days. Due to the natural variability of outdoor illuminance, we continuously monitored ambient light conditions by logging illuminance measurements every minute throughout each outdoor session [77]. Under this approach, the bright condition fell within a range of light measurements 40,000–72,000 lx (mean = 52,037 lx, SD = 7,695 lx).

2) Screen Luminance

Two levels of screen brightness were considered: low (50% brightness: 50 $\mathrm{cd/m^2}$) and high (100% brightness: 100 $\mathrm{cd/m^2}$) [78]. The choice of screen luminance values was informed by prior research on digital screen ergonomics and MR system design, which emphasises the importance of maintaining comfortable contrast ratios under varying ambient lighting [79].

C. DEPENDENT VARIABLES

We followed established research on SIIDs [11], [20] and employed multiple performance measures (dependent variables) tailored to each interaction type. Additionally, we captured visual fatigue through both objective (blink rate) and subjective (Visual Fatigue Questionnaire) indicators to assess the physiological impact of mixed reality interaction.

1) Target Selection Performance Measures

Movement Time (MT): The time (in milliseconds) between two consecutive successful target selections [38]. It is a critical measure of speed, linearly associated with the Index of Difficulty (ID) as described by Fitts' Law [39] as shown in Equation 1, where ID is a function of target width (W) and distance (D) between targets (Equation 2). Lower MT values indicate faster task performance.

$$MT = a + b \times ID \tag{1}$$

$$ID = \log_2\left(\frac{D}{W} + 1\right) \tag{2}$$

Pointing Offset: The Euclidean distance between the target's centre (x_0, y_0, z_0) and the actual point of selection

$(\hat{x}, \hat{y}, \hat{z})$ [15] as shown in Equation 3. This distance quantifies how accurately participants selected the intended target, with a larger pointing offset indicating poorer performance [13].

$$\text{Offset} = \sqrt{(\hat{x} - x_0)^2 + (\hat{y} - y_0)^2 + (\hat{z} - z_0)^2} \tag{3}$$

Error Rate: The percentage of incorrectly selected targets relative to the total number of targets [11] as shown in Equation 4. This measure captures the frequency of mistakes during target selection, including wrong selections and failed attempts to engage with virtual elements. Lower error rates reflect more effective interaction [15], [16].

$$E_{\text{target}} = \frac{\text{Incorrect Selections}}{\text{Total Targets}} \times 100\% \tag{4}$$

2) Text Entry Performance Measures

Throughput: A comprehensive measure of text entry performance accounting for speed-accuracy biases [70], [80]. These biases arise from the users' tendency to prioritise either speed or accuracy, making it difficult to compare fast, error-prone performance with slow, precise input [47]. To resolve this, Zhang et al. [47] proposed throughput, a unified metric grounded in Shannon's information theory [81]. It is computed as the amount of transmitted information, $I(X, Y)$, defined as $I(X, Y) = H(X)$ - $H_Y(X)$, where $H(X)$ is the source entropy of the phrase set, and $H_Y(X)$ is the conditional entropy considering for both UER and CER [47]. The speed is expressed as characters per second (CPS) [46]. Higher values denote a better performance [80].

$$\text{Throughput} = I(X, Y) \times CPS \tag{5}$$

Uncorrected Error Rate (UER) and Corrected Error Rate (CER): UER measures the proportion of characters in the final transcribed text that differ from the intended phrase [46]. In contrast, CER quantifies the percentage of mistyped characters that users successfully identify and rectify during the task (using a backspace) [46]. They are calculated in Equation 6 and 7, respectively, where C = Correct characters in the final transcribed text; INF = errors that remain in the final text; IF = errors that were made but corrected before the final text.

$$\text{UER} = \frac{\text{INF}}{\text{C} + \text{INF} + \text{IF}} \times 100\% \tag{6}$$

$$\text{CER} = \frac{\text{IF}}{\text{C} + \text{INF} + \text{IF}} \times 100\% \tag{7}$$

3) Visual Fatigue Measures

Blink Rate: An objective indicator of visual fatigue, measured in blinks per minute [82]. A reduced blink rate is linked to increased visual discomfort or strain [25]. Blink count was assessed during a passive baseline period and during the active MR tasks. By comparing baseline and task-period blink rates across conditions, we aim to quantify changes in ocular behaviour that reflect accumulating fatigue. All blink data were automatically collected using the headset's integrated eye-tracking, ensuring consistency and minimising participant awareness of the measurement to avoid bias. A lower blink rate indicates severe visual fatigue.

$$\text{Blink Rate} = \frac{\text{Total Blink Count}}{\text{Total Duration (s)}} \times 60 \tag{8}$$

Visual Fatigue Questionnaire (VFQ) Score: A subjective measure using a validated 10-item questionnaire from Hayes et al. [57]. Participants rated symptoms on a 10-point Likert scale (0: none, 10: extreme). The questionnaire was administered pre-task (after baseline) and post-task for each condition to measure the change in perceived fatigue. The final VFQ score is computed by averaging the ratings across all 10 items (Equation 9), with lower scores corresponding to less fatigue.

$$\text{VFQ Score} = \frac{1}{10} \sum_{i=1}^{10} S_i \tag{9}$$

D. PARTICIPANTS AND PROCEDURE

We recruited 24 participants (12 female, 12 male) through our university's mailing lists and assigned unique anonymous identifiers (P01 - P24). This sample size is consistent with established guidelines for within-subjects experimental designs in HCI research [83], [84]. Participants ranged in age from 20 to 29 years (mean = 23.7, SD = 2.5). All participants possessed normal or corrected-to-normal vision, with 9 participants having normal vision and 15 using corrective lenses (glasses or contact lenses). 13 out of 24 participants indicated they had little or no prior experience using MR headsets. 2 participants were left-handed, and their data were mirrored along the sagittal plane to align with right-handed coordinate systems, following established research [11], [20].

Upon arrival, participants received a comprehensive overview of the study's purpose and procedure. After providing informed consent, they completed a brief demographic questionnaire, collecting information on gender, age, dominant hand, height, weight, vision status, and prior experience with MR. Before the main experiment, all participants underwent a training session to familiarise themselves with the headset and all three experimental tasks. This training continued until participants felt comfortable and confident with each interaction type, helping to minimise potential learning effects and performance variations during the actual experiment [10], [85]. The text phrases used in training were excluded from the set used during the main study. Both training and actual experiment sessions were completed while seated. Participants were instructed to complete each task as quickly and accurately as possible [11], [21].

The main experiment consisted of six conditions. Each condition began with the baseline visual fatigue measurement, as described in Section III-A1, followed by completion of the VFQ to assess pre-task visual fatigue. Participants then completed three experimental tasks in a counterbalanced order. Upon completion of all three tasks, they filled out the VFQ again. A mandatory 5-minute rest period was enforced between conditions to alleviate fatigue, with extensions avail-

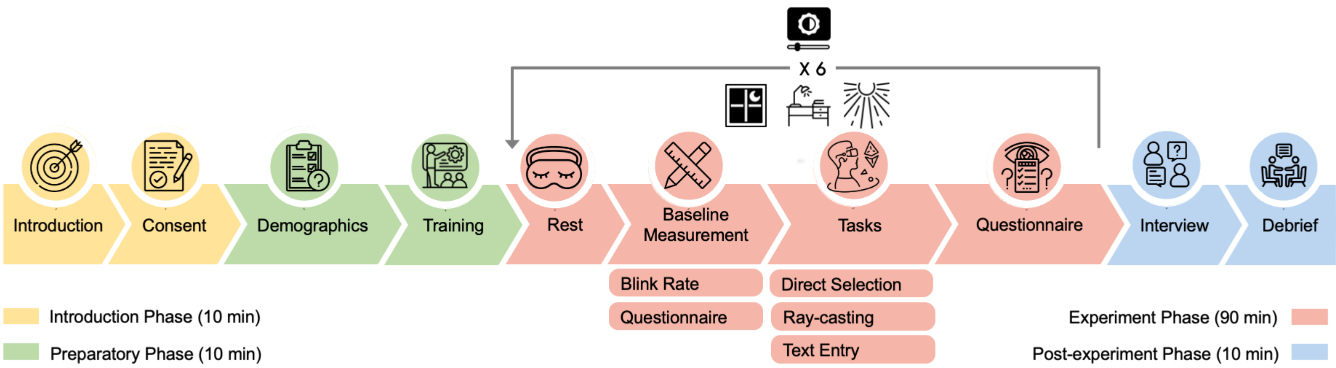


FIGURE 3. The experimental procedure.

able as needed. During breaks, participants were encouraged to look at a distance of over 6 meters to relax their eyes [86]. After the final condition, a brief semi-structured interview was conducted to gather qualitative feedback. The entire session lasted approximately 120 minutes, and participants were compensated with a $20 voucher. The overview of the procedure is presented in Figure 3.

E. DATA ANALYSIS

We analysed our data using the Bayesian statistical method due to its ability to quantify uncertainty, incorporate prior knowledge, and better handle small samples [87]. We refer readers unfamiliar with such methods to McElreath [88] for an accessible introduction and to Schmettow [89] for examples within HCI. We fit a separate model for each dependent variable, with baseline references set to medium ambient illuminance (500 lx) and low screen brightness (50%). Regularising priors were applied to stabilise estimates and avoid overfitting [90]. We also included covariates (ID, ID_e, target width) due to their known influence on performance measures [20]. To control for confounding influences such as learning and fatigue across the experiment [91], we included condition order as a fixed effect. Participant ID was modelled as a random effect to account for individual differences and repeated measures, consistent with best practices in within-subject designs [20].

We fit these models using the brms package in R [92], which provides an interface for fitting Bayesian multilevel models via the Stan probabilistic programming language [93]. We assessed the stability of the Markov Chain Monte Carlo sampling with R-hat, which should be lower that 1.01 [94] and the Effective sample size (ESS), which should be greater than 1,000 [92]. All of our estimates fit these criteria. We report the posterior means with standard errors and 95% compatibility interval (CI) [92]. Furthermore, we evaluate hypotheses by examining the posterior probability, which represents the probability that a condition has a positive effect on performance [95]. Our analysis scripts and detailed model outputs are available in the supplementary materials.

IV. RESULTS

A. TARGET SELECTION

A total of 13,467 records were collected during the target selection tasks. We filtered the dataset to remove accidental activations rather than intentional selections, which are common in pointing tasks [15]. Following the criteria established by Ng et al. [15], trials with a selection time below 0.1 seconds were excluded from our dataset, resulting in a final dataset of 12,877 records (11 targets/task/participant $\times$ 6 conditions (3 illuminance $\times$ 2 brightness) $\times$ 4 IDs $\times$ 2 tasks (direct selection/ray-casting) $\times$ 24 participants + wrong selections) for analysis.

