RESEARCH REPORT

Neural responses to natural visual motion are spatially selective across the visual field, with selectivity differing across brain areas and task

Jason J. Ki[1] | Jacek P. Dmochowski[1] | Jonathan Touryan[2] |
Lucas C. Parra[1]

[1]Department of Biomedical Engineering, City College of the City University of New York, New York, New York, USA

[2]U.S. Army Research Laboratory, Aberdeen, Maryland, USA

Correspondence
Lucas C. Parra, Department of Biomedical Engineering, City College of the City University of New York, NY, USA.
Email: parra@ccny.cuny.edu

Abstract

It is well established that neural responses to visual stimuli are enhanced at select locations in the visual field. Although spatial selectivity and the effects of spatial attention are well understood for discrete tasks (e.g. visual cueing), little is known for naturalistic experience that involves continuous dynamic visual stimuli (e.g. driving). Here, we assess the strength of neural responses across the visual space during a kart-race game. Given the varying relevance of visual location in this task, we hypothesized that the strength of neural responses to movement will vary across the visual field, and it would differ between active play and passive viewing. To test this, we measure the correlation strength of scalp-evoked potentials with optical flow magnitude at individual locations on the screen. We find that neural responses are strongly correlated at task-relevant locations in visual space, extending beyond the focus of overt attention. Although the driver's gaze is directed upon the heading direction at the centre of the screen, neural responses were robust at the peripheral areas (e.g. roads and surrounding buildings). Importantly, neural responses to visual movement are broadly distributed across the scalp, with visual spatial selectivity differing across electrode locations. Moreover, during active gameplay, neural responses are enhanced at select locations in the visual space. Conventionally, spatial selectivity of neural response has been interpreted as an attentional gain mechanism. In the present study, the data suggest that different brain areas focus attention on different portions of the visual field that are task-relevant, beyond the focus of overt attention.

KEYWORDS
evoked neural activity, natural dynamic stimuli, spatial selective attention, stimulus–response correlation

1 | INTRODUCTION

Traditional studies of visual perception employ tightly controlled experimental paradigms. For instance, classic behavioural studies on visual attention present discrete stimuli and attention cues at select areas in the visual field and measure accuracy or response times as a function of location and cues (Posner, 1980). These studies have established a clear difference between overt attention, defined as the location of observable gaze position, and covert attention, which manifests as a performance gain when subjects are given a cue directing their attention to a location different from their gaze position (Moran & Desimone, 1985; Spitzer et al., 1988). Studies on the effects of covert attention on neural response often present discrete stimuli at specific locations in the visual field and evaluate the effect of attentional cues on neural activity (McAdams & Maunsell, 1999; Motter, 1993). Using intracranial recordings, animal and human studies have established that neuronal firing to visual stimuli is selectively enhanced for attended locations in the visual field (Luck et al., 1997; Moore, 1999; Self et al., 2016). Similarly, location-dependent neuronal gains have been found for attended locations with scalp recordings in humans. For instance, discrete visual stimuli produce robust contralateral responses when covertly attending to a selected visual hemisphere (Hillyard & Anllo-Vento, 1998; Luck et al., 1990; Mangun, 1995). Moreover, similar results have been shown in neuroimaging studies with the enhanced contralateral haemodynamic response over the visual cortex during spatial attention tasks (Beauchamp et al., 2001; Mangun et al., 1998; Tootell et al., 1998).

Although a great deal has been learned from these studies, real-world tasks such as driving differ considerably from these traditional experimental paradigms. In traditional studies of covert attention, subjects are often asked to inhibit saccades (e.g. Luck et al., 1990; Mangun, 1995; Moran & Desimone, 1985; Rugg et al., 1987; Spitzer et al., 1988). Yet, existing evidence from evoked potentials suggests that neural processing is altered when saccades are artificially constrained (Ki et al., 2016; Kulke et al., 2016). More recent studies have explored overt attention during free viewing of static natural images (Eckstein et al., 2006; Võ & Henderson, 2010; Võ & Wolfe, 2015) or film (Dorr et al., 2010; Itti & Baldi, 2009). Most recently, eye movements have been explored also in real-world conditions involving head movements and walking (David et al., 2021; Foulsham et al., 2011; Vo & Henderson, 2009). A few studies have also explored the effects of overall attention on neural activity for naturalistic dynamic visual stimuli such as movies, video games and driving (Bavelier et al., 2012;

Dmochowski et al., 2012; Hasson et al., 2008; Ki et al., 2020, 2016; Wang et al., 2015). However, little is known about the spatial aspects of covert visual attention for naturalistic dynamic stimuli and tasks.

In recent years, data-driven methods have developed to quantify the neural activity generated in response to complex naturalistic stimuli. These studies employ multivariate modeling to map low-level stimulus properties to observed neural response (encoding) or vice versa (decoding) (Holdgraf et al., 2017; Naselaris et al., 2011). This stimulus–response modeling is also referred to as system identification (Gallant et al., 2012; Wu et al., 2006) and is particularly advantageous for complex dynamic visual stimuli, which cannot be readily decomposed into discrete events required for conventional event-related analysis (Crosse et al., 2016). One area of research that has popularized the stimulus–response modeling approach is studies on visual representation using fMRI. These studies capture spatial and orientation features of naturalistic images and dynamics movies to predict haemodynamic response at individual voxel regions (Kay et al., 2008; Naselaris et al., 2009; Nishimoto et al., 2011), and the mapping can be reversed to reconstruct the original stimulus with reasonable accuracy (Miyawaki et al., 2008; Naselaris et al., 2009; Nishimoto et al., 2011). A basic finding is that different region of the visual cortex exhibit spatial selectivity to visual input at different locations across the visual field, with selective attention increasing the reliability of these responses (Kay & Yeatman, 2017).

In electrophysiology, similar encoding models have been applied to map temporal visual contrast to neural responses (Gonçalves et al., 2014; Lalor et al., 2006; Lalor & Foxe, 2009). This approach yields a precise estimation of the evoked response dynamics, which is analogous to event-related potentials; however, it does not account for the spatial heterogeneity of visual dynamics. In a similar approach, the strength of evoked activity across multiple electrodes can be measured as the correlation of stimulus and response (Dmochowski et al., 2018). Using this approach, we previously investigated the effects of active perception by comparing the strength of neural response during active play and passive viewing of a video game (Ki et al.,). We found that active engagement in the task enhanced responses to the overall visual contrast dynamic. However, this analysis was conducted on coarsely averaged visual features that represented the dynamics of the entire visual field with a one-dimensional time series.

Based on the spatial selectivity and attention effects observed in EEG and fMRI, we hypothesized that during a natural visual task, the strength of neural responses differs across the visual field and that the task selectively

modulates neural activity. To test this, we reanalyse the EEG recordings from Ki et al. (2020) in which subjects either actively played a kart racing video game or passively viewed a pre-recorded video of a race. We extended the system identification approach by measuring stimulus–response correlation (SRC) *resolved in visual space*. In doing so, we developed a novel analytic approach that recovers the neural response evoked in the brain by the various features embedded in a visual stimulus. Importantly, the recovered responses are themselves varying across cortex, effectively permitting the identification of the mapping from visual to brain space. Our findings show that the neural response to movement is enhanced in task-relevant areas of the visual scene. Remarkably, this spatial selectivity differs across brain areas. Moreover, we find that active play enhances neural responses in areas extending beyond the focus of overt gaze position. Overall, these results suggest that neural response during a naturalistic dynamic visual experience differs across the visual field depending on task demands.

The goal of this exploratory work is to shed new insights about the spatial selectivity of neural response during a naturalistic visual task. This work shows the heterogeneity of stimulus-evoked neural response across the visual space during a continuous stream visual experience. Although we do not explicitly manipulate attention, the kart video game shows that active interaction with the visual input increases the strength of neural responses in visual space in comparison to passive viewing.

2 | RESULTS

Human subjects ($N = 42$) were asked to play a kart race video game (Figure 1), whereas we recorded continuous EEG activity and, in some subjects, also gaze position ($N = 17$). The players control velocity and left/right heading direction with their right hand. Their goal is to complete the course as fast as possible (each race takes approximately 3 min to complete). Good performance implies driving fast without veering off the road or crashing into buildings and competing race karts, which incurs a large time penalty. The visual dynamic is dominated by the translation of the moving vehicle, which produces an outward radial pattern of optical flow (Figure 1a), computed here for each screen location (Horn & Schunck, 1981). The magnitude of optical flow across all races is dominant in the periphery and weakest at the centre (Figure 1b). In contrast, the task-relevant information is on the road and adjacent buildings (see video clip, eye position markers: cyan circle, active play; green circle, passive viewing). Interestingly, judging by the

distribution of gaze positions (Figure 1c), overt attention seems to be narrowly focused on the road ahead. During the kart race, the gaze distribution is similar to that of real-world vehicle control in which the driver's eye gaze is particularly focused on the inner 'tangent point' of the road curvature [Land & Lee, 1994]. The rightward skew of the eye gaze distribution (Figure 1c) indicates that gaze is directed at this curvature point.

2.1 | Spatially resolved SRC

Given that the player's kart is centred on the screen (Figure 1a, dotted circle), absolute locations on the screen take on distinct importance in this driving task. We hypothesized that the strength of the neural response to visual movement is enhanced at select areas on the screen that are task relevant. To test this, we assess the strength of visually evoked activity resolved in visual space. Specifically, we measure SRC between the raw EEG signal and optical flow magnitude at individual screen locations (Figure 2). We will use two different approaches to measure SRC. In the first approach, the multiple EEG electrodes are projected onto a component space (Dmochowski et al., 2018) similar to the principal or independent components—these can be conceived as 'virtual electrodes'—and the stimulus is filtered to predict the neural activity observed in these virtual electrodes (Figure 3). In the second approach, the stimulus is filtered to directly predict neural activity in each of the original EEG electrodes (Lalor et al., 2006) (Figure 4). In both instances, the temporal filters, that is, impulse responses or temporal response functions (TRF), are estimated for each screen location separately, and SRC is the correlation of the predicted neural activity with the actual neural activity. Given that optical flow for this video game is sufficiently distinct at each of the screen locations (Figure 2a), one might expect quite different SRC for each location (Figure 2c).

