

Special Section on CEIG 2026

An analysis of gaze behavior under multisensory cognitive load in immersive environments[☆]

Jaime Bielsa^{id}, Jorge Pina^{id}, Ana Serrano^{id}, Daniel Martin^{id}*

Universidad de Zaragoza – I3A, Spain



ARTICLE INFO

Dataset link: <https://graphics.unizar.es/project/s/Bielsa2026-CLGaze/>

Keywords:

Gaze behavior
Cognitive load
Eye tracking
Visual attention

ABSTRACT

Virtual Reality (VR) enables immersive and realistic experiences in which complex multisensory interactions often impose elevated cognitive demands on users. Excessive levels of cognitive load (CL) have been shown to degrade performance and user experience, motivating the search for robust and non-intrusive indicators for CL monitoring in immersive environments. While recent approaches have leveraged physiological signals for this purpose, many of these signals are sensitive to motion and require complex sensor setups. Given the widespread integration of eye tracking in modern head-mounted displays, gaze behavior emerges as a promising alternative for objective CL assessment. In this work, we present a comprehensive analysis of gaze behavior and its relationship to cognitive load in immersive, task-oriented VR. Using a publicly available dataset collected in a multisensory visual search experience, we first examine how a range of gaze-derived markers, including fixations, saccades, eye eccentricity, and pupil dilation, vary across cognitive load conditions. We then investigate their relationship with biomarkers derived from complementary physiological signals, including electrocardiogram (ECG), electrodermal activity (EDA), and respiration, to better understand how gaze-based markers relate to broader physiological responses associated with cognitive load. Our results reveal consistent changes in oculomotor behavior under high cognitive load, characterized by reduced saccade amplitudes and eye eccentricity, together with increased pupil dilation. These patterns indicate a more centered and less exploratory visual behavior, consistent with attention tunneling under high task demands. In addition, we identify strong correlations between several gaze markers and the phasic component of electrodermal activity, a well-known indicator of mental effort. Together, these findings highlight the potential of gaze-based measures as lightweight and non-intrusive indicators of cognitive load, supporting the development of adaptive, user-aware VR applications.

1. Introduction

Virtual Reality (VR) experiences have lately become increasingly complex, allowing for new forms of multisensory interaction and immersion. This has led to the development of VR applications in domains such as training, education, and entertainment, where performance, comfort, and user experience are key. In these contexts, understanding the cognitive demands imposed on users is essential, as excessive mental workload can affect decision-making, slow task execution, and reduce overall engagement.

The Cognitive Load Theory [1] characterizes cognitive load (CL) as the amount of cognitive resources allocated to task processing, and prior work has shown that elevated levels of CL can degrade performance and hinder user experience [2]. Because of these effects,

being able to estimate users' cognitive load is key to the development of user-aware VR systems, as it enables adapting the experience to maintain performance, engagement, and comfort. For example, learning and training applications may dynamically adapt task difficulty, information density, or feedback to improve retention and skill acquisition. Real-time CL monitoring can also help anticipate overload-related effects such as fatigue or cybersickness, enabling proactive system adaptation. More broadly, integrating cognitive load assessment into VR design facilitates the creation of personalized experiences that balance effectiveness, immersion, and user well-being.

Traditionally, CL assessment has relied on subjective measurement methods, mostly self-reporting questionnaires (e.g., NASA-TLX [3])

[☆] This article is part of a Special issue entitled: 'CAG_CEIG 2026' published in Computers & Graphics.

* Correspondence to: C. de Mariano Esquillor Gomez, s/n, 50018 Zaragoza, Spain.

E-mail addresses: jaimebu@unizar.es (J. Bielsa), jpinac@unizar.es (J. Pina), anase@unizar.es (A. Serrano), danimis@unizar.es (D. Martin).

URLs: <https://jpinacolas.github.io/> (J. Pina), <https://ana-serrano.github.io/> (A. Serrano), <https://webdiis.unizar.es/~danimis/> (D. Martin).

where users provide numerical scores to describe their workload. Despite still being a widely adopted technique, this approach presents several limitations for immersive scenarios: questionnaires lack continuity in time, are often disruptive, and remain purely subjective. As a result, substantial effort has been dedicated to identifying objective markers as indicators of CL. Recent work has explored the correlations between physiological signals and CL variations, linking CL to changes in electrocardiogram (ECG), electrodermal activity (EDA), respiration, or neural signals derived from electroencephalography (EEG) [4,5]. While these modalities allow for continuous and less-disruptive measurement, they are also highly sensitive to confounding factors like physical movement and sensor noise [6]. This is particularly challenging, as most modern VR applications offer dynamic and interactive scenarios, involving frequent head and body motions, making it difficult to disambiguate the effective contributions of CL and movement.

Gaze has also been suggested to be a good proxy of cognitive demands [7,8] and interactions during task completion [8]. Several works have suggested that certain gaze features, such as fixations, saccades, and pupil dilation, are modulated by CL variations [9,10]. With the increasing integration of eye-tracking technology into modern VR headsets, driven in part by the widespread adoption of techniques such as foveated rendering, the availability of high-quality gaze data has grown substantially. The apparent correlation between gaze data and perceptual processes, together with the availability of large datasets and easy real-time gaze capture, has made gaze a compelling and promising candidate for objective CL estimation in VR, bypassing the motion sensitivity of most other physiological signals.

However, interpreting gaze data for CL assessment in VR remains challenging. The relationship between gaze-related features and CL is still not well established for immersive VR environments involving complex tasks, limiting their reliability as general indicators. Further, the impact of multisensory experiences, particularly the role of auditory input, has received limited attention, and existing studies do not always report consistent results regarding its influence on CL. In this work, we address the challenges aforementioned by performing a comprehensive analysis of gaze behavior under different CL conditions in multisensory VR, and by investigating its relationship with physiological signals. To this end, we leverage the publicly available dataset from Pina et al. [6] containing physiological, gaze, and subjective data for different CL levels. This dataset was collected in a visual-search experiment, where high CL was induced by a secondary auditory task performed simultaneously by users. Importantly, the experiment also included two search-area configurations, inducing different visual exploration and movement patterns. This allows us not only to study whether gaze markers vary with cognitive load, but also to better understand the extent to which these effects are influenced by the spatial demands of the task.

Our analysis focuses, first, on characterizing how eye-related features relate to cognitive load, with the goal of evaluating their potential as predictive markers. We extract and analyze gaze features, including fixations (periods where the gaze remains stable), saccades (rapid eye movements between fixations), eye eccentricity (angular deviation of gaze from the center of the field of view), and pupil dilation, and derive corresponding gaze markers including dwell time (accumulated fixation time); number and duration of fixations; number, duration, and amplitude of saccades; eye eccentricity; and head entropy. Through statistical analyses, we examine how these markers relate to both cognitive load and the search area conditions included in the dataset. In addition, we explore the direct relationship between gaze-based markers and other representative biomarkers extracted from the physiological signals available in the dataset, including mean heart rate (HR), mean respiration frequency (RF) and the phasic (EDA_{phasic}) and tonic (EDA_{tonic}) components of the electrodermal activity (EDA). This second analysis aims to assess the extent to which eye-related measures reflect broader changes in users' physiological states, beyond their direct association with cognitive load in the present task.

Our results reveal that search area has a dominant effect on gaze behavior, whereas cognitive load selectively affects only a subset of gaze markers. In particular, higher cognitive load leads to shorter saccade amplitudes, reduced eye eccentricity, and increased pupil dilation. These patterns suggest a more constrained and less exploratory oculomotor behavior under higher task demands, consistent with a narrowing of the effective field of view often referred to as attention *tunneling*. Further, we found that several gaze metrics exhibit meaningful relationships with physiological signals, particularly with the phasic component of the electrodermal activity, which is known to be a good indicator of the triggering of mental effort and task engagement [11–13]. Together, these findings reinforce the potential of integrated eye tracking in head-mounted displays to provide continuous and non-intrusive indicators of cognitive load, while also showing that careful interpretation and further research is needed to separate cognitive effects from task- and environment-driven gaze behavior.

We believe this work contributes to strengthening the study of gaze and visual attention as CL indicators in VR, and that it will foster future work to leverage gaze data for lightweight CL assessment methods in immersive experiences. We will make our complementary metrics extracted from the dataset publicly available upon acceptance.

