Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Sound Improves Peripheral Detection: Under a Narrowed Functional Visual Field in Simulated Visual-Field Impairment

Hikari Takebayashi

hikari.takebayashi@gmail.com

Ritsumeikan University

Yuji Wada

Ritsumeikan University

Article

Keywords: Attention, Audiovisual Interaction, Functional Field of View, Useful Field of View, Virtual Reality, Visual Impairment

Posted Date: March 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8754299/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Observers search for a target by shifting the “Functional Visual Field” (FVF), a limited window around fixation. FVF is defined as a region that elicits a peripheral target conjecture during a search and a saccade toward that location. Using an eye-tracked head-mounted display, we simulated virtual visual constriction—a primary symptom of glaucoma and retinitis pigmentosa—by applying a gaze-contingent peripheral blur. In sighted individuals, the FVF radius during a visual search is ~ 5–10°. We investigated (1) whether the constriction in the visible field narrowed the FVF, and (2) whether synchronized sound provided under the visual constriction facilitated detection of a peripheral target. In Experiment 1, we estimated the FVF radius from scan paths, target landing rates, and detection rates, and found a monotonic decrease with increasing constriction level. In Experiment 2, a brief binaural pure tone was presented when the distance between fixation and a target fell below the estimated FVF radius, and the search display was simultaneously extinguished. Observers judged whether a target was present in the periphery just before extinction. Sound synchronization increased hit rates without changing false-alarm rates relative to a no-sound condition, indicating that auditory cues selectively improve visual detection near the FVF boundary.

Introduction

In visually information-rich environments, humans repeatedly fixate on specific objects. During a search, one’s gaze may fall on an object either randomly or because it attracts attention in peripheral vision. In the latter case, the observer brings the object under the fovea for high-resolution analysis [1, 2]. “Attention” is a broad construct. Relatedly, the “cognitive” visual field around fixation, often termed the Useful Field of View (UFOV) [3, 4] or the Functional Visual Field (FVF) [5], remains imprecisely defined in the attention literature but is commonly considered as the region from which peripheral information can be detected and read out during fixation. This region is typically described as a horizontally elongated ellipse with a radius of approximately 30–60° from the focal point [6, 7]; However, it is not fixed. For example, when radiologists search mammograms, the FVF assumes a vertically elongated ellipse aligned with the portrait orientation of the image [8].

Primarily, “attention” functions as a system for selection in the information-rich environment [1, 9, 10]. According to William James in The Principles of Psychology, “Everyone knows what attention is. It is taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought” [11]. In other words, attention entails choosing some inputs while ignoring others, and the FVF is precisely the region that directly contributes to the selection and filtering of information. Nevertheless, many experiments on the FVF have presented a single light point or letters on a quiet background at some eccentricity from fixation and measured detection performance. Such methods effectively assess form perception or acuity in the periphery and thus conflict with the intended focus of FVF research. Consequently, the widely cited FVF radius of 30–60° may in fact conflate the FVF with the spatial extent of peripheral form perception. Recent studies address this issue by defining the FVF in dynamic visual search as the region around fixation that the observer actively covers with covert

processing and within which the next saccade toward a target is directly triggered [8, 12]. In a simple search task for one “T” among 79 “L”s on a display extending 40.2° horizontally and 22.6° vertically, the FVF radius deployed by observers was estimated to be approximately 5–10° [12]. In a visual search without the constraint of static fixation, the FVF of sighted individuals is much narrower than the primitive visible field (a horizontal radius of 100°).

The present study examined the following two questions:

(1) Does the FVF of individuals with visual-field loss, particularly concentric constriction, contract in conjunction with the size of their visible field? Glaucoma, the leading cause of acquired blindness in many countries, is an optic neuropathy whose primary symptoms often begin with arcuate scotomas and gradually progress to peripheral field constriction, sometimes culminating in blindness. Retinitis pigmentosa, characterized by photoreceptor degeneration, likewise produces marked peripheral constriction. Since progression is typically slow in both conditions, individuals in early stages may be unaware of their symptoms. One hypothesis is that the already small FVF is essentially unchanged if modest losses occur within the ~ 200° horizontal extent of the visible field. Conversely, given the size of the FVF relative to the overall visible field, it is assumed that individuals with constriction have a reduced visible field, leading to further FVF reduction. If the latter hypothesis is supported, standard perimetric measurements of field extent could systematically overestimate patients’ functional visual capacity, including the cognitive components of vision.

(2) If visual-field constriction is accompanied by a concomitant narrowing of the FVF, can multisensory interactions restore it? Audiovisual interactions have driven research on crossmodal perception. Spatiotemporal congruence between vision and audition can influence perception, including auditory and motion aftereffects [13–15]. Nevertheless, strict spatial alignment is not always required for visual facilitation by auditory signals: when a deviant tone is inserted within a continuous binaural tone sequence, the salience of simultaneously presented visual events increases and their detection is enhanced [16–19]. Additionally, audiovisual synchronization is a useful component of visual rehabilitation in hemianopia [20], suggesting that similar mechanisms could benefit individuals with visual field loss.

In the present study, the FVF was defined as the region during dynamic search in which a conjecture that a target may be present in the periphery is formed and triggers a saccade toward that location [12]. By definition, this entails eventual gaze landing on the target. Across two experiments, we used virtual reality (VR) to implement head- and eye-tracking visual-field constriction. Existing studies of the simulation of visual field loss, including those using VR techniques, often occluded parts of images or videos with black or gray patches, effectively creating a display with hidden content [21–24]. Admittedly, some patients with physiological visual-field impairments report small “shadows,” and a subset of individuals with total blindness lack any light perception. However, general clinical reports indicate that scotomas in glaucoma and retinitis pigmentosa are not typically perceived as “black.” Rather, vision commonly begins with partial blur or missing segments with retaining light perception, and the localized

blur connects into a tunnel-like pattern in mid stages [25–27]. Moreover, even individuals with total blindness who lack form vision often retain some awareness of light [28]. Some describe “seeing no images,” yet feeling as if “walking in heavenly light and white” on bright days, and “it's easiest to see something” on cloudy days (words from my friend I. K., who has experienced retinitis pigmentosa). Based on these observations, we simulated mid-stage glaucoma and retinitis pigmentosa by implementing a gradual increase in blur from the fovea to the periphery. Experiment 1 examined whether the FVF is reduced under constricted vision (Fig. 1B) relative to normal-vision (Fig. 1A). Experiment 2 assessed whether presenting a brief sound at the moment that gaze approached a peripheral target—under constriction—improved target detection (Fig. 1C).

