Effects of Transient Loss of Vision on Head and Eye Movements during Visual Search in a Virtual Environment

Erwan David *[ID]**, Julia Beitner** [ID] **and Melissa Le-Hoa Võ** [ID]

Scene Grammar Lab, Department of Psychology, Theodor-W.-Adorno-Platz 6,
Johann Wolfgang-Goethe-Universität, 60323 Frankfurt, Germany; beitner@psych.uni-frankfurt.de (J.B.);
mlvo@psych.uni-frankfurt.de (M.L.-H.V.)
* Correspondence: david@psych.uni-frankfurt.de

Abstract: Central and peripheral fields of view extract information of different quality and serve different roles during visual tasks. Past research has studied this dichotomy on-screen in conditions remote from natural situations where the scene would be omnidirectional and the entire field of view could be of use. In this study, we had participants looking for objects in simulated everyday rooms in virtual reality. By implementing a gaze-contingent protocol we masked central or peripheral vision (masks of 6 deg. of radius) during trials. We analyzed the impact of vision loss on visuo-motor variables related to fixation (duration) and saccades (amplitude and relative directions). An important novelty is that we segregated eye, head and the general gaze movements in our analyses. Additionally, we studied these measures after separating trials into two search phases (scanning and verification). Our results generally replicate past on-screen literature and teach about the role of eye and head movements. We showed that the scanning phase is dominated by short fixations and long saccades to explore, and the verification phase by long fixations and short saccades to analyze. One finding indicates that eye movements are strongly driven by visual stimulation, while head movements serve a higher behavioral goal of exploring omnidirectional scenes. Moreover, losing central vision has a smaller impact than reported on-screen, hinting at the importance of peripheral scene processing for visual search with an extended field of view. Our findings provide more information concerning how knowledge gathered on-screen may transfer to more natural conditions, and attest to the experimental usefulness of eye tracking in virtual reality.

Keywords: visual attention; visual field loss; gaze-contingent protocol; visual search; virtual reality

1. Introduction

In the natural world, we make a succession of fixations by which we process information from the central and peripheral visual fields to explore our environment, usually in pursuit of a behavioral goal. In this study, we aimed at understanding the impact of vision loss during a search task carried out in a virtual environment replicating natural, everyday rooms. Do results from on-screen experiments translate fully to a situation where participants can use more than a screen-worth of visual field, while moving their neck and body along with their eyes? The advent of virtual reality (VR) head-mounted displays with low frame rate latency, acceptable graphic resolution and embedded eye tracking systems, now allows researchers to implement protocols that would not be possible with mobile eye tracking in the real world. We took this opportunity to study how head movements serve visual search and to measure visuo-motor behaviors across search phases.

The central and peripheral fields of view differ by the density of photoreceptor cells [1,2], size of receptive fields [3–5], as well as their cortical representation [6,7]. As a result visual information quality

drops as a function of the eccentricity to the fovea [8,9]. Despite its limitations, peripheral vision helps build a global representation of scenes [10,11]. The work of Boucart et al. [11–13] served to demonstrate that scene categorization stays above chance-level even at high visual eccentricities (70 deg. and farther). Functionally, peripheral vision serves to explore a scene, while central vision is used to sequentially analyze regions of interest. This manifests itself in different types of eye movement control, e.g., local and global processing [14–16], where global processing is exploratory and characterized by long saccades and short fixation durations, whereas local processing is reflected through short saccades and longer fixation durations required to analyze and identify objects or regions of interest. Object features, such as color, help localize search targets in natural scene peripherally [17], whereas, central vision exploits color to identify targets; contrary to textbook knowledge, color perception is indeed effective for scene processing at mid to high eccentricities (approximately 50 deg.) [18,19].

The dual role of the visual field is further demonstrated by the visuo-motor capabilities of participants experiencing transient (e.g., [20–24]) and pathological [25,26] central field defects. The complex consequences of real vision losses notwithstanding, lack of central vision results in diminished performances compared with normal sighted observers. Similarly, transient peripheral vision loss strongly impacts eye movements and search performances [20–24]. The current understanding is that the entire visual field is necessary to carry out visual tasks, but central and peripheral vision most likely have different roles when interacting with naturalistic environments.

Real-world scenes and everyday goals deviate from laboratory experiments in many aspects, for instance with regard to the complexity of the situations [27]. So far there is evidence that overall, visual search findings gathered with computer set-ups translate well to real-world situations [28–31]. In this study, we constructed VR environments by replicating real-world indoor scenes. Combined with the ability to freely move around thanks to the head-mounted display, we intended to approach quasi-natural conditions. In VR and mobile eye tracking experiments, neck and body movements need not be restricted; as a result, observers use the full range of movement permitted by the human body to explore their surroundings.

Due to the addition of head and body movements, tracking eye movements in VR environments reveals strong differences compared to eye movements recorded in restricted on-screen paradigms. The field of fixation is the area within the field of view where fixations can be made only with the eyes, when the head is restricted [32–34]. In cases where the head is unrestricted this becomes the practical field of fixation [34,35]. We do not make use of the entire field of fixation in normal situations, thanks to head movements permitting large saccades without resorting to long eye rotations [36–40]. Head movements are synergistic when they contribute to saccades [39] and compensatory when they help stabilize gaze in the scene during fixations (vestibulo-ocular [39,41] and opto-kynetic [42] (Chapter 2) [43] responses). During synergistic movements, head movements are recorded even during small saccades [39,44–46]. In a free-viewing study in VR [47], we observed that head movements were much less complex compared to eye movements; possibly under the control of behavioral goals related to exploration, whereas eye movements served to analyze regions of interest.

This complex interplay of head, body, and eye movements allows for efficient processing of the visual world that surrounds us. So what role does central vs. peripheral processing play during unrestricted, real-world search of 3D environments? For this purpose, we employed a gaze-contingent protocol to remove part of the field of view in real-time. This protocol [48,49] has been extensively used to segregate central and peripheral visions. The principle idea is to present contents of different natures in the different fields of view. A moving mask (central masking) or moving window (peripheral masking) is centered on the gaze position sampled by an eye tracking device. In the 1970's, it was introduced to study reading [50,51]. It has since been used in conjunction with natural scene viewing tasks to study eye movements, visual perception and visual attention (free-viewing [24,47], scene exploration [20,23,52] and visual search [17,21,53,54]).

While these studies have been essential in furthering our understanding of central vs. peripheral processing, they have done so in very restricted, 2D scenarios. Employing modern VR devices,

however, allows researchers to study visual attention in quasi-natural situations. Advancing to VR resolves some shortcomings of gaze-contingent protocols in 2D: the portion of the field of view excited is significantly increased, masking can be applied independently per eye, and neck and body movements are unrestricted. An eye tracker system embedded in a VR headset allows the implementation of protocols that would not be possible with mobile eye tracking devices, for instance, manipulating the visual fields during natural viewing. Virtual reality devices have significant advantages over mobile eye-tracking. Apart from the possibility to fully manipulate the content of the field of view they also offer the opportunity to generate many complex scenes and easily manipulate them according to various experimental conditions. Additional, contrary to most mobile eye-tracking studies we here report gaze data as the combination of head and eye movements, instead of eye movements alone which would withhold important information, such as synergistic and compensatory head movements (e.g., [39]). Some VR devices possess two cameras on the front that can project a stereoscopic view of the real world inside the headset, allowing a gaze-contingent modification of a projection of a real environment in real-time. On the other hand, gaze-contingent protocols in VR come with new disadvantages compared to on-screen studies: the eye-movement-to-display-update latency increases slightly, stimuli resolution is lowered, eye tracking accuracy can be reduced. Overcoming these obstacles in order to investigate the roles of central vs. peripheral processing during the search of complex 3D environments was the aim of the present study.

The Present Research

In this study, we measured the effect of transient vision losses on visuo-motor behaviors in VR as a mean of studying the role of central and peripheral vision for visual search in more natural settings. We compared our results to previous on-screen experiments to determine how much of what is learned on-screen can be transferred to natural viewing. For that purpose, we implemented a gaze-contingent protocol to display in real-time central and peripheral masks. One novelty in particular, was the use of a head-mounted display retrofitted with a binocular eye tracker to display transient loss of vision while exciting a large part of the field of view. We analyzed the role of central and peripheral visions in a VR environment where users were able to freely move while searching for objects. In a VR context, we report eye (eye-in-head), head (headset tracking in space) and gaze (eye-in-space) movements. The latter are the final combined gaze movements in the 3D environment. In this paper, we focused mainly on reporting the effects of masking on visuo-motor measures during search in 3D environments.

In a first step, we analyzed the impact of central and peripheral masking on visuo-motor behaviors (hypotheses *Ha*). On-screen literature has extensively demonstrated that saccades target areas where information is best [52]. In effect, average saccade amplitudes increase when central information is modified whereas they decrease when peripheral information is modified [8,20–24,47,53,54]. In a VR experiment, where the field of view is extended and the head is free to rotate, the same effects were reported for eye and gaze movements while free-viewing [47]. Based on these findings, we expected average eye and gaze amplitudes during saccades to display the same effects also during visual search in VR. As reported in free-viewing, head movements decreased with a mask [47], maybe in an effort to control an unusual environment; we expected head movements to be similarly impacted in our analyses. Fixation durations increased significantly when part of the visual field was altered, reflecting cognitive processing difficulties when analyzing modified central information and planning new saccades when peripheral information was modified [20]. In pure free-viewing tasks, fixation durations have been shown to decrease when the central field of view was masked due to an increase in return saccades [24,47]. On the other hand, large peripheral masks reduced fixation durations, whereas smaller masks resulted in an increase. Besides the difference in task, the nature of the masking plays an important role. Cajar et al. [23] implemented low and high-pass filtering masks and argued that when processing becomes too difficult, such that information cannot be efficiently used for gaze control, the duration of fixations does not further increase. Average fixation durations increased when

central information was deemed useful (high spatial frequencies in their case). Since we implemented masks removing all visual information, we expected fixation durations to decrease in the absence of central information to process (central masking) and increase without peripheral stimulation to reflect difficulties in saccade planning. Finally, gaze-contingent masks affect saccade directions as well. Namely, central masking increases return saccade rates while peripheral masking diminishes them [24,47,55,56]. This is observed in free-viewing [24,47] and visual search [56], on-screen [24,56] and in VR [47], therefore it appears to be driven by visual stimulation rather than a top-down effect. Moreover, David et al. [47] showed that this effect is exhibited by eye movements rather than head movements. As such, we expected to observe a similar increase of return saccade rates with central masking and a decrease with peripheral masking through eye and gaze movements, as well as a strong tendency toward forward relative movements by the head.

Free viewing and visual search are often used to study visual attention. Visual search is a highly dynamic task, which can be subdivided into three temporal phases [57]. The initiation phase is the duration of the first fixation at trial onset, and seems to represent the period during which scene's gist is being processed. During the subsequent scanning phase, participants explore the scene in search of a predefined target object. This phase ends with the fixation preceding the first saccade made on the target object. The fixation of the target initiates the verification phase. This last phase accounts for the identification of a target, usually done by foveating it and lasts until the participant presses a button to formally select an object as the expected target. Hypotheses *Hb* encompasses the measurements of average saccade amplitude and fixation duration as a function of these last two phases. With regard to main effect of search phases we hypothesize that the scanning phase is driven by exploration and therefore will be mainly represented by short fixation and long saccades, while the verification phase reflects processing of fine information on the fovea and requires longer fixations and shorter saccades. We expected the interaction of masking and search phase to result in a stronger increase in backward saccades when central vision is missing during verification phases; because in this phase, observers need central vision to make a decision, a central mask will make this harder. The fact that this phase relies heavily on fine perception via central vision should logically result in an increase in return saccade rates as participants attempt to perceives objects foveally in spite of the mask. Considering phases of visual search is important to understand the dynamics of this task. To the best of our knowledge, this is the first study identifying exploration and analysis behaviors in relation to search phases. The opportunity to record head movements lets us measure how the eyes and the head contribute independently to these behaviors.

The study plan and analyses were preregistered on the Open Science Framework website (https://osf.io/g9npw/?view_only=53eca7be33684b0cad4d220f59b90615). The second set of hypotheses related to search phases was not initially intended as per the preregistration document. As such, statistical analyses associated to them should be considered unplanned.

2. Method

2.1. Participants

Thirty-two fluent German speakers were recruited for this experiment (24 women, mean age: 24 years old, minimum: 18, maximum: 43). Recruiting was stopped earlier than planned in the pre-registration document (the aim was to test 45) due to COVID-19. Participants with eyeglasses were not included in this study because frames and lenses reduce the eye tracker's accuracies. We asserted normal visual perception by testing acuity (Landolt "c" chart), color perception (Ishihara color blindness test) and depth perception (stereo fly test). In order to improve comfort and viewing conditions we measured interpupillary distance of participants and set the measurement as the distance between the VR headset's lenses [58]. Finally, we asserted the dominant eye with the cardboard or thumb test. All participants gave their written consent before beginning the experiment and were compensated for their time with university credit or eight Euro/h. The experiment took a maximum

of 60 min, conformed to the Declaration of Helsinki, and was approved by the local ethics committee of the Faculty of Psychology and Sport Sciences at Goethe University Frankfurt.