2. Related work

2.1. Gaze behavior analysis

Gaze behavior reflects how visual attention is distributed across a scene. Early research leveraged eye-tracking systems to study gaze behavior when consuming visual content in traditional 2D displays [14–16]. These resorted to gaze-related features such as fixations, saccades, or saliency to characterize how users process visual information. Following work studied whether the insights obtained from traditional 2D displays applied to immersive virtual environments, like the seminal work from Sitzmann et al. [17]. While some patterns translated from one medium to the other, there were notable differences. VR environments are 360°, which limits the content that can fit inside the user's field of view (FoV); give the user complete control of the camera; often provide depth cues; and feature multisensory interactions. This results in more diverse gaze patterns, influenced by many confounding factors [18].

As research has shown, low-level and high-level visual properties of the virtual environment have a strong effect on gaze behavior. Factors such as luminance, contrast, color, size, and depth influence bottom-up (automatic, stimulus-driven) visual attention [19]. At the same time, viewing behavior is also driven top-down (voluntarily) by semantic aspects [20]. To better understand visual attention, and the interplay of such cues, several works have performed user studies to collect comprehensive gaze datasets under different viewing conditions and environments [21–23]. Other works have later leveraged this data to create models able to accurately predict gaze patterns in the form of saliency maps (time static) [24,25] or scanpaths (time dynamic) [26, 27]. Latest approaches have attempted to strive for more realistic applications, accounting for multisensory dynamic scenarios [28,29].

While these works mainly focused on free-exploration scenarios with no user interaction, subsequent research has studied gaze behavior in task-oriented experiences. This is especially relevant, as most real-world VR applications involve some sort of task or goal. Works like Malpica et al. [30] and Hu et al. [31] observed notable differences in gaze patterns while performing tasks. This suggests that the study of gaze behavior should be done consciously of the cognitive processes being performed by the user.

2.2. Cognitive load analysis

Cognitive Load Theory [1] defines CL as the amount of cognitive resources allocated to specific tasks over time. Traditionally, CL has been measured using subjective tools like self-reported user questionnaires [32], with the NASA Task Load Index (NASA-TLX) [3] being the most widely adopted. These methods collect CL levels numerically before and after the experience to detect subjective variations. However, such approaches present limitations: questionnaires lack temporal continuity, they rely on the subjective view of the user, and they usually require breaking immersion. These drawbacks are especially challenging for CL monitoring during VR experiences, where continuity and immersion are key.

Prior work has explored less disruptive methods for CL measurement [33]. These relied on user self-evaluation provided periodically during the experience. However, this poses only a partial solution: while less disruptive, the need of constant feedback from the user remains distracting, which can interfere with task performance and attention. Moreover, the data obtained remains purely subjective.

Beyond self-evaluation, research has investigated the validity of physiological signals as valid CL indicators, following studies linking CL variations with autonomic nervous system shifts [34]. Cardiovascular measures such as heart rate (HR) increments or heart rate variability (HRV) reductions [35], and alterations in the respiratory frequency can be attributed to levels of elevated CL [36]. On the same line, features derived from electrodermal activity (EDA) [4] and neurophysiological features obtained from electroencephalogram (EEG) have also been deemed as strong CL indicators [5].

Nonetheless, these physiological signals can also be influenced by other factors, like movement or sensor noise. Moderate exercise can substantially modulate electrocardiogram (ECG), EDA, or respiration patterns [37,38], making it hard to discern the real source of the variations. This challenge is specially pronounced in VR applications, which often involve interactive and dynamic scenarios. This motivates the exploration of additional complementary measurements that can increase the robustness of CL assessment, like eye and gaze-related data.

2.3. Gaze and cognitive load

Eye-related measurements have been studied as potential indicators of cognitive and emotional states [7,8]. Early work focused on the use of pupil size, correlating increments in pupil diameter to situations of elevated mental effort and high task difficulty [39]. This was later consolidated in pupillometry, reinforcing the phenomenon known as task-evoked pupillary response [40], which links pupil dynamics to cognitive processing through activations of the sympathetic system. Research was first conducted on traditional environments, but more recent work suggests that these findings also extend to immersive settings. In particular, pupil dilation has been consistently shown to increase with task difficulty and mental effort in VR scenarios [41–43], and is often identified as one of the most reliable indicators of cognitive load. However, it is also well known to be sensitive to external factors such as luminance, arousal, and fatigue, which complicates its interpretation in realistic environments.

Beyond pupillometry, other gaze-related features have also been explored as potential indicators of cognitive load, since eye movements carry rich information about task demand. With the increased availability of eye-tracking systems, features such as fixation duration, fixation rate [9], saccade amplitude and velocity [10], or blink frequency [44] have been tested as promising CL indicators. Classic studies suggest that longer fixations reflect increased information processing demands, while saccadic properties are more closely related to global scanning behavior and overall mental workload [45–47]. These trends are supported by more recent works in both traditional and immersive environments, where longer fixations, reduced saccadic

activity, and increased pupil size have been associated with higher cognitive demand [42,48,49].

The use of gaze measurements in CL analysis for VR applications presents advantages: many VR headsets incorporate eye-tracking systems, able to obtain real-time data during the VR experience. This enables easy, non-intrusive data capture. However, it also presents challenges: eye behavior is not only influenced by CL variations, but it can also differ highly on visual elements. Pupil diameter is known to be heavily influenced by changes in the illumination [50], and previous work has studied how gaze patterns can be affected by environmental visual features like color, size, or depth [51]. Furthermore, in immersive environments, recent work shows that gaze behavior is tightly coupled with task dynamics and user interaction: for instance, fixation and saccade patterns may reflect not only cognitive effort but also attentional and search strategies in complex 3D environments [30,48].

Despite these limitations, prior work has highlighted the benefits of using eye-related measurements for tasks like CL estimation. Multimodal approaches combining eye-tracking with other physiological signals have shown improved robustness, providing increased prediction accuracy and outperforming previous state-of-the-art works [52–54]. However, the relative contribution of each individual gaze feature remains unclear when large sets of metrics are considered [55].

Moreover, while some studies demonstrate the feasibility of using eye tracking for cognitive load detection in complex VR tasks, they often focus on predictive performance, with limited analysis of how individual gaze features relate to cognitive processes in dynamic spatiotemporal environments [56].

Overall, although a broad range of gaze metrics have been associated with cognitive load, most existing studies are conducted in non-immersive or weakly immersive settings, rely on simplified tasks, or do not fully account for the complex multisensory interactions and movements characteristic of modern VR experiences. In the context of immersive, task-oriented VR environments, the relationship between gaze behavior, cognitive load, and physiological responses remains insufficiently explored. Moreover, prior work has primarily focused on predictive performance, while providing limited insight into how eye-related markers relate to both cognitive load and other physiological signals in realistic VR scenarios. In this work, we address these gaps by jointly analyzing gaze features, cognitive load, and physiological measures in a multisensory VR setting, and by examining their implications for user-aware application design.

3. Dataset

Several works have collected datasets comprising gaze and physiological signals recorded under different CL conditions in VR [52,53,57]. However, the experimental conditions in which they were captured differ notably, and they use substantially different measuring methods and labeling for the CL levels (e.g., dynamic vs. static, purely visual vs. multisensory, discrete CL levels vs continuous scores). Moreover, most existing studies primarily focus on training predictive models, placing less emphasis on systematic analyses of how gaze and physiological markers relate to cognitive load in these datasets. For this work, we leverage the dataset collected by Pina et al. [6] as the publicly available VR dataset with the largest diversity of physiological signals and derived biomarkers, the one with the most exhaustive analyses on the correlation between cognitive load and physiological markers, and the addition of different levels of movement. We will from now on refer to it as *Pina* dataset.