Results

Experiment 1

Behavioral Responses

In the foraging task, mean detection time increased monotonically from the normal-vision condition to the severe-constriction condition (Fig. 2A). However, the detection rate, i.e., the probability of finding the target within 15 s, did not decrease (Fig. 2B). A repeated-measures ANOVA on mean detection time revealed a significant effect of visual-field condition, $F(2, 34) = 9.83$, $p < .001$, $\eta_p^2 = 0.37$. Bonferroni-corrected pairwise comparisons showed longer detection times in the mild-constriction condition (6.70 s) than in the normal-vision condition (6.01 s), $t(17) = 3.15$, $p_{\text{Bonferroni}} = .02$, Cohen’s $d = 0.72$, and longer times in the severe-constriction condition (6.90 s) than in the normal-vision condition, $t(17) = 4.20$, $p_{\text{Bonferroni}} = .002$, Cohen’s $d = 0.93$. The two constriction conditions did not differ, $t(17) = 1.02$, $p_{\text{Bonferroni}} = .97$, Cohen’s $d = 0.21$. Mean detection rates were ~ 0.77–0.79 and did not differ significantly across conditions, $F(2, 34) = 0.43$, $p = .65$, $\eta_p^2 = 0.03$.

In the search task, both indices exhibited trends similar to those in the foraging task; however, mean values did not differ significantly across the three conditions. Mean detection times were 7.40 s (normal-vision), 7.70 s (mild constriction), and 8.10 s (severe constriction), $F(2, 34) = 2.25$, $p = .12$, $\eta_p^2 = 0.12$ (Fig. 2C). Mean detection rates were 0.83, 0.81, and 0.79, respectively, $F(2, 34) = 1.19$, $p = .32$, $\eta_p^2 = 0.07$ (Fig. 2D). Thus, behavioral performance declined with increasing constriction level in both tasks, but the changes were not dramatic.

Eye movements

The search task contained 50 target-present trials per visual-field condition. After classifying eye movements into the three saccade classes, the resulting sample sizes were insufficient for stable estimates. Therefore, the results of the foraging task are reported. Figure 2E shows heat maps of

saccade start points with all endpoints aligned at the origin, pooling data across participants. For search saccades, the normal-vision condition produced a diamond-shaped distribution with near-equal horizontal and vertical extents, with most start points within ~ 15°. In contrast, under the two constriction conditions the diamond broadened horizontally, with saccades arriving from as far as ~ 20° horizontally. This gross pattern was reflected in other saccade classes, suggesting that the shape of the visual-field loss tunes oculomotor behavior.

A key difference from conventional FVF studies is that we allowed both eye and head movements. Accordingly, Fig. 2F displays the analysis of a composite measure equal to the per-frame saccade length plus the head's yaw rotation. Although this yields larger values than saccade length with a fixed head, it indexes more natural head–eye coordinated movement. On this measure, mean amplitudes were generally smallest in the normal-vision condition across all saccade classes. For example, post-target saccades averaged 2° in the normal-vision condition but increased to 3° in both constriction conditions. Similarly, search and targeting saccades averaged just under 4° in normal-vision and increased to 5–6° under constrictions. Distributions were compared using Kolmogorov–Smirnov tests within each visual-field condition across saccade classes, using α = .017. All pairwise comparisons were significant ($p < .001$); however, differences between search and targeting saccades in severe-constriction condition were weaker than those for the other pairings ($p = .002$).

Head movements

Figure 3A shows the across-participant mean of the total amount of head yaw rotation per trial, and Fig. 3B shows the across-participant mean of the mean angular velocity per trial in the foraging task. A repeated-measures ANOVA revealed a significant effect of visual-field condition on total rotation amount, $F(2, 34) = 10.99$, $p < .001$, $\eta_p^2 = 0.39$. Bonferroni-corrected comparisons showed greater rotation in the mild constriction than in normal-vision, $t(17) = 3.26$, $p_{Bonferroni} = .014$, Cohen's $d = 0.58$, and greater rotation in the severe constriction than in normal-vision, $t(17) = 3.82$, $p_{Bonferroni} = .004$, Cohen's $d = 0.82$; the two constriction conditions did not differ, $t(17) = 1.75$, $p_{Bonferroni} = .295$, Cohen's $d = 0.25$. Additionally, ANOVA was significant for mean angular velocity, $F(2, 34) = 6.79$, $p = .003$, $\eta_p^2 = 0.29$. Pairwise tests indicated higher angular velocity in severe constriction than in normal-vision, $t(17) = 3.33$, $p_{Bonferroni} = .012$, Cohen's $d = 0.46$, with no significant differences between normal-vision and mild constriction, $t(17) = 2.56$, $p_{Bonferroni} = .062$, Cohen's $d = 0.33$, nor between the two constriction conditions, $t(17) = 1.10$, $p_{Bonferroni} = .866$, Cohen's $d = 0.13$. Therefore, as the level of visual-field constriction increased, participants invested additional head movement to inspect the less visible periphery. Indeed, as noted in the behavioral results, foraging-task detection rates did not differ significantly across conditions, consistent with the possibility that participants compensated by moving their heads (and eyes) more drastically to scan a wider area and thereby maintain stable detection performance (see Figs. 2B and 2F).

Figure 3C and 3D illustrate the relative timing of head and eye movement initiation immediately following trial onset during the foraging task. These plots are based on pooled data from all trials across 18 participants. To evaluate the time difference between head and eye movements immediately after the start of trials, this data allows all eye movements and includes all responses, irrespective of accuracy. Moreover, this index allows all three-dimensional head rotations. For both movement types, we applied an angular-acceleration threshold of 0.1 rad/s², corresponding to the minimum level at which the semicircular canals detect head angular acceleration [31]. Using Kolmogorov–Smirnov tests with $\alpha = .017$, we found significant differences in distribution between conditions: normal-vision vs. mild constriction ($D = 0.32$, $p < 2.2\times10^{-16}$), normal-vision vs. severe constriction ($D = 0.27$, $p < 2.2\times10^{-16}$), and mild vs. severe constriction ($D = 0.06$, $p = .007$). Figure 3C plots the proportions of trials in which eye movements led, head movements led, or both began simultaneously. As visible in Fig. 3D, the first peak occurred at a 0 ms lag, indicating synchronous onset. Although the statistical tests show that the distributions differ across conditions, a macroscopic view of Fig. 3D highlights a clear shift in the second peak between the normal-vision and the two constriction conditions. This second peak reflects the head lag relative to eye-movement onset: 21.86 ms in the normal-vision and approximately 33 ms in both constriction conditions. Thus, at the start of search, the eyes typically move first —and visual-field constriction amplified this lag.

Estimation of Attentional FVF radii

This analysis estimated the FVF radius required to achieve the observed detection rates in the search task [12]. The detection rates with a set size of 165 were 0.83 in the normal-vision condition, 0.81 in the mild-constriction condition, and 0.79 in the severe-constriction condition.