1) Movement Time

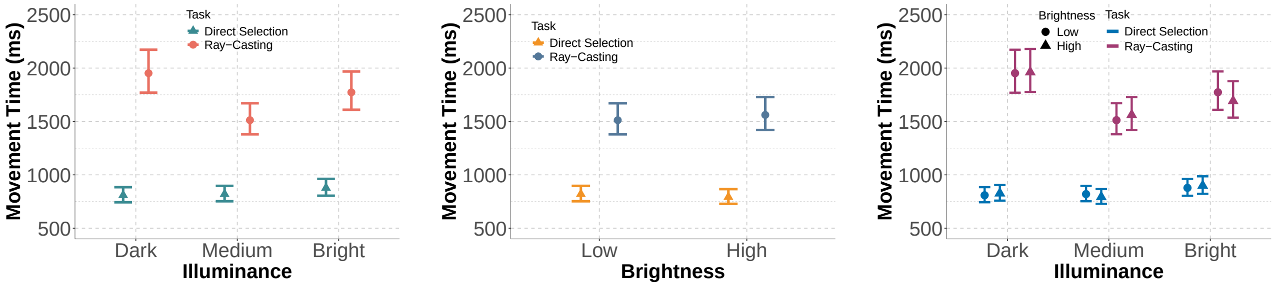
For movement time analysis, we included only the data with correctly selected targets. Erroneous selections were used solely for error rate analysis. We modelled the movement time using a shifted log-normal distribution [96], [97]. Table 1 summarises model coefficients and 95% credible intervals, and Figure 4 visualises posterior predictions.

For direct selection, movement times remained relatively stable across illuminance conditions (Figure 4a). Bright illuminance showed a slight increase in movement time (mean = 0.11, CI = [-0.04, 0.26]) with a posterior probability of 51%, indicating weak evidence for any effect. Dark illuminance demonstrated a small decrease in movement time (mean = -0.11, CI = [-0.26, 0.04]) with a probability of 40%, suggesting the effect is unlikely. Similarly, increasing screen brightness from 50% to 100% did not affect performance (41% probability; mean = -0.02, CI = [-0.17, 0.14]). As indicated in Figure 4c, the movement time remained stable regardless of the combination of ambient illuminance and screen brightness.

For ray-casting, under dark illuminance (mean = 0.10, CI = [-0.08, 0.28]), movement increased by 29% (MT = 1953ms) compared to the average office condition (MT = 1512ms), with a posterior probability of 92%, providing strong evidence supporting the negative effect of dark illuminance on movement time. Similarly, bright illuminance conditions showed a 17% increase (MT = 1773ms) compared to the average office condition, with an 88% probability of having a negative impact on movement time (mean = 0.02, CI = [-0.20, 0.15]). Higher screen brightness appeared beneficial

TABLE 1. Results of the movement time model: MT ~ 1 + (1|participant) + order + ID · (illuminance · brightness). All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Direct Selection		Ray-Casting	
	Estimate (Est.Error)	95% CI	Estimate (Est.Error)	95% CI
Intercept	6.20 (0.08)	[6.04, 6.36]	6.88 (0.09)	[6.71, 7.06]
Bright Illuminance	0.11 (0.08)	[-0.04, 0.26]	0.02 (0.09)	[-0.20, 0.15]
Dark Illuminance	-0.11 (0.08)	[-0.26, 0.04]	0.10 (0.09)	[-0.08, 0.28]
High Brightness	-0.02 (0.08)	[-0.17, 0.14]	-0.11 (0.09)	[-0.29, 0.07]
Bright Illuminance:High Brightness	0.00 (0.11)	[-0.21, 0.21]	-0.02 (0.13)	[-0.28, 0.23]
Dark Illuminance:High Brightness	0.20 (0.11)	[-0.01, 0.42]	0.14 (0.13)	[-0.11, 0.39]



(a) Effect of Illuminance on Movement Time (b) Effect of Brightness on Movement Time (c) Interaction Effect on Movement Time

FIGURE 4. Posterior predictions of movement time (higher is worse) across illuminance and brightness conditions for direct selection and ray-casting tasks. Dots are the mean predictions, and bounds represent 95% CI.

for ray-casting (mean = -0.11, CI = [-0.29, 0.07]) with a posterior probability of 11%, providing moderate evidence for improved performance.

2) Pointing Offset

Pointing offset was analysed in a similar way: only correctly selected targets were included, with erroneous selections reserved for error rate computation. The pointing offset data also followed a shifted log-normal distribution [96], [97]. Table 2 summarises model coefficients and 95% credible intervals, and Figure 5 visualises posterior predictions.

For direct selection, bright illuminance had a substantial negative impact on pointing accuracy for direct selection (mean = 0.32, CI = [0.06, 0.59]), corresponding to a 12% increase in pointing offset (offset = 22.15 mm) compared to the average office condition (offset = 19.74 mm). The posterior probability was 99%, providing strong evidence that higher ambient light impairs direct selection accuracy. Similarly, dark illuminance resulted in a higher pointing offset (mean = 0.12, CI = [-0.14, 0.37]), with an 81% posterior probability, indicating moderate evidence for a detrimental effect. The results also showed that increasing screen brightness appeared to slightly decrease accuracy (mean = 0.09, CI = [-0.16, 0.35]), with moderate evidence for this effect (76% posterior probability).

For ray-casting, both bright and dark illuminance were associated with higher pointing offsets (bright: mean = 0.19, CI = [0.06, 0.45]; dark: mean = 0.19, CI = [0.05, 0.45]), with posterior probabilities of 94% and 93%, respectively. These results provide strong evidence that non-average illuminance conditions negatively affect pointing precision when using ray-casting. In contrast, higher screen brightness was beneficial for ray-casting accuracy (mean = -0.19, CI = [-0.44, 0.06]), with a posterior probability of 6%, providing strong evidence for improved pointing precision under high brightness conditions. As shown in Figure 5c, this improvement was most pronounced under bright ambient illuminance, where higher screen brightness helped counteract the negative effects of the environment.

3) Error Rate

Error rates were modelled using a Poisson distribution, consistent with prior work [20]. Table 3 summarises model coefficients and 95% credible intervals, and Figure 6 visualises posterior predictions.

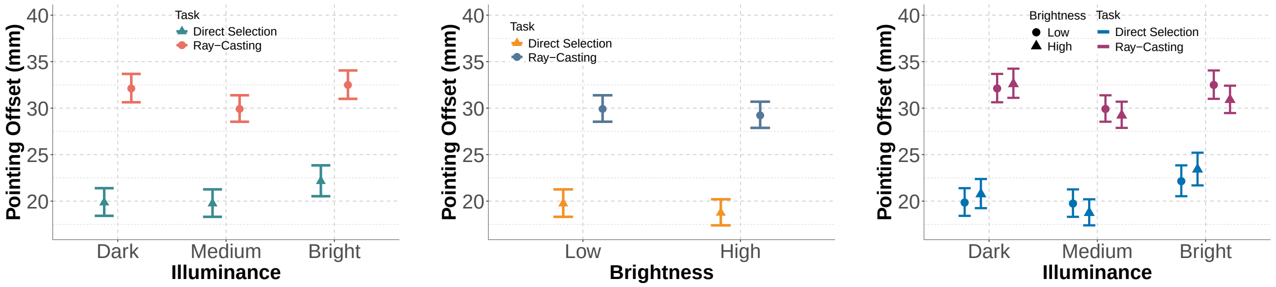
For direct selection, the model suggested that bright illuminance yielded an increase in error rates (mean = 0.89, CI = [-0.52, 2.27]) with a posterior probability of 85%, indicating moderate evidence for increased errors under bright ambient light. The model suggests that the effect of dark illuminance has only a 56% probability of leading to a higher error rate in direct selection (mean = 0.71, CI = [-0.65, 2.07]), which indicates no effect on the error rate. Increasing screen brightness to 100% resulted in a small decrease in error rate (mean = -0.54, CI = [-1.95, 0.85]), but with a low posterior probability (23%), providing moderate evidence for improved accuracy with increased brightness.

For ray-casting, bright light conditions yielded a slight increase in error rates (mean = 0.07, CI = [-0.95, 1.10]) with a posterior probability of only 68%. In contrast, dark illuminance elevated error rates (mean = 0.23, CI = [-0.75, 1.21]) with a posterior probability of 90%, suggesting strong

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3631406

TABLE 2. Results of the pointing offset model: offset ~ 1 + (1|participant) + order + width · (illuminance · brightness). All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Direct Selection		Ray-Casting	
	Estimate (Est.Error)	95% CI	Estimate (Est.Error)	95% CI
Intercept	2.27 (0.10)	[2.08, 2.47]	2.67 (0.09)	[2.49, 2.86]
Bright Illuminance	0.32 (0.13)	[0.06, 0.59]	0.19 (0.13)	[0.06, 0.45]
Dark Illuminance	0.12 (0.13)	[-0.14, 0.37]	0.19 (0.13)	[0.05, 0.45]
High Brightness	0.09 (0.13)	[-0.16, 0.35]	-0.19 (0.13)	[-0.44, 0.06]
Bright Illuminance:High Brightness	-0.29 (0.18)	[-0.64, 0.07]	0.24 (0.18)	[-0.12, 0.59]
Dark Illuminance:High Brightness	-0.15 (0.18)	[-0.51, 0.21]	0.23 (0.18)	[-0.12, 0.59]

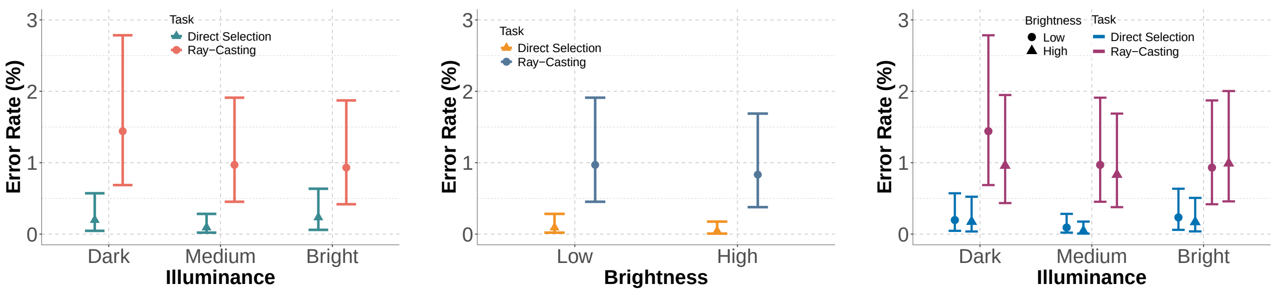


(a) Effect of Illuminance on Pointing Offset (b) Effect of Brightness on Pointing Offset (c) Interaction Effect on Pointing Offset

FIGURE 5. Posterior predictions of pointing offset (higher is worse) across illuminance and brightness conditions for direct selection and ray-casting tasks. Dots are the mean predictions, and bounds represent 95% CI.