Pina dataset contains physiological data from 35 users, out of which 12 were female and none identified as non-binary, not listed, or preferred not to disclose. The data was recorded during a visual-search experiment that was split into four conditions based on two factors: two CL conditions and two search area (SA) conditions. The CL level was separated into *low* and *high*, where the high CL level was elicited by an additional auditory task (pressing a key on the controller when

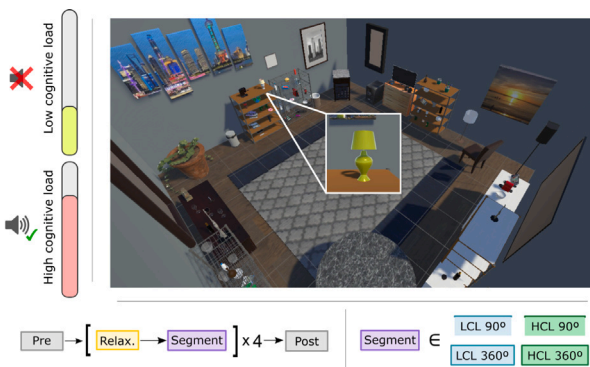


Fig. 1. Overview of the experiment from the dataset used. **Top right:** Virtual scene used during the experiment. **Left:** The experiment included two levels of cognitive load, induced by a secondary auditory task. **Bottom:** Overview of the experimental procedure, which consisted of a visual search task in which participants had to find a highlighted object as fast as possible (**inset**). *Source:* Figure adapted from the original work of Pina et al. [6].

hearing an odd number) performed simultaneously to the main visual search (searching for objects highlighted as golden). The increased mental demand of this multitasking scenario was validated through subjective measurements. The SA condition defined the region where the search task was performed (either 90° or 360°). While the 90° SA presented all the search targets directly in front of the user, inside their field of view, the 360° SA placed the targets around the user, forcing them to rotate to complete the search task. This is a particularly relevant aspect of the dataset, given the dynamic nature of most real VR applications and the influence that movement can have on some physiological signals [58,59]. The experiment followed a full within-subjects design, resulting from the combination of both SA and CL conditions previously mentioned, whose order was counterbalanced to prevent order biases. Additionally, relaxation periods were used both as an initial baseline and to reduce carry-over effects and ensure that the captured signals were not influenced by conditions adjacent in time. Fig. 1 shows the virtual scene and setup used in the experiment, together with an overview of the experimental procedure.

Pina dataset contains physiological data that was obtained with the Shimmer3R ECG Unit and the Shimmer3R GSR+ Unit, comprising Electrocardiogram (ECG), Electrodermal Activity (EDA), Respiration, and Photoplethysmogram (PPG). Additionally, it offers behavioral data including eye tracking, task performance, and subjective evaluations reported by each user. Eye-tracking data was recorded with the Varjo XR-4 built-in eye-tracker at a rate of 200 Hz. This included pupil and iris diameter, eye openness, and the eye forward vector. Pina et al. [6] performed a comprehensive analysis of some of the main physiological signals captured, extracting four representative biomarkers: mean heart rate (HR), mean respiration frequency (RF), median EDA tonic component (EDA_{tonic}), and median EDA phasic component (EDA_{phasic}). However, the eye-related information from the dataset remains unexplored. Therefore, our work aims to leverage the eye-tracking and pupil data available to perform a thorough analysis and explore correlations with the rest of physiological signals.

4. Data processing

Before extracting gaze markers and performing our analyses, we preprocess the raw eye-tracking data to ensure spatial consistency across participants. In particular, we aim to represent all gaze samples in a common reference frame, so that gaze behavior can be compared reliably across users and experimental conditions. Gaze data in *Pina* dataset was recorded using the integrated eye tracker from the Varjo

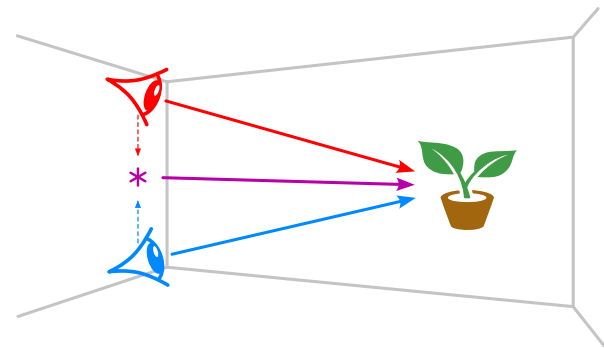


Fig. 2. Overview of the importance of adjusting our computation of gaze UV coordinates to each participant's height and position: different subjects (e.g., red and blue) look at the same object from different points of view. We compensate for their differences (purple asterisk) to normalize all gaze data into a unified coordinate system.

XR-4, which is represented in the HMD local coordinate system. However, local coordinate systems are specific to each user, considering potential differences in height and position shifts during the experiment (see Fig. 2). Therefore, we first aim to normalize all gaze data into a world coordinate system in which they are all comparable; i.e., a gaze UV system common to all users.

To do so, we carry out a ray-casting approach by simulating the experiment room as a prism with the same dimensions. Then, we cast gaze rays from the HMD position and orientation, and store the UV coordinates in which the ray hits the scene. The first component (U, horizontal) represents the longitude, and the second represents the latitude (V, vertical), both normalized between 0 and 1 (west to east and south to north, respectively).

Next, we aim to obtain reliable markers of gaze behavior. We first compute the gaze velocity (i.e., eye movement between frames) from the gaze UV. With such velocities, we can compute the two most prominent gaze movements: fixations and saccades. Fixations are large enough (approx. > 100 ms) spans of time that subjects spend focusing their gaze on an element in the space, while saccades relate to ballistic (approx. > 300 deg/s) gaze motion between fixations. To classify our gaze data into fixations and saccades, we resort to the Velocity-Threshold Identification (I-VT) algorithm [60], which leverages gaze velocity and eye openness to identify those phenomena. In this classification, the type of eye movement is not assumed if eye openness is under a threshold in eye openness (in this case, 0.75). We finally join contiguous equal-tagged frames into unique events. We then store the total number of fixations and their duration, and the number, duration, and angular amplitude for saccades.

4.1. Pupil dilation integrity

Pupil dilation (PD) is known to be influenced by both environmental factors, such as luminance, and cognitive load [41,61]. The original experimental scene (see Fig. 1) featured fixed, static lighting and no dynamic visual elements. Under these conditions, illumination remains constant, allowing us to consider pupil dilation as a candidate marker of cognitive load in our analysis.

However, some current HMDs apply automatic compensation to pupil dilation measurements (e.g., adjusting raw values based on eye rotations or other ocular factors), which may affect the interpretation of pupil changes. Whether such compensation mechanisms are implemented in the Varjo XR-4 headset is not publicly documented. To examine the behavior of the recorded pupil signal, we conducted a preliminary experiment in which participants tracked a highlighted item fixed relative to the headset, inducing controlled variations in eye eccentricity (i.e., angular displacement from the gaze point to the

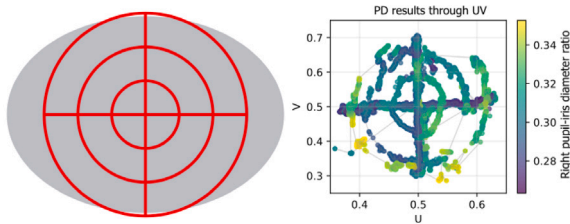


Fig. 3. **Left:** Target trajectories (red lines) followed by participants throughout our experiment on pupil compensation. Both uppermost and lowermost trajectories set the target in the FoV border. **Right:** Eye-tracking results for one participant. X and y axes represent the gaze direction, and color yields pupil diameter ratio (over the iris) in the right eye. PD grows as expected when the target gets away from the center (due to the eye being rotated), which suggests that the used headset did not compensate for pupil dilation automatically.

viewport center). If strong compensation were applied, pupil dilation values would remain invariant despite changes in gaze direction. We analyzed the pupil-to-iris diameter ratio, which provides a relative and bounded measure, as a function of gaze direction. The experiment was conducted with three participants.

We observed systematic variations in the pupil ratio across different eccentricity levels (including lateral gaze shifts; see Fig. 3 for one representative participant and the supplementary material for additional results). This suggests that the recorded signal preserves meaningful variations rather than being fully normalized by headset-level compensation. Therefore, we include pupil dilation in our analyses, while acknowledging that residual device-level processing effects and motion-related influences cannot be entirely ruled out, although their magnitude appears limited under the conditions of our experiment, as illustrated in Fig. 3.

5. Gaze analysis

Many works have studied the use of physiological markers (see Section 2.2) for cognitive load (CL) assessment. However, capturing such markers requires specific devices whose interpretation and analysis are not trivial. On the other hand, recent head-mounted displays (HMD) already incorporate integrated eye-trackers, which allow for the easy capture of gaze behavior during immersive experiences. Since gaze information can also be informative of CL, yet much remains unexplored in this regard, we focus on analyzing (i) the correlation between gaze and CL, and (ii) the correlation between gaze and other physiological markers, both from *Pina* dataset, to ideally study whether gaze can help better understand different levels of cognitive load with integrated eye-trackers, and without resorting to additional devices.