2.2 | The strength of the visually evoked response varies across visual space

We analysed the strength of SRC for different locations across the screen (23×40 patches) for all subjects and races combined (each subject performed two races in active play or passive viewing conditions, yielding a total of $n = 170$ races, 179.91 s $+/- 17.14$ s duration each). Temporal correlations are measured at the 30-Hz frame rate of the video (EEG is down-sampled to this rate).

To quantify the overall SRC, we first used components of the EEG that were optimized to capture the

(a)

(b) optical flow variance   **(c)** eye gaze
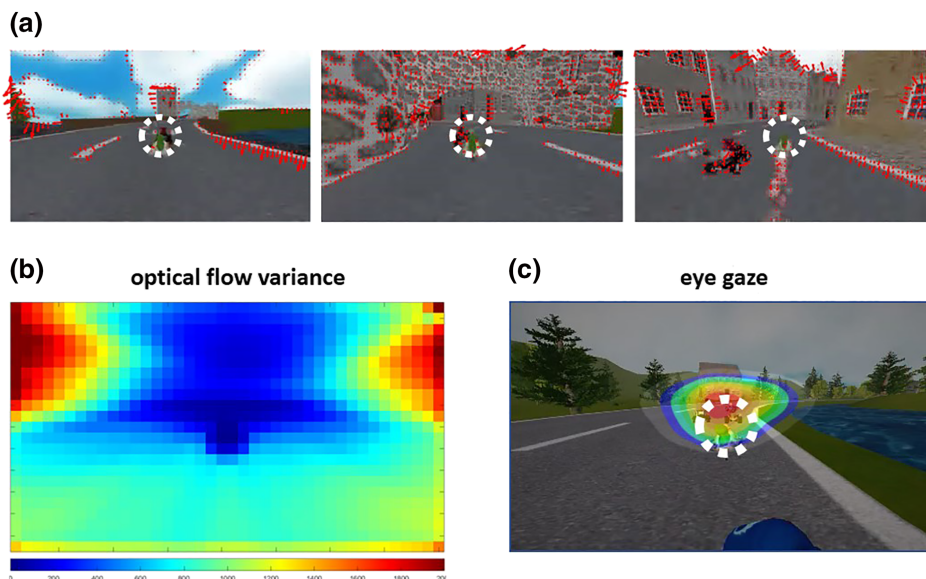
FIGURE 1 Visual dynamics of the race kart video game. (a) the kart's translation creates the overall visual dynamics of the racecourse, which forms optical flow that expands radially outward. red arrows indicated velocity vectors estimated for each location on the screen. We analyse the magnitude of velocity vectors at each location. (b) Variance of optic flow magnitude averaged across all runs. Flow is slow in the Centre and faster at the edge of the screen. (c) Heat map represents distribution of eye gaze position across all runs and subjects. It is superimposed here on a sample image. The probability is high (red) at the Centre of the screen near the road's curvature, extending rightward towards the general turning direction (the race goes clockwise around the track). Dotted white circle is shown here only for reference (players do not see this) and centred in front of the kart

strongest correlation between global optic flow (mean amplitude over the entire screen) and EEG responses following (Ki et al.,). Specifically, $n = 96$ EEG electrodes are combined linearly, generating an EEG component signal that can be thought of as virtual electrodes (see Section 4). The corresponding distribution of these components (Figure 3a) is known as 'forward model' (Haufe et al., 2014). TRFs to predict the EEG component activity are then estimated for each patch on the screen separately (Figure S1). The dependence of these TRFs on visual field and scalp locations will be discussed later (in Figure 6). SRC is the correlation between the observed EEG, in each component, with the activity predicted from optic flow in each screen location (Figure 3b). We consider the top three components as they capture most of the SRC (notice in the colour map that SRC diminishes for the later components). Stronger SRC indicates stronger neural response to local optic flow. For Component 1, which captures the strongest neural response, SRC is highest on the road. The optic flow in this area is dominated by road demarcation as well as the movement of competing cars and other objects. In Component 2, the neural response is strongest for movement on the horizon where buildings appear as well as the road. In Component 3 (right-lateralized scalp topography), the neural response focuses on the horizon, primarily to the left visual hemisphere. These unique spatial patterns suggest

that the strength of visual evoked responses varies with location on the visual field, but also, this spatial preference differs for different brain areas.

2.3 | Visual-spatial dependence of neural responses varies across EEG electrodes

To determine specifically how the visuospatial selectivity of SRC differs across brain areas, in the second approach, we used TRF to map the stimulus to the neural response directly for each scalp electrode. The TRF estimation follows an established system identification approach (Crosse et al., 2016; Lalor et al., 2006), except that we again resolve SRC in visual space by computing unique TRF from the optic flow at individual screen locations (Figure 4, showing a subset of 20 of the 96 electrodes). For midline-frontal channels, SRC is strongest at the top centre of the screen, yet lateral-frontal electrodes respond strongest to movement on the road. We hesitate providing an interpretation for this finding given the proximity of these electrodes to the eyes, despite having removed eye movement artefacts with regression of more lateral and frontal electrodes (Figure S5). At central electrodes, over cortical somatosensory and motor areas, we observe responses lateralized to the contralateral visual field (which is dominated by the movement of buildings, trees
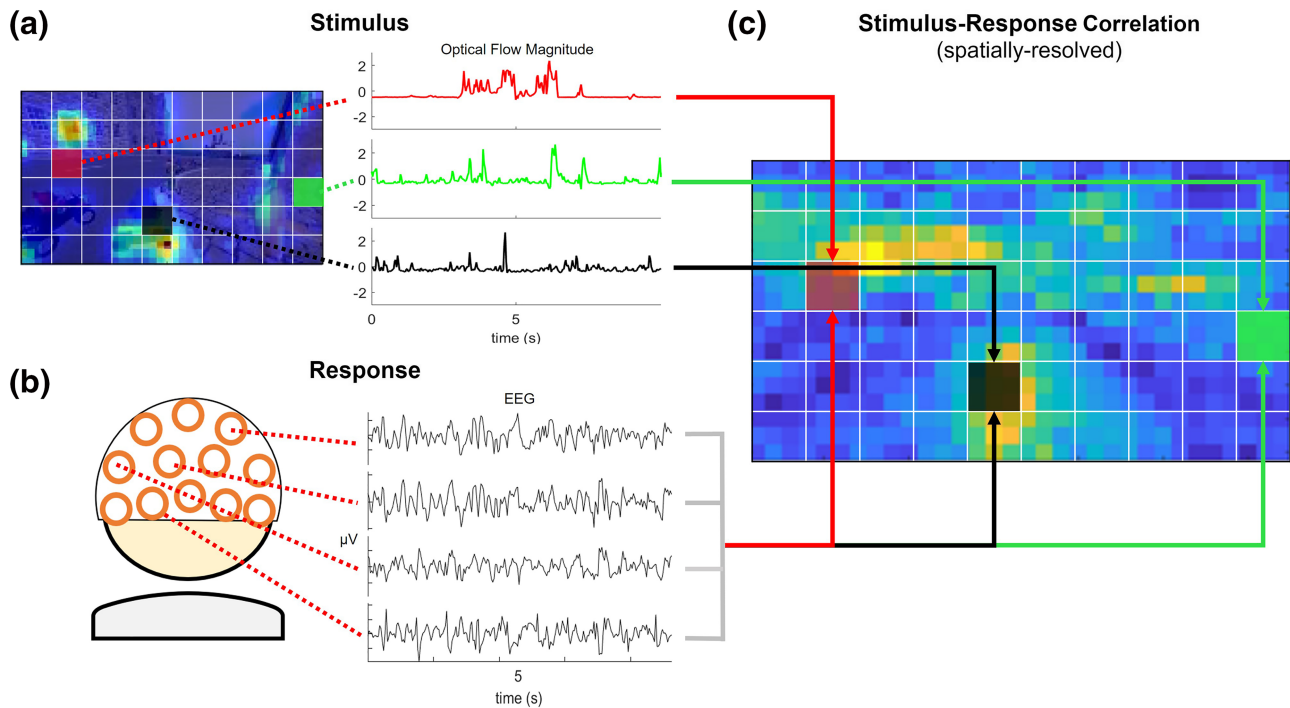
FIGURE 2 Resolving the strength of neural response across visual space. Here, we illustrate the spatially resolved stimulus–response correlation (SRC) approach, which measures the correlation of continuous stream neural response to a temporally filtered stimulus. (a) we record the screen capture of the video game and extract the optical flow at each location on the screen (Horn & Schunck, 1981). The image shows a snapshot of optic flow magnitude (z-scored per patch). Optical flow is averaged over small image patches (7.8 × 8.5 pixels). The coloured patches at different screen locations (red, green, black) show unique optical flow magnitude over time. (b) Neural responses are captured with the scalp EEG during the video game presentation. The independent variable is the SRC, which is computed here in two distinct ways, depending on whether neural activity is confined to a single channel or spatially filtered across multiple electrodes. In both cases, the dynamics of the optical flow of the visual stimulus is correlated with the evoked response. (c) the location dependence of the SRC is assessed at rectangular patches of the screen (i.e. 23 × 40 patches, averaged from a resolution of 180 × 320 pixels)
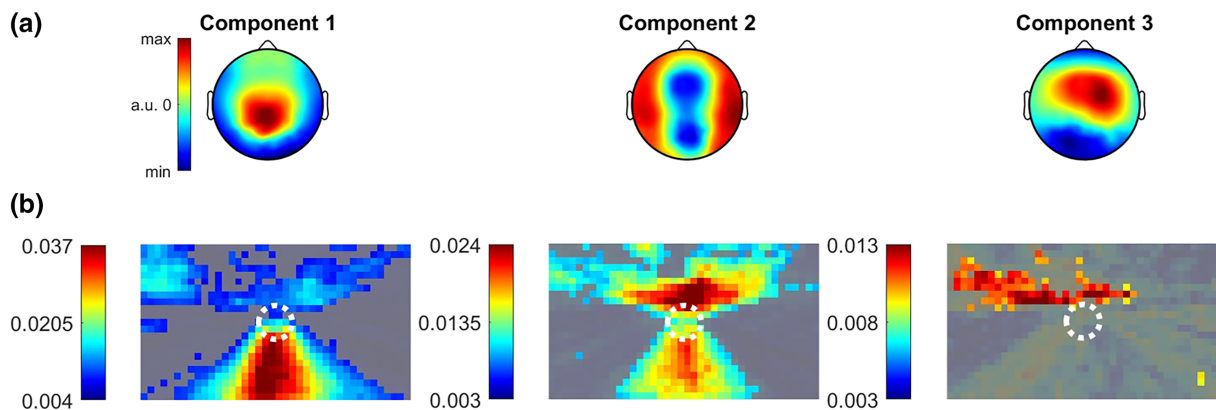
FIGURE 3 The strength of the visually evoked response is location-dependent. (a) for the spatially resolved analysis, we employ the spatial and temporal filters previously computed in Ki et al. (2020) using canonical correlation analysis (CCA). CCA yields a set of components that maximize the correlation between spatially filtered EEG (96-channel) and temporally filtered stimulus. Components of the EEG are constructed by linearly combining the recorded signal across electrodes, with the associated stimulus filters depicting the TRFs (shown in Figure S1). (b) We measure the correlation between spatially filtered EEG and temporally filtered optical flow at individual screen locations (23 × 40 patches, i.e. 920 patches) for the top 3 CCA components. The SRC is equivalent to correlation Pearson coefficients, ranging between −1 and +1. Here, we create a visualization of spatially resolved SRC using a colour map. The colour value at each pixel maps to SRC based on a colour scale. The areas with grey show patch regions in which SRC is below chance correlation ($p > 0.05$, uncorrected)
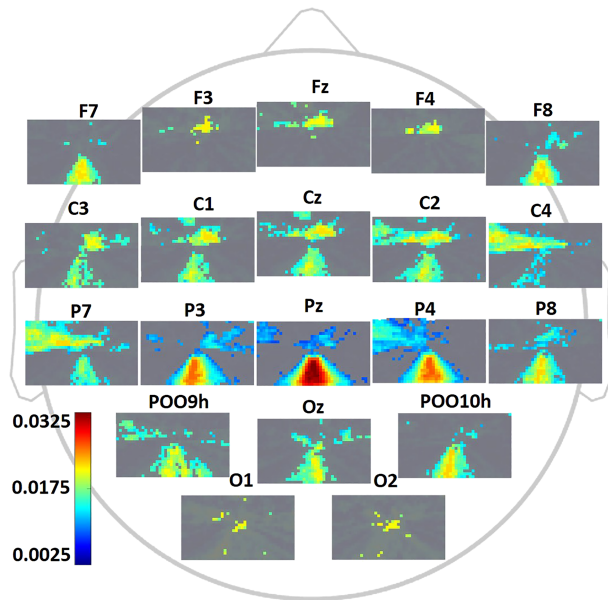
FIGURE 4 Location-dependent neural responses differ across scalp positions. Stimulus-response correlation computed from optic flow at each screen location with each electrode. SRC, represented as false colour, indicates the correlation between actual EEG and EEG predicted by filtering optic flow magnitude with a temporal response function. The areas with grey show patch regions in which SRC is below chance correlation ($p > 0.05$, uncorrected)