TABLE 3. Results of the error rate model: error ~ 1 + (1|participant) + order + ID · (illuminance · brightness). All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Direct Selection		Ray-Casting	
	Estimate (Est.Error)	95% CI	Estimate (Est.Error)	95% CI
Intercept	-4.97 (0.94)	[-6.89, -3.21]	-4.67 (0.56)	[-5.78, -3.59]
Bright Illuminance	0.89 (0.70)	[-0.52, 2.27]	0.07 (0.52)	[-0.95, 1.10]
Dark Illuminance	0.71 (0.69)	[-0.65, 2.07]	0.23 (0.50)	[-0.75, 1.21]
High Brightness	-0.54 (0.71)	[-1.95, 0.85]	0.02 (0.50)	[-0.97, 0.99]
Bright Illuminance:High Brightness	0.30 (0.76)	[-1.18, 1.77]	0.17 (0.59)	[-1.01, 1.31]
Dark Illuminance:High Brightness	0.24 (0.76)	[-1.27, 1.74]	0.04 (0.59)	[-1.11, 1.17]



(a) Effect of Illuminance on Error Rate (b) Effect of Brightness on Error Rate (c) Interaction Effect on Error Rate

FIGURE 6. Posterior predictions of error rate (higher is worse) across illuminance and brightness conditions for direct selection and ray-casting tasks. Dots are the mean predictions, and bounds represent 95% CI.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3631406

evidence. When using ray-casting, dark conditions (error rate = 1.44%) resulted in a 48% increase in error rate compared to the average office condition (error rate = 0.97%). Higher screen brightness had minimal effect on ray-casting error rates (mean = 0.02, CI = [-0.97, 0.99]) with a posterior probability of 51%, indicating the hypothesis is unlikely.

B. TEXT ENTRY

For the text entry task, we collected a total of 432 sentences (3 sentences/participant x 6 conditions (3 illuminance x 2 brightness) x 24 participants). We excluded 11 sentences (2.5%) that were left empty due to participants inadvertently pressing the enter key. This left 421 complete sentences for our text entry analysis.

1) Throughput

We employed Bayesian regression with a Gaussian likelihood to model the effects on sentence text entry throughput. This approach allowed us to quantify typing speed (words or characters per minute) while simultaneously accounting for error rates. Table 4 summarises model coefficients and 95% credible intervals, and Figure 7 visualises posterior predictions.

TABLE 4. Results of the text entry throughput model: TP ~ 1 + (1|participant) + order + illuminance · brightness. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Estimate (Est.Error)	95% CI
Intercept	3.38 (0.26)	[2.87, 3.87]
Bright Illuminance	-0.32 (0.14)	[-0.60, -0.02]
Dark Illuminance	-0.93 (0.14)	[-1.21, -0.65]
High Brightness	0.29 (0.14)	[0.01, 0.57]
Bright Illuminance:High Brightness	-0.50 (0.21)	[-0.92, -0.11]
Dark Illuminance:High Brightness	-0.06 (0.20)	[-0.46, 0.33]

Our results showed that dark illuminance had a substantial negative impact on typing performance (mean = -0.93, CI = [-1.21, -0.65]) with a posterior probability of 100%, providing strong evidence that low-light conditions significantly impair text entry throughput. Specifically, the dark condition (throughput = 3.00 bits/s) resulted in a 30% decrease in throughput compared to the average office condition (throughput = 3.92 bits/s). Bright illuminance also negatively affected performance (mean = -0.32, CI = [-0.60, -0.02]) with a posterior probability of 98%, providing strong evidence. The bright condition (throughput = 3.61 bits/s) resulted in an 8% decrease in throughput compared to the average office condition (throughput = 3.92 bits/s). In contrast, higher screen brightness showed a positive effect on typing performance (mean = 0.29, CI = [0.01, 0.57]) with a posterior probability of 98%, providing strong evidence that increased display brightness enhances text entry throughput.

2) UER and CER

We evaluated both uncorrected error rate (UER) and corrected error rate (CER) of text entry. Both metrics were modelled using a zero-inflated beta distribution, accounting for the possibility of error-free performance in the dataset. Table 5 summarises model coefficients and 95% credible intervals, and Figure 8 visualises posterior predictions.

For UER, there was a reduction in UER under bright illuminance (mean = -0.14, CI = [-0.81, 0.51]) with a posterior probability of 34%, suggesting moderate evidence for a small negative effect. Dark illuminance slightly increased UER (mean = 0.07, CI = [-0.52, 0.67]), but the posterior probability of 59% indicates little to no effect. In contrast, higher screen brightness substantially increased UER (mean = 0.40, CI = [-0.27, 1.08]) with a posterior probability of 87%, providing moderate evidence for more uncorrected errors at higher brightness levels.

For CER, bright illuminance led to a moderate increase (mean = 0.23, CI = [-0.02, 0.48]) compared to normal lighting, with a posterior probability of 96%, providing strong evidence for elevated corrected errors under bright lighting conditions. Dark illuminance had a stronger negative impact (mean = 0.39, CI = [0.15, 0.63]) with a posterior probability of 100%, providing strong evidence for substantially increased corrected errors in low-light conditions. In contrast, higher screen brightness had minimal effect on CER (mean = -0.02, CI = [-0.27, 0.23]) with a posterior probability of 44%, indicating no meaningful effect.

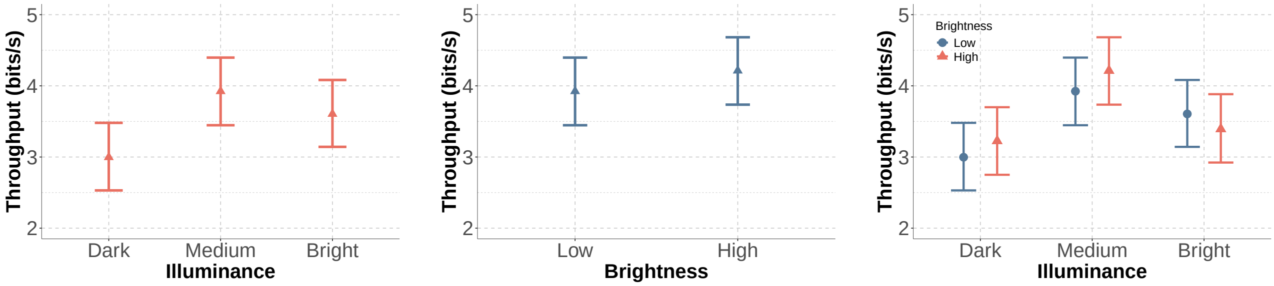
C. VISUAL FATIGUE

The visual fatigue data included blink rate measurements and Visual Fatigue Questionnaire (VFQ) scores. VFQ scores were captured before (baseline) and after (post-task) the interaction phase, while blink rates were measured during the baseline and the task period. Hence, there are a total of 144 pairs of measurements for both metrics. To quantify visual fatigue, we computed the change in each metric from its baseline. For VFQ scores, the change was defined as *post-task − baseline*, such that positive values indicated greater perceived fatigue. For blink rate, since fatigue is associated with a reduction in blink rate, the change was defined as *baseline − task-period*. This inversion ensured that larger positive values consistently represent greater fatigue across both metrics, facilitating clearer comparisons and visualisation.

1) Change in Blink Rate

Changes in blink rate relative to baseline were modelled using Bayesian regression with skew normal distributions to quantify the effects on ocular behaviour while accounting for individual baseline variations. Table 6 summarises model coefficients and 95% credible intervals, and Figure 9 visualises posterior predictions. Figure 10 provides descriptive baseline and task-period blink rates across all conditions.

We found strong evidence that ambient illuminance significantly influenced visual fatigue. Our model suggests that the bright illuminance induced the largest increase in fatigue (mean = 4.93, CI = [3.06, 6.91]) with a posterior probability of a positive effect at 100%. Dark illuminance also showed a notable effect (mean = 1.69, CI = [-0.05, 3.49], posterior probability = 97%). Similarly, higher screen brightness led to

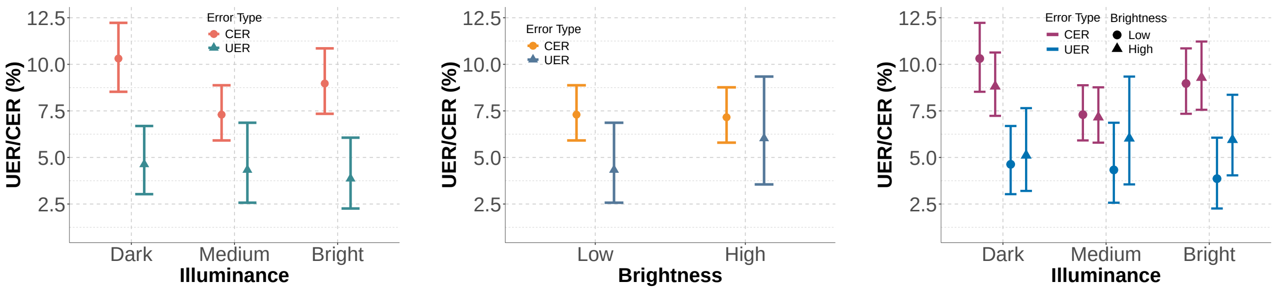


(a) Effect of Illuminance on Throughput (b) Effect of Brightness on Throughput (c) Interaction Effect on Throughput

FIGURE 7. Posterior predictions of text entry throughput (lower is worse) across illuminance and brightness conditions for text entry tasks. Dots are the mean predictions, and bounds represent 95% CI.

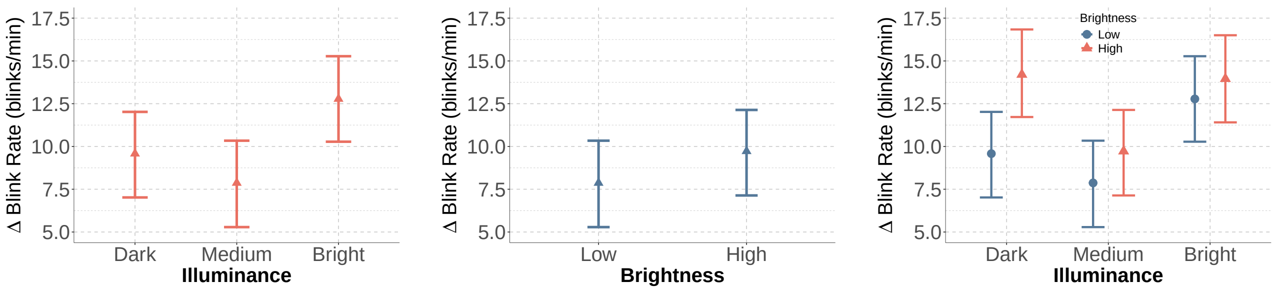
TABLE 5. Results of the UER/CER models: UER/CER ~ 1 + (1|participant) + order + illuminance · brightness. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Uncorrected Error Rate		Corrected Error Rate	
	Estimate (Est.Error)	95% CI	Estimate (Est.Error)	95% CI
Intercept	-1.44 (0.32)	[-2.07, -0.81]	-2.06 (0.13)	[-2.33, -1.81]
Bright Illuminance	-0.14 (0.34)	[-0.81, 0.51]	0.23 (0.13)	[-0.02, 0.48]
Dark Illuminance	0.07 (0.31)	[-0.52, 0.67]	0.39 (0.12)	[0.15, 0.63]
High Brightness	0.40 (0.35)	[-0.27, 1.08]	-0.02 (0.13)	[-0.27, 0.23]
Bright Illuminance:High Brightness	0.13 (0.42)	[-0.70, 0.96]	0.06 (0.18)	[-0.29, 0.41]
Dark Illuminance:High Brightness	-0.28 (0.43)	[-1.13, 0.54]	-0.16 (0.17)	[-0.49, 0.18]



(a) Effect of Illuminance on UER and CER (b) Effect of Brightness on UER and CER (c) Interaction Effect on UER and CER

FIGURE 8. Posterior predictions of UER and CER (higher is worse) across illuminance and brightness conditions for text entry tasks. Dots are the mean predictions, and bounds represent 95% CI.