5.1. Visual representations of gaze

In our first analysis, we aim to find changes in visual behavior under different levels of cognitive load. We first focus on analyzing visual representations of gaze, which are commonly employed to visually inspect participants' gaze behavior. Several works [17,29] have resorted to saliency maps, which can be defined as heatmaps showing which parts of a scene are more conspicuous to viewers, or, similarly, which parts of the scene are more likely to draw viewers' attention. To generate such saliency maps, we leverage the previously computed *Gaze UV* component (see Section 4) to compute discrete fixation maps, which we later convolve with a Gaussian kernel with $\sigma = 1^\circ$, the estimated size of the human fovea [17]. We additionally compute radial plots using only the U (horizontal) component and 1D kernels for better visualization.

Fig. 4 shows both representations for different experimental conditions. Qualitatively, noticeable differences emerge in the saliency

distributions for the different search areas. In the smaller search area (90°), visual attention is strongly concentrated within the restricted field of view. In contrast, in the full search area (360°), attention is more broadly distributed, reflecting the need for head rotations and the wider spatial extent of the task.

We further compare whether saliency distributions differ between different levels of cognitive load for a fixed SA. We resort to the well-known metrics of Pearson's correlation (CC , which retrieves 0 for no correlation, and 1 or -1 for direct or inverse correlations, respectively), and Kullback–Leibler divergence (KLD , which measures divergence between distributions, i.e., the lower the better). Results for Pearson ($CC_{90^\circ} = 0.861 \pm 0.058$, $CC_{360^\circ} = 0.885 \pm 0.060$) show a strong similarity between those maps, and Kullback–Leibler results ($KLD_{90^\circ} = 0.420 \pm 0.240$, $KLD_{360^\circ} = 0.352 \pm 0.219$) yield no significant divergences. These results seem to suggest that there is no prominent variation or pattern in gaze distributions between low and high CL levels, and that if such differences exist, they might come from more low-level gaze characteristics.

5.2. Low-level gaze markers

Since we found no evident differences in visual representations of gaze between different CL levels within a fixed SA, we now focus on studying low-level characteristics of gaze. To do so, we analyze a large set of markers, including all the fixation and saccade markers previously introduced (Section 4) and pupil dilation ratio (Section 4.1). Beyond them, we also leverage our pre-processed data to compute additional gaze behavior markers such as dwell time (the total amount of time the subject is fixating), eye eccentricity (i.e., the angle between head and eye), and head orientation entropy (i.e., Shannon entropy of head orientation), which have been suggested to be good markers of behavioral changes when performing tasks within virtual environments [30]. We compute all these metrics per user and experimental condition (CL \times SA).

Since more than half of our data does not follow a normal distribution (Shapiro–Wilk test, $p \leq 0.05$), we resort to the Aligned Rank Transform (ART) for non-parametric factorial ANOVA [62–64]. We compute several models, so that each gaze metric is a dependent variable, SA and CL are fixed factors, and users are taken into account as random effects: $\text{gaze metric} \sim \text{cognitive load} * \text{search area} + (1 | \text{participant})$, where “*” denotes both the effects of each factor and their interaction. Within the original experiment, condition order was counterbalanced and relaxation periods were included before each condition, effectively mitigating carry-over and order effects [6].

We first observe that the search area (SA) has a dominant effect on gaze behavior (Table 1). Dwell time is lower, and there are more gaze events in 360° SA configurations, where fixations are shorter (half the time in 360° SA), and saccades, while of equal duration, are wider. In addition, pupil dilation increases. We interpret this behavior as a consequence of the larger spatial extent of the task in the 360° SA condition, which requires participants to distribute their attention more broadly and devote less time to individual locations. Furthermore, the 360° configuration actively encourages head movements to explore the environment, which may partially offload visual exploration from the eyes and further shape gaze behavior.

In contrast, the effect of cognitive load (CL) in visual behavior (Table 1) is less prominent, yet still affects some gaze markers. Our results reveal significant effects on saccade amplitude, eye eccentricity, and pupil dilation. Distributions for these metrics can be seen in the violin plots of Fig. 5. In high CL conditions, both saccade amplitude and eye eccentricity decrease, while pupil dilation increases. The decrease in saccade amplitude suggests that less space is scanned between fixations and across time, since the number and duration of both fixations and saccades are similar. This behavior is consistent with previous findings in tasks such as flight simulation and mixed-reality discovery tasks [42,65]. We hypothesize that, when cognitive resources are less

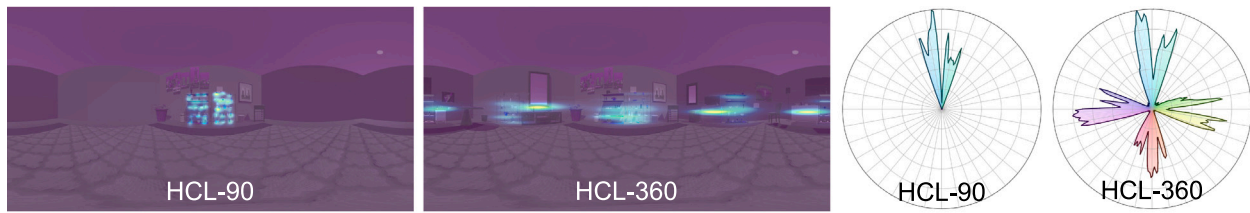


Fig. 4. Left: Saliency maps obtained from *Pina* dataset, averaged for all users by each search area, for high cognitive load settings. Gaze is located in those areas where items appear for each SA. Right: Radial plots built upon the same conditions. Note that visual attention is apparently sharper in the 90° SA condition, and much more spread in the 360° SA condition, which we hypothesize might be because of the latter forcing user to move their head and not always having the highlighted items popping up in their field of view.

constrained, users can afford to explore the environment more broadly, whereas under higher task demands (in this case, due to the secondary auditory task), participants reduce their exploration range, resulting in more localized scanning behavior. This reduction in explored space is accompanied by lower task performance under high CL conditions, as reported in the original experiment [6].

The observed decrease in eye eccentricity suggests that gaze becomes more centered and less exploratory. This aligns with the aforementioned behavior and is consistent with a narrowing of the effective field of view, commonly described as attentional tunneling [66,67]. This suggests that under higher cognitive load, users prioritize central visual information and reduce peripheral exploration. Importantly, this trend holds across both search areas, indicating that it is robust to environmental scale. Together, these results point to a shift towards more conservative and constrained visual sampling strategies under high CL [68], where exploration is reduced and attention becomes more focused. In addition, the increase in pupil dilation that we found might also be related to different work levels and increased stress [69]. Broadly speaking, these results seem to indicate that under high cognitive load, subjects stabilize their gaze by reducing their eye movements, which, in general terms, is consistent with attention research [70].

Interestingly, we did not observe significant differences in fixation duration, which has sometimes been reported as one strong CL indicator [47–49]. Some works have found that longer fixations are usually tied to higher processing demands. In our work, however, we did not find such differences. We believe that, given the nature of the original experiment, fixation behavior may be strongly driven by the task structure and spatial exploration, and that might be overshadowing more subtle cognitive effects.

Taken together, these findings suggest that not all gaze metrics are equally informative of cognitive load in immersive environments. While some features, such as pupil dilation, saccade amplitude, and eccentricity, show consistent sensitivity to CL in our study, other metrics appear to be more strongly influenced by exploration demands (in our case, the search area). This observation highlights the importance of carefully interpreting gaze-based measures in VR, and suggests that disentangling cognitive effects from environment- and task-driven behavior is important for their reliable use.

5.3. Gaze and physiological signals

The original work by Pina et al. [6] studied how different physiological signals related to increases in cognitive demands. They particularly studied four main markers: mean heart rate (HR), mean respiration frequency (RF), and the median tonic (EDA_{tonic}) and phasic (EDA_{phasic}) components of electrodermal activity. We refer the reader to the original work for further insights on these markers. However, whether such markers correlate with gaze, and if thus gaze can be a good predictor of any of those, and by extension, of cognitive load, remained unexplored. Therefore, we perform an exhaustive statistical analysis of the potential correlations between each of them.

Table 1

Results from the Aligned Rank Transform (ART) for non-parametric factorial Analysis of Variance (ANOVA) for different gaze metrics and experimental conditions (cognitive load, CL, and search area, SA). Significant differences are boldfaced, and effect sizes (partial η^2) are included for each of them. Search area (SA) strongly influences gaze behavior. Higher cognitive load (CL) yields shorter saccades, smaller eye eccentricity, and increased pupil dilation, which suggests efforts for gaze stabilization and increases in stress levels from participants. See main text for further discussion.