and hillsides that are to be avoided). To test for the statistical significance of this lateralization, we performed a chi-squared test that compares the proportions of patches with significant SRC between the left and right hemifields. For C4, which showed left contralateral dominance of SRC, we find indeed a greater portion of significant patches at the left side of the screen ($\rho = 10^{-23}, \chi^2 = 166.88$). In contrast, for C3, which showed right-side dominance, we find a greater portion of patches that are statistically significant at the right side of the screen ($\rho = 0.00018, \chi^2 = 14.68$). Note that responses are stronger in the left visual hemifield, possibly due to the right-hand key presses to control the kart. However, the effect is also there during passive play (Figure S2), making this a less likely explanation. At parietal electrodes, we observe strong responses to visual dynamics on the road. This parallels the finding for Component 1 (Figure 3), which had a dominant parietal positivity. Lastly, in the occipital electrodes, we observe a strong response to visual dynamics associated with movements on the road, likely associated with competing karts in the vicinity of the player's vehicle or objects on the road. Overall, there is a predominant bias to the left hemifield, which is also expressed in Component 3. Overall, these results show that spatial selectivity of neural response is different across different brain areas.

2.4 | Neural response is enhanced at task-relevant locations during active play

The manual control of the race kart involves active engagement, requiring selective visual-spatial attention to task-relevant areas. Given the modulation of neuronal gain observed with visual attention (Hillyard et al., 1998; Luck et al., 1997; Maunsell & Cook, 2002; Motter, 1993), we hypothesize that the EEG response will be selectively enhanced during active play for locations that are relevant to the task. To test this, we compared the strength of neural responses to optic flow between active play and passive viewing. In the passive viewing condition, subjects were asked to watch pre-recorded races of the game. The order of active play ($n = 84$ trials; a trial is three laps around the racecourse) and passive viewing ($n = 86$ trials) was counterbalanced across subjects ($N = 42$). We compute the location-dependent SRC as in Figure 3. To obtain an assessment of the overall EEG response, we summed SRC over the Top 3 components (Figure 5). In Figure S2, we resolve this by individual electrodes. The mean across subjects has distinct spatial distribution across the scene for active and passive play (Figure 5a). During active play, neural responses are more robust for movement occurring on the road and beyond the road where buildings and trees could result in collisions. In contrast, during passive viewing, the neural response is strongest only for areas covering the road. Indeed, we find that neural response is significantly enhanced with active play for areas that cover buildings and objects coming ahead on the horizon above the road (Figure 5b highlights locations with $p > 0.05$, Wilcoxon rank-sum, one-tailed, FDR corrected across 920 patches, $N = 42$ subjects). This enhancement extends beyond the focus of overt attention as characterized by gaze position (Figure 5C). Notably, the contrast seems to be driven mostly by central and frontal scalp locations (Figure S2). Importantly, we did not find differences in gaze position between active play and passive viewing (Figure 5d highlights locations with $p > 0.05$, Wilcoxon rank-sum, one-tailed, FDR corrected across patches; eye-tracking data were available in $N = 17$ subjects). The distribution of eye velocity appears to be slightly skewed toward larger saccades in active play (Figure 5e). However, pairwise comparison of the average velocity of individual subjects between active and passive trials showed no significant difference (Figure 5f, $z = 0.355$, $p = 0.72$, df = 16, paired two-tailed Wilcoxon signed-rank test, that is, only including subjects for whom both conditions were available), indicating that eye movement dynamics are comparable in the two conditions. Overall, these results suggest that the neural response is enhanced at task-relevant locations, extending beyond the focus of overt attention, and
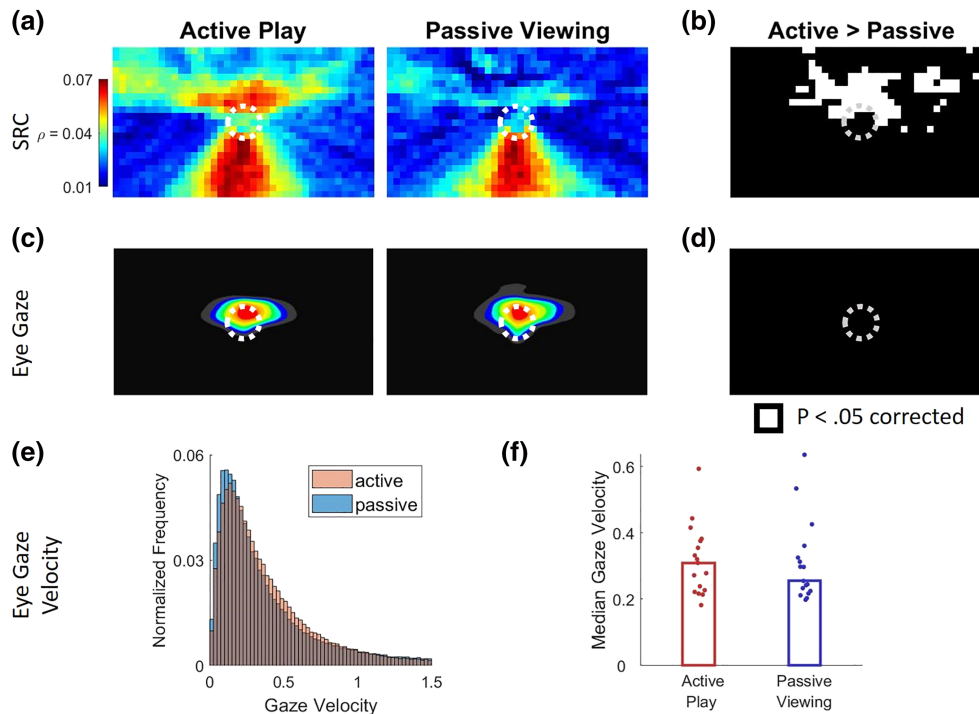
FIGURE 5 Neural response to optic flow is enhanced at task-relevant locations during active play, whereas overt gaze position and saccade velocity are unchanged. (a) Spatially resolved SRC (as in Figure 3) summed over three components but separate for active play and passive viewing conditions. (b) Locations of significant differences in SRC between active and passive conditions ($p > 0.05$, FDR corrected). Enhancement occurs are task-relevant locations. (c) Distribution of gaze position over time-averaged over subjects. The gaze is mostly focused on the heading direction above the stationary position of the subject's own kart (white dotted circle). (d) the contrast in the distribution of gaze position between active play and passive viewing shows no statistical differences (for all locations $p > 0.05$, FDR corrected). (e) Distribution of gaze velocity suggests a slight increase in gaze dynamics during active play. (f) Median velocity of individual subjects, however, shows no significant difference between conditions (each point is a subject)

differences cannot be attributed to changes in fixation position of saccade size.