(a) Effect of Illuminance on ΔBlink Rate (b) Effect of Brightness on ΔBlink Rate (c) Interaction Effect on ΔBlink Rate

FIGURE 9. Posterior predictions of change in blink rate (higher is worse) across illuminance and brightness conditions for text entry tasks. Dots are the mean predictions, and bounds represent 95% CI.

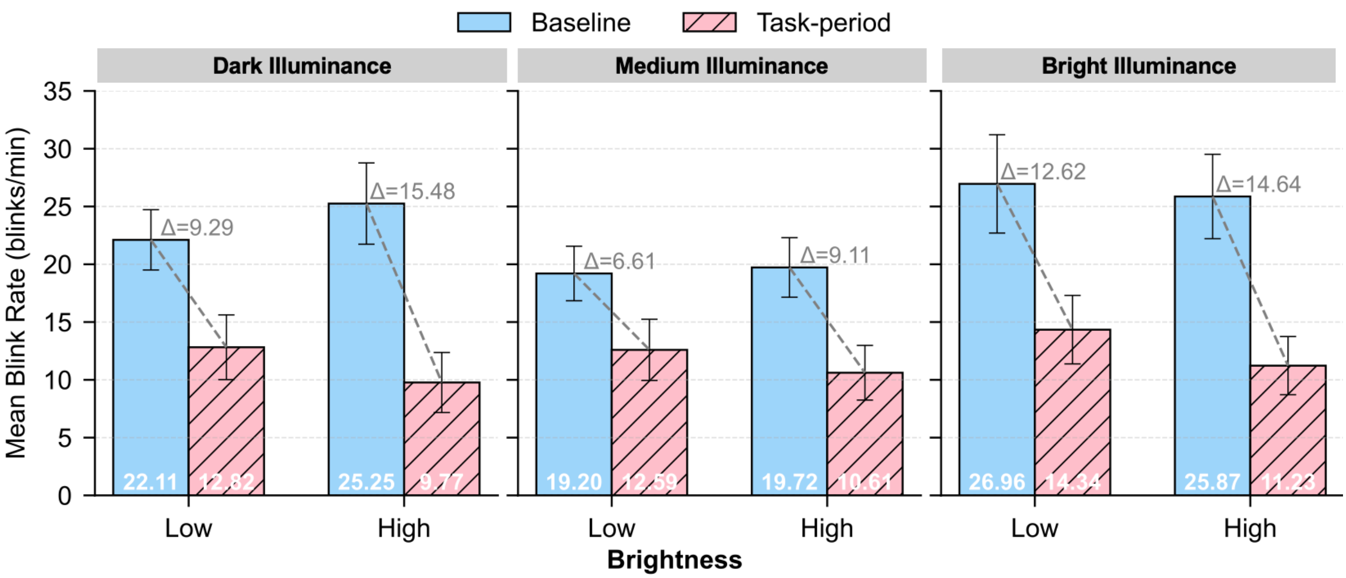


FIGURE 10. Mean blink rate during baseline and task periods across different illuminance and screen brightness conditions. Blue bars represent baseline measurements, while red hatched bars denote task-period rates.

TABLE 6. Results of the change in blink rate model: blink_rate_change ~ 1 + (1|participant) + order + illuminance · brightness. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Estimate (Est.Error)	95% CI
Intercept	8.47 (1.39)	[5.68, 11.10]
Bright Illuminance	4.93 (0.97)	[3.06, 6.91]
Dark Illuminance	1.69 (0.90)	[-0.05, 3.49]
High Brightness	1.86 (0.85)	[0.23, 3.55]
Bright Illuminance:High Brightness	-0.68 (1.24)	[-3.15, 1.78]
Dark Illuminance:High Brightness	2.81 (1.23)	[0.45, 5.31]

TABLE 7. Results of the change in visual fatigue score model: visual_score_change ~ 1 + (1|participant) + order + illuminance · brightness. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Estimate (Est.Error)	95% CI
Intercept	0.37 (0.11)	[0.15, 0.57]
Bright Illuminance	0.32 (0.09)	[0.15, 0.49]
Dark Illuminance	0.21 (0.08)	[0.05, 0.38]
High Brightness	0.26 (0.08)	[0.10, 0.43]
Bright Illuminance:High Brightness	-0.19 (0.12)	[-0.43, 0.04]
Dark Illuminance:High Brightness	0.12 (0.12)	[-0.11, 0.35]

a greater increase in fatigue (mean = 1.86, CI = [0.23, 3.55]), with a 99% probability of leading to higher visual strain.

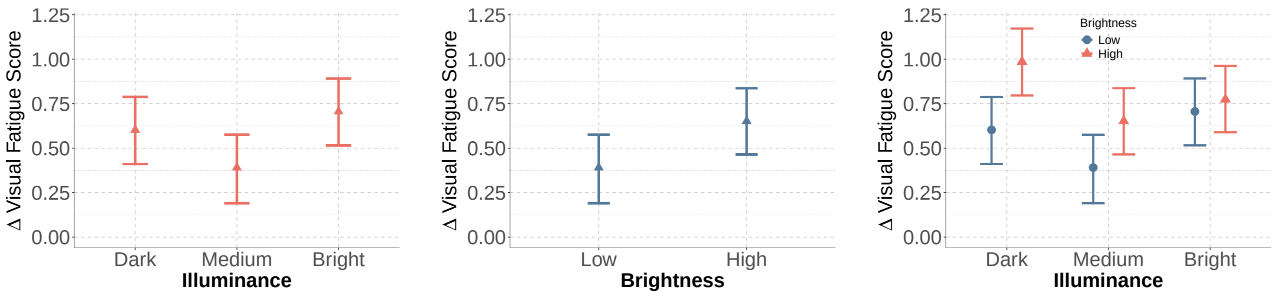
We also observed strong evidence for an interaction between dark illuminance and 100% screen brightness (mean = 2.81, CI = [0.45, 5.31]), with a posterior probability of 99%. This combination of a dark environment and a bright screen resulted in the highest overall change in blink rate. Our model predictions are consistent with the descriptive data in Figure 10, which shows the largest drop in blink rate occurred in the dark illuminance, 100% Brightness condition ($\Delta = 15.48$). By contrast, medium illuminance and low screen brightness resulted in a more modest reduction ($\Delta = 6.61$).

2) Change in Visual Fatigue Score

Visual fatigue score changes were similarly analysed using Bayesian regression with skew normal distributions, allowing for robust estimation of subjective fatigue measures. Table 7 summarises model coefficients and 95% credible intervals, and Figure 11 visualises posterior predictions. Figure 12 provides descriptive baseline and post-task VFQ scores across all conditions.

Our model indicates that, compared to the medium illuminance condition, bright illuminance led to the largest increase in subjective fatigue (mean = 0.32, 95% CI = [0.15, 0.49], probability = 100%), which is consistent with the blink rate findings. This provides strong evidence that higher ambient light levels increased participants' perceived visual fatigue after tasks. Similarly, dark illuminance also resulted in a significant increase in fatigue scores (mean = 0.21, 95% CI = [0.05, 0.38]), with a 99% posterior probability of a positive effect, which constitutes strong evidence. We also found strong evidence for the effect of screen brightness. Increased screen brightness led to a notable rise in visual fatigue scores (mean = 0.26, CI = [0.10, 0.43]), with a posterior probability of 100%. We found moderate evidence for an interaction between dark illuminance and high screen brightness (mean = 0.12, CI = [-0.11, 0.35]), with a probability of 85% that this combination leads to greater increases in fatigue.

These model predictions align well with the descriptive data presented in Figure 12. The largest increase in VFQ score was observed in the dark lighting, 100% brightness condition ($\Delta = 0.98$). Conversely, the smallest increase in fatigue occurred in the medium illuminance, 50% brightness condition



(a) Effect of Illuminance on Δ Fatigue Score (b) Effect of Brightness on Δ Fatigue Score (c) Interaction Effect on Δ Fatigue Score

FIGURE 11. Posterior predictions of change in visual fatigue score (higher is worse) across illuminance and brightness conditions for text entry tasks. Dots are the mean predictions, and bounds represent 95% CI.

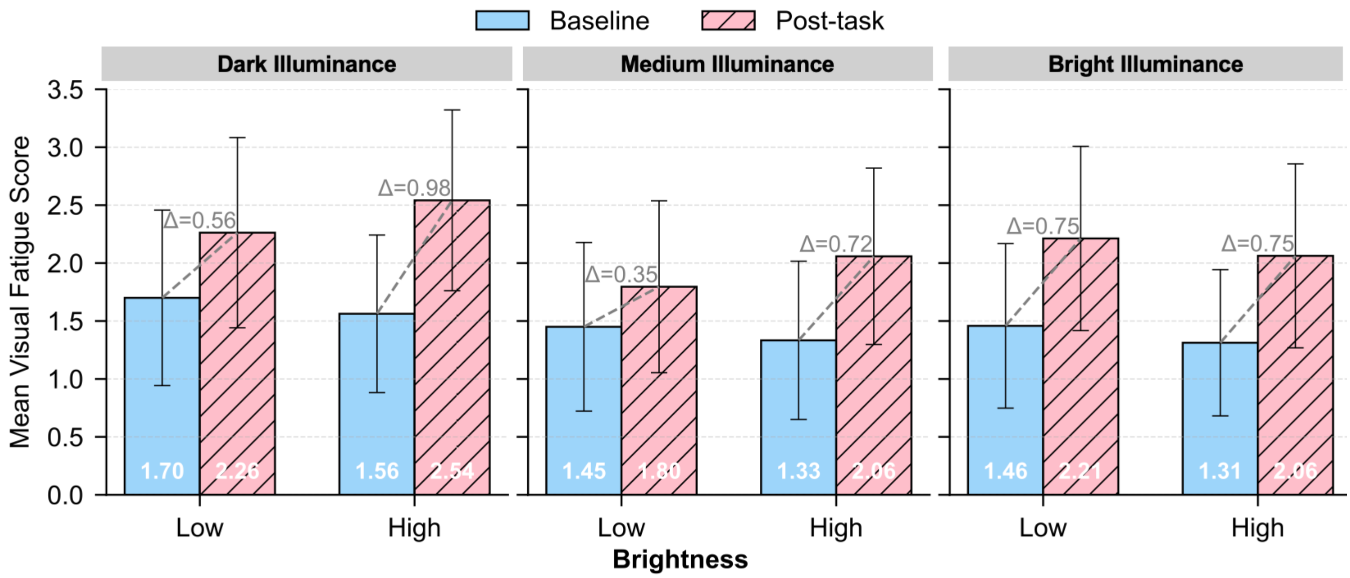


FIGURE 12. Mean visual fatigue scores at baseline and post-task across different illuminance and screen brightness conditions. Blue bars represent baseline measurements, while red hatched bars denote post-task scores.