Metric	p-value		ES - Partial η^2	
	CL	SA	CL	SA
Dwell time	0.821	<0.001	0.001	0.690
Number of fixations	0.398	<0.001	0.007	0.650
Number of saccades	0.172	<0.001	0.020	0.280
Fixation median duration	0.288	<0.001	0.010	0.710
Saccade median duration	0.165	0.183	0.020	0.020
Saccade median amplitude	0.001	<0.001	0.110	0.540
Eye eccentricity	0.024	<0.001	0.050	0.250
Head orientation entropy	0.069	<0.001	0.030	0.810
Pupil dilation	0.001	<0.001	0.100	0.420

For this analysis, we focus only on the data of the 360° SA condition, as we consider it to be the most representative of immersive experiences, while also capturing important aspects such as user movement within the environment. We then run ANOVA tests in which each of the previously explored gaze metrics and the cognitive load level are used as fixed factors. For the random effects, we again account for variability between participants. However, instead of assuming that gaze affects all participants equally, we allow the effect of each gaze marker to vary across individuals by including participant-specific slopes. This allows the model to capture how strongly gaze relates to the physiological signal being evaluated for each participant, thereby accounting for individual differences in gaze–physiology coupling. This leads to models of the form $physiological\ signal \sim gaze\ marker * cognitive\ load + (1 + gaze\ marker | participant)$.

Results for the ANOVA tests (see Table 2) suggest relationships between some gaze markers and some physiological markers. First we found strong correlations between EDA_{phasic} and the number ($p = 0.0001$) and mean duration of fixations ($p = 0.0004$), the number ($p = 0.0007$) and median duration of saccades ($p = 0.0427$), eye eccentricity ($p = 0.0177$), and pupil dilation ($p = 0.0366$). In addition, we found more correlations between HR and the number ($p = 0.0001$) and mean duration of fixations ($p < 0.0001$), and the median saccade duration ($p = 0.0199$). The results for all the other markers and physiological signals are included in the supplementary material.

As for the significant correlations, several studies have suggested that markers derived from EDA and ECG can sometimes act as strong CL indicators [11–13]. While HR is generally considered less reliable (mainly due to motion sensitivity), EDA_{phasic} is a more robust marker that is able to capture CL variations from short-period stimuli. Theoretically, phasic EDA is known to be a good indicator of transient sympathetic arousal associated with mental effort and task engagement, which increases under higher cognitive load. The strong correlation

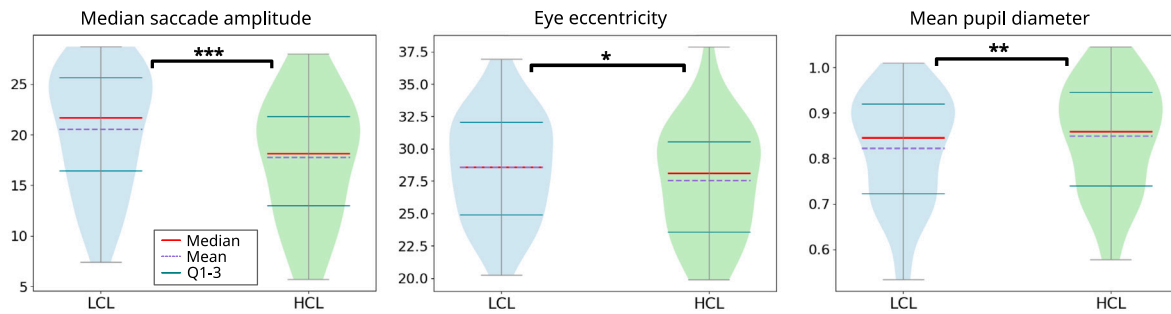


Fig. 5. Violin plots showing the distributions of saccade amplitude, eye eccentricity, and pupil dilation under different cognitive load levels (see Table 1). We detect the strongest dependencies over saccade amplitude, which suggests that, under high CL, subjects tend to stabilize their gaze by doing narrower saccades. This stabilization is also supported by a smaller eye eccentricity, which seems to relate to a more centered and less exploratory visual behavior, consistent with attention tunneling under high task demands (see main text). Lastly, pupil diameter also seems to relate to high CL conditions, probably because of its inherent relation to higher stress levels. (* for $p < 0.05$, ** for $p < 0.01$; *** for $p < 0.001$).

Table 2

Results (p-values) from the Analysis of Variance (ANOVA) between physiological and gaze markers within the 360° SA condition, where cognitive load is included as a factor. Significant values are boldfaced. EDA_{phasic} and more slightly HR , yield strong correlations. This suggests that eye-tracking, nowadays integrated in the latest head-mounted display, could be used as a proxy for cognitive demand triggers. See the main text for further details, and the supplementary material for the table with the statistical results for all statistical variables studied.

Gaze metric	Physiological marker			
	HR	$RESP$	EDA_{tonic}	EDA_{phasic}
Number of fixations	0.0001	0.2735	0.2772	0.0001
Fixation mean duration	<0.0001	0.2361	0.0167	0.0004
Number of saccades	0.4710	0.4289	0.6189	0.0007
Saccade median duration	0.0199	0.8392	0.8748	0.0427
Eye eccentricity	0.7692	0.2947	0.1174	0.0177
Pupil dilation	0.2420	0.2543	0.3547	0.0366

between so many gaze-based markers and EDA_{phasic} suggests that these gaze markers could be a good proxy of such activation mechanisms, and thus of moments in which cognitive demands are triggered.

6. Discussion and applications

We have performed thorough analyses to provide insights into how cognitive load (CL) manifests in gaze behavior within immersive visual-search tasks. Our results suggest that high CL scenarios resulted in changes in gaze behavior that align with the notion of a more centered and less exploratory visual behavior, similar to the theory of tunnel vision, where motor activity is reduced when cognitively burdened in order to preserve resources. The increased pupil dilation during high CL conditions also aligns with previous work in the field of pupillometry, although it also suggests that pupil dilation in immersive environments could reflect overall mental effort, instead of being tied exclusively to visual processing tasks.

We also found that several gaze markers were strongly related to HR and EDA_{phasic} . However, no significant correlation was found with EDA_{tonic} , which is also traditionally considered to be a strong CL indicator. These results point towards the potential of gaze-derived markers for the interpretation of specific forms of sympathetic activations. While EDA_{tonic} is known to reflect slow but constant variations in CL, EDA_{phasic} is capable of capturing fast alterations in cognitive demand.

Taken together, our results could indicate that, while some gaze markers like the saccade amplitude become prominent indicators only for longer experiences, other markers like the eye eccentricity and the pupil dilation might be capable of reflecting rapid CL variations in a

similar pattern to EDA_{phasic} . Overall, we believe that gaze and physiological signals should not be seen as competing explanations of cognitive load, but as complementary ones. Physiological measures reflect broader autonomic responses, while gaze shows how perception and eye movements adapt to task demands. Their partial overlap suggests that both types of signals coexist and together provide a clearer picture of cognitive load in immersive environments. Overall, we believe this correlation is very promising, as gaze-based indicators of cognitive load can be readily obtained with existing head-mounted displays through their integrated eye-trackers, ultimately enabling immediate and sensor-free practical applications.

6.1. Applications

Achieving accurate cognitive load (CL) estimation through gaze-related markers has direct implications for the development of VR applications. The current presence of integrated eye-tracking systems in commercial head-mounted displays would allow for lightweight CL monitoring during VR experiences, enabling adaptive rendering based on the current mental workload, training assistance, or adaptive VR experiences. Here we briefly discuss such applications and how our insights could foster each of them.

Workload-contingent rendering: Visual processing constitutes a permanent mental demand for the user, and increases with visual complexity. Obtaining real-time information of the CL level experienced by the user could allow for more efficient rendering strategies [71,72], such as lowering visual detail for scenarios of high CL, or increasing it whenever the user is not engaged enough. This could result in both an increase in rendering performance and a decrease in CL levels. For instance, our work suggest that whenever sensorimotor economy triggers, and eye movements decrease (i.e., smaller saccades and smaller eye eccentricity), and thus, if applications were to analyze them in real time, they could adaptively adjust their rendering strategies.

Training assistance: As VR applications have become increasingly realistic, virtual environments have been adopted for training purposes in multiple domains, such as medicine, navigation, simulation, among others [73]. Continuous monitoring of CL variations during these training experiences would provide valuable feedback on the user’s mental state at each part of the process. Additionally, recent work has attempted automatic difficulty regulation on training VR applications to improve learning [74,75]. Therefore, analogous to the previous application, our insights suggest that some particular gaze markers could enable easy adjustment of the difficulty of the training experience.