First, using each participant's eye movement data from the foraging task, we binned the distance between the current gaze point and the target for every frame in 1° increments. For each distance bin, we counted the total number of saccades and the number that actually landed on the target. From these indices, we computed the probability that at least one of the next three saccades would land on the target for each distance (Fig. 3E). We adopted the "one out of three" criterion because the probability that the next single saccade landed on the target was too low to be informative. As shown in Fig. 3E, saccades rarely reached the target once the gaze–target distance exceeded 5°. Colored solid lines show across-participant means, which were used in subsequent analyses. Next, we extracted scan paths from all true-negative trials from the search task. We only used true-negative trials to avoid early termination, as in target-present trials, and obtain scan paths that more fully covered the display. We then synthetically generated target-present displays and overlaid them on every scan path, computing the framewise distance between the current fixation and the pseudo-target. Finally, for attentional FVF radii spanning 1–15° (1° bins), we used the mean landing rates (Fig. 3E) to compute the probability of generating a hit in this pseudo-search setting for each radius. When the attentional FVF radius is small, the chance that the scan path covers the target within the FVF is low, yielding few hits. As the radius increases, target coverage improves and, in principle, multiple hits can occur within a single trial.

Consequently, the estimated FVF radius needed to reproduce the observed search-task detection rates decreased monotonically across conditions: 7.64° in the normal-vision, 7.18° in mild constriction, and 6.67° in severe constriction (Fig. 3F).

Experiment 2

Using data from the final sample (N = 36), the mean hit rate was 0.50 and 0.55 in the no-sound and sound conditions, respectively. The mean false-alarm rate was 0.08 in both conditions. Signal detection analysis (Fig. 4A) yielded a mean discriminability (d') of 1.59 in no-sound condition and 1.78 in sound condition; a paired t-test indicated a trend toward significance, $t(35) = -1.83$, $p = .076$, Cohen's $d = -0.31$ (Fig. 4B). The mean criterion (c) was 0.80 in no-sound condition and 0.74 in sound condition, with no significant difference, $t(35) = 1.18$, $p = .247$, Cohen's $d = 0.20$ (Fig. 4C).

A generalized linear mixed model (logit link) was used to comprehensively examine how signal (target) and sound presence affected response probability. Fixed effects were signal presence, sound presence, and their interaction. For participants, we specified random intercepts and random slopes for two main effects, allowing correlations among random effects. The signal × sound interaction was significant ($z = 2.00$, $p = .050$). We then converted the model's logit-scale linear predictors to probability of a "Yes" response and tested simple effects with α = .025. In target-absent trials, sound did not change response probability, Odds Ratio, OR = 1.10, $z = 0.60$, $p = .563$. In target-present trials, sound produced a significant difference, OR = 0.79, $z = -2.82$, $p = .005$ (Fig. 4D). In summary, these results indicate a positive crossmodal effect: inserting a sound during search increased hits while not increasing false alarms.

Discussion

Traditional definitions of the FVF have been grounded in a binary principle (0/1), implying that items outside the FVF are not cognitively accessible [32]. Moreover, some tasks have measured the FVF as the ability to identify peripheral stimuli while looking at a central stimulus on a quiet background [33–35]. Such paradigms, however, can conflate peripheral form perception with the attentional FVF. The present study considered the FVF as the region within a dynamic search wherein peripheral targets are detected probabilistically and from which the observer's gaze is driven toward those targets.

Experiment 1 applied a VR-based simulation of visual field impairment and obtained two findings similar to the earlier FVF study [12] and two novel findings. The first common finding is that, in the normal-vision condition, the estimated FVF radius was essentially the same in our free-head-movement search as in prior fixed-head search: both computations placed the radius at ~ 7–8°. However, the proportion of saccades that actually landed on a target from distances > 5° was very low (Fig. 3E). Therefore, the FVF radius in visual search should be reasonably considered within a 5–10° range. The second commonality

is the flexibility of saccade start-point distributions. Analogous to the vertical elongation observed when radiologists scanned mammograms in portrait orientation [8], our Experiment 1 showed a horizontal elongation of start-point distributions that aligned with the shape of the virtual field constriction. This pattern suggests that oculomotor behavior is adjusted to the demands of the search display and the observer's physiological field geometry.

The first novel finding from Experiment 1 is that the attentional FVF narrowed with the increasing level of virtual field constriction. Notably, the estimated radii in the mild and severe-constriction conditions—7.18° and 6.67°, respectively—remained well within the visible field. Thus, the explanation "a target could not be seen due to strong blur" does not apply here. Although peripheral blur increased gradually with eccentricity, participants still had sufficient acuity to perceive the form of stimuli located 7.18° or 6.67° from fixation if they attempted. Nevertheless, the FVF curtailment may mean that the presence of peripheral blur reduces the motivation of deploying attention to the periphery. The second novel finding is that head–eye coordination varied with visual field condition. Using a VR-based approach, we removed head movement constraints and assessed natural head–eye coordination. Under simulated field constriction, the head lag relative to eye movements increased. In general, during cognitive tasks with gaze shifts ≤ 40°, eyes lead head, whereas situations in which head leads eyes (gaze shifts ≥ 60°) are reportedly rare, context-specific, and voluntary, such as lane changes while driving [36–38]. This asymmetry arises because the inertia acting on head movements is much greater than that on ocular movements [39]. This tendency was consistently observed across all visual field conditions, indicating that participants efficiently used motor resources. One interpretation of the increased head lag under constriction is strategic: anticipating a higher risk of misses with a narrowed field, participants may have exhaustively sampled the current field with their eyes before moving the head to the next region. The near-equivalence of detection rates across conditions in both the foraging and search tasks provides converging evidence that participants attempted to maintain performance despite constriction. Nevertheless, future work should test how the timing of head–eye coordination changes in free viewing without explicit task goals or under exogenous attention driven by stimulus saliency.

In Experiment 2, relative to display offset without sound, display offset with sound yielded more hits, indicating that auditory stimulation facilitated item detection at the edge of the FVF during dynamic search. Notably, the sound was not spatially congruent with a visual target: for example, even when a target lay 6.5° to the right of the current gaze point, the tone was not played to the right ear. To avoid explicitly cueing a specific direction—and thereby artificially helping the form perception at that location—we deliberately presented the sound binaurally, encouraging access to peripheral information in general without indicating the target item. Nevertheless, hit rates increased without an accompanying rise in false alarms, demonstrating a positive auditory effect on the peripheral allocation of visual attention. This pattern aligns with existing evidence that sounds can transiently boost the salience of concurrent visual information [16–19]. A limitation of Experiment 2 was the difficulty in localizing the stage of cognitive processing impacted by the auditory effect. Did the sound enhance encoding of visual information immediately before display offset? Or did elevated arousal make visual memory easier to

read out? Addressing these questions will require targeted tests—for example, probing aftereffects to determine how sound modulates information present during encoding in dynamic search tasks.