($\Delta = 0.35$), which served as the reference level in our model and represents a more balanced lighting environment.

V. DISCUSSION

In this section, we consolidate the observed phenomena from both experiments conducted in our study and previous research to gain a better understanding of the influence of ambient illuminance and screen luminance on MR interaction.

A. EFFECTS OF AMBIENT ILLUMINANCE AND SCREEN LUMINANCE ON MR INTERACTION

1) Target Selection

In the direct selection task, our analysis did not reveal an effect of ambient illuminance on movement time. In contrast, we noticed that ray-casting required more time to complete, as it involves a two-step process: aiming the ray at the target and then confirming selection with a pinch gesture. This indicates that ray-casting has higher demands on motor precision and visual alignment [21], imposing a greater intrinsic cognitive load, which refers to the inherent mental effort required by the task itself [98]. Moreover, ray-casting performance was substantially affected by ambient illuminance: in bright sunlight, movement time increased considerably compared to the average office baseline, and in dark conditions, the increase was even more pronounced. The vulnerability of ray-casting is consistent with findings from mobile device studies, where participants took significantly longer to tap a target under the dimmed light during mobile interaction [26].

Pointing offset showed a similarly differentiated pattern between interaction techniques. In direct selection, dark ambient illumination had minimal impact, with offsets staying within 2 mm (roughly 2–5% of the target size), which is a negligible difference for real-world MR interactions [20]. However, bright illuminance noticeably increased the offset, likely due to glare and reduced contrast, which makes the target's centre harder to discern. In ray-casting, pointing offsets were larger overall due to the wider effective target width. We also observed a strong link between ambient illuminance and ray-casting offsets, with significant increases in both bright and dark environments. This mirrors prior mobile interaction

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3631406

findings that accuracy declines in tapping circles under dim environments compared to normal lighting [26].

Error rates showed only a minor influence from dark illuminance on ray-casting, with a small increase. Effects were even smaller for bright illuminance and across all direct selection conditions. Overall, these changes are negligible, as all error rates remained well below the 4% threshold assumed in Fitts' law [39]. Thus, ambient illuminance appears to have no meaningful impact on target selection error rates.

In summary, these results highlight a key difference between the two interaction techniques. Direct selection demonstrated high resilience to changes in lighting. This resilience is likely due to the near-field nature of the task, where users benefit from a physical sense of their hand's position in space (proprioception) and absolute depth perception [99], [100]. This reduces reliance on high visual fidelity and thereby lowers the cognitive load associated with visual processing, allowing users to maintain effective performance even in suboptimal lighting. Users can draw on natural spatial awareness and motor memory during physical reaches toward virtual objects, minimising the need for precise visual cues.

Conversely, ray-casting requires sustained foveation and fine motor stability to align a narrow ray with distant targets [21]. This demands precise hand control while simultaneously interpreting depth cues and spatial relationships. Research has shown that high ambient illuminance can reduce display contrast and weaken depth cues, making depth judgments less accurate [101]. Under bright sunlight, the additive nature of HMDs causes virtual elements to appear washed out, severely reducing their contrast against the real-world background. As a result, both the ray and target become faint and difficult to discern. This impairs ray visibility and trajectory perception, forcing participants to slow down and struggle with precision, which in turn increases the mental effort. This phenomenon is well-explained by Cognitive Load Theory [98], which identifies this type of environmentally-induced difficulties as extraneous cognitive load. This cognitive burden diverts working memory resources away from the primary task of aiming and selecting, thereby causing slower and less precise interactions. These findings align with prior AR studies by Gabbard et al. [28] and Erickson et al. [31], which highlighted that high ambient illuminance significantly reduces the visibility and contrast of virtual elements. In dark conditions, the lack of real-world visual cues impairs depth perception and spatial awareness, making it harder to gauge distance and maintain a steady hand for precise pointing. This issue was compounded by technical limitations of the headset, as a large number of participants (N=20) experienced hand tracking issues in low-light environments, e.g., *"There was a major tracking issue; I had pinched multiple times but there was no response"* (P01, P18), *"There was a moment when it lost track of my hand and the virtual hand disappeared, so I had to adjust a bit for it to pick my hand up again"* (P10). These observations are consistent with real-world reports that low-light conditions impair the camera-based hand-tracking capabilities of wearable headsets [102], [103]. These tracking instabilities further degraded task performance by forcing users to re-establish hand registration and disrupting their spatial orientation.

Furthermore, our study found that screen brightness adjustments (low vs. high) had negligible effects on all performance for both interaction techniques. This aligns with research on VR displays by Vasylevska et al. [35], who found that substantially lowered screen brightness produced no statistically significant differences in task performance. Our finding is reinforced by participant feedback during interviews, where the majority of participants (N=18) reported no perceived difference in target selection performance when screen brightness changed, e.g., *"I didn't really feel any difference in target selection when the screen brightness changed"* (P16). This suggests that within the tested range, internal display brightness is a far less critical factor for interaction performance in target selection than the ambient lighting environment.

Finally, while our results demonstrate robust effects of ambient illuminance on target selection, the fundamental differences between interaction techniques remain the dominant factor. Ray-casting required 84-141% longer movement times than direct selection across all lighting conditions (medium: direct = 820ms vs. ray-casting = 1512ms; dark: direct = 809ms vs. ray-casting = 1953ms). Similarly, pointing offsets for ray-casting were 52-62% larger (medium: direct = 19.74mm vs. ray-casting = 29.91mm; dark: direct = 19.85mm vs. ray-casting = 32.12mm). These differences are much larger than the effects caused by the situational impairments. Therefore, while adapting MR systems to lighting is important, the initial selection of an efficient interaction modality for a given context is a more critical design consideration.

2) Text Entry

Our study reveals a negative correlation between ambient illuminance and text entry throughput. Specifically, we observed a substantial decrease in throughput under dark conditions and a notable reduction in bright environments. As noted by many participants (N=15), the dark condition significantly hampered their typing performance, e.g., *"Text entry task was the most affected among these three tasks"* (P15). This pronounced sensitivity can be attributed to the high visual demands of text entry, which requires users to continuously shift their gaze between the individual keys of the virtual keyboard and the text output area. This task necessitates sustained visual attention to target small interface elements accurately, making it an intrinsically high-load cognitive task.

In bright sunlight, the diminished contrast between virtual keyboard and real-world background made it more challenging for users to rapidly and accurately identify individual keys and read the text being entered. When the visual distinctiveness of targets is reduced, users must invest more effort in locating keys and are more susceptible to selection errors. As text entry is already an intrinsically high-load task, this diminished contrast added an extraneous cognitive load [98], overwhelming users' cognitive capacity and leading to slower, less accurate typing. This finding is consistent with prior research by Gabbard et al. [28], which demonstrated that bright conditions can "wash out" augmented reality displays, reducing the perceived contrast of overlaid text. Conversely, in dark environments, the lack of ambient light obscured the user's own hands. This forced participants to rely solely on the virtual representations of their hands for input, which can be subject to tracking latency and imprecision [104]. A large number of participants (N=20) reported these issues, e.g., *"When in the dark room, the virtual hand did not align well with my real hand; it was shifted"* (P02), *"There is a tracking issue under the dark condition; I can still see the virtual keyboard clearly, but it's easy to hit the wrong keys"* (P04, P23). These observations align with recent findings by Tran et al. [102], which reported degraded hand-tracking and passthrough quality in dark lighting due to limited camera functionality among widely available wearable headsets. This reduces the productivity and usability in dimly lit settings.

Our results indicate that screen luminance can partially mitigate these challenges. Increasing screen brightness led to an improved throughput. This was particularly effective in counteracting the negative effects of bright ambient light by enhancing the legibility of the virtual content, a finding supported by the work of Yan et al. [79] on optimal screen luminance. However, this increase in brightness also corresponded with a higher UER. This suggests that brighter displays might induce user overconfidence. Participants appeared to prioritise speed over accuracy under these conditions, leading to fewer real-time corrections but more residual errors in the final text output. On the other hand, the CER increased significantly under both bright and dark conditions. This pattern suggests that extreme lighting conditions frequently resulted in mis-taps and reduced input accuracy, necessitating more frequent visual re-checking and corrections, which in turn slows down the overall text entry throughput.

Interestingly, these results contrast with prior work on mobile text entry, which found no significant impact of dim lighting on typing speed or error rate [26]. This discrepancy might be due to the differences in input modalities. MR text entry relies on in-air typing, requiring both spatial alignment and accurate hand-tracking. As our study and others have shown, it is more susceptible to environmental lighting conditions than touchscreen-based mobile typing.

3) Visual Fatigue

Our findings demonstrate that both extreme ambient lighting conditions and high screen brightness substantially increased eye strain, with objective and subjective measures showing consistent patterns. Most notably, the combination of dark ambient illumination with high screen luminance produced the largest fatigue increase across both measures. As noted by Vasylevska et al. [35], a significant mismatch between the brightness of a head-mounted display (HMD) and a dark background is a primary cause of eye strain. This is because the human eye struggles to adapt to the drastic difference in illumination. This extreme contrast forces the iris into a state of constant and strenuous adjustment, leading to symptoms such as eye strain, sensitivity to light (photophobia), and headaches [48], all of which are captured by our VFQ. In post-experiment interviews, a number of participants (N=13) confirmed these findings, e.g., *"It caused very severe eye strain, and my eyes felt incredibly dry"* (P09, P22).

Bright sunlight also proved to be highly fatiguing, producing a significant increase in VFQ scores and a significant drop in blink rate. The reduction in blinking observed in challenging visual conditions is a well-documented physiological response to increased visual demands [105]. When visual processing becomes effortful, the central nervous system suppresses blinking to maximise visual information intake [106]. However, reduced blinking hinders the proper distribution of the tear film across the eye's surface, which can cause symptoms of dry eye, such as a gritty or burning sensation, and contribute to overall visual discomfort [25]. As a few participants (N=9) noted, *"Under bright sunlight, I tend to squint, and want to rub my eyes during the tasks"* (P03, P17).

In contrast, the most comfortable condition for participants was found to be average ambient lighting (500 lx) combined with low screen brightness (50%). This suggests that a balanced lighting environment is optimal for minimising visual fatigue. Our results also align with arguments that reducing screen brightness in HMD can maintain task performance while lowering visual strain [35].