Cybersickness anticipation: Electrodermal activity (EDA), particularly its phasic component, has been repeatedly linked to autonomic arousal and discomfort responses in immersive environments [76,77]. The

correlations we observe between EDA_{phasic} and specific gaze markers suggest that oculomotor behavior could provide an early, non-intrusive proxy for emerging cybersickness. In practical terms, real-time detection of gaze stabilization patterns (such as reduced saccadic amplitude or constrained eye eccentricity) could signal the onset of physiological stress before subjective symptoms become explicit. VR systems could potentially leverage this information to proactively adjust motion, field-of-view, or other factors related to sickness, thereby mitigating discomfort.

7. Conclusions and future work

In this work, we introduce a thorough study of gaze behavior under different cognitive load levels induced by multisensory tasks in immersive environments. We extend a previously existing dataset and perform several studies correlating gaze, cognitive load, and physiological markers. We found changes in gaze markers likely correlated to a gaze stabilization strategy compatible with a less exploratory visual behavior in highly cognitive-demanding scenarios, and also found correlations between some gaze markers and physiological signals. We believe this can help establish some gaze markers' ability to discern rapid variations in cognitive load.

Many interesting future avenues remain open. For instance, there is room for exploring many other gaze markers, such as blinks, smooth pursuits, vestibulo-ocular reflexes, or microsaccades, among others. Some of them are thought to be good CL indicators, yet eye-tracker frequencies (usually under 200 Hz) and noise (usually $\sim 1-2^\circ$ of visual angle) hinder computing them. Similarly, exploring additional levels of cognitive load, either induced visually or by multisensory competition, could extend our insights and provide an even more generalizable baseline for CL assessment. A natural next step following our offline statistical study would be the potential predictive modeling of CL in VR through the use of gaze markers: if gaze markers are obtainable at a good frequency and with enough accuracy, developing lightweight models for CL prediction (which would enable many of the previously discussed applications) would benefit real-time adaptation of VR experiences. To this end, increasing diversity in participant pools and running additional experiments – therefore increasing the data available for such models – would enhance model generalizability and robustness. Overall, we believe that our work will foster the development and refinement of such models, enabling dynamically responsive VR experiences that can account for the cognitive state of each user individually.

CRedit authorship contribution statement

Jaime Bielsa: Formal analysis, Software, Visualization, Writing – original draft, Writing – review & editing. **Jorge Pina:** Formal analysis, Software, Visualization, Writing – original draft, Writing – review & editing. **Ana Serrano:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Daniel Martin:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by the European Union (ERC grant number 101220555, PROXIE). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. This work has also been funded by grant PID2022-141539NB-I00, funded by MICIU/AEI/10.13039/501100011033 and by ERDF, EU; and by the Aragon Institute for Engineering Research (I3A) through the Impulso program. J. Pina was supported by an FPI predoctoral grant (PRE2023-UZ-16).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.cag.2026.104613>.

Data availability

Our data is publicly available at <https://graphics.unizar.es/projects/Bielsa2026-CLGaze/>.

References

- [1] Sweller J. Cognitive load during problem solving: Effects on learning. *Cogn Sci* 1988;12(2):257–85. http://dx.doi.org/10.1207/s15516709cog1202_4.
- [2] Skulmowski A, Xu KM. Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educ Psychol Rev* 2022;34(1):171–96. <http://dx.doi.org/10.1007/s10648-021-09624-7>.
- [3] Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: *Advances in psychology*, vol. 52, Elsevier; 1988, p. 139–83. [http://dx.doi.org/10.1016/S0166-4115\(08\)62386-9](http://dx.doi.org/10.1016/S0166-4115(08)62386-9).
- [4] Buchwald M, Kupiński S, Bykowski A, Marcinkowska J, Ratajczyk D, Jukiewicz M. Electrodermal activity as a measure of cognitive load: a methodological approach. In: *Signal processing: algorithms, architectures, arrangements, and applications*. SPA, 2019, p. 175–9. <http://dx.doi.org/10.23919/SPA.2019.8936745>.
- [5] Yedukondalu J, Sunkara K, Radhika V, Kondaveeti S, Anumothu M, Murali Krishna Y. Cognitive load detection through EEG lead wise feature optimization and ensemble classification. *Sci Rep* 2025;15(1):842. <http://dx.doi.org/10.1038/s41598-024-84429-6>.
- [6] Pina J, Bernal-Berdun E, Martin D, Malpica S, Real C, Barquero A, Armanac-Julian P, Lazaro J, Martin-Yebra A, Masia B, Serrano A. A comprehensive analysis of the influence of cognitive load on physiological signals in virtual reality. In: *2025 IEEE international symposium on mixed and augmented reality*. ISMAR, IEEE Computer Society; 2025, p. 433–43. <http://dx.doi.org/10.1109/ISMAR67309.2025.00054>.
- [7] Wang Q, Yang S, Liu M, Cao Z, Ma Q. An eye-tracking study of website complexity from cognitive load perspective. *Decis Support Syst* 2014;62:1–10. <http://dx.doi.org/10.1016/j.dss.2014.02.007>.
- [8] Zagermann J, Pfeil U, Reiterer H. Measuring cognitive load using eye tracking technology in visual computing. In: *Proceedings of the sixth workshop on beyond time and errors on novel evaluation methods for visualization*. BELIV '16, New York, NY, USA: Association for Computing Machinery; 2016, p. 78–85. <http://dx.doi.org/10.1145/2993901.2993908>.
- [9] Liu J-C, Li K-A, Yeh S-L, Chien S-Y. Assessing perceptual load and cognitive load by fixation-related information of eye movements. *Sensors* 2022;22(3). <http://dx.doi.org/10.3390/s22031187>.
- [10] Stuyven E, Van der Goten K, Vandierendonck A, Claeys K, Crevits L. The effect of cognitive load on saccadic eye movements. *Acta Psychol* 2000;104(1):69–85. [http://dx.doi.org/10.1016/S0001-6918\(99\)00054-2](http://dx.doi.org/10.1016/S0001-6918(99)00054-2).
- [11] Chioffi F, Welsch R, Villa S, Chuang L, Mayer S. Virtual reality adaptation using electrodermal activity to support the user experience. *Big Data Cogn Comput* 2022;6(2). <http://dx.doi.org/10.3390/bdcc6020055>.
- [12] Setz C, Arnrich B, Schumm J, Marca R, Tröster G, Ehlert U. Discriminating stress from cognitive load using a wearable EDA device. *IEEE Trans Inf Technol Biomed* 2010;14:410–7. <http://dx.doi.org/10.1109/TITB.2009.2036164>.
- [13] Liu Y, Siqing D. Psychological stress level detection based on electrodermal activity. *Behav Brain Res* 2017;341. <http://dx.doi.org/10.1016/j.bbr.2017.12.021>.
- [14] Koch C, Ullman S. Shifts in selective visual attention: Towards the underlying neural circuitry. In: *Vaina LM, editor. Matters of intelligence: conceptual structures in cognitive neuroscience*. Dordrecht: Springer Netherlands; 1987, p. 115–41. http://dx.doi.org/10.1007/978-94-009-3833-5_5.