One participant used a VR headset occasionally, one had used a headset twice before, three had once before, the rest of the participants had never used a VR headset before the experiment. One participant reported motion sickness during the experiment, but chose to finish after resting.

2.2. Apparatus

Virtual scenes were displayed inside an HTC Vive (HTC, Valve corporation) VR head-mounted display worn by participants. The display is refreshed 90 times per second and shows, on average, a field-of-view of 90 by 90 deg. binocularly. The headset was retro-fitted with a Tobii eye tracker (Tobii Technology), tracking gaze binocularly at 120 Hz with a precision below 1.1 deg. within a 20 deg. window centered in the viewports.

The experimental protocol was implemented in the Unity game engine (Unity Technologies) with the Steam VR (Valve corporation) and Tobii eye tracking (Tobii Technology) software libraries. We minimized the time between gaze sampling to updating a mask position in a viewport, by updating the display as late as possible in the rendering process with the latest sample received from the eye tracker. The latency problem is critical in moving-mask conditions, because we need to ensure that central vision is modified even at the start and end of saccades [49,59]. We estimated the maximum ("worst-case scenario") latency to be below 15 ms.

2.3. Stimuli

Sixteen natural everyday rooms were used as virtual scenes (Figure 1). The rooms included three living rooms, bedrooms, bathrooms, and kitchens as well as four offices. Their size was chosen to fit in the physical room where the experiment took place (3.8 by 3.5 m), so that there was no risk for participants of colliding with walls or furniture. The room and object dimensions were chosen to be similar to the real-world. Virtual scenes were populated with 36 unique objects keeping in mind scene grammar rules. Scenes were globally illuminated. As a result, objects did not cast shadows and looked similar from any angle. For a complete description of the scenes please refer to [31]. One of the scenes was set aside for a training phase.

A lexicon of German words was built, referencing all objects present in the scenes. Words from this lexicon were used when displaying a target word relating to the object to look for in the scenes. We chose to use target words rather than pictures of the target in order to reduce search guidance based on low-level, visual characteristics of the actual target, in favor of more high-level guidance by scene grammar. For each scene, six target objects were selected.

Figure 1. Views from three virtual rooms. Examples of target items (from left to right: baby bottles, fruit bowl, drinking bottle) appear within red frames. The black display screen appears in the first two rooms as an illustration of its position only (it is not actually visible in rooms).

2.4. Experimental Design

We implemented a gaze-contingent protocol in order to remove central or peripheral visual information where a participant's gaze was located in real-time (Figure 2). In a shader pass, we masked part of the scene by adding a gray circular mask centered on gaze positions to obtain central masking. The inverse situation (only a central circular area remained unmodified) was implemented to remove

information in the peripheral field of view. Alpha-blending was used at the edge of the circle to create a smooth transition between the mask and the scene [60]. This method was applied to both eyes: the left viewport was updated with left gaze data and the right one with right gaze data.

We chose a circle of six deg. of radius for both central and peripheral masking conditions on the basis of results from a previous experiment [24], and a pre-study dedicated to testing the implementation of the protocol (six subjects, 540 trials). A smaller radius was not advised, as we started noticing latency effects [49,59]. Moreover, a bigger radius may provoke results indistinguishable from a no-mask condition as it approaches the visual span required to complete visual search tasks (eight deg. [54], though this figure was obtained on-screen). We selected only one radius in order to keep sufficient statistical power. The experiment therefore contained two masking conditions (central and peripheral masks) as well as a no-mask condition (control).

Participants were tasked with finding six target objects in 16 scenes, one per trial in one of the three mask conditions (90 trials plus six training trials). We controlled for semantic congruity when selecting target objects. The announced target object was always present in the scenes Figure A1 (Appendix B) shows the polar distribution of target directions relative to the observer's starting position and direction in scenes. We counterbalanced our experimental design so that target objects, scenes and mask conditions across all subjects would be presented the same number of times in a different order.

Figure 2. Masking conditions are presented here in a viewport measuring 90 by 90 deg. of field of view, mask radii are proportionally accurate. From left to right: control no-mask condition, central mask of six deg. of radius, peripheral mask of six deg. of radius.

2.5. Procedure

The experiment started with a calibration procedure (nine points), followed by a validation procedure (nine points) in order to calibrate and test eye tracking precision within 20 deg. from the center of the viewports. We did not test past this window because it has been shown that observers within VR headsets very seldom look further in the periphery of the viewports [47,61,62]. Mean successful validation accuracies across subjects was high (95% Confidence Interval (CI) [1.67, 2.21]). More information about the calibration/validation phase can be found in the Appendix A. Participants got accustomed to the material and the protocol in a training phase where they searched for six objects in one dedicated scene that was not used in the rest of the experiment. They experimented with masked and control conditions (twice each in a random order).

A trial went as follows: starting in an empty room, a black screen on a wall displayed instructions. Participants had to walk onto a blue square on the floor, this was the starting search position from which all target objects could be seen and where a user's body did not intersect with any object (close to the center of the rooms). Then, they had to fixate the screen for one second before the target word for that trial appeared (in German). The word was displayed for a second before it disappeared along with the pre-trial room, instantly replaced by the virtual scene. Observers had 30 s to find a target object. They could end a trial early by pressing the trigger on the VR controller (held in their hand,

but invisible in the virtual world). They were told not to aim at the target object with the controller but to look at it. That allowed us to register from their gaze if they had identified the right object or not, while reducing a bias related to the time needed to target an object in the virtual world. Trials were separated by at least one second (as the program was loading the next scene), a calibration/validation phase was triggered every 10 trials, a resting period was inserted every 30 trials but participants could take a break after any trial. Participants went through all trials in 40 min on average.

2.6. Data Preparation

During trials we retrieved the following data: head rotation (quaternions), left and right gaze directions (3D unit vectors), left and right eye positions (2D position in the left and right viewports), and the object hit by left and right gaze rays (lines in 3D space starting from the left or right eye position, going in the direction of the gaze vector). In order to process eye, head and gaze data we adapted the toolbox [63] developed for the Salient360! benchmark [64,65].

Gaze unit vectors are expressed relative to camera rotations. Therefore, we start by multiplying camera rotations with gaze vectors (quaternion-vector multiplication) to obtain combined eye and head gaze directions in 3D space. This representation and head rotations are projected on a unit sphere centered on the user's head position. Eye movements are 2D positions on the viewports (views inside the headset projected on the displays). We will refer to the addition of head and eye rotations as the **gaze** data (*eye-in-space*), whereas the position of the gaze (dominant eye) in the viewport (*eye-in-head*) will be **eye** data, head rotations are **head** data [66,67].

We identified saccades and fixations on the basis of the combined gaze signal with the help of a velocity-based algorithm [68]: we obtained velocity between gaze samples (deg. per ms.) by calculating the orthodromic distance (great-circle distance) and dividing by the time difference; the resulting signal was smoothed with a Savitzky-Golay filter [69]; filtered samples below 120 deg./ms. were identified as part of a fixation. Visual search phases are identified on the basis of the combined gaze data.

We removed trials for which more than 15% of either left or right eye tracking signals were lost (N = 77). In the context of our protocol, loss of signal results in perceiving part of the scene when that should not be possible. We removed trials that lasted less than 500 ms. or for which only one fixation was found, as they were probably the result of a misfire (N = 2). From this dataset we extracted eye, head and gaze movement-related variables which we describe and analyze in the next section.

3. Analyses

Continuous response variables were analyzed with linear mixed models [70] in order to account for fixed and random experimental effects. The random effects present in our experiment are subjects, scenes and objects (present as random intercepts). We tested *Ha* and *Hb* hypotheses with planned contrasts:

- *Ha*: control condition vs. central mask vs. peripheral mask;
- *Hb*: masking conditions as a function of search phases (scanning, verification).

Hypotheses *Hb* were also tested descriptively by plotting measures over time (Appendix E), and we also displayed measures over a normalized time-course expecting regularities to appear (Appendix D). We believe that a short scanning or verification period does not necessarily compare to a longer one or a failed trial (when a target was not found in the allotted time). As a result, we separated scanning and verification periods into phases lasting less than 2.5 s, less than 7 s, over 7 s, and failed trials (Figure A3, Appendix D).

As part of our analyses we report the following values: *b*-value (estimated difference between means), *SE* (estimated standard error) and *t*-value. An asterisk next to a variable's name in Figure 3 means that distributions were log-transformed to improve the normality of the residuals and the reliability of the models. Analyses were run in *R* (The *R* Project for Statistical Computing [71]) using the *lme4* package [72]. Statistical information about the second set of hypotheses is reported in Table 1.

We removed fixations lasting less than 80 ms. (N = 308) as we consider this the minimum amount of time needed to analyze visual information and prepare the next saccade [42,73,74]). Fixations lasting more than 2000 ms were also removed (N = 20) because they have a high probability of being caused by processes independent from the experimental task [75]. The final dataset is made up of 37,111 fixations and 36,950 saccades.

Table 1. Statistical output of the analysis of the effects of search phases on fixation duration and saccade amplitude (hypotheses *Hb*: scanning vs. verification phase).

Variable	Masking	*b*	*SE*	*t*
Saccade amplitude (Eye)	Control	−0.06	0.01	−4.05
	Central	−0.01	0.01	−1
	Peripheral	0.04	0.01	3.43
Saccade amplitude (Head)	Control	−0.2	0.02	−12.14
	Central	−0.21	0.01	−14.96
	Peripheral	−0.16	0.01	−12.25
Saccade amplitude (Gaze)	Control	−0.35	0.01	−24.04
	Central	−0.3	0.01	−23.69
	Peripheral	−0.22	0.01	−18.53
SRR (Eye)	Control	0.02	0.01	1.61
	Central	0.06	0.01	6.56
	Peripheral	0.02	0.01	2.2
SRR (Head)	Control	−0.02	0.01	−2.99
	Central	−0.02	0.01	−2.85
	Peripheral	0	0.01	0.19
SRR (Gaze)	Control	0.04	0.01	4.22
	Central	0.11	0.01	11.8
	Peripheral	0.03	0.01	3.37
Fixation duration	Control	0.44	0.01	42.51
	Central	0.27	0.01	30.04
	Peripheral	0.35	0.01	42.01
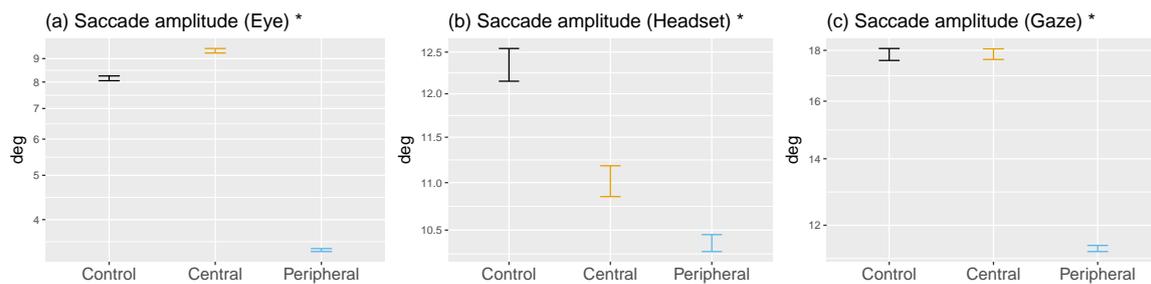
Figure 3. *Cont.*
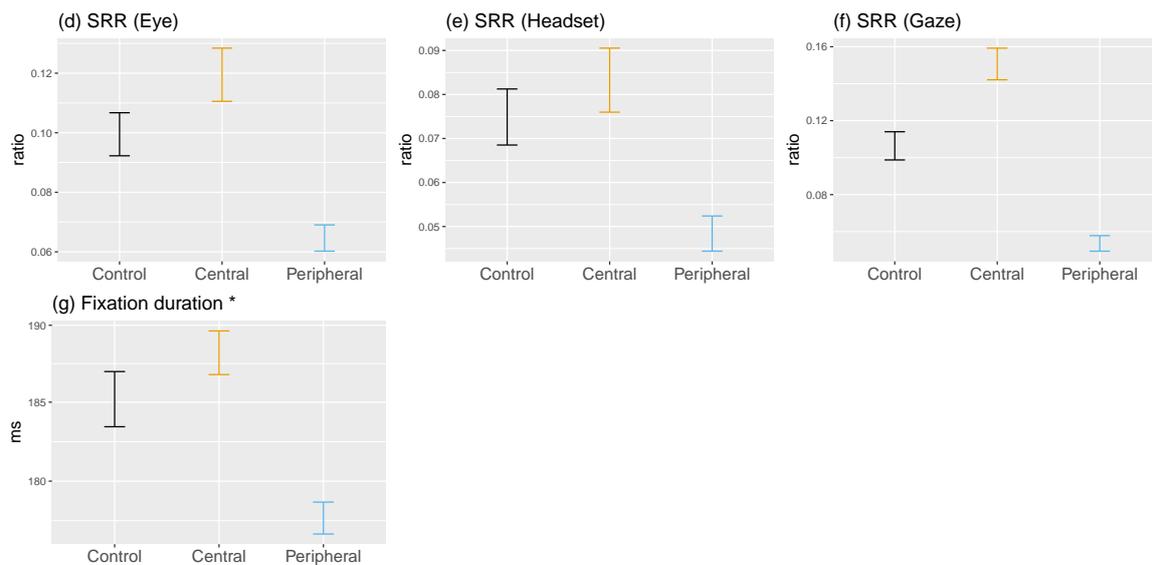
Figure 3. Visuo-motor variables are presented in this figure as a function of mask condition (control in black, central in orange and peripheral in blue) showing means and 95%CI. An asterisk to the right of a variable's name indicates that it was log-transformed in linear mixed models and is presented on a log-scale here. SRR stands for saccadic reversal rate and measures the ratio of backward saccades produced precisely in the 170–190 deg. interval.