2.5 | Neural response occurs faster at the periphery

To determine if the latency of neural responses differs across the visual space, we analyse the TRFs that were extracted separately for each visual location and electrode (for Figure 4). Analogous to ERPs, TRF yields an estimate of evoked response time to a specific visual input. Here, we focused on the parietal Pz electrode and displayed the distinct TRFs obtained by regressing optical flow at each two-dimensional (2D) screen location (Figure 6a). In combination, the collections of TRFs have two visual space dimensions and one time dimension indicating the latency of neural response to optical flow at each screen location (Figure 6b). We find a dependence of latency on location in visual space, which is most obvious when displaying the TRFs in time along a vertical direction (Figure 6d, indicated by the white dashed arrows). Neural

response first peaks at 130 ms for movement at the bottom of the screen, with latency increasing towards the centre of the screen, where it reaches a latency of 300 ms (Figure 6f, arrow). A similar pattern is observed for electrodes Cz and Oz (Figure S3) as well Component 1 (Figure S1). This delay is primarily driven by the movement on the road. In the horizontal direction extending beyond the road, we also see a similar phenomenon albeit with different latency and less pronounced (Figure 6f). Notably, the delay does not seem to depend on whether the subject is actively engaged in the game (Figure S4). In summary, in both horizontal and vertical cross sections, the peak of neural response occurs first at the lateral positions of the screen and finishes last at the centre, and this effect is not dependent on task engagement.

Interestingly, these centripetal time delays closely match the centrifugal movement of objects on the screen (Figure S6). Specifically, the slope of the white arrow in Figure 6d captures the processing delay of 75 patches/ seconds (p/s). The corresponding velocity of objects on the road in this portion of the screen is 76 p/s, that is,

(a)

(b) **(c)** **(d)** **(e)** **(f)** **(g)**
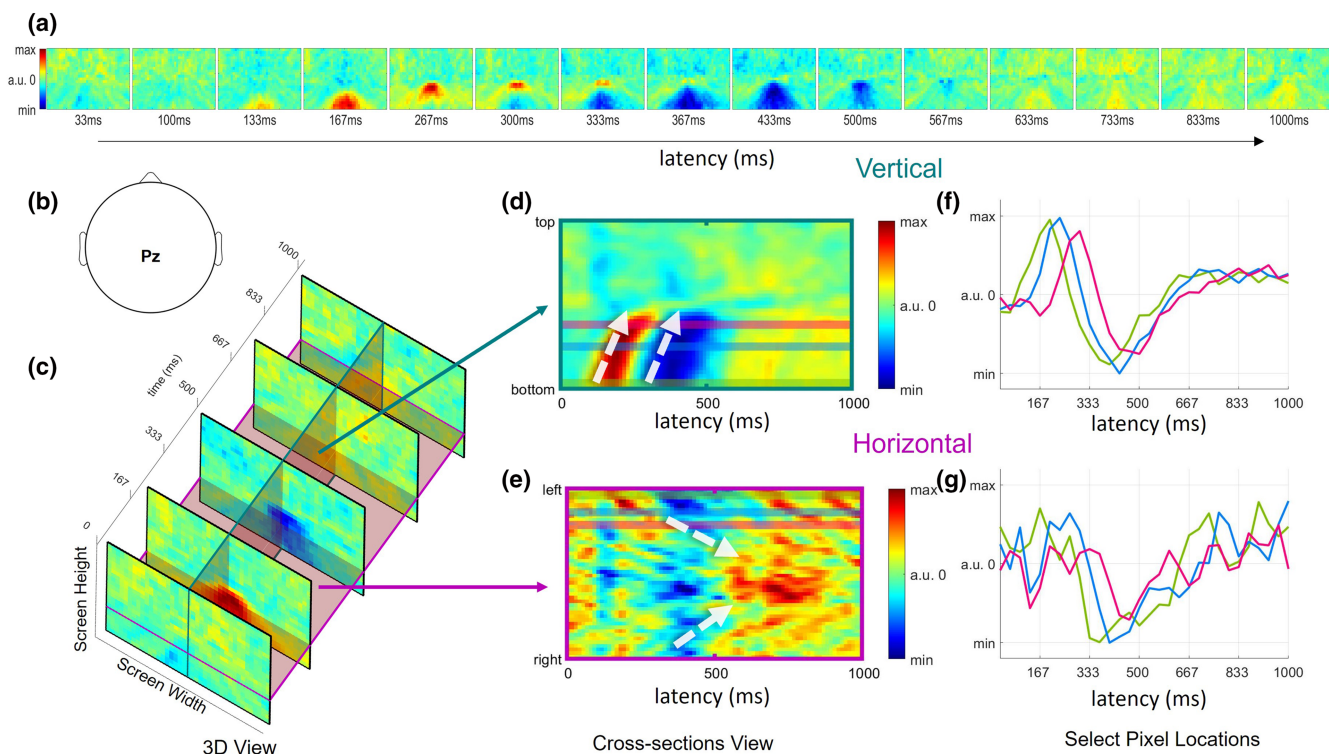
FIGURE 6 Neural response to peripheral movement is faster. (a) Spatially resolved temporal response function (TRF) at select time points. The TRFs are computed by regressing optical flow at each screen location with Pz electrode. Each panel represents the TRF a given latency indicated as number below. (b) Neural response measured at the Pz location. (c) 3D visualization of spatially resolved TRFs (screen width × screen height × latency). The slice along the latency dimension indicates cross sections of temporal response at vertical (dark green) and horizontal (maroon) positions on the screen. (d) TRF along vertical screen direction. White dashed arrows indicate increasing delay towards the Centre of the screen (opposite to the direction of optic flow, which moves from the centre outwards) (e) TRF in horizontal screen direction. (F) Time course at select pixel positions (green, blue, pink), which corresponds with coloured lines along the vertical cross section in (d). (g) Same as (f) but along the coloured lines at the horizontal cross section

outward vertical direction (Figure S6b). Similarly, the slope of the arrows in Figure 6e are 30 and 40 p/s whereas objects move on the left and right in horizontal direction at a speed of 33 and 36 p/s (Figure S6c,d). These slope estimates are only approximate. Nonetheless, they suggest that the TRF compensates for the delays of objects movement on the screen. This centripedal TRF finding holds for the Pz and Oz electrodes but is less pronounced for Cz and disappears for frontal electrodes (Figure S3).

3 | DISCUSSION

In this study, we assessed the strength of neural responses elicited by motion across the visual space during a dynamic video game. The task involved racing a kart while avoiding obstacles, as collisions significantly impact race time, which was the primary objective of the game. This task requires selective attention to specific objects or locations on the screen in order to plan and

control the kart. Because the kart location was maintained at the centre of the screen, and obstacles appeared in similar locations on the screen, we can draw inferences on the task relevance of neural responses to movement occurring at each location of the screen. The need for spatial consistency is perhaps a caveat of the approach; nonetheless, it allows us to derive a visual-spatial representation for EEG, which has traditionally been constrained to the temporal domain.

Similar to real-world driving (Land & Lee, 1994), we found that subjects' gaze was fixated on the curvature of the road while simultaneously shifting the gaze to competing karts and objects. The translation of the kart along the road produces a uniform optical flow that radiates outwardly, whereas other competing karts produce sparse optical flow along the road. Due to this self-motion, the optic flow was strongest in the lateral periphery (Warren & Hannon, 1988). Yet, we find that neural responses were strongest for movement in the middle of the screen, including the road and obstructions coming up on the horizon. Therefore, neural responses are

strongest for task-relevant locations. Importantly, this spatially selective neural response extended beyond the focus of gaze position and thus beyond the focus of overt visual attention.

Perhaps the most salient finding of this work is that optic flow elicited responses across the entire scalp with spatial selectivity differing across scalp locations. For example, parietal electrodes show strong responses for the road, which is dominated by the movement of traffic lines demarcating the road and competing karts. It is worth noting that visuomotor transformations, as may be required in this task, are known to occur in the superior parietal cortex, as demonstrated in non-human primates (Fogassi & Luppino, 2005) as well as parietal components of human EEG recordings (Naranjo et al., 2007). The parietal activity responded earlier to movement in the periphery and later to movement at the centre of the screen. This is intriguing as it is in the opposite direction of objects moving from the centre to the periphery. The effect was most pronounced along the road. Earlier studies on the latency of evoked potentials as a function of visual eccentricity in static visual stimuli give mixed results with latency depending on upper versus lower and contralateral versus ipsilateral hemifield (Busch et al., 2004; Capilla et al., 2016).

During natural vision, different regions are selective to local motion and global flow (Bartels et al., 2008). Given that the perception of self-motion is mainly determined by the magnitude of visual flow in the periphery (Warren et al., 1988), it is possible that peripheral visual motion is processed faster in the brain. Processing of optic flow is ascribed to the dorsal visual stream, particularly in the medial superior temporal (MST) areas (Duffy, 1998; Tanaka & Saito, 1989), which may have contributed to the strong parietal response. An alternative interpretation is that visual content in the centre of the screen had more complex semantics and therefore resulted in longer processing times, consistent with the relatively long response onset time of 200–300 ms (VanRullen & Thorpe, 2001). An important observation of these varying latencies in the TRFs of parietal and occipital electrodes is that they appear to compensate for the outward movement of objects on the screen. Thus, the 'slower' response for movement in the centre may reflect not necessarily slower neural processing, but rather a grouping of the low-level movement features into a coherent object. In short, parietal and occipital neural responses are time aligned to objects, independent of their centrifugal movement on the screen.