- [15] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans Pattern Anal Mach Intell* 1998;20(11):1254–9. <http://dx.doi.org/10.1109/34.730558>.
- [16] Krejtz K, Szmidt T, Duchowski AT, Krejtz I. Entropy-based statistical analysis of eye movement transitions. In: Proceedings of the symposium on eye tracking research and applications. ETRA '14, New York, NY, USA: Association for Computing Machinery; 2014, p. 159–66. <http://dx.doi.org/10.1145/2578153.2578176>.
- [17] Sitzmann V, Serrano A, Pavel A, Agrawala M, Gutierrez D, Masia B, Wetzstein G. Saliency in VR: How do people explore virtual environments? *IEEE Trans Vis Comput Graphics* 2018;24(4):1633–42. <http://dx.doi.org/10.1109/TVCG.2018.2793599>.
- [18] Yoo S, Jeong S, Kim S, Jang Y. Saliency-based gaze visualization for eye movement analysis. *Sensors* 2021;21(15). <http://dx.doi.org/10.3390/s21155178>.
- [19] Zhao Q, Koch C. Learning saliency-based visual attention: A review. *Signal Process* 2013;93(6):1401–7. <http://dx.doi.org/10.1016/j.sigpro.2012.06.014>, Special issue on Machine Learning in Intelligent Image Processing.
- [20] Haskins AJ, Mentch J, Botch TL, Robertson CE. Active vision in immersive, 360° real-world environments. *Sci Rep* 2020;10(1):14304. <http://dx.doi.org/10.1038/s41598-020-71125-4>.
- [21] Rai Y, Gutiérrez J, Le Callet P. A dataset of head and eye movements for 360 degree images. In: Proceedings of the 8th ACM on multimedia systems conference. MMSys '17, New York, NY, USA: Association for Computing Machinery; 2017, p. 205–10. <http://dx.doi.org/10.1145/3083187.3083218>.
- [22] Chao F-Y, Ozcinar C, Wang C, Zerman E, Zhang L, Hamidouche W, Deforges O, Smolic A. Audio-visual perception of omnidirectional video for virtual reality applications. In: 2020 IEEE international conference on multimedia & expo workshops. ICMEW, 2020, p. 1–6. <http://dx.doi.org/10.1109/ICMEW46912.2020.9105956>.
- [23] Bernal-Berdun E, Martin D, Malpica S, Perez PJ, Gutierrez D, Masia B, Serrano A. D-sav360: A dataset of gaze scanpaths on 360° ambisonic videos. *IEEE Trans Vis Comput Graphics* 2023;29(11):4350–60. <http://dx.doi.org/10.1109/TVCG.2023.3320237>.
- [24] Nguyen A, Yan Z, Nahrstedt K. Your attention is unique: Detecting 360-degree video saliency in head-mounted display for head movement prediction. In: Proceedings of the 26th ACM international conference on multimedia. 2018, p. 1190–8. <http://dx.doi.org/10.1145/3240508.3240669>.
- [25] Borji A. Saliency prediction in the deep learning era: Successes and limitations. *IEEE Trans Pattern Anal Mach Intell* 2019;43(2):679–700. <http://dx.doi.org/10.1109/TPAMI.2019.2935715>.
- [26] Le Meur O, Baccino T. Methods for comparing scanpaths and saliency maps: strengths and weaknesses. *Behav Res Methods* 2013;45(1):251–66. <http://dx.doi.org/10.3758/s13428-012-0226-9>.
- [27] Martin D, Serrano A, Bergman AW, Wetzstein G, Masia B. ScanGAN360: A Generative Model of Realistic Scanpaths for 360° Images. *IEEE Trans Vis Comput Graphics* 2022;28(5):2003–13. <http://dx.doi.org/10.1109/TVCG.2022.3150502>.
- [28] Coketek M, Imamoglu N, Ozcinar C, Erdem E, Ardem A. Leveraging frequency based salient spatial sound localization to improve 360° video saliency prediction. In: 2021 17th international conference on machine vision and applications. MVA, 2021, p. 1–5. <http://dx.doi.org/10.23919/MVA51890.2021.9511406>.
- [29] Bernal-Berdun E, Pina J, Vallejo M, Serrano A, Martin D, Masia B. AViSal360: Audiovisual saliency prediction for 360° video. In: 2024 IEEE international symposium on mixed and augmented reality. ISMAR, 2024, p. 1246–55. <http://dx.doi.org/10.1109/ISMAR62088.2024.00141>.
- [30] Malpica S, Martin D, Serrano A, Gutierrez D, Masia B. Task-dependent visual behavior in immersive environments: A comparative study of free exploration, memory and visual search. *IEEE Trans Vis Comput Graphics* 2023;29(11):4417–25. <http://dx.doi.org/10.1109/TVCG.2023.3320259>.
- [31] Hu Z, Bulling A, Li S, Wang G. FixationNet: Forecasting eye fixations in task-oriented virtual environments. *IEEE Trans Vis Comput Graphics* 2021;27(5):2681–90. <http://dx.doi.org/10.1109/TVCG.2021.3067779>.
- [32] Krieglstein F, Beege M, Rey GD, Sanchez-Stockhammer C, Schneider S. Development and validation of a theory-based questionnaire to measure different types of cognitive load. *Educ Psychol Rev* 2023;35(1):9. <http://dx.doi.org/10.1007/s10648-023-09738-0>.
- [33] Schmeck A, Opfermann M, van Gog T, Paas F, Leutner D. Measuring cognitive load with subjective rating scales during problem solving: differences between immediate and delayed ratings. *Instr Sci* 2015;43(1):93–114. <http://dx.doi.org/10.1007/s11251-014-9328-3>.
- [34] Ayres P, Lee JY, Paas F, Van Merriënboer JGG. The validity of physiological measures to identify differences in intrinsic cognitive load. *Front Psychol* 2021;12:702538. <http://dx.doi.org/10.3389/fpsyg.2021.702538>.
- [35] Ahmadi M, Michalka SW, Lenzoni S, Ahmadi Najafabadi M, Bai H, Sumich A, Wuensche B, Billingham M. Cognitive load measurement with physiological sensors in virtual reality during physical activity. In: Proceedings of the 29th ACM symposium on virtual reality software and technology. 2023, p. 1–11. <http://dx.doi.org/10.1145/3611659.3615704>.
- [36] Grassmann M, Vlemingx E, von Leopoldt A, Mittelstädt JM, Van den Bergh O. Respiratory changes in response to cognitive load: A systematic review. *Neural Plast* 2016;2016(1):8146809. <http://dx.doi.org/10.1155/2016/8146809>.
- [37] Posada-Quintero HF, Reljin N, Mills C, Mills I, Florian JP, VanHeest JL, Chon KH. Time-varying analysis of electrodermal activity during exercise. *PLoS One* 2018;13(6):1–12. <http://dx.doi.org/10.1371/journal.pone.0198328>.
- [38] Malakan Rad EA, Karimi M, Momtazmanesh S, Shabanian R, Saatchi M, Asbagh PA, Zeinaloo AA. Exercise-induced electrocardiographic changes after treadmill exercise testing in healthy children: A comprehensive study. *Ann Pediatr Cardiol* 2021;14(4). http://dx.doi.org/10.4103/apc.apc.254_20.
- [39] Gavas R, Chatterjee D, Sinha A. Estimation of cognitive load based on the pupil size dilation. In: 2017 IEEE international conference on systems, man, and cybernetics. SMC, 2017, p. 1499–504. <http://dx.doi.org/10.1109/SMC.2017.8122826>.
- [40] Beatty J. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol Bull* 1982;91(2):276. <http://dx.doi.org/10.1037/0033-2909.91.2.276>.
- [41] Lee JY, de Jong N, Donkers J, Jarodzka H, van Merriënboer JGG. Measuring cognitive load in virtual reality training via pupillometry. *IEEE Trans Learn Technol* 2024;17:704–10. <http://dx.doi.org/10.1109/TLT.2023.3326473>.
- [42] Hebbbar PA, Vinod S, Shah AK, Pashilkar AA, Biswas P. Cognitive load estimation in VR flight simulator. *J Eye Mov Res* 2023;15(3):23. <http://dx.doi.org/10.16910/jemr.15.3.11>.
- [43] Nenna F, Zanardi D, Gamberini L. Enhanced interactivity in VR-based telerobotics: an eye-tracking investigation of human performance and workload. *Int J Hum-Comput Stud* 2023;177:103079. <http://dx.doi.org/10.1016/j.ijhcs.2023.103079>.
- [44] Ledger H. The effect cognitive load has on eye blinking. *Plymouth Stud Sci* 2013;6(1):206–23, doi:<http://hdl.handle.net/10026.1/14015>.
- [45] Just MA, Carpenter PA. Eye fixations and cognitive processes. *Cogn Psychol* 1976;8(4):441–80. [http://dx.doi.org/10.1016/0010-0285\(76\)90015-3](http://dx.doi.org/10.1016/0010-0285(76)90015-3).
- [46] May JG, Kennedy RS, Williams MC, Dunlap WP, Brannan JR. Eye movement indices of mental workload. *Acta Psychol* 1990;75(1):75–89. [http://dx.doi.org/10.1016/0001-6918\(90\)90067-P](http://dx.doi.org/10.1016/0001-6918(90)90067-P).
- [47] Bläsing D, Bornewasser M. Influence of increasing task complexity and use of informational assistance systems on mental workload. *Brain Sci* 2021;11(1):102. <http://dx.doi.org/10.3390/brainsci11010102>.
- [48] Khan R, Vermooij J, Salvatori D, Hierck BP. Assessing cognitive load using EEG and eye-tracking in 3-D learning environments: A systematic review. *Multimodal Technol Interact* 2025;9(9):99. <http://dx.doi.org/10.3390/mti9090099>.
- [49] Barrett RCA, Poe R, O'Camb JW, Woodruff C, Harrison SM, Dolguikh K, Chuong C, Klassen AD, Zhang R, Joseph RB, et al. Comparing virtual reality, desktop-based 3D, and 2D versions of a category learning experiment. *PLoS One* 2022;17(10):e0275119. <http://dx.doi.org/10.1371/journal.pone.0275119>.
- [50] Pan Q, Zhang S, Bian J, Luo T, Wei M, Xue P, Xie J, Liu J. Pupil dynamic adaptation in response to environmental luminance: A controlled laboratory study. *Build Environ* 2026;287:113887. <http://dx.doi.org/10.1016/j.buildenv.2025.113887>.
- [51] Zelinsky G, Zhang W, Yu B, Chen X, Samaras D. The role of top-down and bottom-up processes in guiding eye movements during visual search. In: Weiss Y, Schölkopf B, Platt J, editors. *Advances in neural information processing systems*. vol. 18, MIT Press; 2005. <http://dx.doi.org/10.5555/2976248.2976445>.
- [52] Wei J, Siegel E, Sundaramoorthy P, Gomes A, Zhang S, Vankipuram M, Smathers K, Ghosh S, Horii H, Bailenson J, Ballagas RT. Cognitive load inference using physiological markers in virtual reality. In: 2025 IEEE conference virtual reality and 3D user interfaces. VR, 2025, p. 759–69. <http://dx.doi.org/10.1109/VR59515.2025.00098>.
- [53] Setu JN, Le JM, Kundu RK, Giesbrecht B, Höllerer T, Hoque KA, Desai K, Quarles J. Predicting and explaining cognitive load, attention, and working memory in virtual multitasking. *IEEE Trans Vis Comput Graphics* 2025;31(5):3014–24. <http://dx.doi.org/10.1109/TVCG.2025.3549850>.
- [54] Pavel MI, Mahmud MR, Setu JN, Desai K, Quarles J. PatchFusionVR: Multitask prediction of user gaze, reaction time, and cognitive load in virtual reality from multimodal signals. In: Proceedings of the 2025 31st ACM symposium on virtual reality software and technology. VRST '25, New York, NY, USA: Association for Computing Machinery; 2025. <http://dx.doi.org/10.1145/3756884.3766053>.
- [55] Shojaeizadeh M, Djamasi S, Paffenroth RC, Trapp AC. Detecting task demand via an eye tracking machine learning system. *Decis Support Syst* 2019;116:91–101. <http://dx.doi.org/10.1016/j.dss.2018.10.012>.
- [56] Nasri M, Kosa M, Chukoskie L, Moghaddam M, Harteveld C. Exploring eye tracking to detect cognitive load in complex virtual reality training. In: 2024 IEEE international symposium on mixed and augmented reality adjunct. ISMAR-adjunct, IEEE; 2024, p. 51–4. <http://dx.doi.org/10.1109/ISMAR-Adjunct64951.2024.00022>.
- [57] Bhat SS, Dobbins C, Dey A, Sharma O. Multi-modal classification of cognitive load in a VR-based training system. In: 2023 IEEE international symposium on mixed and augmented reality. ISMAR, 2023, p. 503–12. <http://dx.doi.org/10.1109/ISMAR59233.2023.00065>.
- [58] Reis P, Hebenstreit F, Gabsteiger F, von Tscharnar V, Lochmann M. Methodological aspects of EEG and body dynamics measurements during motion. *Front Hum Neurosci* 2014;Volume 8 - 2014. <http://dx.doi.org/10.3389/fnhum.2014.00156>.
- [59] Fernandez Rojas R, Brown N, Waddington G, Goecke R. A systematic review of neurophysiological sensing for the assessment of acute pain. *Npj Digit Med* 2023;6(1):76. <http://dx.doi.org/10.1038/s41746-023-00810-1>.