3.1. Fixation Durations

Fixation durations are affected by the exploitation of central information, as well as by the exploration of the peripheral field of view needed to prepare the next saccade [17,23,76,77]. Past screen-based studies have shown that degrading or removing information within the field of view resulted in longer fixation times as scene processing is made more difficult and more time is needed to analyze central or peripheral data [17,20,23]. This effect is particularly salient when peripheral vision is modified.

3.1.1. Central vs. Peripheral Masking

When central vision was impaired fixations lasted on average approximately as long as in the no-mask condition (188 ms; $b = 0.01, SE = 0.01, t = 1.48$; Figure 3g). The absence of peripheral stimulation reduced fixation durations to 177 ms ($b = −0.04, SE = 0.01, t = −7.03$). Differences between the scanning and verification phases strongly show an increase in fixation durations reflecting a need for more foveal processing when identifying a target before ending a trial. Figure A6 (Appendix D) shows that over normalized trial time-courses verification trials end with longer fixation durations which are necessary to analyze finely an object of interest and identify it as a target.

3.1.2. Scanning vs. Verification Phase

During scanning, masking had little effect on fixation duration (Figure 4g). However, we observe a decrease in average durations with a mask compared to a no-mask condition during the verification phase. We would have expected the combined effect of masking and verifying to increase fixation durations above the control baseline. In the case of central masking this can be explained by the fact that an increase in return saccades was reported to decrease fixation durations [24,47] because once a region of interest is fixated it is masked and peripheral regions become much more salient in comparison [55], resulting in a fast back-and-forth behavior between a region of interest and the rest of the scene. As for the decrease during peripheral-masking trial, we believe that the mask is large

enough not to hinder the task significantly but still shorten fixation durations due to the absence of peripheral information processing.
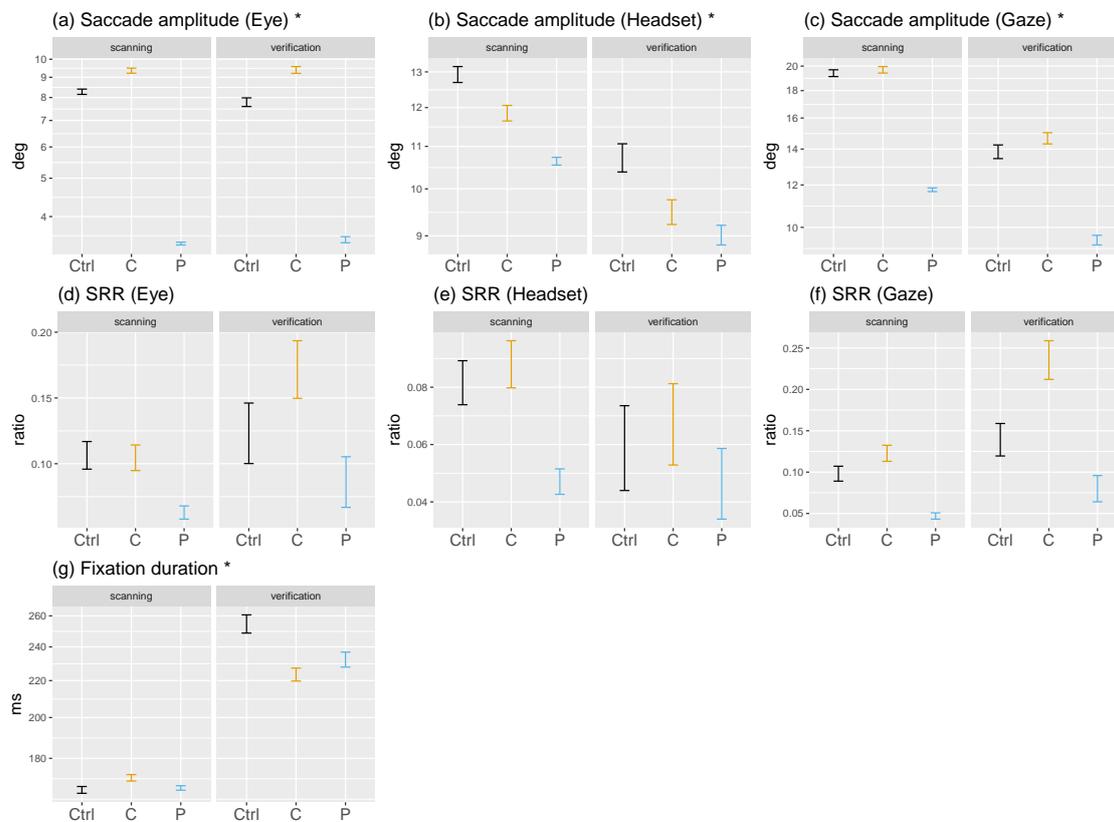
Figure 4. Visuo-motor variables are presented in this figure as a function of search phase (scanning and verification) and mask condition (control in black, central in orange and peripheral in blue) showing means and 95%CI. An asterisk to the right of a variable's name indicates that it was log-transformed in linear mixed models and is presented on a log-scale here. SRR stands for saccadic reversal rate and measures the ratio of backward saccades produced precisely in the 170–190 deg. interval.

As expected, scanning is determined by exploration with small fixation durations, while they increase significantly during verification where potential targets are analyzed. We could have expected to observe longer durations during scanning as well because of difficulties in identifying targets in the periphery with a central mask and the absence of periphery information to direct attention in the second masking condition. But our results indicate that target pre-identification in the visual periphery is truly achieved without central vision. As for peripheral masking, the absence of peripheral regions of interest resulted in a strong saccadic momentum, a mechanism of exploration [78,79] whereby saccades are directed roughly in the same direction, thus away from previous fixations.

3.1.3. Visual Search vs. Free-Viewing and On-Screen vs. Virtual Reality

In free-viewing conditions on-screen a decrease in fixation durations was observed when central vision was removed [24]. This effect increased as mask size grew, probably reflecting the increase in return saccade rate observed jointly: with no central information to analyze and little top-down drive brought by the task, participants quickly produced saccades outside of the masked area. Conversely with peripheral masking, masks left more central information to process as the peripheral masks became smaller, leading to increased fixation durations because saccade planning was made difficult by a lack of peripheral information to target. During a free-viewing task in VR [47] durations decreased with central masks of six and eight deg., whereas the impact of peripheral masks of six deg. was similar to a no-mask condition. Four deg. peripheral masks resulted in an increase of

fixation durations, possibly showing that fixation duration is inversely correlated to peripheral mask size. These results established that for a wide field of view these large masks produce similar effects compared to very small masks on-screen (1.5 and 2.5 deg. of radius), probably because object sizes are smaller in natural scenes presented on-screen rather than in VR. It appears that a peripheral mask of six deg. in our study replicates the effect of large similar masks on-screen (>3.5 deg.). In addition, compared to on-screen gaze-contingent experiments (e.g., [17,20,22]) fixation durations reported in VR previously [47] and in this study seem shorter (approximately [260–300] ms on-screen and [180–200] ms in VR). This may be the result of a more natural environment, but let us remember that fixation durations are impacted by the choice of gaze-parsing algorithm and its parameters [80–82], therefore this result may be an experimental bias and only the within-study differences between conditions should be considered.

3.2. Saccade Amplitudes

Saccade amplitudes are reported to be strongly impacted by the removal of parts of the visual field. As observers plan fixations where information is best at the time of saccade planning [23,24,52,83–85], peripheral masking results in average amplitudes below mask radius and central masking in average amplitude exceeding mask radius. Interestingly, with a peripheral mask in on-screen experiments, observers targeted the edges of the mask and only went significantly beyond in cases where the mask radius was very small (e.g., 1.5 deg. [24]). Therefore, despite building a representation of the scene, participants do not use that knowledge to plan saccades into the masked periphery. Access to eye movement data collected in 3D VR environments now allows asking whether (1) this effect replicates when the head is unrestricted and (2) how eye and head rotations are impacted, if at all? We measured saccade amplitude as the orthodromic distance between two fixation centroids for gaze and head, and as the Euclidean distance on the viewport for eye movements.

3.2.1. Central vs. Peripheral Masking

Contrary to our expectations, removing central vision resulted only in a small increase in the average eye movements during saccades ($b = 0.13, SE = 0.01, t = 16.34$; Figure 3a), while we did not observe an effect in general gaze movements ($b = 0, SE = 0.01, t = 0.36$; Figure 3c). This is the result of a decrease in head movement amplitude ($b = -0.11, SE = 0.01, t = -10.93$; Figure 3b), an effect that was observed in VR previously [47]. This may be explained by the fact that eye and gaze amplitudes during saccades were higher than observed in on-screen experiments ($\approx$4–6 deg.) and in VR ($\approx$6 deg.) to begin with. When peripheral vision was removed, movements were globally reduced: seen via eye ($b = -0.88, SE = 0.01, t = -119.82$), head ($b = -0.18, SE = 0.01, t = -20.2$) and gaze ($b = -0.46, SE = 0.01, t = -59.29$) movements. Once again, gaze-contingent masking lead the eyes to target areas within or at the mask's boundary (Figure 5b). In effect, observers very seldom executed saccades to locations where there was no information, showing a dependency on visual stimulation. In the case of central masking, making longer saccades is unusual, therefore participants produced the shortest saccades they could outside of the mask. However, the no-mask condition showed that they would not have produced particularly smaller saccades by default; as for peripheral masking, targeting the boundary amounts to maximizing information sampling under the constraint of not reaching outside of the mask.
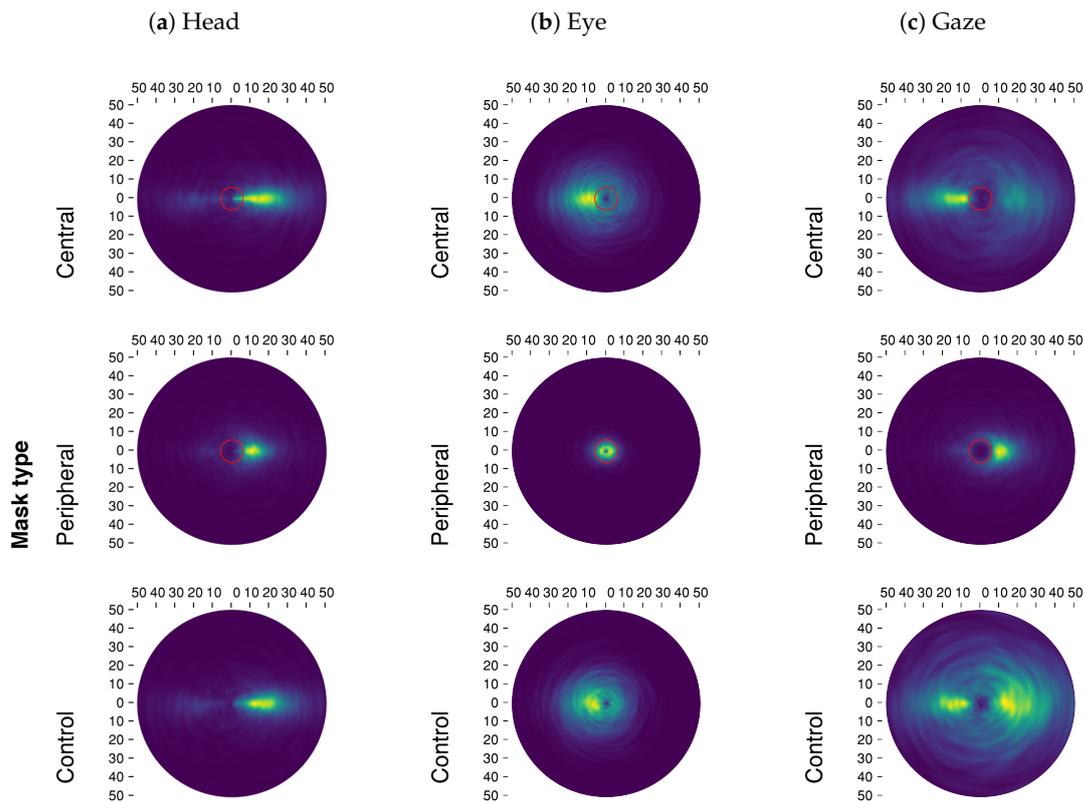
Figure 5. The joint distribution of saccade amplitudes and relative saccade directions are presented in this figure for eye, head and gaze as a function of mask conditions. Red circles at the center of the polar plots denote masks' radius (six deg.). A brighter color represents a higher density of saccades.