Neural responses were significantly enhanced during active play of the game compared with passive viewing of the same visual stimuli. This enhancement was spatially selective, overlapping with the focus of overt attention at the centre of the screen, but also extending to task-relevant locations. It is well established that visual evoked potentials are enhanced when attending to a specific target or locations, even in the absence of overt orienting of eye gaze (Hillyard & Anllo-Vento, 1998; Luck et al., 2000; Luck & Kappenman, 2011; Mangun & Hillyard, 1988). Here similarly, neural responses were enhanced without a change in gaze position, suggesting that this may be the result of task-relevant increase of spatially selective attention. The finding that the enhancement of evoked responses during active viewing manifested in a task-related region of visual space suggests that the enhancement reflects a circuit-level attentional effect. We speculate that the active nature of the task led to the upregulation of pathways specific to processing of the visual stimulus. On the other hand, it seems less likely that changes due to increased concentration, arousal, or effort, which likely manifest as global neuromodulatory effects, would strengthen the evoked response to only the task-related region of visual space. However, a caveat of this study is that we did not explicitly manipulate attention; therefore, the interpretation of the results as selective spatial attention is only suggestive. In particular, the study does not determine whether the enhancement may have resulted from transient shifts of attention, or an overall broadening of the spotlight, or simply an extended duration of attention to task-relevant areas.

We note that this selective spatial enhancement for active play was strongest for central electrodes over the motor cortex, which responded more strongly to visual movement on the horizon in the contralateral visual hemisphere (where upcoming buildings appear alongside the road) for both left- and right-side electrodes. This left/right symmetry cannot be ascribed to the neural activity associated with somatosensory or hand motor control, as driving the kart involved a single hand (the right hand). It is well established that attending to one hemisphere of the visual field elicits stronger visual evoked potentials at the contralateral occipital cortex (Luck et al., 1990; Mangun, 1995). The visual-spatial selectivity of central scalp electrodes suggests that central brain areas 'attend' to visual movement on the contralateral visual field.

Another possibility is that the central components reflect a visuomotor integration required for navigating the kart. Note in this context that there was a dominance of responses to the left visual hemifield, notably stronger for the right hemisphere of the brain. If this asymmetry were the result of somatosensory or motor evoked responses due to the actual right-hand key presses, one would have expected stronger asymmetry for the left (contralateral) brain hemisphere. Note that visual-spatial

attention can modulate evoked potentials at central electrodes when the subject is engaged in a motor task (Praamstra & Oostenveld, 2003). In primates, neurons in the motor cortex are found to respond to the attended visual location (Lebedev & Wise, 2001). However, no topographic correspondence with the visual field is known for the motor cortex. Although selective visual attention may provide the groundwork necessary for the relationship between visual stimulation and central brain activity to emerge, one should be cautious with a conclusion that motor areas are 'attending' to visual movement on the contralateral hemifield.

An alternative interpretation of the spatial selectivity of central electrodes is that participants are moving their body during turns, or perhaps just 'feeling' the acceleration of a turn without overt movement. Additionally, simply observing an action may elicit covert evoked responses over central electrodes (Holländer et al., 2011). It is possible that this motor or sensory activity could cause a correlation between optic flow and neural activity on the contralateral side. But we consider this to be unlikely, as right and leftward turns on the road are reflected in the optic flow of the entire visual scene. Regardless of the interpretation, the finding that central electrodes respond to optic flow on the contralateral visual field is noteworthy.

We found spatial selectivity for occipital electrodes Oz, but not O1 and O2, which might have been expected if the effect is due to visual-spatial attention. One possibility is that the symmetry of centrifugal motion between the upper and lower hemifield cancelled the evoked responses in these electrodes due to opposite polarity of the cruciform structure of early visual cortex (Kelly et al., 2013). Perhaps weaker effects were not resolved when O1 and O2 were analysed individually. Indeed, the component extraction method (Figure 3) does appear to have contribution from all occipital electrodes. Finally, we found some spatial selectivity in frontal electrodes. However, we hesitate providing an interpretation for this finding given the proximity of these electrodes to the eyes despite efforts to remove eye movement artefacts.

The overall enhancement observed with active play relative to passive viewing may result from selective attention, but other non-spatial constructs may be equally relevant such as focus, arousal or game engagement. A possible confound of active play is that button responses elicit motor and somatosensory evoked responses as well as changes in optic flow. These may therefore correlate, which in turn results in an increase of SRC above the values observed in passive viewing. However, such enhancement should be strongest in the periphery of the screen, which is more strongly affected by changes in heading direction. Yet, we find the

enhancement with active play to be strongest in the centre of the visual field.

An alternative interpretation is that enhancement with active play results from bottom-up stimulus-driven reorienting, consistent with the observation that visually salient events shift eye gaze and activate dorsal ventral regions (Nardo et al., 2011). Indeed, a caveat of the present study is that during free viewing, eye position and eye movements are dynamic and uncontrolled. Traditionally, a distinction is made between overt and covert attention in experiments that aim to strictly control eye movements (Kelly et al., 2010; Kulke et al., 2016; Posner et al., 1984; Rugg et al., 1987). Eye movements such as saccades, microsaccades, pursuit and stigmas are distinguished as they contribute uniquely to attentional-related ERPs (Kelly et al., 2010; Nobre et al., 2000). Although we found that the overall eye movement velocity was quite similar between active and passive conditions, we did not analyse the effects of different types of eye movements in this investigation. It is possible that changes in eye movement contributed to enhanced neural responses. However, it has been shown that the magnitude of responses to visual cues remains similar regardless of eye movement condition (Parker et al., 2020). Moreover, covert and overt attention shifts elicit similar early evoked responses (Hanning et al., 2019). Taken together, we believe that the effects of differing eye movement dynamics played only a minor role, although we cannot fully rule this out.

Methodologically, this study combines new and old approaches to studying neural responses to visual stimuli. Visual-spatial attention and feature-based attention have been extensively studied in the last 50 years using traditional event-related paradigms (Hillyard & Anllo-Vento, 1998; Luck et al., 2000; Luck & Kappenman, 2011; Mangun & Hillyard, 1988; Schoenfeld et al., 2007; Van Voorhis & Hillyard, 1977). More recently, decoding methods have been applied for continuous dynamic stimuli such as continuous speech, movies and video games. (de Cheveigné et al., 2018; Dmochowski et al., 2018; Golumbic et al., 2013; Iotzov & Parra, 2019; Lalor & Foxe, 2010; Mesgarani et al., 2014; Mesgarani & Chang, 2012; Naselaris et al., 2011; O'Sullivan et al., 2015). The contribution of this study was to resolve neural responses in visual space while leveraging methods suitable to continuous dynamic stimuli. Thus, we believe that it is the first study to explore selective spatial responses for a naturalistic dynamic stimulus and task.

The system identification approach taken here has focused on a continuous feature of the stimulus. But the same approach can be used if one is interested instead in responses to discrete events in time (Dimigen &

Ehinger, 2021; Gonçalves et al., 2014; Stankov et al., 2021). What was new here is that the continuous feature was resolved in visual space, which in the present task also carried semantic meaning. The same system identification approach can be used when semantics are associated with discrete events in time such as the semantics of words in continuous speech (Broderick et al., 2018). We hope that the new approach and findings of this study motivates future work to explore complex relationships of neural activity with visual dynamics involving naturalistic stimuli. Additionally, although we show spatial selectivity of neural response to various visual dynamics, we do not explicitly disassociate individual visual components (e.g. localized contrast change or global flow), which has shown to elicit a neural response at select brain regions (Bartels et al., 2008; Russ et al., 2016). By further probing other spatial and temporal dimensions of the stimuli (beyond optical flow), or linking the regression to specific objects, time points or features in the scene rather than locations, we may gain new insights into how neural signals convey visual processing during dynamic natural visual experiences.

4 | METHODS

4.1 | Participant

In total, 42 healthy human subjects (18 females) aged 20 ± 1.56 participated in this experiment. All subjects provided written informed consent in accordance with procedures approved by the Institutional Review Board of the City University of New York.

4.2 | Video stimulus

We used SuperTuxKart, which is an open-source game based on Super Mario Kart, a popular video game that emulates driving. The third-person perspective of this game allows the players to see competing karts behind them without head movements or saccades to a rear-view mirror, as required in real-world driving. All experimental trials were conducted on the default course and spanned three laps in 'easy' mode. Game-related graphical interfaces were programmatically altered from the original game in order to minimize the visual elements and to simplify the task. The modified version of the game only involved the race car, track, passive objects on the road and opponents.

4.3 | Experimental procedure

In the original data collection, all subjects played the game in four different task conditions, which is described in Ki et al. (). Here, in the subsequent analysis, we only report trials from 'active play' and 'passive viewing' conditions. During active play, the subject's used a keyboard to control the kart (using right hand): the left and right keys-controlled steering, whereas the up and down keys produced acceleration and braking. During passive viewing, subjects freely viewed a playback of previously recorded games. Here, subjects were instructed to attend to the video playback without a specific task or goals. For active play, participants were encouraged to complete the race as fast as possible. As motivation, race time was shown at the end of each run, ranked together with performance of previous subjects. Passive viewing did not have a performance metric or incentive. The recordings shown in passive viewing were distinct and not previously seen by the participants but were reused across participants (within each condition, all subjects viewed the same two stimuli). For each condition, subjects performed two trials. The ordering of all conditions was randomized and counterbalanced for all subjects. Prior to the recording of data, subjects were given an instruction on how to play the game, and it was followed by a practice trial.