Another study limitation is that it lacks sufficient validation of whether our VR simulation is truly homologous to the perceptual experience of all patients with visual-field constriction. Patients' visual fields are highly diverse, making it difficult to capture all possible profiles. Several future issues remain, including multifarious field-simulation methods, characterizing patterns of head–eye coordination, and conducting basic research on how audiovisual interactions shape visual attention. Nevertheless, virtual reproduction of visual impairment can encompass not only constriction but also scotomas, metamorphopsia, and photophobia [29]. By systematically manipulating levels of field impairment within psychophysical paradigms, the present approach may enable principled evaluation of what observers find easy or hard to do, ultimately contributing to making benchmarks of functional (everyday) vision.

Methods

Experiment 1

Participants

Participants included 21 observers from Ritsumeikan University; however, three participants discontinued the experiment due to VR motion sickness. All experiments were completed by 18 participants (8 men, 10 women; age range: 18–36 years). The final sample size matched that in the existing FVF study [12]. All the participants had normal or corrected-to-normal vision. After the experiment, each participant received an Amazon gift card worth JPY 2,000 for their 2-h participation. This study was approved by the Institutional Review Board of the Ethics for Research Involving Human Subjects at Ritsumeikan University and performed in accordance with these guidelines.

Apparatus and stimuli

We used a head-mounted display (HMD) with a 120 Hz sampling rate of eye-tracking (HTC Vive Pro Eye, HTC Corporation, Taiwan). The HMD provided a horizontal field of view of 110°, a vertical field of view of 70°, a resolution of 2880 × 1600 pixels, and a 90 Hz refresh rate. A white search panel subtended 78° (horizontal) × 58° (vertical), on which 164 black "L"s and one "T" were arranged in a 15 × 11 grid. Each character subtended 2° per side, and the center-to-center spacing between the characters was 5°. The simulation of visual-field constriction was implemented on Windows 11 using Unreal Engine 4.26 with the VARID plugin (https://github.com/VARID-XR/VARID-plugin-ue4) [29, 30]. We assessed three conditions as a within-subject factor: a normal-vision condition (Fig. 1D), a mild-constriction condition (Fig. 1E), and a severe-constriction condition (Fig. 1F). The peripheral blur took an elliptical shape matched to the Vive Pro Eye's field of view, and the VR scene contained 252 "VF Maps"—points specifying blur location and values—covering the entire visual field. Blur was generated via a Gaussian pyramid, with linear interpolation between VF-Map points. In the mild-constriction condition, visibility was held at 34 dB from the center out to 9° horizontally (i.e., the maximum value equivalent to the

normal-vision condition), decreasing to 24 dB at 27°, such that blur strength increased linearly with eccentricity. In the severe-constriction condition, visibility remained at 34 dB out to 3° horizontally and decreased to 24 dB at 20° from the center.

Procedure

Participants wore an HMD and held the accompanying Vive controllers in dominant hand while seated. Using the SRanipal Runtime, an eye-movement calibration was performed for each participant before beginning the tasks. The experiment comprised two tasks: foraging and search tasks. In the foraging task, participants searched for a target “T” presented on every trial among 164 distractors. Using a controller-tip laser pointer, participants indicated the target and then pulled the rear trigger button to “confirm,” advancing to the next trial (Fig. 1G). The top of the search display showed the elapsed time; it changed from white to red after 15 seconds, but the trial did not proceed until the button was pressed. Participants were instructed to find the target as quickly as possible before the time turned red. For familiarization, participants engaged in four practice trials followed by 100 experimental trials with short breaks provided every 20 trials. This procedure was repeated under three visual-field conditions as a within-subject factor, yielding a total of 300 trials for the foraging task. Similar to this, four practice trials were followed by 100 experimental trials in one visual-field condition in the search task; however, half were target-present trials. Using the two handheld controllers, participants pressed the left/right button and right/left button if a target was present and absent, respectively. The search task was repeated under the three visual-field conditions and comprised 300 trials in total. The button mapping and the order of visual-field conditions were counterbalanced across participants; however, all participants performed the foraging task before the search task.

Data Analysis

For both the foraging and search tasks, trials with response times ≥ 15 s were excluded from each participant’s dataset. This criterion was adopted because such trials typically reflected exhaustive item-by-item search or lapses in attention such as sleepiness, rather than deployment of a bounded FVF. Next, we defined saccades as samples with velocities ≥ 40°/s and < 700°/s and removed all data frames outside this range. Values below the lower bound were considered as fixation or very slow movements, whereas values above the upper bound were likely due to tracking loss or other transient errors. All saccades were classified into three types [12]. A targeting saccade in each trial is any saccade for which the distance between fixation and the target fell below 4° within the last five frames preceding the final frame. The other saccades occurring before this frame were labeled search saccades, and the small eye movements from that frame to the final frame were labeled post-target saccades. Thus, search saccades reflect random movements made before the target is found, whereas post-target saccades can be regarded as fine-tuning following gaze landing on the target. However, in the foraging task, when the gaze did not approach within 4° during the final five frames and the participant nonetheless indicated the target with the laser pointer and pulled the trigger, we designated the final frame as a targeting saccade. In the search task, for analogous gaze patterns, we designated the final frame as a targeting saccade only on correct trials.

Experiment 2

Participants

Participants included another 42 observers from Ritsumeikan University; however, two participants discontinued the experiment due to VR motion sickness. Forty participants eventually completed all tasks (14 men and 26 women; age range: 18–30 years). Furthermore, data from 36 participants were used for analysis after outlier exclusion (12 men and 24 women; age range: 18–30 years). This final sample size had been determined in advance by parameters of effect size d = 0.5, α = 0.05, and power = 0.9 from G*Power 3.1. All the participants had normal or corrected-to-normal vision. After the experiment, each participant received an Amazon gift card worth JPY 1,500 for their 1–1.5-h participation. Experiment 2 was conducted under the same institutional ethics approval and guidelines as Experiment 1.

Apparatus and Stimuli

The same HMD and visual search task as in Experiment 1 was conducted. Experiment 2 tested whether auditory stimulation can facilitate peripheral attention even when visual field constriction narrows the FVF. Accordingly, the visual field condition was limited to the severe-constriction condition to directly address the experimental objective and to avoid participants' fatigue. The method for generating peripheral blur was identical to the previous experiment. The auditory stimulus was a 1259 Hz, 50 ms pure tone created with the free audio editor Audacity 3.7.4. The tone's frequency and timing followed prior studies on audiovisual interactions, including research about the "freezing phenomenon [16, 18, 19]." This sound was presented binaurally through headphones attached to the HMD.