3.2.2. Scanning vs. Verification Phase

Across scanning and verification time-courses, we notice that at the beginning of scanning phases head rotations are shorter but increase over the course of the first few seconds (Figures 4b, 6, and A8, Appendix E). Moreover, head movement amplitudes decreased further towards the end of the verification phase as the field of view is stabilized to process a potential target. As for eye rotations, they seem to start scanning phases with higher amplitudes and decrease in the first seconds as the head picks up amplitude (Figure 4a). This dynamic appears to originate from head movements; it seems that participants are in a phase of exploration with long head movements before bringing a target into view. What ensues is a sharp decrease in head movement and increase in fixation durations as participants transition from processing the scene globally to processing locally.

During the scanning phase observers made longer head and gaze movements (Figure 4c) than in the verification phase in order to scan the omnidirectional scene in search of the target. This strong effect does not seem to interact with masking conditions. Interestingly, the average eye movement amplitude did not differ in the two search phases. We found no records of similar phase analyses of eye movements in on-screen studies, therefore we cannot compare our results to past findings. Nevertheless, on-screen we would of course also expect average saccade amplitudes to be longer during scanning compared to verification. In the present experiment large shifts of attention were made more efficiently with the head considering the omnidirectional nature of the stimuli, whereas on-screen what would be a small movement in VR (<15 deg.) could traverse half the scene. In essence, the nature of the VR content may be what elicits exploratory head movements. If we presented non-omnidirectional stimuli in VR with the same relative size as on-screen studies, possibly that eyes movements would show exploratory behaviors and head movements would nearly disappear.
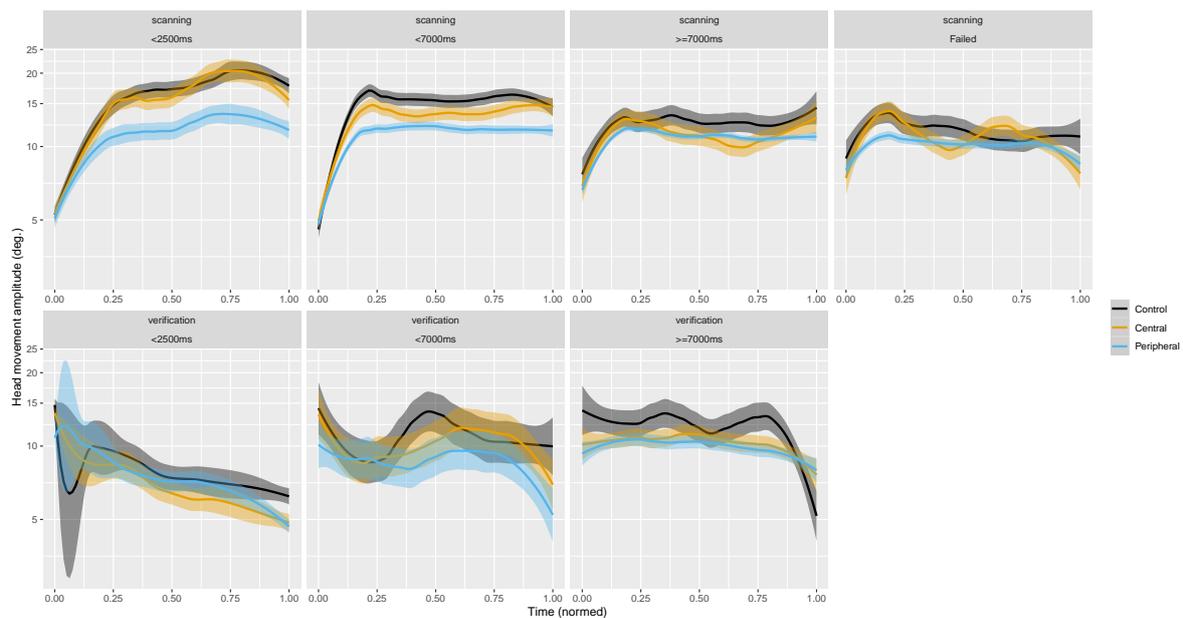
Figure 6. Time-course of average head movements amplitude during saccades as a function of search phase (trial duration subgroups do not overlap, see FIgure A3, Appendix D). Ribbons show 95%CI. Control condition appears in black, central masking in orange and peripheral masking in blue.

3.2.3. Visual Search vs. Free-Viewing

Eye and head movements measured in this experiment were longer than reported in VR for a free-viewing task [47] across all masking conditions (approximately by one deg. for eye amplitude and four deg. for head amplitude). We hypothesize that this reflects a top-down strategy brought out by the search task to quickly find the target by exploring with longer saccades.

3.2.4. On-Screen vs. Virtual Reality

Overall, we replicated the effects of a central versus peripheral gaze-contingent masks on eye movements, however, the ability to use the head in VR to explore resulted in final gaze movements that differed from what we observed on-screen. As can be seen in the saccade direction polar plots (Figure 5), the eyes very rarely escaped the visible areas during peripheral-masking trials, suggesting a strong effect of vision loss on eye movement preparation. Furthermore, it is worth noting that the addition of eye and head movements placed most saccade landing positions just outside of the mask (at the time of saccade planning). If replicated, this would stand in contradiction with theories of trans-saccadic memory requiring an object to target in the periphery [86,87]. On another hand, it seems that deviating from saccadic landing points does not necessarily impact trans-saccadic integration, because attention is not necessarily located at the saccadic landing point [88–90]. This observation raises questions in relation to head and eye saccadic planning and trans-saccadic integration: were participants expecting to saccade within the visible area and felt distraught by mismatched expectations, or were head movements taken into account to purposefully target locations partly void of information at the time of planning? The resulting gaze amplitude distribution (Figure 5c) shows most saccades targeting areas just outside of the mask, which means that after a saccade the new unmasked field of view shares half of its content or more with the preceding one, only displaced. In that regard, there may be enough information shared between fixation locations to permit trans-saccadic integration. This could show an optimal sampling solution to targeting as far as possible into a masked region without incurring negative effects of failed trans-saccadic integration; under the condition that eye movement programming be stimulus-driven but head movement can be modulated by the task (exploration behavior).

3.3. Relative Saccade Directions

"Relative" here describes the angle between two consecutive saccade directions. In contrast to absolute angle informing about the direction of a saccade compared to the horizontal plane (not analyzed but presented in Figure A2, Appendix C), relative angles inform about the dynamics between saccade directions. Relative saccadic angle distributions are presented as polar plots in Figure 5. In this figure the saccadic momentum [78,79] is identified as a mode at 0 deg. of rotation (East; in the same direction as the previous saccade), whereas the facilitation of return [78] appears as a mode positioned at 180 deg. (West; in the inverse direction compared to a previous saccade direction). Past on-screen [24,55] and VR [47] studies have shown that central masking increases return (backward) saccades because observers make back-and-forth motions between regions of interest that they could not properly analyze because of the mask. By contrast, peripheral masking results in fewer backward saccades as the absence of information in the periphery prevents visual attention from lingering on previously fixated regions, decreasing facilitation of return tendencies to the benefit of saccadic momentum.

We obtain angles between saccade direction by projecting gaze position (on a sphere) on to a plane with a Mercator projection. Contrary to an equirectangular projection, the Mercator one is conformal (it preserves angles). This variable is transformed into a ratio in order to study it statistically with linear mixed models. We consider the saccadic reversal rate (SRR [91]): the number of saccades directed within a 170–190 deg. interval divided by the total number of saccades in a given trial. This ratio measures backward (possibly return) saccades, it represents how often a participant looked back precisely in the direction where a fixation was made at $t - 1$.

3.3.1. Central vs. Peripheral Masking

In accordance with past research, we observed an increased number of return saccades when central vision was masked (eye: $b = 0.02, SE = 0, t = 4.05$; gaze: $b = 0.04, SE = 0.01, t = 8.5$; Figures 3d–f). This increase was marginal in the case of head movements ($b = 0.01, SE = 0, t = 2.08$). Participants persevered in trying to analyze foveally a region of interest: a back-and-forth takes place where a fixation is made on a region of interest which is then devoid of information, making peripheral portions of the field of view much more salient, thus a new fixation is made outside unmasking the previous regions. In contrast, when no peripheral information is available, gaze appeared to be directed predominantly by saccadic momentum: the SRR strongly decreased globally in eye ($b = -0.03, SE = 0, t = -7.12$), head ($b = -0.03, SE = 0, t = -6.54$) and gaze ($b = -0.05, SE = 0.01, t = -10.38$) movement directions. However, we observed that eye relative directions with a peripheral mask were almost uniform radially (Figure 5b). The strong bias in forward saccade rates that can be seen in this condition was driven by the head (Figure 5a).

3.3.2. Scanning vs. Verification Phase

Our analyses of search phases (Figures 4d–f) shows that return saccade rates increased particularly during verification phases when central vision was masked. Return saccade rates were at no-mask level during scanning, showing exploring on par between central-masking and control trials, but shot up in the final phase that required a fine perception of objects for analysis. This increase was apparent in eye rather than head movements. Head movements show a general decrease in the second phase reflecting the stability necessary to finely analyze with the eyes. The absence of peripheral stimulation resulted in fewer return saccades, approximately of the same order in both search phases.

3.3.3. Visual Search vs. Free-Viewing

Compared to a previous study implementing free-viewing in VR [47] our general results (aggregated over search phases) overall do not differ. Scanning and verification represented as periods of exploration and analysis, also manifest themselves in free-viewing [14,92]. While it seems

that the general behavior is similar between tasks, more analyses are needed to study global and local scene processing in free-viewing where there is possibly less pressure to explore and/or analyze.

3.3.4. On-Screen vs. Virtual Reality

To our knowledge only one on-screen study, which implemented masking during a free-viewing task, reported a measure of return saccade rates [24]. Therefore, data for a comparison is limited. We can say that the average difference between control and central-masking trials is comparable to on-screen masks of radius 2.5 to 3.5 deg. Whereas the impact of peripheral-masking seems to correspond to a larger mask radius of 4.5 deg.

4. Discussion

With a recent push to study vision under increasingly realistic conditions, it has become crucial to take a closer look at eye, head, and gaze movement characteristics during a real-world task in 3D space. Moreover, there is an increasing need to test which findings taken from on-screen studies actually hold truth when moving into 3D and vice versa. In the study presented here, we therefore implemented a gaze-contingent protocol to remove either central or peripheral vision while observers accomplished a real-world search task immersed in virtual every-day scenes. A particularly original aspect of this study is the measurement and comparison of visuo-motor behaviors across visual search phases (scanning and verification). Our results have important implications regarding the reproducibility of knowledge gathered in on-screen experiments for more natural environments. Additionally, to our knowledge this is the first study reporting on visual search with gaze-contingent masks in 3D, omnidirectional virtual every-day scenes. As such, our analyses of visuo-motor measures inform about search behaviors in more natural settings, in particular about the role of eye and head movements.