4.4 | EEG acquisition

The scalp electroencephalogram (EEG) was acquired with a 96-electrode cap (custom montage with dense coverage of the occipital region; see Figure S5) housing active electrodes connected to a Brain Products ActiChamp system and Brain Products DC Amplifier (Brain Vision GmbH, Munich, Germany). The EEG was sampled at 500 Hz, digitized with 24 bits per sample.

4.5 | Video capture acquisition

The video game was played on a high-definition Monitor (Dell 24inch UltraSharp, 1920 × 1080 pixels, 60 Hz) at a viewing distance of 60 cm in a dark and sound-dampened room with playback sound muted. The video game was rendered at 30 Hz. For each trial, the video frame sequence was captured with the Open Broadcaster Software (open source), which records the screen display in real time at the native resolution and frame rate.

4.6 | Eye-tracking acquisition

Eye gaze (right eye) was measured using EyeLink 1000+ (a video-based 2D eye-tracking device, SR Research, Mississauga, Ontario, Canada) at a sampling rate of 500 Hz. For every trial, the eye tracker was calibrated with a 5-point grid (centre and four corners of the display). Subjects were asked to fix their gaze (in turn) at each of five locations where the dots are presented. This step was repeated until the error between all five eye-tracking positions and dot location were less than $2°$. Of the 42 subjects, eye-tracking data was collected from 24 subjects. The eye data were collected for every subject during the first cohort of the experiment (18 of 18 subjects) and part of the second cohort (six of 24 subjects). Eye-tracking data were not used in the original study (Ki et al., 2020); thus, the collection of the eye data was discontinued in the second cohort. Of the 24 subjects, seven subjects missed eye tracking in one of the conditions; thus, they were excluded in the condition comparison analysis. In total, we had eye-tracking data from 17 subjects (active play: 39 trials, passive viewing: 32 trials).

4.7 | EEG, eye tracking and video synchronization

To synchronize the video stimulus with the EEG, we create flash events on the screen with a 30×30 pixel square placed (programmatically implemented) at the top right corner of the game, which flickered on (white) and off (black) at ~2 Hz from the start to the end of each trial. This flickering was not visible to the participants as it was covered by a photodiode, which converts it into event markers. These events were transmitted to a separate event data acquisition hardware that relayed the signal via parallel port to a recording computer (which collected all incoming data: EEG, eye tracker, photodiode). To synchronize the data, we used Lab Streaming Layer software (Kothe, 2014), which created timestamps (in real time) for all incoming data under a universal clock.

To obtain the event markers of the video capture, we isolated the flickering square patch for individual frames and computed the mean pixel intensity values of this region—the pixel intensity events marked on and off flash events on-screen. The event markers from the video and the auxiliary channels allowed us to detect temporal jitter in the video presentation (X of Y runs had to be excluded for this reason) as well linearly compensate for clock drift to synchronize the video capture with EEG and eye-tracking data.

4.8 | Optical flow extraction

The video capture of each kart race game was down-sampled at a spatial resolution of 320×180 pixels and temporal resolution of 30 frames per second using Video Convert Factory (WonderFox Soft Inc). The video is greyscaled and further down-sampled to a grid of 23×40 patches. The optical flow was computed using MATLAB Computer Vision System Toolbox using the function 'opticalFlowHS' (Horn & Schunck, 1981). This algorithm estimates the direction and speed of object motion for each pixel, assuming that the velocity of the brightness pattern varies smoothly (see Figure 1a for examples of estimated velocity vectors). We averaged the magnitude of this velocity estimate across pixels in each patch and normalized across time (z-scored). This yields 920 (23 height $\times$ 40 width) time series features, $s(t)$ for each trial.

4.9 | EEG processing

EEG data were collected at 500 Hz and high-passed with fourth-order Butterworth filters with a cut-off frequency at 1 Hz to remove slow drifts. After this, the data were down-sampled to 30 Hz to match the frame rate. To remove gross artefacts from the data, we employed the robust principal component analysis (PCA) technique (Candès et al., 2011) using the implementation of Lin et al. (2010) with the default hyperparameter of $\lambda = 0.5$. This robust PCA method provides a low-rank approximation to the data and thereby removes sparse noise from the recordings. To reduce the contamination of eye movement from EEG, we linearly regressed out the activity of four frontal electrodes (F9, Fp1, Fp2, F10, in lieu of EOG electrodes, which were not available; see Figure S5). Regression finds the best linear predictor of each electrode from reference electrodes and then subtracts the predicted signal. This technique is routine in EEG (e.g. Schlögl et al., 2007) and can be used also with combinations of EEG electrodes as reference for the eye movement and blink artefacts (Parra et al., 2005). To further denoise the EEG, we rejected electrodes whose mean power exceeded the mean of all channel powers by four standard deviations. Within each channel, we also rejected time samples (and its adjacent samples) whose amplitude exceeded the mean sample amplitude by four standard deviations. These samples were replaced with zeros. We repeated the channel and sample rejection procedures over three iterations.

4.10 | Eye-tracking processing

The eye-tracking data were down-sampled to 30 Hz to match the frame rate of the video. The gaze velocity is computed by differentiating each gaze point, $v(t) = x(t+1) - x(t)$. For further outliers and jitter artefacts removal, we find and replace samples (with not a number marker) in which the gaze velocity magnitude (squared magnitude of the x and y gaze position derivative) was 3 standard deviations greater than the gaze velocity magnitude for a given trial. For computing the gaze position, we compute a histogram across 2D space, which shows the number of times eye gaze was directed at a patch location for individual trials. For display, we normalize and apply Gaussian smoothing over the 2D histogram image.

4.11 | Spatially resolved SRC

The main goal of this study is to assess the difference in neural response to visual dynamics across the visual space. Specifically, we measure the temporal correlation of continuous EEG to optical flow magnitude across the screen. At each patch location on the screen, optical flow, $s(t)$, is linearly mapped to neural response, $r(t)$, via TRF (Crosse et al., 2016; Lalor et al., 2006). This relationship is defined by linear convolution, $\widehat{r}(t) = \sum_{\tau}^{T} h(\tau)s(t-\tau)$, where $h(\tau)$ is the temporal coefficients at each time lag (τ) with temporal window length, $T = 30$ (equivalent to the video sampling rate). This is a classic linear systems identification approach, which reduces to conventional trial averaging of event-related designs, provided the stimulus is discrete in time and events are uncorrelated (Lalor et al., 2006). In this study, neural response can be represented by either individual EEG electrode $r_i(t)$, where i denotes a channel index (Figure 4), or components $r_v(t) = \sum_{i=1}^{D} w_{vi} r_i(t)$, where, $r_v(t)$ represents a linear combination D-electrodes with spatial weights, w_i, across the scalp (Figure 3). The temporal response filter $h(\tau)$ is computed via ridge regression, separately for each electrode or component. SRC is then the correlation between the estimated response $\widehat{r}(t)$ and the actual neural response $r(t)$. For Figure 6, we summed this SRC across the three components depicted in Figure 3.

When considering the component-level analysis of SRC (as opposed to that observed at individual electrodes), we sought to learn linear combinations of the recorded electrodes that best recover visual evoked activity. To that end, we employed canonical correlation analysis (CCA; Hotelling, 1936, Dmochowski et al., 2018). The essence of CCA is the identification of pairs of projection vectors such that the projection of one data set is most strongly correlated to the projection of a second, related data set. This leads to a generalized eigenvalue problem whose solution spans pairs of projection vectors, with the pairs ordered from highest to least correlation (i.e. the first pair of filter outputs shows the strongest correlation). In our application, CCA seeks to identify a temporal filter on the stimulus feature and a spatial filter on the EEG where the filtered signals exhibit maximal temporal correlation. We have found that the bulk of the SRC is contained within the first three components, which explain more than 34.8% of the total correlation between stimulus and EEG. In order to interpret the spatial distribution of the neural responses recovered by the spatial filters, the computed filter weights are transformed into a 'spatial response function' a_j following established techniques (Haufe et al., 2014; Parra et al., 2005). Example spatial response functions are shown in Figure 3a, where the associated spatially resolved SRC values (Figure 3b) have been summed across the first three projections given by CCA.

4.12 | Statistical significance of SRC

To test the significance of SRC, we compare the probability of SRC for all trials to a set of 100 phase-randomized data. We compute the probability of SRC at individual patches being greater than the SRC computed on the surrogate data. The patch locations that are below chance significance ($p > 0.05$) is greyed out.

4.13 | Statistical comparison between left and right visual field

To test the difference between the strength of neural responses between the left and right visual field for central electrodes, we perform a simple chi-squared test that compares the proportions of significant SRC between the left and right patches. The SRC across the left and right visual field is divided at the horizontal centre, resulting in 460 patches for each side, and we counted the number of patches that yielded statistically significant SRC for the two sides.

4.14 | Statical comparison between active play and passive viewing (Figure 5)

For each condition, we computed SRC for every trial (active play: 84 runs, passive viewing: 84 runs) across the 23×40 patch location. To compare the statistical difference between the active play and passive viewing runs,

we performed a Wilcoxon rank-sum test for each patch. We correct for multiple comparisons across the 920 patches using false discovery rate (FDR), that is, by limiting family-wise FDR to 0.05. For comparing average gaze position and saccade velocity, we account for the missing trials by counterbalancing the samples. Here, we averaged the trials for individual subjects, resulting in 17 samples for each condition. For this, we apply pairwise comparison (Wilcoxon signed rank), and for gaze position, we again correct for multiple comparisons across the patches using FDR.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

AUTHOR CONTRIBUTIONS

JPD and JJK designed the experiment. JJK collected data. JJK and LCP analysed the data. LCP, JJK, JDP and JT wrote the paper.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.15503.