Procedure

Participants performed the search task after gaze calibration. As shown in Fig. 1H, the key difference from Experiment 1 was that participants were prevented from looking directly at the target during search. In the task of searching for one "T" among 164 "L"s, the search display was extinguished with presentation of a binaural 50-ms pure tone at the instant that the distance between the current gaze point and the target center fell below 6.5°. Participants then used the two controllers to report the detection of the target (yes- left/right button; no-right/left button) in the periphery immediately after the display offset. However, if the gaze did not approach the target within 15 s, the search display automatically disappeared, and the participants were forced to respond regarding the presence or absence of a peripheral target. Unlike Experiment 1, the search was terminated after 15 s to reduce fatigue due to the experimental session's duration. Half of the trials were target-present. In target-absent trials, one of the 165 "L"s served as a pseudo-target: when the gaze point approached within 6.5° of that letter, the display was extinguished with the sound; the pseudo-target location was randomized across trials. The 6.5° criterion approximately matched the attentional FVF radius for the severe-constriction condition estimated in Experiment 1. The presence vs. absence of sound at display offset was a within-

subject factor. A blocked design was used with 150 trials each in the sound and no-sound conditions; block order and button mapping were counterbalanced across participants.

Data Analysis

Trials in which gaze failed to approach the target within 15 s (i.e., search terminated by timeout) and trials with total search durations ≤ 3 s were excluded. In the latter cases, the participant's initial gaze point happened to be deployed near the target at trial start, causing an immediate termination. Since participants were not mentally prepared to detect a peripheral target in such cases, these trials were omitted from analysis. Next, data of three participants whose accuracies were below 40% in both the sound and no-sound conditions (condition-wise mean accuracies of 25%, 23%, and 20%) were excluded, interpreting this as evidence of task misunderstanding or confusion about the response buttons. Finally, one participant's data whose discriminability (d') was negative, which indicates a systematic reversal in signal detection analysis —i.e., classifying signal as noise and vice versa—, was treated as an outlier and excluded.

Declarations

Acknowledgement

We wish to express our sincere appreciation to I. K., whose insightful comments on the state of the visual field in retinitis pigmentosa greatly contributed to this study. We are also deeply grateful to Dr. Satoshi Nakadomari for his valuable comments on the definition of the functional visual field.

Author contributions

Hikari Takebayashi and Yuji Wada developed the study concept. Additionally, Hikari Takebayashi contributed to the study design and performed the experiments, data collection, analysis, interpretation, and drafting of the manuscript. Yuji Wada contributed to funding acquisition and furnished resources such as experimental equipment. Both authors have approved the final version of the manuscript for submission.

Availability of data and materials

The experimental data used in this study can be found online at https://osf.io/2e6m3/overview?view_only=74732466174746de8cb92587c6315841.

Funding

This research was supported by the Ritsumeikan-Global Innovation Research Organization Fourth Phase Program and JST, CRONOS, Japan Grant Number JPMJCS24K3.

Competing Interest Statement

The authors declare no potential conflicts of interest regarding the research, authorship, and/or publication of this paper.

Ethical approval and informed consent statements

The Institutional Review Board of the Ethics for Research Involving Human Subjects at Ritsumeikan University approved the present study protocol. Written informed consent was obtained from all the participants in advance.

References

1. Neisser, U. *Cognitive Psychology* (Appleton-Century-Crofts, 1967).
2. Treisman, A. M. & Gelade, G. A feature-integration theory of attention. *Cogn. Psychol.* **12**, 97–136 (1980).
3. Mackworth, N. H. Visual noise causes tunnel vision. *Psychon Sci.* **3**, 67–68 (1965).
4. Sanders, A. F. Some aspects of the selective process in the functional visual field. *Ergonomics* **13**, 101–117 (1970).
5. Ikeda, M. & Takeuchi, T. Influence of foveal load on the functional visual field. *Percept. Psychophys*. **18**, 255–260 (1975).
6. Harada, Y. & Ohyama, J. Spatiotemporal characteristics of 360-degree basic attention. *Sci. Rep.* **9**, 16083 (2019).
7. Itoh, N., Sagawa, K. & Fukunaga, Y. Useful visual field at a homogeneous background for old and young subjects. *Gerontechnology* **8**, 42–51 (2009).
8. Wolfe, J. M., Wu, C. C., Li, J. & Suresh, S. B. What do experts look at and what do experts find when reading mammograms? *J. Med. Imaging (Bellingham)*. **8**, 045501 (2021).
9. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
10. Wolfe, J. & Horowitz, T. Five factors that guide attention in visual search. *Nat. Hum. Behav.* **1**, 0058 (2017).
11. James, W. *The Principles of Psychology, Vol. 1* (Henry Holt and Company, 1890).
12. Wu, C. C. & Wolfe, J. M. The Functional Visual Field(s) in simple visual search. *Vis. Res.* **190**, 107965 (2022).
13. Kitagawa, N. & Ichihara, S. Hearing visual motion in depth. *Nature* **416**, 172–174 (2002).
14. Park, M., Blake, R., Kim, Y. & Kim, C. Y. Congruent audio-visual stimulation during adaptation modulates the subsequently experienced visual motion aftereffect. *Sci. Rep.* **9**, 19391 (2019).
15. Park, M., Blake, R. & Kim, C. Y. Audiovisual interactions outside of visual awareness during motion adaptation. *Neurosci. Conscious.* niad027 (2024). (2024).