Overall, we have found that the impact of gaze-contingent windows and masks on visual search in VR were quite similar to those reported previously in the literature using on-screen experiments. Eye rotation movements increased in amplitude when we removed central vision and decreased without peripheral information, as was expected from previous findings [8,20–24,47,53,54]. This effect translated to gaze movements as well, though the increase in eye rotation amplitude was mediated by the baseline increase observed without masking. During no-mask trials observers made longer saccades than observed in free-viewing in VR [47], the impetus to accomplish the task pushed them to explore their omnidirectional environment with more distance between fixations. The absence of central information increased backward saccade rates as observers made back-and-forth movements between regions of interest [21,24,47,55]. Either in an effort to verify a target exra-foveally or because the lack of central information leads visual attention to be grasped in a bottom-up fashion by peripheral stimulations. On the contrary, with no peripheral stimulation gaze was guided predominantly by scene exploration, showing a strong saccadic momentum [78,79]. We observed that fixation durations were only slightly affected by masking, significantly so when peripheral vision was missing. Contrary to gaze-contingent experiments manipulating spatial frequencies (e.g., [8,20–23,52–54]) or color [17], our protocol removed all stimulation within masked areas. As a result, there was no difficulty related to processing modified information since it was entirely removed. There may exist difficulties in planning new saccades in the peripheral masking condition but this did not transpire in this study. As was observed previously [24,47], participants made shorter fixations of longer amplitudes generally directed in a common direction, characteristic of a global scanning phase [14,15].

In addition to merely analyzing eye rotations in 3D, VR gave us the opportunity to study how the head contributes to visual search in 3D. Head movements serve exploration by extending saccade amplitude's length, being fairly simple and moreover showing a very strong forward momentum bias. Note that this bias also appeared because we defined saccades on the basis of gaze movements, in a case where the head rotates to span a big part of one's surroundings there may be several saccades happening. Because we segmented all signals into fixations on the basis of the overall gaze movement data, a large head movement would be segmented into several ones corresponding to the motion

of the head in-between two fixations. Therefore, one head movement made to explore a big part of the scene will be recorded as several head movements in the same direction (forward relative directions). Additionally, head movement is measured by the VR headset's sensors not differentiating rotations stemming from feet, torso or neck movements. The effect of transient vision loss on head movements is new and shown as a global decrease in average amplitude, we assume that this effect reflects the abnormality of the situation. If we consider that effects of vision loss on eye movements is stimulus-driven and almost uncontrollable, reducing head movements may reflect difficulty through uncertainty and lack of proper control of the situation.

We have reported differences between visual search and free-viewing tasks by comparing our results to previous research on-screen and in VR. In particular, we show longer saccade amplitudes in this study compared to free-viewing in VR; possibly reflecting a bias toward exploration, which is needed to scan a large environment for a target with a time-constraint. This partly replicates the findings of Mills et al. [93] comparing four different visual tasks. The authors measured shorter fixations and longer saccades in the visual search task compared to free-viewing. Smith and Mital [94] reported the same effects comparing free-viewing to a "spot-the-view" task; free-viewing resulted in longer fixations and shorter saccades. Considering the literature, it makes sense to discern free-viewing from goal-oriented tasks. Importantly, our results do not show shorter fixation durations compared to a previous VR free-viewing experiment [47], instead we possibly showed a tendency of immersive contents to decrease average fixation durations. A dedicated study comparing visual search and free-viewing is needed to make clear interpretations with regard to task-related differences of eye movements in VR.

Studying the time-course of fixation duration and movement amplitudes during saccades we showed that participants started trials with the head hardly moving. We believe this mirrors the initiation phase of visual search: the time needed to start significantly shifting the field of view to explore may be positively correlated with the difficulty to control the situation and build a suitable representation of the scene. Right after the first fixation on a target (start of the verification phase), observers made shorter head movements because there was no need to make big visual field shifts once a potential target was in view. Our analyses confirm our hypothesis by showing that scanning phases are exploratory with long saccades and short fixation durations, while verification phases are characterized by short saccades and long fixation durations. Fixation durations particularly increased at the end of the second phase as participants made their final decision before pressing the controller's button to end the trial. Our data did not show a clear interaction effect between search phases and masks. Interestingly, eye rotation amplitudes were not significantly impacted by search phases. We noticed that scanning phases began with eye movement amplitudes larger than observed during the rest of the phase, which had been observed on-screen previously [95,96] and is thought to reflect a strategy to extract the most informative part of the scene at the start of a trial. This strategy turns into exploration of the omnidirectional scene when head movements increase in amplitude as eye movements become smaller. Considering return saccade rates (SRR), we observed that central masking only affected this measure during verification phases. This indicates that scene exploration is not hindered by a large central mask.

The strong effect of transient peripheral vision loss on almost all measures studied reiterates the importance of peripheral vision in visual search—particularly in real-world scenes—to build a representation of the scene and to detect potential targets. Masking central vision did not have as strong an impact on visuo-motor behaviors as we would expect from past literature considering the large mask size. This implies that peripheral vision (6 deg. and up) is in some measure sufficient to explore a complex environment.

The differential effects of vision loss on eye and head movements potentially indicate a difference in control and planning during fixations and saccades. Eye movements appear to be strongly impacted by available visual information, while head rotations are dedicated to exploring the environment. In peripheral-masking trials the head even allows planning saccades into the masked area. This shows

that there can be a strong effect of vision loss and at the same time a pursuit of a task goal (exploration), both dependent on different elements of the visuo-motor system. Further work on the topic of visual search in virtual reality could be beneficial to the quality of life of patients suffering from visual field deficits [97–99] for example, by recommending efficient search strategies based on the remaining useful field of view.

5. Conclusions

In this study, thanks to a fully integrated VR eye tracking apparatus, we were able to gain new insights regarding gaze behaviors during an object search task in quasi-natural settings. We replicated on-screen gaze-contingent experiments and made clear that the impact of gaze-contingent masking is strongly reflected on eye movements, less so on head rotations. We have shown that the head, not the eyes, drives exploration and analyses during the scanning and verification phases, respectively. We hope this study will motivate the community to look into VR to simulate more natural conditions and implement new experimental designs.

## Appendix A. Calibration and Validation Phase

The calibration of the embedded Tobii eye tracker was done using the vendor's proprietary procedure, we replaced the default five points procedure with a nine points one to increase accuracy. We implemented a validation protocol displaying concentric rings on a grid centered in the viewports. The grid was displayed with a shader program on a quad object. This quad was configured so as to always fill exactly the camera's field of view. Subsequently, during a validation phase we projected left and right gazes onto the validation grid quad (2D plane) and measured the Euclidean distance (in deg.) between the projected gaze points and the current validation target position. We acquired binocular gaze samples for a maximum of 3 s, for each eye the best accuracy measured within a 200 ms window was returned by the subroutine. This procedure was repeated for all nine validation target positions. We tolerated one value on the upper part of the viewport to be less accurate, because the Tobii eye tracker was demonstrably less reliable in this area.

Appendix B. Target Objects Placement
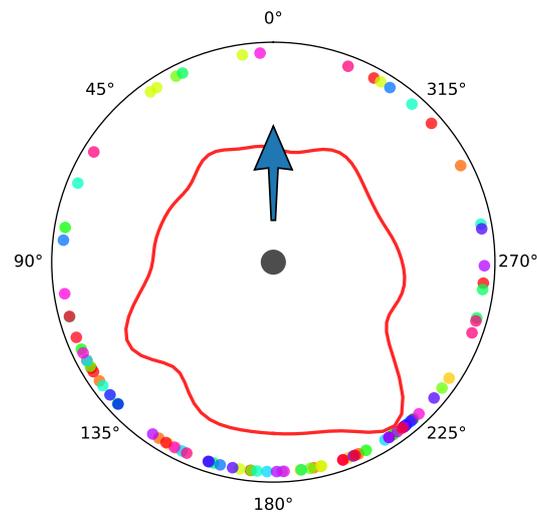
Figure A1. Directions of target objects in scenes relative to the observer's starting directions (blue arrow pointing north) after collapsing the up axis (view from above). The dots represent individual target objects, their color refers to the scene they belong to. The polar density distribution of directions is shown as a red curve (estimated with a von Mises kernel). Color-to-scene legend: ● B01, ● B02, ● B03, ● K01, ● K02, ● K03, ● L01, ● L02, ● L03, ● O01, ● O02, ● O03, ● O04, ● S01, ● S02, ● S03 (where *B* means *bathroom*, *K* means *kitchen*, *L* means *living-room*, *O* means *office* and *S* means *bedroom*).

Appendix C. Saccade Absolute Directions
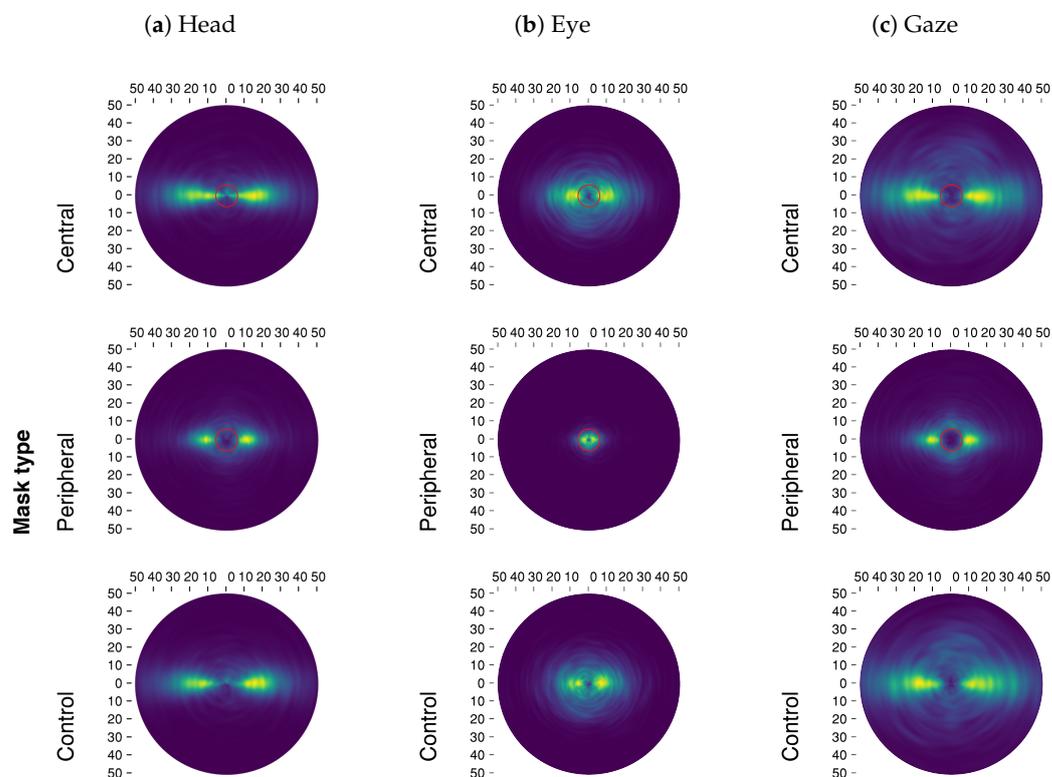
Figure A2. The joint distribution of saccade amplitudes and saccade absolute directions are presented in this figure for eye, head and gaze movements as a function of mask conditions. Red circles at the

center of the polar plots show masks' radius (six deg.). A brighter color represents a higher density of saccades.