DATA AVAILABILITY STATEMENT

The data and code used to generate the results of this study are publicly available on Github repository (https://github.com/JasonJKi/Spatially-Resolved-SRC-SuperTuxKart).

ORCID

Jason J. Ki https://orcid.org/0000-0002-4599-5019
Jacek P. Dmochowski https://orcid.org/0000-0001-6482-5691
Jonathan Touryan https://orcid.org/0000-0001-7343-7869
Lucas C. Parra https://orcid.org/0000-0003-4667-816X

REFERENCES

Bartels, A., Zeki, S., & Logothetis, N. K. (2008). Natural vision reveals regional specialization to local motion and to contrast-invariant, global flow in the human brain. *Cerebral Cortex*, *18*, 705–717. https://doi.org/10.1093/cercor/bhm107

Bavelier, D., Achtman, R. L., Mani, M., & Föcker, J. (2012). Neural bases of selective attention in action video game players. *Vision Research*, *61*, 132–143. https://doi.org/10.1016/j.visres.2011.08.007

Beauchamp, M. S., Petit, L., Ellmore, T. M., Ingeholm, J., & Haxby, J. V. (2001). A parametric fMRI study of overt and covert shifts of visuospatial attention. *NeuroImage*, *14*, 310–321. https://doi.org/10.1006/nimg.2001.0788

Broderick, M. P., Anderson, A. J., Di Liberto, G. M., Crosse, M. J., & Lalor, E. C. (2018). Electrophysiological correlates of semantic dissimilarity reflect the comprehension of natural, narrative speech. *Current Biology*, *28*, 803–809.e3. https://doi.org/10.1016/j.cub.2018.01.080

Busch, N. A., Debener, S., Kranczioch, C., Engel, A. K., & Herrmann, C. S. (2004). Size matters: Effects of stimulus size, duration and eccentricity on the visual gamma-band response. *Clinical Neurophysiology*, *115*, 1810–1820. https://doi.org/10.1016/j.clinph.2004.03.015

Candès, E. J., Li, X., Ma, Y., & Wright, J. (2011). Robust principal component analysis? *Journal of the ACM*, *58*(3), 1–37. https://doi.org/10.1145/1970392.1970395

Capilla, A., Melcón, M., Kessel, D., Calderón, R., Pazo-Álvarez, P., & Carretié, L. (2016). Retinotopic mapping of visual event-related potentials. *Biological Psychology*, *118*, 114–125. https://doi.org/10.1016/j.biopsycho.2016.05.009

Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The multivariate temporal response function (mTRF) toolbox: A MATLAB toolbox for relating neural signals to continuous stimuli. *Frontiers in Human Neuroscience*, *10*, 604. https://doi.org/10.3389/fnhum.2016.00604

David, E. J., Beitner, J., & Võ, M. L.-H. (2021). The importance of peripheral vision when searching 3D real-world scenes: A gaze-contingent study in virtual reality. *Journal of Vision*, *21*, 3–3. https://doi.org/10.1167/jov.21.7.3

de Cheveigné, A., Wong, D. D., Di Liberto, G. M., Hjortkjaer, J., Slaney, M., & Lalor, E. (2018). Decoding the auditory brain with canonical component analysis. *NeuroImage*, *172*, 206–216. https://doi.org/10.1016/j.neuroimage.2018.01.033

Dimigen, O., & Ehinger, B. V. (2021). Regression-based analysis of combined EEG and eye-tracking data: Theory and applications. *Journal of Vision*, *21*, 3–3. https://doi.org/10.1167/jov.21.1.3

Dmochowski, J. P., Ki, J. J., DeGuzman, P., Sajda, P., & Parra, L. C. (2018). Extracting multidimensional stimulus-response correlations using hybrid encoding-decoding of neural activity. *NeuroImage*, *180*, 134–146. https://doi.org/10.1016/j.neuroimage.2017.05.037

Dmochowski, J. P., Sajda, P., Dias, J., & Parra, L. C. (2012). Correlated components of ongoing EEG point to emotionally laden attention – A possible marker of engagement? *Frontiers in Human Neuroscience*, *6*, 112. https://doi.org/10.3389/fnhum.2012.00112

Dorr, M., Martinetz, T., Gegenfurtner, K. R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision*, *10*, 28–28. https://doi.org/10.1167/10.10.28

Duffy, C. J. (1998). MST neurons respond to optic flow and translational movement. *Journal of Neurophysiology*, *80*, 1816–1827. https://doi.org/10.1152/jn.1998.80.4.1816

Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science*, 17, 973–980. https://doi.org/10.1111/j.1467-9280.2006.01815.x

Fogassi, L., & Luppino, G. (2005). Motor functions of the parietal lobe. *Current Opinion in Neurobiology*, 15, 626–631. https://doi.org/10.1016/j.conb.2005.10.015

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, 51, 1920–1931. https://doi.org/10.1016/j.visres.2011.07.002

Gallant, J. L., Nishimoto, S., Naselaris, T., & Wu, M. C. (2012). System identification, encoding models, and decoding models: A powerful new approach to fMRI research. In N. Kriegeskorte & G. Kreiman (Eds.), *Visual population codes: Toward a common multivariate framework for cell recording and functional imaging* (pp. 163–188). MIT Press.

Golumbic, E. M. Z., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., Goodman, R. R., Emerson, R., Mehta, A. D., Simon, J. Z., Poeppel, D., & Schroeder, C. E. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron*, 77(5), 980–991. https://doi.org/10.1016/j.neuron.2012.12.037

Gonçalves, N. R., Whelan, R., Foxe, J. J., & Lalor, E. C. (2014). Towards obtaining spatiotemporally precise responses to continuous sensory stimuli in humans: A general linear modeling approach to EEG. *NeuroImage*, 97, 196–205. https://doi.org/10.1016/j.neuroimage.2014.04.012

Hanning, N. M., Szinte, M., & Deubel, H. (2019). Visual attention is not limited to the oculomotor range. *Proceedings of the National Academy of Sciences*, 116, 9665–9670. https://doi.org/10.1073/pnas.1813465116

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. J. (2008). Neurocinematics: The neuroscience of film. *PRO*, 2, 1–26. https://doi.org/10.3167/proj.2008.020102

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., & Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*, 87, 96–110. https://doi.org/10.1016/j.neuroimage.2013.10.067

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences*, 95, 781–787. https://doi.org/10.1073/pnas.95.3.781

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 353, 1257–1270. https://doi.org/10.1098/rstb.1998.0281

Holdgraf, C. R., Rieger, J. W., Micheli, C., Martin, S., Knight, R. T., & Theunissen, F. E. (2017). Encoding and decoding models in cognitive electrophysiology. *Frontiers in Systems Neuroscience*, 11, 61. https://doi.org/10.3389/fnsys.2017.00061

Holländer, A., Jung, C., & Prinz, W. (2011). Covert motor activity on NoGo trials in a task sharing paradigm: Evidence from the lateralized readiness potential. *Experimental Brain Research*, 211, 345–356. https://doi.org/10.1007/s00221-011-2688-x

Horn, B. K., & Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence*, 17(1-3), 185–203. https://doi.org/10.1016/0004-3702(81)90024-2

Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28(3–4), 321–377. https://doi.org/10.2307/2333955

Iotzov, I., & Parra, L. C. (2019). EEG can predict speech intelligibility. *Journal of Neural Engineering*, 16, 036008. https://doi.org/10.1088/1741-2552/ab07fe

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, 49, 1295–1306. https://doi.org/10.1016/j.visres.2008.09.007

Kay, K. N., Naselaris, T., Prenger, R. J., & Gallant, J. L. (2008). Identifying natural images from human brain activity. *Nature*, 452, 352–355. https://doi.org/10.1038/nature06713

Kay, K. N., & Yeatman, J. D. (2017). Bottom-up and top-down computations in word-and face-selective cortex. *eLife*, 6, e22341. https://doi.org/10.7554/eLife.22341

Kelly, S. P., Foxe, J. J., Newman, G., & Edelman, J. A. (2010). Prepare for conflict: EEG correlates of the anticipation of target competition during overt and covert shifts of visual attention. *The European Journal of Neuroscience*, 31, 1690–1700. https://doi.org/10.1111/j.1460-9568.2010.07219.x

Kelly, S. P., Vanegas, M. I., Schroeder, C. E., & Lalor, E. C. (2013). The cruciform model of striate generation of the early VEP, re-illustrated, not revoked: A reply to ales et al. (2013). *NeuroImage*, 82, 154–159. https://doi.org/10.1016/j.neuroimage.2013.05.112

Ki, J. J., Kelly, S. P., & Parra, L. C. (2016). Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. *The Journal of Neuroscience*, 36, 3092–3101. https://doi.org/10.1523/JNEUROSCI.2942-15.2016

Ki, J. J., Parra, L. C., & Dmochowski, J. P. (2020). Visually evoked responses are enhanced when engaging in a video game. *The European Journal of Neuroscience*, 52, 4695–4708. https://doi.org/10.1111/ejn.14924

Kothe, C. (2014). Lab streaming layer (LSL). Github.com/sccn/labstreaminglayer. https://github.com/sccn/labstreaminglayer

Kulke, L. V., Atkinson, J., & Braddick, O. (2016). Neural differences between covert and overt attention studied using EEG with simultaneous remote eye tracking. *Frontiers in Human Neuroscience*, 10, 592. https://doi.org/10.3389/fnhum.2016.00592

Lalor, E. C., & Foxe, J. J. (2009). Visual evoked spread spectrum analysis (VESPA) responses to stimuli biased towards magnocellular and parvocellular pathways. *Vision Research*, 49, 127–133. https://doi.org/10.1016/j.visres.2008.09.032

Lalor, E. C., & Foxe, J. J. (2010). Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *The European Journal of Neuroscience*, 31, 189–193. https://doi.org/10.1111/j.1460-9568.2009.07055.x