- [60] Salvucci DD, Goldberg JH. Identifying fixations and saccades in eye-tracking protocols. In: Proceedings of the 2000 symposium on Eye tracking research & applications. 2000, p. 71–8. <http://dx.doi.org/10.1145/355017.355028>.
- [61] Palinko O, Kun AL. Exploring the effects of visual cognitive load and illumination on pupil diameter in driving simulators. In: Proceedings of the symposium on eye tracking research and applications. ETRA '12, 2012, p. 413–6. <http://dx.doi.org/10.1145/2168556.2168650>.
- [62] Kim D, Kim J, Shin J-E, Yoon B, Lee J, Woo W. Effects of virtual room size and objects on relative translation gain thresholds in redirected walking. In: 2022 IEEE conference on virtual reality and 3D user interfaces. VR, IEEE; 2022, p. 379–88. <http://dx.doi.org/10.1109/VR51125.2022.00057>.
- [63] Wobbrock JO, Findlater L, Gergle D, Higgins JJ. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In: Proceedings of the SIGCHI conference on human factors in computing systems. 2011, p. 143–6. <http://dx.doi.org/10.1145/1978942.1978963>.
- [64] Martin D, Sun X, Gutierrez D, Masia B. A study of change blindness in immersive environments. *IEEE Trans Vis Comput Graphics* 2023;29(5):2446–55. <http://dx.doi.org/10.1109/TVCG.2023.3247102>.
- [65] López P, Bernardos AM, Casar JR. Eye-tracking analysis for cognitive load estimation in wearable mixed reality. In: Proceedings of the 2024 ACM symposium on spatial user interaction. 2024, p. 1–2. <http://dx.doi.org/10.1145/3677386.368888>.
- [66] Reimer B. Impact of cognitive task complexity on drivers' visual tunneling. *Transp Res Rec* 2009;2138(1):13–9. <http://dx.doi.org/10.3141/2138-03>.
- [67] Williams LJ. Tunnel vision or general interference? Cognitive load and attentional bias are both important. *Am J Psychol* 1988;171–91. <http://dx.doi.org/10.2307/1422833>.
- [68] Pomplun M, Reingold EM, Shen J. Investigating the visual span in comparative search: The effects of task difficulty and divided attention. *Cognition* 2001;81(2):B57–67. [http://dx.doi.org/10.1016/S0010-0277\(01\)00123-8](http://dx.doi.org/10.1016/S0010-0277(01)00123-8).
- [69] Hess EH, Polt JM. Pupil size in relation to mental activity during simple problem-solving. *Science* 1964;143(3611):1190–2. <http://dx.doi.org/10.1126/science.143.3611.1190>.
- [70] Pannasch S, Velichkovsky BM. Distractor effect and saccade amplitudes: Further evidence on different modes of processing in free exploration of visual images. *Vis Cogn* 2009;17(6–7). <http://dx.doi.org/10.1080/13506280902764422>.
- [71] Krajancich B, Kellnhöfer P, Wetzstein G. Towards attention-aware foveated rendering. *ACM Trans Graph* 2023;42(4). <http://dx.doi.org/10.1145/3592406>.
- [72] Hu X, Wu X, Ma M, Xu X, Gu Y, Wang G, Xu Y, Meng X, Wang L. Efficient VR rendering: Survey on foveated, stereo, cloud, and low-power rendering techniques. *Virtual Real Intell Hardw* 2025;7(5):421–52. <http://dx.doi.org/10.1016/j.vrih.2025.08.003>.
- [73] Martin D, Malpica S, Gutierrez D, Masia B, Serrano A. Multimodality in VR: A survey. *ACM Comput Surv* 2022;54(10s):1–36. <http://dx.doi.org/10.1145/3508361>.
- [74] Matam B, Mann A, Studer K, Gabbianelli C, Castelo S, Liu J, Silva C, Turakhia D. CLAD-Vr: Cognitive load-based adaptive training for machining tasks in virtual reality. In: 2025 IEEE international symposium on mixed and augmented reality adjunct. ISMAR-adjunct, 2025, p. 432–5. <http://dx.doi.org/10.1109/ISMAR-Adjunct68609.2025.00086>.
- [75] Nasri M. Towards intelligent VR training: A physiological adaptation framework for cognitive load and stress detection. In: Proceedings of the 33rd ACM conference on user modeling, adaptation and personalization. UMAP '25, 2025, p. 419–23. <http://dx.doi.org/10.1145/3699682.3727575>.
- [76] Gavgani AM, Nesbitt KV, Blackmore KL, Nalivaiko E. Profiling subjective symptoms and autonomic changes associated with cybersickness. *Auton Neurosci* 2017;203:41–50. <http://dx.doi.org/10.1016/j.autneu.2016.12.004>.
- [77] Yang M, Huang J. Phasic and tonic electrodermal activity in emotional responses to acoustic environments. *J Environ Psychol* 2024;96:102282. <http://dx.doi.org/10.1016/j.jenvp.2024.102282>.