16. Takebayashi, H. & Wada, Y. Effects of audiovisual temporal synchronization on visual experience of the non-dominant eye. *Cogn. Process.. Advance online publication*. (2025).
17. Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W. & Theeuwes, J. Pip and pop: nonspatial auditory signals improve spatial visual search. *J. Exp. Psychol. Hum. Percept. Perform.* **34**, 1053–1065 (2008).
18. Vroomen, J. & de Gelder, B. Sound enhances visual perception: cross-modal effects of auditory organization on vision. *J. Exp. Psychol. Hum. Percept. Perform.* **26**, 1583–1590 (2000).
19. Wada, Y. et al. Sound enhances detection of visual target during infancy: a study using illusory contours. *J. Exp. Child. Psychol.* **102**, 315–322 (2009).
20. Rowland, B. A., Bushnell, C. D., Duncan, P. W. & Stein, B. E. Ameliorating hemianopia with multisensory training. *J. Neurosci.* **43**, 1018–1026 (2023).
21. Lavalle, L. K., Pourhashemi, N. & Cleworth, T. W. The relationship between a simulated glaucoma impairment and postural threat on quiet stance. *Virtual Real.* **29**, 31 (2025).
22. Neugebauer, A., Stingl, K., Ivanov, I. & Wahl, S. Influence of systematic gaze patterns in navigation and search tasks with simulated retinitis pigmentosa. *Brain Sci.* **11**, 223 (2021).
23. Ricci, F. S. et al. Virtual reality as a means to explore assistive technologies for the visually impaired. *PLOS Digit. Health*. **2**, e0000275 (2023).
24. Veerkamp, K., Müller, D., Pechler, G. A., Mann, D. L. & Olivers, C. N. L. The effects of simulated central and peripheral vision loss on naturalistic search. *J. Vis.* **25**, 6 (2025).
25. Crabb, D. P., Smith, N. D., Glen, F. C., Burton, R. & Garway-Heath, D. F. How does glaucoma look? patient perception of visual field loss. *Ophthalmology* **120**, 1120–1126 (2013).
26. Peli, E., Goldstein, R. & Jung, J. H. The invisibility of scotomas I: The carving hypothesis. *Optom. Vis. Sci.* **100**, 515–529 (2023).
27. Taylor, D. J., Edwards, L. A., Binns, A. M. & Crabb, D. P. Seeing it differently: self-reported description of vision loss in dry age-related macular degeneration. *Ophthalmic Physiol. Opt.* **38**, 98–105 (2018).
28. Lettieri, G., Calce, R. P., Giraudet, E. & Collignon, O. Visual experience shapes bodily representation of emotion. *Emotion* **25**, 657–670 (2025).
29. Jones, P. R., Somoskeöy, T., Chow-Wing-Bom, H. & Crabb, D. P. Seeing other perspectives: evaluating the use of virtual and augmented reality to simulate visual impairments (OpenVisSim). *npj Digit. Med.* **3**, 32 (2020).
30. Jones, P. R. & Ometto, G. Degraded Reality: using VR/AR to simulate visual impairments. IEEE Workshop on Augmented and Virtual Realities for Good (VAR4Good). (Reutlingen, Germany, 1–4. (2018).
31. Meiry, J. L. The vestibular system and human dynamic space orientation. NASA CR-628. *NASA Contract. Rep., NASA CR, US, NASA*, 1–192 (1966).
32. Andersen, G. J., Ni, R., Bian, Z. & Kang, J. Limits of spatial attention in three-dimensional space and dual-task driving performance. *Accid. Anal. Prev.* **43**, 381–390 (2011).

33. Park, G. D. & Reed, C. L. Nonuniform changes in the distribution of visual attention from visual complexity and action: A driving simulation study. *Perception* **44**, 129–144 (2015).
34. Seya, Y. & Watanabe, K. The minimal time required to process visual information in visual search tasks measured by using gaze-contingent visual masking. *Perception* **41**, 819–830 (2012).
35. Shamsi, F., Chen, V., Liu, R., Pergher, V. & Kwon, M. Functional field of view determined by crowding, aging, or glaucoma under divided attention. *Transl Vis. Sci. Technol.* **10**, 14 (2021).
36. Bischof, W. F., Anderson, N. C. & Kingstone, A. A tutorial: analyzing eye and head movements in virtual reality. *Behav. Res. Methods*. **56**, 8396–8421 (2024).
37. Doshi, A. & Trivedi, M. M. Head and eye gaze dynamics during visual attention shifts in complex environments. *J. Vis.* **12**, 9 (2012).
38. Solman, G. J. F., Foulsham, T. & Kingstone, A. Eye and head movements are complementary in visual selection. *R Soc. Open. Sci.* **4**, 160569 (2017).
39. Freedman, E. G. Coordination of the eyes and head during visual orienting. *Exp. Brain Res.* **190**, 369–387 (2008).

Figures

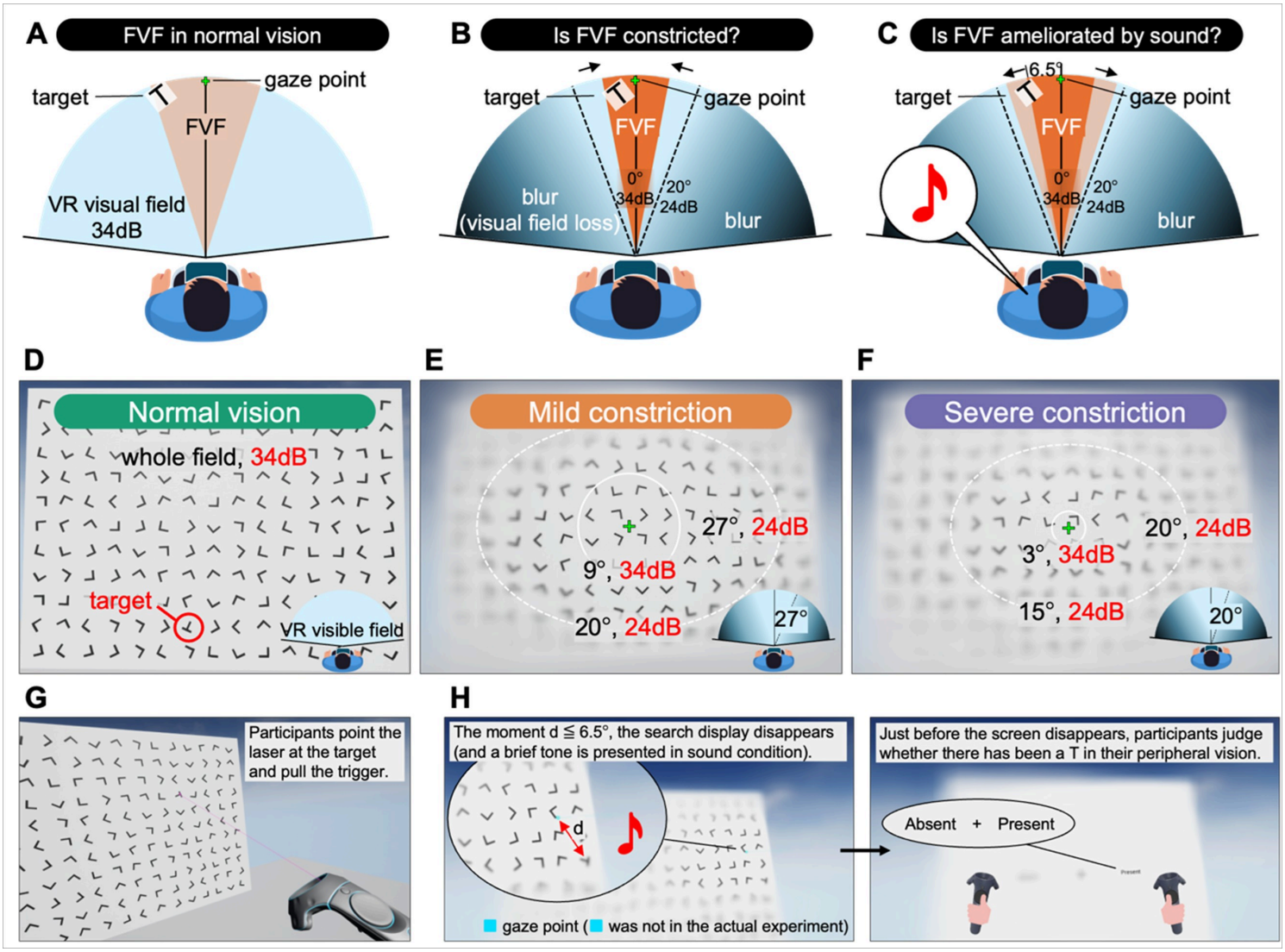


Figure 1

The hypotheses and task settings.