Lalor, E. C., Pearlmutter, B. A., Reilly, R. B., McDarby, G., & Foxe, J. J. (2006). The VESPA: A method for the rapid estimation of a visual evoked potential. *NeuroImage*, 32, 1549–1561. https://doi.org/10.1016/j.neuroimage.2006.05.054

Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature*, 369, 742–744. https://doi.org/10.1038/369742a0

Lebedev, M. A., & Wise, S. P. (2001). Tuning for the orientation of spatial attention in dorsal premotor cortex. *The European Journal of Neuroscience*, 13, 1002–1008. https://doi.org/10.1046/j.0953-816x.2001.01457.x

Lin, Z., Chen, M., & Ma, Y. (2010). The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. *arXiv*, 1009.5055. https://arxiv.org/abs/1009.5055

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, 77, 24–42. https://doi.org/10.1152/jn.1997.77.1.24

Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components. *Electroencephalography and Clinical Neurophysiology*, 75, 528–542. https://doi.org/10.1016/0013-4694(90)90139-B

Luck, S. J. & Kappenman, E. S. (Eds.) (2011). *The Oxford handbook of event-related potential components*. Oxford University Press.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, 4, 432–440. https://doi.org/10.1016/S1364-6613(00)01545-X

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, 32, 4–18. https://doi.org/10.1111/j.1469-8986.1995.tb03400.x

Mangun, G. R., Buonocore, M. H., Girelli, M., & Jha, A. P. (1998). ERP and fMRI measures of visual spatial selective attention. *Human Brain Mapping*, 6, 383–389. https://doi.org/10.1002/(SICI)1097-0193(1998)6:5/6<383::AID-HBM10>3.0.CO;2-Z

Mangun, G. R., & Hillyard, S. A. (1988). Spatial gradients of visual attention: Behavioral and electrophysiological evidence. *Electroencephalography and Clinical Neurophysiology*, 70, 417–428. https://doi.org/10.1016/0013-4694(88)90019-3

Maunsell, J. H., & Cook, E. P. (2002). The role of attention in visual processing. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 357(1424), 1063–1072. https://doi.org/10.1098/rstb.2002.1107

McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *The Journal of Neuroscience*, 19, 431–441. https://doi.org/10.1523/JNEUROSCI.19-01-00431.1999

Mesgarani, N., & Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, 485, 233–236. https://doi.org/10.1038/nature11020

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science*, 343, 1006–1010. https://doi.org/10.1126/science.1245994

Miyawaki, Y., Uchida, H., Yamashita, O., Sato, M., Morito, Y., Tanabe, H. C., Sadato, N., & Kamitani, Y. (2008). Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. *Neuron*, 60, 915–929. https://doi.org/10.1016/j.neuron.2008.11.004

Moore, T. (1999). Shape representations and visual guidance of saccadic eye movements. *Science*, 285, 1914–1917. https://doi.org/10.1126/science.285.5435.1914

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229, 782–784. https://doi.org/10.1126/science.4023713

Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, 70, 909–919. https://doi.org/10.1152/jn.1993.70.3.909

Naranjo, J. R., Brovelli, A., Longo, R., Budai, R., Kristeva, R., & Battaglini, P. P. (2007). EEG dynamics of the frontoparietal network during reaching preparation in humans. *NeuroImage*, 34, 1673–1682. https://doi.org/10.1016/j.neuroimage.2006.07.049

Nardo, D., Santangelo, V., & Macaluso, E. (2011). Stimulus-driven orienting of visuo-spatial attention in complex dynamic environments. *Neuron*, 69, 1015–1028. https://doi.org/10.1016/j.neuron.2011.02.020

Naselaris, T., Kay, K. N., Nishimoto, S., & Gallant, J. L. (2011). Encoding and decoding in fMRI. *NeuroImage*, 56, 400–410. https://doi.org/10.1016/j.neuroimage.2010.07.073

Naselaris, T., Prenger, R. J., Kay, K. N., Oliver, M., & Gallant, J. L. (2009). Bayesian reconstruction of natural images from human brain activity. *Neuron*, 63, 902–915. https://doi.org/10.1016/j.neuron.2009.09.006

Nishimoto, S., Vu, A. T., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. L. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology*, 21, 1641–1646. https://doi.org/10.1016/j.cub.2011.08.031

Nobre, A. C., Sebestyen, G. N., & Miniussi, C. (2000). The dynamics of shifting visuospatial attention revealed by event-related potentials. *Neuropsychologia*, 38, 964–974. https://doi.org/10.1016/S0028-3932(00)00015-4

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., Slaney, M., Shamma, S. A., & Lalor, E. C. (2015). Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cerebral Cortex*, 25, 1697–1706. https://doi.org/10.1093/cercor/bht355

Parker, S., Heathcote, A., & Finkbeiner, M. (2020). Establishing the separable contributions of spatial attention and saccade preparation across tasks with varying acuity demands. *Journal of Experimental Psychology. Human Perception and Performance*, 47, 172–188. https://doi.org/10.1037/xhp0000881

Parra, L. C., Spence, C. D., Gerson, A. D., & Sajda, P. (2005). Recipes for the linear analysis of EEG. *NeuroImage*, 28, 326–341. https://doi.org/10.1016/j.neuroimage.2005.05.032

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, 32, 3–25. https://doi.org/10.1080/00335558008248231

Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1984). Effects of parietal injury on covert orienting of attention. *The Journal of Neuroscience*, 4(7), 1863–1874. https://doi.org/10.1523/JNEUROSCI.04-07-01863.1984

Praamstra, P., & Oostenveld, R. (2003). Attention and movement-related motor cortex activation: A high-density EEG study of spatial stimulus–response compatibility. *Cognitive Brain Research*, 16, 309–322. https://doi.org/10.1016/S0926-6410(02)00286-0

Rugg, M. D., Milner, A. D., Lines, C. R., & Phalp, R. (1987). Modulation of visual event-related potentials by spatial and nonspatial visual selective attention. *Neuropsychologia*, 25, 85–96. https://doi.org/10.1016/0028-3932(87)90045-5

Russ, B. E., Kaneko, T., Saleem, K. S., Berman, R. A., & Leopold, D. A. (2016). Distinct fMRI responses to self-induced versus stimulus motion during free viewing in the macaque. *The Journal of Neuroscience*, 36, 9580–9589. https://doi.org/10.1523/JNEUROSCI.1152-16.2016

Schlögl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., & Pfurtscheller, G. (2007). A fully automated correction method of EOG artifacts in EEG recordings. *Clinical Neurophysiology*, *118*, 98–104. https://doi.org/10.1016/j.clinph.2006.09.003

Schoenfeld, M. A., Hopf, J.-M., Martinez, A., Mai, H. M., Sattler, C., Gasde, A., Heinze, H.-J., & Hillyard, S. A. (2007). Spatio-temporal analysis of feature-based attention. *Cerebral Cortex*, *17*, 2468–2477. https://doi.org/10.1093/cercor/bhl154

Self, M. W., Peters, J. C., Possel, J. K., Reithler, J., Goebel, R., Ris, P., Jeurissen, D., Reddy, L., Claus, S., Baayen, J. C., & Roelfsema, P. R. (2016). The effects of context and attention on spiking activity in human early visual cortex. *PLoS Biology*, *14*, e1002420. https://doi.org/10.1371/journal.pbio.1002420

Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science*, *240*, 338–340. https://doi.org/10.1126/science.3353728

Stankov, A. D., Touryan, J., Gordon, S., Ries, A. J., Ki, J., & Parra, L. C. (2021). During natural viewing, neural processing of visual targets continues throughout saccades. *Journal of Vision*, *21*(10), 7. https://doi.org/10.1101/2021.02.11.430486

Tanaka, K., & Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, *62*, 626–641. https://doi.org/10.1152/jn.1989.62.3.626

Tootell, R. B. H., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., & Dale, A. M. (1998). The Retinotopy of visual spatial attention. *Neuron*, *21*, 1409–1422. https://doi.org/10.1016/S0896-6273(00)80659-5

Van Voorhis, S., & Hillyard, S. A. (1977). Visual evoked potentials and selective attention to points in space. *Perception & Psychophysics*, *22*, 54–62. https://doi.org/10.3758/BF03206080

VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*, *13*, 454–461. https://doi.org/10.1162/08989290152001880

Vo, M. L.-H., & Henderson, J. M. (2009). Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision*, *9*(3), 24. https://doi.org/10.1167/9.3.24

Võ, M. L.-H., & Henderson, J. M. (2010). The time course of initial scene processing for eye movement guidance in natural scene search. *Journal of Vision*, *10*, 14–14, 13. https://doi.org/10.1167/10.3.14

Võ, M. L.-H., & Wolfe, J. M. (2015). The role of memory for visual search in scenes. *Annals of the New York Academy of Sciences*, *1339*, 72–81. https://doi.org/10.1111/nyas.12667

Wang, Y.-K., Jung, T.-P., & Lin, C.-T. (2015). EEG-based attention tracking during distracted driving. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *23*, 1085–1094. https://doi.org/10.1109/TNSRE.2015.2415520

Warren, W., Morris, M., & Kalish, M. (1988). Perception of translational heading from optical flow. *Journal of Experimental Psychology. Human Perception and Performance*, *14*, 646–660. https://doi.org/10.1037//0096-1523.14.4.646

Warren, W. H., & Hannon, D. J. (1988). Direction of self motion is perceived from optical flow. *Nature*, *336*, 162–163. https://doi.org/10.1038/336162a0

Wu, M. C.-K., David, S. V., & Gallant, J. L. (2006). Complete functional characterization of sensory neurons by system identification. *Annual Review of Neuroscience*, *29*, 477–505. https://doi.org/10.1146/annurev.neuro.29.051605.113024

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.