B. ADDRESSING THE EFFECTS OF AMBIENT ILLUMINANCE AND SCREEN LUMINANCE ON MR INTERACTION

To effectively mitigate the negative impact of extreme lighting conditions on MR interaction, it is necessary first to detect ambient illuminance before adaptation can take place [13]. Integrating light sensors capable of continuously monitoring illuminance enables MR systems to identify challenging contexts in real time and dynamically adapt the interaction accordingly [17]. As proposed by Li et al. [21], adapting interaction modalities to the current context can benefit performance. For instance, given our findings that direct selection is less sensitive to lighting extremes than ray-casting, MR applications could dynamically switch the primary or recommended interaction modality to be direct selection in challenging lighting conditions.

To counteract the "washing out" of virtual elements in high-illuminance settings, MR systems can employ dynamic visual enhancements. Similar to previous studies for mobile devices that suggest increasing contrast or button sizes to accommodate situational visual impairments [107], [108], MR systems could automatically modify virtual element properties (e.g., size, contrast, and opacity). For ray-casting, dynamically thickening the ray mesh or adding a high-contrast outline could significantly improve its visibility against a bright background, addressing the precision issues we observed.

Furthermore, our findings highlighted the critical failure of camera-based hand-tracking in low-light conditions. By detecting these challenging conditions, an MR system could automatically augment or switch to more robust input modal-

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3631406

ities. This multimodal approach would allow users to remain productive even when a primary interaction technique is ineffective. For example, eye-tracking is a viable alternative for selection tasks, recognised for its efficient input capabilities [109]. In dark conditions where hand-tracking fails, the system could enable gaze-based selection. Similarly, voice command technology could serve as a powerful complementary input method [110], particularly for text entry, which our results showed was highly susceptible to performance degradation in the dark.

To address the significant visual fatigue induced by extreme lighting and brightness mismatches, MR devices could implement proactive fatigue monitoring systems. Research has shown that monitoring spontaneous blink rate with eye aspect ratio mapping can reliably identify the onset of eye strain [111]. Therefore, the MR system could leverage its integrated eye-tracking feature to detect blink rate in real-time. Upon detecting signs of fatigue, the system could trigger adaptive interventions, such as displaying subtle reminders to blink, suggesting short rest breaks, or automatically dimming the screen brightness to a more comfortable and eye-friendly level, as suggested by Vasylevska et al. [35].

VI. LIMITATIONS AND FUTURE WORK

Our work has several limitations. First, our study focused on three basic MR tasks and restricted participants to using only the index finger of their dominant hand for interaction. While these controls ensured fair comparisons by minimising variability in interaction strategy [11], [20], we acknowledge that real-world scenarios involve more complex interactions. We hypothesise the observed effects may be more pronounced in such scenarios, which future work could explore. Second, although our bright illuminance experiments were conducted outdoors to capture realistic conditions, this came at the cost of reduced experimental control. Despite continuously logging illuminance and restricting sessions to stable weather, we could not fully account for environmental factors, such as temperature and wind.

Third, our objective assessment of visual fatigue relied solely on blink rate. A more comprehensive evaluation could incorporate additional physiological measures such as pupil diameter [48] and EEG [112]. However, implementing such measures presents practical challenges, as optometric instruments typically require specialised expertise and equipment that may not be feasible for large-scale user studies [113]. Furthermore, our participant sample was limited to university students within a certain age range, which limits the generalizability of our findings, as factors such as visual acuity and motor skills can vary substantially across demographics. Future research should incorporate a more diverse participant sample to address this issue.

Fourth, we tested only three representative ambient lighting levels and two display brightness levels based on common MR usage scenarios. We acknowledge that these discrete levels represent a subset of the continuous and nuanced range of conditions users experience in reality. Future work could expand on this by examining intermediate illuminance levels (e.g., cloudy daylight) and various brightness levels.

Fifth, our findings are limited to a video see-through (VST) MR headset. We acknowledge that transfer to other headsets may vary with camera and display characteristics. In VST systems, virtual content is rendered over a camera feed, which allows the system to control the entire visual scene's contrast [114]. This differs significantly from optical see-through (OST) devices such as HoloLens 2, which blend virtual content directly with the user's natural view of the world through transparent displays [114]. Prior work has shown that OST displays face distinct challenges, as they are more sensitive to background brightness and can lose readability in high illuminance compared to VST displays [115]. Future work could replicate our study across multiple VST devices spanning camera and display capabilities, as well as conduct comparative studies between VST and OST HMDs to better understand technology-specific vulnerabilities.

Finally, our study revealed a confound in low-light conditions: it is difficult to separate performance degradation caused by human perceptual challenges from that caused by the hardware's technical limitations. This highlights the need for future work to focus on fundamental advances in hardware and tracking algorithms to ensure robust performance across all conditions.

VII. CONCLUSION

In this study, we investigated the effects of ambient illuminance and screen luminance on performance and visual fatigue in MR interaction. Our results reveal that the impact of lighting varies significantly across different interactions. Direct selection proved robust across all conditions. In contrast, ray-casting was highly sensitive to ambient extremes, with both movement time and pointing offset increasing substantially in bright and dark ambience. Text entry was also significantly impaired, with throughput dropping substantially in dark and bright conditions. Screen luminance had a minimal effect on target selection; however, higher luminance improved text entry throughput, which came at the cost of increased uncorrected error rates. Importantly, both extreme ambient lighting and high screen brightness substantially increased visual fatigue. The most fatigue was caused by dark environments combined with maximum screen brightness. Overall, our findings extend the understanding of situational impairments in MR interaction and highlight the need for adaptive solutions to support comfortable and effective MR experiences across diverse lighting contexts.

REFERENCES

[1] N. B. Milman, "Defining and conceptualizing mixed reality, augmented reality, and virtual reality," *Distance Learning*, vol. 15, no. 2, pp. 55–58, 2018.

[2] L. Kent, C. Snider, J. Gopsill, and B. Hicks, "Mixed reality in design prototyping: A systematic review," *Design Studies*, vol. 77, p. 101046, 2021.

[3] C. E. Hughes, C. B. Stapleton, D. E. Hughes, and E. M. Smith, "Mixed reality in education, entertainment, and training," *IEEE computer graphics and applications*, vol. 25, no. 6, pp. 24–30, 2005.

[4] B. John and N. Wickramasinghe, "A review of mixed reality in health care," *Delivering Superior Health and Wellness Management with IoT and Analytics*, pp. 375–382, 2019.

[5] A. Sears, M. Lin, J. Jacko, and Y. Xiao, "When computers fade: Pervasive computing and situationally-induced impairments and disabilities," in *HCI international*, vol. 2, no. 3, 2003, pp. 1298–1302.

[6] Z. Sarsenbayeva, "Situational impairments during mobile interaction," in *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, 2018, pp. 498–503.

[7] G. W. Tigwell, Z. Sarsenbayeva, B. M. Gorman, D. R. Flatla, J. Goncalves, Y. Yesilada, and J. O. Wobbrock, "Addressing the challenges of situationally-induced impairments and disabilities in mobile interaction," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '19. New York, NY, USA: Association for Computing Machinery, May 2019, pp. 1–8.

[8] J. O. Wobbrock, "Situationally-induced impairments and disabilities," in *Web Accessibility: A Foundation for Research*, Y. Yesilada and S. Harper, Eds. London: Springer, 2019, pp. 59–92.

[9] J. Goncalves, Z. Sarsenbayeva, N. van Berkel, C. Luo, S. Hosio, S. Risanen, H. Rintamäki, and V. Kostakos, "Tapping task performance on smartphones in cold temperature," *Interacting with Computers*, vol. 29, no. 3, pp. 355–367, May 2017.

[10] Z. Sarsenbayeva, J. Goncalves, J. García, S. Klakegg, S. Rissanen, H. Rintamäki, J. Hannu, and V. Kostakos, "Situational impairments to mobile interaction in cold environments," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. Heidelberg Germany: ACM, Sep. 2016, pp. 85–96.

[11] Z. Sarsenbayeva, N. van Berkel, E. Velloso, V. Kostakos, and J. Goncalves, "Effect of distinct ambient noise types on mobile interaction," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 2, pp. 82:1–82:23, Jul. 2018.

[12] T. Li, D. Peng, E. Velloso, A. Withana, K. Minamizawa, and Z. Sarsenbayeva, "Estimating the effects of ambient noise on mixed reality interaction," in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. New York, NY, USA: ACM, 2025.

[13] Z. Sarsenbayeva, N. van Berkel, D. Hettiachchi, W. Jiang, T. Dingler, E. Velloso, V. Kostakos, and J. Goncalves, "Measuring the effects of stress on mobile interaction," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 1, pp. 24:1–24:18, Mar. 2019.

[14] M. Goel, L. Findlater, and J. Wobbrock, "WalkType: Using accelerometer data to accomodate situational impairments in mobile touch screen text entry," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Austin Texas USA: ACM, May 2012, pp. 2687–2696.

[15] A. Ng, S. A. Brewster, and J. H. Williamson, "Investigating the effects of encumbrance on one- and two- handed interactions with mobile devices," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: Association for Computing Machinery, Apr. 2014, pp. 1981–1990.

[16] A. Ng, J. Williamson, and S. Brewster, "The effects of encumbrance and mobility on touch-based gesture interactions for mobile phones," in *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '15. New York, NY, USA: Association for Computing Machinery, Aug. 2015, pp. 536–546.

[17] Z. Sarsenbayeva, N. van Berkel, C. Luo, V. Kostakos, and J. Goncalves, "Challenges of situational impairments during interaction with mobile devices," in *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, ser. OzCHI '17. New York, NY, USA: Association for Computing Machinery, Nov. 2017, pp. 477–481.

[18] P. Ravasio, S. G. Schär, and H. Krueger, "In pursuit of desktop evolution: User problems and practices with modern desktop systems," *ACM Transactions on Computer-Human Interaction*, vol. 11, no. 2, pp. 156–180, Jun. 2004.

[19] F. Heller, D. Vanacken, E. Geurts, and K. Luyten, "Impact of situational impairment on interaction with wearable displays," in *22nd International Conference on Human-computer Interaction with Mobile Devices and Services*. Oldenburg Germany: ACM, Oct. 2020, pp. 1–5.

[20] T. Li, E. Velloso, A. Withana, and Z. Sarsenbayeva, "Estimating the effects of encumbrance and walking on mixed reality interaction," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Yokohama Japan: ACM, Apr. 2025, pp. 1–24.

[21] Y. Li, J. Liu, J. Huang, Y. Zhang, X. Peng, Y. Bian, and F. Tian, "Evaluating the effects of user motion and viewing mode on target selection in augmented reality," *International Journal of Human-Computer Studies*, vol. 191, p. 103327, Nov. 2024.

[22] T. Li, E. Velloso, A. Withana, and Z. Sarsenbayeva, "Weight-induced consumed endurance (wice): A model to quantify shoulder fatigue with weighted objects," in *The 38th Annual ACM Symposium on User Interface Software and Technology*. Busan, Republic of Korea: ACM, Sep. 2025.