## Appendix D. Time-Course of Visuo-Motor Measures over Search Phases, Normalized
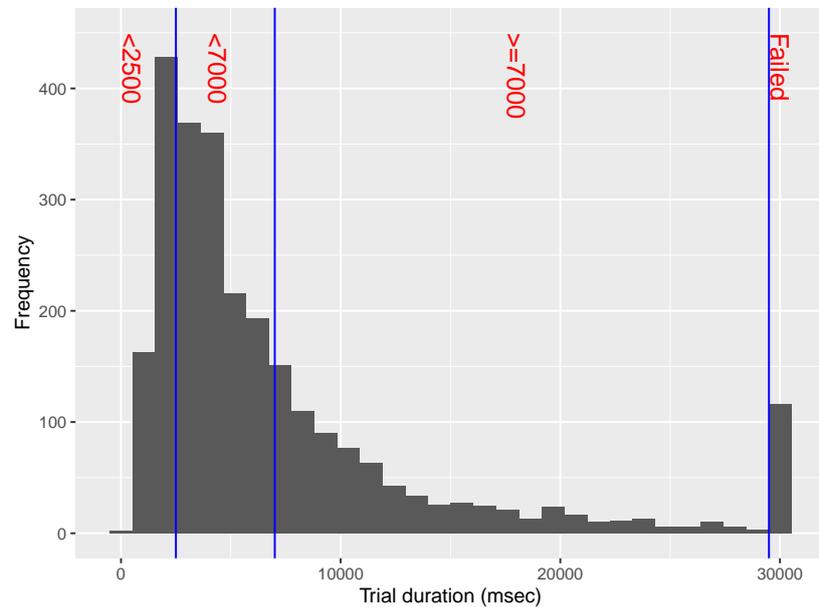
Figure A3. Frequency of trial durations. The blue vertical lines delimit the non-overlapping subgroups reported in Figure 6, and Appendices D and E. The first two subgroups divide the distribution at its mode, the third subgroup aggregates samples located at the tail of the distribution.
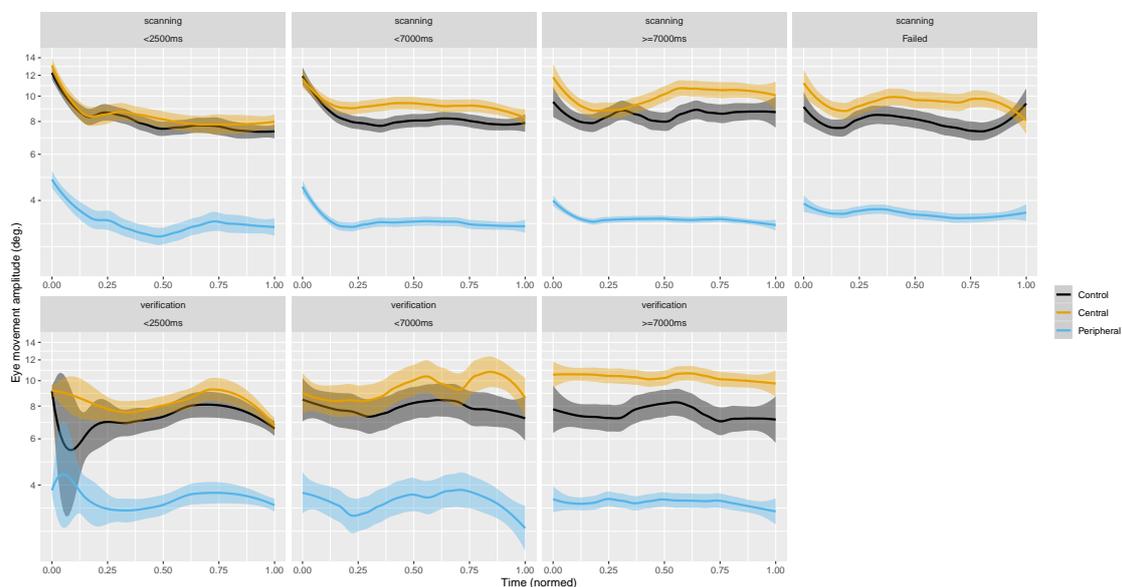
Figure A4. Time-course of average eye movements amplitude during saccades as a function of search phase (trial duration subgroups do not overlap). Ribbons show 95%CI. Control condition appears in black, central in orange and peripheral in blue.
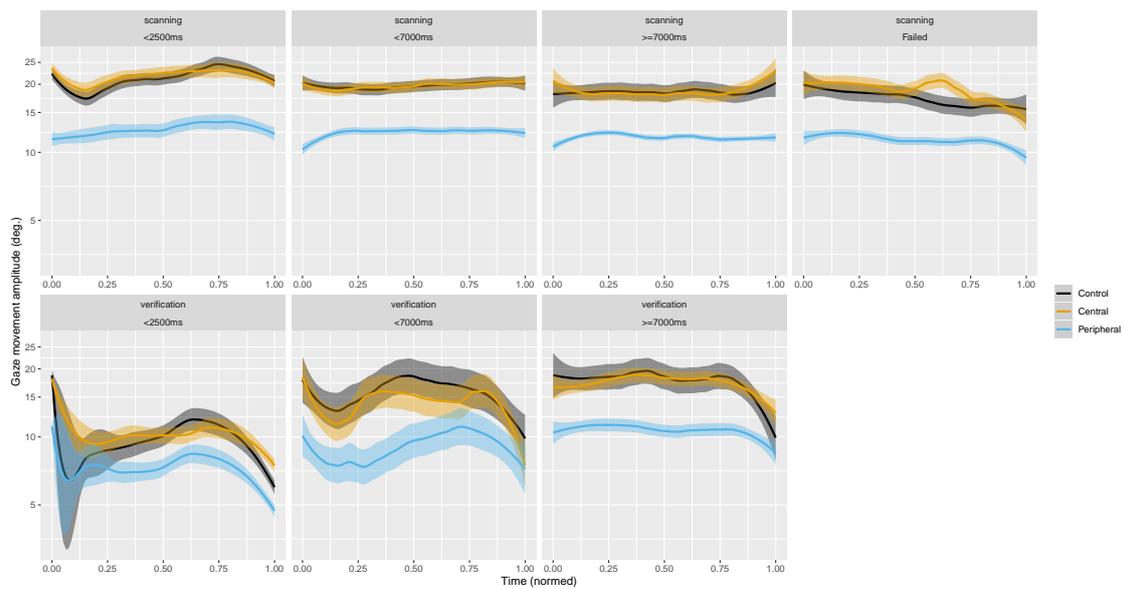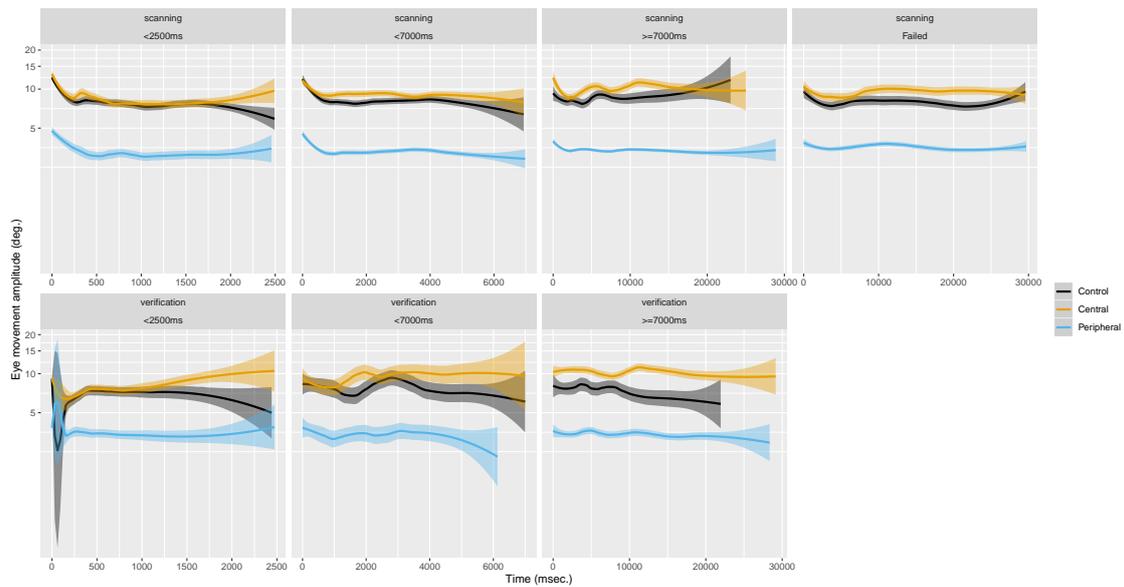
Figure A5. Time-course of average gaze movements amplitude during saccades as a function of search phase (trial duration subgroups do not overlap). Ribbons show 95%CI. Control condition appears in black, central in orange and peripheral in blue.
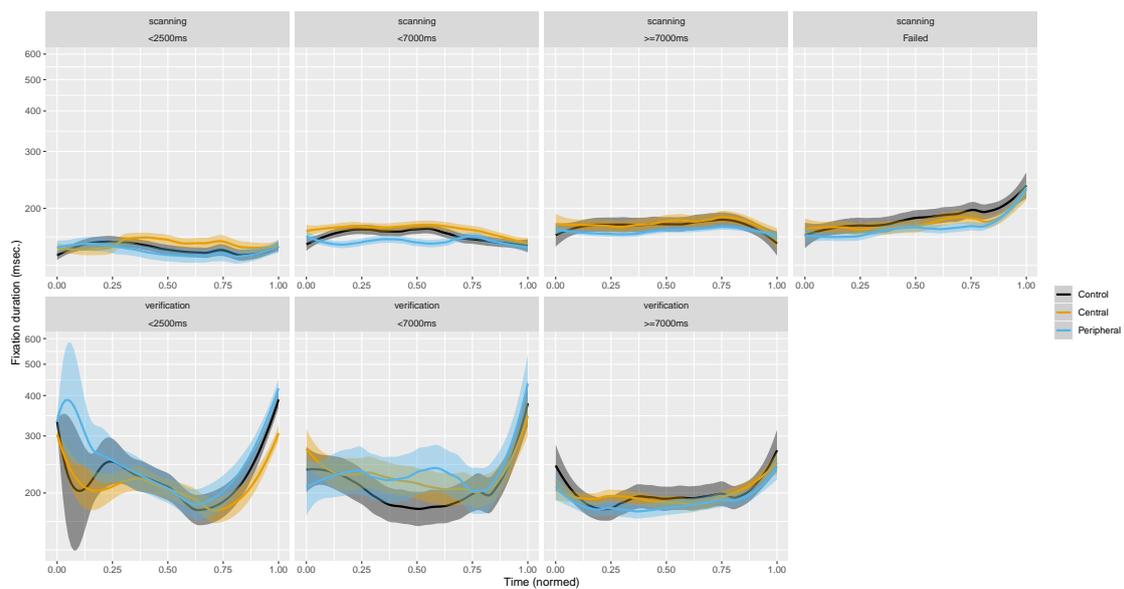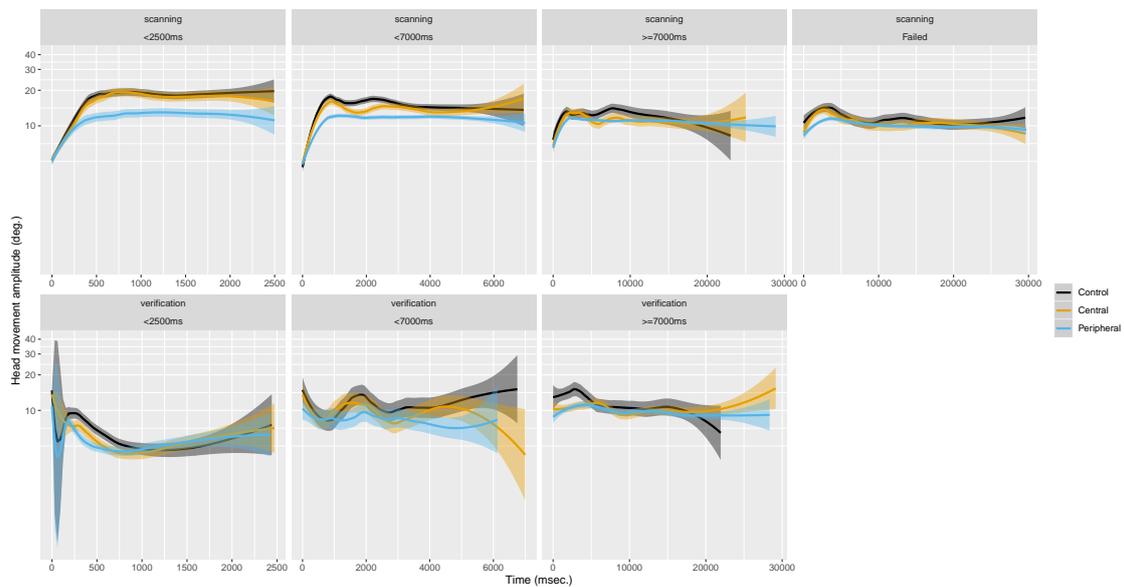

Figure A6. Time-course of average fixation duration during saccades as a function of phase search (trial duration subgroups do not overlap). Ribbons show 95%CI. Control condition appears in black, central in orange and peripheral in blue.
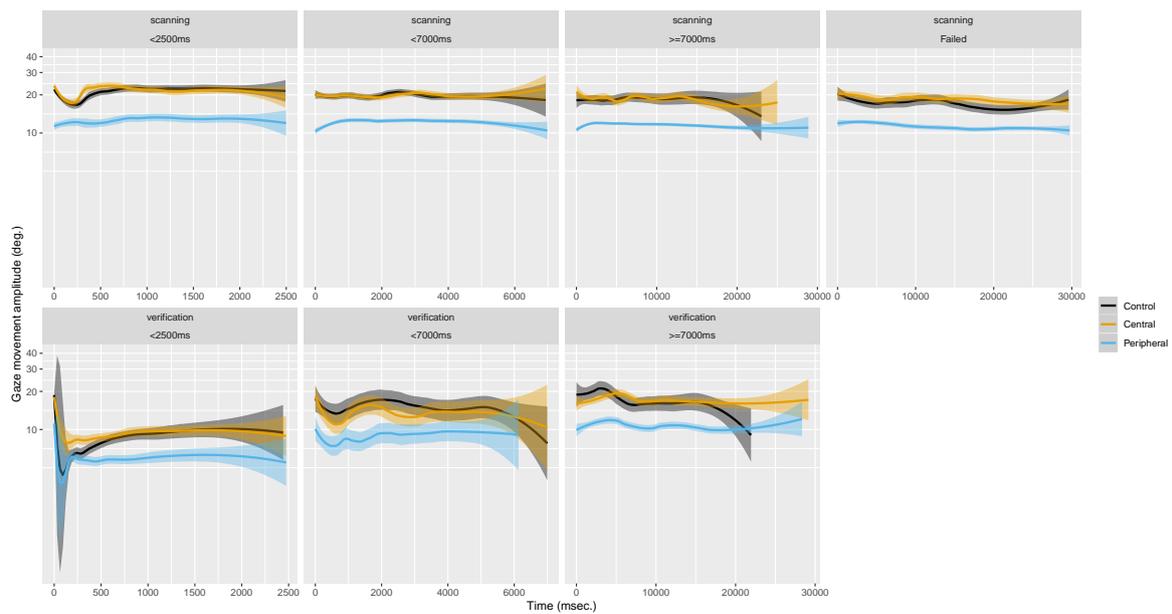
## Appendix E. Time-Course of Visuo-Motor Measures over Search Phases

Figure A7. Time-course of average eye movements amplitude during saccades as a function of search phase (trial duration subgroups do not overlap). Ribbons show 95%CI. Control condition appears in black, central in orange and peripheral in blue.