Note. A-C illustrate hypotheses. (A) FVF is a limited region in one's physiological field of view. A target inside this region is probabilistically detected. (B) The peripheral blur caused by visual field impairments constricts the FVF, leading to the difficulty of peripheral target detection even if its location is laid within a relatively high-visibility area. (C) The moment the gaze point approaches a target sufficiently (within the estimated FVF radius), a brief tone is presented binaurally, and the search screen disappears. Even with a narrowing FVF, audio-visual interaction helps the peripheral target detection. E-F represent virtual visual fields made by 252 points of VF map. The 15 × 11 invisible grids have 164 Ls and one T. (D) Normal-vision condition had sufficient clarity in the whole visual field. (E) In mild-constriction condition, the sensitivity to light gradually decreased from the fovea of 34 dB to 27° horizontal distance of 24 dB. (F) In severe-constriction condition, the depression of sensitivity was relatively drastic. The 20° horizontal distance already reached 24dB. (G) The VR environment of foraging task in Experiment 1. (H) The trial procedure in Experiment 2. When the distance between the present gaze point and a T fell below 6.5° (i.e., almost analogous to the estimated attentional FVF radius in severe-constriction condition), the

search display disappeared. Participants were required to scan the whole display within 15 s, or the display automatically disappeared. After searching, they judged the awareness of the presence of the peripheral target.

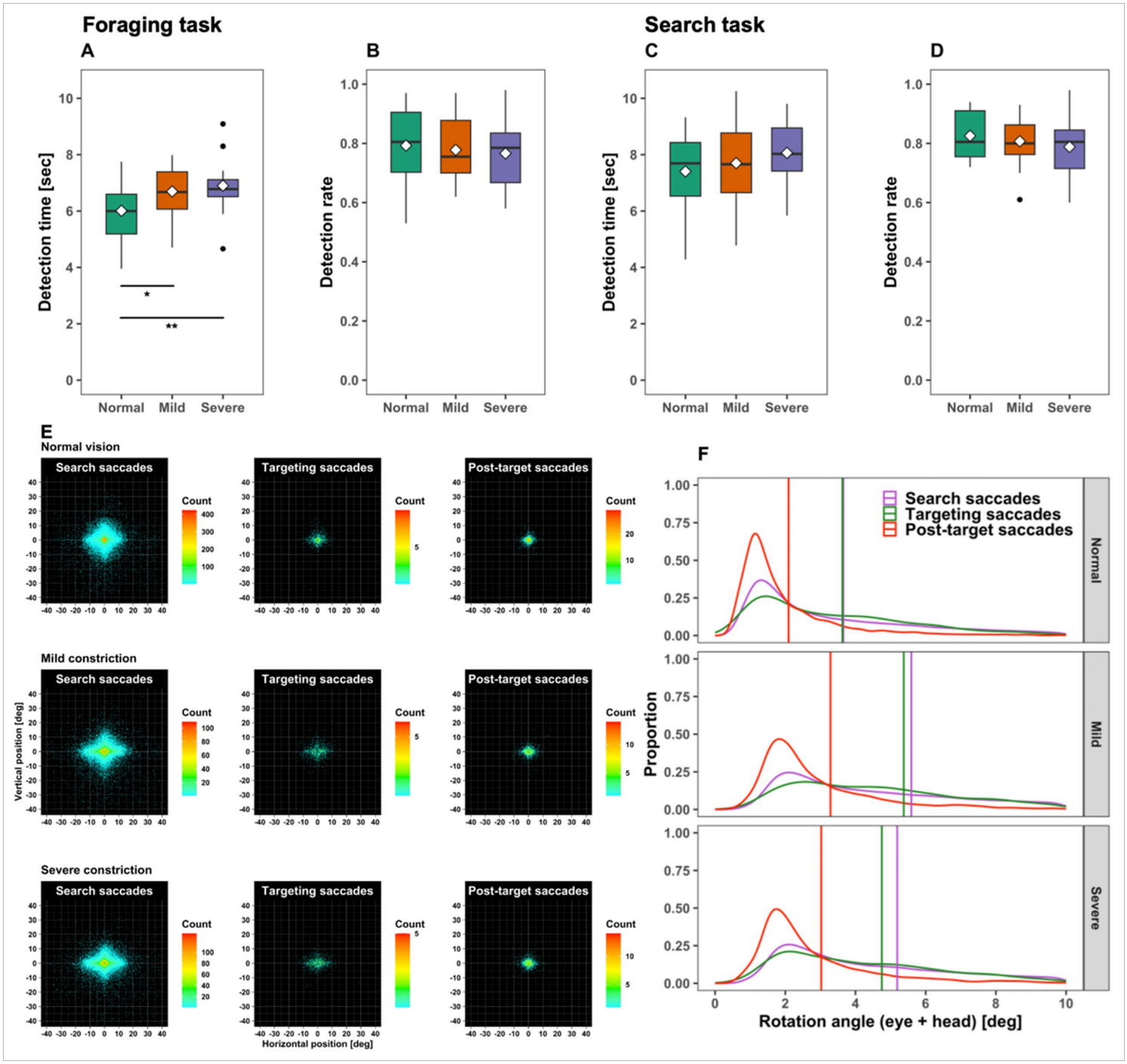


Figure 2

Results of behavioral responses and eye movements in Experiment 1.

Note. (A) Detection time and (B) detection rate in the foraging task. (C) Detection time and (D) detection rate in the search task. Diamonds represent the mean. (E) The [0, 0] is fixed as the endpoint of saccades. Colored points represent heatmaps of the start point from pooled participants' data in the foraging task.

Search saccades are random movements before the target detection. Targeting saccades are movements directed toward a peripheral target. Post-target saccades are fine-tuning after gaze landing on the target. (F) Distributions represent the sum of saccade length and head yaw angle per frame in the foraging task. The vertical lines are drawn on the mean.