[23] P. Liu, F. Zafar, and A. Badano, "The effect of ambient illumination on handheld display image quality," *Journal of Digital Imaging*, vol. 27, no. 1, pp. 12–18, Feb. 2014.

[24] H.-P. Huang, M. Wei, H.-C. Li, and L.-C. Ou, "Visual Comfort of Tablet Devices under a Wide Range of Ambient Light Levels," *Applied Sciences*, vol. 11, no. 18, p. 8679, Jan. 2021.

[25] S. Benedetto, A. Carbone, V. Drai-Zerbib, M. Pedrotti, and T. Baccino, "Effects of luminance and illuminance on visual fatigue and arousal during digital reading," *Computers in Human Behavior*, vol. 41, pp. 112–119, Dec. 2014.

[26] Z. Sarsenbayeva, N. van Berkel, W. Jiang, D. Hettiachchi, V. Kostakos, and J. Goncalves, "Effect of ambient light on mobile interaction," in *Human-Computer Interaction – INTERACT 2019*, D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler, and P. Zaphiris, Eds. Cham: Springer International Publishing, Sep. 2019, pp. 465–475.

[27] D.-S. Lee, Y.-H. Ko, I.-H. Shen, and C.-Y. Chao, "Effect of light source, ambient illumination, character size and interline spacing on visual performance and visual fatigue with electronic paper displays," *Displays*, vol. 32, no. 1, pp. 1–7, Jan. 2011.

[28] J. L. Gabbard, J. E. Swan, II, and D. Hix, "The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality," *Presence: Teleoperators and Virtual Environments*, vol. 15, no. 1, pp. 16–32, Feb. 2006.

[29] S. M. U. Arif, M. Brizzi, M. Carli, and F. Battisti, "Human reaction time in a mixed reality environment," *Frontiers in Neuroscience*, vol. 16, p. 897240, Aug. 2022.

[30] R. Liu, T. Li, and Z. Sarsenbayeva, "Understanding and addressing ambient illumination as a situational impairment for digital devices," in *The 37th Australian Conference on Human-Computer Interaction*, ser. OzCHI '25. Sydney, Australia: ACM, 2025.

[31] A. Erickson, K. Kim, G. Bruder, and G. F. Welch, "Exploring the limitations of environment lighting on optical see-through head-mounted displays," in *Symposium on Spatial User Interaction*. Virtual Event Canada: ACM, Oct. 2020, pp. 1–8.

[32] H. Choi, Y. Kim, and G. Kim, *Presence, Immersion and Usability of Mobile Augmented Reality*, Aug. 2019.

[33] X. Xie, Y. , Suihuai, and D. and Chen, "Effects of screen color mode and color temperature on visual fatigue under different ambient illuminations," *International Journal of Human–computer Interaction*, vol. 41, no. 2, pp. 821–833, Jan. 2025.

[34] G. W. Tigwell, D. R. Flatla, and R. Menzies, "It's not just the light: Understanding the factors causing situational visual impairments during mobile interaction," in *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*, ser. NordiCHI '18. New York, NY, USA: Association for Computing Machinery, Sep. 2018, pp. 338–351.

[35] K. Vasylevska, H. Yoo, T. Akhavan, and H. Kaufmann, "Towards eye-friendly VR: How bright should it Be?" in *2019 IEEE Conference on Virtual Reality and 3d User Interfaces (Vr)*, Mar. 2019, pp. 566–574.

[36] P. Kalra and V. and Karar, "Effect of screen switching and brightness on visual fatigue in AR environments," *IETE Technical Review*, vol. 40, no. 3, pp. 303–311, May 2023.

[37] D. Yu, Q. Zhou, B. Tag, T. Dingler, E. Velloso, and J. Goncalves, "Engaging participants during selection studies in virtual reality," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Mar. 2020, pp. 500–509.

[38] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement." *Journal of experimental psychology*, vol. 47, no. 6, p. 381, 1954.

[39] I. S. MacKenzie, "Fitts' law as a research and design tool in human-computer interaction," *Human-computer interaction*, vol. 7, no. 1, pp. 91–139, 1992.

[40] J. Bergström, T.-S. Dalsgaard, J. Alexander, and K. Hornbæk, "How to evaluate object selection and manipulation in VR? Guidelines from 20 years of studies," in *Proceedings of the 2021 CHI Conference on Human*

Factors in Computing Systems. Yokohama Japan: ACM, May 2021, pp. 1–20.

[41] F. Argelaguet and C. Andujar, "A survey of 3D object selection techniques for virtual environments," *Computers & Graphics*, vol. 37, no. 3, pp. 121–136, May 2013.

[42] J. Petford, M. A. Nacenta, and C. Gutwin, "Pointing all around you: Selection performance of mouse and ray-cast pointing in full-coverage displays," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Montreal QC Canada: ACM, Apr. 2018, pp. 1–14.

[43] J. D. Hincapié-Ramos, K. Ozacar, P. P. Irani, and Y. Kitamura, "GyroWand: IMU-based raycasting for augmented reality head-mounted displays," in *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*. Los Angeles California USA: ACM, Aug. 2015, pp. 89–98.

[44] J. Huang and B. Lee, "Modeling error rates in spatiotemporal moving target selection," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow Scotland Uk: ACM, May 2019, pp. 1–6.

[45] I. S. MacKenzie and R. W. Soukoreff, "Text entry for mobile computing: Models and methods, theory and practice," *Human–Computer Interaction*, vol. 17, no. 2-3, pp. 147–198, 2002.

[46] R. W. Soukoreff and I. S. MacKenzie, "Metrics for text entry research: An evaluation of msd and kspc, and a new unified error metric," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2003, pp. 113–120.

[47] M. R. Zhang, S. Zhai, and J. O. Wobbrock, "Text entry throughput: Towards unifying speed and accuracy in a single performance metric," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow Scotland Uk: ACM, May 2019, pp. 1–13.

[48] A. L. Sheppard and J. S. Wolffsohn, "Digital eye strain: Prevalence, measurement and amelioration," *BMJ Open Ophthalmology*, vol. 3, no. 1, Apr. 2018.

[49] J. Kim, Y. Sunil Kumar, J. Yoo, and S. Kwon, "Change of blink rate in viewing virtual reality with HMD," *Symmetry*, vol. 10, no. 9, p. 400, Sep. 2018.

[50] C. Blehm, S. Vishnu, A. Khattak, S. Mitra, and R. W. Yee, "Computer vision syndrome: A review," *Survey of Ophthalmology*, vol. 50, no. 3, pp. 253–262, May 2005.

[51] Y. Wang, G. Zhai, S. Zhou, S. Chen, X. Min, Z. Gao, and M. Hu, "Eye Fatigue Assessment Using Unobtrusive Eye Tracker," *IEEE Access*, vol. 6, pp. 55 948–55 962, 2018.

[52] J. K. Portello, M. Rosenfield, and C. A. Chu, "Blink rate, incomplete blinks and computer vision syndrome," *Optometry and Vision Science*, vol. 90, no. 5, p. 482, May 2013.

[53] A. Erickson, K. Kim, G. Bruder, and G. F. Welch, "Effects of dark mode graphics on visual acuity and fatigue with virtual reality head-mounted displays," in *2020 IEEE Conference on Virtual Reality and 3d User Interfaces (Vr)*, Mar. 2020, pp. 434–442.

[54] Q. Fan, J. Xie, Z. Dong, and Y. Wang, "The effect of ambient illumination and text color on visual fatigue under negative polarity," *Sensors (basel, Switzerland)*, vol. 24, no. 11, p. 3516, May 2024.

[55] 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*, vol. 3, no. 3, pp. 203–220, 1993.

[56] T. Hirzle, M. Cordts, E. Rukzio, J. Gugenheimer, and A. Bulling, "A critical assessment of the use of SSQ as a measure of general discomfort in VR head-mounted displays," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Yokohama Japan: ACM, May 2021, pp. 1–14.

[57] J. R. Hayes, J. E. Sheedy, J. A. Stelmack, and C. A. Heaney, "Computer Use, Symptoms, and Quality of Life," *Optometry and Vision Science*, vol. 84, no. 8, p. E738, Aug. 2007.

[58] C. Chu, M. Rosenfield, J. K. Portello, J. A. Benzoni, and J. D. Collier, "A comparison of symptoms after viewing text on a computer screen and hardcopy," *Ophthalmic and Physiological Optics*, vol. 31, no. 1, pp. 29–32, 2011.

[59] G. Charness, U. Gneezy, and M. A. Kuhn, "Experimental methods: Between-subject and within-subject design," *Journal of Economic Behavior & Organization*, vol. 81, no. 1, pp. 1–8, Jan. 2012.

[60] J. V. Bradley, "Complete counterbalancing of immediate sequential effects in a latin square design," *Journal of the American Statistical Association*, vol. 53, no. 282, pp. 525–528, Jun. 1958.

[61] L. Rao, Y. Park, A. Klement, C. Kang, E. Park, J. Zhuang, C. Kim, H.-Y. Chiu, R. Ning, D. Wang *et al.*, "5-1: Invited paper: Infinite display for meta quest pro," in *SID Symposium Digest of Technical Papers*, vol. 54, no. 1. Wiley Online Library, 2023, pp. 32–35.

[62] S. Wei, D. Bloemers, and A. Rovira, "A preliminary study of the eye tracker in the meta quest pro," in *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences*. Nantes France: ACM, Jun. 2023, pp. 216–221.

[63] A. Della Greca, A. Ilaria, C. Tucci, N. Frugieri, and G. Tortora, "A user study on the relationship between empathy and facial-based emotion simulation in virtual reality," in *Proceedings of the 2024 International Conference on Advanced Visual Interfaces*, 2024, pp. 1–9.

[64] E. Rocchi, A. Ferrarotti, and M. Carli, "A comparison of the meta quest pro and HTC vive focus 3 eye-tracking systems: Analysis of data accuracy and spatial precision," *IEEE Access*, vol. 13, pp. 71 995–72 010, 2025.

[65] Meta, "Eye tracking for movement sdk for unity," 2024, accessed: 2025-07-24. [Online]. Available: https://developers.meta.com/horizon/documentation/unity/move-eye-tracking

[66] 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*, vol. 17, p. 1173127, Mar. 2023.

[67] A. R. Bentivoglio, S. B. Bressman, E. Cassetta, D. Carretta, P. Tonali, and A. Albanese, "Analysis of blink rate patterns in normal subjects," *Movement Disorders*, vol. 12, no. 6, pp. 1028–1034, 1997.

[68] S. Patel, R. Henderson, L. Bradley, B. Galloway, and L. Hunter, "Effect of visual display unit use on blink rate and tear stability," *Optometry and Vision Science*, vol. 68, no. 11, p. 888, Nov. 1991.