Figure A8. Time-course of average head movements amplitude during saccades as a function of search phase (trial duration subgroups do not overlap). Ribbons show 95%CI. Control condition appears in black, central in orange and peripheral in blue.

Figure A9. Time-course of average gaze movements amplitude during saccades as a function of search phase (trial duration subgroups do not overlap). Ribbons show 95%CI. Control condition appears in black, central in orange and peripheral in blue.
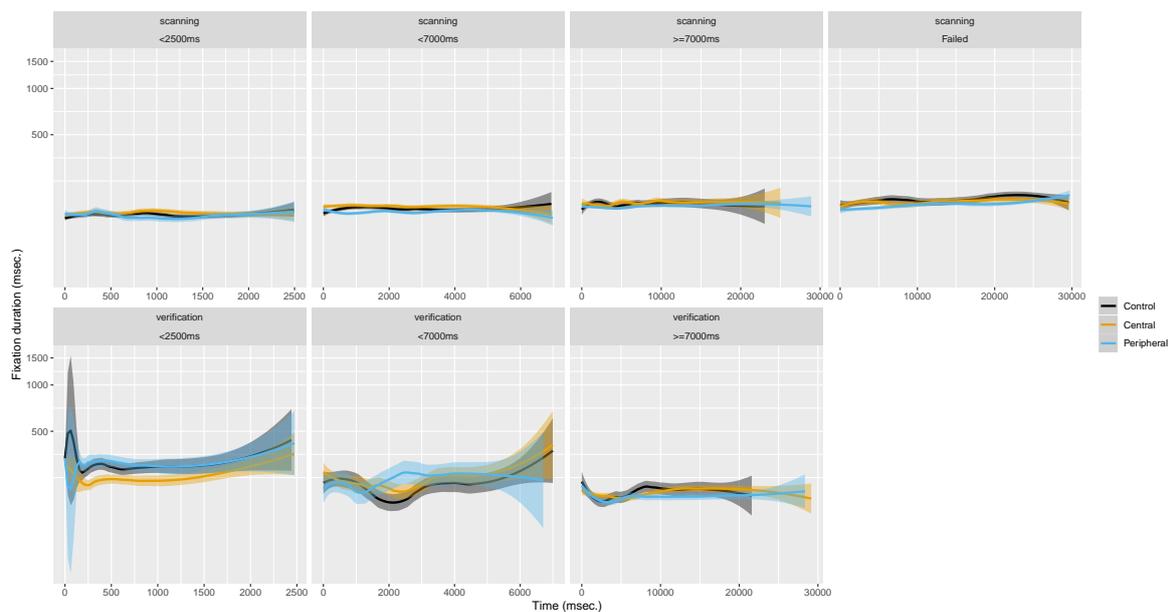
Figure A10. Time-course of average fixation duration during saccades as a function of phase search (trial duration subgroups do not overlap). Ribbons show 95%CI. Control condition appears in black, central in orange and peripheral in blue.

References

1. Curcio, C.A.; Sloan, K.R.; Kalina, R.E.; Hendrickson, A.E. Human photoreceptor topography. *J. Comp. Neurol.* **1990**, *292*, 497–523. [CrossRef] [PubMed]
2. Oyster, C.W. *The Human Eye: Structure and Function*; Sinauer Associates: Sunderland, MA, USA, 1999.
3. Dacey, D.M. The mosaic of midget ganglion cells in the human retina. *J. Neurosci.* **1993**, *13*, 5334–5355. [CrossRef] [PubMed]
4. Croner, L.J.; Kaplan, E. Receptive fields of P and M ganglion cells across the primate retina. *Vis. Res.* **1995**, *35*, 7–24. [CrossRef]

5. Nassi, J.J.; Callaway, E.M. Parallel processing strategies of the primate visual system. *Nat. Rev. Neurosci.* **2009**, *10*, 360–372. [CrossRef]

6. Schira, M.M.; Wade, A.R.; Tyler, C.W. Two-dimensional mapping of the central and parafoveal visual field to human visual cortex. *J. Neurophysiol.* **2007**, *97*, 4284–4295. [CrossRef]

7. Schira, M.M.; Tyler, C.W.; Spehar, B.; Breakspear, M. Modeling magnification and anisotropy in the primate foveal confluence. *PLoS Comput. Biol.* **2010**, *6*, e1000651. [CrossRef]

8. Loschky, L.; McConkie, G.; Yang, J.; Miller, M. The limits of visual resolution in natural scene viewing. *Vis. Cogn.* **2005**, *12*, 1057–1092. [CrossRef]

9. Rosenholtz, R. Capabilities and limitations of peripheral vision. *Annu. Rev. Vis. Sci.* **2016**, *2*, 437–457. [CrossRef]

10. Larson, A.M.; Loschky, L. The contributions of central versus peripheral vision to scene gist recognition. *J. Vis.* **2009**, *9*, 6. [CrossRef]

11. Loschky, L.; Szaffarczyk, S.; Beugnet, C.; Young, M.E.; Boucart, M. The contributions of central and peripheral vision to scene-gist recognition with a 180 visual field. *J. Vis.* **2019**, *19*, 15. [CrossRef]

12. Boucart, M.; Moroni, C.; Thibaut, M.; Szaffarczyk, S.; Greene, M. Scene categorization at large visual eccentricities. *Vis. Res.* **2013**, *86*, 35–42. [CrossRef] [PubMed]

13. Boucart, M.; Lenoble, Q.; Quettelart, J.; Szaffarczyk, S.; Despretz, P.; Thorpe, S.J. Finding faces, animals, and vehicles in far peripheral vision. *J. Vis.* **2016**, *16*, 10. [CrossRef] [PubMed]

14. Tatler, B.W.; Vincent, B.T. Systematic tendencies in scene viewing. *J. Eye Mov. Res.* **2008**, *2*. [CrossRef]

15. Godwin, H.J.; Reichle, E.D.; Menneer, T. Coarse-to-fine eye movement behavior during visual search. *Psychon. Bull. Rev.* **2014**, *21*, 1244–1249. [CrossRef] [PubMed]

16. Velichkovsky, B.M.; Korosteleva, A.N.; Pannasch, S.; Helmert, J.R.; Orlov, V.A.; Sharaev, M.G.; Velichkovsky, B.B.; Ushakov, V.L. Two Visual Systems and Their Eye Movements: A Fixation-Based Event-Related Experiment with Ultrafast fMRI Reconciles Competing Views. *Med Technol. Med./Sovrem. Tehnol. V Med.* **2019**, *11*, 7–18. [CrossRef]

17. Nuthmann, A.; Malcolm, G.L. Eye guidance during real-world scene search: The role color plays in central and peripheral vision. *J. Vis.* **2016**, *16*, 3. [CrossRef]

18. Naïli, F.; Despretz, P.; Boucart, M. Colour recognition at large visual eccentricities in normal observers and patients with low vision. *Neuroreport* **2006**, *17*, 1571–1574. [CrossRef]

19. Hansen, T.; Pracejus, L.; Gegenfurtner, K.R. Color perception in the intermediate periphery of the visual field. *J. Vis.* **2009**, *9*, 26. [CrossRef]

20. Laubrock, J.; Cajar, A.; Engbert, R. Control of fixation duration during scene viewing by interaction of foveal and peripheral processing. *J. Vis.* **2013**, *13*, 11. [CrossRef]

21. Nuthmann, A. How do the regions of the visual field contribute to object search in real-world scenes? Evidence from eye movements. *J. Exp. Psychol. Hum. Percept. Perform.* **2014**, *40*, 342. [CrossRef]

22. Cajar, A.; Schneeweiß, P.; Engbert, R.; Laubrock, J. Coupling of attention and saccades when viewing scenes with central and peripheral degradation. *J. Vis.* **2016**, *16*, 8. [CrossRef] [PubMed]

23. Cajar, A.; Engbert, R.; Laubrock, J. Spatial frequency processing in the central and peripheral visual field during scene viewing. *Vis. Res.* **2016**, *127*, 186–197. [CrossRef] [PubMed]

24. David, E.; Lebranchu, P.; Da Silva, M.P.; Le Callet, P. Predicting artificial visual field losses: A gaze-based inference study. *J. Vis.* **2019**, *19*, 22. [CrossRef]

25. Thibaut, M.; Tran, T.H.C.; Szaffarczyk, S.; Boucart, M. The contribution of central and peripheral vision in scene categorization: A study on people with central vision loss. *Vis. Res.* **2014**, *98*, 46–53. [CrossRef] [PubMed]

26. Thibaut, M.; Delerue, C.; Boucart, M.; Tran, T. Visual exploration of objects and scenes in patients with age-related macular degeneration. *J. Fr. D'ophtalmologie* **2016**, *39*, 82–89. [CrossRef]

27. Malcolm, G.L.; Groen, I.I.; Baker, C.I. Making sense of real-world scenes. *Trends Cogn. Sci.* **2016**, *20*, 843–856. [CrossRef]

28. Howard, C.J.; Pharaon, R.G.; Körner, C.; Smith, A.D.; Gilchrist, I.D. Visual search in the real world: Evidence for the formation of distractor representations. *Perception* **2011**, *40*, 1143–1153. [CrossRef]

29. Mack, S.C.; Eckstein, M.P. Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. *J. Vis.* **2011**, *11*, 9. [CrossRef]

30. Foulsham, T.; Chapman, C.; Nasiopoulos, E.; Kingstone, A. Top-down and bottom-up aspects of active search in a real-world environment. *Can. J. Exp. Psychol. Can. de Psychol. Expérimentale* **2014**, *68*, 8. [CrossRef]

31. Helbing, J.; Draschkow, D.; Võ, M. Anchors aweigh! Semantic and syntactic anchor object information interact to make visual search in immersive scenes efficient. *J. Vis.* **2020**, *20*, 573. [CrossRef]

32. Asher, W. Monoculares und binoculares Blickfeld eines Myopischen. *Graefe's Arch. Clin. Exp. Ophthalmol.* **1898**, *47*, 318–338. [CrossRef]

33. Hofmann, F. *Die Lehre vom Raumsinn des Auges*; Springer: Berlin/Heidelberg, Germany, 1925.

34. Von Noorden, G.K.; Campos, E.C. *Binocular Vision and Ocular Motility: Theory and Management of Strabismus*, 6th ed.; Mosby: St. Louis, MO, USA, 2002.

35. Rötth, A. Über das praktische Blickfeld. *Graefe's Arch. Clin. Exp. Ophthalmol.* **1925**, *115*, 314–321. [CrossRef]

36. Stahl, J.S. Amplitude of human head movements associated with horizontal saccades. *Exp. Brain Res.* **1999**, *126*, 41–54. [CrossRef] [PubMed]

37. Freedman, E.G.; Sparks, D.L. Coordination of the eyes and head: Movement kinematics. *Exp. Brain Res.* **2000**, *131*, 22–32. [CrossRef]

38. Malinov, I.V.; Epelboim, J.; Herst, A.N.; Steinman, R.M. Characteristics of saccades and vergence in two kinds of sequential looking tasks. *Vis. Res.* **2000**, *40*, 2083–2090. [CrossRef]

39. Einhäuser, W.; Schumann, F.; Bardins, S.; Bartl, K.; Böning, G.; Schneider, E.; König, P. Human eye-head co-ordination in natural exploration. *Network Comput. Neural Syst.* **2007**, *18*, 267–297. [CrossRef] [PubMed]

40. Freedman, E.G. Coordination of the eyes and head during visual orienting. *Exp. Brain Res.* **2008**, *190*, 369. [CrossRef] [PubMed]

41. Collewijn, H.; Smeets, J.B. Early components of the human vestibulo-ocular response to head rotation: Latency and gain. *J. Neurophysiol.* **2000**, *84*, 376–389. [CrossRef]

42. Leigh, R.J.; Zee, D.S. *The Neurology of Eye Movements*; Oxford University Press: Oxford, UK, 2015.

43. Robinson, D. Control of eye movements. In *Handbook of Physiology, Vol. III*; Brooks, V., Ed.; American Physiological Society: Washington, DC, USA, 1981; Chapter 28, pp. 1275–1320.