* Adjusted $p < .05$, ** Adjusted $p < .01$

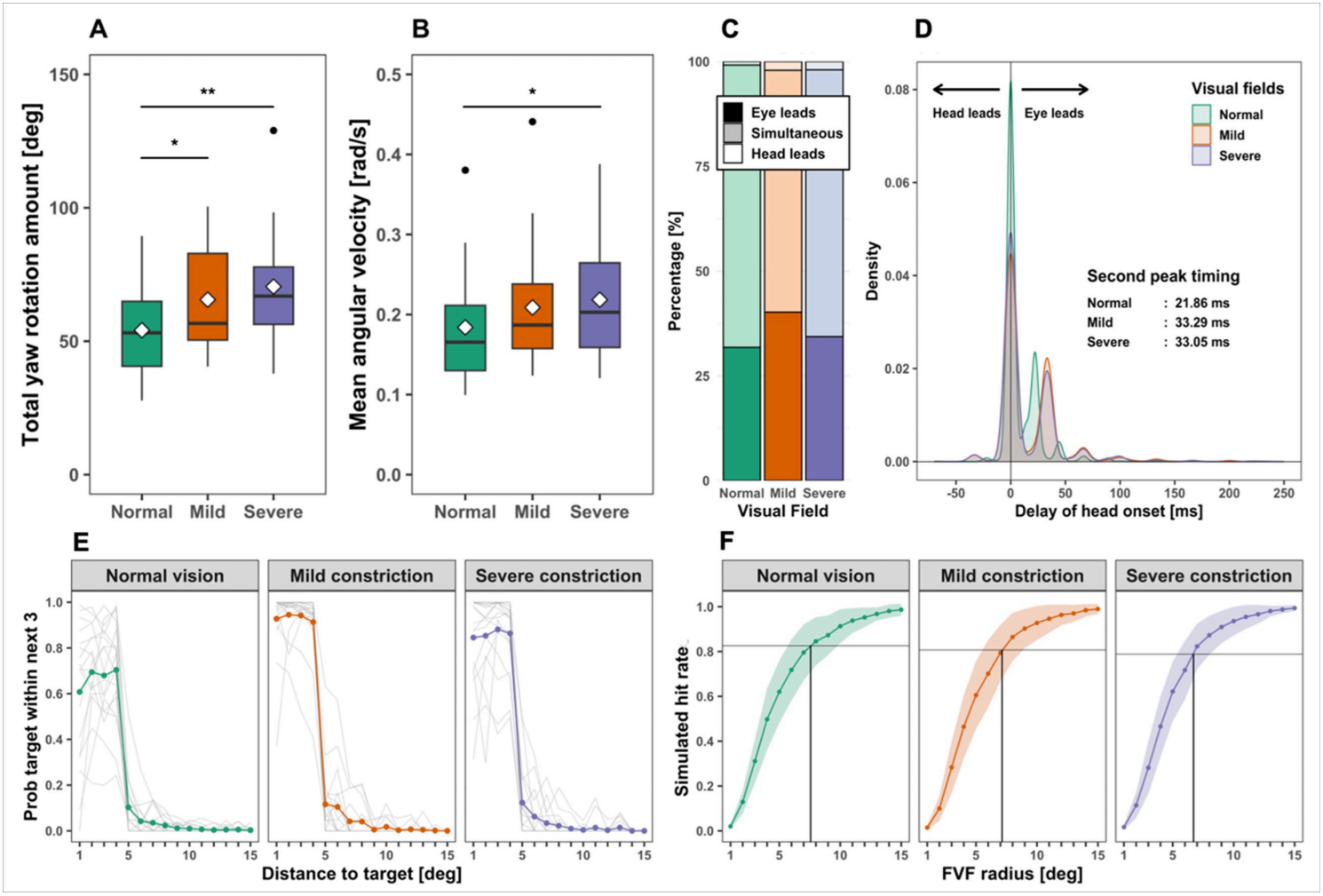


Figure 3

Results of head movements, landing probability of saccades, and estimated attentional FVF.

Note. (A) (B) Only yaw rotation, which is the head movement from side to side, is extracted in the foraging task. The indices represent mean values per single trial. (C) The proportions of trials in which eye movements led, head movements led, or both began simultaneously in the foraging task. (D) The smoothed distribution of onset lag between head and eye movements immediately after the foraging starts. (E) The probability that at least one of the next three saccades would land on the target for each distance in the foraging task. The colored solid line represents the mean, while light-gray solid lines represent individual participants. The landing probability drastically decreases over a 5°-distance. (F) Scan paths of true negative trials in the foraging task are overlaid on the pseudo-search display. Under this simulation, the landing probability (i.e., plot E) is applied to calculate pseudo-hit rates, which are

represented by colored solid lines. The ribbons show the standard deviation. Black horizontal lines are drawn on the actual detection rates in the search task. Black vertical lines indicate estimated attentional FVF radii that are needed to achieve the actual hit rates.

* Adjusted $p < .05$, ** Adjusted $p < .01$

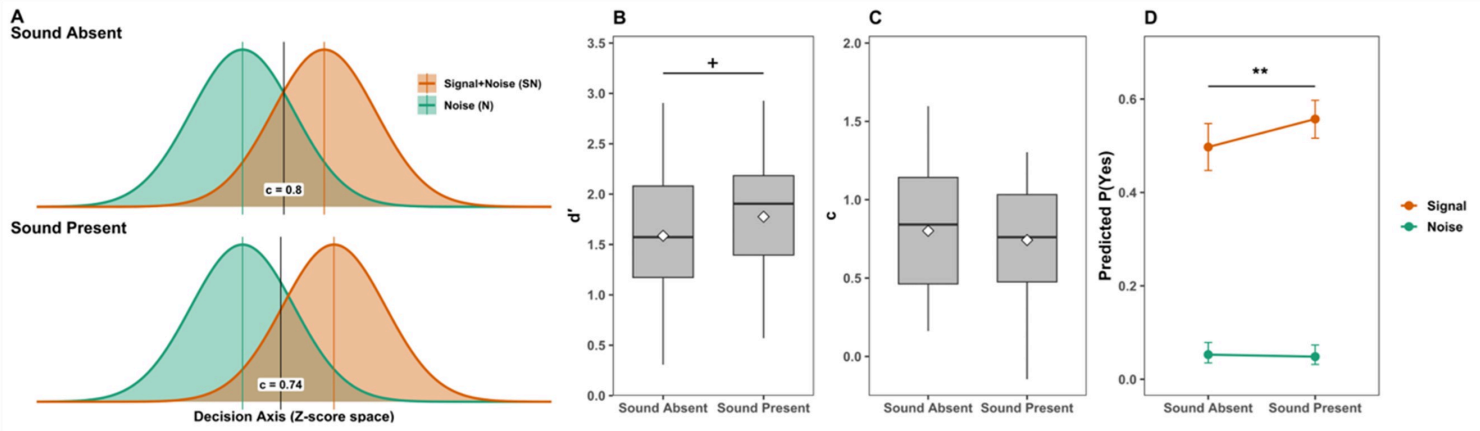


Figure 4

Results of Experiment 2.

Note. (A) Colored vertical lines show the peaks of signal detection distributions. Noise distribution means target-absent trials, while Signal + Noise distribution means target-present trials in the search task. (B) (C) Discriminability and criterion are illustrated with the mean values as diamonds. (D) The index converted from the logit to the probability scale in the GLMM analysis is plotted.

+ $p < .10$, ** $p < .01$