[69] Meta, "Key considerations - mr design guideline," 2025. [Online]. Available: https://developers.meta.com/horizon/design/mr-design-guideline/

[70] T. J. Dube and A. S. Arif, "Text entry in virtual reality: A comprehensive review of the literature," in *International Conference on Human-Computer Interaction*. Springer, 2019, pp. 419–437.

[71] I. S. MacKenzie and R. W. Soukoreff, "Phrase sets for evaluating text entry techniques," in *CHI '03 Extended Abstracts on Human Factors in Computing Systems - CHI '03*. Ft. Lauderdale, Florida, USA: ACM Press, 2003, p. 754.

[72] M. R. Zhang, S. Zhai, and J. O. Wobbrock, "Text entry throughput: Towards unifying speed and accuracy in a single performance metric," in *Proceedings of the 2019 HI conference on human factors in computing systems*, 2019, pp. 1–13.

[73] I. L. Handbook, "Ies lighting handbook," *New York: Illuminating Engineering Society*, 1966.

[74] S. K. Bhandary, R. Dhakal, V. Sanghavi, and P. K. Verkicharla, "Ambient light level varies with different locations and environmental conditions: Potential to impact myopia," *PLOS One*, vol. 16, no. 7, p. e0254027, Jul. 2021.

[75] D. Nikitenko, J. Evans, D. R. Flatla, T. Driscoll, G. Quinlan, and K. Lukaszek, "Situational visual impairments on mobile devices - modeling the effects of bright outdoor environments," in *Proceedings of the 50th Graphics Interface Conference*, ser. GI '24. New York, NY, USA: Association for Computing Machinery, Sep. 2024, pp. 1–10.

[76] C. P. Halsted, "Brightness, luminance, and confusion," *Information display*, vol. 9, pp. 21–21, 1993.

[77] F. Zafar, M. Choi, J. Wang, P. Liu, and A. Badano, "Visual methods for determining ambient illumination conditions when viewing medical images in mobile display devices," *Journal of the Society for Information Display*, vol. 20, no. 3, pp. 124–132, 2012.

[78] Meta, "Horizon design: Display," 2024, accessed: 2025-07-27. [Online]. Available: https://developers.meta.com/horizon/design/display/

[79] Y. Yan, J. M. Han, and H.-J. Suk, "Optimal brightness of head mounted display for virtual reality contents," in *50th Congress of International Association of Color, AIC2017*. International Association of Color, 2017.

[80] A. T. Welford, "Fundamentals of skill." 1968.

[81] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[82] J. A. Stern, D. Boyer, and D. Schroeder, "Blink rate: a possible measure of fatigue," *Human factors*, vol. 36, no. 2, pp. 285–297, 1994.

[83] K. Caine, "Local standards for sample size at CHI," in *Proceedings of the 2016 Chi Conference on Human Factors in Computing Systems*. San Jose California USA: ACM, May 2016, pp. 981–992.

[84] H. Zhou, T. Kip, Y. Dong, A. Bianchi, Z. Sarsenbayeva, and A. Withana, "Juggling extra limbs: Identifying control strategies for supernumerary

multi-arms in virtual reality," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025, pp. 1–16.

[85] H. Zhou, T. Ayesh, C. Fan, Z. Sarsenbayeva, and A. Withana, "Coplayingvr: Understanding user experience in shared control in virtual reality," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 8, no. 3, pp. 1–25, 2024.

[86] C. Talens-Estarelles, A. Cerviño, S. García-Lázaro, A. Fogelton, A. Sheppard, and J. S. Wolffsohn, "The effects of breaks on digital eye strain, dry eye and binocular vision: Testing the 20-20-20 rule," *Contact Lens and Anterior Eye*, vol. 46, no. 2, p. 101744, 2023.

[87] J. Brailsford, F. Vetere, and E. Velloso, "Exploring the association between moral foundations and judgements of AI behaviour," in *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM, May 2024, pp. 1–15.

[88] R. McElreath, *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, 1st ed. Chapman and Hall/CRC, 2016.

[89] M. Schmettow, *New statistics for design researchers: A Bayesian Workflow in Tidy R*. Springer, 2021.

[90] S. Wadinambiarachchi, R. M. Kelly, S. Pareek, Q. Zhou, and E. Velloso, "The effects of generative AI on design fixation and divergent thinking," in *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM, May 2024, pp. 1–18.

[91] T. A. Snijders and R. Bosker, "Multilevel analysis: An introduction to basic and advanced multilevel modeling," 2011.

[92] P.-C. Bürkner, "**Brms** : An *R* package for bayesian multilevel models using *stan*," *Journal of Statistical Software*, vol. 80, no. 1, 2017.

[93] B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell, "Stan: A probabilistic programming language," *Journal of Statistical Software*, vol. 76, pp. 1–32, Jan. 2017.

[94] A. Vehtari, A. Gelman, D. Simpson, B. Carpenter, and P.-C. Bürkner, "Rank-normalization, folding, and localization: An improved r^ for assessing convergence of MCMC (with discussion)," *Bayesian Analysis*, vol. 16, no. 2, Jun. 2021.

[95] B. Lambert, "A student's guide to bayesian statistics," 2018.

[96] K. Nieuwenhuizen and J.-B. Martens, "Advanced modeling of selection and steering data: Beyond fitts' law," *International Journal of Human-Computer Studies*, vol. 94, pp. 35–52, Oct. 2016.

[97] R. Plamondon and A. M. Alimi, "Speed/accuracy trade-offs in target-directed movements," *Behavioral and brain sciences*, vol. 20, no. 2, pp. 279–303, 1997.

[98] J. Sweller, J. J. Van Merrienboer, and F. G. Paas, "Cognitive architecture and instructional design," *Educational psychology review*, vol. 10, no. 3, pp. 251–296, 1998.

[99] S. Shioiri, T. Sasada, and R. Nishikawa, "Visual attention around a hand location localized by proprioceptive information," *Cerebral Cortex Communications*, vol. 3, no. 1, p. tgac005, Jan. 2022.

[100] W. G. Darling, B. I. Zuck, L. Mikhail, and J. Adhikari, "Proprioceptive acuity for landmarks on the hand and digits," *Experimental Brain Research*, vol. 242, no. 2, pp. 491–503, Feb. 2024.

[101] M. Pölönen, M. Salmimaa, and J. Häkkinen, "Effect of ambient illumination level on perceived autostereoscopic display quality and depth perception," *Displays*, vol. 32, no. 3, pp. 135–141, Jul. 2011.

[102] T. T. M. Tran, S. Brown, O. Weidlich, S. Yoo, and C. Parker, "Wearable AR in everyday contexts: Insights from a digital ethnography of YouTube videos," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Yokohama Japan: ACM, Apr. 2025, pp. 1–18.

[103] Meta, "Hands - Meta Horizon Design Documentation," 2025, accessed: 2025-08-11. [Online]. Available: https://developers.meta.com/horizon/design/hands

[104] T. Li, P. Somarathne, Z. Sarsenbayeva, and A. Withana, "Ta-gnn: Physics inspired time-agnostic graph neural network for finger motion prediction," *arXiv preprint arXiv:2503.13034*, 2025.

[105] A. A. Abusharha, "Changes in blink rate and ocular symptoms during different reading tasks," *Clinical optometry*, pp. 133–138, 2017.

[106] J. D. Rodriguez, K. J. Lane, G. W. Ousler III, E. Angjeli, L. M. Smith, and M. B. Abelson, "Blink: characteristics, controls, and relation to dry eyes," *Current Eye Research*, vol. 43, no. 1, pp. 52–66, 2018.

[107] G. W. Tigwell, R. Menzies, and D. R. Flatla, "Designing for situational visual impairments: Supporting early-career designers of mobile content," in *Proceedings of the 2018 Designing Interactive Systems Conference*. Hong Kong China: ACM, Jun. 2018, pp. 387–399.

[108] K. Reinecke, D. R. Flatla, and C. Brooks, "Enabling Designers to Foresee Which Colors Users Cannot See," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. San Jose California USA: ACM, May 2016, pp. 2693–2704.

[109] D. Yu, X. Lu, R. Shi, H.-N. Liang, T. Dingler, E. Velloso, and J. Goncalves, "Gaze-supported 3D object manipulation in virtual reality," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Yokohama Japan: ACM, May 2021, pp. 1–13.

[110] F. L. I. Dutsinma, D. Pal, S. Funilkul, and J. H. Chan, "A systematic review of voice assistant usability: An ISO 9241–11 approach," *SN Computer Science*, vol. 3, no. 4, p. 267, May 2022.

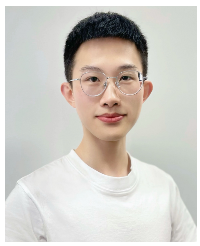
[111] A. Kuwahara, K. Nishikawa, R. Hirakawa, H. Kawano, and Y. Nakatoh, "Eye fatigue estimation using blink detection based on eye aspect ratio mapping(EARM)," *Cognitive Robotics*, vol. 2, pp. 50–59, Jan. 2022.

[112] B.-W. Hsu and M.-J. J. Wang, "Evaluating the effectiveness of using electroencephalogram power indices to measure visual fatigue," *Perceptual and motor skills*, vol. 116, no. 1, pp. 235–252, 2013.

[113] T. Hirzle, F. Fischbach, J. Karlbauer, P. Jansen, J. Gugenheimer, E. Rukzio, and A. Bulling, "Understanding, addressing, and analysing digital eye strain in virtual reality head-mounted displays," *ACM Trans. Comput.-hum. Interact.*, vol. 29, no. 4, pp. 33:1–33:80, Mar. 2022.

[114] J. P. Rolland, R. L. Holloway, and H. Fuchs, "Comparison of optical and video see-through, head-mounted displays," in *Telemanipulator and Telepresence Technologies*, vol. 2351. SPIE, 1995, pp. 293–307.

[115] S. Debernardis, M. Fiorentino, M. Gattullo, G. Monno, and A. E. Uva, "Text readability in head-worn displays: Color and style optimization in video versus optical see-through devices," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 1, pp. 125–139, Jan. 2014.



RUILIN LIU received the B.S. degree in computer science and software development from the University of Sydney, Australia. He is currently pursuing the B.S. (Hons.) degree in computer science. His research interests include human–computer interaction, extended reality, ubiquitous computing, and accessibility.

TINGHUI LI received the B.S. degree in data science and biochemistry, and the B.Adv.Stud. (Hons.) degree in computer science from the University of Sydney, Australia, in 2022 and 2023, respectively. She is currently pursuing the Ph.D. degree in human–computer interaction. Her research interests include human–computer interaction, extended reality, ubiquitous computing, and accessibility.

ZHANNA SARSENBAYEVA received the Ph.D. degree in computer science from the University of Melbourne, Australia, in 2020. From 2020 to 2022, she was a postdoctoral research fellow at the University of Melbourne. She is currently a Lecturer with the School of Computer Science, Faculty of Engineering, University of Sydney, Australia. Her research interests include human–computer interaction, accessibility, affective computing, and ubiquitous computing.