44. Ritzmann, E. Ueber die Verwendung von Kopfbewegungen bei den gewöhnlichen Blickbewegungen. *Graefe's Arch. Clin. Exp. Ophthalmol.* **1875**, *21*, 131–149. [CrossRef]

45. Bizzi, E.; Kalil, R.E.; Morasso, P. Two modes of active eye-head coordination in monkeys. *Brain Res.* **1972**, *40*, 45–48. [CrossRef]

46. Hu, B.; Johnson-Bey, I.; Sharma, M.; Niebur, E. Head movements during visual exploration of natural images in virtual reality. In Proceedings of the 2017 51st Annual Conference on Information Sciences and Systems (CISS), Baltimore, MD, USA, 22–24 March 2017; pp. 1–6.

47. David, E.; Lebranchu, P.; Da Silva, M.P.; Le Callet, P. What visuo-motor tendencies for the free-viewing of omnidirectional scenes in virtual reality? 2020, in preparation.

48. Duchowski, A.T.; Cournia, N.; Murphy, H. Gaze-contingent displays: A review. *CyberPsychol. Behav.* **2004**, *7*, 621–634. [CrossRef] [PubMed]

49. Aguilar, C.; Castet, E. Gaze-contingent simulation of retinopathy: Some potential pitfalls and remedies. *Vis. Res.* **2011**, *51*, 997–1012. [CrossRef] [PubMed]

50. McConkie, G.W.; Rayner, K. The span of the effective stimulus during a fixation in reading. *Percept. Psychophys.* **1975**, *17*, 578–586. [CrossRef]

51. Rayner, K.; Bertera, J.H. Reading without a fovea. *Science* **1979**, *206*, 468–469. [CrossRef] [PubMed]

52. Foulsham, T.; Teszka, R.; Kingstone, A. Saccade control in natural images is shaped by the information visible at fixation: Evidence from asymmetric gaze-contingent windows. *Atten. Percept. Psychophys.* **2011**, *73*, 266–283. [CrossRef]

53. Loschky, L.; McConkie, G.W. Investigating spatial vision and dynamic attentional selection using a gaze-contingent multiresolutional display. *J. Exp. Psychol. Appl.* **2002**, *8*, 99. [CrossRef]

54. Nuthmann, A. On the visual span during object search in real-world scenes. *Vis. Cogn.* **2013**, *21*, 803–837. [CrossRef]

55. Henderson, J.M.; McClure, K.K.; Pierce, S.; Schrock, G. Object identification without foveal vision: Evidence from an artificial scotoma paradigm. *Percept. Psychophys.* **1997**, *59*, 323–346. [CrossRef]

56. Cornelissen, F.W.; Bruin, K.J.; Kooijman, A.C. The influence of artificial scotomas on eye movements during visual search. *Optom. Vis. Sci.* **2005**, *82*, 27–35.

57. Malcolm, G.L.; Henderson, J.M. The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *J. Vis.* **2009**, *9*, 8. [CrossRef]

58. Best, S. Perceptual and oculomotor implications of interpupillary distance settings on a head-mounted virtual display. In Proceedings of the IEEE 1996 National Aerospace and Electronics Conference NAECON 1996, Dayton, OH, USA, 20–22 May 1996; Volume 1, pp. 429–434.

59. Arabadzhiyska, E.; Tursun, O.T.; Myszkowski, K.; Seidel, H.P.; Didyk, P. Saccade landing position prediction for gaze-contingent rendering. *ACM Trans. Graph. (TOG)* **2017**, *36*, 50. [CrossRef]

60. Reingold, E.M.; Loschky, L. Saliency of peripheral targets in gaze-contingent multiresolutional displays. *Behav. Res. Methods Instrum. Comput.* **2002**, *34*, 491–499. [CrossRef] [PubMed]

61. Sitzmann, V.; Serrano, A.; Pavel, A.; Agrawala, M.; Gutierrez, D.; Masia, B.; Wetzstein, G. How do people explore virtual environments? *arXiv* **2017**, arXiv:1612.04335.

62. Rai, Y.; Le Callet, P.; Guillotel, P. Which saliency weighting for omni directional image quality assessment? In Proceedings of the 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), Erfurt, Germany, 31 May–2 June 2017; pp. 1–6.

63. David, E.J.; Da Silva, M.P.; Le Callet, P. A new toolbox to process gaze and head motion in virtual reality. 2020, in preparation.

64. Gutiérrez, J.; David, E.J.; Coutrot, A.; Da Silva, M.P.; Le Callet, P. Introducing UN Salient360! Benchmark: A platform for evaluating visual attention models for 360° contents. In Proceedings of the 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), Cagliari, Italy, 29 May–1 June 2018; pp. 1–3.

65. David, E.J.; Gutiérrez, J.; Coutrot, A.; Da Silva, M.P.; Callet, P.L. A dataset of head and eye movements for 360° videos. In Proceedings of the 9th ACM Multimedia Systems Conference, Amsterdam, The Netherlands, 12–15 June 2018; pp. 432–437.

66. Larsson, L.; Schwaller, A.; Nyström, M.; Stridh, M. Head movement compensation and multi-modal event detection in eye-tracking data for unconstrained head movements. *J. Neurosci. Methods* **2016**, *274*, 13–26. [CrossRef] [PubMed]

67. Lappi, O. Eye movements in the wild: Oculomotor control, gaze behavior & frames of reference. *Neurosci. Biobehav. Rev.* **2016**, *69*, 49–68. [PubMed]

68. Salvucci, D.D.; Goldberg, J.H. Identifying fixations and saccades in eye-tracking protocols. In Proceedings of the 2000 symposium on Eye tracking research & applications, Gardens, FL, USA, 6–8 November 2000; pp. 71–78.

69. Nyström, M.; Holmqvist, K. An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behav. Res. Methods* **2010**, *42*, 188–204. [CrossRef]

70. Baayen, R.H.; Davidson, D.J.; Bates, D.M. Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* **2008**, *59*, 390–412. [CrossRef]

71. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.

72. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting linear mixed-effects models using lme4. *arXiv* **2014**, arXiv:1406.5823.

73. Salthouse, T.A.; Ellis, C.L. Determinants of eye-fixation duration. *Am. J. Psychol.* **1980**, *93*, 207–234. [CrossRef]

74. Manor, B.R.; Gordon, E. Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. *J. Neurosci. Methods* **2003**, *128*, 85–93. [CrossRef]

75. Inhoff, A.W.; Radach, R. Definition and computation of oculomotor measures in the study of cognitive processes. In *Eye Guidance in Reading and Scene Perception*; Elsevier: Amsterdam, The Netherlands, 1998; pp. 29–53.

76. van Diepen, P.; d'Ydewalle, G. Early peripheral and foveal processing in fixations during scene perception. *Vis. Cogn.* **2003**, *10*, 79–100. [CrossRef]

77. Nuthmann, A.; Smith, T.J.; Engbert, R.; Henderson, J.M. CRISP: A computational model of fixation durations in scene viewing. *Psychol. Rev.* **2010**, *117*, 382. [CrossRef] [PubMed]

78. Smith, T.J.; Henderson, J.M. Facilitation of return during scene viewing. *Vis. Cogn.* **2009**, *17*, 1083–1108. [CrossRef]

79. Smith, T.J.; Henderson, J.M. Does oculomotor inhibition of return influence fixation probability during scene search? *Atten. Percept. Psychophys.* **2011**, *73*, 2384–2398. [CrossRef] [PubMed]

80. Shic, F.; Scassellati, B.; Chawarska, K. The incomplete fixation measure. In Proceedings of the 2008 Symposium on Eye Tracking Research & Applications, Savannah, GA, USA, 26–28 March 2008; pp. 111–114.

81. Blignaut, P. Fixation identification: The optimum threshold for a dispersion algorithm. *Atten. Percept. Psychophys.* **2009**, *71*, 881–895. [CrossRef] [PubMed]

82. Komogortsev, O.V.; Jayarathna, S.; Koh, D.H.; Gowda, S.M. Qualitative and quantitative scoring and evaluation of the eye movement classification algorithms. In Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, Austin, TX, USA, 22–24 March 2010; pp. 65–68.

83. Bruce, N.D.; Tsotsos, J.K. Saliency, attention, and visual search: An information theoretic approach. *J. Vis.* **2009**, *9*, 5. [CrossRef]

84. Sağlam, M.; Lehnen, N.; Glasauer, S. Optimal control of natural eye-head movements minimizes the impact of noise. *J. Neurosci.* **2011**, *31*, 16185–16193. [CrossRef]

85. Peterson, M.F.; Eckstein, M.P. Looking just below the eyes is optimal across face recognition tasks. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E3314–E3323. [CrossRef]

86. Currie, C.B.; McConkie, G.W.; Carlson-Radvansky, L.A.; Irwin, D.E. The role of the saccade target object in the perception of a visually stable world. *Percept. Psychophys.* **2000**, *62*, 673–683. [CrossRef]

87. Drissi-Daoudi, L.; Öğmen, H.; Herzog, M.H.; Cicchini, G.M. Object identity determines trans-saccadic integration. *J. Vis.* **2020**, *20*, 33. [CrossRef]

88. Van Der Stigchel, S.; De Vries, J.P. There is no attentional global effect: Attentional shifts are independent of the saccade endpoint. *J. Vis.* **2015**, *15*, 17. [CrossRef] [PubMed]

89. Wollenberg, L.; Deubel, H.; Szinte, M. Visual attention is not deployed at the endpoint of averaging saccades. *PLoS Biol.* **2018**, *16*, e2006548. [CrossRef] [PubMed]

90. Schut, M.J.; Van der Stoep, N.; Fabius, J.H.; Van der Stigchel, S. Feature integration is unaffected by saccade landing point, even when saccades land outside of the range of regular oculomotor variance. *J. Vis.* **2018**, *18*, 6. [CrossRef]

91. Asfaw, D.S.; Jones, P.R.; Mönter, V.M.; Smith, N.D.; Crabb, D.P. Does Glaucoma Alter Eye Movements When Viewing Images of Natural Scenes? A Between-Eye Study. *Investig. Ophthalmol. Vis. Sci.* **2018**, *59*, 3189–3198. [CrossRef] [PubMed]

92. Pannasch, S.; Velichkovsky, B.M. Distractor effect and saccade amplitudes: Further evidence on different modes of processing in free exploration of visual images. *Vis. Cogn.* **2009**, *17*, 1109–1131. [CrossRef]

93. Mills, M.; Hollingworth, A.; Van der Stigchel, S.; Hoffman, L.; Dodd, M.D. Examining the influence of task set on eye movements and fixations. *J. Vis.* **2011**, *11*, 17. [CrossRef]

94. Smith, T.J.; Mital, P.K. Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. *J. Vis.* **2013**, *13*, 16. [CrossRef]

95. Unema, P.J.; Pannasch, S.; Joos, M.; Velichkovsky, B.M. Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Vis. Cogn.* **2005**, *12*, 473–494. [CrossRef]

96. Fischer, T.; Graupner, S.T.; Velichkovsky, B.M.; Pannasch, S. Attentional dynamics during free picture viewing: Evidence from oculomotor behavior and electrocortical activity. *Front. Syst. Neurosci.* **2013**, *7*, 17. [CrossRef]

97. Ehrlich, J.R.; Ojeda, L.V.; Wicker, D.; Day, S.; Howson, A.; Lakshminarayanan, V.; Moroi, S.E. Head-mounted display technology for low-vision rehabilitation and vision enhancement. *Am. J. Ophthalmol.* **2017**, *176*, 26–32. [CrossRef]

98. Bennett, C.R.; Bailin, E.S.; Gottlieb, T.K.; Bauer, C.M.; Bex, P.J.; Merabet, L.B. Virtual reality based assessment of static object visual search in ocular compared to cerebral visual impairment. In Proceedings of the International Conference on Universal Access in Human-Computer Interaction, Las Vegas, NV, USA, 15–20 July 2018; Springer: Cham, Switzerland, 2018; pp. 28–38.

99. Chow-Wing-Bom, H.; Dekker, T.M.; Jones, P.R. The worse eye revisited: Evaluating the impact of asymmetric peripheral vision loss on everyday function. *Vis. Res.* **2020**, *169*, 49–57. [CrossRef] [PubMed]